Improving non-destructive techniques for stone weathering research *in situ*

Katrin Wilhelm

Kellogg College

Thesis submitted for the degree of

Doctor of Philosophy of Science

University of Oxford

01/06/2016

Abstract

"In time, and with water, everything changes" (Leonardo da Vinci in Kemp, 2006). In the field of cultural built heritage these changes often eventually result in loss of unique irreplaceable sites. This loss is considered to have an effect on societies as heritage is an important part of cultural identity and future development. In order to prolong the life of built heritage structures and preserve the collective memory they represent the weathering behaviour of the materials needs to be understood. Stone is a very common component of built heritage, especially limestone which is the focus of this thesis.

Stone weathering behaviour can be investigated under controlled laboratory conditions, but results do not entirely reflect its behaviour under real world conditions (because of complex weathering histories and spatial heterogeneity found on real built heritage). Therefore, it is necessary to complement the laboratory approach with *in situ* investigations. For *in situ* investigation a variety of methods is available ranging from destructive to non-destructive (NDT) and sophisticated and expensive to more simple and economical. This thesis is based on the key principle of built heritage conservation i.e. to preserve as much original fabric as possible and keep destructive sampling to a minimum. Furthermore, to allow for wider application on a bigger scale and more frequently, the focus has been on non-destructive, portable, and economical methods. However, standards and good practice guides for these methods have not yet been developed. Thus, the overall aim of this thesis was to develop reliable methodologies for these methods in order to quantify the extent and rate of limestone heritage decay *in situ* under real world conditions.

The thesis has three objectives. Objective 1 improved the application of selected NDT methods under laboratory conditions, focusing on sampling protocols (e.g. sample sizes) and reliability of data generated. Innovative aspects of research for this objective include extending their application (converting some drawbacks into advantages), combining them and applying modern statistical methods to the data evaluation. With this approach information on stone surface and subsurface properties was gained. This assists to capture stone weathering behaviour trajectory more holistically by investigating processes preceding total stone mass loss (erosion). Objective 2 applied the improved NDT methods to a time series of dated Portland limestone gravestones covering 1 to 248 years of exposure in order to evaluate the changing rate of surface property changes. The method proposed here provides a novel application of surface hardness data for quantifying stone deterioration rates over short- and long-term. Further, QC_{50} (the regression coefficient for 0.50 quantile regression) is introduced as novel robust measure for surfaces property changes. Is was found that depending on the time scale of investigation weathering behaviour is either defined as non-linear (whole period of 248 years) or linear (periods <100 years). It was found that stone weathering behaviour in cases needs to be investigated below block scale due to spatial variances. Objective 3 applied the NDT methods to diagnose the nature and causes of catastrophic limestone deterioration observed after a harsh winter at the archaeological site of Dülük Baba Tepesi, South Turkey. The cause for catastrophic stone decay *in situ* were reconstructed using NDT techniques and past climate data reports. This provides a novel application to infer the cause of catastrophic decay *in situ* by combining moisture uptake characteristics with robust data evaluation for surface and subsurface hardness data with past meteorological data. It was concluded that the Hellenistic-Roman structures are too vulnerable to be exposed to the prevalent environment without any further preservation measures.

Similar to the 'scientific toolkit' recommended by Meneely et al. (2009) for more sophisticated methods (e.g. 3D laser scan, ground penetrating radar etc.) the methods evaluated in this thesis are seen as a contribution to a potential 'scientific toolkit of low-cost methods' which could be complemented with other methods like ultrasound velocity measurements, drilling resistance etc. Thus, this study shows that the improved methods may assist in both 1) understanding heritage stone weathering under real world conditions (without damaging them by sample taking, whilst capturing surface/subsurface changes); and 2) more frequent investigation of the state of preservation/deterioration of stone heritage on-site in order to detect ongoing deterioration at an early stage.

Acknowledgments

Heartfelt thanks are due to my supervisor, Professor Heather Viles, who inspired me to undertake this DPhil. Her unmatched passion for weathering and gadgets was central to many of the ideas in this thesis, and her unfaltering support, patience and enthusiasm were truly appreciated every step of the way.

I acknowledge the financial and material help provided by Proceq ©, the Engineering and Physical Sciences Research Council (EPSRC) and the German DAAD, the Oxford University Estates Services and the British Geological Survey without which I would not have been able to conduct this research.

Thanks are also given to Dr Martin Coombes, for his scientific rigour and encouragement, and for very fruitful conceptual and methodological discussions. Troy Sternberg provided a fantastic platform off which to bounce ideas, and Jerome Mayaud is thanked deeply for his infinite help and support in the field, with formatting and for being always there when needed.

The loyal support of all my friends, particularly The Ladies, never faltered to believe in me, and for that I thank them from the bottom of my heart. I would also like to acknowledge all my fellow departmental postgraduates, who helped forge a real sense of community within OUCE. Finally, thanks go to my sister and chosen family, for the enduring love that has helped see me through.

Table of contents

List of figures

List of tables

List of acronyms, notations and abbreviations

δ	Density of water = 0.998 g/cm^3 at ~20°C
(ρ or rs	Spearman's rank correlation coefficient
°C	Celsius
μ	Average
μm	Micrometre
3D	Three dimensional
A	Area through which water penetrated
Al_2O_3	Aluminium oxide
ANOVA	Analysis of variance
ASTM	American Society for Testing and Materials
ave	Average
BAT	Bath limestone
Bca	Bias corrected and accelerated
BET	Brunauer, Emmett and Teller
BGS	British Geological Survey
BRE	Building Research Establishment
BS	Bending strength
BS-EN	British Standards – European Norm (Standard)
BSI	British Standards Institute
B-value	Water penetration coefficient
$C_2H_2O_4$	Oxalatic acid
$C_2O_4^{2-}$	Oxalate
$Ca(C_2H_3O_2)_2$	Calcium acetate
Ca(NO3)2	Calcium nitrate
$Ca(OH)_2$	Calcium hydroxide
Ca_2	Calcium
Ca^{2+}	Calcium
$CaCO_3$	Calcium carbonate
$CaSO_4$	Calcium sulphate (gypsum)
CCT	Churches Conservation Trust
Cfb1	Temperate, without dry season, warm summer (Koeppen climate map)
CH_3COO^-	Acetate
Cl^-	Chloride
CLI	Clipsham limestone
cm^3	Cubic centimetre
CO_3^{2-}	Carbonate
Csa1	Temperate, without dry season, hot summer (Koeppen climate map)
CT	X-ray tomography
CWGC	Commonwealth War Grave Commission
D	Diameter of Karsten tube
df	Degrees of freedom
DIN	Deutsche Industrie Norm (German industrial standard)
Equ	Equilibrium
ERH	equilibrium relative humidity
EU	European Union
Fe_2O_3	Iron oxide
g	Gram

GPa	Gigapascal
GPR	Ground penetrating radar
GUI	Guiting limestone
H	Hydrogen
H_o	Null Hypothesis
H_2O	Water
HCO_3	Bicarbonate
$HCOO^-$	Formate
$HLDL_{R.med}$	DL-probe, median of the 3 highest values in each of the 3 repeated impact method (RIM) datasets of 20 readings
$HLDL_{S.MAD}$	DL-probe, single impact method, median absolute deviation
$HLDL_{S.mean}$	DL-probe, single impact method, mean
$HLDL_{S.med}$	DL-probe, single impact method, median
$HLDL_{S.SD}$	DL-probe, single impact method, standard deviation
$HLD_{R.med}$	D-probe, median of the 3 highest values in each of the 3 repeated impact method (RIM) datasets of 20 readings
$HLD_{S.MAD}$	D-probe, single impact method, median absolute deviation
$HLD_{S.mean}$	D-probe, single impact method, mean
$HLD_{S.med}$	D-probe, single impact method, median
$HLD_{S.SD}$	D-probe, single impact method, standard deviation
HMC	hygroscopic moisture content
HNO_3	Nitric acid
IQR	Inter quartile range
IR	Infrared
k or εr	dielectric constant or permittivity
kΩcm	kilo-ohm cm
K^+	Potassium
kN/m^3	Kilo Newton per cubic metre
KNO_3	Potassium nitrate
L	Litre
LED	Light emitting diode
LQS	Least squares
M	Penetrated water
m/s	Metre per second
m^2	Square metre
Ma	Million years
MAD	Median absolute deviation
MEM / TMEM	Micro erosion meter / Traverse micro erosion meter
Mg^{2+}	Magnesium
$MgCO_3$	Magnesium carbonate
MgO	Magnesium oxide
$MgSO_4 \bullet 7H_2O$	Magnesium sulphate
$MgSO_4 \bullet 7H_2O$	Epsomite
MIP	Mercury porosimetry
MIT	Minimally invasive technique
ml	Millilitre
Mm	Millimetre
MPa	Mega Pascal
N m	Newton metre
Na(HCOO)	Sodium formate
Na^+	Sodium

Na_2CO_3	Natrite
Na_2SO_4	Sodium sulphate
NaCl	Sodium chloride
NaCl	Sodium chloride
NDT	Non-destructive technique
NH_4^+	Ammonium
NMEP	National Materials Exposure Program
NO_3^-	Nitrate
NO_x	Nitrogen compounds
O_3	Ozone
OxRBL	Oxford Rock Breakdown Laboratory
pH	Power of hydrogen (*potentia hydrogenii*)
PLI	Point load strength index
POR	Portland limestone
p-value	Significance level for α (the probability of making a Type I error in statistics)
QC_{50}	Gradient quantile regression for median
R	Radius
R^2	LQS Pearson's coefficient
Rel%	Relative moisture measurement for capacitance meter
RH%	Relative humidity
RILEM	Recherches sur les Materiaux et les Construction
RIM	Repeated impact method
s.m.	Similar moisture
S0, S1, S2	Salt contamination levels
SAS	Statistical Analysis Software
SD	Standard deviation
SHT	Surface hardness testing
SIM	Single impact method
SiO_2	Silicium oxide
SO_2	Sulphur dioxide
SO_4^{2-}	Sulphate
SPSS	Statistical Package for the Social Sciences
S-SW	South Southwest
SW	Southwest
t	Time
Tmf	Fırat formation
Tmga	Gaziantep formation
UCS	Unconfined compressive strength
UK	United Kingdom
UPV	Ultrasonic pulse velocity
V_{kar}	Volume of water taken up through the Karsten tube
W/mK	Watt per Kelvin
WAAP	Water absorption under low (atmospheric) pressure
WAK	Water uptake capacity
WME	Wood moisture equivalent
wt%	Weight percentage
w-value	Water uptake coefficient
XRF	X-ray fluorescence

1. INTRODUCTION

1.1. INTRODUCING THE CHALLENGE –HERITAGE AT RISK

An astounding amount of cultural heritage is literally 'written in stone'. Monuments like the Taj Mahal in Agra and St. Paul’s Cathedral in London are just two prominent examples of thousands of built heritage sites around the world which have survived for centuries (Figure 1.1 and Figure 1.2). Stone was often the first choice building material as it is more durable than many other materials like wood, glass or textiles (Steiger et al., 2010). However, looking in more detail, all stone monuments inexorably deteriorate over time due to climatic impact, air-pollution and other anthropogenic activity.

Figure 1.1 St Paul's cathedral in London (UK) (source: Inkpen et al., 2012b)

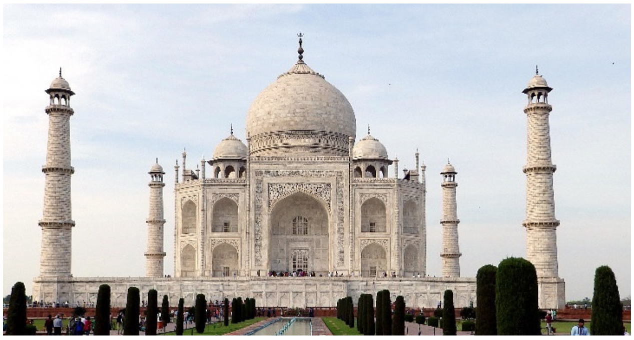


Figure 1.2 Taj Mahal in Agra (India) (source: diver, 2013)

The decay of stone build heritage is progressive and irreversible (Figure 1.3 and 1.4). Losing cultural heritage has a severe impact on societies as it is an integral part of people's cultural identity and practices and thus, plays an important role as a foundation for future development (e.g. Jha and Duyne, 2010; Hall, 2011;

European Union, 2013). Stone weathering has been recognized as a problem for millennia (Vitruvius 2:VII Granger, F., 1970, Herodotus Thomas and Rawlinson, 1997). Yet, only since the Industrial Revolution (~ 1850) has it been accelerating to an extent which is seriously concerning (e.g. Doehne and Price, 2010; Smith et al., 2011a). Considering the accelerated progressive decay of irreplaceable cultural heritage, its preservation should be a high priority (Fitzner, 2002).

Figure 1.3 (above) Germany, Munich, Königsplatz (Simon 2001)

Figure 1.4 (right) Italy, Pompeii, (excavation site), Casa del Labirinto, oecus 1998, collapsed ceiling is leaned against wall painting (Heritage at Risk 2000)

As Petzet (2010) concludes, the recent dramatic changes in the cultural heritage landscape (which are not comparable to gradual changes experienced over past centuries) puts increased pressure on the built heritage conservation and stone weathering science community to protect and preserve built heritage. The interactions of heritage stone material with the environment in the field are very complex and still not fully understood. Yet, only a thorough understanding of factors involved will allow for the right conclusions to be drawn and appropriate preservation measures to be undertaken (e.g. Svahn, 2006; Auras 2011a; Inkpen et al., 2012b Přikryl, 2013).

1.2. LIMESTONE

A common practice in the past was to source building materials locally. Therefore, the surrounding geology often determined the used building stone (Adam, 1999). Accordingly, stone built heritage consists of all types of stone available like igneous, metamorphic and sedimentary stone. Sedimentary rocks cover more than 50% of the earth's surface and thus, play a fundamental role in human cultural history as building material (Schön, 2011). Across Europe limestone has been the material of choice for many centuries (Smith et al., 2010). Limestone is found in many stages of the Earth's geological time scale. In the UK, for example, the main building limestones date from the Jurassic period (201.3 Ma – 145 Ma) (Leary, 1983; Huang, 2012)(Figure 1.5). Furthermore, usually the most widely used limestones are worked easily due to their less intense diagenesis processes (Smith et al., 2010). Though, the history of sedimentation and diagenesis of a limestone determines its resilience or vulnerability to weathering (Mosch and Siegesmund, 2007; Steiger et al., 2010). Accordingly, O'Brien (1995) found limestone to be the most vulnerable of the tested stones in his study of stone durability. However, Mosch and Siegesmund (2007) point out the high variability within limestones due to inherent natural inhomogeneity and anisotropy. Therefore, generalizations of weathering behaviour across limestones are difficult to make and May (1998) suggests that limestone types with their subcategories (formations, beds etc.) need to be investigated separately.

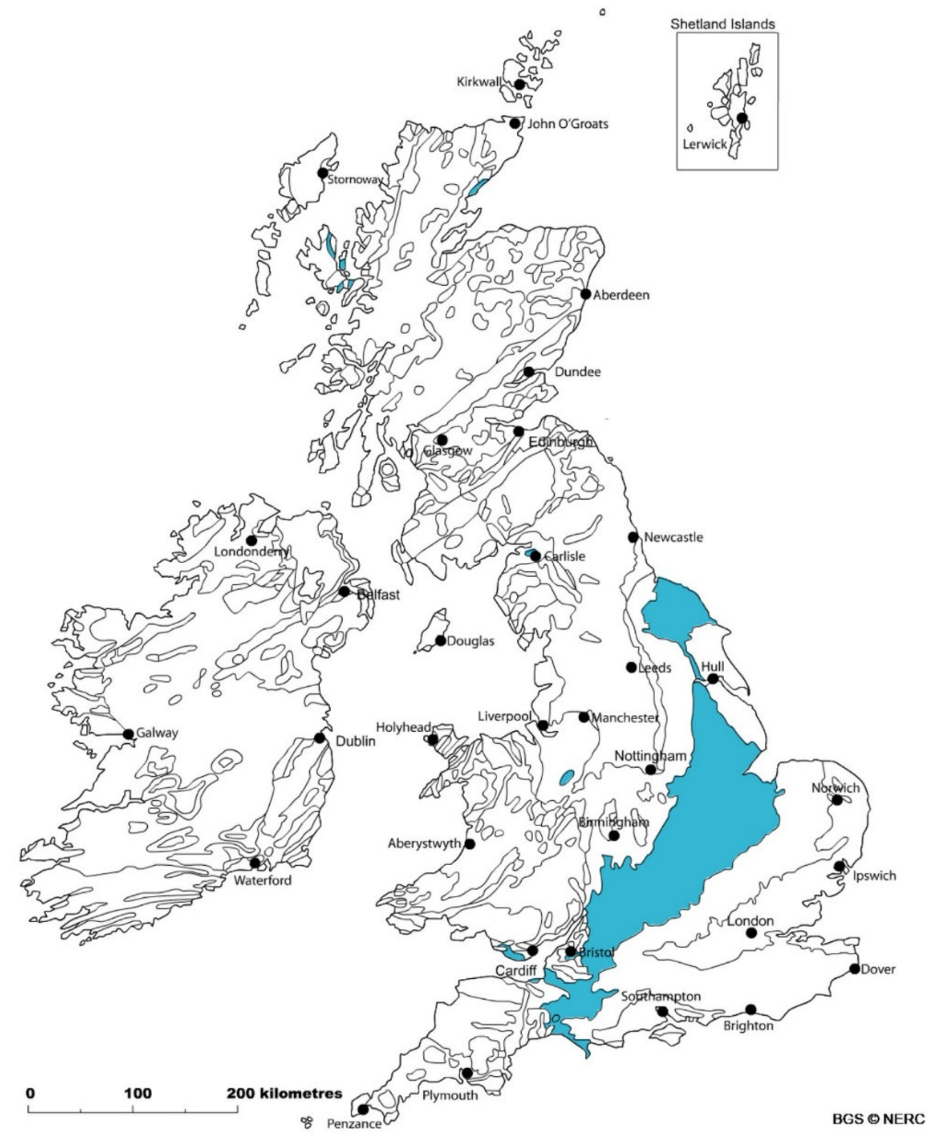


Figure 1.5 Map of Britain with the Jurassic period belt highlighted in blue (British Geological Survey BGS © NERC. Exported from interactive Make-a-map resource 10/2015)

Limestone is particularly prone to acidic chemical attack induced by air pollution and has been the focus of many weathering studies in the past (Charola and Ware 2002; Mitchell and Searle 2004; Brimblecombe and Grossi 2009). Current trends in conservation research are impacts of climate change (Doehne and Price, 2010) and much importance is given to threats of increased air pollution to cultural heritage (e.g. Brimblecombe and Grossi, 2009; Doehne and Price, 2010; Ruddiman, 2010; Zalasiewicz et al., 2011; Fuente et al., 2013; Howard, 2013).

1.3. CHALLENGES TO QUANTIFYING STONE WEATHERING BEHAVIOUR

"The whole is greater than the sum of its parts- dilemma" – The bigger picture: complex character of deterioration

To sufficiently protect and preserve built heritage understanding and quantifying the nature, causes and controls of stone weathering behaviour over time is key (Smith et al., 2008). The nature of stone weathering behaviour can be slow and steady or rapid and catastrophic (e.g. Sass and Viles, 2010; Smith et al., 2010). The change of initial inherent stone properties over time in interaction with external physical, biological and chemical impacts through climate, air-pollution and anthropogenic activity leads ultimately to decay. The intensity, duration and frequency of these internal changes and external impacts control the stone weathering behaviour. Stone weathering behaviour has been investigated under controlled laboratory conditions or *in situ* with either exposing samples or sampling destructively and non-destructively from real built heritage. Figure 1.6 gives an overview of the different approaches to investigate stone weathering behaviour.

Stone weathering behaviour in laboratory tests

The majority of research on stone weathering behaviour to date has been based on laboratory experiments. Fresh or artificially weathered stones are usually investigated (e.g. Yavuz et al., 2006; Bourges, 2006; Ahmad et al., 2009). To obtain results in a reasonable time the standard resistance tests run accelerated, extreme frost cycles (-20° under full saturation, BS-EN 12371-2010)

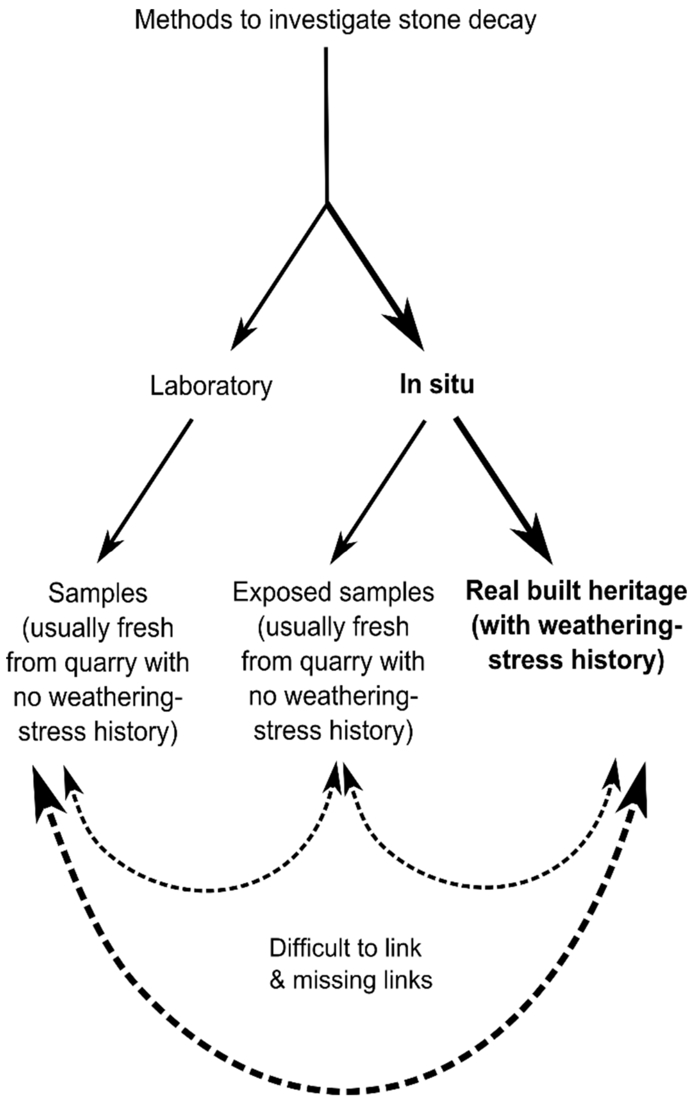


Figure 1.6 Overview of methods to investigate stone decay. The two main strands are 1) in the laboratory under controlled conditions using samples with known weathering stress history (mainly fresh stone from the quarry) and 2) in situ using exposed samples with known weathering stress history or investigating heritage structures directly (either taking samples or investigating the structure directly). The core interest for this this is marked bold: non-destructive methods for in situ testing of real heritage

and use heavy loads of highly damaging salt (Sodium sulphate, BS-EN 12370-1999) on test specimens. Laboratory experiments also use small samples such as for example suggested by the BS-EN 1926:2006 standard for unconfined compressive strength with required sample dimensions of 50 x 50 x 50 mm. Both, the high surface area to volume ratio and lack of constraints of the samples where stone is often part of a built structure and thus only one or two surfaces are exposed to weathering. Furthermore, the focus of standard tests is on testing the durability of fresh stone material in order to evaluate their suitability for engineering applications. However, for stone weathering behaviour of real heritage architecture resilience should be the focus. Resilience describes the ability of a system (here built heritage structure) to remain functionally consistent under and recover from weathering impacts (Giesen et al., 2011).

Weathering-stress history

Laboratory experiments cannot simulate heterogeneity of patterns, whereas at larger scales these are evident as built heritage such heterogeneity arises from the often complex histories. Consequently, laboratory tests frequently fail to account for the weathering-stress history of the stone (e.g. Bell, 1993; Warke et al., 2003; Moroni and Pitzurra, 2008; Inkpen et al., 2012a) as shown by McGreevy and Smith (1982) in salt weathering experiments, Trudgill and Viles (1998) for chemical weathering of limestone and Ingham (2005), who observes noticeably different frost weathering behaviour of stone when compared laboratory experiments and *in situ* results. Therefore, such tests have been heavily criticized (e.g. Siegesmund and Kirchner, 2003; Goudie, 1999). Instead, in terms of predicting weathering behaviour of natural stones scientists agree that tests should be conducted on naturally weathered stones (Bourges 2006).

Sampling real heritage is critical

The interactions of heritage stone with the environment in the field are very complex and still not fully understood and difficult to simulate under laboratory conditions. An alternative approach is to take samples from real heritage and investigate them in the laboratory. So for example Meinhardt-Degen (2005) investigated drill cores taken at heritage sites to gain insight into performance of consolidation materials and their re-application. Samples in the laboratory are usually investigated using standard methods like microscopy (VIS and SEM), water uptake under atmospheric pressure and vacuum (e.g. BS-EN 13755:2008 and 1936:2006 respectively), unconfined compressive strength (BS-EN 1926:2006) etc. (e.g. Ahmad, 2011). Meinhardt-Degen (2005) however points out that drill cores often lack to represent the whole heterogeneity of the historic structure (though she mentions that this may be balanced with an experienced practical restorer, similarly to Svahn (2006) who emphasizes that "more precise conservation documentation in an organized and systematic way" needs to be provided as knowing the conservation history of the particular stone is crucial in order to successfully preserve it). In addition, Viles (2001) emphasizes, that depending on the scale of investigation weathering process-response systems may be characterised as ordered or chaotic. Figure 1.7 shows a comparison of weathering patterns and their scales where for example at the micro and macro scale the responses might be described at ordered where at the meso scale it appears complex (chaotic) and/or change from ordered to chaotic depending on the temporal scale of investigation (e.g. Arbona et al. 2014).

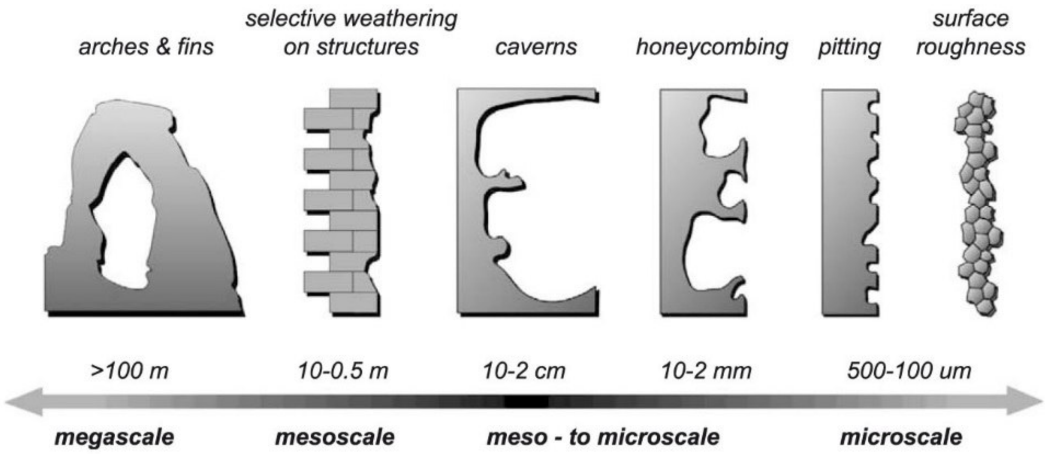


Figure 1.7 Scale dependent weathering morphology (Turkington et al., 2005)

Despite the deeper insight sampling of real heritage provides, there are clear drawbacks as destructive tests i) require special apparatus and expertise, ii) can only represent a small area of an object (due to sample taking restrictions often associated with cultural heritage objects), iii) are costly and iv) most importantly involve sample taking, which often violates key principles of built heritage conservation i.e. to preserve as much original fabric as possible (e.g. the Venice Charter, 1964; the Malta Convention, 1992; Petzet, 2010).

Stone weathering behaviour of exposed samples

Another approach is to quantifying stone weathering samples in exposure trials, deliberately exposed. Such exposure trials have been used to investigate limestone weathering rates mainly in response to air pollution (e.g. Lipfert, 1989; Trudgill et al., 1991; Butlin et al., 1992; O'Brien, 1995; Bonazza et al., 2009; Brimblecombe and Grossi, 2009). This approach poses similar scale problems as laboratory experiments. Indeed, the use of small samples (e.g. 50 x 8 x8 mm for samples from the National Materials Exposure Program (NMEP) (Butlin et al. , 1992)) complicates the upscaling of results to meaningfully larger scales, such as built heritage (e.g. Bell, 1993; Trudgill and Viles, 1998). In addition, the temporal scale of investigation of many previous weathering studies on stone in cultural heritage is often limited to at most a few years (e.g. Butlin et al. 1992; McIlroy de la Rosa et al., 2014). Upscaling from short-term weathering information to long-term weathering behaviour has proved to be difficult due to unaccounted effects influencing the weathering-stress history of the stone like potential extreme weather events etc. (e.g. Bell, 1993; Warke et al., 2003; Moroni and Pitzurra, 2008; Inkpen et al., 2012a).

Focus on erosion only

To date the majority of limestone weathering rate studies have an overwhelming focus on the erosion (mass loss) of stone. However, erosion is merely understood the final step in a series of decay mechanisms preceding this loss stage, such as surface hardening (redeposition of solutional products) or softening (induced by both climate and biological activity), which lead to stone surface property alterations including increased porosity and the formation of superficial layers (Pope et al., 2002; Hoke and Turcotte, 2004; Smith and Viles 2006; Inkpen et al., 2012b; McIlroy de la Rosa et al., 2014). Therefore, to improve understanding of limestone breakdown the entire weathering trajectory needs quantifying (Wilhelm et al., 2016c).

The same principles apply described as disturbance regimes for process domains in geomorphology, where spatial variability on a multi-scale governs temporal patterns of disturbance defined by magnitude, frequency and duration (Montgomery, 1999; Figure 1.8). Similarly, stone weathering behaviour occurs on a multi-scale with spatial variability, which determines the temporal patterns of 'disturbances' that in turn influences the structure and dynamic of weathering patterns.

Therefore, stone weathering behaviour is ideally investigated under real world conditions on real heritage considering a range of time scales and past environmental conditions.

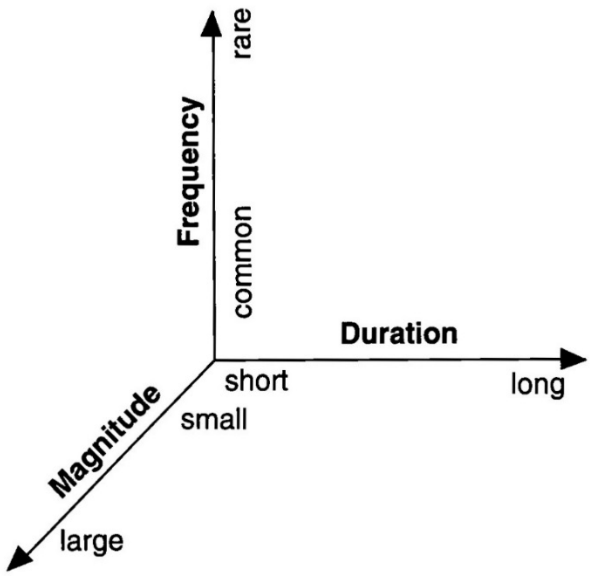


Figure 1.8 Disturbance regimes for geomorphological processes can be defined in terms frequency, magnitude, and duration of associated impacts (Montgomery, 1999)

For stone weathering research and cultural heritage conservation science, there is a need to understand and investigate stone weathering behaviour *in situ* under real world conditions (e.g. Trudgill and Viles, 1998; Doehne and Price, 2010; Moses et al., 2014). Stone conservation scientists and restorers are mostly confronted with stone built heritage with largely unknown weathering-stress history and inhomogeneous weathering patterns (Trudgill and Viles, 1998; Gomez-Heras et al., 2010; Viles, 2013) (Fig. 1.8). Many limestones exposed to the environment exhibit non-linear dynamical weathering behaviour (e.g. Smith and Viles, 2006; Smith and Gomez-Heras et al., 2010). Thus, the spatial complexity and temporal variability of stone decay need to be understood in order to inform appropriate conservation strategies (2010 Svahn, 2006; Auras 2011a; Inkpen et al., 2012b).

1.4. METHODS TO QUANTIFY WEATHERING BEHAVIOUR IN SITU

1.4.1 Long-term time series

In the field of stone heritage previous research has utilized different dated structures to provide a datum point for long-term time series (Moses et al., 2014). Thus, rock art, lead lettering/plugs and cemeteries in general constitute unique repositories for investigating stone weathering behaviour under real world conditions over a variety of timescales (e.g. Cooke et al. 1995; Meierding, 1993b; Inkpen and Jackson, 2000). In addition, these structures also overcome the scale problem discussed above. An example of this approach is the 30-year (1980–2010) investigation of limestone erosion on the balustrade at St Pauls Cathedral in London (Trudgill et al. 1989, 2001; Inkpen et al., 2012a, b). These approaches are limited to structures, which a) provide a datum point and b) require the datum points to not be affected by weathering itself (e.g. Inkpen and Jackson, 2000).

1.4.2 Portable methods

To investigate real heritage structures destructive, minimally invasive and non-destructive methods are applied (an overview of available methods is given in Figure 1.9 and Table 1.1).

Table 1.1 Selection of portable destructive, minimally invasive and non-destructive methods for stone weathering research in situ; NDT=non-destructive technique, MIT=minimally invasive technique (modified after Fitzner, 2002; Auras 2011; Moses, 2014). The methods relevant for this thesis are marked in bold italics.

Method		**Surface**	**Subsurface**	**Operation**	**Expenditure**	**References (selection)**
Monument mapping	Visual assessment, damage classification indices and	x		NDT	low	Fitzner and Heinrichs (2002); Warke et al. (2003);
	Lichenometry					Winchester (1988); Osborn et al. (2015); Upreti et al. (2015);
Surface measurement	Profile measurement	x		NDT	low to moderate	Jaynes and Cooke (1987)
	Roughness measurement	x		NDT	low to moderate	Simon (2001)
	Photogrammetry	x		NDT	high	Bruno et al. (2011); Haubeck and Prinz (2013)
	Laser-optical measuring	x		NDT	high	Meneely et al. (2009); Costanzo et al. (2015)
	Fissure measuring	x	x	NDT	low	Fitzner (2002)
Acoustic methods	Ultrasound velocity	x	x	NDT	moderate	Bellopede and Manfredotti (2006); Fort et al. (2013); Martínez-Martínez (2013);
	Hollow area detector		x	NDT	low	Auras (2011); Meinhardt-Degen (2011, 2012)
Electromagnetic methods	IR-Thermography	x	x	NDT	moderate	Nava et al. (2010); Costanzo et al. (2015)
	Radar (GPR)		x	NDT	moderate	Cosentino et al. (2011); Leucci et al. (2012)
X-ray methods	Raman	x		NDT	high	Gómez-Laserna et al. (2012)
	XRF	x		NDT	high	Prieto-Taboada et al. (2013)
Geoelectric methods	***Resistivity measurement (1D and 2D)***	***x***	***x***	***NDT, MIT***	***low to moderate***	***Mol and Viles (2010); Sass and Viles (2010)***
Water based methods	***Karsten tube***	***x***	***x***	***NDT***	***low***	***Vandevoorde et al. (2012); Hendrickx (2013)***
	Mirowski tube	x	x	NDT	low	Vandevoorde et al. (2012)
	Water drop	x		NDT	low	Bläuer-Böhm et al. (2012)
	Strip off	x		MIT	low	Auras (2011)
Strength testing methods	Drilling resistance	x	x	MIT	moderate	Ferreira et al. (2008); Pamplona et al. (2008)
	Low impact rebound hardness	***x***	***x***	***NDT***	***low to moderate***	***Aoki and Matsukura (2007); Yilmaz (2013)***
	High impact	x	x	NDT,	low to	Viles et al. (2011)

	rebound hardness			MIT	moderate	
	Penetration hardness measurement	x	x	NDT, MIT	low to moderate	Hachinohe et al. (2000)
Bore-hole investigation	Endoscopy		x	MIT	high	Pierce et al. (2011)
Chemical methods	Colouring test	x		NDT	low	Cutler et al. (2013)
	Respiration test (metabolism of biological growth)	x		NDT	low	Warscheid and Braams (2000)
	Heptane drop test (surface charge indicator)	x	x	NDT	low	Bläuer-Böhm et al. (2012)

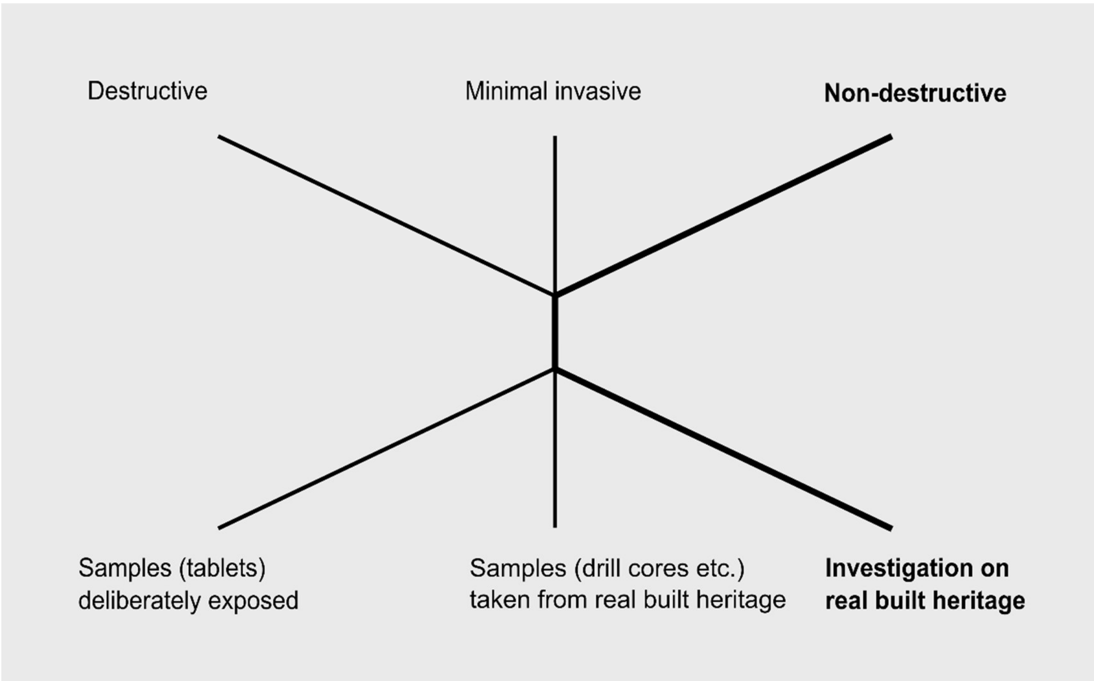


Figure 1.9 Analyses methods and sampling types to investigate stone weathering *in situ*. They can and have been used both individually and in conjunction. The methods and sampling type for this thesis are marked bold.

1.4.3 Non-destructive methods

Costanzo et al. (2015) emphasize that wherever possible non-destructive techniques should be preferred over destructive techniques in order to preserve the integrity of a given historic structure. Thus, this thesis focusses on portable non-destructive methods (marked bold in Figure 1.9). In comparison to destructive methods non-destructive methods can be applied on a larger scale and more frequently as no historic material is damaged or interfered with. Data obtained from non-destructive measurements assist in:

- Detecting stone decay at an early stage and thus inform preventive conservation measures
- Monitoring and control former conservation interventions and therefore, determine a) quality of applied consolidants and b) the frequency of necessary preservation campaigns
- Describing and predicting stone weathering behaviour

Meneely et al. (2009) recommend a 'non-destructive scientific toolkit' including 3D Laser scanning, digital photography, colorimetry, measuring of permeability and ground penetrating radar (GPR), thermography and X-ray fluorescence (XRF) for built heritage preservation. Cataldo et al. (2009) combine ground penetrating radar (GPR), microclimate monitoring and analysis of biological growth to assess the state of preservation and causes for deterioration of Abbey S Salvatore of Montecorona, an important Benedictine monastery of the eleventh century in Northwest Italy. Gómez-Laserna et al. (2012) and Prieto-Taboada et al. (2013) used Raman spectroscopy and X-ray fluorescence (XRF) to investigate the impact of recent air pollution, biological growth and water on stone build heritage. Yet, these sophisticated methods are costly and require special apparatus and expertise, and so cannot easily be used by conservators. Alternatively, Doehne and Price (2010) and Bläuer-Böhm (2012) point out that accessible and economic methods are available which similarly serve purposes of practical built heritage conservation and stone weathering research *in situ*. Auras et al. (2011) published a comprehensive guide on stone monitoring of architectural heritage for sustainable heritage preservation. The guide introduces a wide range of portable non-destructive methods from simple power strip test (similarly used by Maierhofer, 2010), over capillary water uptake with Karsten tube to salt contamination analysis, microscopy and spectrophotometry for colour analysis. The latter technique has also been employed by Cutler et al. (2013), who combined it with surface hardness testing and permeability tests to infer the deterioration status of architectural heritage. The main advantages of these inexpensive non-destructive testing methods are:

- Covering a greater number of heritage assets (including various grades of hierarchy of significance)
- More frequent application
- Straightforward handling and immediate interpretation of results
- Application with greater spatial coverage at built heritage site

Research advancing various non-destructive portable economic methods for on-site application has been conducted by a range of researchers in recent years. Bellopede (2006), Myrin (2006) and Vasconcelos et al. (2007) developed and have all improved *in situ* ultrasound velocity measurements. Bayer et al. (2010) and Eklund et al. (2013) investigated the use of electronic handheld moisture meters to measure moisture content of heritage stone. D'ham et al. (2011), Vandevoorde et al. (2012), Drdácký and Slížková (2013) and Hendrickx (2013) improved measurements of capillary water uptake under low pressure. For the latter a new standard was published in 2013 (BS EN Standard 16302:2013 Conservation of cultural heritage – Test methods – Measurement of water absorption by pipe method). Aoki and Matsukura (2007), Viles et al. (2011) and Yilmaz (2013) have all developed improved surface hardness testing methods for on-site applications. Similar to the ‚scientific toolkit' recommended by Meneely et al. (2009) for more sophisticated methods (e.g. 3D laser scan, ground penetrating radar etc. see section 1.4.3.) it would be beneficial to establish the same for economic methods. This would fill the gap between highly sophisticated and specialised *in situ* studies whilst also providing tools for applied (preventive) conservation.

1.4.4 Missing standards and guides for good practice (Methodology improvements required)

However, quantifying stone weathering behaviour and rates has proved hard especially *in situ* and using portable non-destructive techniques. Some reasons have been already mentioned (i.e. natural variance of stone and unknown weathering-stress history), but limitations also arise from the way methods are applied due to missing standards and guides for good practice and how data analysis is conducted.

Sampling protocols

Despite advances in the field such as a new introduced BS EN standard (16302:2013) for portable water uptake in 2013, there is a need for improved sampling protocols (especially sample sizes) for economic non-destructive methods. At present there is no consensus on methodology for methods like low impact surface hardness testing, portable moisture meters etc. in the field or laboratory, nor in the evaluation of the data obtained (Viles et al., 2011; Yilmaz, 2013).

Sampling size

The number of readings taken has bearing on the meaningfulness of subsequent statistical tests (e.g. for Schmidt Hammer (Niedzielski et al., 2009)). Therefore, a sufficient sampling size needs to be determined for the individual non-destructive methods in order to reliably represent the "population" of weathered stone *in situ* with heterogonous weathering patterns.

Impact of water, porosity and salt on non-destructive testing - confounding effects on non-destructive methods – drawbacks turned into advantages

There is a range of confounding effects on non-destructive methods which are mainly seen as drawbacks as they affect the accuracy of methods. Factors such as inherent material moisture can have a confounding effect on some of the non-destructive techniques used to diagnose stone deterioration. For example, weathering of stone often results in an increase in porosity and in turn results in lower ultrasound velocities when dry and higher when wet (up to 35% for limestone) as the ultrasound waves now travel through a higher proportion of air or water in the pores, which reduces or increases the velocity (travel velocity of sound through air ~ 340 m/s and water ~ 1500 m/s (Bourges, 2006; Simon, 2001).

Furthermore, non-destructive methods based on electrical currents and fields like a range of moisture meters can be influenced by the mineralogy, homogeneity and density of the measured material, temperature and moisture distribution within the material, the presence of contaminants (e.g. salt), the application pressure used, as well as the type of measuring voltage or frequency (Arendt and Seele, 2000; Martinez and Byrnes, 2001; Eklund et al., 2013). Of these factors, one of the most important is the presence of salts – which are nearly ubiquitous in historic buildings and structures (Wilhelm et al., 2016b). Only a few studies correlated the effect of present salts on moisture meters (e.g. Bayer et al., 2010; Schuh et al., 2011). However, despite these publications being in German (and thus not really accessible) the studies still consider the effect of salt on the methods as drawback. This thesis made the attempt to convert this drawback into an advantage by testing whether these meters could actually be used to detect salt and diagnose, reliably, both moisture and salt problems in heritage stone.

Furthermore, operator variance always needs to be considered (e.g. Moore et al., 1989; Viles et al., 2011; Eklund et al., 2013). Finally, the surface condition of the investigated stone may influence the measurement as has been reported for surface hardness measurements (e.g. Aoki and Matsukura, 2008; Feal-Pérez and Blanco-Chao, 2012).

Reliability of generated data

Variability in stone weathering data is a key challenge (e.g. Trudgill et al., 1989; Van de Wall and Ajalu, 1997; Hansen et al., 2013; Alberti et al., 2013). Quantification of stone weathering behaviour is complicated by the inherently high variance observed even in fresh stones and expected to increase for longer weathering-stress histories (e.g. Cooper et al., 1992; Siegesmund and Dürrast, 2010; Fort et al., 2013; McCabe et al., 2015). In addition, data variability has been observed for index non-destructive methods like surface hardness testing (e.g. Alberti et al., 2013; Hansen et al., 2013). Finally, given the requirement to test built heritage *in situ* to understand true stone weathering behaviour with the given sampling limitations requires the data evaluation, which accounts for the experienced data variability and still is reliable and robust. Given these circumstances, there is a need for improved data evaluation methods as already previously stated by Burkinshaw (2002), Svahn (2006) and Viles et al. (2011).

Multidisciplinarity

Costanzo et al. (2015) highlight the need for a multidisciplinary approach when tackling the conservation of architectural heritage. Stone weathering research has gained from other disciplines such as geology, climatology, geo-statistics, engineering etc. Nevertheless, there is a need to exploit the potential of cross- and multidisciplinary approaches to problems in stone weathering research (e.g. McGreevy et al., 2000; Pope et al., 2002; Doehne and Price, 2010; Moses, 2012; Kilic, 2015). For example, Palmer (2008) states, that despite clear advances in understanding the durability of natural stone by geologists and stone technologists, such knowledge is rarely transferred to architects and restorers. Yet, such knowledge transfer with architects and stone restorers observing stone weathering behaviour *in situ* would allow for better linking back of phenomena such as climate change and air pollution to standard durability laboratory tests, in order to improve their accuracy (e.g. Ross and Butlin, 1989; Viles, 2002a; Ingham, 2005; Smith et al., 2011; Viles and Cutler, 2012).

Taking on board all the considerations reviewed above the motivation for this thesis stems from the need to develop reliable methodologies for the collection of reliable data from easy to use and comparatively inexpensive, portable non-destructive testing methods in order to quantify the extent and rate of *in situ* limestone heritage decay under real world conditions over time (time series) with implications for stone conservation strategies.

1.5. Thesis aims, objectives and Research questions

The overall aim of this thesis is to improve selected low-cost non-destructive methods for the diagnosis of deterioration and weathering behaviour of stone built heritage *in situ*. The selection of low-cost non-destructive methods in this thesis was informed by the funding body (Proceq, which produces non-destructive testing equipment like surface hardness testing, ultrasound velocity and moisture measurement) and low-cost, portable devices commonly used by architects and practical conservators like handheld moisture meters and Karsten tubes.

The improvement of these methods contributes to a challenging, but common situation in heritage preservation, where despite disadvantageous conditions cultural heritage deterioration needs to be understood and accordingly preserved i.e. unique heritage site, unknown weathering-stress history, financial constraints, sampling not permitted, rapid decay and thus, high pressure to develop preservation strategies and undertake conservation measures. With the extended, combined and improved application of the selected methods in this thesis stone weathering behaviour can be investigated *in situ* and contribute to theoretical scientific advances.

The study developed from simple to complex in stages as outlined in Figure 1.10. Figure 1.11 shows how the three objectives are linked. The approach taken has been to follow a developmental sequence of evaluation and testing selected portable non-destructive methods under controlled conditions in the laboratory and moving on to more complex field settings to investigate stone weathering behaviour. To make the problem tractable the focus has been narrowed down to limestone.

The three objectives of this thesis divide into two main strands of investigation one laboratory based and the other field based. The first objective improved selected non-destructive methods on fresh porous heritage limestone in the laboratory under controlled conditions for their eventual *in situ* application. Sampling protocols (e.g. sufficient sampling sizes) and reliability of data generated by portable non-destructive methods have been improved by partly converting some drawbacks into advantages like testing handheld moisture meters for salt detection, combining methods and applying modern statistical methods to the data evaluation. This included:

- Determining sufficient sampling sizes to reflect on porous stone characteristics
- Addressing operator variance
- Determining and increasing reliability of results with alternative (modern) statistical methods
- Determining and utilizing the effect of salt content on moisture meter measurements
- Extending method application to gain both surface and subsurface information on stone properties including combination of methods and modern statistics

Objectives 2 and 3 were field based and investigated heritage limestone as time series for short- and long-term weathering behaviour under real world conditions at real heritage sites in the UK and Turkey. Objective 1 informed objectives 2 and 3 and the improved non-destructive methods were applied following the developed guides of good practices *in situ*. The nature of limestone weathering behaviour and deterioration problems *in situ* under real word conditions with varying complexities of weathering stress-histories (increasing from objective 2 to 3) was investigated and rates of limestone weathering behaviour established.

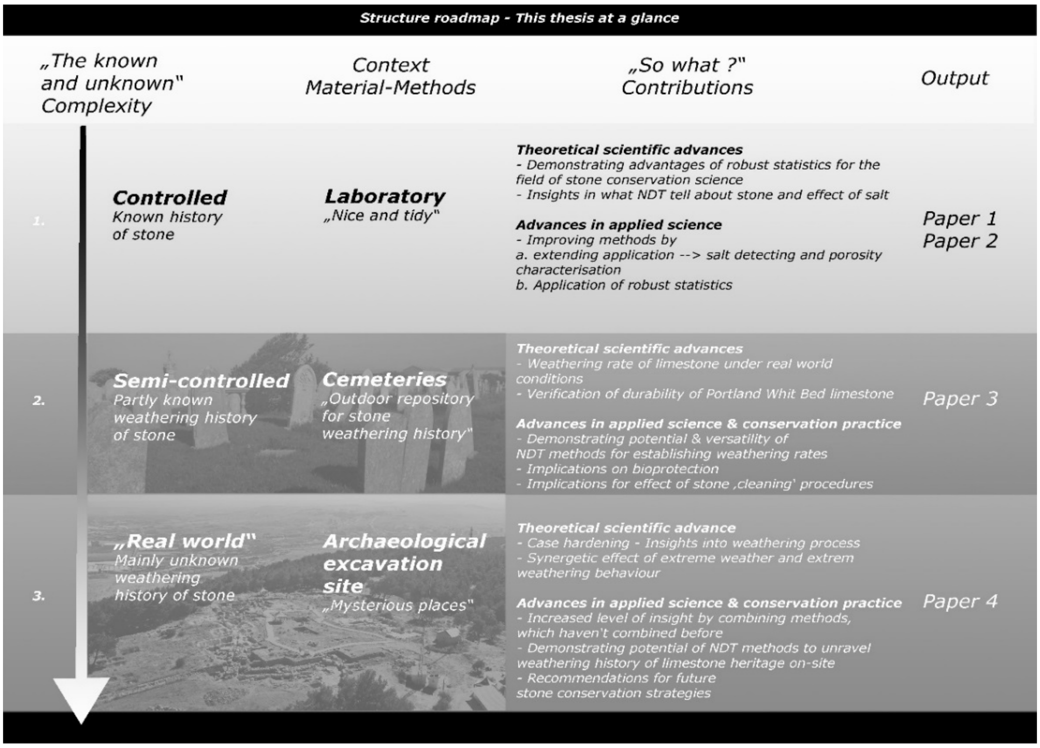


Figure 1.10 Structure "roadmap" – This thesis at a glance. Overview on how the research project developed from 'simple' to 'complex' in three stages (3 objectives). Each context area produced a range of contributions in theoretical science, applied science and conservation practice. The contributions are published in four scientific papers

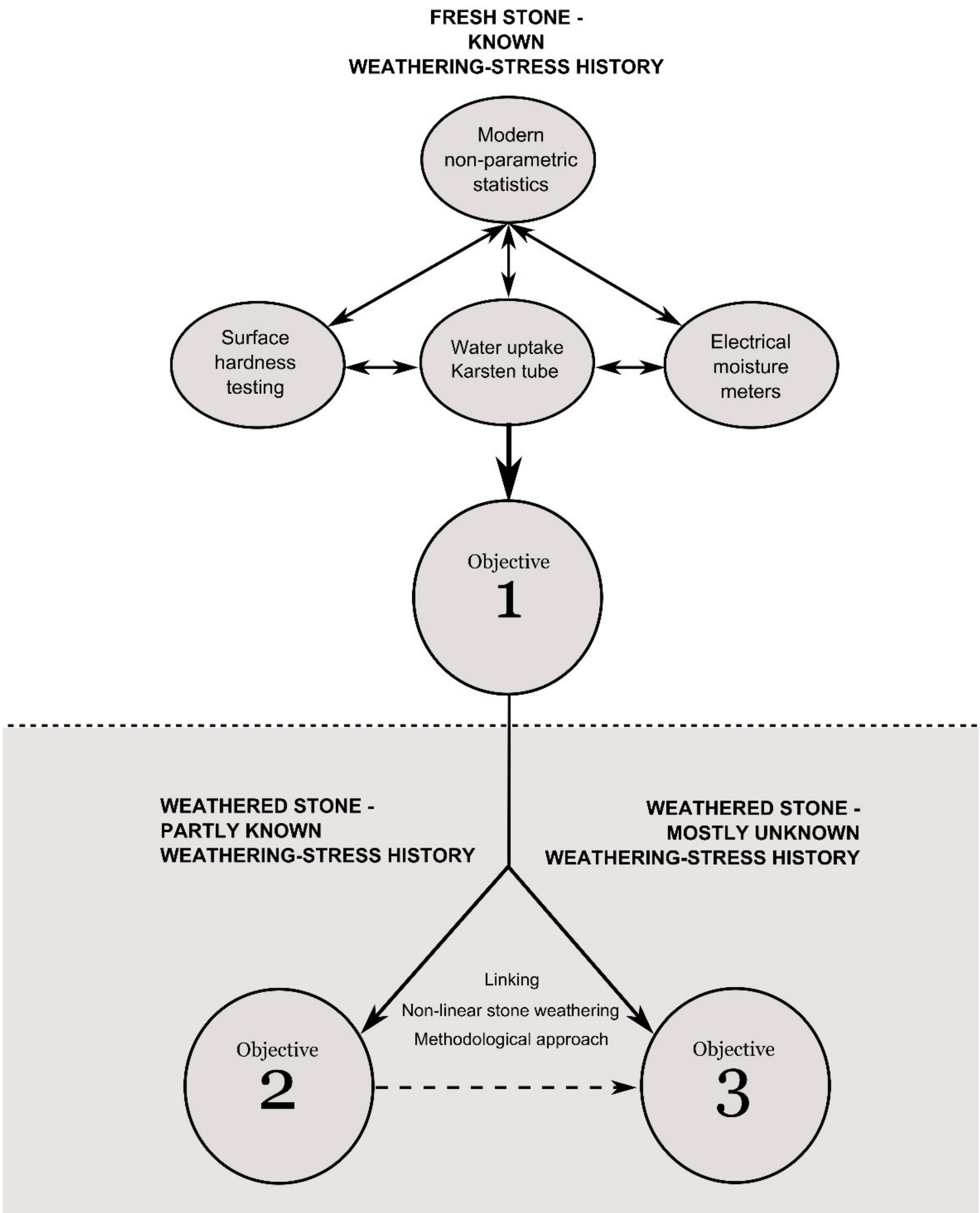


Figure 1.11 Three objectives of this thesis dividing into two main strands of investigation. Objective 1 = Improving selected non-destructive methods on fresh porous heritage limestone in the laboratory under controlled conditions for eventual in situ application, Objectives 2 & 3 = Investigating heritage limestone weathering status as well as time series for short- and long-term weathering behaviour under real world conditions

1.5.1 Objective 1 – Low impact hardness testing and handheld moisture meters – Improving and developing guides for good practice under controlled laboratory conditions on limestone samples

Objective 1 focuses on addressing gaps and problems in the application of two promising sets of non-destructive devices for diagnosis stone deterioration and weathering behaviour i.e. low impact surface hardness testers and handheld electric moisture meters.

To tackle this objective, these two methods have been applied under laboratory conditions to four UK limestones often used in heritage buildings i.e. Portland limestone, Bath (Hartham Park) limestone, Clipsham limestone and Guiting limestone. The effect of stone porosity, moisture and salt content on the data output, which is usually considered a drawback has been used to infer more information about their effect on the non-destructive methods. Surface hardness testing was modified following the approach of Yilmaz (2013). Therefore, not only surface, but also sub-surface information of limestone was gained. Furthermore, appropriate sample size for surface hardness testing was determined using robust statistical data evaluation. For selected handheld electronic moisture meters, the effect of concentrations of sodium chloride as found under real world conditions in porous limestone on the readings was quantified and further used to detect salt (under laboratory conditions). The main findings of objective 1 are the basis of two articles. The results of objective 1 informed the experimental set up of objective 2 and 3.

Paper 1 and research questions

Paper 1 (Objective 1): Low impact surface hardness testing (Equotip) on porous surfaces – advances in methodology with implications for rock weathering and stone deterioration research

Paper 1 has been published in the journal Earth Surface Processes and Landforms The paper addresses the following research questions: How do the Equotip D and DL probes compare? Is the Equotip appropriate for application on porous stone? How to address effects like surface roughness? What are the most appropriate statistical methods to handle Equotip data? How should outliers be treated? And what is an adequate sample size to collect?

Paper 2 and research questions

Paper 2 (Objective 1): The influence of salt on handheld electrical moisture meters: Can they be used to detect salt problems in porous stone?

Paper 2 has been published in The International Journal for Architectural Heritage. The aim of this paper is to shed some light on the influence of salt contamination on selected handheld moisture meters, and to evaluate the potential of these effects to be used to diagnose salt and moisture problems in stone heritage. The paper addressed the following research questions: Are resistance and capacitance mode based moisture meter are equally affected by salt contamination in porous stone? Can the selected moisture meters be utilized to detect salt in porous stone? Can the level of insight be increased by the combination of certain moisture meters?

1.5.2 Objective 2 – Determine deterioration rates of limestone heritage with partly-known weathering history with the improved methods developed in objective 1

Objective 2 focused on applying one of the non-destructive methods improved in objective 1 (i.e. low impact surface hardness testing) to a suite of dated Portland limestone gravestones *in situ* (with partly known weathering history) in order to evaluate the character of changing rate of surface properties. This provides a novel application of surface hardness data for quantifying stone surface changes and deterioration rates.

Paper 3 and research questions

Paper 3 (Objective 2): Surface hardness as a proxy for weathering behaviour of limestone heritage: A case study on dated headstones on the Isle of Portland, UK

Paper 3 has been accepted by the journal Environmental Earth Science. The paper addresses the following main questions: Can stone property change rates of Portland limestone monoliths be developed by means of surface hardness changes? How fast and in what ways deteriorate Portland limestone monoliths over a period of 250 years? Are there spatial differences in deterioration over time? How do results compare to limestone recession rates derived from former studies?

1.5.3 Objective 3 – Diagnosing the cause and nature of catastrophic deterioration of limestone under complex field conditions with the improved methods from objective 1 and 2

Objective 3 focuses on applying a set of non-destructive methods improved in objective 1 to diagnose the nature and causes of catastrophic limestone deterioration observed after a harsh winter at the archaeological site of Dülük

Baba Tepesi, South Turkey. Catastrophic limestone decay after a harsh winter at an archaeological site in South Turkey was reported by collaborating archaeologists. The suddenness and severeness of heritage stone decay determined the focus of objective 3 and the cause for catastrophic stone decay *in situ* were reconstructed using non-destructive measuring techniques and past climate data reports. This provides a novel application to infer the cause of catastrophic decay *in situ* by combing moisture uptake characteristics with robust data evaluation for surface and subsurface and surface hardness data with past meteorological data.

Paper 4 and research questions

Paper 4 (Objective 3): Catastrophic limestone decay at the central sanctuary of Iupiter Dolichenus at Dülük Baba Tepesi in South Turkey: causes and implications for future conservation

Paper 4 has been published in the Journal Conservation and Management of Archaeological Sites. The paper addresses the following research questions: What caused the catastrophic limestone decay? What are the implications for conservation interventions and future site management?

1.6. Thesis structure

This thesis follows the 'thesis by paper' approach with four submitted papers forming the core of the research output. Following this current introductory chapter, the thesis divides into six substantive chapters. Chapter 2 reviews the state of the art of scientific and conservation literature underpinning this thesis. It covers causes and effects of limestone weathering relevant to cultural heritage;

non-destructive methods applied to limestone heritage; and explores appropriate statistical methodologies for data handling and evaluation. Chapter 3 summarises the materials and methods used in the thesis. Chapter 4 starts with a brief linking statement which sets the context for the two papers which both address objective 1. Chapter 5 contains material which addresses objective 2 and comprises paper 3. Chapter 6 addresses objective 3. After a brief linking statement, this chapter is made up of paper 4. Chapter 7 provides a discussion and conclusion, which link together the findings of the four papers into the wider scope of the thesis. The chapter reviews the main contributions made by the thesis, its implications for the fields of stone weathering research and heritage conservation, and makes some recommendations for future research.

2. Literature Review

2.1. Terminology

Knowledge about rock weathering and stone deterioration is generated by, and of interest to, both geomorphologists and those involved with cultural stone heritage conservation (including scientists and conservators). The different fields of research use different terminology with different terms having similar meaning (like deterioration, weathering, surface diagenesis, degradation, decay, stone pathology (Pope et al., 2002)) and sometimes with the same terms meaning different things in the respective field (like rock and stone). This might cause confusion and further impede interdisciplinary approaches in terms of knowledge exchange (e.g. literature, mutual understanding) and, thus preventing transfer of methods (e.g. statistics, weathering behaviour) (e.g. Viles et al., 1997; Price, 2010). Furthermore, Price (2010) points out a lack of common language even within the field of cultural heritage science. Several publications tried to tackle the issue, but there is no standard, which could be applied. Table 2.1 summarizes a range of approaches to address the issue of varying terminology. Usually, limestone in the natural environment is referred to as 'rock' and the processes affecting it are known as 'weathering', whereas in the built environment the terms 'stone' and 'deterioration' are used (Figure 2.1).

Pope et al. (2002) define the term 'cultural stone' as stone having been altered by humans including quarries, rock art, architectural stone and sculptures. This study however suggests to use the term ‘cultural rock’ as the term ‘cultural stone’ is defined by the material having been removed from its original location, which is not the case for rock art. Although this thesis is concerned with cultural stone, the term rock is used when i) the cited study applied the term and ii) aspects and methods can be transferred to rock weathering.

Table 2.1 Overview of publications addressing the issue of terminology in the field of stone weathering research (modified after Price, 2010)

Title	Authors/editors
Glossary of decay terms	Italian Commissione NORMAL (UNI 2006)
Natural Stone Glossary	Stone Federation of Great Britain (1991)
A Glossary of Historic Masonry Deterioration Problems and Preservation Treatments	Grimmer (1984)
Weathering forms at natural stone monuments: Classification, mapping and evaluation	Fitzner, Heinrichs, and Kownatzki (1997)
The ICOMOS-ISCS *Illustrated Glossary on Stone Deterioration Patterns*	Vergès-Belmin (2008)
Building Stone Decay: Observations, Experiments and Modeling	Warke et al. (2003)

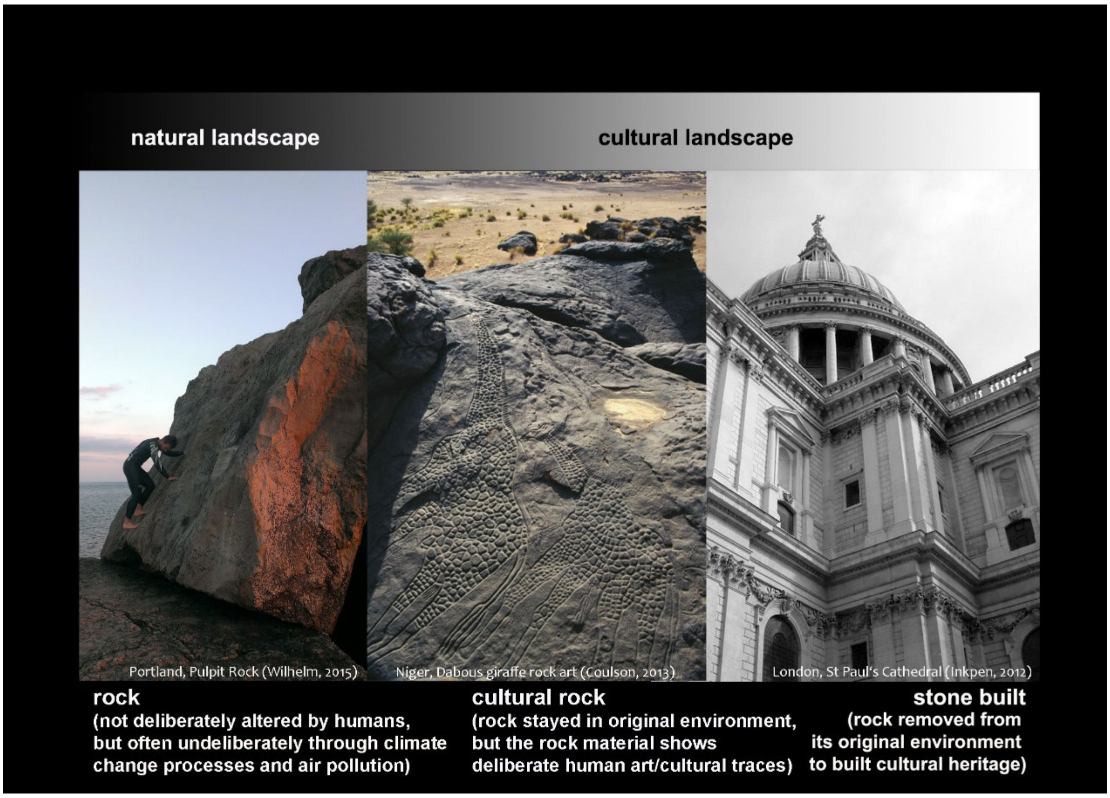


Figure 2.1 Distinction between natural and cultural landscape and the terms for rock and stone

related to the respective landscape

Rock and stone have the same origin as stone is removed from rock in its original (natural) place. Thus, both have in common the initial intrinsic stone properties determining weathering behaviour (i.e. high microporosity in Portland Base Bed will increase water retention for both a rock and building stone and thus increase chances for accelerated decay). However, their weathering-stress history often shows significant differences and results in different weathering behaviour on a range of scales. The first divide in weathering-stress history occurs when stone is quarried. Removing stone from its natural formation puts it in an instable place from which it will always thrive to reach equilibrium again (Přikryl, 2013). Furthermore, the geometry of the worked stone differs from the shape of a natural outcrop, which again results in different weathering patterns.

This puts limitations on transferring methods and linking knowledge between the fields of geomorphology and heritage stone weathering research. Manipulating the stone surface is "zeroing" the weathering clock, which may provide the baseline to investigate alterations since (Pope et al., 2002). Hoke and Turcotte (2004) observed a lag of decay provided by polished surface of marble tomb stones. Furthermore, built heritage structures often exhibit a mixture of materials (joint mortar, different stone types be it for the artistic intention or replacement material). Finally, cultural stone gets different attention and is exposed differently (higher exposure pressure due to human activity (walking, touching sculptures and using architectural structures). Consequently, built heritage structures are subject to conservation measures which sometimes may have an adverse effect or introduce new material etc. This results in a rather complex weathering stress history and cannot be compared to natural limestone outcrops and thus, landform development.

2.2. Nature, Causes and controls for limestone weathering behaviour

The core research interest of this thesis is on non-destructive methods to investigate limestone weathering behaviour of built heritage *in situ*. Data generation i*n situ* with non-destructive methods is not without difficulties due to a range of impacts on the measuring procedures, which will be discussed in this chapter. Further, common threats to built limestone heritage are introduced as well as factors promoting limestone decay processes and the significance for quantifying weathering behaviour under real world conditions.

Limestone weathering behaviour is complex and so is the development of weathering rates. Only a thorough understanding of factors involved will allow for the right conclusions to be drawn and appropriate preservation measures to be undertaken (Přikryl, 2013). Thus, the literature review starts with introducing the mechanisms for limestone decay. This is followed by a discussion of common approaches to investigate limestone weathering *in situ* using non-destructive methods. The chapter concludes with a review on common and novel data evaluation methods.

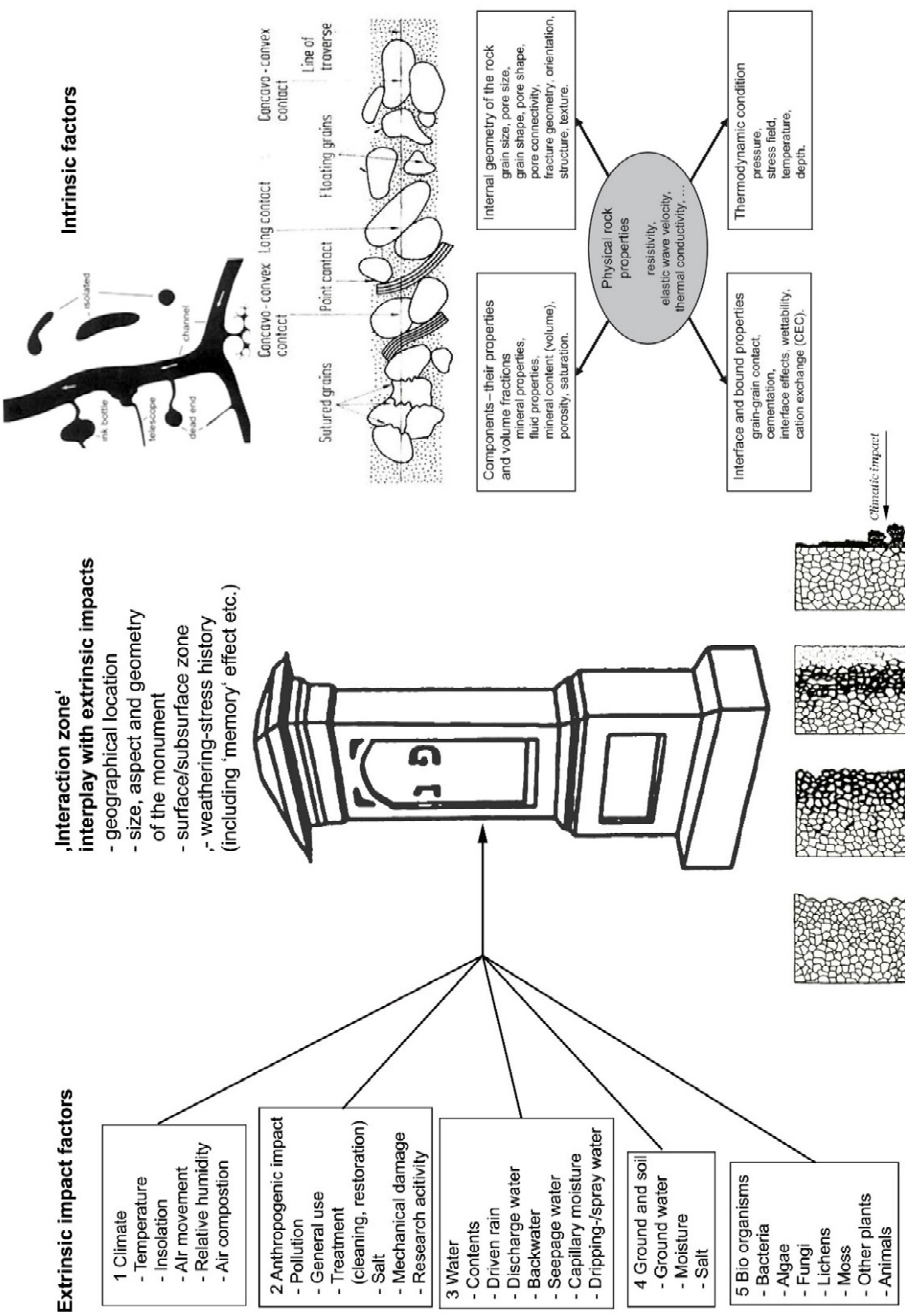


Figure 2.2 Overview of extrinsic and intrinsic factors which are relevant for stone deterioration and decay (modified after Grimm, 2010; Bourges, 2006; Schön, 2011). On the left 5 groups of extrinsic impact factors are listed. The middle displays the 'interaction zone' where intrinsic and extrinsic impacts interact. The right shows intrinsic factors i.e. stone properties such as porosity characteristics, mineral composition, grain characteristics, intergranular bonds (cementation), thermodynamic condition (e.g. pressure, temperature etc.).

Figure 2.2 provides an overview of the extrinsic (climate, anthropogenic impact, water, salt and aquaisms) and intrinsic (stone properties, monument characteristics and inheritance from the past like accumulation of air pollutants ('memory' effect) and mechanical stresses (e.g. crack propagation) which control stone deterioration. It is the interplay of extrinsic and intrinsic factors (or force and resistance) which controls both the nature and rate of weathering/deterioration.

2.2.1 Intrinsic factors affecting stone weathering – Limestone properties

Geology

May (1998) states that in order to understand the weathering behaviour of limestone good geological knowledge is indispensable. Geological aspects of importance are age, geochemistry, mineralogy, petrology and structural characteristics.

Distinguishing beds

The importance of distinguishing beds of the same rock formation is demonstrated with the Portland limestone formation (Figure 2.3). The two Portland limestone beds (varieties) most relevant for built heritage are Portland Base Bed and Portland Whit Bed. The latter has been found to be more durable and weathering rates are given in the literature of 1 - 2 mm surface recession per 100 years (with greater extent under severe exposures) for the UK climate (Leary, 1983; Building Research Establishment, 1997a,b; Viles et al., 2002a; Dubelaar et al., 2003; Godden, 2012). In contrast, Portland Base Bed has a higher weathering rate of 3 - 4 mm surface recession per 100 years (with greater extent under severe exposures or on the edges of stonework) (Building Research Establishment, 1997a,b).

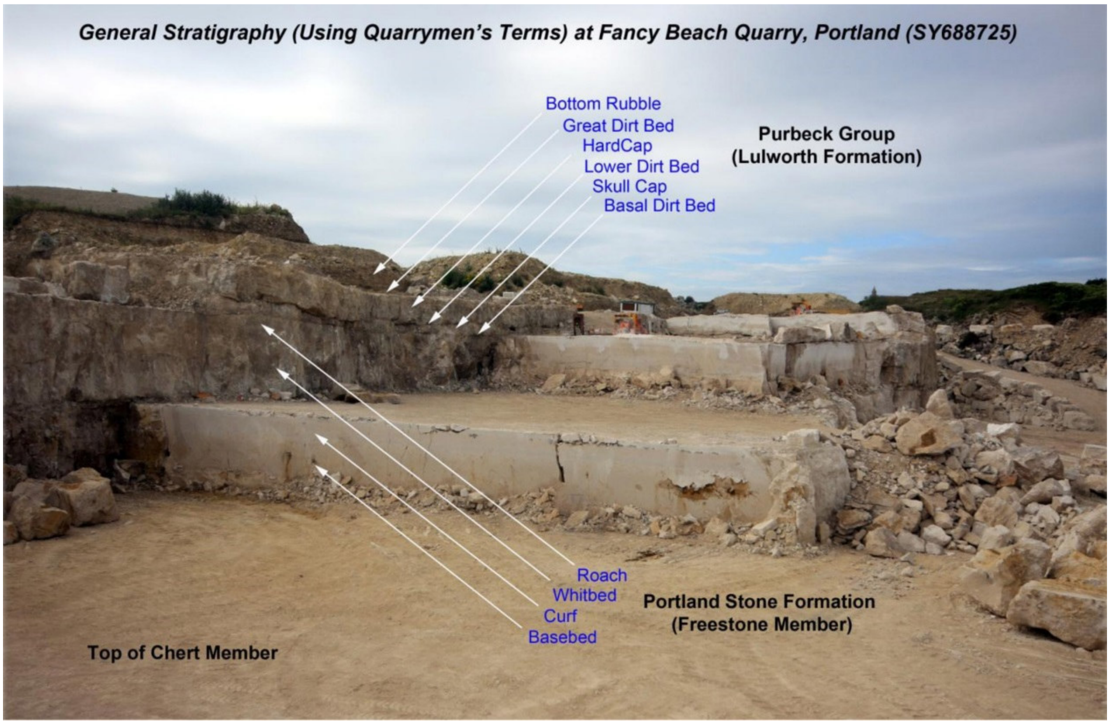


Figure 2.3 Portland limestone beds, general stratigraphy (using quarrymen's terms) at Fancy Beach Quarry, Portland (UK, SY688725; source: Godden, 2012, p.8)

Despite these crucial differences affecting durability and decay the two varieties are rarely distinguished in the literature on weathering and built heritage. The two varieties are also used interchangeably in many buildings. A mixed usage for both varieties is reported where on the one hand Portland Base Bed has been favoured over Whit Bed for aesthetic and workability reasons (the lack of visible shell content), but on the other hand confusion in nomenclature ("Best Bed" for Base Bed) and order phrasing ("Whit Bed without shells") might have caused mixed usage of both Portland Base Bed and Whit Bed in the past (Gray, 1861-1862; Edmunds and Schaffer, 1932). Figure 2.4 shows an example for natural variability in Portland limestone formation. Such variability within and between beds of the same limestone type is found in many other examples.

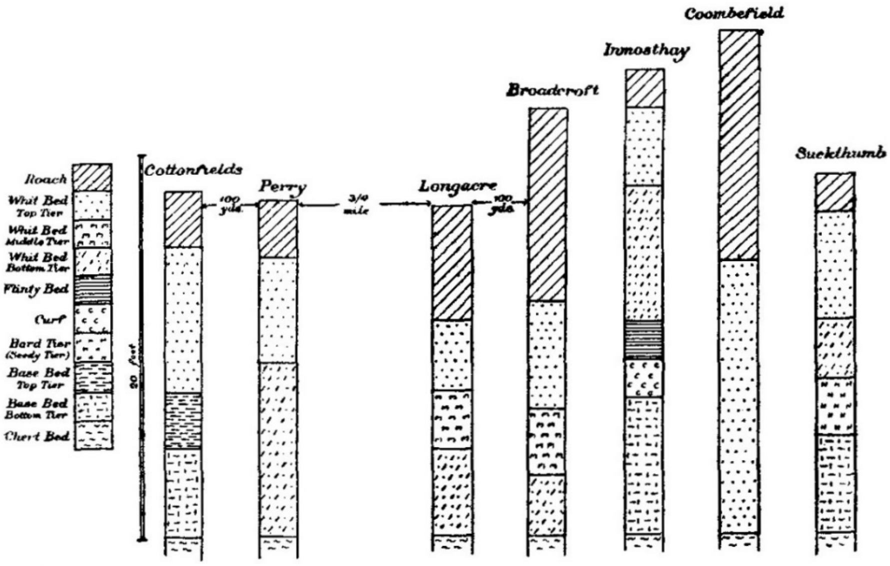


Figure 2.4 Complex comparative sections showing lithological variation in Portland limestone. From top to bottom: Roach, Whit Bed Top Tier, Whit Bed Middle Tier, Whit Bed Bottom Tier, Fancy Bed, Curf, Hard Tier, Base Bed Top Tier, Base Bed Bottom Tier, Chert Bed. From left to right: Cottonfields, Perry, Longacre, Broadcroft, Inmosthay, Coombefield, Suckthumb (source: Edmunds and Schaffer, 1932, p. 231)

Classification limestone

Mosch and Siegesmund (2007) distinguish limestone, oolitic limestone, lime breccia, travertine and dolomite of which oolitic limestone are of interest for this study as discussed in chapter 3. Limestone is characterised by its composition, texture, apparent density (or 'raw' density including pore space) and porosity.

Composition and texture

Folk (1959) and Dunham (1962) present the most common classification systems for carbonate rock based on textural (grain) properties (e.g. shells, crystals, allochems), and grain-support (sparite and micritic cement). Figure 2.5 shows Folk' and Dunham's modified carbonate classification by Embrey and Klovan 1971. Folk's system is based on a multistage division: 1st presence or lack of allochem (either fossils bioclast, oolids or ooids, pellets an intraclasts (Ahmad, 2011)) and

2nd whether these allochems embeded in a micritic or sparite cement matrix (Figure 2.5).

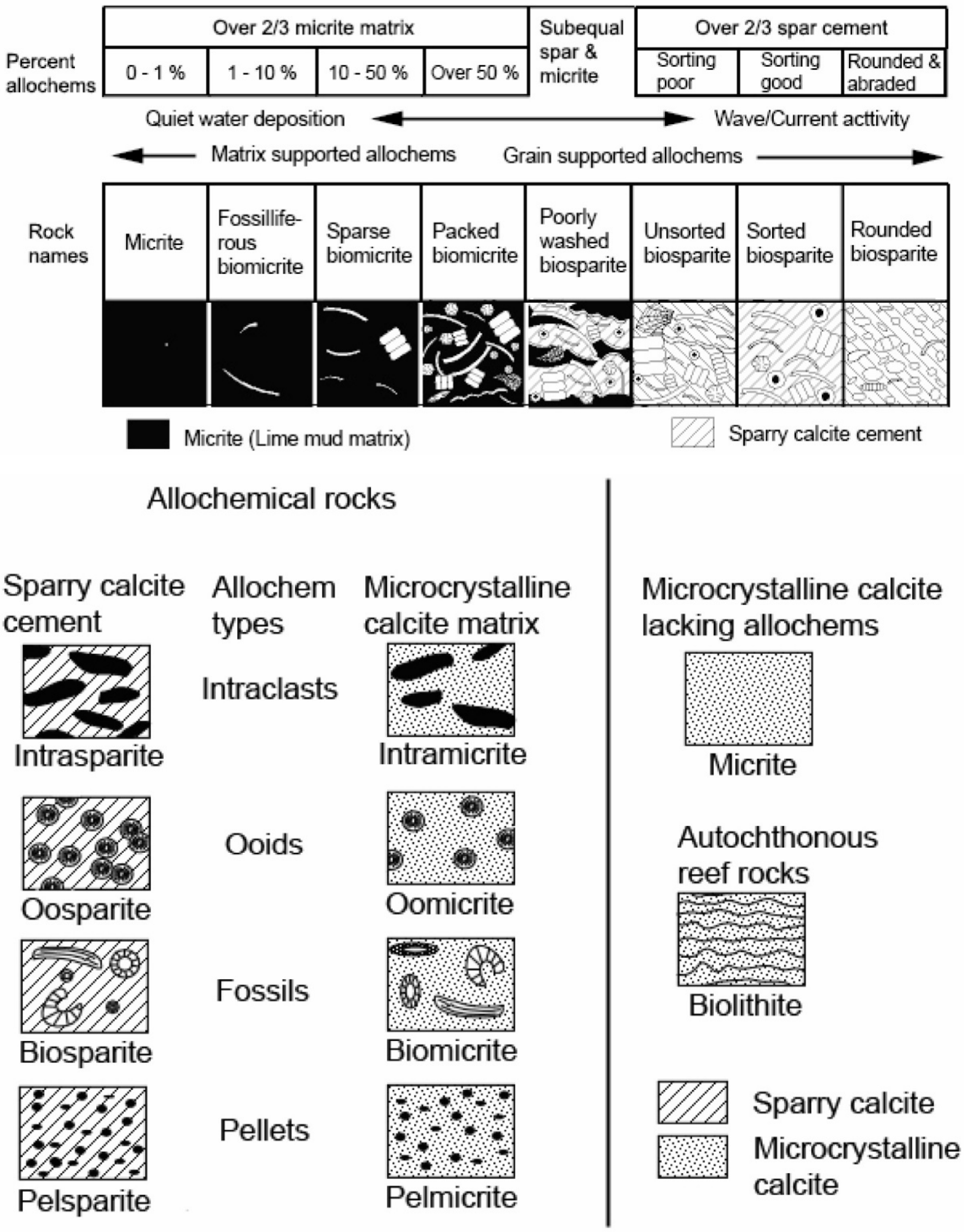


Figure 2.5 Folk and Dunham carbonate (modified after Dunham, 1962 and Folk 1959 by Embrey and Klovan 1971)

Dunham system describes the depositional, which may or may not be recognizable and further diagenesis processes where binding either occurs during deposition or not. Coarse-grained limestone can be more resistant to weathering than fine-grained (micritic) (Emmanuel and Levenson, 2014). Mosch and Siegesmund (2007) further establish a sub-categorization depending on limestone density. They found in a meta-analysis of stone property data for 2100 building stones that correlation coefficients between density and strength are improved, when dividing the limestone group into two subgroups according to their average density, < 2.6g/cm³ and > 2.6g/cm³. Improving these correlations improves further correlations of non-destructive testing results to unconfined compressive strength established in a range of studies for a range of stone types (Table 2.2).

Table 2.2 Selection of studies which correlated unconfined compressive strength testing to surface hardness testing. UCS = unconfined compressive strength (destructive), BS= bending strength (destructive), PLI = point load strength index (destructive), SHT = surface hardness testing (non-destructive), UPV =ultrasonic pulse velocity (non-destructive)

Study	Tested stone types	Testing methods
Aliabdo et al. 2012	Marble, limestone, basalt, bricks	SHT, UCS, UPV
Aydin and Basu 2005	Granite	SHT, UCS
Bruno et al. 2013	Limestone	SHT, UCS
Güney et al. 2005	Limestone, Travertine, Marble	SHT, UCS, UPV, BS, PLI
Alvarez Grima and Babuška (1999)	Sandstones, limestones, dolomitic limestones, dolomites, granites and granodiorites	SHT, UCS
Aoki and Matsukura (2008)	Tuff, sandstone, granite, andesite, gabbro, limestone	SHT, UCS
Yilmaz (2013)	Limestone, marble, travertine, dolomite	SHT, UCS

Figure 2.6 shows boxplots for bulk density, effective porosity and water absorption (under atmosphere). The limestones interesting to this thesis (shaded grey) show a wide asymmetric spread for all three measured parameters. Thus, it

is expected that weathering behaviour is equally ‘asymmetric’ (i.e. unpredictable) and thus, care needs to be taken when transferring/comparing weathering behaviour between studies. So, for example for Portland limestone (Jordans Base bed) unconfined compressive strengths show a huge data variety for example in Albion’s Quarry report 56 (2012) with min: 28.49MPa, max: 57.63 MPa, ave: 41.15 MPa compared to this study with min: 43.2 MPa, max: 75.73 MPa, ave: 55.98 MPa. The same is true for porosity again in Albion’s Quarry report 56 (2012) with min: 17.59%, max: 20.67%, ave: 19.08 compared to this study with min: 13.12%, max: 13.82, ave: 13.47%. Nevertheless, the BGS derives a weathering rate of 1-2 mm per 100years which may vary with varying climatic conditions. Yet, the variability of the stone itself is not taken into account. Given that porosity and composition are determining weathering behaviour, their high variance poses difficulties on predictions of weathering rates.

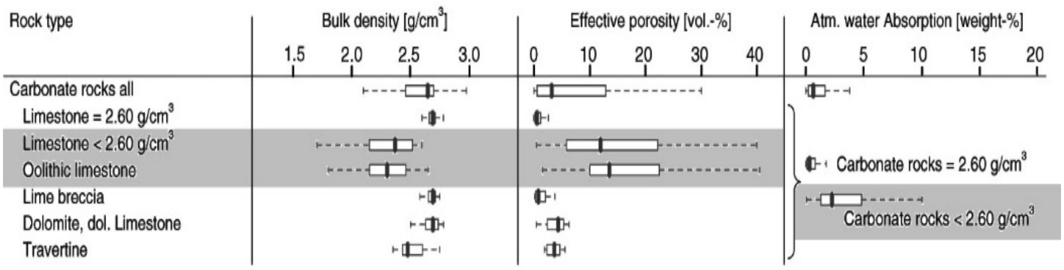


Figure 2.6 Value ranges for limestone bulk density, effective porosity and water absorption under low pressure (WAAP) (modified after Siegesmund and Dürrast, 2010) limestone types relevant for this study (density <2.6 g/cm³ and oolitic limestone) are marked grey.

Porosity

Porosity is commonly given in percentage and summarizes the fraction of stone bulk volume, which is pore space. That includes all pores, fractures, cracks and fissures (Schön, 2011). Limestone is a sedimentary stone derived from biological deposit like fossil fragments and other particles with varying complex morphology resulting in complex porosity (Schön, 2015). This complexity is further increased over scales by aspects of weathering stress history like dissolution, reprecipitation and chemical alterations (Schön, 2011). Siegesmund and Dürrast (2010) state that due to the complexity of pore shapes it is difficult to establish a consistent pore classification system.

The first important distinction is between noneffective and effective porosity (also 'active', 'apparent' or 'open' porosity; accessible for water under low/atmospheric pressure). Only effective porosity is relevant for stone weathering behaviour research as it determines internal water regimes. Effective porosity has been identified as a key factor in deterioration. It functions as a transport system for water, which is a key deteriorative agent (e.g. Poschlod 1990; Nicholson, 2001; Franzen and Mirwald, 2004; Meinhardt-Degen, 2005; Mosch and Siegesmund, 2007; Palmer, 2008). The second distinction is between primary and secondary porosity dependent whether the voids formed during deposition or through diagenetic processes (Tucker and Wright 1990; Fitzner and Basten, 1994; Schön, 2015).

The interconnectedness of pores is a further determining characteristic for stone weathering behaviour. Pore space characteristics can be divided in pore space between grains (interparticle porosity) and all other pore space (this might be dissolved grains, fossil chambers, biological skeletons (Fitzner and Basten, 1994; Palmer, 2008). Figures 2.7 and 2.8 give an overview of different pore shapes and connectivity.

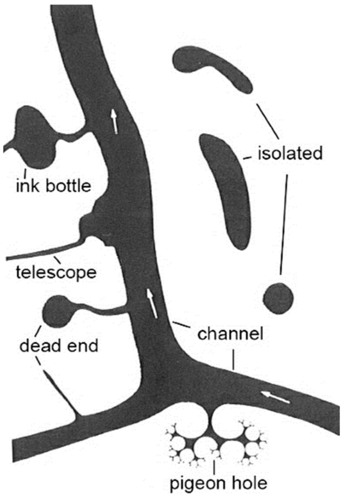


Figure 2.7 Pore types (Fitzner and Basten, 1994).

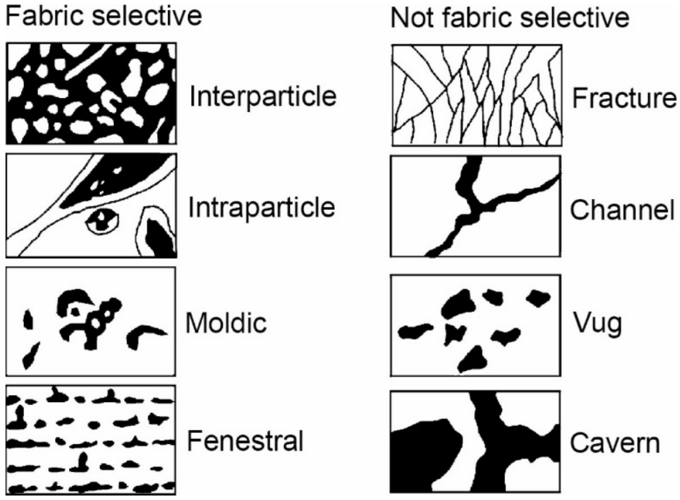


Figure 2.8 Porosity types (Choquette and Pray, 1970 in Tuğrul, 2004).

Table 2.3 shows common pore size classifications with micro-, meso- and macro pores (where the diameter is based on a simple model of a pore as a cylinder). A simplified classification of limestone according to the amount of (open) pores is given in Siegesmund et al. (2010) in terms of non-porous or compact limestone, porous limestone and travertine. They further present a classification scheme from Moos and Quervain (1948) with < 1% compact, 1- 2.5% a few pores, 2.5-5% slightly porous, 5- 10% significantly porous, 10-20% many pores, and more than 20% significantly high amount of pore space.

Table 2.3 Selected pore size classifications (r = radius). Note the differences in ranges. (modified after Siegesmund and Dürrast, 2010)

	Micro pores (µm)	'Microcapillary active' pores (µm)	'Capillary active' pores (µm)	Mesopores (µm)	Macropores (µm)	Large pores (µm)
Klopfer (1985)	<0.1	n.a.	0.1–1,000	n.a.	>1,000	n.a.
Ahmad (2011)	<0.1	0.1<r<5	5<r<1,000	n.a.	>1,000	n.a.
Quervain (1967)	<5	n.a.	n.a.	5-200	200–2,000	>2,000

Water penetrates a structure either as liquid or vapour controlled by capillary forces, diffusion processes, flow and hydrostatic pressures (e.g. Charola, 2000; Ahmad, 2011) as illustrated in Figure 2.9.

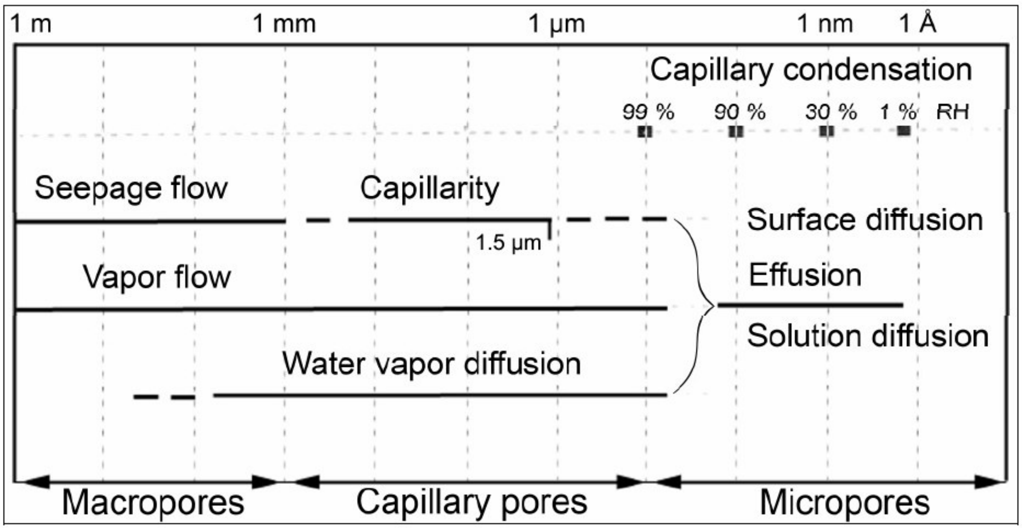


Figure 2.9 Water transport mechanisms depending on pore size (source: Ahmad, 2011; Snethlage, 1984 and Neisel, 1995; after Klopfer 1979, 1980)

Changes in porosity are used as a descriptor for weathered and unweathered stone (Bernal and López, 2000). A study for the relationship between porosity, saturation coefficient and durability by Honeybourne and Harris (1958) redrawn by Palmer (2008) suggests although no simple tendency for increase of weathering susceptibility with porosity increase, a marked tendency for weathering susceptibility with an increase of small pores within the overall porosity (Palmer 2008). Similarly, Dubelaar et al. (2003) and Yu and Oguchi (2010) found in durability tests (freeze-thaw and salt crystallisation) that microporosity relates to lower durability. A decrease in porosity (and superficial increased increase of density) occurs during crust forming and case hardening processes, caused by dissolution and re-precipitation of calcite cements (Smith and Viles, 2006; Hendrickx, 2013). The stone behaves as a two material system (composite) as the indurated surface now has different physical properties compared to the subsurface and core of the stone (which may be further softened). Such surface crusts have been associated with episodic catastrophic deterioration on vulnerable limestones, often in combination with freeze-thaw weathering induced by harsh winter conditions (e.g. Smith and Viles, 2006; Martínez-Martínez et al., 2013).

Under outdoor exposure pore space is changed over time due to chemical, biological and physical impact. The changes might either lead to a decrease or an increase on overall effective porosity and/or alter the shape and connectivity and thus initiate and promote different weathering behaviour. These effects take place in synergy with environmental conditions and are introduced in section 2.2.2.

Of particular interest for this study is microporosity (<5 μm), which has been documented to increase water retention and plays a crucial role in limestone weathering behaviour as it correlates with poor durability (Palmer, 2008). Capillary force results in higher water retention and thus may result in a) reduction of strength (e.g. reported by Çanakci (2007) for collapsed limestone caves in Gaziantep) and b) reduces resistance to weathering impacts like crystallisation-dissolution processes through salt and frost.

Chemical mechanisms causing the dissolution/reprecipitation of limestone

As a sedimentary stone limestone is defined as consisting of at least 50% calcium carbonate ($CaCO_3$) and other stone minerals like chert, quartz, sand, silt and clay, etc. (Missouri Department of Natural Resources, Division of Geology and Land Survey, 2011). Compared to sandstone (clastic sediment) limestone is chemically instable (Schön, 2011). The high amount of $CaCO_3$ renders it particularly prone to acidic chemical attack induced by air pollution, and dissolution/induration processes are common weathering processes (Charola and Ware 2002; Mitchell and Searle 2004; Brimblecombe and Grossi 2009).

$CaCO_3$ reacts with acids by releasing CO_2 (Equation 1):

$$CaCO_3 + 2H^+ \rightarrow Ca_2+ + CO_2 + H_2O \quad \text{(Equation 1)}$$

Pure rain water is already slightly acidic with a pH of around 5.6 (as usually atmospheric CO_2 is dissolved in it (Benedix, 1999), with increased CO_2 rain water acidity increases (Vallero, 2008). The involved chemical dissolution process can be described as follows (Equation 2):

$$CaCO_3 + CO_2 + H_2O \rightarrow Ca_2^+ + 2HCO_3 \quad \text{(Equation 2)}$$

The $CaCO_3$ is being transformed into calcium bicarbonate $Ca(HCO_3)_2$, which is far more soluble in water compared to the initial calcium carbonate (0.013 g/L (25 °C) versus 166 g/L (20°C)(Tegethoff, 2001; Vallero, 2008). Dissolution may on the one hand increase porosity within the stone matrix (e.g. Camuffo, 1995; Çanakci, 2007); on the other hand promote the formation of superficial layers with a higher density (Hoke and Turcotte, 2004; Schmidt and Viles 2006; Inkpen et al., 2012b). $Ca(HCO_3)_2$ only exists in water, thus after evaporation $CaCO_3$ reprecipitates and may fill pores or forms (white) crusts depending on the water regime (i.e. internal moisture movement vs run-off)(e.g. Camuffo, 1995; Török, 2003) (Equation 3).

$$Ca_2^+ + Ca(OH)_2 + 2HCO_3^- \rightarrow 2CaCO_3 + 2H_2O \quad \text{(Equation 3)}$$

The above equations are only part of a range of chemical processes apparent under outdoor conditions. Further chemical interactions for example with air pollutants like sulphur dioxide (SO_2) or microbiological activity may take place at the same time and location and are discussed further in section 2.2.3.

2.2.2 Extrinsic factors and impacts

Anthropogenic impact

The current period of the Anthropocene poses a range of pressures and threats to cultural stone heritage due to human activity. Mechanical damage caused by general usage of heritage structures are common in historic public buildings, churches and tourist sites (e.g. Duval and Smith, 2013; Cámara et al., 2014). For Al-Khazneh, Petra (Jordan) Pope et al. (2002) related surface recession at a level of 1.50 to 2.00 m above ground to mechanical abrasion from site visitors. Preservation measure e.g. cleaning (Camuffo, 1995) and consolidation may have a detrimental effect (Gomez-Heras and McCabe, 2015). Research activity like archaeological excavation creates vulnerable, exposed heritage sites (Warke et al., 2010). Although the Convention of Malta (ratified by countries like the UK, Germany and Turkey) states in Article 3: "To preserve the archaeological heritage and guarantee the scientific significance of archaeological research work, each Party undertakes: i. to apply procedures for the authorisation and supervision of excavation and other archaeological activities in such a way as: b. – the elements of the archaeological heritage are not uncovered or left exposed during or after excavation without provision being made for their proper preservation, conservation and management; [...]"(Council of Europe, 1992, p. 2), in practice archaeological remains are often exposed to the elements and pollution and prone to deterioration and therefore in need of conservation measures (Jackson et al., 2005; Bonazza et al., 2007; Al-Houdalieh, 2009; Warke et al., 2010; Tapete et al., 2012).

Much importance is given to threats of increased air pollution and climate change to cultural heritage (e.g. Brimblecombe and Grossi, 2009; Doehne and Price, 2010; Ruddiman, 2010; Zalasiewicz et al., 2011; Fuente et al., 2013; Howard, 2013). The potential risks of climate change to cultural heritage in Europe are summarized in 'The atlas of climate change impact on European cultural heritage' (a result from The Noah's Ark Project) (Sabbioni et al., 2010) (Table 2.4).

Table 2.4 Summary of predicted/modelled climate change impacts on European cultural heritage (Sabbioni et al., 2010)

Increase of annual salt crystallisation frequency over next century
Reduction of wet-frost (wet frost index = the number of rainy days where T > 0°C, followed immediately by days with a mean temperature below -1°C in a year) for most of Europe (exception may apply for "Northern Europe and some areas of European landmass, such as Russia may experience more frequent wet-frost")
No excessive effects of climate change on biomass stock (this only applies to horizontal surfaces of hard acid stones in non urban environments) with exceptions for boreal areas such as North of Russia, Scandinavia or Scotland where a great increase of biomass is expected
Lichen species richness: with increase in temperature a decline in richness is expected and has effects on biodiversity
Surface recession of low porosity carbonate stones: karst effect expected to be the dominant weathering process with a general risk increase of max 6μm $year^{-1}$
Thermoclastism (microcracking and exfoliation of stone): the Mediterranean Basin at highest risk including Central Europe in the near and far future. Major impact is expected in the far future, with a max of 200 events/a as forecast for Southern Spain and Greece
Clay containing materials: a 100% increased risk of damage is predicted for clay containing sandstones in the northern part of Europe (due to less frost and increased precipitation)

The patterns of deterioration are likely to shift, with some threats reducing (e.g. freeze-thaw) in some areas and increasing in others. However, considering changes of rainfall and temperature including extreme events, it is quite likely that the overall threats to cultural heritage in Europe will rise (Smith et al., 2011a). Smith et al. (2010) describe the decrease of atmospheric sulphur dioxide (SO_2) in cities as an "ironic" (Smith et al. 2010, p.8) shift in threats, which although perceived as positive development might lead to increased stone deterioration by limiting superficial crust-forming which has the potential to (temporarily) affect the long-term stability of built heritage structures as for example Zehnder (1996) observed for wall paintings (quoted in Charola et al., 2007).

Concerns over the deteriorative effect of air pollution has led to numerous weathering studies with a particular focus on the interaction of sulphur dioxide (SO_2) levels and limestone surface recession rates (e.g. Jaynes and Cooke 1987, Butlin et al. 1992, Fuente et al. 2013). Although recent decrease of SO_2 levels in urban areas limit the effect (Bonazza et al., 2009), Kucera (2002) describes an even more complex scenario with a multi-pollutant situation (synergistic effects of ozone (O_3), nitrogen compounds (NO_x) leading to an increased acidity of precipitation) resulting again in limestone surface recession and other forms of decay.

However, Bonazza et al. (2009) and Brimblecombe and Grossi (2009) suggest, that pollution controlled damage overall is decreasing and point the focus towards the impact of climate change on stone weathering. Changes to air quality and climate change are likely to promote biological growth, whose protective or destructive character still needs to be investigated (Viles and Cutler 2012, Cutler et al. 2013). Further, an increased frequency of salt crystallisation processes will lead to an increase of salt weathering processes (Sabbioni et al., 2010).

Weather and climate

The three main climatic factors inducing limestone weathering and promoting decay are thermal impact, wind and water in form of snow and rain as well as relative humidity. Limestone dissolution is the process whereby $CaCO_3$ is congruently dissolved in water acidified by CO_2 (the underlying chemistry is described in detail in section 2.2.1). This process of dissolution is responsible for the production of karst landscapes, and also affects limestone built heritage. The rate of dissolution is thereby related to the effective precipitation (precipitation minus evatranspiration).

Limestone is also affected by freeze-thaw and heating and cooling, which can lead to physical breakdown such as fatigue, granualar disintegration and cracks (Hall, 2004; Ingham, 2005; Yavuz, 2006; Smith et al., 2011b; Martínez-Martínez et al., 2013). Thermal regimes have an influence on water regimes and in turn weathering behaviour. Camuffo (1995) Reports that when structures a completely dried out (e.g. during summer in Southern Europe) the effect of sudden rain will be limited due to pores not being lined with monolayer of water to allow for the water to penetrate.

Frost action resulting in damage are still not entirely understood (Ingham, 2005) and three theories exist: i) When water freezes it expands and similar to salt, when crystallisation pressures exceed tensile strength of the respective material it results in structural breakdown, ii) 'ice lense theory' where an ice lense forms and continues growing by attracting water from adjacent pores resulting in higher expansion compared to ice growth without additional water available (Tabor, 1929,1930, Everett, 1961), iii) frost is the indirect cause of damage, where solid ice forces remaining liquid water into smaller pores where it gets trapped and may produce pressures up to 2100kg/cm³ (e.g. Cooke and Dornkamp, 1974; Ross et al., 1991). For both processes, temperature cycling is crucial and high frequency will increase the rate of decay. For freeze-thaw additionally the presence of adequate water is required. Wind- and water-bourne sediment can also be an extrinsic factor in stone deterioration, causing abrasion to sensitive materials (as reported by Kahraman and Gunaydin, 2007; Feal-Pérez and Blanco-Chao, 2012).

Water

Water is key in promoting limestone deterioration as it facilitates the majority of deterioration processes including the transport of salts and aggressive pollutants into the stone structure. Water is vital for limestone dissolution. Further, it enables surface induration and crystallisation processes of ice and salt in limestone surfaces/subsurface-zones.

Several authors discuss different aspects of water's role in stone deterioration. For example, crust forming processes initiated by the presence of water (Camuffo, 1995). Smith and Viles (2006) point out that for weathered stone especially in relation to crust forming processes and catastrophic decay even slight modifications in rainfall frequency and amount, evaporation rate "could trigger disproportionate changes" in decay patterns and rate of stone decay. The determining factors are time of wetness, penetration depth and frequency of wetting and drying. Time of wetness describes how long a stone block is wet whereby McCabe et al. (2013) emphasize the importance of distinguishing the surface zone and deep wetting as usually only surface moisture is considered, but deep sitting moisture can have an effect on weathering behaviour. On the one hand pollutants might be transported deeper into the structure, but it also determines the deteriorative effect of wet and dry deposition (air pollution) as well as biological colonization (e.g. Haneef et al., 1992; Lewry et al., 1994; Camuffo, 1995; Charola and Ware 2002; Smith et al., 2011a; McCabe et al., 2013). Bjelland and Thorseth (2002) for example describe that CO_2 produced by lichens respiration forms carbonatic acid with water and is thus, detrimental to the host stone of the lichen as it dissolves the substrate. Water/moisture is further necessary to initiate salt damage (Charola, 2000). Water also enhances biological growth and changes the overall temperature regime, (e.g. thermal conductivity, which is increased with increase amount of moisture), which affects the stone's response to other external impacts like thermal and water impact and wind abrasion (e.g. Benavente et al., 2008).

Sources of water are precipitation (rain, snow etc.), fog, mist, sea spray, surface run-off, relative humidity (RH%), groundwater (rising damp), the sea and rivers (potentially rising), leaks (pipes, gutters and drains) and agricultural activity (watering) (e.g. Snethlage and Wendler, 1997; Siedel et al., 2008; Posas, 2011; Gómez-Laserna et al., 2012; Cassar et al., 2013). In terms of understanding stone weathering behaviour it is thus relevant to know whether the water regime is mainly driven by for example changes in relative humidity or driven rain (e.g. Erkal et al., 2013).

Despite the many sources of moisture, its ubiquity and importance to stone deterioration there are many gaps in knowledge. One key issue is that it is difficult to measure surface, near surface and deep seated moisture within porous building materials.

Bioorganisms

Biochemical alteration of stone can be induced chemically by metabolism products of microorganisms (like oxalic acid (Adamo and Violante, 2000)) and biophysical damage by organisms like bacteria, algae, lichens, moss and higher plants. As surface-modifiers (Caneva et al., 2008; McCabe et al., 2015) they contribute to complex (heterogeneous) stone behaviour in response to environmental impacts by affecting surface moisture (examples for potential scenarios are an increase of surface moisture by absorption or a reduction of penetration of water run-off) and temperature regimes (e.g. mitigating temperature gradients) (McCabe et al., 2015). The predicted increase of rainfall for certain regions in Europe (Sabbioni et al., 2010, see Table 2.4) is likely to encourage the growth of algae, and though they are mainly seen as an aesthetic problem, they can encourage bacteria growth which in turn have been demonstrated to be powerful agents of biodeterioration of limestone (Lyalikova and Petushkova, 1991; Cutler et al., 2013).

Some organic growth on building stone surfaces may be bioprotective rather than biodeteriorative. Controversy exists, for example, on the role of lichens on limestone (McIlroy de la Rosa et al., 2014). Their destructive role is associated with disaggregation of the stone surface, dissolution processes, precipitation and formation of new minerals like oxalates (e.g. Chen et al., 2000; Bjelland and Thorseth, 2002; St. Clair and Seaward, 2004). However, Adamo (2000) and Seaward (2001) suggest that in fact some lichens might not necessarily be detrimental to the stone surface. Further, in view of extreme environments with extremely high abrasion like coastal areas the protective role of lichens has been reported (Bjelland and Thorseth, 2002; Caneva et al., 2008; Matthews and Geraint, 2008). They can mediate thermal stresses (keeping surfaces hot and dry more constantly), reduce chemical reactions (water and pollutants) and decrease physical impacts (wind and wind driven rain) and stabilise grains on the surface as demonstrated by, for example Ariño et al., 1995; Seaward, 2001; Garcia-Vallès et al., 2003; Mottershead et al., 2003; Carter and Viles, 2005; Beierkuhnlein, 2011; Özvan et al., 2015). A recent study by McIlroy de la Rosa et al. (2014) shows in particular the potential reduction of dissolution rates on limestone surfaces covered by endolithic lichens (through hyphal binding) in comparison with bare stone, especially in the winter month when metabolic activity of lichens is low.

Salt

Salts, which are nearly ubiquitous in historic buildings and structures are a major agent of deterioration when water is present (Arnold, 1984; Charola, 2000; Steiger et al., 2010). Common sources of salts are air pollutants, de-icing salt (usually sodium chloride (NaCl), soil (relevant for bases of buildings or monuments like gravestones and archaeological heritage), fertilizers, sea spray, inappropriate conservation treatments, modern building materials (e.g. cement), metabolic processes (e.g. metabolism products from bio-organisms). Salts in porous stone can directly cause damage through frequent crystallisation-dissolution processes, which cause fatigue if pressures exceed the specific tensile strength of the respective stone (e.g. Charola, 2000; Doehne, 2000). Furthermore, salt via hygroscopicity increases water retention and thus, affects moisture behaviour (Winkler and Wilhelm, 1970; Snethlage and Wendler, 1997; Rodriguez-Navarro et al., 2000; Steiger et al., 2008; Noiriel et al., 2010). As a positive effect, this might decrease freezing temperatures, but as a negative effect it might cause permanent superficial dampness, which reduces durability performance of stone as described above in the example of the collapsed limestone in Gaziantep.

The solubility of salts is influenced by temperature and the mixture of ions present (Steiger et al., 2010). Figure 2.10 shows the changing solubility of sodium acetate depending on the temperature. Natural salts commonly occur as mixtures, which might lower the deliquescence point *points* (Wexler and Seinfeld, 1991; Bionda, 2004; Price, 2007).

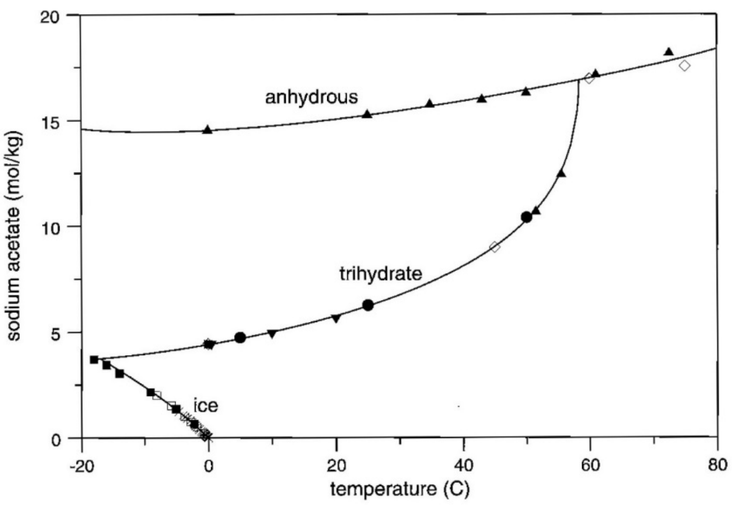


Figure 2.10 Experimental solubility and freezing point data derived from a range of studies for the system $NaCH_3COO-H_2O$ (source: Price, 2000, p. 30)

The most damaging soluble salts for historic structures are formed of the following ions: Sodium (Na^+), potassium (K^+), magnesium (Mg^{2+}), calcium (Ca^{2+}), chloride (Cl^-), nitrate (NO_3^-), sulphate (SO_4^{2-}), hydrocarbonate (HCO_3^-), carbonate (CO_3^{2-}). Further ions are acetate (CH_3COO^-), formate ($HCOO^-$), oxalate ($C_2O_42^-$), which can be related to the deteriorative effect of lichens (Adamo, 2000)), ammonium ((NH_4^+), e.g. relevant for built heritage formerly used as animal stall (or other sources of urine)(Schwarz, (accessed 2015); Cámara et al., 2014, Table 2.5).

Table 2.5 Common ions which in various combination form deteriorative agents (saltwiki http://193.175.110.91/saltwiki/index.php/Home; Barger, 1989; Massey, 1999; Doehne and Price, 2010; Abdelhafez et al., 2012)

Ion	Chemical symbol	Example for salt	Chemical formula
Sodium	Na^+	Sodium sulphate	Na_2SO_4
Potassium	K^+	Potassium nitrate	KNO_3
Magnesium	Mg^{2+}	Magnesium sulphate	$MgSO_4 \bullet 7H_2O$
Calcium	Ca^{2+}	Calcium sulphate (gypsum)	$CaSO_4$
Chloride	Cl^-	Sodium chloride	NaCl
Nitrate	NO_3^-	Calcium nitrate	Ca(NO3)2
Sulphate	SO_4^{2-}	Epsomite	$MgSO_4 \bullet 7H_2O$
Carbonate	CO_3^{2-}	Natrite	Na_2CO_3
Acetate	CH_3COO^-	Calcium acetate	$Ca(C_2H_3O_2)_2$
Formate	$HCOO^-$	Sodium formate	Na(HCOO)
Oxalate	$C_2O_4^{2-}$	Oxalatic acid	$C_2H_2O_4$
Ammonium	NH_4^+	Nitric acid	HNO_3

2.2.3 Synergy between intrinsic and extrinsic factors: limestone weathering behaviour and weathering-stress history

It is of course, the interaction between extrinsic and intrinsic factors, which determines the rate and nature of stone deterioration. This section reviews the complexities of the relations between the two sets of factors.

Smith and Viles (2006) discuss the role of complex interplay of moisture regime, salt and dry deposition of atmospheric pollutants in promoting catastrophic limestone decay.

Micro – Meso –Macro – The matter of scale

Scale matters when interpreting limestone weathering behaviour, as it influences measurement and understanding. What is defined as deterioration varies with scale. Thus, at the micro-scale etching of individual calcite crystals is a quantifiable and obvious manifestation of deterioration; however, at the meso to macro scale this might not result in any observable deterioration or weathering as neither aesthetics nor static problems result. On the other hand, meso and macro scale cracking may be an obvious feature of deterioration on some limestone monuments, but this might not be associated with any micro scale alteration (Viles, 2001; Moses et al., 2014). The differences in scale are also one key issue in the difference between laboratory and field investigations of weathering (Ingham, 2005; Viles, 2013). For example, do the small blocks usually used in laboratory weathering experiments fully represent the behaviour of stone in buildings?

Scale issues do not simple concern spatial scale. Further criticism on stone weathering studies under real world condition refers to their short-term character and small scale (Doehne and Price, 2010; Stephenson et al., 2010; Smith et al., 2011b). It has been proved to be difficult to upscale from short-term weathering observation to long-term weathering behaviour due to unaccounted weathering-stress history effects and potential extreme weather events (Inkpen and Jackson, 2000). The same is true for upscaling from small samples commonly used in stone exposure studies (Bell, 1993; Cooke and Gibbs, 1994; Trudgill and Viles, 1998; Moroni and Pitzurra, 2008; Moses et al., 2014). On the other hand, rates inferred from very long term cosmogenic surface dating studies in geomorphology (usually thousands of years) have been found to be curvilinear and are equally problematic to downscale (Stahl, 2013).

Scale issues also influence measurement of stone deterioration and comparability of results. For example, some measurement devices only collect data from millimetre to centimetre sized areas of a surface. How representative is this of the weathering status of an entire building stone block, element of façade? The scale issues reviewed above and the choice of scales for investigation will affect the interpretation and potential to link to a) other scales and b) other studies.

Location, aspect and orientation characteristics

Limestone weathering and its interpretation is highly dependent on geographical location, aspect, degree of orientation and height. So, Inkpen and Jackson (2000) find contrasting weathering rates of gravestones made of the same stone in urban and rural environments. However, there are multiple sources of differences between places in terms of their environmental conditions. Not only variations in air quality (where urban is generally thought of as more polluted), but also in proximity to the coast, altitude etc. (Bell, 1993). Furthermore, rural does not always equate to non-polluted, and O'Brien (1995) finds a higher stone loss rate in rural areas. At a single site, aspect differences can have a major impact on weathering processes and rates. In the UK southwest (SW) being the prevailing weather direction limestone facing this aspect are generally thought to experience higher rates of dissolution as impact of wind and water is more pronounced. Accordingly, Sharp et al. (1982), found significantly higher rates of limestone weathering on the Portland limestone balustrade of St Paul's Cathedral, London on SW (exposed). In terms of degree of orientation of exposed surfaces Paradise (1998) observes accelerated limestone weathering on horizontal vs vertical surfaces of the Great Temple of Amman (Jordan).

The effect of geometry on weathering behaviour is best described with the well-known forming of gypsum crusts on monuments (Figure 2.11). Gypsum crusts on limestone surfaces are a result of the sulphation process in which water induced dissolution of the limestone ($CaCO_3$) forms calcium hydroxide ($CaCO_3 \rightarrow Ca(OH)_2$) which chemically interacts with sulphur oxides (SO_4^{2-} and SO_3^{2-} to gypsum (Ca_2SO_4) (Gomez-Heras et al., 2008, Equation 4)

$$CaCO3 + H2SO4 \rightarrow CaSO4 + CO2 + H2O \quad \text{(Equation 4)}$$

A noticeable fact with regards to the *in situ* investigation of gypsum crusts is that gypsum crusts can be white, grey and black or not visible at all as the colouration is not related to gypsum itself, but from soot, carbon etc. (Siegesmund et al., 2007). Thus, where urban air quality is improving, especially where particulates are declining, gypsum crusts may be present but harder to diagnose (Searle and Mitchell, 2006; Siegesmund et al., 2007).

Figure 2.11 Gargoyle with black gypsum crust, facing S-SW at St Mary's Church, Oxford, UK (2012)

Figure 2.11 shows a black gypsum crust on the lower parts of a gargoyle (facing S-SW) which usually form on lee side areas sheltered from rainwash and driven rain (Williams and Robinson, 2000; Smith et al., 2003; Török, 2003). This shows that stone weathering behaviour is, a) dependent on geometry of the heritage structure and b) is often heterogeneous at block size level as described by McCabe et al. (2015). As a consequence, Smith et al. (2011) and McCabe et al. (2013) emphasize the importance of investigating stone response at a local level in order to understand spatial variability.

Weathering-stress histories

The weathering-stress history of a building stone encompasses all relevant past processes and events which have an ongoing influence on deterioration rates and processes. The key elements are quarrying, dressing, climatic impacts (bio, physico-chemical) and dissolution and pollution histories, and finally history of use (e.g determines exposure to rain when roof is maintained or not) and conservation treatments. Weathering-stress history is a crucial influence on the present day interactions between extrinsic and intrinsic factors. To understand weathering-stress history it is first necessary to look at the 3 dimensional nature of the weathering zone (surface/subsurface).

Recent research acknowledges the importance of considering the dimension of subsurface depth to stone weathering behaviour on-site (Pope, 2002). McCabe et al. (2015) find 'surface/subsurface-to-depth heterogeneity' at block scale caused by surface-modifiers (biological growth). Similarly, Hoke and Turcotte (2004) describe the 'formation of a dissolution layer' as pre-surface recession weathering in relation to an 'incubation' time before surface recession sets in. The dimension of this 'zone' might be from less than a few millimetres to a few centimetres. The level of heterogeneity (number of distinguishable layers) depends on the material and its weathering-stress history (e.g. Stahl et al., 2013). Figure 2.12 illustrates different types of such surface modifications. The implications for limestone weathering are large.

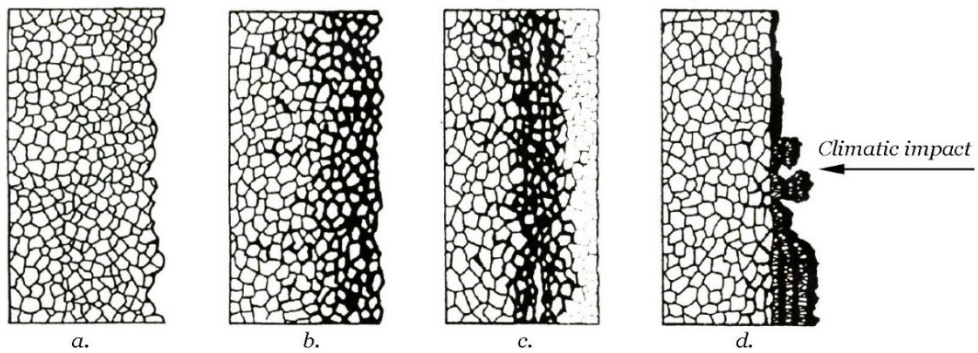


Figure 2.12 Schematic description for common weathering profiles of stone surfaces; a. superficial (granular) disintegration & erosion, b. increased porosity & decline of intergranular bonds, c. case hardening (indurated surface) with increased superficial density followed by zone of increased porosity, d. crust (altered after Wolf Dieter Grimm, 2010; p 176).

With the modified surface/subsurface zone the stone behaves as a two material system (composite) as the surface layer now has different physical properties compared to the subsurface and core of the stone (which may be further softened). As a consequence, the deterioration potential of climatic impacts like thermal fluctuations, water ingress and crystallisation events (frost and salt) is increased. Instead of freeze-thaw events causing minor granular disintegration of the soft limestones, catastrophic exfoliation and blistering of the indurated layer and the exposure of a 'core softened' zone underneath is likely. Thus, the stone behaves in a more complex fashion than expected.

The exposure history of a building limestone on heritage monuments starts with quarrying. As soon as the stone, which was in equilibrium with its surroundings of either the rock formation or soil with its specific pressures, moisture content and chemical set up (pH etc.), is exposed to (subaerial) environmental impacts it weathers in a complex, dynamic, metastable way until equilibrium with the environment is reached (Přikryl, 2013) (i.e. total loss of stone). In terms of quarrying of building limestone, and in fact for limestone structures exposed after archaeological excavation as well, the process of 'curing' is another aspect to consider (Přikryl, 2013). Limestone having been covered (within its formation or soil etc.) and experienced dissolution processes contains calcium hydroxide ($Ca(OH)_2$) and the stone is softened ($Ca(OH)_2$ is chemically less stable than $CaCO_3$). As soon as exposed to the atmosphere reaction with carbon dioxide (CO_2) to calcium carbonate ($CaCO_3$) takes place resulting in hardening of the surface. Therefore, for building stone purposes quarried stone might (or should) have been used after a sufficient period of 'curing' or 'hardening'; usually 2 to 3 years. This was already known by architects like Vitruvius and Wren (in Ashurst and Dimes, 1998). Another potentially protective layer is the surface 'finishing' or dressing of building stones, which Hoke and Turcotte (2004) relate to a lag ('incubation time') of marble gravestones before dissolution takes place as the smooth surface for example lessens the effect of water run-off effects.

The opposing mechanisms are dissolution processes and chemical attack through dry and wet deposition of air pollutants. The accumulation of these depositions has been described as the 'memory effect' (Cooke and Gibbs, 1994). The hypothesis supporting the 'memory effect' is a lag in stone weathering response to air pollution exposure, meaning past elevated air pollution levels (1965-1980) may affect recent stone weathering behaviour (Trudgill et al., 1991). The implication is that despite decreasing air pollution levels over the last decades a respective decrease in limestone weathering rate cannot necessarily be expected. However, no clear evidence has been found for the 'memory effect' neither in short-term studies (Vleugels, 1993) nor in longer studies (30 years) (Inkpen, 2012b). It seems to be problematic to differentiate the 'memory effect' from further contributing weathering mechanisms like, for example, synergistic effects between accumulated atmospheric pollutants and microbial contaminants as found by Moroni and Pitzurra (2008). In the scope of this thesis the ‘memory effect’ is seen as part of the whole weathering-stress history of a stone, which may include other deteriorative processes. Therefore, ‘weathering-stress history’ is the term of choice in this thesis to give a more holistic image of potential past factors affecting present stone weathering behaviour.

Past conservation measures might add to the weathering-stress history of the stone. Thus, cleaning with aggressive detergents, consolidation with inappropriate materials (non-reversible or insufficient penetration depth etc.), insertion of inappropriate stone or incorrect use of replacement mortars. Svahn emphasizes, that in order to achieve successful diagnosis of the state of preservation of any investigated stone heritage its conservation history needs to be known (Svahn, 2006).

The synergistic action of intrinsic and extrinsic factors culminates in the creation of complex weathering-stress histories. Moses et al. (2014) emphasize the need to measure and monitor weathering behaviour in order to understand interaction of extrinsic impacts with intrinsic stone properties. Advances in understanding these deterioration processes will inform future conservation strategies (Smith et al. 2010).

2.3. QUANTIFYING STONE DECAY – NON-DESTRUCTIVE TESTING ON-SITE

Establishing meaningful estimates of the rate of limestone weathering and its variation over time and space is important for geomorphology and heritage conservation. Decisions on heritage conservation strategies can be based on the results such as developing risk maps in order to define the levels of urgency for remedies to be undertaken. Fuente et al. (2013) for example in course of the CULT-STRAT project (EU 6^{th} Framework Programme) mapped past, present and future air pollution effects on cultural heritage in cities to determine where weathering processes like corrosion of bronze and recession of Portland limestone exceed established tolerable levels.

2.3.1 Time series

Further, stone response to impacts like climate change and air pollution might be linked back to standard durability tests in order to improve their accuracy (e.g. Ross and Butlin, 1989; Meierding, 1993a; Viles, 2002b; Smith et al., 2011; Viles and Cutler, 2012). The complexity of the three dimensional system of stone response to its environment is further increased by taking the 4th dimension into consideration i.e. time. As discussed earlier weathering-stress histories can influence current deterioration.

Historic cemeteries are an example for what Gomez-Heras and McCabe (2015) introduce as concept of stone-built heritage as "large scale laboratory" and 'recorder' for past environmental evolution. Thus, cemeteries are ideal for investigating stone weathering behaviour under real world conditions over a variety of timescales. Comparing headstones installed at different dates allows weathering rates to be established through time, and their relatively large size allows comparative information to be collected from different sections of the stones (e.g. Cooke et al., 1995; Inkpen and Jackson, 2000).

2.3.2 Laboratory vs in situ, standard tests, durability, resilience, time and scale

This section reviews how by combining the given non-destructive methods a deeper insight in stone weathering behaviour in situ can be gained and suggests ways to improve reliability of data evaluation.

2.3.3 Measuring weathering extent and rate

The important question in terms of limestone weathering rates is whether they are decelerating or accelerating. Secondly, the character of weathering behaviour is crucial to know in order to potentially make predictions and decisions especially in view of climate change impact (Smith et al., 2008). Answers would have implications for the urgency of indirect and direct preservation measures (like changing air quality levels (European Union, 2008) or applying chemical consolidation products (e.g. Pinto and Delgado Rodrigues, 2008) to be undertaken and further advance understanding of interactions between stone and environment. Limestone weathering rates are described by the temporal change of a measureable parameter as a proxy for limestone weathering behaviour (like surface recession or surface property changes.

Methods to measure rates of stone surface changes are summarized in Table 1.1. The majority of studies investigated limestone weathering rates in response to air pollution relied on deliberately exposed samples (e.g. Lipfert, 1989; Trudgill et al., 1991; Butlin et al., 1992; O'Brien, 1995; Bonazza et al., 2009; Brimblecombe and Grossi, 2009). The related scale problems are discussed in Chapter 1. In order to overcome the scale problem built heritage can be studied *in situ* using both contact measurements (e.g. micro-erosion meters (MEM)) and direct measurements relative to a datum point (Moses et al., 2014). Relative measurement points include artificially introduced structures such as lead plugs and lead letters, or parts of a historic structure itself, such as unweathered surfaces and quartz veins. An example of this approach is the 30-year (1980–2010) investigation of limestone erosion on the balustrade at St Pauls Cathedral in

London (Trudgill et al. 1989, 2001; Inkpen et al., 2012a, b), where both lead plug index and MEM methods were applied" (Wilhelm et al., 2016c).

For all methods to provide appropriate accuracy crucial points need to be considered such as for example interference from airborne Calcium to limestone recession determination in microcatchment field studies. McIlroy et al. (2014) present a carefully designed study, where the weathering of limestone is measured as calcium run-off from exposed samples and the results are corrected for calcium obtained through the air. Not all microcatchment studies consider or clearly account for such confounding factors. Another example (the lead plug or lead lettering index) reveals further issues.

The limitation of lead plug index measurements was a) bias towards intact lead-plugs, b) its sensitivity. Inkpen and Jackson (2000) find weathering rates of marble beyond the sensitivity of the callipers and investigated time period (< 100 years). Micro erosion meter measurements also have their challenges. Inkpen et al. (2012b) find that different ways of quantifying rates (surface change and surface recession) yield different quantitative results (though show a similar trend).

Besides the described limitations, the majority of methods measure superficial changes only and non-destructive methods capturing weathering processes in both the surface and subsurface zone would be beneficial.

In general, it is found that combining methods and evaluating more than one stone parameter yields deeper insight and allows for more reliable inference and prediction (e.g. Aliabdo et al., 2012; Breysse, 2012). Thus, in terms of the determination of limestone decay rates existing methods (with focus on chemical weathering) would benefit from being complemented by methods describing physical changes like surface scanning techniques (Emmanuel, 2015) and surface hardness testing (a long-established method in geomorphology for relative dating).

2.3.4 Non- destructive testing *in situ*

In contrast to destructive methods for stone weathering research, non-destructive on-site testing methods require no sample taking, often generate immediate results, can be applied on a larger scale and more frequently (relevant for time-series studies) as no historic material is damaged or interfered with and thus, key principles of built heritage conservation are maintained (i.e. to preserve as much original fabric as possible (e.g. the Venice Charter, 1964; the Malta Convention, 1992; Petzet, 2010).

A range of accessible, economic and portable non-destructive methods is available for stone weathering research on-site (e.g. Doehne and Price, 2010; Bläuer-Böhm, 2012). However, since stone weathering research focuses on the resilience and vulnerability of stone with weathering-stress histories most non-destructive methods have to be adapted and often standards or guidelines for good practice are not available or applicable. For example, Burkinshaw (2002) reports that handheld moisture meters are often regarded with suspicion by conservators. Yet, they are convenient, simple, inexpensive tools and are applied frequently by professional surveyors, geomorphologists and heritage conservation scientists.

Thus, there is a need to develop reliable methodologies for simple and inexpensive, portable non-destructive testing methods in order to quantify the extent and rate of limestone heritage decay under real world conditions. Especially with regards to improved sampling protocols (e.g. sample sizes) and improved reliability of data generated by non-destructive methods with more appropriate data evaluation methods (e.g. Burkinshaw, 2002; Svahn, 2006; Viles et al., 2011; Wilcox, 2012).

In this section three non-destructive methods are introduced and reviewed 1) low impact surface hardness testing, 2) moisture measurement with handheld moisture meters and 3) capillary water uptake measurement methods. When used in combination, these methods are able to provide linked surface and subsurface information. Furthermore, it is shown how effects on the data output like porosity and salt content in the stone structure considered as limitations could be utilized to increase insight into limestone weathering behaviour.

Portable surface hardness testing

Research on stone weathering based on non-destructive methods benefits from cross-fertilizations with fields like engineering and geomorphology. Thus, non-destructive index tests like surface hardness tests are frequently applied in the field of concrete and stone engineering as an alternative to destructive tests like unconfined compressive strength (e.g. Aliabdo et al., 2012).

The most popular rebound device for geomorphological applications is the Schmidt Hammer (e.g. Aydin and Basu, 2005; Goudie, 2006; Fort et al., 2013; Stahl et al., 2013). Due to its high impact energy (Type L = 0.735 N m and type N = 2.207

N m (Proceq©, 2006) which can damage the surface of the stone and the required surface preparation with carborundum – its application is unacceptable for most cultural stone heritage (Pope, 2000; Viles et al., 2011). In contrast, the Equotip family of devices offers a practical alternative. The impact energy of the Equotip D is 0.0115 N m which is a fraction of that of the Schmidt Hammer and thus it is more suitable for valuable and vulnerable materials. It measures a wide range of stone or rock surfaces (e.g. gypsum, tuff, limestone, granite) at different stages of weathering, as well as detecting subtle changes in surface hardness (e.g. Hack et al., 1993; Verwaal and Mulder, 1993; Aoki and Matsukura, 2007; Viles et al., 2011; Alberti et al., 2013; Coombes et al., 2013; Hansen et al., 2013). Low rebound values indicate soft, porous and/or weathered stone surfaces, higher values less weathered or case hardened surfaces.

Natural stone property variations like porosity are known to affect Equotip data generation, and this effect is expected to increase with degree of stone weathering. Nevertheless, McCaroll (1991), who observed a similar effect for Schmidt Hammer measurements, states that surface roughness and weathering are intimately related and suggest utilizing it for comparison for stone with similar surface textures prior to the influence of weathering.

Surface hardness testing devices can be applied in two ways. The single impact method (SIM) is most common and involves applying the device randomly across the stone surface (Aoki and Matsukura, 2007). In contrast, the repeated impact method (RIM) collects surface hardness data repeatedly on the same point of the stone surface (Aoki and Matsukura, 2007).

Different information is obtained, where the SIM testing reflects on the stone surface elastic and plastic properties when applied with the single impact method (measurements are randomly distributed over the surface). In contrast, RIM method reflects on the elastic and plastic properties of the surface and subsurface of the stone.

Siedel and Siegesmund (2010) point out high variability of stone properties for low density limestone. This variety will affect Equotip readings similar to the Schmidt Hammer for which studies have shown that the number of readings taken has bearing on the meaningfulness of subsequent statistical tests (Niedzielski et al., 2009). Thus, the key issues that need to be addressed when applying Equotip devices to stone or rock surfaces are a) the number of readings that should be taken, and how this affects the reliability of statistical tests applied to the data collected; only a sufficiently big sample size will reflect the true surface hardness of a material, and this may be somewhat dependent on the material being tested and its weathering status, b) how outliers should be treated and c) the effects of surface roughness and varying porosity.

Handheld moisture measurement

In view of water being a key deterioration agent for limestone weathering, assessing moisture regimes of immovable heritage on-site is indispensable. A range of handheld resistivity and capacitance type moisture meters alongside infrared (IR) and microwave based methods are available and frequently used by professional surveyors, geomorphologists and heritage conservation scientists (e.g. Viles, 2013; Cutler et al., 2013).

Electrical moisture meters apply either a current or electrical field to a porous material (e.g. stone). Thus, they do not directly measure actual moisture content, but detect a change in either resistivity or capacitance. Most devices automatically convert the results to an estimated or calibrated moisture content value usually based on a specific material like wood (Eklund et al., 2013), which might not be representative for the material being tested. Therefore, care needs to be taken and it is necessary to understand the measuring principle and factors affecting it in order to interpret the results reliably (Arendt and Seele, 2000; Burkinshaw, 2002).

Difficulties arise from a range of influences on the measurement (e.g. mineralogy, homogeneity and density of the measured material, temperature and moisture distribution within the material, presence of contaminants, application pressure, type of measuring voltage or frequency, operator variance and surrounding factors like the presence of metal (e.g. reinforcement in concrete structures) (Arendt and Seele, 2000; Martinez and Byrnes, 2001; Eklund et al., 2013)).

"Of these factors, one of the most important is the presence of salts – which are nearly ubiquitous in historic buildings and structures. Dissolved salts increase conductivity of (pore) water and, therefore, reduce resistivity in stone (Loke, 1999). However, it remains unclear how these devices are affected by the presence of salt and how best to interpret the data they provide (Wilhelm et al., 2016b). A lack of knowledge on the exact interactions between moisture meters and building stone condition is evident.

Portable capillary water uptake testing

As mentioned in section 2.2.1 the stone pore system in combination with water ingress is one of the driving forces of stone weathering behaviour. Therefore, investigating pore characteristics is crucial in order to understand limestone weathering behaviour. Single aspects of porosity (like pore diameter distribution, effective and ineffective porosity, interconnectivity etc.) are commonly investigated in the laboratory by mercury porosimetry (MIP), BET ((BET=Brunauer, Emmett and Teller) a method, which uses the physical adsorption of gas to determine surface areas of solid substrates), thin section analysis and X-ray tomography (CT) to name a few (Benavente et al., 2004). However, to describe water uptake behaviour capillary water uptake under atmospheric pressure is the method of choice as it reflects on the behaviour of the pore space as a body (3D) and thus, allows for better estimates of its role in the stone weathering behaviour (Vandevoorde et al., 2012).

Similar to the methods discussed in the sections above a range of portable and non-destructive water uptake testing methods is available, of which the Mirowski and Karsten tubes are the most common (Auras, 2011; Vandevoorde et al., 2009 and 2012). Figure 2.14 shows the *in situ* application of the Karsten tube. The Karsten tube is a glass cylinder open on one side with a tube with indicated gradation on the other side, which is applied to the stone surface using putty (Plastic Fermit) (Auras et al., 2011). It is filled once with distilled water and subsequently the time it takes for the water to penetrate the stone through the open side of the glass cylinder is measured.

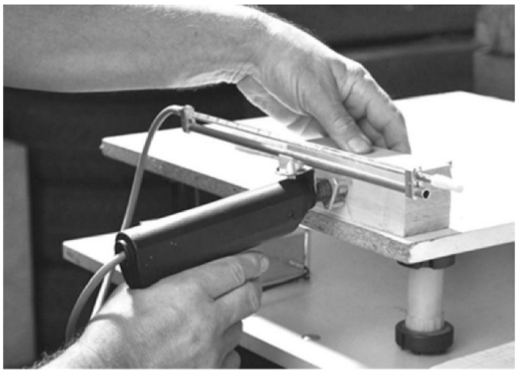


Figure 2.13 Microtube method prototype (Drdácký, 2012)

Figure 2.14 In situ application of Karsten tube to a limestone headstone during field work at St George's Church cemetery in Portland, Isle of Portland (UK)

Drdácký (2012) points out four problems related to the use of the Karsten tube 1) operators find it frequently difficult to fix the tube to the stone surface, 2) as a result tube tends to leak, 3) two operators are required where one monitors the water uptake and time and the second records the results and 4) the sealant leaves residue on the stone surface. Alternatively, he provides the 'microtube method', a new promising invention, where the water uptake measurements are partly automated (Drdácký, 2012, Figure 2.13). Unfortunately, the device is not commonly available yet.

A range of data evaluation methods for Karsten tube results is available but, there appears to be no consensus over i) which method to use, ii) what an appropriate sample size was, iii) the level of replicability and iv) how to best evaluate the data.

Four possible data evaluation methods have been proposed in the literature. The most straight forward approach is described in the BS EN Standard 16302:2013. It suggests dividing the measurement into time intervals of 10 seconds up to one minute (depending on the porosity of the material) and to take readings until either a constant value is reached or for maximum in an hour. Results should be expressed on a water absorption graph; with volume of water absorbed plotted against time. The Standard provides solely a procedure to describe water uptake with a Karsten tube (or a similar pipe), but does not give any recommendations on interpretation of the data or in depth evaluation. Vandevoorde (2012) follows another approach and applies data evaluation recommended in RILEM II.4 (Reunion Internationale des Laboratoires D'Essais et de Recherches sur les Materiaux et les Constructions (RILEM), 1980) where measurements are recorded after 5 and 15 minutes. For highly porous stone this approach is not appropriate as distinctive water absorbing behaviour might be evident before the 5 minute measuring threshold as found by Vandevoorde (2012) and the author; this under the premise of not refilling the tube, which has been found to be nonessential as the pressure which changes with the retreating water column are not significantly affecting the results (Vandevoorde, 2013; Hendrickx, 2013).

Alternatively, D'ham et al. (2011) introduced two evaluation procedures using Microsoft Excel©. The first, Calkarow V3.2 was developed by Wendler and Pfefferkorn (1989) and the second, more recent, by Niemeyer (2013). The programs add a level of information (spatial dimension), because both programs use algorithms, which model the water 'body' entering the stone structure as an (idealized) geometric cylinder with surrounding quarter torus. Based on this assumption the penetration depth of the water is calculated and the water penetration coefficient (B-value) and water uptake coefficient (w-value) can be calculated accordingly. The Niemeyer algorithm provides a better approximation than Calkarow. However, the limitation here is given by water absorption coefficients, which is part of the calculation and needs to either be determined on samples or retrieved from literature, thus if not available the calculation becomes inaccurate (see equations 5-8) for calculating water uptake coefficient, water penetrating coefficient, water uptake capacity and actual volume of taken up water).

$$w = \frac{m}{A \cdot \sqrt{t}} \quad \text{(Equation 5)}$$

w = water uptake coefficient
m = penetrated water
A = area through which water penetrated
t = time

$$B = \frac{x}{\sqrt{t}} \quad \text{(Equation 6)}$$

B = water penetrating coefficient

x = penetration depth of water

$$WAK = \frac{w}{B \cdot \delta} \quad \text{(Equation 7)}$$

WAK = water uptake capacity
δ= density of water = 0.998 g/cm³ at ~20°C

$$V_{kar} = \frac{\pi \cdot d^2 \cdot w \cdot \sqrt{t}}{4 \cdot \delta} + \frac{\pi^2 \cdot d \cdot w^2 \cdot t}{4 \cdot WAK \cdot \delta^2} + \frac{2 \cdot \pi \cdot w^3 \cdot t \cdot \sqrt{t}}{3 \cdot WAK^{2} \cdot \delta^3} \quad \text{(Equation 8)}$$

V_{kar} = volume of water taken up through the Karsten tube based on the assumption that the penetrating water body has the following shape: a cylinder with a rotated quarter of a torus

d = diameter of Karsten tube

A further limitation is the assumption of water penetration being a linear process. So for example, surface to depth heterogeneities (McCabe et al., 2015)) like induration or a 'dissolution layers' might cause a higher or lower rate of water uptake for the first few millimetres and change for subsequent areas in the bulk of the stone. This has been pointed out by Svahn (2006), who also misses a quantification of non-linear water uptake behaviour over time.

2.4. DATA EVALUATION – PARAMETRIC VS. NON-PARAMETRIC STATISTICS

The main requirements for the statistical approach in this thesis were applicability and robustness for data derived from non-destructive index methods in the field of stone weathering research and heritage preservation *in situ*. The nature of the subject brings certain challenges such as inherent variability of stone as a natural product, data variability of index methods due to non-destructive character and limited sampling sizes (due to heritage protection, financial and time constrains).

This section assesses statistical measures and procedures currently used in the field of stone weathering research in order to determine an approach with high reliability (accuracy) and comparability of results (linking to other studies and fields of research).

Descriptive statistics of results and hypothesis testing are at the core of stone weathering research and therefore an inevitable part of the overall methodology. Parametric statistical measures (mean, standard deviation) and null hypothesis

(H_0) significance testing (t-test, ANOVA) are commonly used in rock weathering and stone deterioration studies (e.g. Hack et al., 1993; Aoki and Matsukura, 2007; Erceg-Hurn and Mirosevich, 2008; Cutler et al., 2013; Yilmaz, 2013). To successfully apply parametric statistical tests a range of assumptions have to be met like a) normal data and error distribution and b) homogeneity of variance between groups (homoscedasticity), (e.g. Erceg-Hurn and Mirosevich, 2008; Field, 2009; Wilcox, 2012). In cases of non-normal data distribution, the reliability of statistical estimates based on the assumption of normality may be affected and parametric tests are largely inappropriate (Tukey, 1977; Fowler et al., 1998; Filzmoser and Todorov, 2013).

Linear regression analysis with least square (LQS) is commonly used to correlate standard tests like unconfined compressive strength to non-destructive methods like surface hardness testing, water uptake and ultrasound velocity measurements (e.g. Pamplona, 2008; Vasconcelos et al., 2008; Aliabdo et al., 2012; Yilmaz, 2013). Furthermore, it is used for describing and estimating weathering rates of limestone. In stone weathering research usually the exposure time is associated with the quantification of a parameter reflecting on stone property changes like surface recession/surface loss (Inkpen and Jackson, 2000), weight loss (Trudgill et al., 1994) or in geomorphology Schmidt Hammer surface hardness (e.g. Fort et al., 2013).

Non-normal and heteroscedastic data can cause a high probability (up to 50% (Wilcox, 2003)) for Type I error for a (typical) α-value of 0.05 i.e. falsely rejecting H_0, although there was no significant difference between hypothesis and findings ('finding an effect which is not existent'). Secondly, the probability (power) of Type II error is effected i.e. falsely not rejecting H_0 although there was a significant difference between hypothesis and findings ('missing an effect which is existent')(Erceg-Hurn and Mirosevich, 2008). In practice these assumptions are rarely met when handling real world data (Erceg-Hurn and Mirosevik, 2008; Wilcox, 2012). Wilcox (2012) states: "[…] distributions are never normal" (Wilcox, 2005, p 2) and finds "that departures from normality that have practical importance are rather common in applied work" (Wilcox, 2012, p 108). Reimann (2008) states that environmental data usually deviates from normal distribution. Similarly, in stone weathering research non-normally distributed data is commonly found as for example stone surface hardness by Alberti et al. (2013) and Hansen et al., 2013 or natural stone properties as found by Siedel and Siegesmund (2010). As mentioned before stone as a natural product, and limestone in particular, displays a huge variance in inherent properties (e.g. porosity) and with accumulated weathering history the effect is thought to be even more pronounced, which in turn will affect statistical data evaluation.

For LQS Pearson's coefficient of determination (R^2 with values ranging from 0 and 1) illustrates the strength of association i.e. how close data points are to a fitted linear line. Assuming there is a real association between the correlated values, then a low R^2 value indicates a poor fit and effects the precision of prediction. The dataset needs to be investigated further using residual plots and t-test to test for significant differences. Figure 2.15 shows regression example from a study on weathering rates of marble gravestones by Hoke and Turcotte (2004).

The interesting finding of this study was that through a 'dressed' surface the contact time for precipitation was reduced and thus an 'incubation time' of c. 20 years

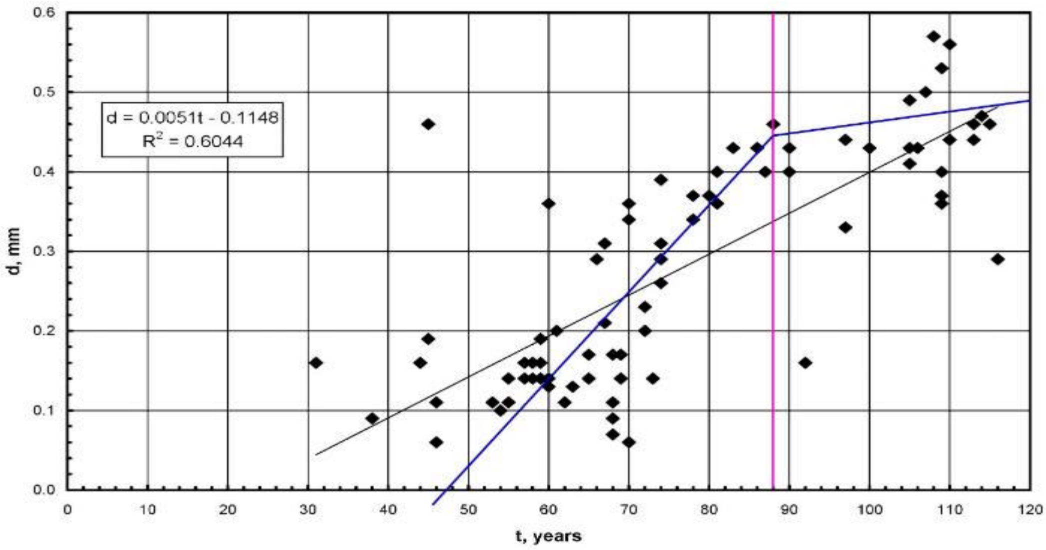


Figure 2.15 Regression graph demonstrating 'incubation time' for marble gravestones, where surface 'dressing' results in a 'protection' layer retarding weathering for about 20 years. Mean values (depth of weathering) are plotted against weathering exposure time. The black line derives from least square correlation. R^2 fit is moderate (see for classification). Alternatively (suggested by the author), the pink line (qualitative) could indicate a potential break-point resulting in two new regression lines (blue) with varying gradients (coefficient/rates). With the alternative approach the original interesting information of a non-zero intercept is maintained, only the period for the 'incubation time' has shifted towards a higher value (~48 years) (modified after Hoke and Turcotte 2004)

established before surface recession weathering would show an effect. There are clear positive implications for built heritage preservation as it indicates that 'simple mechanical' surface finishing of gravestones could prolong the period of a healthy state of preservation. (opposed to chemical consolidation treatments common to preventively protect built stone). However, Pearson's R^2 only shows a moderate fit with 0.6. A closer look at the distribution of data points shows that another than linear fit (e.g. sigmoidal (s-shaped)) might have been more appropriate. Alternatively, in order to apply linear regression, whilst potentially improving the data output and establishing an even longer 'incubation' time breakpoint analysis might have been beneficial with segmented regression (Muggeo, 2003; Crawley, 2005; Muggeo, 2008). Figure 2.15 shows a visually fitted potential breakpoint at ~48 years and two new regression lines (blue) with varying gradients (coefficient/rates) where between ~48 and 88 years an increased rate of weathering (surface recession) occurs, which after 88 years might even out. This alternative would thus 'add' 8 more years to the originally calculated incubation time and thus predict a longer protective effect of smooth (dressed) surface.

2.4.1 Detecting non-normality

Non-normality in datasets can visually be detected with density plots (where skewness indicates a shift from the non-normal distributions) or boxplots, which show clearly data distribution and potential outliers as presented by e.g. Lednicá (2012) for Schmidt Hammer hardness and Siegesmund and Dürrast (2010) for natural stone variability. The Shapiro-Wilk test helps to assess normality numerically where values <0.05 indicate non-normality.

As mentioned in section 2.2.1 fresh porous limestone as a natural product can show a high variance in stone properties. This potential variance is expected to increase with extended weathering history (Cooper, 1992). The variance of porous limestone will have an effect on data evaluation as found for surface hardness testing *in situ* on weathered rock by Hansen et al. (2013) and Alberti (2013). The common approach to increase n in order to address high variability is not applicable for *in situ* testing of protected heritage sites (where sample sizes around 30 readings per unit of investigation are common) due to time, money and surface area constraints.

2.4.2 Addressing non-normality in datasets

A few stone deterioration researchers addressed the problem of non-normally distributed data in stone research. Mosch and Siegesmund (2007) and Van de Wall (1997) employ boxplots (with which non-normality and outliers become transparent) to display the natural variance of stone.

Semi-parametric tests (a hybrid of parametric and non-parametric (Powell, 1996) are one solution, however require data modification. To modify the data needs to be normalized and/or trimmed, thresholds defined and outliers removed before analysis (Reimann, 2008; Good and Hardin, 2009). The transformation and modification of data does not always lead to an evaluable dataset. For example, Alberti et al. (2013) modified 24 Equotip datasets using two methods (in one instance using only the 50% highest values and in another removing the 8 extreme values from datasets), and yet some datasets remained non-normally distributed (Wilhelm et al., 2016a). Feal-Pérez and Blanco-Chao (2012) dealt with the problem of potential outliers in regression analysis for on-site rock testing by calculating the Huber M-Estimator, which is robust against outliers (maximum likelihood type (Huber, 1981, page 43)).

2.4.3 Outliers

Outliers are one factor associated with non-normally distributed data, which may be present in a dataset as a result of human and/or instrument error, or due to natural deviations in the sample population (Hodge and Austin, 2004). Viles et al. (

2011) frequently find outliers in their surface hardness testing data. Bruno (2013) evaluates the effect of outliers on surface hardness data sets (Schmidt Hammer) from limestone samples. A common approach is to remove outliers from the dataset to prevent restrictions on subsequent data evaluation (Rosner, 1983). In order to remove outlier(s) they need to be identified, which is a matter of definition and various outlier detection methods are available (Wilcox, 2005). Outliers are extreme values, which are noticeable different from the majority of the data and are defined by the respective researcher and the chosen method to detect them. There is a range of approaches and the most common is to define the boundaries via the formula (e.g. Reimann et al., 2008, Equation 9):

$$Mean \pm 2 \cdot SD \qquad \text{(Equation 9)}$$

SD=Standard deviation

The same procedure can be applied to robust measures (median and median absolute deviation). Equation 10 shows an example applied in this thesis following the moderately conservative recommendation of Leys et al. (2013):

$$Median - 2.5 * MAD < x_i < Median + 2.5 * MAD \qquad \text{(Equation 10)}$$

MAD = Median absolute deviation

Identifying outliers is an important part of any statistical evaluation as they can provide useful information about the sample in their own right (Lipfert, 1989; Banerjee and Iglewicz, 2007).

Their removal however, is only recommended when it is clear that their occurrence resulted from errors in the data gathering process and is not related to the population characteristics (Field, 2009). In stone weathering research outliers may reflect for example inherent, true variability in the hardness of a deteriorating stone. Furthermore, in cases of already small samples size the removal of outliers would further reduce the dataset. However, where outliers are to be retained a new approach to statistical evaluation is required (Wilhelm et al., 2016a)

To modify the data, detect outliers and potentially remove them in order to apply parametric statistical test is one solution to deal with non-normally distributed data. A second alternative solution is non-parametric statistics. Niedzielski et al. (2009) already state that for on-site surface hardness testing of rock robust non-parametric statistical methods may be more appropriate. Similarly, Mottershead et al. (2003) use the median and Kruskal-Wallis test to evaluate sandstone weathering rate of historic structures on-site. Erceg-Hurn and Mirosevich (2008) stress the huge potential of non-parametric tests to improve data analysis and investigate why these methods despite clear advantages are not popular amongst researchers. They found a lack of exposure to and misconceptions of modern robust statistical methods such as that the software to perform modern statistics is not readily available because it is not built into widely used statistical software such as SPSS and SAS (Erceg-Hurn and Mirosevich, 2008). Furthermore, procedures like trimming and ranking are regarded with suspicion as it seems counterintuitive that the accuracy of a test could be improved by removing information (Wilcox, 2001). Nevertheless, research has shown that the use of modern methods aids to control Type I error and narrowing confidence intervals (e.g. Keselman et al., 1998; Lix et al., 1996) Fortunately, Erceg-Hurn and

Mirosevich (2008) provide a non-technical introduction to modern non-parametric methods (e.g. modern rank statistics and bootstrap).

Bootstrapping generates a predefined (large) number of new datasets from the original dataset to derive an empirical estimate of the distribution of a statistic such as parameter estimation, regression, prediction models, estimation of unknown variability and any analysis of a small representative sample (Mooney and Duval, 1993; Kelley, 2005; Erceg-Hurn and Mirosevich, 2008; Uraibi et al., 2009). Thus, it offers a solution to both the natural variability of stone affecting generated data and determining sufficient sample sizes reflecting on specific characteristics of any investigated stone type (Wilhelm et al., 2016a).

When non-parametric statistical measures are used data transformation is not necessary. Non-parametric summary statistics like median and median absolute deviation (MAD) are less affected by deviations from normality (Filzmoser and Todorov, 2013). When combined with non-parametric statistical techniques like bootstrapping, robust regression, Kruskal-Wallis and Mann-Whitney U tests, they may provide an appropriate solution to some of the challenges faced in stone weathering research such as variance of data derived from non-destructive index methods and natural stone with heterogeneous weathering patterns.

2.5. LITERATURE REVIEW CONCLUSION

The two main strands of investigation of this thesis are (1) improving selected non-destructive methods on fresh porous heritage limestone under controlled in the laboratory conditions for their eventual *in situ* application in objective 1 and (2) applying the improved non-destructive methods *in situ* to investigated heritage limestone weathering status and short- and long-term weathering at real heritage sites in the UK and Turkey in objective 2 and 3.

The topics discussed in the literature review will be addressed as follows:

- Applying modern statistical methods to determining sufficient sampling sizes, overcome operator variance, increasing reliability of results including outlier detection resulting in method improvement (Chapter 3, Paper 1)
- Investigating and overcoming potentially confining effects of salt content on moisture measurements and surface roughness on surface hardness testing resulting in method improvement (Chapter 3, Paper 1 and Paper 2)
- Adding more levels of information and gaining both surface and subsurface information on conditions of stone properties by extending applications of the three selected methods (Chapter 3, Paper2; Chapter 4, Paper 3; Chapter 5, Paper 4)
- Investigating and quantifying the interaction of intrinsic and extrinsic factors contributing to stone weathering behaviour using the improved non-destructive methods *in situ* by investigating the spatio-temporal character of short- and long-term time series (Chapter 4, Paper 3; Chapter 5, Paper 4)

3. Materials and methods

3.1. Introduction

3.1.1 Thesis structure

The four scientific papers in this study (chapter 5-6) cover the respective materials and methods sections for the three objectives. Therefore, this chapter introduces the underlying methodology of the whole project with 1) experimental set up and choice of scale, 2) choice of stone material, 3) destructive and non-destructive testing (not covered in the papers) and 4) data evaluation. Table 3.1 shows an overview of the individual parts of the overall methodology including the three objectives with their specific choice of spatio-temporal scale, limestone type, sample dimensions, locations, weathering-stress history, applied destructive and non-destructive methods and statistical data analysis.

3.2. Methods

3.2.1 Surface hardness testing

This method has been used throughout all 3 objectives and continuously improved over the course of the study by adapting data evaluation methods from Yilmaz (2013) to porous limestone in this study and combinations with water uptake and moisture meter measurements *in situ*. With these modifications information of the condition of surface and subsurface and the effect of porosity characteristics on the data was obtained. The results are reported in Paper 1 (chapter 4), Paper 3 (chapter 5) and Paper 4 (chapter 6).

3.2.2 Handheld electronic moisture meters

Similarly, to surface hardness testing, handheld electronic moisture meters have been used throughout all 3 objectives. In objective 1 the focus was on quantifying the effect of salt (sodium chloride) on the obtained moisture meter data. The results are reported in paper 2 (chapter 6). For objective 2 and 3 they functioned as an accompanying method as it is crucial for *in situ* measurements to determine the material moisture conditions in order to account for potential effects on the measurement methods (i.e. surface hardness testing and ultrasound).

3.2.3 Karsten tube

The Karsten tubes used in this study had an inner diameter of 26mm and were attached to the stone surface using putty (Plastic Fermit) (Figure 2.14). It was filled once with distilled water (4 ml) and subsequently the time measured (stopwatch) and recorded in 0.1 ml steps with the time noted accordingly (t). A minimum of 7 data pairs (ml and t, of D'ham et al., 2011; BS EN Standard 16302:2013) was collected and the application stopped after 60 minutes.

The overall rate of water uptake was determined. Furthermore, segmented linear regression models were fitted iteratively to detect break-points using the package 'segmented' in RStudio (Muggeo, 2003; Crawley, 2005; Muggeo, 2008) as trends in water uptake over time for the individual blocks were not linear.

Table 3.1 Overview methodology. Methods, test sites, stone types and data evaluation of this thesis.

	Objective 1		Objective 2	Objective 3
Research output	Paper #1	Paper #2	Paper #3	Paper #4
Testsite(s)/ Environment	Laboratory (OxRBL, UK)	Laboratory (OxRBL ,UK)	Isle of Portland (UK)	Gaziantep (Turkey)
Stone type(s)	Oolitic limestones	Oolitic limestones	Oolitic limestone	Limestone
Portland (Base Bed)	x	x	x	-
Portland (Whit Bed)	-	-	x	-
Bath (Hartham Park)	x	x	-	-
Clipsham	x	x	-	-
Guiting	x	x	-	-
Firat formation	-	-	-	x
Gaziantep formation	-	-	-	x
Temporal character	-	-	time-series	time-series
Time-scale	-	-	1 -c. 250 years	c. 1,800 years (origin) environmental exposure (< 10 years)
Weathering history (Complexity)	known	known	partly-known	mostly unknown
Weathering impacts	-	-	outdoor weathering Cfb1 (Koeppen climate map)	subterranean and outdoor weathering Csa1 (Koeppen climate map)
Spatial character	Samples	Samples	Blocks in situ	Blocks in situ
Sample dimensions/measuring area (per group)	30 x 8 x 5 cm 5 x 5 x 5 cm	30 x 8 x 5	0.2 - 0.4 m²	~ 0.75-1.00 m²
Salt	-	Sodium chloride	-	-
Equotip/Piccolo D&DL probe (surface	x	-	x	x

hardness)				
Protimeter (resistivity) (moisture testing)	-	x	x	x
Protimeter (capacitance) (moisture testing)	-	x	x	x
CEM (capacitance) (moisture testing)	-	x	x	x
Resipod (resistivity) (moisture testing)	-	x	-	-
Ultrasound	x	-	-	x
Karsten tube (water uptake)	-	-	x	x
UCS (BS EN 1926:2006)	x	-	-	-
Density / open porosity (BS EN 1936: 2006)	x	-	-	-
WAAP (BS EN 13755:2008)	x	-	-	-
Ion chromatography	-	x	x	x
In situ climate (temp./RH%)		x	x	x
Statistical evaluation				
Shapiro-Wilk	x	x	x	x
Outlier detection	x	x	x	x
Kruskal Wallis	x	-	x	
Mann-Whitney U	x	x	x	x
Spearman rank	x	-	x	-
Break-point analysis/ piecewise regression	-	-	x	x
Pearson's R^2	x	-	x	-
Quantile regression	-	-	x	-
Bootstrap	x	-	x	-
Confidence intervals	x	-	x	-
Model fit (linear/non-linear)	-	-	x	x

3.2.4 Data evaluation

RStudio (version 0.97.551) was used for statistical analysis throughout the whole thesis. In each study the Shapiro-Wilk test was used to test for non-normal distribution of the datasets.

Frequently non-normal data distribution was found as expected for porous and/or weathered limestone (Mosch and Siegesmund, 2007; Palmer, 2008). Porous limestone can exhibit non-normally distributed data even when fresh and tested under controlled laboratory conditions. To apply common statistics a non-normal distributed dataset needs to be modified. Furthermore, as discussed in section 2.4.2, transformation and modification of data does not always lead to an evaluable dataset (e.g. Alberti et al., 2013). Consequently, in order to account for inherent variability in natural stone properties (on-site) and to avoid the need for data transformation (no outlier removal), this study employed a range of non-parametric statistical measures and methods for data obtained from non-destructive testing throughout the whole study. Therefore, non-parametric measures and methods are more appropriate for both laboratory tests and *in situ*.

This study followed the approach of Aydin (2009) and no values were removed from the datasets. Instead outliers were identified in order to determine their number and gain potentially interesting information about inherent stone properties (i.e. porosity). To detect outliers the MAD was used and (x_i) the boundary for extreme values (outliers) was specified using (moderately conservative) 2.5*MAD following the recommendation of Leys et al. (2013) and shown in Equation 11:

$$Median - 2.5 * MAD < x_i < Median + 2.5 * MAD \quad \text{(Equation 11)}$$

Paper 1 (excerpt from Chapter 4)

For the first study the surface hardnesses for selected limestone were determined by mean and median with SD and MAD (respectively) for the two probes (D and DL). Single impact and repeated impact method was conducted and a range of surface hardness data collected (Table 3.2). Based on the results the robust hybrid dynamic hardness (HDH_{robust}) was calculated adapting Yilmaz' (2013) approach to porous limestone (Equation 12 and 13).

$$DR_{robust} = HLDL_{S.med} \,/\, HLDL_{R.med} \qquad \text{(Equation 12)}$$

The robust hybrid dynamic hardness (HDH_{robust}) is calculated as follows:

$$HDH_{robust} = DR_{robust} \; x \; HLDL_{S.med} = (HLDL_{S.med})^2 / \; HLDL_{R.med} \qquad \text{(Equation 13)}$$

Table 3.2 Overview of surface hardness data collected and calculated in this study (source: Paper 1)

Hardness unit	Definition
$HLD_{S.mean}$	D-probe, single impact method, mean
$HLD_{S.SD}$	D-probe, single impact method, standard deviation
$HLD_{S.med}$	D-probe, single impact method, median
$HLD_{S.MAD}$	D-probe, single impact method, median absolute deviation
$HLDL_{S.mean}$	DL-probe, single impact method, mean
$HLDL_{S.SD}$	DL-probe, single impact method, standard deviation
$HLDL_{S.med}$	DL-probe, single impact method, median
$HLDL_{S.MAD}$	DL-probe, single impact method, median absolute deviation
$HLD_{R.med}$	D-probe, median of the 3 highest values in each of the 3 repeated impact method (RIM) datasets of 20 readings
$HLDL_{R.med}$	DL-probe, median of the 3 highest values in each of the 3 repeated impact method (RIM) datasets of 20 readings
$HDH_{D.robust}$	D-probe, robust hybrid dynamic hardness (combination of SIM and RIM)
$HDH_{DL.robust}$	DL-probe, robust hybrid dynamic hardness (combination of SIM and RIM)

Pearson's R^2 and Spearman's rank correlation coefficient (ρ or rs) as a non-parametric version of the Pearson correlation coefficient) were used to evaluate which calculated hardness would best reflect on the porous character of the tested limestone. The Kruskal-Wallis test was used as a robust alternative to one-way ANOVA to evaluate significant differences between the tested limestone types and the two probes (D and DL) (Hodges and Lehmann, 1963). This was followed by further specifying the differences between the individual stone types using the Mann-Whitney U test (two-tailed test with a significance level of p-value 0.05, unpaired) as an alternative to the t-test (Hodges and Lehmann, 1963). The data were visualised using boxplots and density plots in order to determine skewness and detect outliers.

In a second step, the appropriate sample sizes for Equotip data collection on limestone was determined using the bootstrap technique to calculate confidence intervals for surface hardness median values. The desired sampling size would sufficiently reflect the true stone surface hardness, but also needed to be practical for on-site application. 120 surface hardness readings represented the true stone surface hardness ('population'). The original dataset was resampled without replacement (for each sample size this process was repeated a 100 times to simulate variation) using bootstrap for a range of smaller sample sizes (5, 10, 20, 45 and 60 readings). Finally, confidence intervals for the medians of the individual modelled sample size datasets were obtained through bootstrapping using the bias corrected and accelerated (bca) bootstrap for confidence intervals in R (10,000 times) with 95% confidence level. The appropriate sample size was determined by comparing the bootstrapped confidence intervals for the medians of the modelled sample size datasets to the original sample confidence intervals (using the original 120 readings) and calculating the differences of confidence interval widths in percentages.

Paper 2 (Excerpt from Chapter 4)

Similar to the approach in the first study here the median values and median absolute deviation (MAD, a robust measure for variance) were used for the moisture meter readings. Moisture meter readings of the same salt contaminated group (S0, S1, S2) under different RH% climates and the different NaCl contamination levels were evaluated with the two-tailed Mann-Whitney U test (significance level of p-value 0.05).

Paper 3 (Excerpt Chapter 5)

In the third study the approach from the first study was adapted and median of 30 surface hardness single impact measurements is expressed as $HLD_{S.med}$. To display surface hardness changes the third study introduces QC_{50} (gradient of the 0.50 quantile (median)) as a novel proxy for determining the rate of surface change. Non-crossing quantile regression for 0.25, 0.50, 075 quantile (bootstrapping 1,000 iterations) was applied. Quantile regression (in contrast to least-squares regression) is robust against outliers and heteroscedasticity (Cade and Noon, 2003; Koenker, 2005; Crawley, 2007; Dette and Volgushev, 2008). The 0.50 quantile shows the rate of change in median surface hardness over time. In addition, the 0.25 and 0.75 quantiles encompass the rate of change of the inter quartile range (IQR) of the datasets. This allows to investigate how the dispersion of surface hardness values changes over time develop and whether this happens in a homo- or heterogeneous manner. Understanding variance in changes over time can add to the understanding of stone weathering behaviour.

Paper 4 (Excerpt Chapter 6)

The fourth study used median and the median absolute deviation (MAD), boxplots and the Mann-Whitney U test to determine any significant differences in surface hardness between the two tested stone types and the different exposure periods (2005, 2007 and 2013).

3.3. LIMESTONE

The choice of the four oolitic limestone types tested in objective 1 was informed by the prevalent building stones for architectural heritage in Oxford and London, UK. Both cities hold an abundance of iconic buildings like Buckingham Palace and St Paul's Cathedral in London and the Radcliffe Camera and Sheldonian Theatre in Oxford (Figure 3.1). The majority of the original building stone for Oxford's architectural heritage is not quarried anymore (like Taynton or Headington limestone), instead for restoration campaigns alternative replacement stone is used.

Many oolitic limestones used for construction are relatively recent in the geological column, therefore, less crystalline, less dense and less resilient to many decay processes compared to geologically older limestones (Leary, 1983; Smith and Viles, 2006).

Figure 3.1 Oxford (UK), Radcliffe Camera (view from NW) patchwork of Taynton (yellowish-orange) and other Oxfordshire stone (buff coloured) and Clipsham (columns front, darker coloured stone)

Figure 3.2 illustrates the changing use of replacement limestone types in Oxford. The original material for the Sheldonian Heads (put in place in 1669) was Taynton limestone, which lasted for about 200 years. In the mid of the late 19th century (~1868) they were replaced a first time using Milton limestone. Inferior quality (strength) coincided with the uprise of the Industrial Revolution and led to rapid decay of the second set (Figure 3.3 middle).

Figure 3.2 Three generations of Sheldonian heads in Oxford (UK). Left probably 1st set from ~1669, possibly made from Taynton (the applied device is a handheld moisture meter Resipod (Proceq©) in 2014) now in Worcester College garden; Middle probably 2nd set ~1868 made from Milton stone according to Arkell 1947 (source: Eric de Mare 1970); Right 3rd (recent) set put in place 1972 made from Clipsham

To give an indication of the weathering-stress history limestones in Oxford were exposed to, in 1850 over 34.000 tons of coal were burnt in Oxford alone. The associated air pollution (sulphur dioxide (SO_2)) results in the formation of black crust on limestone and, despite potentially stabilizing the surface temporarily, it finally results in exfoliation and blistering. Thus, in the 19th century Oxford was well known for its black facades. In 1972 the third recent set of Sheldonian heads was installed, this time made from Clipsham limestone, which is a widespread replacement stone for architectural heritage (e.g. in 2014 used at the Radcliffe Camera during a restoration campaign of the middle storey). The preservation history of the Sheldonian heads shows, that knowledge about weathering resistance of stone under a given climate is inevitable, when concerned with sustainable architectural heritage preservation. All tested stones in objective 1 are either still relevant as building stone or as replacement stone.

3.3.1 Oolitic Limestone for architectural heritage sites in the UK (→ objective 1 and 2)

"The bones of the landscape are the rocks at the surface" (Powell, 2005, p. 8). One might as well add the rocks of the landscape provide the bones for architectural heritage. The surrounding geology often determined the used building stone (Adam, 1999). Cities like Oxford, Bath and London are located on a long belt of Jurassic limestone (Figure 3.3), which has shaped their appearance for centuries.

Figure 3.2 Geology map for the south UK with the quarry places of the four oolitic limestones tested in objective 1. London was added for orientation (Geological Map Data NERC 2015. ©Crown Copyright and Database Right 2015. Ordnance Survey (Digimap Lisence)

The oolitic limestones of objective 1 contain more than 95% $CaCO_3$ and according to Mosch and Siegesmund's classification are low density (<2.6g/m^3) and highly porous stones (>10%). Table 3.3 summarizes the lithology and index properties of the investigated stone varieties in objective 1 and 2 and Figure 3.5 shows the respective quarry locations. As discussed in chapter 2 stone variety and especially porosity is determined for resistance and weathering behaviour. As can be clearly seen in Table 3.3 noticeable variability in natural properties is evident for the tested limestone types.

3.3.2 Portland limestone

The most prominent of the four tested stones is Portland limestone. Known since Roman times, Christopher Wren initiated its renaissance in the 18th century when London was rebuilt after the great fire. It was recently nominated as a "Global Heritage Stone Resource" (international recognition of those natural stone resources that have achieved widespread utilisation in human culture; Hughes et al., 2013) and has national and international reputation; prominent examples are St Paul's Cathedral in London and Exeter Cathedral in Exeter. It further was the material of choice for the Commonwealth War Grave Commission gravestones with over 300,000 alone in the UK of which 20,000 either repaired or replaced per year (Bell and Coulthard, 1990; Godden, 2012; Viles, 2013; CWGC 2015).

Two Portland limestone varieties are relevant for built heritage, Portland Base Bed and Portland Whit Bed. As both varieties are discussed in detail in Paper 3 (objective 2), this section summarizes important facts about the two varieties. Although deriving from the same formation, it is crucial to distinguish Portland Base Bed from Portland Whit Bed. The latter has a reputation of being the more durable building stone, which is linked to beneficial pore characteristics (e.g. (Leary, 1983; Dubelaar et al., 2003; Godden, 2012). Of particular interest here is microporosity (<5 µm), which has been documented to increase water retention and plays a crucial role in limestone weathering behaviour as it correlates with poor durability (Palmer, 2008; Yu and Oguchi, 2010; Dubelaar et al. 2003). Dubelaar et al. (2003) determined (with mercury porosimetry) a high proportion of micropores (~ 75%) for Portland Base Bed.

3.3.3 Clipsham limestone

Clipsham belongs to the Lincolnshire Limestone Formation (Bajocian Age, <165 Mio), which is the most thickly developed Middle Jurassic limestone unit in the East Midlands. Two units are distinguished, lower and upper Lincolnshire Limestone, of which Clipsham belongs to the latter (British Geological Survey). Clipsham limestone has significance as replacement stone and has been extensively used in Oxford (e.g. All Soul's, Christ Church and New College) and elsewhere like Windsor Castle in the 14th Century and the Palace of Westminster in the 20th Century, but similarly to Portland limestone has been known since Roman times and still has significance as building stone (BGS 1997abc).

3.3.4 Bath limestone

Bath limestone belongs to the Middle Jurassic (late Bajocian to early Bathonian age, y170 Mio); Great Oolite Group. It is quarried around Bath, Avon (Ashurst, 1998). Similarly, to Portland and Clipsham limestone, Bath limestone has also been used since Roman times (Ashurst, 1998). In terms of stone property characteristics Bell (1993) classifies it as moderately strong and complements well the range of different porosities and unconfined compressive strengths for this study.

3.3.5 Guiting limestone

Guiting limestone is an oolithic limestone and belongs to the Inferior Oolite of middle Jurassic age (late Bajocian to early Bathonian age, y170 Mio) (BRE, 1997c). The stone is sourced from a quarry at Ford which is north-west of Stow-on-the-Wold (BRE, 1997). The stone had been subjected to weathering studies before by Aliha (2012) and Bell (1993) and was chosen in this study for its low unconfined compressive strength and high porosity, which was thought to generate non-normally distributed data. Thus, in order to quantify the effect of stone variability on the non-destructive methods this soft stone was included.

Table 3.3 Physico-mechanical properties of the tested stone in objective 1 and 2 (derived using standard procedures) and surface hardness results D probe (HDL) and DL probe (HLDL). Water absorption under atmospheric pressure (WAAP) was tested using BS EN 13755. Unconfined compressive strength (UCS) was tested using BS-EN 1936:2006

Stone type	UCS [MPa] (min-max of n=10)	Open porosity [%wt] (min-max of n=6)	WAAP [Mass %] (min-max of n=6)	Apparent density (min-max of n=6)	HDL median MAD (Leeb)	HLDL median MAD (Leeb)
Portland Jordans Basebed	(43.20-75.73) μ 55.98	(13.12 – 13.82) μ 13.5	(6.49 – 6.87) μ 6.71	(2177.65-2223.31) μ 2205.99	469 27	525 20
Bath Hartham Park	(14.32-20.09) μ 16.04	(21.11-23.51) μ 22.2	(11.07-12.68) μ 11.84	(1954.6-2017.51) μ 1984.45	241 38	297 40
Clipsham	(17.37-50.65) μ 26.71	(12.48-17.97) μ 15.63	(6.23-9.53) μ 7.89	(1975.66-2284.59) μ 2123.27	318 62	422 76.5
Guiting	6.15 – 17.15 μ 11.15	(16.1 – 24.96) μ 21.3	(7.94 – 14.23) μ 11.55	(1796.41-2376.79) μ 2004.54	215 22.5	266 30.5

3.3.6 Firat and Gaziantep limestone (Objective 3)

The choice of Firat and Gaziantep limestone for objective 3 was informed by a collaboration with archaeologists from the University Münster at Dülük Baba

Tepesi, an archaeological excavation site in Southern Turkey. At the site dramatic deterioration of soft limestone Hellenistic-Roman remains recently excavated has been observed following the cold, wet winter of 2011/2012.

Firat and Gaziantep limestone are quarried in the Gaziantep region in Southern Turkey. The Firat formation belongs to the lower middle Oligocene (Chattian age, <28Ma) and the Gaziantep formation derives from the middle upper Eocene (<37 Ma) (Baykasoglu et al., 2008). The two varieties the main types found at the archaeological excavation site in Dülük Baba Tepesi, where this study conducted *in situ* stone weathering research with focus on catastrophic decay and extreme climatic impact.

The English written literature often discusses both types though mostly under geological considerations (e.g. Coskun, 2000; Dagistan and Simsek, 2005) and thus with limited implication for stone weathering problems in heritage conservation. However, the Gaziantep formation recently gained research interest due to problems with collapsing caves in the city of Gaziantep. Çanakci (2007) discusses strength issues of the Gaziantep formation, which loses around 50% of compressive strength, when tested fully saturated. This information is invaluable for this study with regards to investigating the same stone type, but as archaeological temple remains being exposed to the environment where rainfall might cause a similar (if though temporary) saturation of the surface zone of the archaeological remains (Table 3.4).

Table 3.4 Existing research on lithology and index properties of limestone Gaziantep and Fırat Formation. Modified after [1]Kaymakci, 2010; [2]Robertson et al., 2015; [3]Dagistan, 2005; [4]Coskun, 2000; [5]Çanakci et al., 2007; [6]Türkkan, 2011; [7]Baykasoglu, 2008; [8]Çanakci, 2007; [9]Özvan et al., 2010; (*Karabakir, **Hamdi Kutlar (investigated collapsed caves in Gaziantep (Çanakci, 2007)).

Index test	Gaziantep formation (Tmga)	Fırat formation (Tmf)
Lithology	"Limestone with cherty intervals and cherty nodules"[1]; "Chalky"[2] ; "argillaceous limestones, white, grey"[4]; "heterogeneous rock"[5]; "contains large gravel particles [...] crystalline silica"[5]	"Chalky"[1]; "cream-grey coloured, hard and brittle reefbank type limestones"[3] "weathered surfaces are dark yellow- reddish , hard, medium - weak strength, freshly broken surface beige"[6]
Mineralogy	97% $CaCO_3$, rest: SiO_2, MgO, Al_2O_3 [7]	96.46% $CaCO_3$, 0.28 SiO_2, 0.08 Fe_2O_3, 1.48 $MgCO_3$, Al_2O_3, rest 1.7[6]
Dry unit weight (kN/m³)	16.76*[8], 16.99[5], 17.3[7] , 18.64**[8], 19.1[7], 23.21[5]	26.8[6]
Saturated unit weight (kN/m³)	20.2[7], 20.6**[8], 20.79*[8]	-
Water absorption by weight (%)	11**[8], 13[7], 18[7], 24*[8]	0.8[6]
Compressive strength, dry (MPa)	10.2**[8], 10.7[7], 25.51*[8], 25-68[9]	72.12[6]
Compressive strength, saturated (MPa)	5.36**[8], 11.53*[8]	-
Tensile strength, dry (MPa)	2.41**[8], 3.12*[8], 3.8[7]	-
Tensile strength, saturated (MPa)	0.31**[8], 0.65*[8]	-
UPV, dry (m/s)	2656**[8], 2637[7], 2906*[8] , 3380[7]	-
Modulus of elasticity, dry (GPa)	4.45**[8], 11.3*[8]	-
Porosity (%)	-	1.7[6], 10[3]
Schmidt Hammer	-	50.5[6]
Thermal conductivity	0.9264-2.5158 W/mK[5]	-

4. Objective 1: Low impact hardness testing and handheld moisture meters – Improvement and developing of guide for good practice under controlled laboratory conditions

This chapter presents the main findings of objective 1, which are summarized and published in two papers (part of this chapter).

Paper 1 (Objective 1): Improving Equotip hardness testing methodology in rock weathering and stone deterioration research

Paper 1 has been published in Earth Surface Processes and Landforms (re-submitted and in revision process). The paper addresses the following research questions: How do the Equotip D and DL probes compare? Is the Equotip appropriate for application on porous stone? How to address effects like surface roughness? What are the most appropriate statistical methods to handle Equotip data? How should outliers be treated? And what is an adequate sample size to collect?

Paper 2 (Objective 1): The influence of salt on handheld electrical moisture meters: Can they be used to detect salt problems in porous stone?

Paper 2 has been published in The International Journal for Architectural Heritage (re-submitted and in revision process). The aim of this paper is to shed some light on the influence of salt contamination on selected handheld moisture meters, and to evaluate the potential of these effects to be used to diagnose salt and moisture problems in stone heritage.

4.1. PAPER1_LOW IMPACT SURFACE HARDNESS TESTING (EQUOTIP) ON POROUS SURFACES – ADVANCES IN METHODOLOGY WITH IMPLICATIONS FOR ROCK WEATHERING AND STONE DETERIORATION RESEARCH

Katrin Wilhelm[1*], Heather Viles[1], and Órlaith Burke[2]

[1] Oxford Rock Breakdown Laboratory (OxRBL), School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, UK

[2] Nuffield Department of Population Health, University of Oxford, Oxon OX3 7LF, UK

[*] katrin.wilhelm@ouce.ox.ac.uk

Abstract

The Equotip surface hardness tester is becoming a popular method for rock and stone weathering research. In order to improve the reliability of Equotip for on-site application this study tested four porous limestones under laboratory conditions. The range of stone porosity was chosen to represent likely porosities found in weathered limestones in the field. We consider several key issues: (i) its suitability for soft and porous stones; (ii) the type of probe required for specific on-site applications; (iii) appropriate (non-parametrical) statistical methods for Equotip data; (iv) sufficient sampling size. This study shows that the Equotip is suitable for soft and porous rock and stone. From the two tested probes the DL probe has some advantages over the D probe as it correlates slightly better with open porosity and allows for more controlled sampling in recessed areas and rough or curved areas. We show that appropriate sampling sizes and robust non-parametric methods for subsequent data evaluation can produce meaningful measures of rock surface hardness derived from the Equotip. The novel Hybrid dynamic hardness, a combination of two measuring procedures (single impact

method (SIM) and repeated impact method (RIM)), has been adapted and is based on median values to provide a more robust data evaluation. For the tested stones in this study we propose a sample size of 45 readings (for a confidence level of 95%). This approach can certainly be transferred to stone and rock with similar porosities and hardness. Our approach also allows for consistent comparisons to be made across a wide variety of studies in the fields of rock weathering and stone deterioration research.

Keywords: rock and stone surface hardness testing; Equotip; limestone; non-parametric statistics; outliers

4.1.1 Introduction

Weathering manifests itself in the near surface zone as changes in stone properties such as porosity and intergranular bonds (McCabe et al., 2015). Quantifying these changes is important for rock weathering and stone deterioration research to understand spatio-temporal weathering behaviour and establishing decay rates (e.g. Meierding, 1993; Inkpen et al., 2012). Results may further inform decision making on heritage conservation strategies and provide hard evidence of stone response to impacts such as climate-change and air-pollution (e.g. Ross and Butlin, 1989; Smith et al., 2011; Viles and Cutler, 2012).

Surface induration or weakening are common property changes induced through environmental impacts (e.g. Inkpen et al., 2012; Moses et al., 2014). A common method to investigate such surface changes on-site is surface hardness testing. As a portable, non-destructive method it avoids the need to take samples as required to perform other common destructive tests like unconfined compressive strength. Although a long-established proxy method for relative dating of surface exposure in geomorphology (e.g. Aydin and Basu, 2005; Goudie, 2006; Fort et al., 2013; Stahl et al., 2013), only a few studies in built heritage science have used this method to quantify the state of preservation or deterioration of monuments (e.g. Török 2003, 2007, 2008; Cutler et al. 2013; Fort et al., 2013).

The most popular device for geomorphological applications is the Schmidt Hammer (e.g. Aydin and Basu, 2005; Goudie, 2006; Fort et al., 2013; Stahl et al., 2013). However, due to its high impact energy (Type L = 0.735 N m and type N = 2.207 N m its application on soft and porous or easily damaged stone is limited (Pope, 2000; Viles et al., 2011). In contrast, the impact energy of the Equotip with probe D is 0.0115 N m which is only a fraction of that of the Schmidt Hammer (probe versions with similarly low impact energy are Type C = 0.003 N m and Type G = 0.090 N m (Proceq© SA, 2010). Therefore, the Equotip is suitable for measuring a wide range of stone and rock surfaces (e.g. gypsum, tuff, limestone, granite) at different stages of weathering, as well as detecting subtle changes in surface hardness (e.g. Hack et al., 1993; Verwaal and Mulder, 1993; Aoki and Matsukura, 2007; Viles et al., 2011; Alberti et al., 2013; Coombes et al., 2013; Hansen et al., 2013). Low rebound values indicate soft, porous and/or weathered stone surfaces, higher values less weathered or case hardened surfaces.

The overall aim of this study is to develop a reliable methodology for using the Equotip for rock weathering and stone deterioration research. This paper answers the following questions: How do the Equotip D and DL probes compare? What are the most appropriate statistical methods to handle Equotip data? How should outliers be treated? And what is an adequate sample size to collect on porous stone?

The Equotip family of devices and probes

The Equotip devices relevant for this paper are Equotip 3 and Equotip Piccolo 2, which come with a range of different probes (Table 4.1). They measure the difference between impact and rebound velocity of a (small) hard metal impact body traveling in a probe and propelled by spring force against the surface (Proceq© SA, 2010). The D probe is the most commonly used in stone weathering research to date with a small impact body (27 mm) measuring 3 mm in diameter (Figure 4.1). In contrast, the DL probe has a slim long (82 mm) front section and slightly smaller diameter end (2.78 mm) (Figure 4.2), and is suitable for confined spaces and recessed surfaces (Proceq© SA, 2010). To our knowledge the DL probe has not been trialled for rock weathering or stone deterioration research. It may provide a useful addition to weathering studies for collecting data on rough and / or porous surfaces.

Figure 4.1. Equotip Piccolo 2 with impact body D on-site at Radcliffe Camera, Oxford.

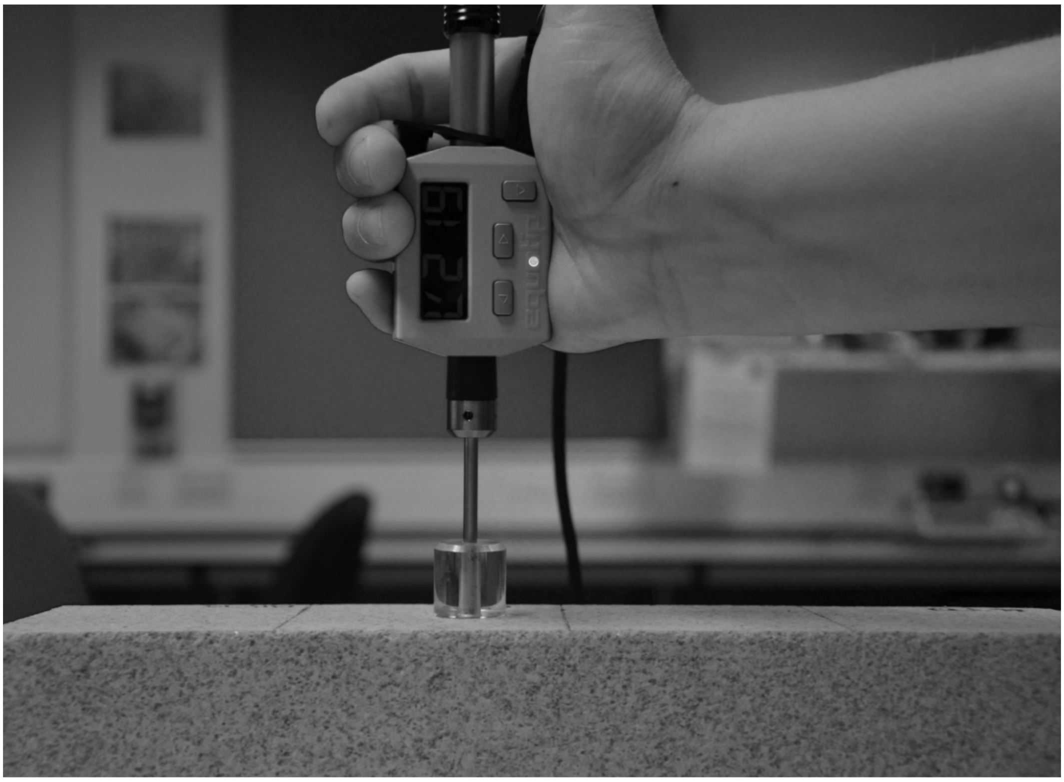


Figure 4.2 Equotip Piccolo 2 with impact body DL in the Oxford Rock Breakdown Laboratory (OxRBL).

Table 4.1 Key characteristics of the Equotip 3 and Equotip Piccolo 2 mobile hardness testing devices

Key characteristics of probes			Equotip 3	*Equotip Piccolo 2*
Impact body	Ball intender type and diameter	Impact energy (Nmm)		
C	tungsten carbide, 3.0 mm	3.0	x	
D	*tungsten carbide, 3.0 mm*	*11.5*	*x*	*x*
DC	tungsten carbide, 3.0 mm	11.5	x	
DL	*tungsten carbide, 2.78 mm*	*11.1*	*x*	*x*
E	polycrystalline diamond	11.5	x	
G	tungsten carbide, 5.0 mm	90.0	x	
S	ceramics, 3.0 mm	11.5	x	
Measuring range			1 – 999 HL	*150 – 950 HLD / 250 – 970 HLDL*
Measuring accuracy			± 4 HL (0.5% at 800 HL)	*±4 HL (0.5% at 800 HLD / HLDL)*
Impact direction automatic compensation			Yes (except DL probe)	*Yes*
Software			x	*x*
Weight			780 g	*142 g*

Note: Devices and probes tested in this study are shown in italic typeface.

The main differences between the Equotip devices are the range of impact bodies that can be attached to them and the evaluation software. The Equotip 3 is more versatile and comes with a separate recording unit. Obtained data is directly comparable for both devices, when using the same probe type, whereas the probes are not comparable among themselves (i.e. the DL probe gives higher readings than the D). Hardness data is expressed on the 'Leeb hardness' scale (1 - 999) and can be converted directly to all common hardness scales (e.g. Vickers, Rockwell etc.(Proceq©)). Furthermore, data is stored automatically and Equotip

3 and Equotip Piccolo 2 calculate and record basic descriptive statistics such as mean values and standard deviation (SD) during the measurement process.

Challenges for using Equotip devices on rock and stone surfaces

While Equotip devices offer a useful non-destructive means of testing the hardness of stone and rock surfaces, there are several challenges associated with the use of this equipment.

Effect of natural variability of rock and stone and weathered surfaces on Equotip data

Feal-Pérez and Blanco-Chao (2012) find surface roughness of weathered clasts affects Equotip measurements on-site. Similarly Aoki and Matsukura (2008) report data scatter obtained from unweathered limestone and andesite surfaces due to subtle roughness and large pores of the particular stone types. Thus, natural property variations of fresh stone have an effect on Equotip data and such variations are likely to increase as weathering proceeds. Nevertheless, McCarroll (1991), who observed a similar effect for Schmidt Hammer measurements, states that surface roughness and weathering are intimately related. Therefore, instead of defining it as a limitation he suggested it could be utilized for comparison in cases where "surfaces have displayed similar surface textures prior to the influence of weathering" (McCarroll 1991, p. 479).

Previous research has found good correlations between Equotip measurements and unconfined compressive strength (Alvarez Grima and Babuška, 1999; Aoki and Matsukura, 2008; Yilmaz, 2013). Yilmaz (2013) tested a range of unweathered carbonate rocks (dolomite, limestone, travertine and marble) with densities between 2.24 up to 2.80 g/cm^3 and open porosities between 0.14 and 7.00 %. Aoki and Matsukura (2007) tested weathered sandstone on-site, which originally had a density of 2.69 g/cm^3 and open porosity of 6.9%. Both Yilmaz (2013) and Aoki and Matsukura (2007) utilized two Equotip application methods. For the single impact method (SIM) individual measurements are randomly distributed over the stone surface. Obtained values reflect on the elastic and plastic properties of the stone surface. In contrast, with the repeated impact method (RIM) repeated measurements on one point are taken, which reflects the elastic and plastic properties of the surface and subsurface of the stone. Yilmaz (2013) and Aoki and Matsukura (2007) combined both methods to gain deeper insight in stone surface and subsurface characteristics. Aoki and Matsukura (2007) introduced the *k*-value, whereas Yilmaz (2013) calculated the hybrid dynamic hardness (HDH) measure. For both porosity characteristics of the stone are taken into account and thus, natural stone variations are better reflected.

Methodology gaps

At present there is no consensus on methodology for the use of Equotip devices in the field or laboratory, nor in the evaluation of the data obtained (Viles et al., 2011; Yilmaz, 2013). This is a major limitation if reliable and comparable data are to be collected by different studies. Table 4.2 summarises the approaches taken by a range of researchers using the Equotip within the geomorphology and heritage science fields, and illustrates the need for further investigations into the most efficient sample size, and the best approach to statistical analysis given variable and often non-normal data, with outliers. As explained earlier several different methods can be used to quantify surface hardness with Equotip devices, including the SIM, RIM, and combinations of the two using *k*-value (Aoki and Matsukura, 2008) or hybrid dynamic hardness (HDH) (Yilmaz, 2013). These offer solutions to address problems like surface roughness and porous stone and have been utilized and adapted for porous limestone using alternative statistical approaches in this study.

What sample size is needed to get reliable data from rock and stone surfaces?

One key issue that needs to be addressed when applying Equotip devices to stone and rock surfaces is the number of readings that should be taken, and how this affects the reliability of statistical tests applied to the data collected. For example, studies with the Schmidt hammer have shown that the number of readings taken has bearing on the meaningfulness of subsequent statistical tests (Niedzielski et al., 2009). The implication is that only a sufficiently big sample size will reflect the true surface hardness of a material, and how big is sufficient depends on the material being tested and its weathering-stress history. Table 4.2 shows sample sizes used in a selection of previous studies that have applied the Equotip 3 and Piccolo 2 devices. The number of readings taken ranges from 10 (Aoki and Matsukura, 2007) up to 80 (Coombes et al., 2013). It is not clear, however, how well any of these sample sizes used reflect the true surface hardness and Viles et al. (2011) suggest that a sample size of > 50 is needed in some circumstances. No consistent approach has been taken in previous research, and no justification has been given for the choice of sample sizes in most of these studies. How can researchers cope with natural variability of stone and the need for large sample sizes?

Table 4.2 Existing research on Equotip in the field of rock and stone testing

Study	Device - Probe	Tested stone types / location	Sample size SIM/(RIM)	Application method/ Surface preparation (Y/N)	Data evaluation / Test for normality (Y/N) / Outliers (Y/N) Modification (Y/N)
Aoki and Matsukura (2007)	Equotip 3 - D	Sandstone /On-site	10 (20)	SIM, RIM, k-value / N	Parametric / n.a. / n.a. / n.a.
Viles et al. (2011)	Equotip 3 - D Piccolo 2 - D	Limestone, sandstone, dolerite, basalt /On-site	50	SIM / Y & N	Parametric / n.a. / n.a. / n.a.
Mol and Viles (2012)	Equotip 3 - D	Sandstone /On-site	10	SIM / N	Parametric / n.a ./ n.a. / n.a.
Yilmaz (2013)	Equotip 3 - D	Limestone, dolomite, marble, travertine /Laboratory	20 (10-20)	SIM, RIM, HDH / Y	Parametric / n.a. / n.a. / n.a.
Coombes et al. (2013)	Equotip 3 - D	Limestone, granite, concrete /On-site	80	SIM / N	Parametric / Y / n.a. / n.a.
Alberti et al. (2013)	Equotip 3 - D	Quartzite /On-site	600 total on 25 clasts at each of 24 outcrops	n.a./ Y	Parametric and non-parametric/ Y / Y / Y
Hansen et al. (2013)	Equotip 3 - D	Dolerite /On-site	15 per aspect, per clast (210 values in total)	SIM / N	Parametric / n.a./ n.a./ Y

[1] SIM = single impact method, RIM = repeated impact method with combinations of the two (SIM and RIM) being k-value and HDH = Hybrid Dynamic Hardness, Y=Yes, N=No (table modified after Yilmaz (2013).

What is the best statistical methodology to handle Equotip data?

As well as being more variable than data from the Schmidt Hammer (Viles et al., 2011), it is likely that Equotip data obtained from porous and/or weathered rock and stone surfaces will be non-normally (asymmetrically) distributed. Accordingly, Hansen et al. (2013) and Alberti et al. (2013) find that Equotip data derived from on-site measurements on weathered stone are affected. Nevertheless, standard parametric statistical methods were employed (e.g. t-test, analysis of variance (ANOVA), Fisher`s least significant difference (LSD)), whereas for the evaluation of Equotip (and Schmidt Hammer) data robust methods may have been more beneficial (i.e. Mann-Whitney U, Kruskal-Wallis test and Spearman correlation)(Niedzielski et al., 2009). In cases of non-normal data, the reliability of statistical estimates based on the assumption of normality may be affected and parametric tests are largely inappropriate (Tukey, 1977; Fowler et al., 1998; Filzmoser and Todorov, 2013).

Data transformation

Semi-parametric tests (a hybrid of parametric and non-parametric (Powell, 1996) are one solution to treat non-normal data, and have been applied to Equotip data by Alberti et al. (2013). However, semi-parametric tests often require data modification. This involves decision making (i.e. normalization, defining thresholds, trimming, outlier-detecting, outlier removal etc.) before analysis using appropriate methods (Reimann, 2008; Good and Hardin, 2009). Depending on the statistical program used to define outliers, different procedures can be applied, and these are not always obvious or consistent between different studies. Furthermore, transformation and modification of data does not always lead to an evaluable dataset. For example, Alberti et al. (2013) modified 24 Equotip datasets using two methods (in one instance using only the 50% highest values and in another removing the eight extreme values from datasets), and yet some datasets remained non-normally distributed.

Outliers

One factor associated with non-normal data is the occurrence of outliers, which may be present in a dataset as a result of human and / or instrument error, or due to natural deviations in the sample population (Hodge and Austin, 2004). Outliers are frequent in Equotip datasets (Viles et al., 2011). A common approach to outliers in classical statistics is to remove them entirely from the sample, as they place restrictions on subsequent data evaluation (Rosner, 1983). However, outliers should only be removed when it is clear that their occurrence is not related to the population characteristics but have resulted from errors in the data gathering process (Field, 2009). Identifying outliers is an important part of any statistical evaluation (Lipfert, 1989; Banerjee and Iglewicz, 2007), including Equotip data, as they can provide useful information about the sample in their own right (Iglewicz and Hoaglin, 1993). However, where outliers are to be retained, such as when they are deemed to reflect inherent, true variability in the hardness of a deteriorating stone for example, a new approach to statistical evaluation is required.

Statistical analysis using robust measures and bootstrap

Data transformation is not necessary when robust statistical measures are used. Robust summary statistics like median and median absolute deviation (MAD) are less affected by deviations from normality (Filzmoser and Todorov, 2013). When combined with non-parametric tests like Kruskal-Wallis and Mann-Whitney U test, they may provide an appropriate solution to some of the challenges associated with the analysis of Equotip datasets. Furthermore, the bootstrap technique as robust statistical techniques offer a solution to both the natural variability of stone affecting generated data and determining sufficient sample sizes reflecting on specific characteristics of any investigated stone type. Bootstrapping generates a predefined (large) number of new datasets from the original dataset to derive an empirical estimate of the distribution of a statistic like mean, median or confidence intervals (Mooney and Duval, 1993; Kelley, 2005). Mooney and Duval (1993) state that bootstrapping has advantages over traditional parametric statistical approaches. The latter derive probability based inferences from a sample by distributional assumptions (usually normal distribution assumed) and analytic formulas (Mooney and Duval, 1993). In contrast bootstrapping replaces those theoretical formulations by resampling with replacement from the original dataset (Erceg-Hurn and Mirosevich, 2008; Uraibi et al., 2009). Thus, rather than drawing conclusions from potentially unrealistic assumptions (using traditional approaches) bootstrapped empirical estimates of statistical quantities of interest (mean, median or confidence intervals) can further improve statistical analyses such as parameter estimation, regression, prediction models, estimation of unknown variability and any analysis of a small representative sample (Erceg-Hurn and Mirosevich, 2008; Uraibi et al., 2009). Bootstrapping is unaffected by non-normality in the original dataset to which it is applied, as is common for surface hardness data obtained from porous and weathered stone on-site. Therefore, robust bootstrapping may be used to reduce bias in statistical estimations derived from porous stone.

4.1.2 Materials and methods

Stone samples

Stone types tested

The tests were conducted on four porous (oolitic) limestones that have been widely used in built heritage in the City of Oxford, including the Radcliffe Camera and the University Church of St. Mary the Virgin. Table 4.3 summarizes the limestone properties of the following types Portland (Jordans Base Bed), Bath (Hartham Park), Clipsham and Guiting. Stone samples were obtained fresh from quarries and cut to 300 mm x 80 mm x 50 mm dimensions. Porosity has been found to influence surface hardness testing (Aoki and Matsukura, 2008) and thus limestones with a wide range of porosity values were used in this study (13.5 – 22.2%). Unconfined compressive strength (UCS), open porosity and water absorption under atmospheric pressure were determined following BS-EN standards 1926:2006, BS-EN 1936:2006 and BS-EN 13755:2008, respectively (British Standards Institute, 2006q, 2006b, 2006c). UCS was determined with 10 cubes per stone type in order to determine the correlation with surface hardness values as regression can vary for different rock types (Dinçer et al., 2004). Open porosity was determined using six cubes (50 mm x 50 mm x 50 mm dimensions) for each limestone type.

Table 4.3 Physico-mechanical properties of the tested stone (derived using standard procedures) and surface hardness results D probe (HLD) and DL probe (HLDL).

Stone type	UCS [MPa] (min-max of n=10)	Open porosity [%] (min-max of n=6)	WAAP [Mass %] (min-max of n=6)	Apparent density [kg/m³] (min-max of n=6)
Portland Base Bed	µ 55.98 med 52.65 (43.20-75.73)	µ 13.5 med 13.63 (13.12 – 13.82)	µ 6.71 (6.49 – 6.87)	µ 2205.99 (2177.65-2223.31)
Bath Hartham Park	µ 16.04 med 15.76 (14.32-20.09)	µ 22.2 med 22.11 (21.11-23.51)	µ 11.84 (11.07-12.68)	µ 1984.45 (1954.6-2017.51)
Clipsham	µ 26.71 med 26.19 (17.37-50.65)	µ 15.63 med 16.33 (12.48-17.97)	µ 7.89 (6.23-9.53)	µ 2123.27 (1975.66-2284.59)
Guiting	µ 11.15 med 11.82 (6.15 – 17.15)	µ 21.3 med 21.94 (16.1 – 24.96)	µ 11.55 (7.94 – 14.23)	µ 2004.54 (1796.41-2376.79)

[1] Water absorption under atmospheric pressure (WAAP) was tested using BS EN 13755. Unconfined compressive strength (UCS) was tested using BS EN 1926:2006 and for porosity BS EN 1936:2006; µ=mean, med=median.

Sample dimensions and preparation

Three replicate blocks for each limestone type were tested with the Equotip. The measurement surface (top face, 300 mm x 80 mm in dimensions) of each specimen was finished with P120 sandpaper prior to measurement, in order to minimise measurement error and to make sure that all values obtained were 'true' values and any outliers were due to the inherent, true variability in the hardness of the limestone (i.e. porosity rather than roughness). The device was applied perpendicular to the bedding of the blocks, which were placed on a solid limestone base to prevent interference from vibration.

Equotip Piccolo 2 with D- and DL-probe

Most previous geomorphological studies have employed the Equotip 3 in combination with the D probe (e.g. Hack et al., 1993; Coombes et al., 2013). This study used the Equotip Piccolo 2 (referred to as Equotip in this paper), which in terms of impact energy and measurement scale is comparable to the Equotip 3 but more portable. The principles tested in this paper for the Equotip Piccolo 2 are equally applicable to the Equotip 3 device. The DL probe was used as well as the D probe given its advantage of being able to obtain readings in confined spaces. The Equotip was frequently checked for calibration and all measurements (except for the assessment of operator variance) were conducted by the same operator (first author) under laboratory conditions.

Surface hardness test procedure

In this study SIM and RIM were applied and HDH calculated. For SIM the Equotip randomly applied 120 times distributed over an area covering about 720 cm^2 (total of surface area of three blocks per group). For RIM this study followed the approach of Aoki and Matsukura (2008) and collected 20 RIM values. For further data analysis the median of the highest values in each of the three RIM testing dataset per limestone type was calculated.

Operator variance

Within the scope of this study a pilot study was conducted to assess operator variance. Three operators with varying experiences towards the Equotip device (experienced and inexperienced) applied the Equotip with the D probe 20 times to a metal test block provided by Proceq (type: calibration block for D probe for high hardness range ~55.2HRC). Two different standards to assess Leeb hardness tester accuracy can be applied, DIN 50156 and ASTM A956 (Pollok and Mennicke, 2010). Depending on the standard the Equotip with D probe is considered to be calibrated when the mean value of >three readings on the test block are HLD 765 with a tolerance of ± 6 (ASTM A956) or ± 15 (DIN 50156). For this study the latter tolerance was used.

Statistical data analysis and sample size determination

The statistical data analysis was two-fold. In a first step SIM mean and median with SD and MAD (respectively) were determined for the two probes (D and DL). Based on these values the HDH was calculated. The hardness data collected and calculated in this study are shown in Table 4.4. In view of potential on-site Equotip application to porous and weathered stone (which might display increased porosity) regression analysis (Pearson's R^2 and Spearman's rank correlation coefficient (ρ or rs) as a non-parametric version of the Pearson correlation coefficient) were used to evaluate which calculated hardness would best reflect on the porous character of the tested limestone.

In a second step, the appropriate sample sizes for Equotip data collection on limestone was determined using the bootstrap technique to calculate confidence intervals for surface hardness median values. For statistical analysis RStudio (version 0.97.551) was used.

Table 4.4 Overview of surface hardness data collected and calculated in this study

Hardness unit	Definition
$HLD_{S.mean}$	D-probe, single impact method, mean
$HLD_{S.SD}$	D-probe, single impact method, standard deviation
$HLD_{S.med}$	D-probe, single impact method, median
$HLD_{S.MAD}$	D-probe, single impact method, median absolute deviation
$HLDL_{S.mean}$	DL-probe, single impact method, mean
$HLDL_{S.SD}$	DL-probe, single impact method, standard deviation
$HLDL_{S.med}$	DL-probe, single impact method, median
$HLDL_{S.MAD}$	DL-probe, single impact method, median absolute deviation
$HLD_{R.med}$	D-probe, median of the 3 highest values in each of the 3 repeated impact method (RIM) datasets of 20 readings
$HLDL_{R.med}$	DL-probe, median of the 3 highest values in each of the 3 repeated impact method (RIM) datasets of 20 readings
$HDH_{D.robust}$	D-probe, robust hybrid dynamic hardness (combination of SIM and RIM)
$HDH_{DL.robust}$	DL-probe, robust hybrid dynamic hardness (combination of SIM and RIM)

Adapting Yilmaz' (2013) approach for porous limestone this study combined SIM and RIM based on median hardness to calculate the deformation ratio (DR) and HDH (see Equations 14 and 15).

$$DR_{robust} = HLDL_{S.med} \,/\, HLDL_{R.med} \qquad \text{(Equation 14)}$$

The robust hybrid dynamic hardness (*HDH*$_{robust}$) is calculated as follows:

$$HDH_{robust} = DR_{robust} \; x \; HLDL_{S.med} = (HLDL_{S.med})^2 / \; HLDL_{R.med} \quad \text{(Equation 15)}$$

Normality – parametric and non-parametric statistics

Outliers

Following the approach of Aydin (2009) all measured values were used in the evaluation and outliers were not removed from the datasets. Nevertheless, outliers were identified in order to determine their number and gain potentially interesting information about individual stone properties (i.e. porosity). To detect outliers the MAD was used and (x_i) the boundary for extreme values (outliers) was specified using (moderately conservative) 2.5*MAD following the recommendation of Leys et al. (2013) and shown in Equation 16:

$$Median - 2.5 * MAD < x_i < Median + 2.5 * MAD \quad \text{(Equation 16)}$$

Kruskal-Wallis and Mann-Whitney U

The Kruskal-Wallis test was used as a robust alternative to one-way ANOVA to evaluate significant differences between the tested limestone types and the two probes (D and DL) (Hodges and Lehmann, 1963). This was followed by further specifying the differences between the individual stone types using the Mann-Whitney U test (two-tailed test with a significance level of p-value 0.05, unpaired) as an alternative to the t-test (Hodges and Lehmann, 1963). The data were visualised using boxplots and density plots in order to determine skewness and detect outliers.

Sample size determination

In addition to evaluating data using robust statistical measures, the second aim of the study was to determine an appropriate sample size for the Equotip that would sufficiently reflect the true stone surface hardness, but that was also practical for on-site application. For this, the 120 readings obtained for each stone type were taken to represent the true stone surface hardness ('population'). A range of smaller sample sizes (5, 10, 20, 45 and 60 readings) were then modelled by resampling the original dataset without replacement (for each sample size this process was repeated a 100 times to simulate variation) using bootstrap in RStudio. Finally, confidence intervals for the medians of the individual modelled sample size datasets were obtained through bootstrapping.

Our assumption was that the width of the confidence intervals would vary for the different sample sizes (i.e. a small sample size would result in wider confidence interval), taken to reflect the degree of variation of the median. These intervals were calculated with 95% confidence level using the bias corrected and accelerated (bca) bootstrap for confidence intervals in R (10,000 times), the most robust version for analysing non-normal data (Efron, 1987). The bootstrapped confidence intervals for the medians of the modelled sample size datasets were then compared to the original sample confidence intervals (using the original 120 readings) by calculating the differences of confidence interval widths in percentages. Based on the results an appropriate sample size was determined.

4.1.3 Results and discussion

This section firstly evaluates the performance of two Equotip probes (D and DL) on porous limestone under laboratory conditions on four porous limestone types. It is shown that for general data analysis for data obtained with the Equotip on porous limestone it is more beneficial to use robust measures and methods in order to account for natural variability of porous stone. Furthermore, an appropriate sample size for Equotip readings to be collected on porous limestone are determined to gain meaningful results.

Despite controlled laboratory conditions, fresh and smooth stone surfaces and a large sample size of 120 readings per stone type, the majority of the Equotip data sets show non-normal distribution (Shapiro-Wilk test, Table 4.5), caused by outliers and skewness (Figures 4.3, 4.4, 4.5, 4.6 and Tables 4.5 and 4.6).

Table 4.5 Surface hardness results for this study (120 readings per stone type

Stone type	$HLD_{S.med}$ ($HLD_{S.MAD}$)	Conf.int. $HLD_{S.med}$ low	Conf.int. $HLD_{S.med}$ high	$HLD_{R.med}$	$HDH_{D.robust}$	$HLD_{S.mean}$ ($HLD_{S.SD}$)	Shapiro-Wilk test (p-value)	Skewness	Kurtosis
P	469.00 (27)	454.00	479.00	681.00	318.32	462.99 (42.29)	0.012	-0.423	1.848
B	241.50 (38)	233.00	254.00	583.00	99.28	258.72 (68.90)	<0.000	0.272	2.909
C	318.50 (62)	302.00	344.50	623.00	161.63	333.18 (90.46)	0.029	1.506	0.321
G	215.50 (22.5)	210.01	222.00	622.00	78.68	217.89 (41.84)	0.000	1.873	18.798

[1] HLD=values obtained with D-probe. P=Portland Jordans Base Bed, B=Bath Hartham Park, C=Clipsham, G=Guiting; Conf.int=confidence interval, low=lower boundary, high=upper boundary. Subscript key: Med=median. MAD=median absolute deviation, S=SIM (single impact method) and R=RIM (repeated impact method). (See also Figure 5a)

Table 4.6 Surface hardness results for this study (120 readings each stone type)

Stone type	$HLDL_{S.med}$ ($HLDL_{S.MAD}$)	Conf.int. $HLDL_{S.med}$ low	Conf.int. $HLDL_{S.med}$ high	$HLDL_{R.med}$	$HDH_{DL.robust}$	$HLDL_{S.mean}$ ($HLDL_{S.SD}$)	Shapiro-Wilk test (p-value)	Skewness	Kurtosis
P	525.00 (20)	516.00	500.00	766.00	363.94	527.35 (36.78)	0.294	-0.423	0.300
B	297.00 (40)	284.50	315.50	637.00	138.98	315.15 (83.88)	<0.001	0.272	5.893
C	422.00 (76.5)	399.00	448.80	758.00	267.93	414.15 (112.93)	0.117	1.506	-0.588
G	266.00 (30.5)	252.24	276.50	631.00	106.40	270.53 (51.86)	0.019	1.873	0.801

[1] HLDL=values obtained with DL-probe. P=Portland Jordans Base Bed, B=Bath Hartham Park, C=Clipsham, G=Guiting; Conf.int=confidence interval, low=lower boundary, high=upper boundary. Subscript key: Med=median. MAD=median absolute deviation, S=SIM (single impact method) and R=RIM (repeated impact method). (See also Figure 5b).

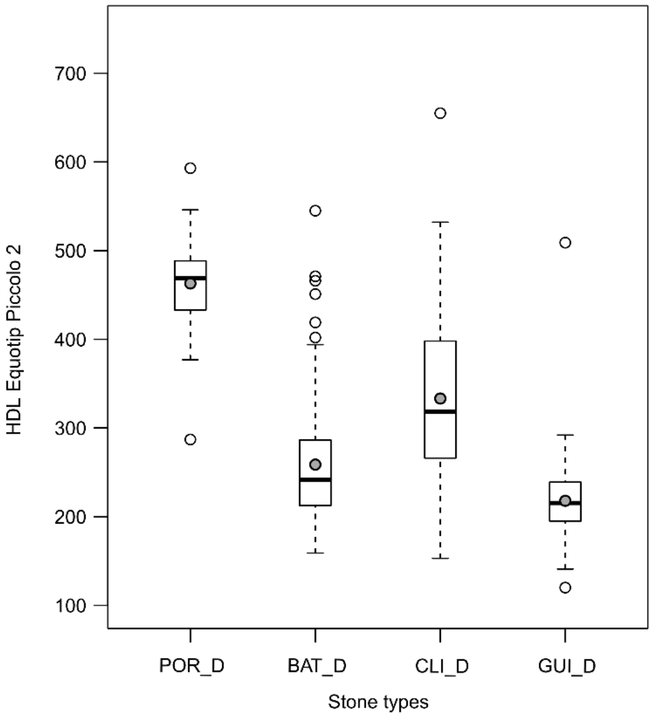


Figure 4.3 Boxplot of surface hardness values, with median (black line) and mean (grey dot) and outliers (white dots), four different stone types (Portland=POR, Bath=BAT, Clipsham=CLI, Guiting=GUI) with smooth surfaces (ground with sandpaper P.120), Equotip Piccolo 2 probe D, n=120.

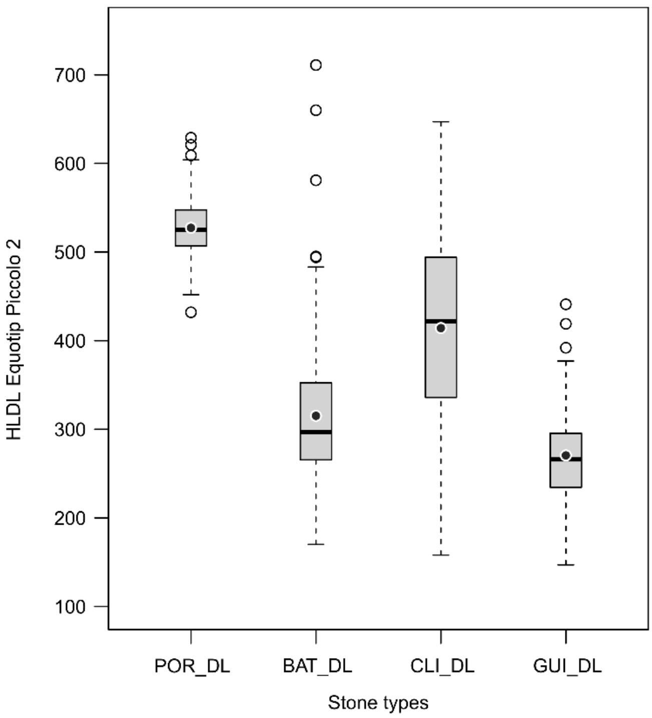


Figure 4.4 Boxplot of surface hardness values, with median (line) and mean (black dot) four different stone types (Portland=POR, Bath=BAT, Clipsham=CLI, Guiting=GUI) with smooth surfaces (ground with sandpaper P.120), Equotip Piccolo 2 probe DL, n=120.

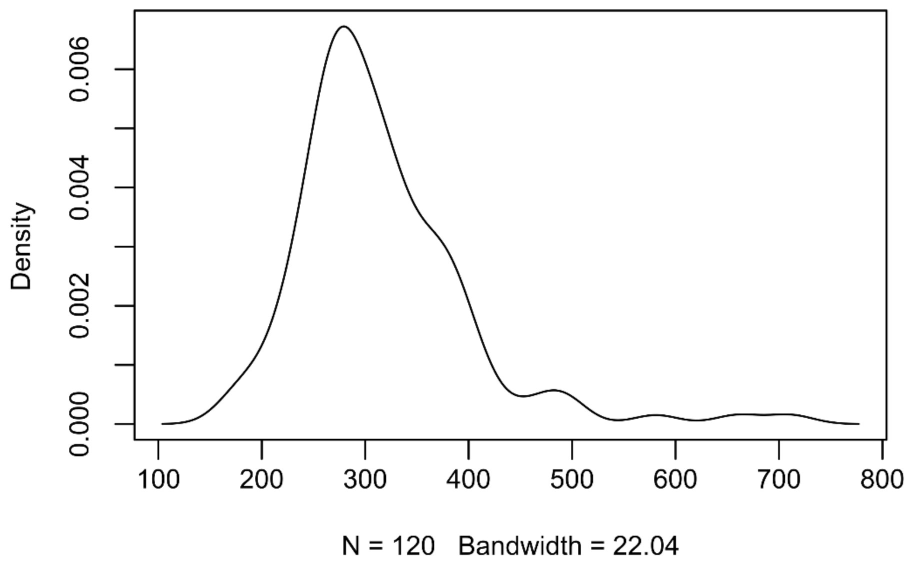


Figure 4.5 Example for skewed data in this study, density plot for distribution of surface hardness values (HLDL) for Bath limestone showing positive skew, Equotip Piccolo 2 with impact body DL, 120 readings.

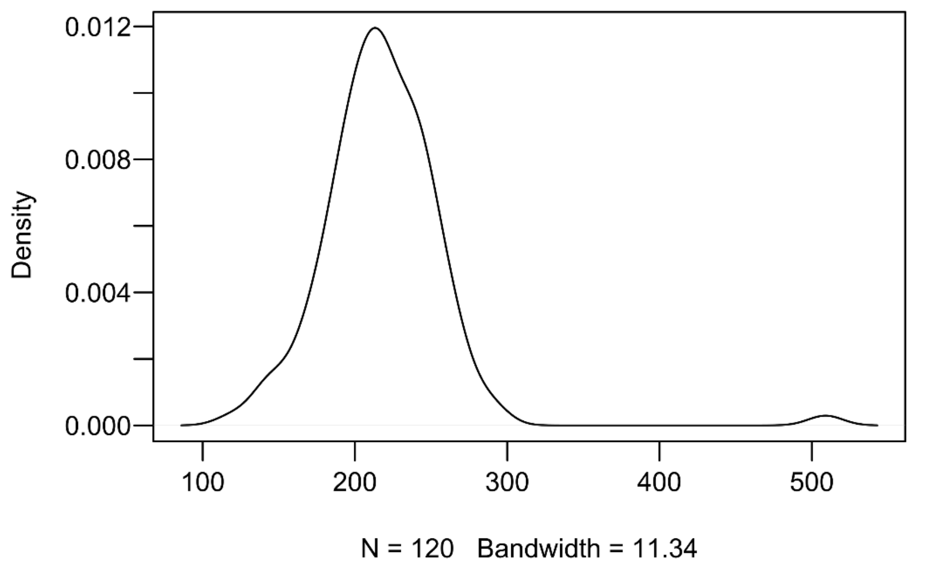


Figure 4.6 Example for skewed data in this study, density plot for distribution of hardness values (HLD) for Guiting limestone showing positive skew, Equotip Piccolo 2 with impact body D, 120 readings.

Probes

Figure 4.3 and 4.4 and Tables 4.5 and 4.6 show the data collected using the two probes (D and DL). As expected, they are not directly comparable. In every case, the DL probe produced higher hardness values, which was confirmed by Proceq©

as being usual (Personal communication 28/11/2013). It would have been useful to be able to convert HLD (hardness values obtained with the D probe) values into HLDL (hardness values obtained with the DL probe) and vice versa, but due to differing variances (probably caused by limestone characteristics) in the individual probe datasets this is not possible. The coefficient for the HLD and HLDL values ranged between 1.12 and 1.32. The DL probe produced a wider data spread than the D probe (except for Portland limestone, where D obtained a wider data spread) (Figures 4.3 and 4.4).

The Kruskal-Wallis test revealed significant differences in Equotip data for the stone types for both probes, D (df=3, chi-squared=305.904, p-value < 0.001) and DL (df=3, chi-squared=282.881, p-value < 0.000) probes. The following Mann-Whitney U tests showed significantly different hardness values (p < 0.001) for both probes on all four limestone types (Tables 4.7 and 4.8). This shows that Equotip can be used to distinguish the stone types used in this study using either probe.

Table 4.7 Results of the Mann-Whitney U test for the **D** probe and the single limestone types (Portland=POR, Bath=BAT, Clipsham=CLI, Guiting=GUI). All stones of this study can significantly be distinguished from each other

Groups D probe		**U**	**p-value**
POR	BAT	335	<0.001
POR	CLI	353.5	<0.001
POR	GUI	32	<0.001
BAT	CLI	3415.5	<0.001
BAT	GUI	4440.5	<0.001
CLI	GUI	1439	<0.001

Table 4.8 Results of the Mann-Whitney U test for the **DL** probe and the single limestone types (Portland=POR, Bath=BAT, Clipsham=CLI, Guiting=GUI). All stones of this study can significantly be distinguished from each other

Groups	DL probe	U	p-value
POR	BAT	411.5	<0.001
POR	CLI	2556	<0.001
POR	GUI	1	<0.001
BAT	CLI	3371	<0.001
BAT	GUI	4607	<0.001
CLI	GUI	2062.5	<0.001

Surface hardness data – Stone variance – Operator variance

Figure 4.7 shows no significant variance for HLD_S values (obtained on a metal test block) between the two experienced operators. The data range is well within the Equotip calibration requirements (HLD 765 ± 15). In contrast, the HLD_S values generated by the inexperienced operator show three outliers and thus, a noticeable shift of the mean. Nevertheless, it can also be seen that the median is not affected by the three outliers and within the calibration requirements. Therefore, using the median improves the reliability of Equotip data even if an inexperienced person is using the device. Since all further measurements in this study were conducted by the same operator (first author) and robust measures are used, operator variance is not considered to be an issue. As a consequence, the variance of surface hardness data observed in this study is attributed to the natural variability of the tested limestone as reported by Palmer (2008) and findings from Siedel and Siegesmund (2010) especially for limestone with low density and high porosity.

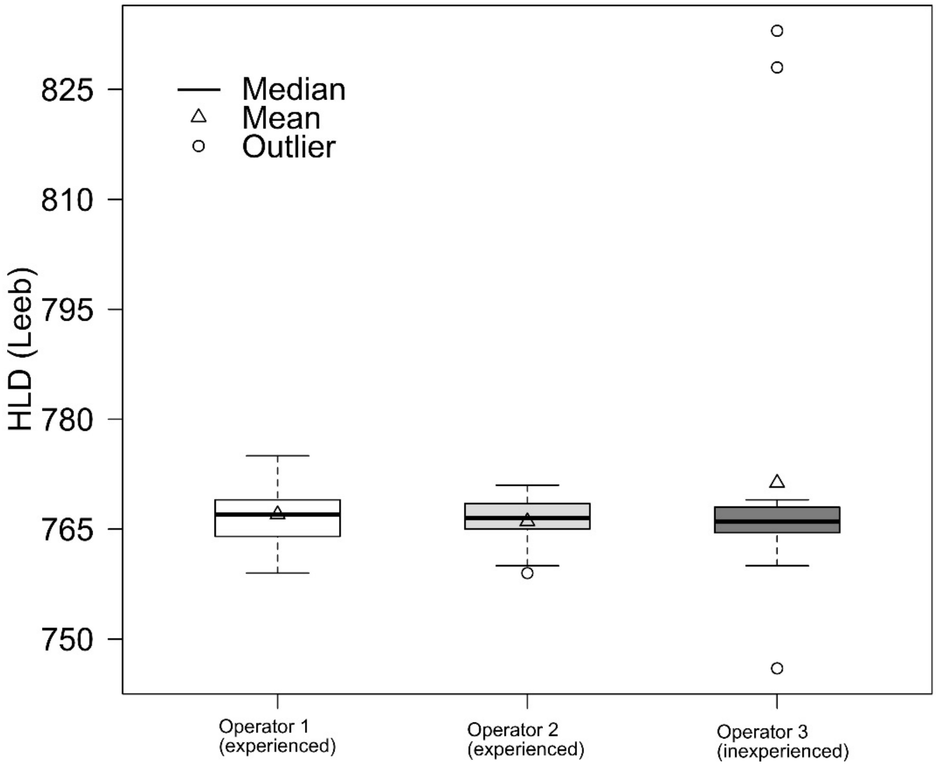


Figure 4.7 Boxplot showing 3 surface hardness datasets (20 single impact readings on a metal test block) generated by three different operators. Operator 1 and 2 had experience with the Equotip and operator 3 was using the device for the first time.

It was found that median values ($HLD_{S.med}$ and $HLDL_{S.med}$) showed lower variance compared to mean values ($HLD_{S.mean}$ and $HLDL_{S.mean}$) (Tables 4.5 and 4.6). The strength of the correlation for Pearson's R^2 and Spearman is categorised following Dancey and Reidy (2004), where the association with 1 = perfect, 0.7 - 0.9 = strong, 0.4 - 0.6 = moderate, 0.1 - 0.3 = weak, 0 = zero. Both the Pearson's R^2 and Spearman correlation coefficient show strong association of UCS median values with the median surface hardness values of both probes (Pearson's R^2: D probe R^2= 0.99 and DL probe R^2= 0.95; Spearman: D probe (r_s (2) = 1, p = 0.0833) and DL probe (r_s (2) = 1, p = 0.0833). The correlation of all surface hardness data (see Table 4.9) with the median open porosity shown in Table 4.3. All hardness data for both tested probes show a strong correlation, therefore reflect sufficiently on the respective porosity. Given the range of tested high porosities

(13.5 – 22.2%) the implications for on-site studies on weathered limestones are, that a) high porosity can be determined using Equotip and further b) porosity changes (increase or decrease) over time through weathering could be investigated. This has implications for the potential application of the Equotip in weathering rate studies.

Table 4.9 Pearson's R^2 and Spearman correlations for varying surface hardness data calculations and median open porosity of the limestones tested in this study

	R^2	Spearman
$HLD_{S.mean}$	0.9082	-0.8
$HLD_{S.med}$	0.8971	-0.8
$HDH_{D.robust}$	0.8626	-0.8
$HLDL_{S.mean}$	0.9452	-0.8
$HLDL_{S.med}$	0.9757	-0.8
$HDH_{DL.robust}$	0.9785	-0.8

Although the DL probe showed higher data spread, it correlates slightly better with open porosity values of limestone in this study compared to D probe shown by the R^2 values in Table 4.9. The best fit is gained with the $HDH_{DL.robust}$. Furthermore, the DL probe might be more advantageous in the field, because it offers a more controlled way of sampling in recessed, rough or curved areas (typical for weathered stone and architectural geometry of built heritage). Also, it offers protection from dust for the Equotip device itself due to the long slim front section, which prevents the impact body from transporting particles into the body of the device.

Outliers

Almost every dataset contained more than one outlier (Table 4.10). In the case of porous limestone outliers may occur due to the heterogeneity (e.g. porosity, shells) of the stone as discussed earlier. Thus, outliers are likely to be part of the natural deviation in the population and should not be removed. For this study it is particularly noticeable that most of the outliers are higher hardness values, as might be found when the Equotip impact body strikes a hard fossil for example. Figures 4.3 and 4.4 show the effect of outliers and skewness on the mean values, which are different from the medians in the majority of cases. The difference between mean and median values is most notable for Clipsham and Bath limestone and might be due to their particular pore size distribution and inherent material variability.

Table 4.10 Results of outlier detection using equation 16 (section 5.5.2). Notice the majority of outliers is beyond the upper bound (i.e. extreme high hardness values) indicating the presence of fossils and other harder elements

Stone	Probe	Total outliers	Beyond lower bound	Beyond upper bound
Portland	D	2	1	1
Portland	DL	6	1	5
Bath	D	8	0	8
Bath	DL	7	0	7
Clipsham	D	1	0	1
Clipsham	DL	0	0	0
Guiting	D	2	1	1
Guiting	DL	4	1	3

For weathered rock and stone surfaces variability in Equotip data is likely to be even higher and thus Equotip data are rarely likely to be normal. In cases of non-normal data, statistical estimates based on common statistical descriptors may be affected (Tukey, 1977) so that parametric tests are largely inappropriate (Fowler et al., 1998). Consequently, in order to account for inherent variability in surface hardness measurements caused by natural stone properties (on-site), and to avoid the need for data transformation, the robust statistical methods used in this paper are preferable to classic statistical measures and methods previously used for Equotip data evaluation.

Appropriate sample size

How many readings should be taken when applying Equotip devices to stone and rock surfaces? This study aimed to determine a sample size big enough to portray reliably the median surface hardness of the four tested stone types, but small enough to also be practical for on-site applications. Based on the original 120 readings ('population') collected per stone type several smaller sample sizes were modelled (5, 10, 20, 45 and 60) and for each the width of the confidence interval for the median was calculated to reflect the degree of variation of the median. A small degree of variation would result in narrower confidence interval widths and thus show high accuracy in prediction of the median. For all stone types a bigger sample size resulted in narrower confidence interval widths (Figures 4.8 and 4.9, Tables 4.5, 4.6), confirming Niedzielski et al.'s (2009) statement that accuracy increases with increased sample size.

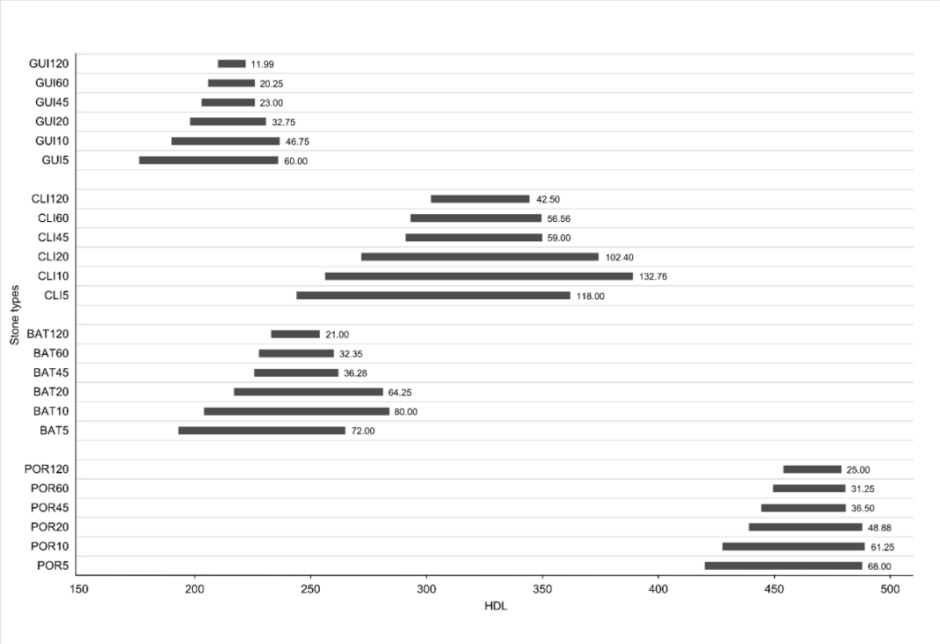


Figure 4.8 Predicted confidence intervals for medians of Equotip Piccolo 2 D probe data for different modelled sample sizes (numbers on the y-axis) on four different limestone in this study (Portland=POR, Bath=BAT, Clipsham=CLI, Guiting=GUI). Confidence intervals are obtained applying bias-corrected accelerated bootstrap to datasets of 120 readings. Modelled samples sizes are 5, 10, 20, 45, and 60 readings, resampled from the original dataset (120). Bars show confidence interval width (numeric value indicated) for median to occur within at 95% confidence level (See also Table 4.5).

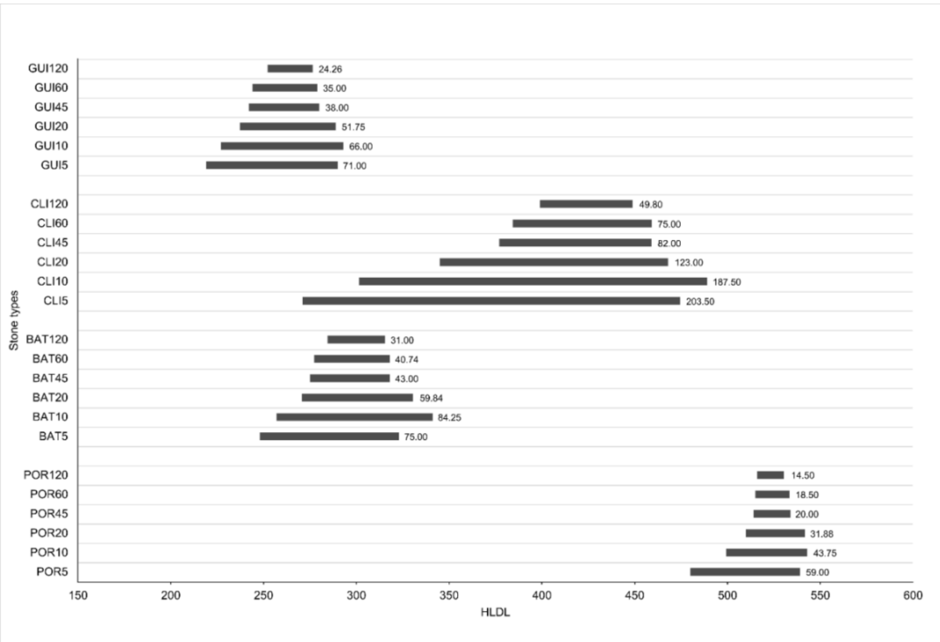


Figure 4.9 Predicted confidence intervals for medians of Equotip Piccolo 2 DL probe data for different modelled sample sizes (numbers on the y-axis) on four different limestone in this study (Portland=POR, Bath=BAT, Clipsham=CLI, Guiting=GUI). Confidence intervals are obtained applying bias-corrected accelerated bootstrap to datasets of 120 readings. Modelled samples sizes are 5, 10, 20, 45, and 60 readings, resampled from the original dataset (120). Bars show confidence interval width (numeric value indicated) for median to occur within at 95% confidence level (See also Table 4.6).

In this study the majority of confidence intervals for the median are wider for the DL probe compared to those of the D probe (exceptions are confidence intervals for a sample size of 5). The confidence intervals obtained for different stone types are noticeably distinct from each other, with wider intervals for stone with complex porosities like Clipsham and narrower intervals for stones with higher compressive strength like Portland. This indicates that an appropriate sample size is heavily dependent on the stone type and consequently its state of preservation, where changes in those properties indicate ongoing weathering processes. Therefore, either a bigger sample size is necessary for porous stone with complex pore size distributions, or a wider confidence interval needs to be tolerated for the median to fall into. Alternatively, a lower confidence level could be accepted (i.e. 90%) with narrower confidence intervals.

For this study the appropriate sample size was determined using a confidence level of 95% accepting that different stone types would therefore have wider confidence intervals. Comparing the predicted confidence intervals for the different modelled sample sizes of this study (Figures 4.8 and 4.9) with sample sizes used by previous studies (e.g. 10 and 20 as reported by Aoki and Matsukura (2007) and Yilmaz (2013) respectively), it becomes clear that here the predicted confidence intervals within which a median is expected to appear are rather wide. For example, given 20 readings for Clipsham and Bath the confidence intervals are not substantially different and therefore the median surface hardness could not be sufficiently distinguished. In contrast, the confidence intervals of the original datasets (120 readings) are very narrow, reflect the median stone surface hardness well, and all four stone types are clearly distinguishable.

However, taking 120 readings is often not practical in the field because it is (a) time consuming and (b) would require a larger measurement area, which may limit potential subsequent investigations. Consequently, to find a good compromise between accuracy and practicality, it was aimed to define smaller appropriate sample sizes and accepting potentially wider confidence intervals, whilst ensuring that the confidence intervals of one single stone type tested in this study should not overlap with one of another stone type.

It is therefore necessary to define a general sample size that would be appropriate for all stone types tested, and that would be transferable to on-site application on stone with unknown history. As stated earlier the modelled sample size of 20 for the Clipsham limestone and Bath limestone did not show clearly separated confidence intervals, which makes it impossible to distinguish the stone types using surface hardness. Therefore, an appropriate sample size was estimated by evaluating further the width difference (as a percentage) of modelled 20, 45 and 60 sample size datasets relative to the original datasets of 120 readings.

Table 4.11 and 4.12 show that the confidence interval differences between 120 and 60 and 45 readings are smaller than between 120 and 20 readings. Although 60 readings better reflect the original data set (i.e. a smaller % difference), it would be sufficient to take 45 readings in order to obtain a narrow enough confidence interval with all confidence intervals for the four different stone types clearly distinguishable. Furthermore, Tables 4. 5 and 4.6 show the lower and higher boundaries for the calculated confidence intervals for the surface median hardness. Thus, for this study every median surface hardness obtained will fall into the respective confidence interval and can clearly be attributed to a stone type. Using the range of a confidence interval rather than a single value like the median to represent stone and rock surface hardness is more applicable (i.e. versatile) for in situ applications where natural stone and rock variance is expected.

Table 4.11 **D** probe with Equotip Piccolo 2, percentage (%) differences of confidence interval widths for sampling sizes of 20, 45 and 60 readings (resampled) in comparison to a sample size of 120 (original 'population')

D probe	**% difference in confidence interval width**			
Readings	POR [%]	BAT [%]	CLI [%]	GUI [%]
120 to 60	25.00	54.06	33.09	68.86
120 to 45	46.00	72.75	38.82	91.79
120 to 20	95.54	205.95	140.94	173.09

Table 4.12 **DL** probe with Equotip Piccolo 2, percentage (%) differences of confidence interval widths for sampling sizes of 20, 45 and 60 readings (resampled) in comparison to a sample size of 120 (original 'population')

DL probe	**% difference in confidence interval width**			
Sample sizes compared	POR [%]	BAT [%]	CLI [%]	GUI [%]
60 and 120	27.59	31.40	50.62	44.25
45 and 120	37.93	38.71	64.67	56.62
20 and 120	119.85	93.02	147.01	113.29

4.1.4 Conclusions and recommendations

On the basis of the results in this study we propose a number of considerations when using Equotip testing:

1) Scope of application of the Equotip: This study shows that the Equotip is suitable for soft and porous rock and stone. It is however, beneficial to

calibrate the Equotip on fresh stone before using it on-site on weathered stone surfaces in order to establish weathering rates. Nevertheless, the Equotip application works as well as relative measure e.g. for the comparison of surfaces exposed to different aspects and/or degree of orientation and height or for quality assessment before and after stone consolidation treatment.

2) D and DL probe: Although the DL probe showed higher data spread, it correlates slightly better with open porosity values of limestone in this study. Further advantages are more controlled sampling in recessed areas, rough or curved areas (typical for weathered stone and architectural geometry of built heritage). The long slim front section of the probe, which prevents the impact body from transporting particles into the body of the device, offers further protection from dust for the Equotip device itself.
3) Non-normal data: In this study, data obtained from four different limestone under controlled conditions yielded non-normal data in the majority of cases, as a result of inherent variability in material properties such as porosity. This paper argues that Equotip data from weathered stone and rock surfaces are rarely normal and thus parametric tests are largely inappropriate and would either require data transformation to gain meaningful results or the application of robust statistical measures and methods.
4) Robust (non-parametric) statistical measures and methods, and outliers: Outliers and skewness were the main cause for the unsymmetrical distributed data in this study. The paper proposes to include outliers in the data analysis as their occurrence is linked to natural stone characteristics – in the case of the limestones tested they indicate the presence of fossils and other harder elements. However, including outliers in data analysis necessitates the new approaches to statistical analyses addressed in this study. The presented alternative, robust statistical approach requires no data transformation (e.g. removing

outliers and more) and is more reliable for non-normally distributed data as well as being adequate for normal data. We recommend to apply robust statistical methods unaffected by non-normal data (e.g. median and MAD as alternative measures of central tendency and variance as well as Kruskal-Wallis and Mann-Whitney test as alternatives to ANOVA and t-test.).

5) HDH: The combination of two measuring procedures (SIM and RIM) based on median values accounts for potential effect of pores/weathering especially when used with DL probe values and thus complements SIM and RIM.
6) Sufficient sample size: A big enough sample size needs to be collected and is highly dependent on the respective porosity of the tested stone. Thus, the more porous (heterogeneous) and weathered the stone the higher the sample size should be. Nevertheless, for practical reasons for on-site applications on stone with unknown history the aim was to determine a general sample size that would be a) appropriate for all stone types tested in this study, and thus include a variety of high porosities, while b) differentiate the respective stone surface median hardness and c) allow to distinguish the tested stone types. Therefore, for the tested stones in this study we propose a sample size of 45 readings (for a confidence level of 95%). It is worth mentioning that calculating sample sizes using a 95% confidence level is a conservative approach. In view of the expected variances for in situ measurements and unknown weathering-stress histories of heritage stone, it might be justified to reduce the confidence level to 90%. This would still provide reliable data output when robust measures are used, but allow for a smaller sample size to be collected. This approach can certainly be transferred to stone and rock with similar porosities and hardness.

While the study was conducted in the laboratory and took variation of natural stone into account, it used fresh, smooth stone samples and thus, research on-site is desirable to link back to results obtained in the laboratory. This study has shown that the Equotip provides valuable measures of surface hardness of porous stone which can be related to other measures such as unconfined compressive strength as found by Hack et al. (1993) and Verwaal and Mulder (1993), but also demonstrates that data evaluation can be improved by using robust measures, applying robust statistical methods and increasing sample size. The proposed methodology requires no data modification (e.g. removing outliers), is more accurate for non-normally distributed data and adequate for normal data, and thus provides a timesaving general approach to data evaluation including on-site measurements. This methodology allows for consistent comparability between different on-site research projects across the fields of rock weathering and stone deterioration research.

Acknowledgments

K. Wilhelm is in receipt of an EPSRC studentship (funded via grants EP/P504287/1, EP/P505216/1 and EP/P505666/1) with additional support from Proceq.

Dr Dan Lunn is thanked for invaluable help with the program R and introduction to bootstrapping techniques. The authors want to thank Dr Helen Reeves and Marcus Dobbs from the British Geological Survey (BGS), who provided technical assistance with compressive strength testing. Dr Martin Coombes and Jerome Mayaud provided very helpful comments on an earlier draft of this paper.

4.2. PAPER 2_ THE INFLUENCE OF SALT ON HANDHELD ELECTRICAL MOISTURE METERS: CAN THEY BE USED TO DETECT SALT PROBLEMS IN POROUS STONE?

K. Wilhelm1*, H. Viles2, and Ò. Burke2

1 Oxford Rock Breakdown Laboratory (OxRBL), School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK; Tel: +44 (0)7919551141; katrin.wilhelm@ouce.ox.ac.uk

2 Oxford Rock Breakdown Laboratory (OxRBL), School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK; Tel: +44 (0)1865 285072; heather.viles@ouce.ox.ac.uk

3 Senior Statistician, Nuffield Department of Population Health, University of Oxford, Richard Doll Building, Old Road Campus, OX3 7LF, UK; Tel: +44(0)1865 289371; orlaith.burke@ndph.ox.ac.uk

Abstract

Salt contamination in heritage stone affects handheld moisture meter measurements on-site. This poses a problem when the readings indicate erroneously higher levels of moisture than actually present. For decision making with regards to moisture prevention treatment it is therefore crucial to distinguish between actual dampness and the hygroscopic action of salts. This study investigated the effect on moisture meter readings of both increased conductivity of pore water and the increased water retention caused by the presence of sodium chloride (NaCl) in artificially contaminated Portland limestone samples. The influence of NaCl contamination on selected handheld moisture meters was quantified. As a result, the paper proposes that under certain circumstances moisture meters could be used to diagnose, reliably, both moisture and salt problems in heritage stone.

Keywords: non-destructive testing; limestone; stone heritage; sodium chloride

4.2.1 Introduction

The presence of water in porous building materials like stone is a key factor in the decay of built heritage (Poschlod, 1990; Meinhardt-Degen, 2005; Smith and Viles 2006). Thus, assessing moisture regimes of immovable heritage on-site is important, yet not without difficulty. Although handheld electrical moisture meters are convenient tools for the task and are applied frequently by professional surveyors, geomorphologists and heritage conservation scientists, they are often regarded with suspicion (Burkinshaw, 2002; Viles, 2013; Cutler et al., 2013). A lack of knowledge on the exact interactions between resistivity- and capacitance-type electrical moisture meters and building stone condition is evident. Electrical moisture meters do not directly measure moisture content, but detect a change in either resistivity or electrical field that is applied to a porous material (e.g. stone). Most devices automatically convert the results to an estimated or calibrated moisture content value (Eklund et al., 2013). Thus, care needs to be taken when interpreting results as the pre-set conversion might not represent the actual material tested (Burkinshaw, 2002). It is further necessary to understand the measuring principle of moisture meters, and factors affecting the data output, in order to interpret the results reliably (Arendt and Seele, 2000).

The measurement of both resistivity- and capacitance-type electrical moisture meters can be influenced by the mineralogy, homogeneity and density of the measured material, temperature and moisture distribution within the material, the presence of contaminants (e.g. salt), the application pressure used, as well as the type of measuring voltage or frequency (Arendt and Seele, 2000; Martinez and Byrnes, 2001). Furthermore, operator variance and surrounding factors like the presence of metal (e.g. reinforcement in concrete structures) also need to be considered (Eklund et al., 2013). Of these factors, one of the most important is the presence of salts – which are nearly ubiquitous in historic buildings and structures. However, it remains unclear how these devices are affected by the presence of salt and how best to interpret the data they provide. This paper begins to address this gap in knowledge.

4.2.2 Electrical moisture meters and the impact of salt

Resistivity-type measurement

Electrical resistivity is an intrinsic material property defining the degree to which that material impedes the flow of an applied electric current. Its reciprocal is conductivity. The term resistance is used more specifically to describe the resistivity of a material of given dimensions. The (electrical) resistivity of a porous building stone is defined by its mineral composition, porosity and the electrolyte contained in pores (pore water) (Arendt and Seele, 2000). The resistivity decreases with increasing internal moisture content (Flint et al., 1999; Loke, 1999) and thus can be assessed with resistivity-type electrical moisture meters, like the Protimeter Surveymaster™ in resistivity mode. However, such meters are mostly calibrated for wood (Burkinshaw, 2002) and results cannot simply be transferred to other materials as studies on brick have shown (Howell, 1995; Burkinshaw and Parrett, 2003; Trotman et al., 2004). The Protimeter SurveymasterTM manual suggests that for investigating porous building material the 'wood moisture equivalent' (WME) value should be used. The term WME is misleading, however, since the (equilibrium) moistures of building materials can be very different from each other and hence need to be calibrated accordingly

(Burkinshaw, 2002). Therefore, results from stone and other building materials apart from wood using resistivity-type meters are only relative and do not represent absolute moisture values (Burkinshaw and Parrett, 2003).

Salts in historic buildings and structures are only active when water or moisture is present (Charola, 2000). One way for built structures to take up water is by absorbing moisture from the air. Thus, depending on the level of moisture, or relative humidity in the air, the salts in the structure are more or less active. Accordingly relative humidity changes are thought to have an important influence on stone deterioration. Accordingly, Erkal et al. (2013) find stone deterioration processes related to frequent relative humidity (RH%) changes in historical stone masonry (whilst excluding rising damp in their study as the investigated areas at 2.75m are above capillary-rise level). Further, Franzen and Mirwald (2009) and Diaz Gonçalves and Delgado Rodrigues (2006) simulate sorption behaviour of salt contaminated stone samples under laboratory conditions and find varying internal moisture, expressed as hygroscopic moisture content (HMC), depending on varying relative humidity levels (RH%).

Depending on its nature, stone itself can absorb moisture to a certain degree (Charola, 2000). In addition salts like sodium chloride ($NaCl$) increase hygroscopic behaviour and moisture retention thus, increase the total amount of water content in stone (Weber, 1988; Camuffo and Sturaro, 2001). Furthermore, when RH% is above the respective equilibrium relative humidity (ERH) salts may deliquesce and form saturated solution (Charola, 2000). Dissolved salts increase conductivity of (pore) water and, therefore, reduce resistivity in stone (Loke, 1999). For resistivity based moisture meters that automatically convert readings to some (arbitrary) unit of moisture this results in higher readings. This is a crucial point, as measurements based on resistivity may indicate erroneously higher levels of moisture when salt is present. Accordingly, Burkinshaw (2002) and Eklund et al. (2013) observe that, in general, inflated readings are obtained using resistivity-type meters when salts are present in the investigated structure. Indeed, some authors have gone further and suggested that salt content associated with moisture in masonry might be investigated via conductivity measurements (Kraska, 1998; McCann and Forde, 2001). However, because of the great variability of possible salt mixtures in buildings and the large range of resistivity-type meters available, careful calibration would be required (Kraska, 1998; Schuh et al., 2011).

Capacitance-based measurement

For capacitance-based moisture meters the dielectric constant or permittivity (i.e. electrical charge storage properties) of materials (k or εr) is central. Water has a high permittivity (80 εr), whereas porous building materials like dry limestone show very low permittivity (7 εr) (Martinez and Byrnes, 2001). Such differences in permittivity can thus be utilised to investigate moisture content in porous building materials, whereby an increase in moisture results in increased capacitance measurement values. Varying statements can be found in the literature regarding the influence of salt on electrical capacitance meters, which appears to depend on the individual characteristics of each particular device (e.g. frequency of the meter) and / or material differences and varying salt concentrations (Trotman et al. (2004), Blakemore et al. (2005), Pinchin (2008), Bayer et al. (2010) and Schuh et al. (2011).

4.2.3 Aims and objectives

Based on the above review, the reliability and utility of electrical moisture meters could be improved by quantifying the effects of selected stone properties and conditions (i.e. porosity and salt contamination) on the readings obtained using both resistivity- and capacitance-type moisture meters. This study tested whether the increased conductivity of pore water caused by the presence of salt can be detected with moisture meters by measuring samples with similar moisture content, but different NaCl contamination levels (before sorption equilibrium). It is further tested whether increased water retention in limestone (induced by changes of relative humidity) results in significantly different meter readings for salt contaminated vs clean samples (at sorption equilibrium). Furthermore, applying new data evaluation methods (robust statistics) could enhance the data output. The aim of this paper is to shed some light on the influence of NaCl contamination on selected handheld moisture meters, and to evaluate the potential of these effects to be used to diagnose salt and moisture problems in stone heritage.

4.2.4 Material and methods

General approach

In this study the effect on electrical moisture meters of varying levels of sodium chloride (NaCl) contamination has been quantified for fresh Portland limestone. We investigated three different handheld moisture meters operating in resistivity or capacitance mode: one resistivity-type meter, the Resipod (Proceq©), which to our knowledge has not been tested before on limestone; one capacitance-type meter, the CEM; and the GE Protimeter Surveymaster, which operates in both resistivity and capacitance mode (Figure 4.10, Table 4.13). Each of these devices is detailed in the following sections.

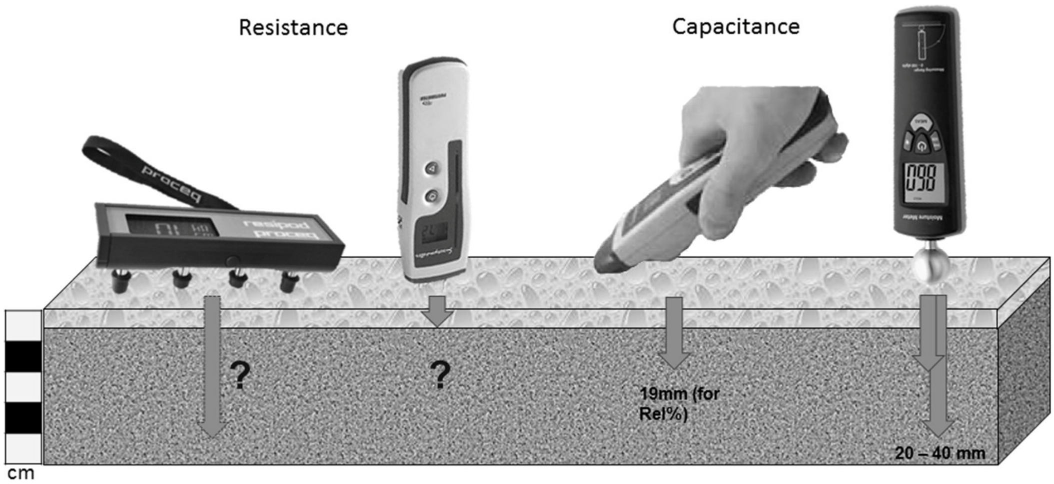


Figure 4.10 Overview of electrical moisture meters used in this study with indicative measuring depths, from left to right: Resipod, Protimeter Surveymaster TM in resistivity mode and in capacitance mode and CEM (modified after GE Sensing 2009; Proceq©).

Table 4.13 Specifications of the electric moisture meters employed in this study according to the manufacturers' datasheets

Moisture meter	Price £	Meas principle	Meas unit / range	Meas depth
Resipod	2236	Resistance	kΩcm / 0 - ~1000	n/a
CEM	50 - 100	Capacitance	Digits from 0 -100	20 -40 mm
Protimeter	244 - 400	Resistance and Capacitance	WME / 6 -90 Rel% / 60 -999	19 mm for Rel%

Moisture meters

GE Protimeter SurveymasterTM (resistivity and capacitance mode)

The GE Protimeter SurveymasterTM offers both resistivity (WME, units 6-90) and capacitance mode (Rel%, units 60-999) (Figure 4.10, Table 4.13). Resistivity is measured with two, 10 mm metal pins set at a spacing of 14 mm which are pressed firmly against the stone. When used in capacitance mode, the pins are covered with a lid and the back of the device applied flat on the surface. The Protimeter SurveymasterTM reports measurements in three ways – with a numerical value, LED lights and sound. It is mainly designed to test damp in wood and is calibrated accordingly. Thus, all values generated by the Protimeter SurveymasterTM only represent relative measurements when used on other materials (Burkinshaw, 2004; Eklund et al., 2013).

Resipod from Proceq©

The second resistivity-type device used was the Resipod from Proceq©, a 4-point Wenner probe (Figure 4.10, Table 4.13). The Resipod has primarily been designed to test concrete (Proceq© 2014) and to our knowledge its performance on limestone has not previously been investigated. The contact pins of this device are spring-loaded to allow for thorough contact when pressed against the surface (which has advantages on uneven surfaces). We used pins that were fixed at 50 mm apart, but different spacings are available. Values recorded are given in kilo-ohm cm (kΩcm) and are stored internally. The device measures a range from 0 to 1000 kΩcm and can be checked for function with a test strip provided by the manufacturer. With the Proceq© software, the data can be converted into a .csv file and simple statistics are provided (e.g. mean and standard deviation). In contrast to the Protimeter SurveymasterTM, the values are not converted to units of 'relative moisture' (i.e. higher readings for higher moisture content) and thus with an increase of material moisture lower values are obtained with Resipod (i.e. actual lower resistivity and higher conductivity).

CEM

The CEM DT-128 (Shenzhen Everbest Machinery Industry Co., Ltd, Shenzhen, China) is based on the capacitance principle of measurement (Figure 4.10, Table 4.13). Its handling is the easiest of all tested devices and it can be calibrated by taking a mid-air reading (Eklund et al., 2013). It displays values from 0 to 100. Like the Protimeter SurveymasterTM, it does not have any internal data storage capability.

Stone specimens

For this study fresh samples of Portland limestone (Jordans Base Bed) were investigated. Porosity (range = 13.12 – 13.82%, mean = 13.5%), unconfined compressive strength (range = 43.20 –75.73 MPa, mean = 55.98 MPa) and water absorption (range = 6.49 – 6.87 Mass %, mean = 6.71 Mass %) under atmospheric pressure were determined following the British standards (BS-EN 1936:2006, BS EN 1926:2006 and BS EN 13755:2008, respectively with n=10). The average open porosity of the Portland Base Bed (13.5%) is in accordance with the findings of Dubelaar et al. (2003). Furthermore, Dubelaar et al. (2003) determined (with mercury porosimetry) a high proportion of micropores (~ 75%) for this stone type. Nine specimens were cut (300 x 80 x 50 mm) each with smooth (cut) surfaces to reduce any potential confounding effect of surface roughness on moisture measurements. The size of the blocks was selected based on the length of the measuring area of the Resipod (200 mm) and the measuring depth of the

CEM (up to 40 mm). The 'fresh' samples in this study were not desalinated in contrast to the study of Eklund (2013), which found a 13% difference for Protimeter (in resistivity mode) and CEM readings before and after desalination of samples. Their study however investigated saturated stone samples and thus, a significantly higher amount of material moisture (water absorption for Portland Whit Bed ~ 6.3 Mass % (BRE, 1997) vs. 0.3 Mass % in sorption equilibrium for this study). Assuming a linear correlation between moisture content and salt effect the error introduced by the naturally contained salts for this study would be < 1%. Therefore, this study considered the % difference for desalinated samples being marginal.

Salt contamination

In the field the interactions of the stone-salt-HMC (hygroscopic moisture content) system with the environment are very complex and still not fully understood and therefore, difficult to simulate under laboratory conditions (e.g. Franzen and Mirwald, 2009). Thus, laboratory experiments usually limit complexity by reducing the number of parameters involved. Accordingly, this study limited complexity of a) hygroscopic behaviour of salt mixtures and b) hygroscopic behaviour of salts with more than one hydration phase (e.g. sodium sulphate (Na_2SO_4) (Rodriguez-Navarro et al., 2000; Yu and Oguchi, 2009) by contaminating the stone samples with a single salt only, sodium chloride (NaCl). NaCl has been used in stone weathering experiments before as it is considered to be deteriorative especially for heritage located near the sea (e.g. Colston et al., 2001; Andriani and Walsh, 2007; Gomez-Heras and Fort, 2007). Furthermore, it is a simple salt without hydration phases (Steiger et al., 2010) and therefore easier to control in an experiment.

The limestone samples were artificially contaminated with two realistic levels of NaCl representative of those found in heritage structures (Arendt and Steele, 2000; Table 4.14). A third set of uncontaminated stones was used as control. With this approach the effects of HMC alone and HMC-salt combination on the performance of the three electrical moisture meters were investigated. The limestone samples were divided into three groups with three replicate blocks each. The control group (S0) was not NaCl contaminated. The remaining two groups were contaminated to a medium (S1) and high (S2) level by saturating them in two concentrations of NaCl solutions. Samples were soaked for 3 weeks to ensure even penetration of the salt.

Table 4.14 Classification of deterioration potential for different levels of salt (anions) known to deteriorate built heritage (Arendt and Seele 2000). Level 2 and 3 (bold) are relevant for this study

Deterioration level	Sulphate in wt%	Chloride in wt%	Nitrate in wt%	Concentration in mmol/kg
0 – unloaded	< 0.024	< 0.009	< 0.016	< 2.5
1 – low	**< 0.077**	**< 0.028**	**< 0.05**	**< 8.0**
2 – medium	**< 0.24**	**< 0.09**	**< 0.16**	**< 25.0**
3 – high	< 0.77	< 0.28	< 0.50	< 80.0
4 – extreme	> 0.77	> 0.28	> 0.50	> 80.0

Given the importance of the effect of changing relative humidity on sorption behaviour of salt contaminated stone, this study induced phase changes of NaCl by manipulating ambient relative humidity (%RH), which causes partial crystallisation and dissolution of NaCl in the structure (Charola, 2000). Thus, this study simulated on-site moisture measurement situations in historic structures with salt contamination at medium and higher levels and where salt deterioration processes are driven by changing relative humidities (e.g. Colston et al., 2001; Linnow et al., 2007). In such situations it is crucial to distinguish between actual moisture content in the built structure and hygroscopic interaction of salts. Dampness, which is caused by hygroscopic salt might be mistaken for rising damp. For the latter the resulting implication for a subsequent conservation intervention is substantially different from desalination procedures or climate control to manage salt contamination problems (Charola, 2000).

Salt phase-changes

NaCl has only one solid phase (no crystal water) and it deliquesces at a relative humidity of about 75.4 % at 25°C (Steiger et al., 2010) (Figure 4.11). To investigate the effect on moisture meters of NaCl below and above its deliquescence, its phase changes from liquid to crystalline were induced by exposing contaminated stone samples to a relative humidity either far below or sufficiently above the deliquescence relative humidity of NaCl at 20°C, with 38% RH (referred to as dry in this study) and 95% RH (referred to as damp in this study) respectively. This was achieved using an environmental cabinet (Binder KBF115).

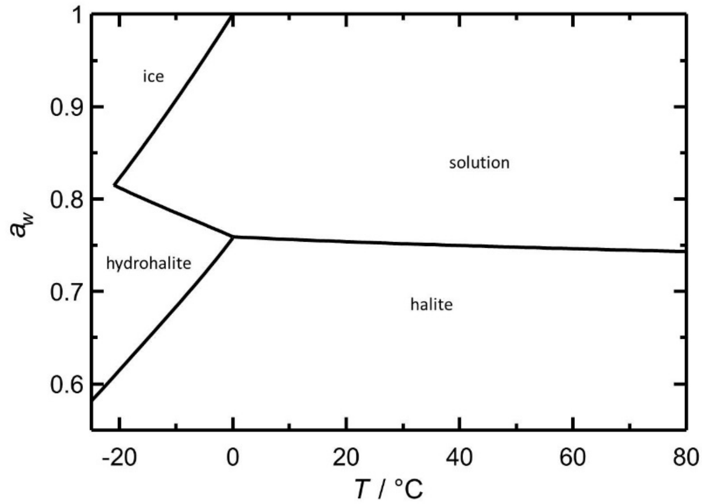


Figure 4.11 Deliquescence behaviour of sodium chloride. Water activity aw is plotted versus temperature (Steiger, 2004).

As mentioned earlier salt has two effects on apparent moisture readings of electrical moisture meters, through hygroscopic behaviour it increases water retention in the stone structure, and it also increases conductivity of the pore water. Both effects were examined in this study. To investigate whether the increased conductivity of pore water related to the presence of NaCl can be detected with moisture meters the stone samples in the 3 groups would need to have the same amount of moisture content, which when in sorption equilibrium is not the case, because of the increased water retention effect with increased NaCl contamination level. To isolate this effect, the samples were measured before they had reached sorption equilibrium in 95% RH, but had similar moisture (s.m.) content (compared to the oven dry weight). After that samples were left to equilibrate and weighed every 24 hours until the difference in mass between successive weighings was no greater than 0.1 g for 38% RH and 95% RH. Thus, the combined effect of increased conductivity and hygroscopicity was investigated.

Moisture Meter Measurements

Blocks were weighed once before moisture measurements were taken and their actual moisture content was calculated as a percentage of the oven-dry weight. Measurements using each of the moisture meters were taken in a consistent way, following a modified version of the 'optimized experimental protocol' of Eklund et al. (2013). All measurements were conducted by the same operator, and the devices were stored in the room where measurements took place. The environmental conditions of the measuring area (temperature and RH%) were monitored using two climate data logger systems (Gemini Tinytag and i-Button®). The moisture meters were always applied to the stone in the same order and calibrated before each measurement. All readings were taken perpendicular to the bedding of the stone, and the meters allowed to stabilize for a few seconds before each reading was taken. Application pressure is thought to affect data obtained using both resistivity and capacitance type meters (Forsén and Tarvainen, 2000; Trotman et al., 2004; Pinchin, 2008). However, pilot testing showed that application pressures ranging from 5 to 40 N resulted in only minor variations in values obtained. For this study meters were always applied to the stone with a pressure of between 20 – 40 N for consistency (samples placed on a balance). Five readings were taken with each meter per block, obtaining fifteen values in total for each treatment group (three blocks per group).

Statistical evaluation

The datasets (15 readings per group) were tested for normality (Shapiro-Wilk test) and for about half of the datasets non-normal distribution was found (Table

4.15). Non normality might be attributed to the nature of the stone, which as a natural product can display inhomogeneity and anisotropy (Mosch and Siegesmund, 2007; Palmer, 2008). As a consequence, moisture and salt might be unevenly distributed and cause non-normally distributed moisture meter readings. To avoid data transformation and account for those effects robust statistics (non-parametric) were used as they are less affected by deviations from normality (Filzmoser and Todorov, 2013).

Accordingly, median values and median absolute deviation (MAD, a robust measure for variance) were used to summarize the data. The Mann-Whitney U test was applied to determine any significant differences between datasets of the individual meters for samples exposed to the three conditions of relative humidity (38% and 95% before and 95% at sorption equilibrium).

Datasets of the same group (S0, S1, S2) under different RH% climates and evaluate significant differences in meter readings between the different NaCl contamination levels were evaluated with the two-tailed Mann-Whitney U test (significance level of p-value 0.05).

4.2.5 Results and discussion

Salt influences on moisture meter readings

All moisture meters tested in this study were affected by NaCl content in the stone samples and resulted in higher readings as compared to the non-contaminated samples. However, the magnitude of this effect was not consistent between the different relative humidity conditions or between the different moisture meters and their measurement modes. Resistivity-type meters seem to be

Table 4.15 Descriptive statistics for this study. Equ stands for equilibrium and sm for similar moisture (before equilibrium).

Device	RH%	Salt	Shapiro .Wilk	Mean	Median	MAD	SD	Sample Variance	Kurtosis	Skewness
CEM	38	S0	0.01	33.07	33.00	0.00	0.70	0.50	-0.67	-0.09
CEM	95% sm	S0	0.00	34.47	33.00	1.48	2.23	4.98	-1.39	0.65
CEM	95% equ	S0	0.29	32.93	33.00	1.48	1.03	1.07	0.01	0.15
CEM	38	S1	0.05	32.33	32.00	1.48	0.90	0.81	-0.68	-0.10
CEM	95% sm	S1	0.01	37.6	39.5	1.48	3.17	10.04	-1.96	-0.48
CEM	95% equ	S1	0.14	37.20	37.00	4.45	2.68	7.17	-1.55	-0.07
CEM	38	S2	0.08	34.53	35.00	1.48	0.92	0.84	-0.48	-0.11
CEM	95% sm	S2	0.76	47.27	47.00	1.48	1.58	2.50	0.12	-0.38
CEM	95% equ	S2	0.02	55.33	55.00	0.00	0.90	0.81	-0.01	0.58
Rel%	38	S0	0.01	169.73	172.00	2.97	3.94	15.50	-0.48	-0.94
Rel%	95% sm	S0	0.06	183.27	184.00	2.97	4.88	23.78	0.13	-0.88
Rel%	95% equ	S0	0.29	189.00	189.00	2.97	6.43	41.29	-0.18	-0.40
Rel%	38	S1	0.57	172.27	172.00	2.97	4.73	22.35	0.12	0.58
Rel%	95% sm	S1	0.07	223.80	213.50	24.46	28.99	840.18	-1.18	0.67
Rel%	95% equ	S1	0.02	220.80	215.00	22.24	23.68	560.89	0.44	1.15
Rel%	38	S2	0.09	180.80	183.00	2.97	4.06	16.46	-1.31	-0.31
Rel%	95% sm	S2	0.00	533.60	579.00	40.03	110.98	12315.83	3.05	-1.96
Rel%	95% equ	S2	0.12	812.20	816.00	19.27	13.51	182.60	-0.97	0.20
WME	38	S0	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
WME	95% sm	S0	0.19	7.93	8.10	0.59	0.56	0.32	0.24	-0.90
WME	95% equ	S0	0.34	8.29	8.40	0.59	0.88	0.77	1.02	0.71
WME	38	S1	0.56	8.05	8.00	1.48	1.43	2.03	-0.75	0.40
WME	95% sm	S1	0.79	10.55	10.60	1.33	1.42	2.02	-0.89	-0.05
WME	95% equ	S1	0.39	9.67	9.80	2.08	1.83	3.34	-1.18	0.07
WME	38	S2	0.40	11.83	12.00	0.74	0.75	0.56	-1.31	-0.12
WME	95% sm	S2	0.97	11.47	11.80	1.19	1.55	2.39	-0.20	-0.25
WME	95% equ	S2	0.05	14.72	15.80	2.22	2.64	6.96	-1.27	-0.59

more affected by NaCl, whereas capacitance meters appear more sensitive to variations in moisture. As the moisture content of the stone samples was modified only indirectly, by varying relative humidity, changes in moisture content (i.e. pore water) occurred within a very narrow range; the maximum moisture uptake was 0.3% (mass), measured for the high (S2) contaminated samples at sorption equilibrium (95% RH). These damp conditions simulate stone in the built environment in equilibrium with the environmental conditions. However, this might not be the case under most real-world scenarios, when environmental conditions are subjected to quick changes. In this instance it is recommended to undertake several measuring campaigns to investigate how the built stone relates to the specific climatic environment.

Resistivity mode – Protimeter Surveymaster™

The Protimeter SurveymasterTM (in resistivity mode) was affected by NaCl in the stone samples regardless of whether the sample was dry (38% RH) or damp (95% RH) both before and at sorption equilibrium. In general readings were higher with higher salt contamination (Table 4.16 and Figure 4.12), which confirms earlier findings (Burkinshaw, 2002; Eklund et al., 2013). Table 4.16 shows that all samples in sorption equilibrium in 95% RH experienced significantly different Protimeter SurveymasterTM readings from dry (38%) readings (p-values < 0.02). For non-contaminated (S0) samples in dry (38% RH) condition the Protimeter SurveymasterTM gave no readings as the stone was too dry and below the detection limit, and so readings on non-contaminated samples (S0) were only obtained for damp samples (95% RH), although without significant difference before and at sorption equilibrium (p-value 0.333, Mann-Whitney U=47.5).

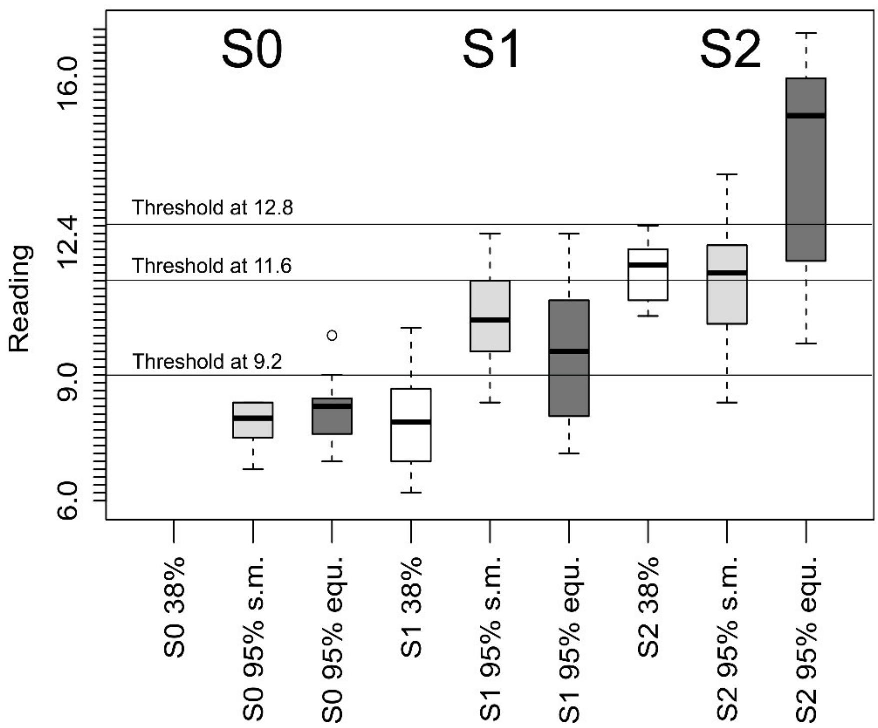


Figure 4.12 Boxplot GE Protimeter Surveymaster TM resistivity mode readings on fresh Portland limestone with three levels of NaCl contamination (no added salt (S0), medium salt (S1), high salt (S2)) under three environmental conditions (38% RH and 95% RH before sorption equilibrium (s.m.) and at sorption equilibrium (equ.), at 20°C). Whisker length=Interquartile Range (IQR) + IQR*1.5. Three thresholds marking ranges are indicated distinguishing particular environmental conditions and levels of salt contamination.

Table 4.16 Mann-Whitney U test for GE Protimeter® (resistance mode) readings on fresh Portland limestone. The table shows significant (grey boxes) versus non-significant differences in readings depending on the environmental condition (38% RH and 95% RH before and after equilibrium).

WME	RH	38%	38%	38%	95% s.m.	95% s.m.	95% s.m.	95% equ.	95% equ.
RH		S0	S1	S2	S0	S1	S2	S0	S1
38%	S0	-							
38%	S1	n.a.	-						
38%	S2	n.a.	0.000	-					
95% s.m.	S0	n.a.	0.9156	0.0000	-				
95% s.m.	S1	n.a.	0.0059	0.0262	0.000	-			
95% s.m.	S2	n.a.	0.001	0.4144	0.000	0.174	-		
95% equ.	S0	n.a.	0.000	0.000	0.333	0.0015	0.000	-	
95% equ.	S1	n.a.	0.000	0.000	0.024	0.183	0.0143	0.000	-
95% equ.	S2	n.a.	0.000	0.000	0.000	0.001	0.002	0.000	0.000

The group of medium S(1) contamination samples showed no significant difference between the condition before and at sorption equilibrium at 95% RH (p-value 0.183, Mann-Whitney U=54), indicating that the effect of increased conductivity by either S(1) or additional moisture is too low to be detected by the device or beyond the measuring depth of Protimeter SurveymasterTM. High contaminated samples S(2) show no significant difference between dry (38% RH) and before sorption equilibrium in 95% RH (p-values 0.4144, Mann-Whitney U=99).

Medium (S1) contaminated samples in dry (38% RH) condition show no significant difference to non-contaminated (S0) before sorption equilibrium in 95% RH (p < 0.9156, Mann-Whitney U=53). Thus, hypothetically assuming NaCl content was unknown to the operator, it could only be concluded that the sample was either damp with S(0) contamination or dry with medium NaCl contamination. However, for all other combinations of different contamination levels in dry (38% RH) and before sorption equilibrium in 95% RH condition significant differences are found (p-values < 0.02). Thus, the isolated effect of increased conductivity due to salt, as measured using samples before sorption equilibrium, is revealed and increased conductivity is in fact detectable. Transferred to on-site measurement this indicates that, even if the building stone is not in sorption equilibrium due to fast changing environmental conditions (reported as dynamic sorption behaviour (Franzen and Mirwald, 2004)), salt contamination might cause higher meter readings and thus, mislead interpretation of results (i.e. indication of higher moisture content as than actually found). This demonstrates the need for care to be taken when carrying out field moisture surveys especially for situations with fast changing environmental conditions.

No significant difference was found between samples with medium (S1) and high (S2) contamination before sorption equilibrium ($p < 0.174$, Mann-Whitney U=50), suggesting that differences in conductivity between these two treatment groups could not be detected by the Protimeter SurveymasterTM. Therefore, in fresh Portland limestone treated with NaCl the Protimeter may be used to distinguish non-contaminated samples from contaminated ones, but not for NaCl level quantification when the sample is not in sorption equilibrium.

Two clear thresholds and a range can be distinguished in Figure 4.12, which may have valuable on-site applications. The first threshold is marked by the maximum value of the S0 95% RH in equilibrium dataset at the reading of 9.2. All values above indicate NaCl-contaminated samples in damp conditions (95% RH regardless of whether the values recorded before or at sorption equilibrium). If a reading below 9.2 is obtained this could indicate that either a non-salt contaminated sample is damp (95% RH before or at sorption equilibrium), or that a dry (38% RH) sample is contaminated with a medium level of NaCl. The limitation here is that these two conditions cannot be distinguished from each other with this device alone, yet this still provides an indication of what may be present in the stone (moisture, salt or a combination of both) and thus that further investigation is required. A second threshold is marked by the maximum value of the dataset S2 38% at the reading of 12.8 (Figure 4.12), with values above indicating a high (S2) NaCl contamination under damp conditions (95% RH at sorption equilibrium). Thirdly a range below 12.8 and above 11.6 indicates a combination of NaCl contamination and damp condition, with a high chance of the sample being highly (S2) NaCl contaminated. Table 4.17 summarises the three indicative thresholds based on our data for Portland limestone, which may be used to determine the nature of the environmental conditions samples have been exposed to and their level of salt contamination.

Table 4.17 Classification and indicative threshold values for Portland limestone using a Protimeter (WME, resistance mode)

WME	Indication	Action
No reading	Dry sample + no salt	None
< 9.2	Either damp sample + no salt, or dry + salt	Further investigation or check using capacitance mode
11.6 - 12.8	With salt, either dry or damp	Test for salt
> 12.8	High salt + damp	Test for salt

The data is particularly interesting in view of the crystallization behaviour of sodium chloride. NaCl is expected to be found in crystal form under conditions of 20°C and 38% RH. Thus, there should be no effect on conductivity and Protimeter readings. However, readings were obtained for sample groups with medium (S1) and high (S2) NaCl contamination in dry condition (38% RH). There are two possible explanations for this effect: (1) firstly, as mentioned in section 2.3 the stone samples may contain salt by natural default. The effect was considered to be marginal, however might be a possible explanation for the obtained meter readings under dry conditions (38% RH). Assuming that the natural salts are a mixture (common in practice (Doehne, 2010)) this would imply that a lower deliquescence point could result as salt mixtures are known to lower deliquescence points (Wexler and Seinfeld, 1991; Bionda, 2004; Price, 2007).

Secondly, high microporosity can lead to increased water retention in stone (Palmer, 2008) and this can enable moisture detection even for ‘dry’ stone. For example, at ambient conditions of 38% RH mono- and multi-molecular layers of water are present in stone pores and capillary condensation starts to take place (Franzen and Mirwald, 2004). This effect is enhanced when microporosity is high. The microporosity of the Portland limestone used in this study is relatively high (~75%, Dubelaar et al., 2003), which likely explains the higher moisture readings within the stone samples relative to the ambient RH of the air. This in turn may lead to salt in pores being (partially) dissolved, conductive, and hence contributing to a positive meter reading.

Resistivity meter – Resipod from Proceq©

The Resipod only produced readings for NaCl contaminated blocks in sorption equilibrium at 95% RH. There was a significant difference in measurement values between samples with medium (S1) and high (S2) levels of contamination ($p < 0.001$, Mann-Whitney $U=0$) (Figure 4.13). All values for medium (S1) contaminated blocks range above a threshold of 619 kΩcm, which is the minimum value of the dataset for S1 95% RH in equilibrium. In comparison, the Protimeter SurveymasterTM (in resistivity mode) is far more sensitive to NaCl contamination for samples under dry 38% RH and damp (95% RH) conditions before sorption equilibrium. This implies that for NaCl-contaminated samples the combination of both devices could be used to determine whether samples are in sorption equilibrium at 95% RH, and contain either medium (S1) or high (S2) levels of NaCl.

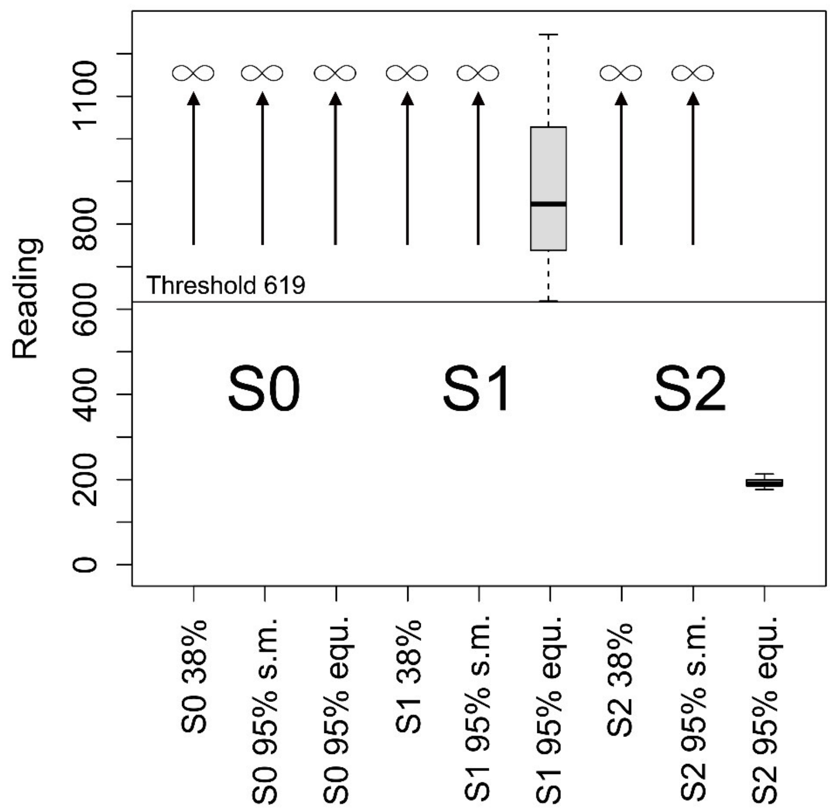


Figure 4.13 Boxplot Resipod from Proceq® readings for fresh Portland limestone with three levels of NaCl contamination (no added salt (S0), medium salt (S1), high salt (S2)) under three environmental conditions (38% RH and 95% RH before sorption equilibrium (s.m.) and at sorption equilibrium (equ.), at 20°C). Whisker length=Interquartile Range (IQR) + IQR*1.5. A threshold is indicated distinguishing particular environmental conditions and levels of salt contamination. Note: the values show actual resistivity (Ω) thus lower values indicate higher moisture/NaCl content.

Capacitance mode – Protimeter Surveymaster TM

The Protimeter Surveymaster TM in capacitance mode was less affected by the presence of NaCl, but more so by moisture; with increasing NaCl levels and dampness, each treatment group displays consecutively significantly higher values (Figure 4.14). In contrast, when used in resistivity mode the highest values for high NaCl contamination (S2) are obtained, regardless of whether the specimens were damp (95% RH) or dry (38% RH).

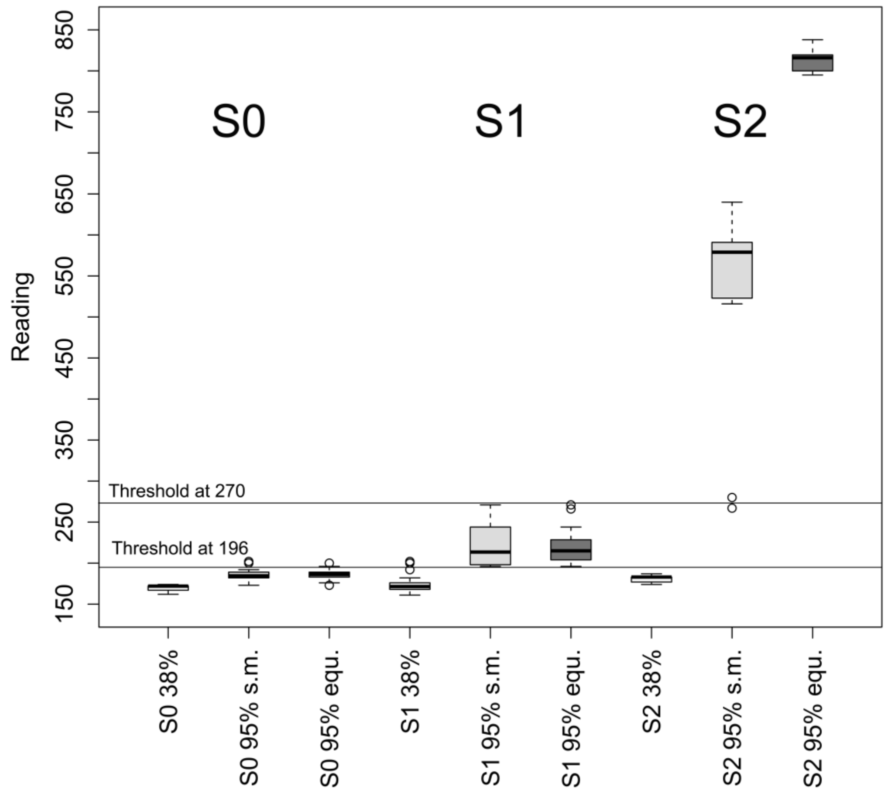


Figure 4.14 Boxplot GE Protimeter Surveymaster TM capacitance mode readings on Portland limestone with three levels of NaCl contamination (no added salt (S0), medium salt (S1), high salt (S2)) under three environmental conditions (38% RH and 95% RH before sorption equilibrium (s.m.) and at sorption equilibrium (equ.), at 20°C). Whisker length=Interquartile Range (IQR) + IQR*1.5. Two thresholds are indicated distinguishing particular environmental conditions and levels of salt contamination.

In only two instances values were not significant from each other. Samples with S(0) contamination are not significantly different from medium S(1) in dry (38%) condition (p-value 0.341, Mann-Whitney U=84, Table 4.18). Further, for medium S(1) contaminated samples no significant difference was found before and at sorption equilibrium damp conditions (95 %RH) (p-value 0.933, Mann-Whitney U=75) indicating that the moisture changes were too small to be detected. In contrast, every other level of NaCl contamination and relative humidity yielded significantly different values (p-value < 0.05). A threshold for damp samples (95% RH before and at sorption equilibrium) was identified, with values above 270 indicating high NaCl level (S2) (Figure 31). A second threshold is apparent for medium (S1) NaCl contamination above a value of 196 for damp samples (95% RH before and at sorption equilibrium). The indicative thresholds are summarized in Table 4.19.

Table 4.18 Mann-Whitney U test for GE Protimeter® (capacitance mode) readings on fresh Portland limestone. The table shows significant (grey boxes) versus non-significant differences in readings depending on the environmental condition (38% RH and 95% RH before and after equilibrium).

Rel	RH	38%	38%	38%	95% s.m.	95% s.m.	95% s.m.	95% equ.	95% equ.
RH		S0	S1	S2	S0	S1	S2	S0	S1
38%	S0	-							
38%	S1	0.3407	-						
38%	S2	0.0000	0.0001	-					
95% s.m.	S0	0.001	0.000	0.001	-				
95% s.m.	S1	0.000	0.006	0.000	0.000	-			
95% s.m.	S2	0.000	0.000	0.001	0.000	0.000	-		
95% equ.	S0	0.0007	0.0000	0.0006	0.024	0.000	0.000	-	
95% equ.	S1	0.0000	0.0007	0.0000	0.000	0.933	0.000	0.0000	-
95% equ.	S2	0.0000	0.0000	0.0007	0.000	0.000	0.001	0.0000	0.0000

Table 4.19 Classification and indicative threshold values for Portland limestone using a Protimeter (Rel%, capacitance mode)

Rel%	Indication	Action
161 - 186	Dry sample + potentially salt	Test for salt
186 - 196	No salt + damp	None
> 196	Medium salt + damp	Test for salt
> 270	High salt + damp	Test for salt

Data obtained using the capacitance mode of the Protimeter SurveymasterTM has been described as being more difficult to interpret than resistivity-mode data, and therefore less useful to surveyors, at least with regards to wood testing (Burkinshaw, 2002). However, here we found that the capacitance mode has potential as a complementary method. For example, for our stone type, when a Protimeter reading in resistivity mode is obtained between about 11.6

and 12.8, indicating probable high (S2) NaCl contamination, additional readings using the capacitance mode can help clarify the result. In this instance, when a capacitance-based reading was above 270 high (S2) NaCl contamination was indicated, and if below, medium (S1) contamination was indicated.

Capacitance meter – CEM

The CEM returned readings for every sample regardless of the humidity conditions they were exposed to (Figure 32). However, neither non-contaminated (S0) samples in dry (38% RH) condition could be significantly distinguished from either before or at sorption equilibrium at 95% RH (p-value 0.094 / 0.675, Mann-Whitney U=79 / 102 respectively, Table 4.20). The same is true for the group S(1) between before and at sorption equilibrium (p-value=0.586, Mann-Whitney U=69.5). Further, dry medium (S1) samples cannot significantly be distinguished from S(0) in sorption equilibrium at 95% RH (p-value 0.127, Mann-Whitney U=77). This indicates that the hygroscopicity of the NaCl is similar to material moisture in sorption equilibrium without additional NaCl contamination. Dry (38% RH) high S(2) contaminated samples show no significant difference from S(0) before sorption equilibrium (p-value 0.352, Mann-Whitney U=90). The results are not conclusive as one would rather expect the hygroscopic effect of increased water retention through high S(2) contamination to be similar to S(0) at sorption equilibrium. This might either be caused by variance in measuring accuracy of the device or variability of the stone sample. This further shows that this device is not sensitive to minor differences in stone moisture. It is, however, affected by NaCl content. Apart from the non-significant results all other combinations of contamination levels and conditions showed significant differences.

Table 4.20 Mann-Whitney U test for CEM readings on fresh Portland limestone. The table shows where significant (grey boxes) versus non-significant differences in readings occur depending on the environmental condition (38% RH and 95% RH before and after equilibrium).

CEM	RH	38%	38%	38%	95% s.m.	95% s.m.	95% s.m.	95% equ.	95% equ.
RH		S0	S1	S2	S0	S1	S2	S0	S1
38%	S0	-							
38%	S1	0.028	-						
38%	S2	0.000	0.000	-					
95% s.m.	S0	0.094	0.003	0.352	-				
95% s.m.	S1	0.000	0.002	0.006	0.006	-			
95% s.m.	S2	0.000	0.000	0.001	0.000	0.000	-		
95% equ.	S0	0.675	0.127	0.000	0.081	0.000	0.000	-	
95% equ.	S1	0.000	0.002	0.006	0.006	0.586	0.000	0.000	-
95% equ.	S2	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000

An indicative threshold of 41 is shown in Figure 4.15, readings above which were all obtained for samples with high (S2) NaCl contamination before and at sorption equilibrium (95% RH). Another range is indicated between the readings 36 and 41, in which the measured values are certainly related to dampness, with a higher probability of medium (S1) NaCl contamination. Clearly all values obtained above the threshold of 50 are related to high (S2) NaCl contamination and dampness (95% RH in equilibrium). Below a value of 36 it cannot be determined whether the sample is dry (38% RH) and NaCl contaminated or damp (95% RH) and not contaminated. In such situations using the Protimeter in capacitance mode could help to clarify the results (see previous sections). The indicative detection thresholds are presented in Table 4.21.

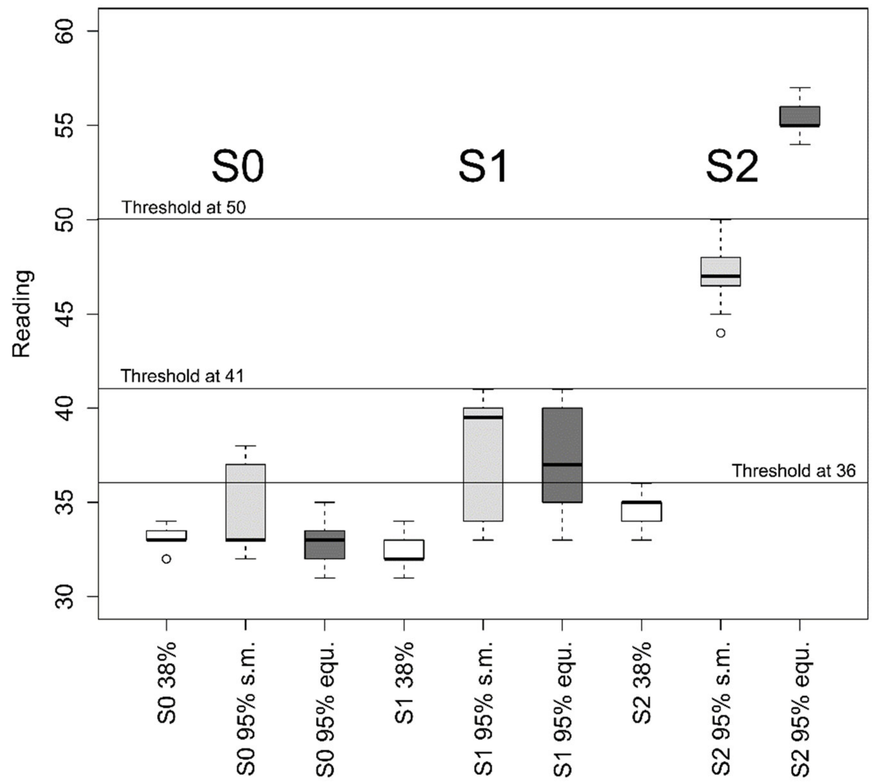


Figure 4.15 Boxplot CEM readings on Portland limestone with three levels of NaCl contamination (no added salt (S0), medium salt (S1), high salt (S2)) under three environmental conditions (38% RH and 95% RH before sorption equilibrium (s.m.) and at sorption equilibrium (equ.), at 20°C). Whisker length=Interquartile Range (IQR) + IQR*1.5. Three thresholds are indicated marking ranges for distinguishing particular environmental conditions and levels of salt contamination.

Table 4.21 Classification and indicative threshold values for Portland limestone using a CEM moisture meter

CEM	Indication	Action
< 36	Either damp + salt, or dry + salt	Further investigation required
36 - 41	Damp + high probability of medium (S1) salt	Test for salt
> 41	High salt + damp	Test for salt
> 50	High salt + damp equilibrium	Test for salt

4.2.6 Conclusions

This study has assessed the effect of varying sodium chloride (NaCl) contamination of fresh Portland limestone on measurements obtained using a range of non-invasive handheld moisture meters (both capacitance- and resistivity-type). The results show that all moisture meters tested are affected to varying degrees by NaCl content. These effects are more pronounced under 95% RH conditions and after the samples reached sorption equilibrium. We isolated the effect of NaCl increasing conductivity (in pore water), alongside the combined influence of increased conductivity and water retention (at sorption equilibrium) by taking measurements both before (same moisture content, but different levels of NaCl contamination) and after samples reached sorption equilibrium,

We found that when used in resistivity mode the Protimeter SurveymasterTM is able to detect high NaCl content (S2) on samples in both dry (38% RH) and damp (95% RH) conditions. In contrast, with the Resipod from Proceq© (resistivity-type) readings could only be obtained for damp samples in sorption equilibrium, which had medium (S1) and high (S2) levels of NaCl contamination. Thus, the Protimeter SurveymasterTM appears to be far more sensitive to NaCl contamination at 38% RH, and before sorption equilibrium in damp conditions (95% RH). However, when the Resipod did return a reading this was a clear indicator that the sample was in sorption equilibrium and contaminated.

The capacitance mode of the Protimeter SurveymasterTM was found useful to complement measurements based on resistivity, in order to help clarify results and discriminate moisture from salt effects. For the CEM (capacitance-type) we found no significant differences in moisture measurements in the absence of any NaCl for samples exposed to 38% RH and 95% RH. This device is therefore less sensitive to slight variations in stone moisture. However, the CEM is affected by

NaCl content and in this respect, like the Protimeter Surveymaster™ in capacitance mode, may also be used to complement resistivity-type meter readings, but the latter was better for distinguishing between salt and moisture conditions.

These results are promising with regards to detecting salt on site, but further research is needed to verify our findings for heritage materials under on-site conditions. Other salts and salt mixtures as well as different stone types and stone with different weathering status need to be examined, as they may have varying effects on data obtained using these types of meters. In general, care needs to be taken when moisture meter data are interpreted; the factors affecting the data output for these devices need to be carefully assessed. Nevertheless, understanding of these factors, such as the relative influences of moisture (including relative humidity) and salts both in isolation and in combination, coupled with appropriate knowledge of the material being investigated (e.g. porosity) can enhance the value of hand-held moisture meter applications.

Acknowledgements

K. Wilhelm thanks Proceq© and EPSRC studentship (funded via studentship grants EP/P504287/1, EP/P505216/1 and EP/P505666/1). Dr Martin Coombes for helpful comments on an earlier draft of this paper, and Dr Mona Edwards for technical support.

5. OBJECTIVE 2: ON-SITE APPLICATION OF NON-DESTRUCTIVE TECHNIQUES TO LIMESTONE HERITAGE *IN SITU*

5.1. SURFACE HARDNESS AS A PROXY FOR WEATHERING BEHAVIOUR OF LIMESTONE HERITAGE: A CASE STUDY ON DATED HEADSTONES ON THE ISLE OF PORTLAND, UK

K. Wilhelm[1]*, H. Viles[2], O. Burke[3], J. Mayaud[4]

[1]Oxford Rock Breakdown Laboratory (OxRBL), School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK; Tel: +44 (0)7919551141; *katrin.wilhelm@ouce.ox.ac.uk

[2]Oxford Rock Breakdown Laboratory (OxRBL), School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK; Tel: +44 (0)1865 285072; heather.viles@ouce.ox.ac.uk

[3]Nuffield Department of Population Health, University of Oxford, Old Road Campus, OX3 7LF, UK; orlaith.burke@ndph.ox.ac.uk

[4]Oxford Rock Breakdown Laboratory (OxRBL), School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK; jerome.mayaud@ouce.ox.ac.uk

Abstract

This study estimates stone weathering rates using in situ surface hardness testing. Surface hardness changes are precursors to erosion and may be utilized to describe stone weathering behaviour. The method proposed here complements previous approaches to determining stone weathering rates by surface loss/change. A time series covering 1 to 248 years of exposure is investigated using a sample of 12 headstones in two nearby cemeteries. Using an Equotip D surface hardness tester, rates of change in surface hardness for top and bottom sections of the headstones were determined and the data evaluated using robust, non-parametric statistical methods.

When considering all headstones as one time series, non-linear behaviour is observed with rates of decline in surface hardness slowing over time. However, breakpoint analysis shows a breakpoint at c. 100 years, with higher rates of surface hardness decline (as measured by QC_{50} – the regression coefficient for 0.50 quantile regression) up to that point, and lower rates thereafter. Up to c. 100 years surface hardness declines more rapidly in the top vs bottom sections. Possible explanations for the differing rates in surface hardness changes are: a) inherent natural stone variability and/or different weathering stress history; b) the use of two different Portland limestone varieties; c) synergistic effects of microclimates and lichen cover. In order to gain a deeper insight into stone weathering behaviour, future studies could combine surface hardness measurements with surface change methods such as micro-erosion meter and lead plug index over short and long-term time series on architectural heritage under real world conditions.

Keywords: Portland limestone; cemetery headstones; time-series; non-parametric statistics; weathering rates

5.1.1 Introduction

Establishing meaningful estimates of limestone weathering behaviour across a variety of spatial and temporal scales is crucial for stone weathering research (Viles, 2001). Improving our understanding of stone weathering behaviour under real world conditions allows us to predict potential responses to environmental and anthropogenic impacts, and therefore to estimate the future life span of cultural stone (Warke et al., 2003; Smith and Prikryl, 2007). It further informs decision-making on stone heritage conservation strategies (Svahn, 2006; Auras 2011b; Inkpen et al., 2012b). Moreover, stone responses to phenomena such as climate change and air pollution can be linked back to standard durability laboratory tests, in order to improve their accuracy (e.g. Ross and Butlin, 1989; Viles, 2002b; Ingham, 2005; Smith et al., 2011; Viles and Cutler, 2012).

The majority of studies that have investigated limestone weathering rates in response to air pollution have relied on deliberately exposed samples (e.g. Lipfert, 1989; Trudgill et al., 1991; Butlin et al., 1992; O'Brien, 1995; Bonazza et al., 2009; Brimblecombe and Grossi, 2009). This approach poses two difficulties: (i) the restrictions on its spatial scale; and (ii) an overwhelming focus on the erosion (loss) of stone. Indeed, the use of small samples (e.g. 50 x 8 x8 mm for samples from the National Materials Exposure Program (NMEP) (Butlin et al. , 1992)) complicates the upscaling of results to meaningfully larger scales, such as built heritage (e.g. Bell, 1993; Trudgill and Viles, 1998). Furthermore, erosion is understood to be the final step in a series of decay mechanisms resulting in the total loss of a material. However, there is strong evidence that several weathering mechanisms precede this loss stage, such as surface hardening (redeposition of solutional products) or softening (induced by both climate and biological activity), which lead to stone surface property alterations including increased porosity and the formation of superficial layers (Pope et al., 2002; Hoke and Turcotte, 2004; Smith and Viles 2006; Inkpen et al., 2012b; McIlroy de la Rosa et al., 2014). Erosion or surface recession can therefore be considered a relatively coarse measure of weathering, and an improved understanding of limestone breakdown relies on quantifying the entire weathering trajectory.

In order to overcome the scale problem discussed above, built heritage can be studied in situ using both contact measurements (e.g. micro-erosion meters (MEM)) and direct measurements relative to a datum point (Moses et al., 2014). Relative measurement points include artificially introduced structures such as lead plugs and lead letters, or parts of a historic structure itself, such as unweathered surfaces and quartz veins (see Table 5.1). An example of this approach is the 30-year (1980–2010) investigation of limestone erosion on the balustrade at St Pauls Cathedral in London (Trudgill et al. 1989, 2001; Inkpen et al., 2012a, b), where both lead plug index and MEM methods were applied. These joint methods allowed the authors to analyse multi-directional surface change (i.e. surface lowering and elevation), rather than purely recession. Based on their results, Inkpen et al. (2012b) introduced a novel index of weathering accounting for the rate of surface change. While the rate of surface change index allows weathering behaviour to be described more comprehensively than before, it does not account for related stone property changes, such as decreases/increases of stone porosity and strength. Combining the index with methods that describe changes in the physical properties of stone, such as water uptake or surface hardness, would be beneficial.

Table 5.1 Examples of use of non-destructive contact methods to investigate stone surface changes on built heritage in-situ

Method	Principle	Surface orientation	Reference	Stone type	Time period
Lead plug index	Surface lowering or elevation in relation to datum point = lead plug	Horizontal	Trudgill 1989, 2001; Inkpen 2012a, b	Portland limestone	Short-term (1980-1987); long-term (1817-1990s)
Lead lettering index	Surface lowering or elevation in relation to datum point = lead letter(s)	Vertical	Cooke et al. 1995; Inkpen and Jackson 2000	Marble	
MEM / TMEM	Surface lowering or elevation in relation to preceding measurements	Mainly horizontal (one test area vertical)	Trudgill 1989, 2001 Inkpen et al. 2012a, b	Portland limestone	Short-term (1980-1987); long-term (1817-1990s)
Max. point recession	Surface lowering or elevation in relation to datum point = unweathered feature in the historic structure	Vertical	Mottershead 1997, 2000	Greenshist, siltstone, shale, slate, sandstone, mudstone, granite, quarz porphyry	Long-term (450-600 years)

Accurate quantification of stone property changes is complicated by the inherently high variance observed even in fresh stones (e.g. Siegesmund and Dürrast, 2010). This variability is expected to increase for longer weathering-stress histories (Cooper et al., 1992; Fort et al., 2013; McCabe et al., 2015), and has a significant impact on data evaluation (Trudgill et al., 1989; Van de Wall and Ajalu Msc, 1997; Hansen et al., 2013; Alberti et al., 2013). In cases where data are not normally distributed, the reliability of statistical estimates based on the assumption of normality may be affected, and it may be inappropriate to apply parametric testing (e.g. Tukey, 1977; Fowler et al., 1998; Filzmoser and Todorov, 2013). While few studies have addressed the issue of non-normality in stone weathering research, Niedzielski et al. (2009) and Wilhelm et al. (2016a) propose that non-parametric ('robust') statistical methods may be more appropriate for on-site testing of stone. Accordingly, Mosch and Siegesmund (2007) employ boxplots (with which non-normality and outliers become more evident) to display the natural variance of stone, and Feal-Pérez and Blanco-Chao (2012) deal with the problem of potential outliers in data from on-site rock testing by calculating the Huber M-Estimator. Mottershead (2000) and Wilhelm et al. (2016a) employ a set of robust measures (median, median absolute deviation (MAD)) and methods (Kruskal-Wallis, Mann-Whitney U, bootstrapped confidence intervals) in the field of stone weathering research. Erceg-Hurn and Mirosevich (2008) stressed the huge potential of non-parametric tests to improve data analysis, and found that a lack of exposure to, and misconceptions of, modern robust statistical methods were the principal reasons for their lack of popularity amongst researchers despite their clear advantages. Non-parametric methods may provide an appropriate solution to some of the challenges faced in stone weathering research on site at built heritage.

In order to complement existing methods for quantifying stone weathering rates, this study introduces a novel proxy that describes the rate of stone property change using surface hardness testing. This proxy, the gradient coefficient of quantile regression for median surface hardness change (QC_{50}), is used to account for the non-normality of stone surface hardness data. The QC_{50} parameter is employed to illustrate short- and long-term weathering behaviour of cemetery headstones in situ. The 12 Portland limestone headstones examined in this study cover an exposure period of 1 to 248 years, and their surface hardness is assessed using an Equotip D. The top and bottom sections of the headstones are differentiated to account for potential spatial variations. Robust statistical measures and methods are employed to increase the reliability of data evaluation.

5.1.2 Material and methods

Study sites

Historic cemeteries constitute unique repositories for investigating stone weathering behaviour under real world conditions over a variety of timescales. Comparing headstones installed at different dates allows weathering rates to be established through time, and their relatively large size allows comparative information to be collected from different sections of the stones (e.g. Cooke et al., 1995; Inkpen and Jackson, 2000). This study presents data from a time series of headstones on the Isle of Portland, Dorset, England (Figure 5.1). The Isle of Portland has a temperate maritime climate, with westerly or southwesterly prevailing winds (Wang et al., 2013). For the period 1981–2010 the annual average rainfall was 667.9 mm and annual average maximum and minimum temperature ranged from 13.4°C to 8.9°C (Met Office, accessed 10/08/2015). The main weathering agents for this area are thought to be salt spray, wind, and precipitation- and soil-based water (Viles, 2002b; Urosevic et al., 2013).

Figure 5.1 The Isle of Portland, showing the two cemeteries sampled in this study.

The study headstones were located in two coastal cemeteries: (i) Royal Naval Cemetery, cared for by the Commonwealth War Graves Commission (CWGC) (Figure 5.2); and (ii) the graveyard of St George's Church, maintained by the Churches Conservation Trust (CCT) (Figure 5.3). Both sites are heritage sites of cultural significance. St George's Church was built in the 18th century by John Gilbert (an apprentice of Christopher Wren), and is reminiscent of the tower of St

Paul's Cathedral (Colvin, 2008; CCT, 2015). As an outstanding example of 18th-century churches in Dorset, it is Grade I-listed (National Heritage List for England). Of interest to this study are the 2500 headstones in the graveyard, of which the vast majority is of made of Portland limestone. The CWGC was founded in 1917 and commemorates casualties of the Commonwealth forces from WWI and II around the world (1.7 million overall). In the UK, 308,000 CWGC headstones are maintained at 13,000 locations, and Portland limestone was the material of choice for the majority of the headstones (Bell and Coulthard, 1990; Godden, 2012; Viles, 2013). The Portland Royal Naval Cemetery on the Isle of Portland contains war graves of both World Wars, 154 casualties in total (CWGC, 2015).

Figure 5.2 The Royal Naval Cemetery on the Isle of Portland (CWGC, 2015).

Figure 5.3 St George's Church cemetery on the Isle of Portland.

Headstones investigated in this study are oolitic limestone from Portland limestone formation (Upper Jurassic) at the Isle of Portland (Figure 5.4). The limestone is white with micritic mass and the dominant grains are small to medium-sized (Leary, 1983; Jaynes et al, 1987; Palmer, 2008). Portland Base Bed and Portland Whit Bed are the two varieties most relevant for built heritage, but in this study the specific variety that made up the headstones could not be established. Due to its natural characteristics Portland limestone shows some variance in stone properties (Table 5.2; Gray 1861-1862; Townson, 1975).

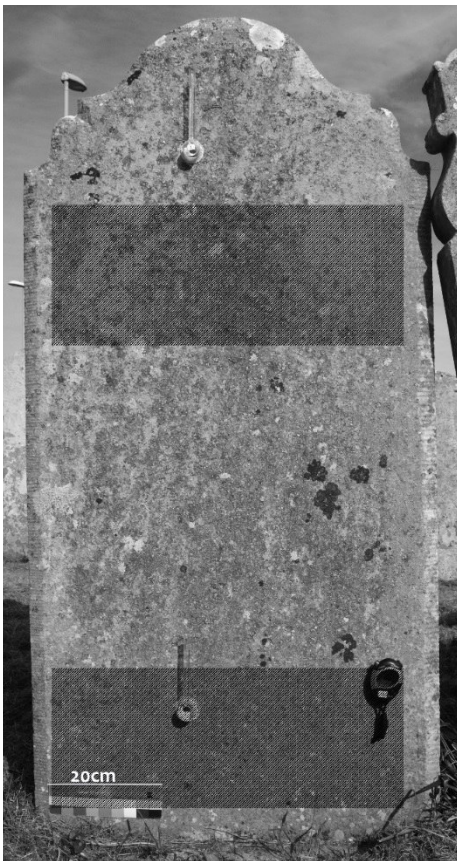


Figure 5.4 Headstone from set B (non-CWGC) at St George's Church cemetery, exposure 212 years, top and bottom area are indicated (shaded).

Table 5.2 Stone properties of two varieties of Portland Limestone (Base Bed and Whit Bed), [1]BRE test results Portland Whit Bed Bowers Quarry (1995-1997), [2]BRE test results Portland Coombefield Whit Bed, [3]BRE test results Portland Base Bed Bowers Quarry (1995-1997), [4]Wilhelm et al. (2016a), [5]Dubelaar et al., 2003.

Portland Limestone	Whit Bed	Base Bed
UCS [MPa]	42 (38-47)[1], 39[2]	52.8 (41.30-64.3)[3], 55.98 (43.2-75.73)[4]
Open porosity [%]	21.5 (20.8-22.1)[1], 21.47-25.37[5]	15.4 (13.7 – 16.8)[3]
Microporosity	30%[5]	75%[5]
Water absorption [wt%] (mean values)	6.3[1]	4.2[3]
Saturation coefficient	0.63[1], 0.60[2]	0.79[3]
Bulk specific gravity [kg/m³]	2128 (2110 - 2146)[1]	2320[3]
Sodium Sulphate Crystallisation [%wt loss]	8[1]	62.8[3]

The headstones were divided into two sets according to whether they were official CWGC headstones or not. Set A was a collection of 5 CWGC headstones based at both the Royal Naval and St George's cemeteries. The data for the Set A headstones were collected in June 2010. Set B was a collection of 7 non-CWGC headstones based only at St George's Church cemetery. The data for the Set B headstones were collected in May 2015. The two sets covered different periods of exposure: 1–91 years of exposure (Set A) and 145–248 years of exposure (Set B). Furthermore, the maintenance scheme differed between sets. Set A headstones have been maintained on a regular basis, with biocide being applied every year to prevent biological growth and cleaning occurring every 3–5 years (personal communication, CWGC). In contrast, Set B headstones have to our knowledge not been cleaned. For comparability reasons, all the tested headstones in both sets: (i) were located in open areas (no tree cover nor in line of surrounding buildings); (ii) had similar vertical axes (90°±8); (iii) faced East-West; and (iv) were similar in size (0.85-1.35 m). The backs of the headstones faced west, so no carving was evident on these surfaces.

Surface hardness testing

Although high impact surface hardness testing is a long-established method for relative dating of surface exposure in geomorphology (Goudie, 2006), very few studies have applied the technique to cultural stone. The main reason is that the usually employed device, the Schmidt Hammer (e.g. Aydin and Basu, 2005; Goudie, 2006; Fort et al., 2013; Stahl et al., 2013), has a high impact energy (Type L = 0.735 N m and type N = 2.207 N m (Proceq©, 2006)) that can damage the surface of the stone, and also the technique requires carborundum cleaning of the surface – both of which are unacceptable for most cultural heritage applications (Pope, 2000; Viles et al., 2011). As a result, low impact hardness testing (using devices such as Equotip and Duroscope) has been applied to evaluate the state of preservation/deterioration of cultural stone, for instance on a sandstone bridge in Japan (Aoki and Matsukura, 2007), at the Ta Nei Temple in Angkor Wat, Cambodia (Futagami et al. (2008; 2010) on public buildings and monuments in Hungary (Török, 2003, 2007, 2008). However, assessments of surface hardness change on cultural stone remain rare (e.g. Kamh and Koltuk, 2014; Matsui et al., 2014; Wedekind et al., 2014).

In this study the Equotip Piccolo 2 device with probe D was employed to measure surface hardness. This is comparable to Equotip 3 used in previous studies (Wilhelm et al., 2016a). This device measures the difference between impact and rebound velocity of a small, hard metal impact body traveling in a probe and propelled by spring force against the surface (Proceq© SA, 2010). The impact energy of the Equotip D is 0.0115 N m, which is much lower than that of the Schmidt Hammer (versions with yet lower impact energy are Equotip Type C = 0.003 N m and Type G = 0.090 N m (Proceq©, 2010)). Data are recorded in 'Leeb hardness' units (1 to 999), are stored automatically, and can be converted directly to all common hardness scales (e.g. Vickers, Rockwell etc.). Weathered stone surfaces are represented by low rebound values, whereas less weathered or case-hardened surfaces result in high rebound values.

5.1.3 Experimental setup

The surface moisture of each stone was measured using handheld moisture meters in resistivity and capacitance mode (Protimeter Surveymaster™ and CEM respectively) to make sure no significant moisture differences between the individual headstones and top and bottom sections were apparent. The measurement campaigns were preceded by dry weather period and conducted on dry and sunny days. Each headstone in the study was divided into top (maximum height from ground 100 cm) and bottom (minimum height 10 cm from ground) sections, and within each section 30 measurements were taken over an area covering about 0.2–0.4 m^2, 5 cm from the edge of the headstones (Figure 5.3). The Equotip was applied at random locations in each top and bottom section using the single impact method (expressed as HLD_S). It has been previously noted that Equotip data tend to underestimate surface hardness in shallow zones of weathering with discontinuities, when the surface is not cleaned beforehand (Hack and Huismann, 2002). However, since surface roughness and weathering are intimately related, comparing hardness across different stones with similar surface textures is a legitimate approach (McCarroll, 1991). Therefore, the tested surfaces in this study were deliberately not manipulated (i.e. scrubbed or cleaned with a carborundum) prior to using the Equotip. The authors are aware of the effects of non pre-treated surfaces on low impact hardness testing, but aim to utilize these effects to describe surface changes. Therefore, the results should be considered as expressions of surface change as opposed to changes in absolute surface hardness.

5.1.4 Data evaluation

The Shapiro-Wilk test for normality (e.g. Park, 2008; Razali, 2011) revealed that some surface hardness datasets were not normally distributed, as has been observed before for heterogeneous natural stone (Mosch and Siegesmund, 2007; Palmer, 2008; Hansen et al., 2013; Alberti et al., 2013; Emmanuel, 2015). Thus, to provide a practical and robust approach, and to account for the natural variability of stone, this study followed the approach of Wilhelm et al. (2016a) and avoided data transformation by using mainly non-parametric measures and methods, which are more accurate for non-normally distributed data whilst adequate for normal data (e.g. Erceg-Hurn and Mirosevich, 2008; Niedzielski et al., 2009; Filzmoser and Todorov, 2013).

The median of 30 single impact measurements, which describes the average surface hardness, is expressed as $HLD_{S.med}$. The Spearman's rank correlation coefficient (r_s (degrees of freedom), p = p-value) was employed to correlate $HLD_{S.med}$ values with the exposure time of each headstone cf. Hauke and Kossowski, 2011). Confidence intervals per respective exposure year were also calculated using bias corrected and accelerated (bca) bootstrapping (10,000 iterations) at a 95% confidence level (Efron, 1987). Our assumption was that the width of the confidence intervals would increase with exposure time, and also that they would provide a reliable interval over which to relate exposure age for Portland limestone surface changes.

This study introduces QC_{50}, the regression coefficient for the 0.50 quantile (median), as a novel proxy for determining the rate of surface change. Quantile regression (in contrast to least-squares regression) is robust against outliers and heteroscedasticity (Cade and Noon, 2003; Koenker, 2005; Crawley, 2007; Dette and Volgushev, 2008). Non-parametric non-crossing quantile regression (0.25, 0.50, 075 quantile, bootstrapping 1,000 iterations) was applied to derive robust estimates of surface change rates (Bondell et al., 2010). The 0.50 quantile shows the rate of change in median surface hardness over time, whereas 0.25 and 0.75 quantiles mark the rate of change of the inter quartile range (IQR) of the datasets, which provides information about how the dispersion of surface hardness values changes over time and whether this happens in a homo- or heterogeneous manner. Understanding variance changes over time can add to the understanding of stone weathering behaviour.

5.2. Results

5.2.1 Spatio-temporal differences between Set A and Set B

Figure 5.5 displays boxplots of HLD_S (single values) according to exposure year for all the study headstones, for the top (a) and bottom (b) sections. Spearman's Rank testing shows a moderate (negative) trend for $HLD_{S.med}$ over time on the top sections of the headstones (r_s = –0. 56 (p = 0.050)). $HLD_{S.med}$ values declined until 91 years of exposure whereas the decline for stones with longer exposure time stagnated. For the bottom sections of the headstones, Spearman's Rank tests show a strong (negative) association (r_s = -0. 73 (p = 0.006)), indicating a strong downwards trend. A clear break between Set A (1-91 years) and Set B (145-248 years) for both top and bottom sections is apparent and supported by the results of the breakpoint analysis (piecewise regression was applied using the package 'segmented' in RStudio (Muggeo, 2003, 2008; Crawley, 2005). Thus, there is strong evidence that Set A and Set B headstones have different surface change rates as measured by surface hardness changes. For the following data evaluation the two datasets are treated separately.

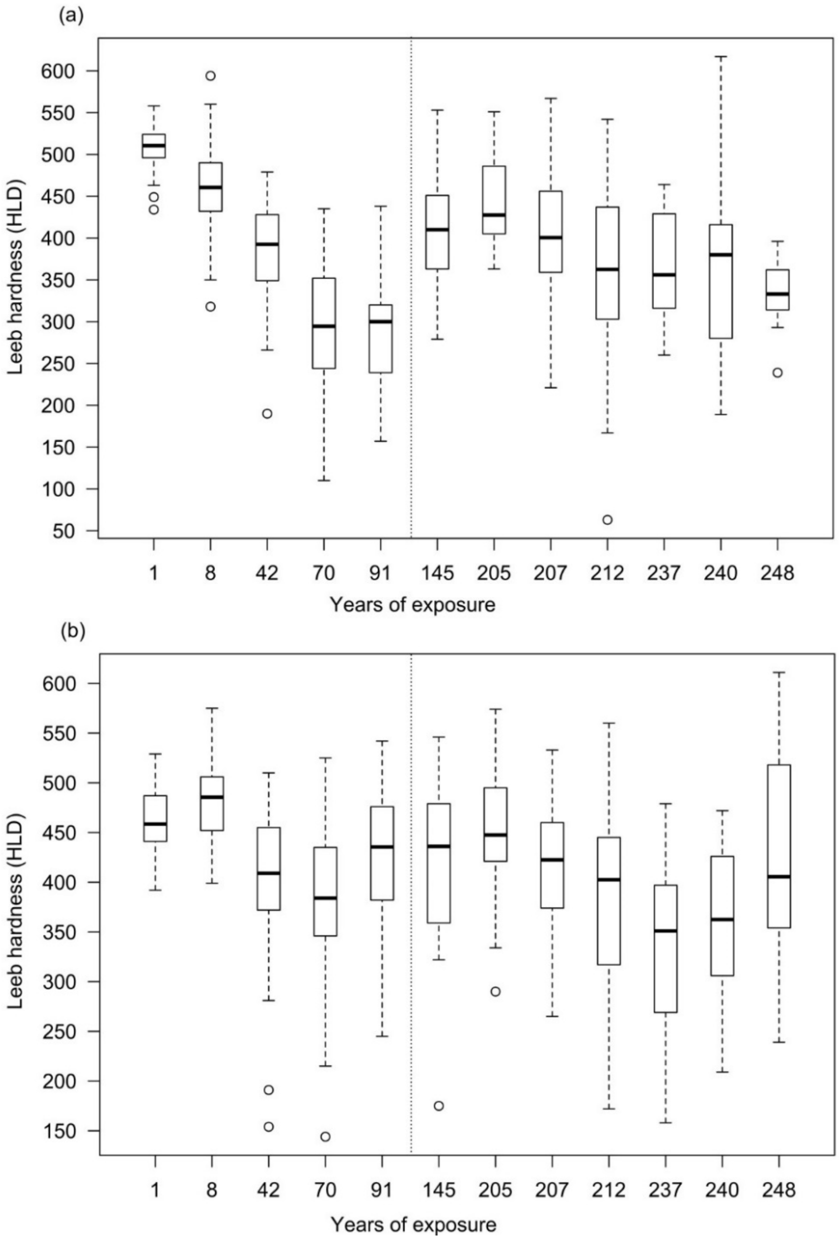


Figure 5.5 Boxplots of surface hardness (HLD_S) values for the top (a) and bottom (b) sections of the tested limestone headstones of this study versus exposure years. Dashed line divides Set A and Set B

5.2.2 Differences within individual sets

Set A – Commonwealth War Graves Commission headstones (CWGC)

Table 5.3 summarizes surface hardness data for Set A. Interestingly, with one exception (8 years of exposure), the majority of headstones in Set A show lower surface hardness on the top section compared to the bottom section. Accordingly, Mann-Whitney U results are significant for the majority of the datasets (Table 5.4). Figure 5.6 and 5.7 show scatterplots and quantile regression for Set A against exposure time, for the top and bottom sections (respectively). The rate of surface hardness change for the top sections of Set A are the highest recorded in this study (QC_{50} = -2.42) (Table 5). This is considerably higher than rate found for the bottom sections (QC_{50} = -0.72). 0.25 and 0.75 quantile regression coefficients varied from 0.50 quantile regression coefficient and from each other, suggesting that the data are heteroscedastic. Similarly, the bootstrapped confidence intervals of $HLD_{S.med}$ tend to widen as exposure time increases (Figure 5.6, 5.7 and Table 5.3). This reflects the general increase of data variability caused by extended weathering (e.g. Fort et al., 2013; McCabe et al., 2015).

Table 5.3 Surface hardness data for Set A, $HLD_{S.med}$ (MAD) = median (median absolute deviation) of 30 single impact readings per headstone. Dashed line divides top and bottom sections. Bootstrapped upper and lower ci = upper and lower confidence interval limit for $HLD_{S.med}$.

Year of death	Years of exposure	$HLD_{S.med}$ (MAD)	Lower ci	Upper ci	$HLD_{S.med}$ (MAD)	Lower ci	Upper ci
		Top section			Bottom section		
2009	1	510.5 (20.01)	500	518.5	458.5 (11.6)	444	481.5
2002	8	460.5 (43.0)	438	477.5	485.5 (7.2)	467	503
1968	42	392.5 (54.11)	361	418.5	409.0 (21.8)	385	445.5
1940	70	294.5 (79.32)	250	331	384.0 (26.6)	364	424
1919	91	300.0 (75.61)	254	310.5	435.5 (16.7)	397	462.5

Table 5.4 Mann-Whitney U results (two-tailed with a significance level of p-value 0.05) to investigate significant spatial differences in surface hardness (*HLDs*) between top and bottom sections of single headstones in Set A (1-91 years). Significant differences are marked bold.

Year of death	Years of exposure	p-value	U
2009	1	**0.0001**	165.0
2002	8	0.1557	309.0
1968	42	0.2301	357.0
1940	70	**0.0001**	200.0
1919	91	**<0.0001**	72.0

Table 5.5 Set A coefficients of quantile regression for association of exposure years and change of surface hardness (SIM single values). Key: Intercpt = intercept (std error). qr25.coef., QC50 and qr0.75.coef. are the coefficients/gradients (std error) of the respective quantile. QC50 is the novel proxy introduced in this study and marked bold.

Set A	Intercpt	qr25.coef.	Intercpt	**QC_{50}**	Intercpt	qr75.coef.
Top	466.93 (9.59)	-2.93 (0.19)	494.35 (8.74)	**-2.42 (0.21)**	516.21 (5.68)	-2.21 (0.18)
Bottom	440.03 (5.74)	-1.03 (0.15)	463.72 (10.92)	**-0.72 (0.20)**	487.79 (8.83)	-0.35 (0.20)

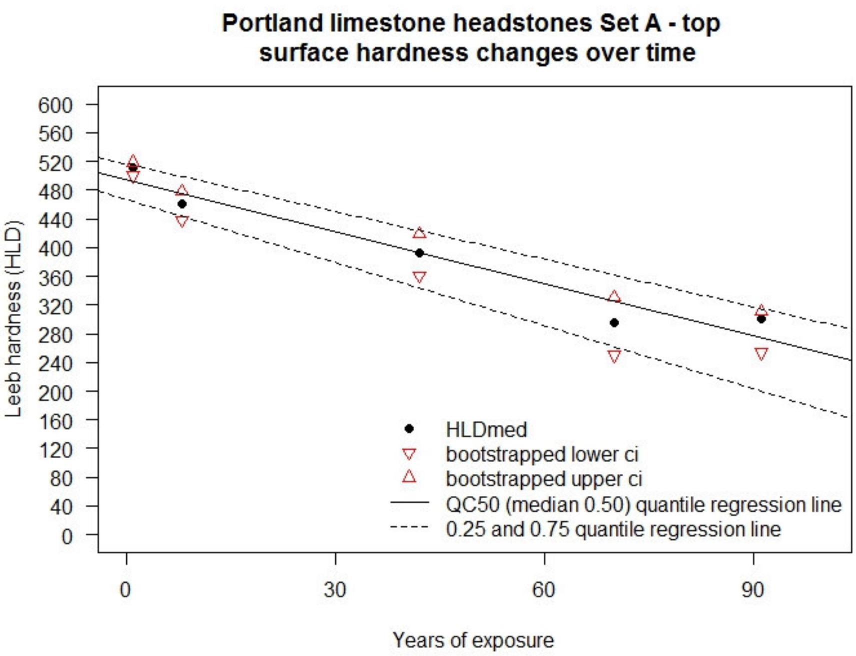


Figure 5.6 Scatterplot of surface hardness median values ($HLD_{S.med}$) for the top sections of Set A (CWGC headstones) and years of exposure. Each point represents a dataset of 30 Equotip readings. The solid line represents the quantile regression for the 0.50 quantile (median) and the lower and upper dashed line represent 0.25 quantile and 0.75 quantile respectively. The lower and upper triangles mark the confidence interval range for bootstrapped $HLD_{S.med}$. Overall time period 91 years.

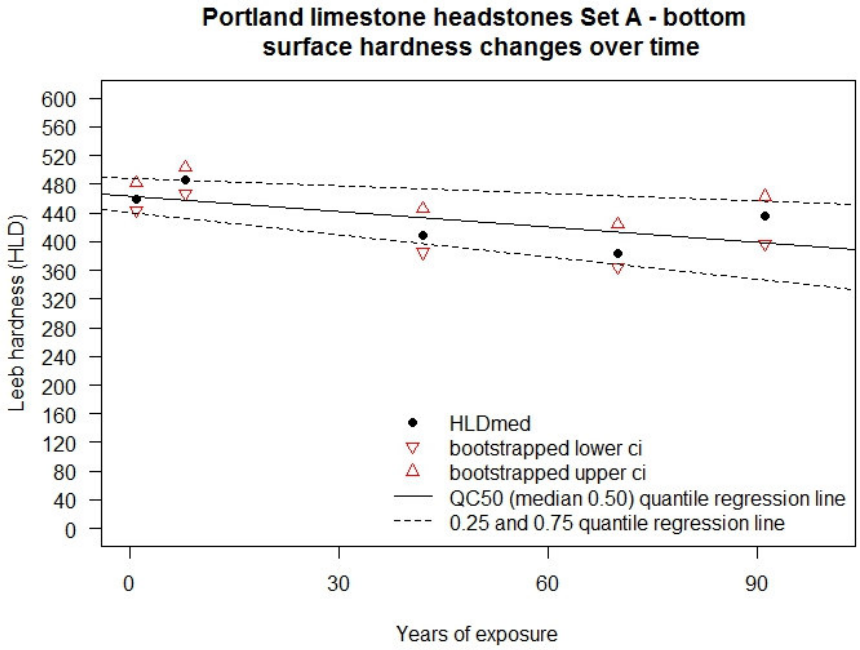


Figure 5.7 Scatterplot of surface hardness median values ($HLD_{S.med}$) for the bottom sections of Set A (CWGC headstones) and years of exposure. Each points represent a dataset of 30 Equotip readings. The solid line represents the quantile regression for the 0.50 quantile (median) and the lower and upper dashed line represent 0.25 quantile and 0.75 quantile respectively. The lower and upper triangles mark the confidence interval range for bootstrapped $HLD_{S.med}$. Overall time period 91 years.

Set B – Non-CWGC headstones

Table 5.6 summarizes surface hardness data for Set B. Similarly to Set A, with two exceptions (237 and 240 years of exposure), the majority of headstones in Set B showed lower surface hardness on the top section compared to the bottom section. In contrast to Set A, this spatial difference was only significant for the exposure time of 248 years (Mann-Whitney U, Table 5.7). However, higher MAD values are observed for the top sections compared to the bottom sections and indicate a higher variance of surface hardness at the top of the headstones. Although the MAD values are higher in comparison to the top sections of set A, they do not increase over time. In contrast to set A this indicates a different weathering behaviour where the data variability does not increase naturally over time, but remains high throughout. Figure 5.8 and 5.9 show scatterplots and quantile regression for Set B according to exposure time, for the top and bottom sections (respectively). The rate of surface hardness change is similar for the top and bottom sections (QC_{50} = -0.81 and QC_{50} = - 0.83 respectively) (Table 5.8). The rates determined for both top and bottom sections of Set B are slightly higher than the rate determined for the bottom sections of Set A, but considerably lower than the rate for the top sections of Set A.

Table 5.6 Surface hardness data for Set B, $HLD_{S.med}$ (MAD) = median (median absolute deviation) of 30 single impact readings per headstone. Dashed line divides top and bottom sections. Bootstrapped upper and lower ci = upper and lower confidence interval limit for $HLD_{S.med}$.

Year of death of the deceased	Years of exposure	$HLD_{S.med}$ (MAD)	Lower ci	Upper ci	$HLD_{S.med}$	Lower ci	Upper ci
		Top section			Bottom section		
1870	145	410.0 (60.79)	375	439	436.0 (16.6)	386	467
1810	205	427.5 (61.53)	412	469	447.5 (14.4)	433	480.5
1808	207	400.5 (71.91)	371	446.5	422.5 (19.2)	391	440.5
1803	212	362.5 (99.33)	305	411.5	402.5 (23.0)	352	438.5
1778	237	356.0 (74.13)	340	398	351.0 (32.9)	291	382
1775	240	380.0 (30.39)	326	398	362.5 (30.7)	329	421
1767	248	333.0 (82.28)	320	353.5	405.5 (22.5)	358	486.5

Table 5.7 Mann-Whitney U results (two-tailed with a significance level of p-value 0.05) to investigate significant spatial differences in surface hardness (*HLDS*) between top and bottom sections of single headstones in Set B (145-248 years) . Significant differences are marked bold.

Year of death	Years of exposure	p-value	U
1870	145	0.2249	367.0
1810	205	0.5928	392.5
1808	207	0.4590	395.5
1803	212	0.2796	377.0
1778	237	0.1528	350.0
1775	240	0.7813	449.5
1767	248	**0.0016**	213.0

Table 5.8 Set B coefficients of quantile regression for association of exposure years and change of surface hardness (SIM single values). Key: Intercpt = intercept (std error). qr25.coef., QC50 and qr0.75.coef. are the coefficients/gradients (std error) of the respective quantile. QC50 is the novel proxy introduced in this study and marked bold.

Set B	Intercpt	qr25.coef.	Intercpt	**QC_{50}**	Intercpt	qr75.coef.
Top	523.05 (52.51)	-0.82 (0.22)	563.25 (29.7)	**-0.81 (0.15)**	579.45 (67.96)	-0.70 (0.30)
Bottom	450.26 (73.91)	-0.84 (0.33)	581.63 (55.94)	**-0.83 (0.26)**	542.47 56.57)	-0.39 (0.28)

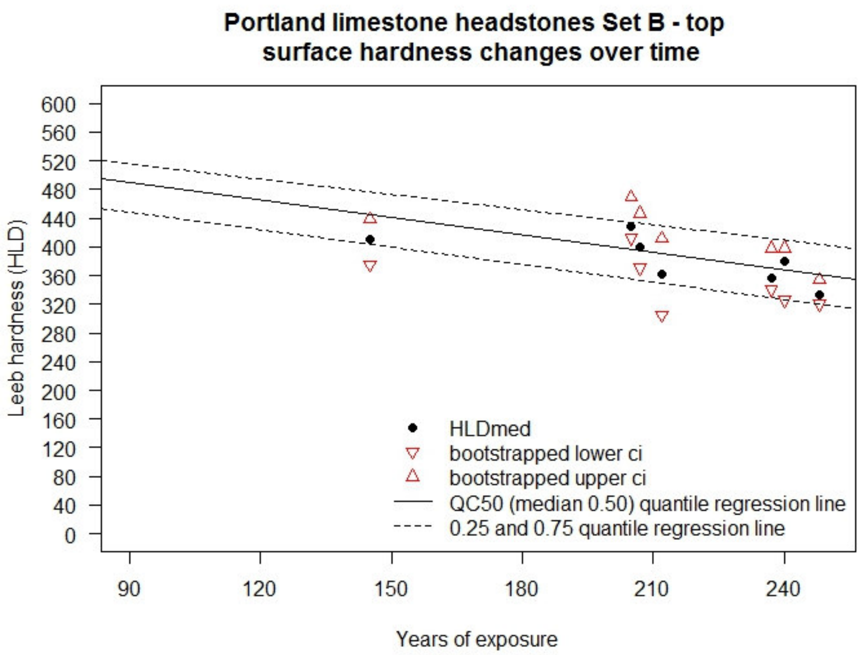


Figure 5.8 Scatterplot of surface hardness median values ($HLD_{S.med}$) for the top section of Set B (non-CWGC headstones) and years of exposure. Each point represents a dataset of 30 readings. Solid line represents quantile regression for 0.50 quantile (median) and the lower and upper dashed line represent 0.25 quantile and 0.75 quantile respectively. The lower and upper triangles mark the

confidence interval range for bootstrapped $HLD_{S.med}$. Overall time period 248 years.

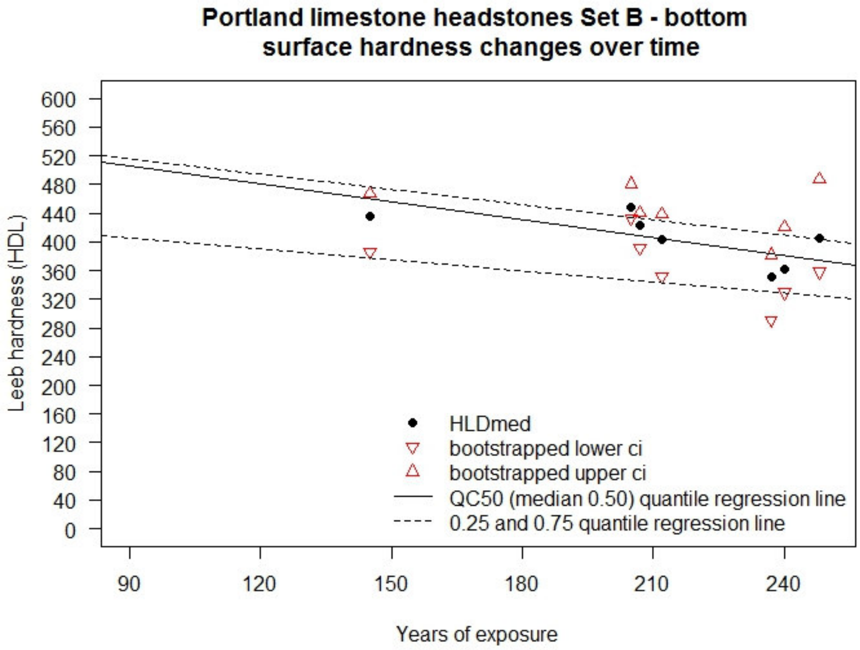


Figure 5.9 Scatterplot of surface hardness median values ($HLD_{S.med}$) for the bottom section of Set B (non-CWGC headstones) and years of exposure. Each point represents a dataset of 30 readings. Solid line represents quantile regression for 0.50 quantile (median) and the lower and upper dashed line represent 0.25 quantile and 0.75 quantile respectively. The lower and upper triangles mark the confidence interval range for bootstrapped $HLD_{S.med}$. Overall time period 248 years.

5.3. DISCUSSION

It is interesting that Set B, which has a longer weathering history than Set A, shows a lower rate in surface hardness change compared to the rates determined for the top sections of Set A. A longer weathering-stress history with accumulated impacts such as air pollution, salt spray and wind abrasion would be expected to result in a higher susceptibility to weathering, especially in a coastal environment (e.g. Cooke and Gibbs, 1994; Mottershead et al., 2003). Furthermore, the existence of strong differences in QC_{50} rates between the top and bottom sections of Set A headstones, and the relative lack of difference in

QC_{50} rates between the top and bottom sections of Set B headstones, must also be explored. A range of potential factors could explain the differences between the two datasets, and these are examined. Macro-scale weather effects were excluded as a differentiating factor since the two cemeteries are only 2 km apart and both are close to the coast (CWGC naval cemetery ~1.3 km and St George's cemetery ~0.5 km).

5.3.1 Portland limestone varieties – Base Bed vs Whit Bed

The two Portland limestone varieties most relevant for built heritage are Portland Base Bed and Portland Whit Bed. The latter has a reputation for being the more durable building stone, due to its beneficial pore characteristics (e.g. Leary, 1983; Dubelaar et al., 2003; Godden, 2012). However, occasionally Portland Base Bed has been favoured over Whit Bed for aesthetic and workability reasons (the lack of visible shell content), and thus has been used "extensively where a faster rate of weathering is acceptable or where its working qualities were required" (Edmunds and Schaffer, 1932; BRE, 1997a, p.1). Previous research established a weathering rate in the UK for Portland Whit Bed limestone of 1–2 mm surface recession per 100 years, with greater recession under severe exposures (BRE, 1997b). In contrast, Portland Base Bed has a higher weathering rate of 3–4 mm surface recession per 100 years, with greater recession under severe exposures or on the edges of stonework (BRE, 1997a,b). Despite these crucial differences affecting durability and decay of the two varieties, they are rarely distinguished in the weathering and built heritage literature. Additionally the beds themselves display a natural variability which might complicate attempts to distinguish them (Gray, 1861-1862; Edmunds and Schaffer, 1932). Thus, it is possible that the Set A headstones (CWGC) were made from Portland Base Bed, which would explain the higher rate in hardness change as this variety is known to be less durable (Leary, 1983; Dubelaar et al., 2003; Godden, 2012). This variety may have been chosen deliberately (although this cannot be confirmed through records, CWGC personal communication 07/2015) or accidentally, as confusion in the nomenclature ("Best Bed" for Base Bed) and phrasing ("Whit Bed without shells") caused mixed usage of both Portland Base Bed and Whit Bed in the past (Gray, 1861-1862; Edmunds and Schaffer, 1932). Nevertheless, this does not explain the lack of spatial differences between top and bottom for Set B.

5.3.2 Weathering-stress history and stone variability

Another factor potentially affecting weathering behaviour of the headstones in this study is the different time periods covered by the two datasets. The last 91 years may have seen more rapid weathering rates and/or greater differences in rates between top and bottom because of altered environmental, climate or air quality factors, whilst longer exposure times would even out spatial differences for Set B. Another explanation for lower surface hardness change rates on the older headstones could be the natural variation in texture and composition resulting in different stone quality over 240 years of quarrying (Gray, 1861-1862; Bell, 1993; Logan et al., 1993; Van de Wall and Ajalu, 1997; McCabe et al., 2015). Increased complexity of stone decay systems with exposure has been described by McCabe et al. (2015), who finds both stone and micro-environmental climate with accumulative exposure history results in surface-to-depth heterogeneity at the stone block scale.

5.3.3 Stone maintenance and biological growth

A likely explanation for the higher rate of surface hardness change for set A cf set B headstones might be the maintenance scheme; frequent cleaning, application of biocides and removal of biogrowth like lichens. The destructive or protective role of lichens on stone is the subject of much scientific controversy. Their destructive role is associated with disaggregation of the stone surface, dissolution processes, precipitation and formation of new minerals like oxalates (e.g. Chen et al., 2000; Bjelland and Thorseth, 2002; St. Clair and Seaward, 2004). Nevertheless, some studies have reported that lichens can play a protective role in environments otherwise experiencing rapid weathering by mediate thermal stresses (keeping surfaces hot and dry more constantly), reduce chemical reactions (water and pollutants) and decrease physical impacts (wind and wind driven rain) and stabilise grains on the surface (as demonstrated by, for example Ariño et al., 1995; Seaward, 2001; Garcia-Vallès et al., 2003; Mottershead et al., 2003; Carter and Viles, 2005a; Özvan et al., 2015).

Thus, for the Set B headstones in this study exhibiting lichen cover, the slower rate of surface hardness change might be attributed to bio-protection through lichens as opposed to physical and chemical deterioration processes through other extrinsic factors such as anthropogenic (cleaning and biocide application) or environmental (wind and sea spray) impacts. This supports existing evidence that lichens are an important factor to be considered when it comes to investigating weathering behaviour of stone built heritage. However, it is not clear why this difference in cleaning regime (and thus lichen cover) would lead to the large differences in weathering rates between the top and bottom sections of the Set A headstones and the much closer top and bottom trends in Set B headstones.

5.3.4 Synergistic effects of microclimate, surface condition and stone maintenance

Microclimatic effects may explain the existence of distinct spatio-temporal differences in rate of surface hardness change between the top and bottom sections of some of the headstones. Higher weathering rates observed on the top sections of Set A headstones could result from higher exposure to precipitation and as well as greater evaporation rates resulting in a higher frequency of crystallisation-dissolution processes (salt and ice), which in turn increase the rate of decay (Pope, 2000; Viles, 2005; Smith et al., 2008; Camuffo, 2014). In contrast, a different water regime is expected to impact the bottom sections. Water run-off from the top may transport dissolved $CaCO_3$ down the stone surfaces and precipitate it at the bottom sections (forming crusts), but also competing with rising water and evaporation processes and potentially resulting in lower frequency of crystallisation-dissolution processes while increasing time of wetness (and thus retarding the weathering rate). Therefore, the synergistic effects of lack of bio-protection and therefore more pronounced microclimatic effects might explain the spatial differences in surface hardness changes in the case of Set A headstones, which are generally absent for the Set B headstones because they have not undergone a rigorous and regular cleaning regime.

5.4. CONCLUSION

In this study, a time series of headstones covering a weathering history period of 1 to 248 years was investigated on the Isle of Portland (UK). The use of a low impact device (the Equotip D) allowed surface hardness of the headstones to be tested in situ. Data were analysed using robust, non-parameteric statistical methods, because surface hardness datasets of heterogeneous natural stone are rarely normally distributed (Mosch and Siegesmund, 2007; Palmer, 2008; Hansen et al., 2013; Alberti et al., 2013; Emmanuel, 2015). We introduce a novel robust proxy, the 0.50 quantile regression coefficient of surface hardness (QC_{50}), to describe rates of stone surface property change under real world conditions of stone built heritage. This provides a complementary approach to methods in previous studies that describe stone weathering behaviour using observations of surface change (Inkpen et al. 2012a, b).

When considered the weathering history of set A and set B as coherent time series at comparable geographical locations, non-linear stone weathering behaviour is observed. However, a range of explanations for non-linearity is possible. Accordingly, piecewise regression showed a clear break between Set A (1-91 years) and Set B (145-248 years) for both top and bottom sections. As a result set A and B were compared to each other and rates of surface hardness change for different time segments were established using the QC_{50} parameter. On average, headstones with c. 100 year exposure history displayed the greatest rates surface hardness change in the top section of each headstone (QC_{50} = -2.42), which is significantly higher compared to the rate of the bottom section

(QC_{50} = -0.72). Headstones with an exposure history of between 100 and c. 250 years, showed low rates of surface hardness changes (QC_{50} = -0.81 for top sections and QC_{50} = - 0.83 for bottom sections). Interestingly, the difference between rates for the top and bottom sections is not significant for headstones with longer weathering histories. Suggested interpretations for the observed weathering behaviour differences of the limestones in this study are:

a. Natural variation in texture and composition in Portland limestone has resulted in different stone quality and/or different weathering stress history during 240 years of quarrying;
b. Two Portland limestone varieties (Portland Base Bed and Portland Whit Bed) may have been jointly used for producing headstones, with the latter being considered more durable;
c. Different stone maintenance schemes may have resulted in different bioprotective lichen cover, which, coupled with synergistic effects on the dynamism of the microclimatic weathering system at the top of the headstones, could have led to different weathering rates due to changes in the frequency of crystallisation-dissolution processes.

In order to gain a deeper insight into stone weathering behaviour, it would be beneficial to determine both rates of stone property changes by means of surface hardness with rates of surface changes with methods such as those outlined in Table 5.1 or laser scanning methods. This would drastically improve our understanding of the short and long-term evolution of weathered surfaces on built heritage under real world conditions and furthermore, have implications for future conservation strategies.

Acknowledgements

K. Wilhelm is in receipt of an EPSRC studentship (funded via grants EP/P504287/1, EP/P505216/1 and EP/P505666/1) with additional support from Proceq. We thank the Leverhulme Trust for funding for the earlier stages of this project through Research project grant NBo. F/08 773/F.

We thank Michael Greet (CWGC) for providing us with helpful information on the CWGC headstones. Roger Stone (quarryman) gave invaluable insight into Portland quarry procedures, history and build stone choices. We thank three anonymous reviewers for their comments, which undoubtedly helped to improve this paper.

6. OBJECTIVE 3: INVESTIGATION OF STONE WEATHERING BEHAVIOUR OF ANCIENT LIMESTONE HERITAGE WITH MAINLY UNKNOWN HISTORY ('REAL WORLD' CONDITIONS) USING IMPROVED NON-DESTRUCTIVE TECHNIQUES

Catastrophic limestone decay after a harsh winter at an archaeological site in South Turkey was reported by collaborating archaeologists. The suddenness and severeness of heritage stone decay determined the focus of objective 3 and the cause for catastrophic stone decay in situ were reconstructed using non-destructive measuring techniques and past climate data reports. Paper 4 reports the main finding and is the core of this chapter. It has been submitted to the Journal of Archaeological Science. The following main research questions are addressed: What caused the catastrophic limestone decay? What are the implications for conservation interventions and future site management?

6.1. CATASTROPHIC LIMESTONE DECAY AT THE CENTRAL SANCTUARY OF IUPITER DOLICHENUS AT DÜLÜK BABA TEPESI IN SOUTH TURKEY: CAUSES AND IMPLICATIONS FOR FUTURE CONSERVATION

K. Wilhelm[1]*, H. Viles[1], E. Winter[2], O. Burke[3], S. Engelstaetter[4] , K.Z. Coyte[5]

[1] Oxford Rock Breakdown Laboratory (OxRBL), School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK
[2] Research Centre Asia Minor, University WWU Münster, Georgskommende 25, 48143 Münster, Germany
[3] Nuffield Department of Population Health, University of Oxford, Old Road Campus, OX3 7LF, UK
[4] Climate Research Lab, School of Geography and the Environment, University of Oxford, South Parks Road, Oxford, OX1 3QY, UK
[5] Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, UK
* katrin.wilhelm@ouce.ox.ac.uk

Keywords: Non-destructive testing; archaeological excavation; stone heritage deterioration; stone conservation science; climate impacts.

Abstract

Dramatic deterioration of Hellenistic-Roman limestone remains recently excavated at Dülük Baba Tepesi (Southern Turkey) has been observed following the cold, wet winter of 2011/2012. A conceptual model is presented to explain the dramatic deterioration in which case hardening develops and initially strengthens the stone against deterioration, but then makes it more prone to exfoliation and blistering. Data collected using non-destructive techniques (Equotip surface hardness tester and Karsten tube for water uptake) on Fırat and Gaziantep formation limestone time series excavated in 2005, 2007 and 2013 demonstrates the progress of case hardening and deterioration after excavation. In combination with meteorological data from Gaziantep weather station the results are used to test and revise the model taking into account non-linear and threshold effects. Future excavation and conservation efforts should take into account the often complex interactions between post-excavation case hardening and extreme winter conditions which can cause catastrophic deterioration.

6.1.1 Introduction

Catastrophic decay of soft limestones within Hellenistic-Roman archaeological remains was observed at the sanctuary site of Iupiter Dolichenus, Dülük Baba Tepesi in South Turkey after the winter of 2011/2012 (Figure 6.1 and 6.2). Such a dramatic event had not previously occurred during long-term excavations carried out since 2001 at the site by the Research Centre Asia Minor, University of Muenster. Notable loss of archaeological stone heritage was caused, which is of considerable concern given the planned sequence of future excavation work at the site. This paper combines in situ non-destructive testing methods and climate data evaluation to address the nature and causes of the catastrophic deterioration, in order to aid future excavation and conservation planning.

Figure 6.1 Limestone (probably Gaziantep formation) excavated in 2005. Image taken in 2011 (Photo© Engelbert Winter 2011).

Figure 6.2 Catastrophic decay of limestone (probably Gaziantep formation) in figure 40 in 2012 (Photo© Engelbert Winter 2011).

Preliminary investigations on site in 2014 identified two distinctive types of limestone which had been affected by dramatic deterioration in the winter of 2011/2012. The Fırat and Gaziantep Formation, both of which are soft limestones and prone to develop case hardened surfaces. Previous research has demonstrated the importance of such surface crusts to the development of episodic catastrophic deterioration on vulnerable limestones, often in combination with freeze-thaw weathering induced by harsh winter conditions (e.g. Smith and Viles, 2006; Martínez-Martínez et al., 2013). Using data from a time-series of stones excavated at different dates (2005, 2007 and 2013) in combination with climatic records this study reconstructs the interlinked history of case hardening and deterioration.

The site of Dülük Baba Tepesi (6.3) is the origin of the Roman god Iupiter Dolichenus, one of the most important oriental deities of the Imperium Romanum in 2nd to 3rd century AD (Blömer, 2012). The significance of the archaeological site stems from it being one of the few sites in the region of South-East Anatolia (Turkey) where sacrificial activity is continuously evident from 1000 BC to late antiquity (Blömer, 2011 and Winter, 2014). Consequently, in 1997 Anïtlar Yüksek Kurulu Adana defined the site as a 'First class archaeological protection zone' (birinci derece sit alanı) (Blömer & Winter, 2005). Future plans of the Turkish General Directorate of Cultural Heritage involve further excavation and the development of the site into a public archaeological park.

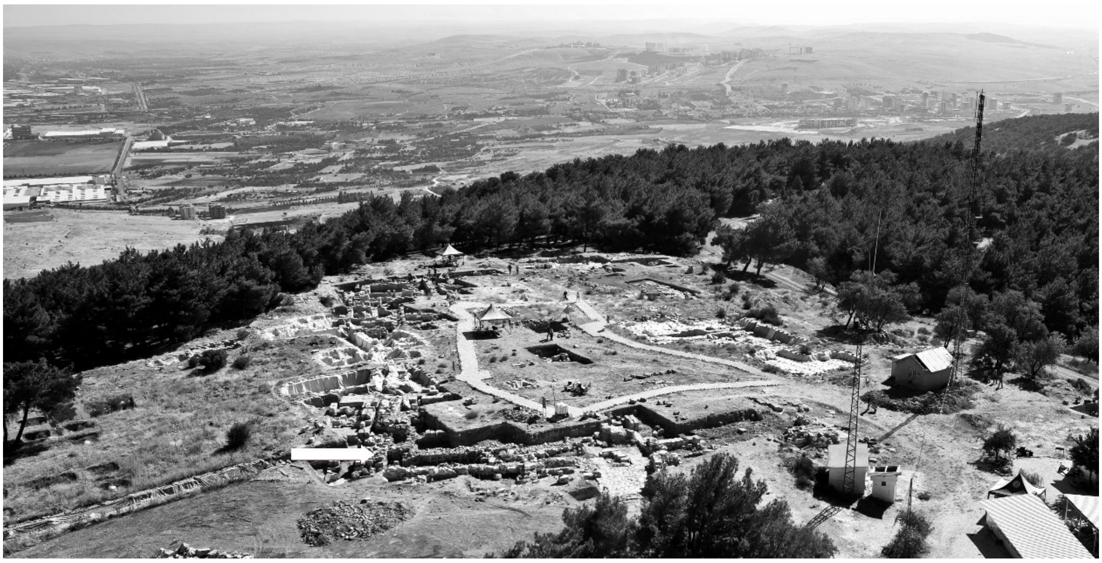


Figure 6.3 Aerial image of Duluk Baba Tepesi, a significant archaeological site in southern Turkey, looking towards the East (white arrow indicates trenches of interest to this study) (Photo© Peter Jülich 2014).

The architectural remains at Dülük Baba Tepesi are constructed from a range of materials including basalt, brick and several varieties of limestone depending on the historic period. This paper focuses on the soft limestone blocks of the Hellenistic-Roman structures as they show particularly serious deterioration. A range of deteriorative mechanisms including freeze-thaw and salt weathering, dissolution and biological weathering is known to affect such soft limestones, but freeze-thaw weathering is likely to be of particular importance here given the climatic and geographic setting of the site. These soft limestones are also prone to the development of superficial indurated layers, called 'case hardening' (schematically shown in 6.4), which are often associated with dissolution and re-precipitation of calcite cements (Smith and Viles, 2006; Hendrickx, 2013). Calcite within the porous near-surface zone becomes dissolved in rainwater acidified with CO_2 (Lipfert, 1989; Smith and Viles, 2006).

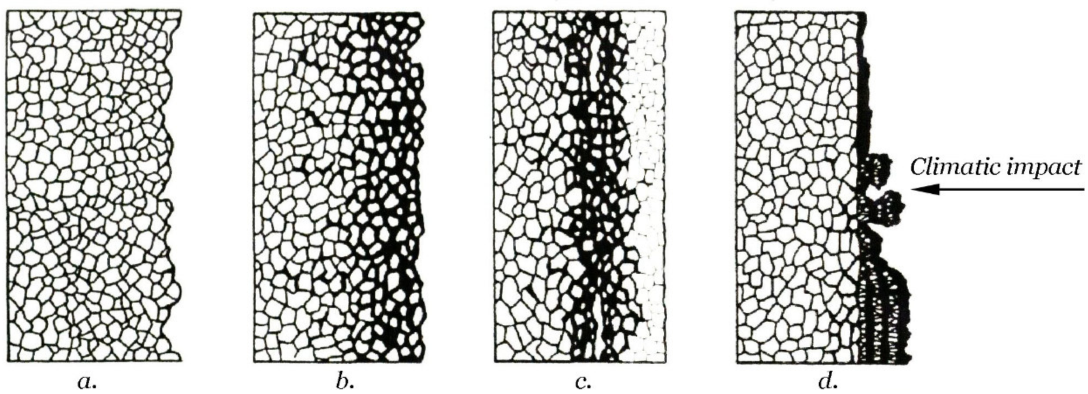


Figure 6.4 Schematic description for common weathering profiles of stone surfaces; a. superficial (granular) disintegration & erosion, b. increased porosity & decline of intergranular bonds, c. case hardening with increased superficial density followed by zone of increased porosity, d. crust (altered after Grimm, 2010; p 176).

Moisture regimes within the limestone, especially in areas with high evaporation rates, favour the re-precipitation of the calcite in a narrow band (a few millimetres thick) close to the stone surface (Winkler, 1994). Once case hardening has been initiated, the stone behaves as a two material system (composite) as the indurated surface now has different physical properties compared to the subsurface and core of the stone (which may be further softened) as shown in the flowchart in Figure 6.5**Error! Reference source not found.**. As a consequence, the deterioration potential of (extreme and/or frequent) climatic impacts like thermal fluctuations, water ingress and crystallisation events (frost and salt) on the case hardened stone is increased and can lead to catastrophic decay (Smith and Viles, 2006). Instead of freeze-thaw events causing minor granular disintegration of the soft limestones, catastrophic exfoliation and blistering of the indurated layer and the exposure of a 'core softened' zone underneath is likely. As new surfaces of soft limestone become exposed to the air (either through excavation or exfoliation of surface layers) the episodic weathering cycle starts again. The rates of the processes in Figure 6.5 are determined by the properties of the material involved and the climatic conditions experienced.

Figure 6.6 illustrates how the conceptual model of the interplay of case hardening and episodic freeze-thaw events in Figure 6.5 might be applied to Fırat and Gaziantep Formation limestones which have been excavated at different dates (2005, 2007, 2013) at Dülük Baba Tepesi. The model assumes a linear development of case hardening (5 Leeb surface hardness units per year) over time applicable to both limestones, and uses surface hardness values drawn from the literature. An episodic period of intense freeze-thaw events in the winter of 2011/2012 results in the removal of all case hardened surfaces and core softened material from stones excavated in 2005 and 2007, effectively setting the stone surface properties back to its state on excavation. The presence of comparable stones excavated at different dates (2005, 2007 and 2013) provides an excellent opportunity to test this model.

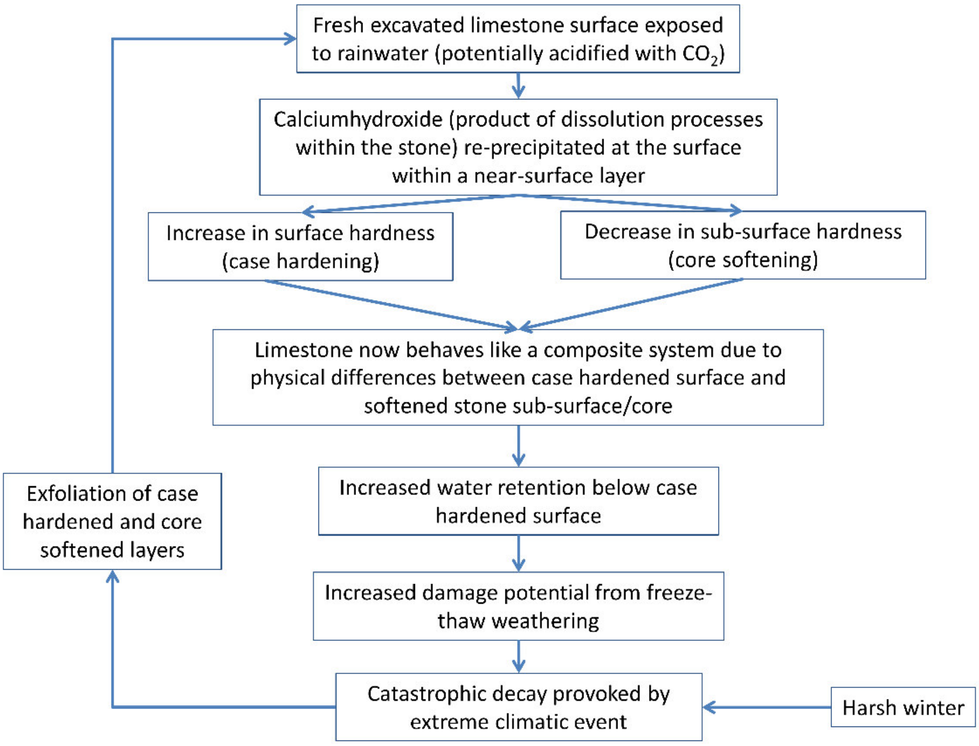


Figure 6.5 Flow chart of conceptual model of the interplay of case hardening and episodic freeze-thaw events.

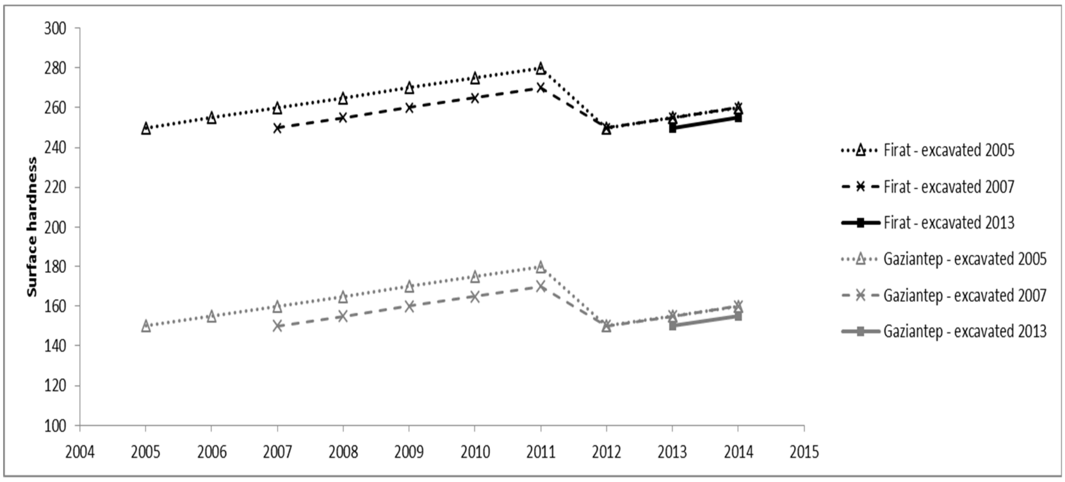


Figure 6.6 Conceptual model hypothesised interplay of case hardening and episodic freeze-thaw weathering producing catastrophic deterioration.

As a significant archaeological site the taking of samples is generally forbidden, and non-destructive and field-based techniques are preferable. As proxies for the degree of case hardening/weathering we used the Equotip device to evaluate surface hardness, and the Karsten tube method to investigate water uptake capacity. The methods are modified (extended application and data analysis) in order to gain additional information about porosity characteristics in the near-surface zone. As case hardening develops we assume that the surface hardness will increase, whilst the water uptake capacity will decrease.

6.1.2 Dülük Baba Tepesi

Dülük Baba Tepesi is located about 10 km north-west of Gaziantep in South Turkey (south eastern Anatolia) at 1100 m above sea level (Figure 6.7). The remains are mostly stone-built structures like foundations, wall structures (without roofing), floors and staircases deriving from Hellenistic to Medieval periods. Once excavated, the remains are usually covered with water permeable textiles to provide some protection from the elements. The investigated area of this study is located on the south west central plateau of the site. Three trenches from different excavation years were investigated (2005, 2007 and 2013) as shown in Figure 6.8. Together, the trenches reveal a coherent foundation structure built of limestone blocks and are believed to have formed the foundation for a temple complex. The two stone varieties were used seemingly at random, with the Gaziantep formation more frequently used. The blocks are arranged in one to three layers above the ground (although the structure might continue underneath) and vary in size with an average size of roughly 90 x 50 x 40 cm.

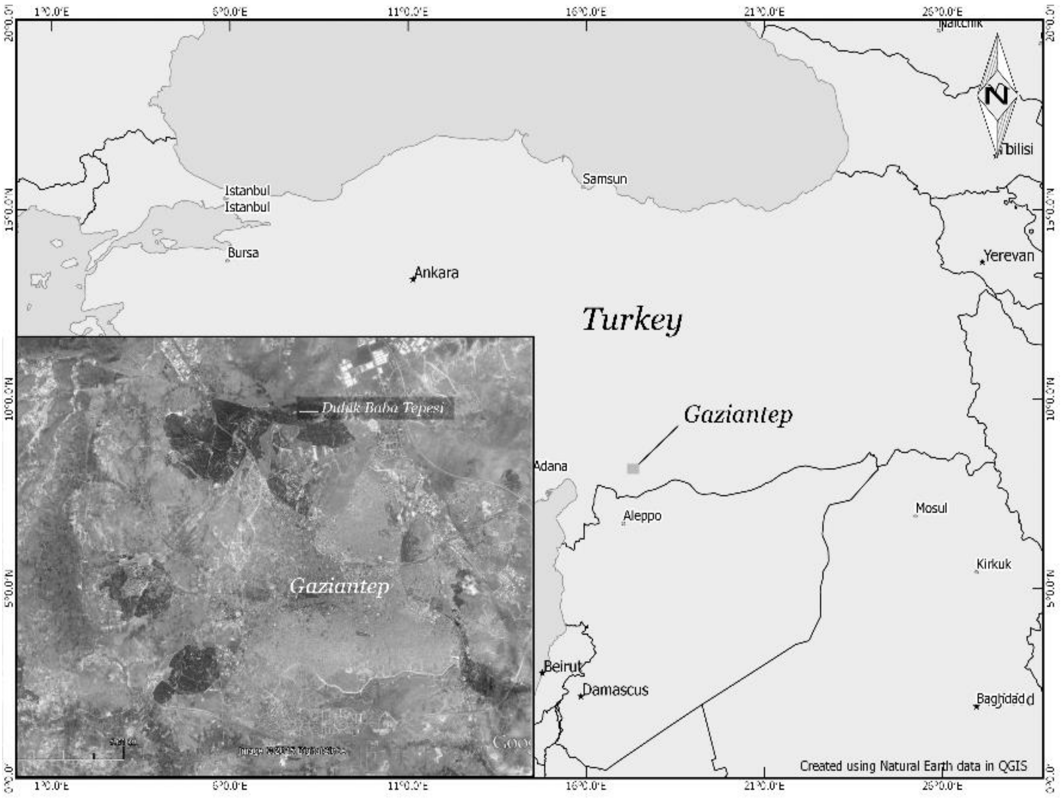


Figure 6.7 Location of Gaziantep in Turkey (arrow), archaeological excavation site is 10 km away from Gaziantep city to the north at 37°07’38.96” N 37°20’44.15” E (Photo© Google Earth 06/2015)

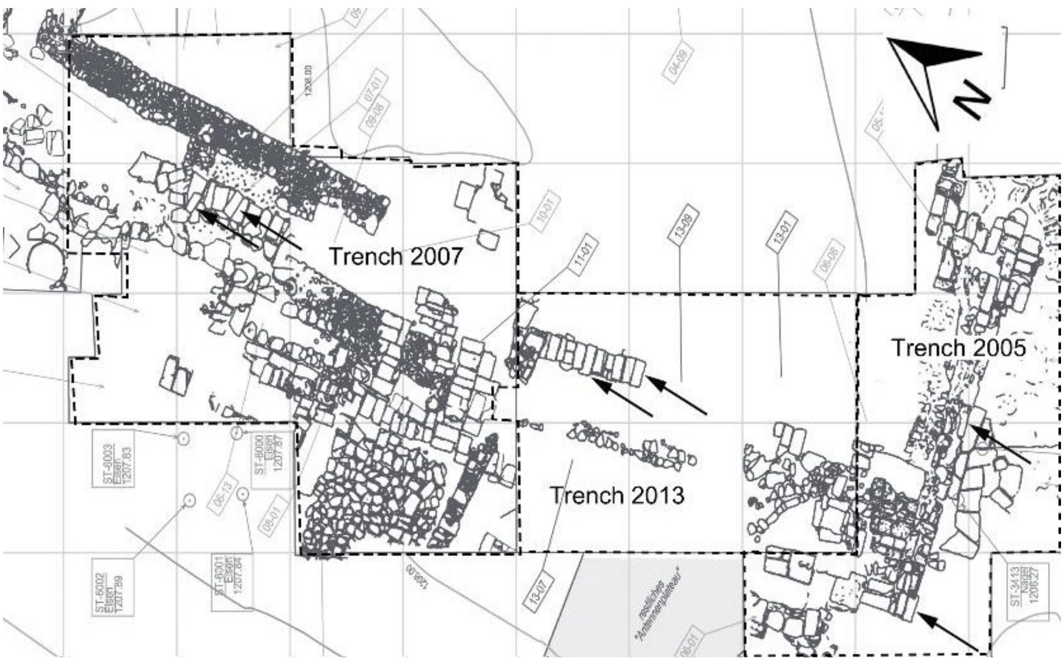


Figure 6.8 Floor plan detail, excavated Roman-Hellenistic structures of investigated trenches from 2005, 2007 and 2013. Investigated sample areas for this study indicated with black arrows, dashed lines mark the borders between the trenches. (one grid square = 5 m).

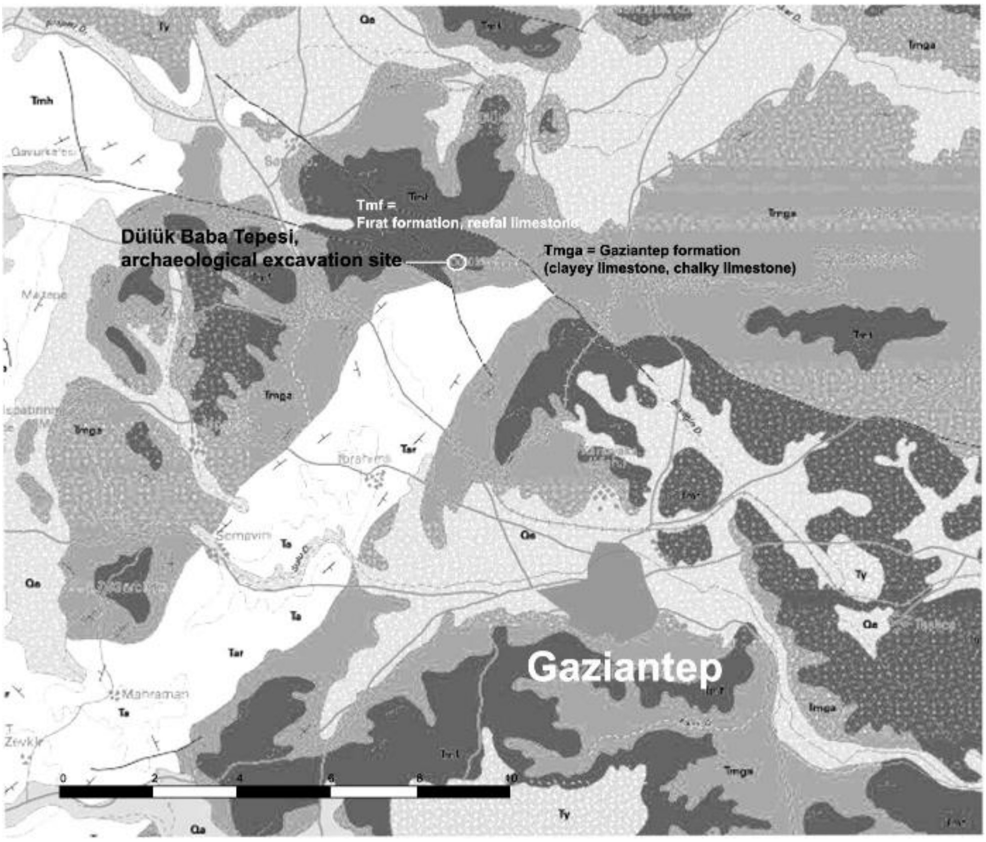


Figure 6.9 Geology map matched with Google earth map to investigate the most probable ancient building stone sources. Note that Duluk Baba Tepesi (red circle) is located at the intersection of Gaziantep (tmga) and Fırat Formation (tmf) (General Directorate of Mineral Research and Exploration, Turkey 1997)

A common practice during the Hellenistic and Roman period was to source building stone locally (Adam, 1999). In order to establish the exact nature of the limestone used in this section of the remains, the stones were observed in detail and compared with outcrops in six adjacent historic quarries whose locations were mapped using a GPS device (Garmin Dakota 20). The GPS data were then matched with the local geological map (General Directorate of Mineral Research and Exploration, Turkey 1997) to identify which limestone types had been used at the site (Figure 6.9). Two main types of limestone have been identified. The first, the Fırat formation (Tmf), Lower to Middle Miocene age, which is a chalky;

cream-grey coloured, hard and brittle reefbank type limestone (Dagistan & Simsek, 2005; Kaymakci et al., 2010; Robertson et al., 2015). Türkkan (2011) describes the weathered surface of Fırat formation as dark yellow-reddish and the freshly broken surface as beige. Visual inspection of the Hellenistic-Roman structures on-site confirms this description. However, the majority of the stones in the structures appear grey-beige in colour rather than reddish. These are likely to be cut from the second stone type, the Gaziantep Formation (Tmga), Upper Eocene age, a white-grey coloured, argillaceous, heterogeneous limestone with cherty intervals and cherty nodules (Coskun, 2000; Çanakci et al., 2007; Kaymakci, 2010).

Table 6.1 shows an overview of the existing research on the lithology and index properties of the two stone varieties. The dominant mineral in both varieties of limestone is $CaCO_3$ with traces of SiO_2, MgO, Al_2O_3 5 Fe_2O_3, and $MgCO_3$ (Baykasoglu et al., 2008; Türkkan, 2011). The main differences are compressive strength and water absorption, with the Fırat formation being the harder stone (72 MPa cf 10-25 MPa commonly recorded for Gaziantep) with lower water uptake characteristics (0.8% cf 11-24% recorded for Gaziantep). Both varieties of limestone are described as heterogeneous and, especially the Gaziantep formation, as soft limestone (Çanakci et al., 2007; Baykasoglu et al., 2008), which renders them susceptible to weathering (e.g. Goudie, 1999; Benavente et al., 2008). Moreover, May (1998) claims for limestone frost decay to be the more severe mechanism compared to dissolution processes. Additionally Brimblecombe (2010) reports enhanced damage potential when precipitation and frost occur in sequence. This is especially relevant in view of the immense strength reduction recorded for saturated Gaziantep limestone (between 44% and 55% according to Çanakci, 2007).

Table 6.1 Existing research on lithology and index properties of limestone Gaziantep and Fırat Formation. Modified after 1Kaymakci, 2010; 2Robertson et al., 2015; 3Dagistan, 2005; 4Coskun, 2000; 5Çanakci et al., 2007; 6Türkkan, 2011; 7Baykasoglu, 2008; 8Çanakci, 2007; 9Özvan et al., 2010; (*Karabakir, **Hamdi Kutlar (investigated collapsed caves in Gaziantep (Çanakci, 2007)).

Index test	Gaziantep formation (Tmga)	Fırat formation (Tmf)
Lithology	"Limestone with cherty intervals and cherty nodules"1; "Chalky"2 ; "argillaceous limestones, white, grey"4; "heterogeneous rock"5; "contains large gravel particles [...] crystalline silica"5	"Chalky"1; "cream-grey colored, hard and brittle reefbank type limestones"3 "weathered surfaces are dark yellow- reddish , hard, medium - weak strength, freshly broken surface beige"6
Mineralogy	97% $CaCO_3$, rest: SiO_2, MgO, Al_2O_3 7	96.46% $CaCO_3$, 0.28 SiO_2, 0.08 Fe_2O_3, 1.48 $MgCO_3$, Al_2O_3, rest 1.76
Dry unit weight (kN/m³)	16.76*8, 16.995, 17.37 , 18.64**8, 19.17, 23.215	26.86
Saturated unit weight (kN/m³)	20.27, 20.6**8, 20.79*8	-
Water absorption by weight (%)	11**8, 137, 187, 24*8	0.86
Compressive strength, dry (MPa)	10.2**8, 10.77, 25.51*8, 25-689	72.126
Compressive strength, saturated (MPa)	5.36**8, 11.53*8	-
Tensile strength, dry (MPa)	2.41**8, 3.12*8, 3.87	-
Tensile strength, saturated (MPa)	0.31**8, 0.65*8	-
UPV, dry (m/s)	2656**8, 26377, 2906*8 , 33807	-
Modulus of elasticity, dry (GPa)	4.45**8, 11.3*8	-

Porosity (%)	-	1.76, 103
Schmidt Hammer	-	50.56
Thermal conductivity	0.9264-2.5158 W/mK5	-

6.1.3 Materials and methods

Climatic data

Weather records covering the period 1984 to 2013 were obtained for a nearby meteorological station in Gaziantep (station ID 17600; extracted from the Met Office Integrated Data Archive System (MIDAS) Global Weather Observation (GL Table) data). Of particular interest for the present study are data for the winter months (October to April) on daily maximum, minimum and average temperature (6-8 measurements per day provided), daily and monthly precipitation, the number of frost days, and the number of individual freeze-thaw cycles. A freeze-thaw cycle is defined by a drop of temperature below 0°C followed by an increase of temperature above 0°C (Grossi et al., 2007). This study also calculated a modified version of the Wet-Frost Index, as developed by The Noah's Ark Project (Sabbioni et al., 2010). It was found that averaging the temperature of the given 6 to 8 measurements per day (provided by the weather records) missed occurring frost events, which are expected to have a weathering effect as well. Therefore, the Wet-Frost Index as used in this study is defined by the number of rainy days with P > 2mm and a maximum temperature >0° immediately followed by days with a minimum temperature -1°C, instead of the suggested mean temperature of -1°C for the following day as suggested by the original Wet-Frost-Index.

Mapping deterioration

Visual inspection and comparison with images from an image archive from earlier excavation campaigns was used to evaluate decay patterns across the Fırat and Gaziantep Formation limestones excavated within the three trenches in 2005, 2007 and 2013.

Surface hardness testing

An Equotip Piccolo 2 with DL probe (referred to as Equotip in this paper) was used in-situ to determine the surface hardness of representative sample sections of Fırat and Gaziantep Formation limestone from trenches excavated in 2005, 2007 and 2013 in order to evaluate any trends in case hardening following excavation. Given the conceptual model presented in Figure 6.6 it was expected that longer exposure times would lead to more intense case hardening and higher surface hardness values until episodic catastrophic deterioration occurred (in this case, likely to have occurred after the winter of 2011/2012) when lower surface hardness values would be found. For comparability reasons each sample area was tested on the south facing side, which was the only side accessible for all trenches.

The Equotip device measures the difference between impact and rebound velocity of a (small) hard metal impact body traveling in a probe and propelled by spring force against the surface (Proceq© SA, 2010). Hardness data were recorded on the Leeb hardness scale (1 to 999). Surface hardness data for this study was collected in 2014 using both the single impact method (SIM) (commonly used in the field e.g. Viles et al., 2011; Alberti et al., 2013; Coombes et al., 2013; Hansen et al., 2013) and the repeated impact method (RIM) (Aoki & Matsukura, 2008; Yilmaz, 2013).

Previous research has indicated that, for the SIM-method, at least 45 individual readings are required to gain a representative sample on heterogeneous stone surfaces (Viles et al., 2011; Wilhelm et al., 2016a). In this study Equotip was randomly applied with the SIM-method at unique points distributed over the stone surface of the sampling areas in the three trenches giving 120 readings per exposure group. For further data analysis the median was used and results are expressed as $HLDL_{S.med}$. In 2013 no Fırat Formation limestone blocks were excavated, thus no data was collected. For the RIM-method, 20 measurements were taken on exactly the same spot, using six sample spots for each exposure group and results are expressed as $HLDL_{R.med}$. For dense stones RIM-method values have been demonstrated to show an initial increase after only c. 10 readings (Yilmaz, 2013). However, given the soft stone at Dülük Baba Tepesi this study followed the approach of Aoki and collected 20 RIM values (Aoki & Matsukura, 2008). For further data analysis $HLDL_{R.med}$ was calculated from the median of the highest value in each of the six RIM-method datasets per area.

The SIM-method gives a simple measure of surface hardness, whilst the RIM-method reflects the plastic deformation potential of a stone, thus relates to its porosity. Yilmaz (2013) combined SIM- and RIM-method results to calculate the deformation ratio (DR) and Hybrid Dynamic Hardness (HDH) for best correlation to unconfined compressive strength of carbonate stone ($R^2 \cong 0.77$). A high DR value is related to a low compaction ratio, correlates to high UCS and to a low porosity. Therefore, for this study we expected despite the soft porous limestone the DR to be high due to surface hardening.

To evaluate the Equotip data in this study, modified methods of calculating DR and HDH were used, because we expected non normally distributed data as a result of heterogeneous stone types having been exposed to outdoor climate (Palmer, 2008; Tiryaki, 2008). Using robust, non-parametric statistical methods allowed us to preserve outliers and avoided the need for data transformation. Thus, in contrast to Yilmaz (2013) this study calculates the deformation ratio (DR_{robust}) and HDH_{robust} from the median of the single ($HLDL_{S.med}$) and the repeated hardness measurements ($HLDL_{R.med}$) (equations 17 and 18).

$$DR_R = HLDL_{S.med} / HLDL_{R.med} \qquad \text{(Equation 17)}$$

Hybrid dynamic hardness is calculated as follows:

$$HDH_R = DR_R \; x \; HLDL_{S.med} = (HLDL_{S.med})^2 / \; HLDL_{R.med} \qquad \text{(Equation 18)}$$

Capillary water uptake

In order to characterise the water uptake capabilities of the Fırat and Gaziantep Formation limestones from the three different excavation trenches on-site, Karsten tubes were applied vertically to the stone surface. The Karsten tube is a glass cylinder (inner diameter = 26mm) open on one side, which is applied to the stone surface using putty (Plastic Fermit) (Auras et al., 2011). Connected with the glass cylinder is a tube with indicated gradation. It is filled once with distilled water (4 ml) and subsequently the time measured (stopwatch) it takes for the water to penetrate the stone through the open side of the glass cylinder.

Water uptake was recorded in 0.1 ml steps and the time noted accordingly (t). Following recommendations of D'ham et al. (2011) and the BS EN Standard

16302:2013 a minimum of 7 data pairs (ml and t) was collected or the application stopped after 60 minutes. A varying number of measurements per exposure period were taken as only later during data evaluation it became clear, that two stone formations had been tested (instead of two varieties of the same stone type which was the initial assumption). Following the conceptual model presented in 6.6, we assumed that longer exposure times would lead to more intense case hardening and lower water uptake rates.

Water uptake values (ml) were plotted against time (min : sec) and the overall rate for each graph determined. It became obvious that the trends in water uptake over time for the individual blocks were not linear. Thus, to further investigate changes in water uptake rates, segmented linear regression models were fitted iteratively to detect break-points using the package 'segmented' in RStudio (Muggeo, 2003; Crawley, 2005; Muggeo, 2008) to show how the water uptake rate would change (either increase or decrease). This data analysis gives insight into porosity conditions of the stone subsurface. So for example, if the initial water uptake rate was slower, compared to the rate followed after the detected breakpoint, this would indicate that the subsurface porosity is higher compared to the surface. This information is valuable in that it reflects on surface crust forming and subsurface softening behaviour.

Statistical evaluation

RStudio (version 0.97.551) was used for statistical analysis of both surface hardness and water uptake values. The Shapiro-Wilk test showed non-normal distribution for some datasets, as expected for weathered limestone (Mosch and

Siegesmund, 2007; Palmer, 2008). As explained in section 3.2 this study avoided data transformation and outlier removal through using robust statistics (non-parametric) (Erceg-Hurn and Mirosevich, 2008; Filzmoser and Todorov, 2013). The median and the median absolute deviation (MAD) were calculated as measures for central tendency and variance respectively and data displayed using boxplots. The Mann-Whitney U test was applied to determine any significant differences in surface hardness between the two stone types and the different exposure periods (2005, 2007 and 2013).

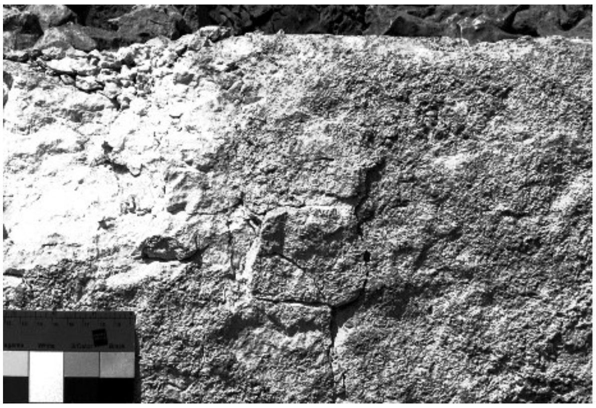
6.1.4 Results and discussion

Deterioration patterns

On excavation, blocks from both Fırat and Gaziantep Formation limestones already show evidence of deterioration (including cracks and rounded edges) as shown in Figure 6.10 to 12. As it is assumed that the investigated blocks are part of the temple foundations, it is likely that they have been affected by chemical weathering in soil during several centuries of burial. After excavation the structures are exposed to outdoor environment for extended periods of time covered only with a permeable textile and start to develop a range of deterioration patterns including spalling, disintegration, crumbling, chalking and fragmentation (Figure 6.10, 6.11 and 6.12). Although no bedding is visible to the naked eye stone show partly a preferred direction of disintegration with more serious deterioration results when orientated vertically (Figure 6.1).

Figure 6.10 Blocks excavated in 2007 showing rounded edges (Photo© Engelbert Winter 2007).

Figure 6.11 Freshly excavated limestone block in 2014 showing macro crack (Photo© Engelbert Winter 2014).



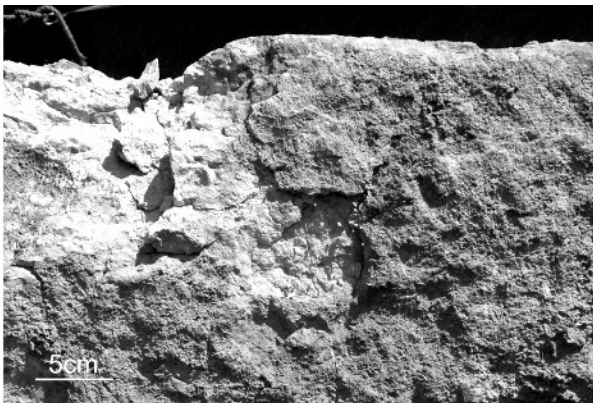


Figure 6.12 Detail Gaziantep formation block in trench 0701, examples for material loss through spalling, left in 2013 and right in 2014.

Climatic trends

Table 6.2 summarises the climatic parameters for the cold periods (2005-2013) relevant for this study with precipitation, number of frost days, the number of individual freeze-thaw cycles and Wet-Frost index. Figure 6.13 illustrates the number of frost days, and the average monthly winter rainfall for the Gaziantep area from 1984-2014. Wetter and colder winters with more frequent freeze-thaw episodes are likely to cause higher rates of frost weathering on vulnerable stones. Over the 30 year period of record, a long term mean of 41.4 frost days is recorded with a standard deviation (SD) of 17.8 days. The winter of 2011/2012 had 63 frost days, which is the maximum for the period of interest in this study (2005

- 2013). However, the cold season of 2007/2008 shows a similar high number of 60 frost days, though no catastrophic decay was reported after that winter. Figure 6.13 shows that whilst both years had similar numbers of freeze-thaw cycles (126 vs 128) differences in winter rainfall may explain the situation. The average of the total precipitation for the cold periods from October to April (of the following year) over the 30 year period 1984 to 2013 was 516.75 mm with a SD of 120.65 mm. As Figure 6.13 illustrates, the winter of 2007/2008 was characterised by a relatively low monthly winter rainfall average (313.44 mm) in comparison with the wetter conditions in 2011/2012 (687.50 mm). The evaluation of the Wet-frost Index reveals 3 and 7 occurrences for 2007/2008 and 2011/2011 respectively, also illustrating the unusually damaging nature of the 2011/2012 winter.

Table 6.2 Overview climatic parameters for cold periods relevant for this study. 2011/2012 marked bold is the period after which catastrophic decay was observed.

Cold periods	Precipitation (mm)	Days of Frost	Freeze-thaw cycles[1]	Wet-Frost Index
Oct 2005 - April 2006	518.1	34	24	4
Oct 2006 - April 2007	494.9	63	56	4
Oct 2007 - April 2008	313.4	60	63	3
Oct 2008 - April 2009	390.4	34	38	1
Oct 2009 - April 2010	516.3	11	13	2
Oct 2010 - April 2011	461.2	21	27	0
Oct 2011 - April 2012	687.5	63	64	7
Oct 2012 - April 2013	696.0	25	27	1

1 (drop of temperature below <0 followed by rise of temperature above >0); 2 modified after Sabbioni et al. 2010 with P > 2mm & max temp. >0°C, followed by days with min temp. -1°C.

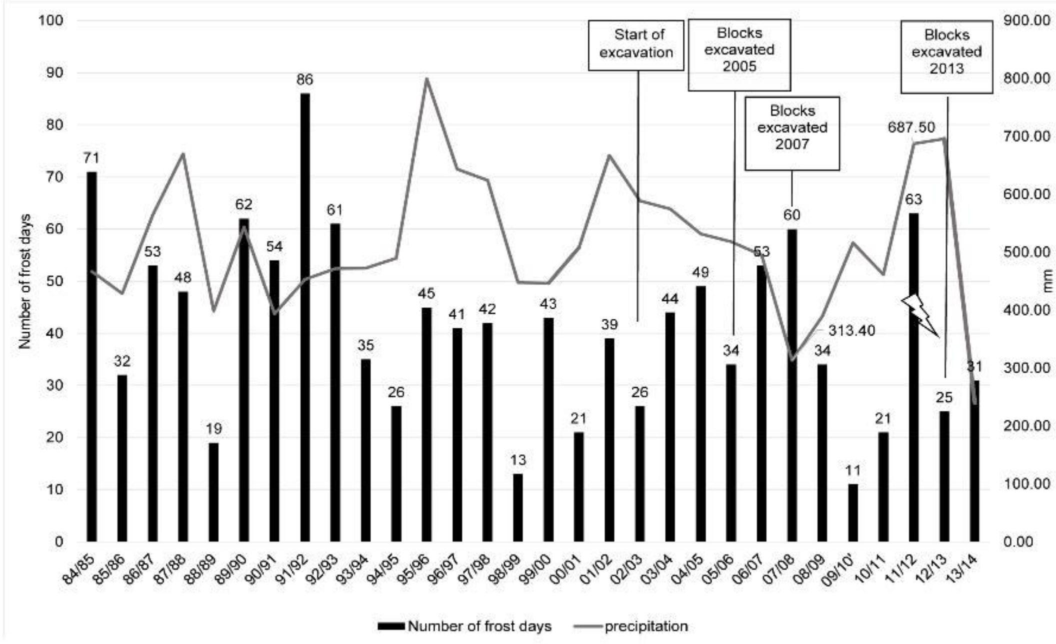


Figure 6.13 Number of frost days (bars) and average total precipitation [mm] (line) of the cold season (October to April) (Oguzeli Airport (Gaziantep, Turkey). Relevant cold season for this study is marked with a flash sign.

Surface hardness and capillary water uptake

Table 6.3 summarises the surface hardness data collected for this study. The boxplot graph in Figure 6.14 illustrates the distribution of the single surface hardness (SIM) values for the two stone types from the three exposure periods. As expected for Firat formation, with noticeably higher compressive strength when un-weathered (Table 6.1), blocks excavated in 2005 and 2007 are significantly harder than those from the Gaziantep formation of the same years (2005: p = <0.000, Mann-Whitney U = 1021.5; 2007: p = <0.000, Mann-Whitney U = 1049) (Table 6.4). Furthermore, for both the Fırat and Gaziantep formations, blocks excavated in 2005 show significantly higher surface hardness than those excavated in 2007, as predicted by the model (Figure 6.6) where a longer exposure time leads to more intense case hardening (instead decrease of surface hardness). However, Gaziantep blocks excavated in 2005 do not show any significant difference in surface hardness in comparison to Gaziantep blocks excavated in 2013 as shown in table 6.3. Therefore, the linear model in Figure 6.6 fails to explain the lack of significant difference in surface hardness between Gaziantep blocks excavated in 2005 and 2013 (the conceptual model predicts that block from 2005 should be much harder).

Table 6.3 Summary of the index data collected in 2014 on limestone blocks for this study.

Formation	Ex-cavation year	HLDLS.med	HLDLS.MAD	HLDLR.med	DRR	HDHrobust
Fırat	2005	403.0	93.4	590.0	0.68	275.27
Fırat	2007	318.5	86.0	682.5	0.47	148.63
Fırat	2013	n.a.	n.a.	n.a.	n.a.	n.a.
Gaziantep	2005	316.5	54.1	654.0	0.48	153.17
Gaziantep	2007	269.0	80.1	623.5	0.43	116.06
Gaziantep	2013	317.5	97.9	638.5	0.50	157.88

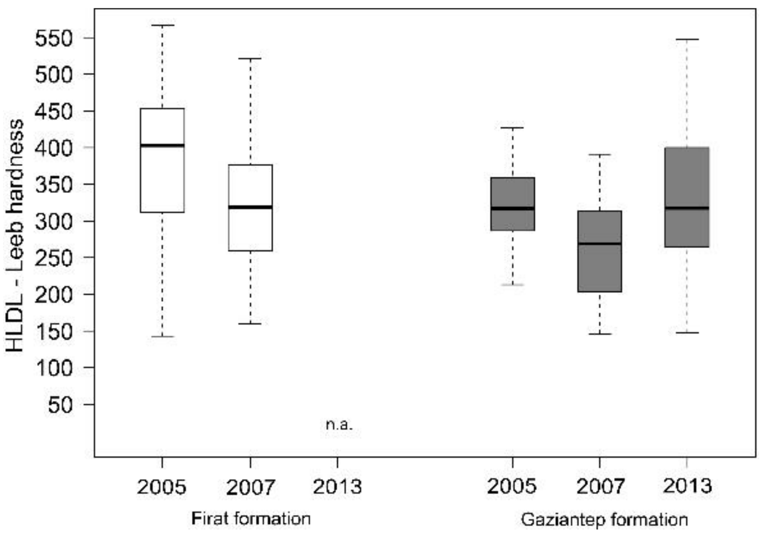


Figure 6.14 Comparison of surface hardness values (SIM) of Gaziantep and Fırat limestone over different exposure periods (no values for Fırat 2013 were collected). The years mark the time of exaction.

Table 6.4 Mann-Whitney U test for significant differences in surface hardness depending on exposure time and between the two stone formations (Firat and Gaziantep (Gaz) Formation; excavated in 2005, 2007, 2013). Significant p-values are bold.

		p-value	U
Fırat 2005	Fırat 2007	0.001	1144
Gaz 2005	Gaz 2007	0.000	897
Gaz 2005	Gaz 2013	0.6880	3408.5
Gaz 2007	Gaz 2013	0.000	1976.5
Fırat 2005	Gaz 2005	0.000	1021.5
Fırat 2007	Gaz 2007	0.000	1049
Fırat 2005	Gaz 2007	0.000	565.5
Fırat 2007	Gaz 2005	0.9580	1789.5

Water uptake data (Karsten tube) are presented in Table 6.5. There is a trend for the Firat formation to show lower water uptake compared to the Gaziantep formation. This would be expected as the two stone varieties already show noticeable differences in water absorption when un-weathered (Table 6.1). This further supports the surface hardness observations, where deformation ratio (DR) is the highest for Firat blocks excavated in 2005 and correlates well with the low water uptake found in those blocks (Table 6.5). However, it can be seen, that the overall rates do not reflect on the differentiated, heterogeneous water uptake behaviour in the subsurface zones presented by the breakpoint analysis and inferred rates. Therefore, for this study the evaluation of the spatial character of water uptake behaviour is of more interest than the actual amount of water uptake.

Table 6.5 Overview of recorded overall and segmented water uptake rates (Karsten tube) for Firat and Gaziantep (Gaz) Formation. Breakpoints (Bp) and respective time (min:sec) at which rate changed are shown. Rates, which increased again after a breakpoint are of special interest and marked bold.

Formation Excavation year	Total rate	Breakpoints							
		Start time	start rate	1st Bp time	2nd rate	2nd Bp time	3rd rate2	3rd Bp time	4th rate
Firat 2005	0.02	00:00	0.04	05:00	0.03	15:00	0.02	-	-
Firat 2005	0.00	00:00	0.00	-	-	-	-	-	-
Firat 2007	1.05	00:00	4.99	00:05	1.81	00:45	0.62	-	-
Firat 2007	0.02	00:00	0.03	04:00	0.01	-	-	-	-
Gaz 2005	0.47	00:00	3.18	00:10	1.94	00:45	1.38	-	-
Gaz 2005	1.98	00:00	3.17	00:10	1.94	00:45	1.39	-	-
Gaz 2007	0.85	00:00	0.04	05:00	0.02	15:00	0.01	35:00	0.00
Gaz 2007	0.39	00:00	2.32	00:20	0.35	01:00	0.63	02:00	0.26
Gaz 2007	0.33	00:00	1.80	00:15	0.57	00:50	0.23	03:40	0.20
Gaz 2007	0.70	00:00	1.40	00:50	0.21	01:00	0.91	01:45	0.37
Gaz 2013	6.84	00:00	13.33	00:05	6.13	-	-	-	-
Gaz 2013	2.38	00:00	2.50	00:35	0.67	-	-	-	-
Gaz 2013	0.19	00:00	0.59	00:30	1.00	03:00	1.03	-	-
Gaz 2013	0.69	00:00	1.10	00:40	0.38	03:00	4.63	-	-
Gaz 2013	0.21	00:00	0.67	00:20	0.22	04:00	0.16	-	-
Gaz 2013	0.19	00:00	1.25	00:15	0.23	02:00	0.17	-	-

For the majority of the tested blocks one or more breakpoints were detected. Only Gaziantep blocks excavated in 2007 show three breakpoints. Usually the water uptake rate declines further after each break-point. Yet, of special interest are rates, which after initial decrease increase again as they indicate a higher porosity in the subsurface zone. It can be seen that some of the Gaziantep formation blocks excavated in 2007 and 2013 show such reoccurring increase. Interestingly the increase of rate for the blocks from 2013 appears after 3 minutes compared to 1 minute for 2007 blocks, showing that the zone of increased porosity is closer to the surface for the latter. Accordingly those blocks also show the lowest DR with 0.43. With the highest number of breakpoints and varying water-uptake rates this renders their subsurface zone as most heterogeneous. The thickness and heterogeneity of this zone is suspected to play a key role in causing catastrophic decay after severe winters and further research on that aspect is highly desirable. In this instance relatively small exposure time differences (2-5 years) seem to have distinct effects on sub-surface porosity and case hardening development, where a longer exposition might be beneficial as for Gaziantep blocks excavated in 2005 display a more homogeneous water uptake behaviour and higher DR. Both the surface hardness and Karsten tube data suggest that the case hardening trend had been interrupted by the cold period of 2011/2012 and resulted in material loss, which 'reset the clock' of the limestone resulting in new even softer surfaces (former sub-surfaces of varying depths).

Synergistic effect of climate and probability of stone decay

The in situ, non-destructive test results of this study suggest that catastrophic decay at Dülük Baba Tepesi is a product of a synergistic effect between crust forming behaviour (case hardening) and extreme weather events which cause intense episodes of freeze-thaw weathering. This finding is consisting with Brimblecombe's (2010) statement of enhanced damage potential when precipitation and frost occur in sequence. The results show progressive case hardening with stones excavated in 2005 showing higher surface hardness compared to stones excavated in 2007. Further, the crust forming behaviour of the two stone varieties is different as for stones excavated in the same year the Fırat formation displays higher surface hardness and lower water uptake than the Gaziantep formation. The data supports the hypothesis of increasing surface hardness (and decreasing water uptake) with longer exposure to outdoor climate. In view of the catastrophic decay observed after the winter of 2011/2012, it is suggested that the extreme climatic conditions 'reset' the blocks excavated in 2005 and 2007. In this sense 'reset' describes the severe impact of combined high amount of precipitation and frost events causing sufficient stone surface (crust) material loss and exposure of softened stone core to the elements.

Modified model

This study found the stone weathering behaviour to be more complex than shown in the simplified model in Figure 6.6. Instead Figure 6.15 gives a more realistic picture of the sequence of events. Both non-linearity and thresholds are simulated as they are found to have affected the progression of surface hardness before and after the severe cold period in 2011/2012. It takes the different progression of deformation ratios (related to porosity) of the two stone varieties after excavation and after the severe climatic event into account. The model presented suggests a logarithmic (base 10) development of increase of surface hardness over the years 2005 to 2011 until the extreme climatic event. Further, the model multiplies surface hardness values with assumed deformation ratios (to simulate HDH), where freshly excavated Gaziantep limestone is assigned a lower DR of 0.7 for its natural lower strength and higher porosity compared to Firat limestone, which is ascribed a DR of 1 (see also Table 6.1). Depending on the exposure history of the stones and the resulting surface-to-depth heterogeneity (i.e. varying thickness of a sub-surface zone of increased porosity or hardened case), the model assumes that the DR decreases after a severe climatic event for Firat blocks excavated in 2005 to a DR of 0.5, for Firat blocks excavated in 2007 and Gaziantep blocks excavated in 2005 to a DR of 0.2 and finally for Gaziantep blocks excavated in 2007 to a DR of 0.1. Thus, the model shows that both stone variety and length of exposure may determine the magnitude of the impact of the severe climatic event. In this instance the harder stone (Firat) and the longest exposure period (9 years) are least affected compared to the softer stone (Gaziantep) and a shorter exposure time (7 years). These variations are thought to be key to understand complex interactions between post-excavation case hardening and extreme winter conditions which can lead to catastrophic deterioration. Such increased complexity of stone decay systems with exposure has been described by McCabe (2015), who finds both stone and micro-environmental climate with accumulative exposure history results in surface-to-depth heterogeneity at the stone block scale. Further data collection is needed to provide a fuller test of the refined conceptual model proposed in Figure 6.15, but the severity of the effect of the cold winter has been demonstrated to depend on both stone type and exposure time since excavation.

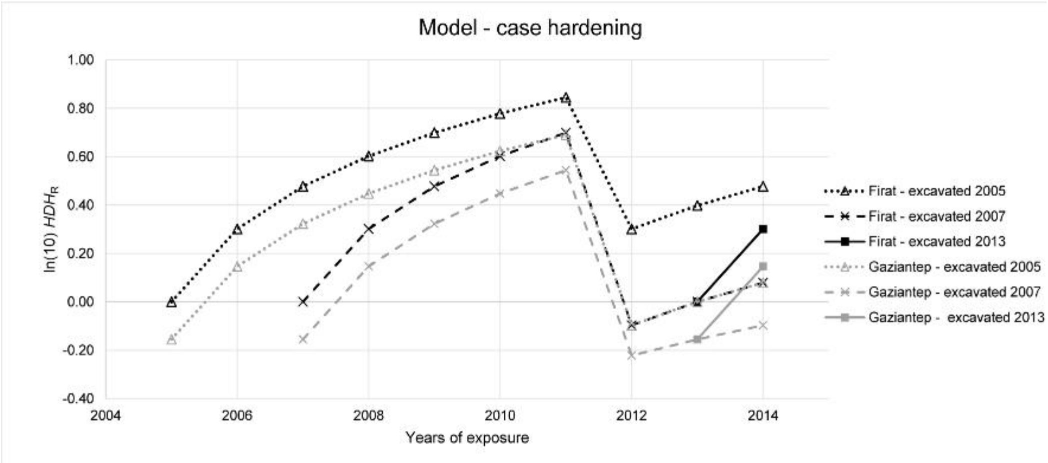


Figure 6.15 Conceptual model (based on the results of this study) of the interplay of case hardening and episodic freeze-thaw weathering producing catastrophic deterioration. Years of exposure on the x-axis and ln(10) of predicted robust hybrid dynamic surface hardness (considering deformation ratio i.e. reflects on stone porosity). At 2012 a severe climatic event is simulated affecting the progress of weathering behaviour.

6.1.5 Conclusion

The results of this study illustrate the importance of temporal sequences of case hardening processes in limestones and episodic harsh winter condition to catastrophic deterioration. The causes of the catastrophic limestone decay at the archaeological excavation site at Dülük Baba Tepesi in South Turkey were investigated in order to aid future excavation and conservation planning. This study reconstructed the interlinked history of limestone case hardening and deterioration from a time-series of stones excavated in three different recent years with climatic records. The combination of non-destructive modified surface hardness testing and water uptake analysis on-site gave insights into property changes of surface and subsurface zones of the limestone. A multi-layered system of zones with varying porosity acts like a composite system with different physical properties for surface, subsurface and core of the stone enhancing the system response to extreme climatic impact. Non-linear weathering behaviour was found for both stone types (Firat and Gaziantep formation) with different crust forming characteristics. In agreement with McCabe (2015) we found that accumulative exposure history results in 'surface-to-depth heterogeneity' at the stone block scale.

Of particular interest for this study were winter (October to April) climatic data on monthly precipitation, the number of frost days, the number of individual freeze-thaw cycles and the Wet-Frost Index. The main differences between the two cold periods of interest to this study (2007/2008 and 2011/2012) are the amount of precipitation, which was twice as high for 2011/2012 (687.50 mm vs 313.44 mm) and the occurrences of Wet-Frost events, which were twice as often for 2011/2012 (7 vs 3 Wet-Frost events). Therefore, we propose the extreme cold and wet conditions in 2011/2012 to have been the key trigger for catastrophic decay.

This study introduces a crust forming model in which case hardening develops and initially strengthens the stone against deterioration, but ultimately makes it more prone to environmental impact and eventually results in material loss. This process happens over years until a harsh winter (or other extreme event) causes catastrophic decay which ‘resets’ the system. Although the probability of an extreme cold period like in 2011/2012 might be low, the long term climatic dataset in Figure 6.13 illustrates that such events are likely over a 30-50 year period. The particularly dramatic response of the stones excavated in 2007 (especially the Gaziantep formation) to the harsh winter of 2011/2012 show that specific care should be taken to protect stones 3 to 4 years after excavation when they might be extremely sensitive to catastrophic deterioration.

We conclude that the Hellenistic-Roman structures are too vulnerable to be exposed to the prevalent environment without any further preservation measures. In terms of preventive conservation it is strongly recommended to prevent water ingress into the structures (further potential conservation measures might be reburial, roof construction, protective coping etc.) in order to limit frost damage and loss of significant cultural heritage.

The wider implications of this study are that weathering rates can be established using the inexpensive and easy to handle methods introduced and improved in this study. In case of the catastrophic decay at the archaeological excavation site establishing weathering rates at an early stage (directly after excavating) might have prevented catastrophic decay from happening. Crust forming processes would have been detected and appropriate preservation measures could have been undertaken (i.e. protect structures from water ingress).

Acknowledgments

K. Wilhelm is in receipt of an EPSRC studentship (funded via grants EP/P504287/1, EP/P505216/1 and EP/P505666/1) with additional support from Proceq.

We want to thank the archaeologists Dr Michael Blömer and Dilek Cabanoglu for providing invaluable insight in excavation history of the site. Dr Martin Coombes we are grateful for helpful comments on an earlier draft of this paper.

7. OVERALL THESIS DISCUSSION – ISSUES, IMPLICATIONS AND FUTURE RESEARCH – CONCLUSION

7.1. INTRODUCTION

The scientific realm of this thesis is heritage science with focus on low-cost, non-destructive methods for the diagnosis of limestone weathering behaviour and deterioration problems of stone built heritage *in situ*. The overall aim of this thesis was to improve selected economic, non-destructive methods for the diagnosis of deterioration and investigating weathering behaviour of stone built heritage *in situ* with implications for both conservation strategies and stone weathering research.

The study divided into three objectives with two main strands of investigation, one laboratory and the other field based. The tested methods in this thesis were chosen to serve a complex, yet very common situation in the field of immovable cultural heritage preservation and research:

- Rapid and progressive decay of immovable heritage with the need for urgent preservation action
- No sample taking allowed
- Mostly unknown weathering-stress and preservation history
- Heterogeneous weathering patterns on various scales are evident
- Financial and time constrains (reduced sampling size)

To address the challenges mentioned above the thesis divided into three objectives. The first objective focussed on selected non-destructive methods (low impact surface hardness testing, handheld electronic moisture meter testing and water uptake measurement), which were improved under laboratory conditions on fresh porous heritage limestone for their eventual *in situ* application. Objectives 2 and 3 then applied the improved methods to investigate limestone weathering behaviour *in situ* at architectural and archaeological heritage sites in the UK and Turkey.

This thesis followed the 'thesis by paper' approach. Each of the four published papers provides a specific discussion section (see chapter 4-6). Thus, this chapter links the findings and discussion of the individual papers into the broader scope of the thesis following the structure established in chapter 1. This is followed by a section which reviews answered research questions and main findings individually. Table 7.1 summarizes the contributions to the field of this thesis.

In order to preserve irreplaceable cultural heritage its weathering behaviour, which is determined by its initial properties and subsequent weathering stress history as well as current environmental conditions, needs to be understood and quantified accordingly. As stated before limestone weathering behaviour is complex due to a) inherent stone variability with respective heterogeneous behaviour (system response) becoming more complex over time with accumulated weathering-stress history and b) a dynamic environment where process domains on a range of scales may affect each other and the stone in ordered and complex ways. To account for complexity the thesis approach addresses the "*The whole is greater than the sum of its parts- dilemma*".

Table 7.1 Overview how this study advanced the field and filled gaps in the field of stone weathering research

Approach/method	Gap / issues	Solution	Implications
Drawback converted into an advantage/useful application	Moisture meters affected by salt contamination in stone	Utilize moisture meters as salt detectors (Paper 2)	Baseline for future research. Method improvement
Combination of methods	Unknown potential for selected methods to be enhanced by combination	Low impact surface hardness and Karsten tube water uptake have been combined to detect short-term crust forming behaviour at a vulnerable archaeological excavation site in Turkey (Paper 4)	Base for 'scientific toolkit of low-cost methods' (following a suggestion by Meneely et al.(2009)) Implications for preventive preservation practices. Method improvement. Cross-disciplinary exchange as inevitable part of (preventive) conservation.
Surface/subsurface information gain	Unknown potential for the selected methods to gain subsurface information	Two application methods of low impact surface hardness (single and repeated impact measurement) were combined to the Hybrid Dynamic Hardness based on Yilmaz' (2013) approach but modified with modern statistics to adapt the application to porous stone (Paper 1 and Paper 4)	Extended application of the improved method.
		The application of breakpoint analysis to Karsten tube water uptake allowed for stone subsurface heterogeneity to be displayed (Paper 4)	Potential to extend investigation of whole stone weathering behaviour trajectory. Method improvement. Better understanding of stone weathering behaviour
Temporal scale of time-series	Majority of studies investigated on short-term timescales	Utilizing both cemeteries and archaeological excavation site as outdoor exposition laboratory and archive with	Potential to extend investigation of whole stone weathering behaviour trajectory. Method improvement.

Approach/method	Gap / issues	Solution	Implications
		different datum points of investigated stones (Paper 3 and Paper 4)	Better understanding of stone weathering behaviour
Spatial scale of *in situ* investigations	Present recommendation is to investigate stone weathering behaviour *in situ* at block scale (e.g. Meneely et al., 2009; McCabe et al.,2015)	The study published in Paper 3 shows that in cases it is necessary to investigate at 'sub-block' scale to capture the whole spatial variety	Baseline for future research. Potential to extend investigation of whole stone weathering behaviour trajectory. Better understanding of stone weathering behaviour
Extent of capturing stone weathering behaviour trajectory	The majority of limestone weathering behaviour studies focused on erosion, which is understood to be the final step in a series of decay mechanisms	Investigating decay mechanisms preceding erosion, such as surface hardening (redeposition of solutional products) or softening (induced by both climate and biological activity). All three improved methods in this thesis are suitable for this more holistic approach (Paper 3 and Paper 4)	Baseline for future research. Method improvement. Better understanding of stone weathering behaviour.
Character of stone weathering behaviour	Though the majority of studies suggests non-linear weathering behaviour for limestone heritage exposed to the outdoors (e.g. Smith and Viles, 2006; Smith and Gomez-Heras et al., 2010) deeper insights for the individual causes of non-linearity are needed.	The study published in Paper 3 explains, that the temporal scale of investigation will determine whether stone weathering behaviour is characterised as linear or non-linear.	Baseline for future research. Method improvement. Better understanding of stone weathering behaviour.
		The study published in Paper 4 shows, that even on short-term weathering trajectories non-linear weathering behaviour can occur and thus, the type of weathering (here case hardening) is also determining linear or non-linear behaviour characteristics.	Baseline for future research. Better understanding of stone weathering behaviour.

Single stone parameters investigated with small samples on a short-term scale simulated under laboratory conditions often fail to reflect true stone weathering behaviour *in situ* as found by a range of former studies (e.g. Bell, 1993; McGreevy and Smith, 1982; Trudgill and Viles, 1998; Warke et al., 2003; Ingham, 2005; Moroni and Pitzurra, 2008; Smith et al., 2008; Sass and Viles, 2010; Smith et al., 2010; Inkpen et al., 2012a). Furthermore, we learnt in the introduction that the majority of limestone weathering behaviour studies *in situ* had an overwhelming focus on the erosion (mass loss) of stone. Yet, erosion is understood to be the final step in a series of decay mechanisms preceding this mass loss stage, such as surface hardening (redeposition of solutional products as shown in the study published in Paper 4 (Chapter 6) with case hardening processes of archaeological stones) or softening (induced by both climate and biological activity. Compare Paper 3 (Chapter 5)). These preceding decay mechanisms lead to stone surface property changes including increased porosity and the formation of superficial layers (Pope et al., 2002; Hoke and Turcotte, 2004; Smith and Viles 2006; Inkpen et al., 2012b; McIlroy de la Rosa et al., 2014). Due to the complexity of stone weathering behaviour it has been shown that to gain a holistic understanding its entire weathering trajectory needs to be quantified. Therefore, this study proposes to investigate stone weathering behaviour ideally (frequently) under real world conditions on real heritage considering a range of time scales and past environmental conditions.

This thesis is committed to the key principles of built heritage conservation i.e. to preserve as much original fabric as possible, whilst investigating the whole heritage structure as a system with its responses to the respective environments and impacts over time (opposed to single sampling or exposed samples). Consequently, the focus was on portable non-destructive techniques which can be applied *in situ*. For *in situ* investigation a variety of methods is available ranging from destructive to non-destructive and sophisticated and expensive to more simple and economical. Yet, as discussed in the introduction the sophisticated methods are costly and require special apparatus and expertise. This has implication on i) that they cannot easily be used by conservators and ii) when the entire weathering trajectory should be quantified a certain frequency of applying methods is necessary, where often time and finances determine the frequency. Therefore, the overall aim of this thesis was to improve selected economic, non-destructive methods for the diagnosis of deterioration and investigating weathering behaviour of stone built heritage *in situ* with implications for both conservation strategies and stone weathering research. The set of methods focused upon (surface hardness testing, moisture measurement and water uptake) is a small selection of a bigger pool of available low-cost, portable methods (compare Table 1.1). Similar to the 'scientific toolkit' recommended by Meneely et al. (2009) for more sophisticated methods (e.g. 3D laser scan, ground penetrating radar etc. see section 1.4.3.) the methods evaluated in this thesis are seen as a contribution to a potential 'scientific toolkit of low-cost methods' which could be complemented with other methods like ultrasound velocity measurements, drilling resistance etc. For example, water uptake with Karsten tubes could be combined with drilling resistance and ultrasound measurements to get a better resolution of layer heterogeneity (e.g.

Bellopede, 2006; Myrin, 2006 and Vasconcelos et al., 2007). Extending this suite of methods and transferring the modern statistical methods developed in this thesis will provide an even deeper examination. Further research on linking different techniques is certainly desirable.

7.1.1 Non-destructive methods improvements and guide for good practice

Reliability of data

The studies of this thesis show, that the accuracy and reliability of the selected low-cost, non-destructive methods can be improved by using modern statistical methods. Especially in view of the special constraints provided by built heritage mentioned above like small sampling sizes and heterogeneous weathering patterns which have impacts on the data output. Outliers are expected, but have been found to reflect on natural variability of the stone and thus are considered to be part of the dataset and should be retained. Yet, this requires an alternative data evaluation approach where the dataset either needs to be modified to be able to apply common statistics or alternatively a non-parametric statistical approach. The advantage of the latter is that no data modification/transformation is necessary and still be adequate for normally distributed data. This thesis therefore, applied a range of robust non-parametric statistical methods (e.g. Kruskal Wallis, Mann-Whitney U, Spearman rank, Break-point analysis/piecewise regression, Quantile regression, Bootstrap (compare Table 3.1)).

Three low-cost, non-destructive methods have been tested under controlled conditions and improved to address the first objective of this thesis. For all three non-destructive methods (in terms of their application *in situ* on weathered stone) it has been shown that the reliability of data output can be improved significantly by applying non-parametric statistics, which is in accordance for example with Niedzielski et al., 2009. For *in situ* applications where natural stone and rock variance is expected outliers and skewed data are expected to be part of the population characteristics (Field, 2009). To keep outliers and to deal with operator variance and non-normal data modern statistics offer solutions as they are unaffected by non-normal data and no data transformation is required. This approach is more reliable for non-normally distributed data as well as being adequate for normal data.

For low impact surface hardness testing the hybrid dynamic hardness (HDH) developed in former studies as a combination of single and repeated impact method (e.g. Aoki and Matsukura, 2007; Yilmaz,2013) was improved in this study with non-parametric statistics. This study introduced the robust hybrid dynamic hardness (HDH_R) to reduce the effect of stone variability and operator variance on the data output.

It was further found that using the range of a confidence interval rather than a single value like the median to represent stone and rock surface hardness is more applicable and representative (i.e. versatile). Bootstrap techniques were employed to calculate confidence interval for the medians to fall into but also to determine sufficient sample sizes reflecting on specific characteristics of any investigated stone type (Paper 1 and Paper 3). In terms of sample sizes for low impact surface hardness testing it was found that the confidence intervals for different stone types are noticeably distinct from each other, with wider intervals for stone with complex porosities like Clipsham and narrower intervals for stones with higher compressive strength like Portland (Wilhelm et al., 2016a). Each individual stone would therefore require a specific sampling size, however this study aimed to provide a general sampling size that would be appropriate for all stone types tested, and that would be applicable for on-site application considering a range of porosities. A general number of 45 readings produced narrow enough confidence intervals with all confidence intervals for the four tested stone types clearly being distinguishable.

Throughout this thesis a confidence level of 95% for data evaluation was adapted. However, this is a very conservative approach and for *in situ* measurements and unknown weathering-stress histories of heritage stone, it might be justified to reduce the confidence level to 90%. This would still provide reliable data output when robust measures are used, but allow for smaller sampling sizes to be collected (Wilhelm et al., 2016a). The approaches discussed above can certainly be transferred to stone and rock with similar porosities and hardnesses as well as other non-destructive methods.

Block scale – Surface and subsurface information

Meneely et al. (2009) and McCabe et al. (2015) emphasize the importance of investigating stone weathering behaviour at block scale with attention to surface-to-depth heterogeneity. The three non-destructive methods of this thesis were improved in that both surface and subsurface information is gained.

This study developed further Yilmaz' approach (the combination of two measuring procedures (single impact method (SIM) and repeated impact method (RIM)→ robust hybrid dynamic testing (HDH_R)) and based the calculation on median values to account for potential effect of pores (Paper 1) and additionally can be utilized to detect crust forming behaviour (Paper 4). Therefore, using HDH_R reflects more comprehensively on stone characteristics for the surface and subsurface zone (Wilhelm et al., 2016a). For future research it would be beneficial to combine HDH_R and drilling resistance measurements (e.g. Pamplona et al., 2008) and Karsten tube water uptake to get a more precise resolution of crust thickness and its development over time.

A similar increase of depth in insight was obtained by applying breakpoint analysis to the Karsten tube water uptake data output. Varying trends in water uptake during individual measurements indicate heterogeneous subsurface zones (compare Table 6.5). Again, further research needs to be done in combining this method with for example ultrasound measurements and drilling resistance to quantify the depth of the varying subsurface layers. Furthermore, it would be desirable to implement the more into depth data evaluation into the standards. The need for alternative water uptake data evaluation techniques was as already suggested by Svahn (2006). Finally, the combination of different moisture meters with different penetration depths has an advantage over applying only one meter as again an increase into depth of information is gained.

Impact of water, porosity and salt on non-destructive testing - confounding effects on non-destructive methods – drawbacks turned into advantages

This thesis showed that perceived drawbacks (i.e. effect of porosity, surface roughness and salt content) of the tested non-destructive methods can be turned in to advantages. This study finds that low impact surface hardness testing is suitable for application on porous limestone and thus for generating reliable data for *in situ* applications provided that a sufficient sample size is collected and non-parametric data evaluation is applied surface roughness impacts operator variance can be overcome. Further, the pore characteristics are indirectly reflected in data derived from the repeated impact method (RIM) and robust hybrid dynamic hardness (HDH_R). It is therefore highly recommended to include this method extension, when low impact surface hardness is applied.

One key outcome with regards to the moisture meter testing was that meters are more or less affected by salt contamination in the stone structure, which bears the potential to utilize them as salt detectors. Paper 2 started to investigate this application, but further research needs to be done to calibrate the meters for a) different stone types and b) different salts and salt mixtures. In terms of guide for good practice and reliable *in situ* measurements the tested moisture meters should be combined with microwave moisture meters (>1GHz), which are not affected by salt (e.g. Maierhofer and Wöstmann, 1998) to investigate both salt and moisture problems.

7.1.2 Limestone weathering behaviour

The two Portland limestone varieties, focus of Paper 3 (Chapter 5), Portland Base Bed and Portland Whit Bed, are rarely distinguished in the weathering and built heritage literature. Yet, they show crucial differences in weathering behaviour, where Portland Base Bed has a higher weathering rate of 3–4 mm surface recession per 100 years (compared to 1–2 mm surface recession per 100 years for Portland Whit Bed)(BRE, 1997a,b). Reasons for not distinguishing the two varieties might be confusion in the nomenclature ("Best Bed" for Base Bed) and phrasing ("Whit Bed without shells") (Gray, 1861-1862; Edmunds and Schaffer, 1932). Additionally, the beds themselves display a natural variability which complicate attempts to distinguish them (Gray, 1861-1862; Edmunds and Schaffer, 1932). The implications for this study are limitations to correlate investigated weathering behaviour to one variety or the other. Deeper insight would only be provided through sample taking followed by laboratory tests like unconfined compressive strength, determining porosity etc. and was beyond the scope of this thesis, which focussed on non-destructive *in situ* measurements and methods. Nevertheless, it has been shown, that non-destructive in situ methods can determine stone properties and their changes sufficiently to find significant differences.

Stone weathering behaviour is complex and it has been shown, that to gain a holistic understanding its entire weathering trajectory needs to be quantified. Rather than focussing on erosion only (i.e. National Materials Exposure Program (NMEP) (Butlin et al., 1992)), preceding weathering mechanisms such as surface hardening (redeposition of solutional products) or softening (induced by both climate and biological activity) need to be investigated. Paper 3 introduced a complementary method to detect surface hardness property changes over time and provides a complementary approach to methods in previous studies that describe stone weathering behaviour using observations of surface change (Inkpen et al. 2012a, b). In paper 4 it is shown how this method can be utilized for both stone weathering research *in situ* and for preventive conservation for archaeological excavation sites to detect ongoing case hardening processes and thus, inform timely conservation measures and prevent catastrophic decay.

Whether stone weathering behaviour is defined as linear or non-linear depends highly on the scale of observation. Accordingly, Viles (2001) emphasizes, that depending on the scale of investigation weathering process-response systems may be characterised as ordered or chaotic. Both, temporal and spatial scale need to be considered and due to implications for heritage preservation the magnitude of relevant scales is noticeably different from geomorphology scales; i.e. much smaller as property changes might occur already after one year of exposure, especially for archaeological excavated stones and damage might happen after 5 years as shown in paper 4 (Chapter 6).

Paper 3 (Chapter 5) showed non-linear stone weathering behaviour for Portland limestone (i.e. surface hardness property changes), but only when considered the whole period of 250 years of exposure. For the evaluation the time-series were divided into two individual sets comprising ~90 and 250 years and QC_{50} (robust regression coefficients (0.50 quantile = median)) were determined to compare rates of surface hardness changes.

In the literature review the following issue was raised: “For example, some measurement devices only collect data from millimetre to centimetre sized areas of a surface. How representative is this of the weathering status of an entire building stone block, element of façade?" It was further pointed out, that the choice of scales depends on the ultimate aim of the research and Smith et al. (2011) and McCabe (2013) already stressed the importance of investigating stone response at a local level in order to understand spatial variability. McCabe (2015) finds stone weathering behaviour often heterogeneous at block size level. However, in this study it has been shown that to understand stone weathering behaviour *in situ* it is necessary to investigate below block size scale as significant spatial differences have been found for the top and bottom sections of investigated headstones in Paper 3 (Chapter 5).

At the archaeological excavation site in South Turkey similarly non-linear weathering has been observed though on a much shorter time scale. The crucial differences between the two sites are weathering stress histories which vary with climate, vulnerability of stone type and exposure time (250 years in the outdoors vs. ~1800 years mostly covered in soil and then being exposed for ~12 years). For parts of the archaeological remains those differences resulted in a weathering-history which lead to catastrophic decay. This maybe partly attributed to the young geological age (>37Ma) of the Gaziantep limestone formation and the specific stone properties (e.g. clay and chalk content, see section 3.4.6), but more importantly to the fact, that archaeological excavation can be seen as 'disturbing force' to a stable system (i.e. burried archaeological remains), where suddenly phyisco-chemical surroundings (e.g. pressure, temperature and chemism) change resulting in morphological evolution and structural rearrangement (Brunsden, 2001, Figure 7.3).

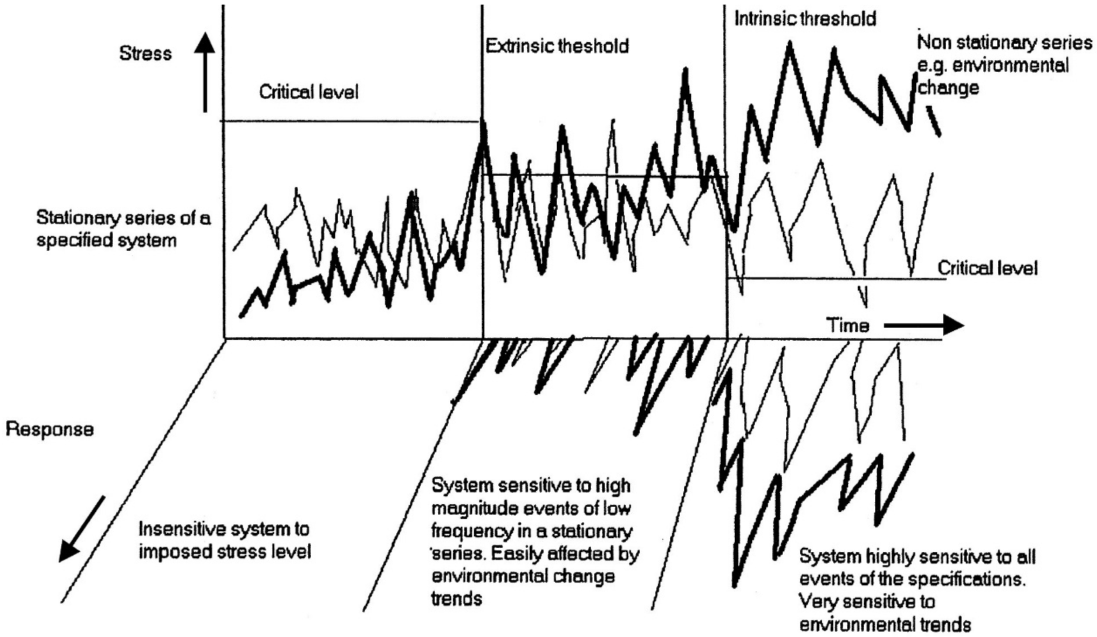


Figure 7.1 Schematic portray of stress-response sequences with thresholds, where crossing each threshold changes the critical level of response. The extrinsic threshold is exceeded at a critical level by a change in the number of events (e.g. extreme winter at archaeological site in Turkey). The intrinsic threshold is not related to increase in stress, but to internal mechanisms that change the resistance in a way that affects the response. Both stationary and non-stationary series are shown. The catastrophic stone decay at the archaeological excavation site in Turkey is an example for a non-stationary series, where the stone system due to intrinsic property changes (i.e. case hardening) became sensitive to high magnitude events (extreme winter) and responded accordingly (i.e. catastrophic decay), which left the system initially sensitive to all events until a quasi-equilibrium is reached again (i.e. stabilising crust forming after ~4 years) (Brunsden, 2001 redrawn after Church, 1980)

Figure 7.1 explains schematically the catastrophic stone decay at the archaeological excavation site in Turkey as a non-stationary series. The initial excavation is the first impulse of stress with an abrupt change in exposure conditions that force to archaeological stone to react and readjust. The stone system due to intrinsic property changes reacted with case hardening processes with a threshold of extreme vulnerability between 3 to 4 years post excavation where before and after this threshold a higher resistance to external impact is evident. During the period of extreme vulnerability the system is sensitive to high magnitude events such as the extreme winter as another major stress impulse.

To which the system responded accordingly (i.e. catastrophic decay of blocks at threshold of extreme vulnerability), which left the system initially sensitive to all events until a quasi-equilibrium was reached again (i.e. stabilising crust forming after ~4 years).

The study at the Turkey excavation site has shown that not only the weathering stress history of the stones play a role, but as well their interaction with extreme weather events. This puts limitations on predictability. Thus, stone weathering behaviour might be investigated into more depth and understood using non-destructive methods *in situ* yet the unpredictability of weather and climate leaves the conservation community with the difficult task to provide appropriate preservation measures. A potential solution would be to determine the vulnerability of an archaeological excavation site over the first 10 years as part of preventive preservation and depending on the rate of stone property changes to schedule the frequency of monitoring accordingly to provide an early warning system.

7.1.3 Multidisciplinarity

This study followed a highly interdisciplinary approach and contributed actively to scientific and practical exchange between a wide range of scientific fields. Part of this work involved learning the respective jargon and understanding different concepts common within individual disciplines. The beneficial effect of mutual understanding cannot be underestimated as it allows for knowledge exchange synergies which otherwise would not be possible.

Below are examples of jargon barriers which were overcome in this thesis; different terms to describe surface/subsurface-processes a key principle to understanding stone weathering behaviour and its often non-linear character:

- ‘Interaction zone’, surface/subsurface changes (this thesis)
- Dissolution layer or zones (near surface) (e.g. Hoke and Turcotte, 2004; Dewanckele et al., 2012)
- 'Karst weathering processes' (e.g. Inkpen et al., 2012b; Ghobadi and Torabi-Kaveh, 2014)
- Crust forming (e.g. Alexandrowicz et al., 2014; Bednarik et al., 2014)
- Weathering rind (e.g. Robinson and Moses, 2011; Stahl et al., 2013)

Furthermore, the work published in paper 4 shows that interdisciplinary collaboration between archaeologists and heritage conservation scientists can and should be part of preventive conservation. Further work is necessary especially in terms of communication, unifying language (not only in terms of jargon but also actual language (i.e. English as scientific language) as pointed out by Doehne and Price, 2010) as this will enhance transferability of results and be of great value for heritage preservation.

The improved methods may contribute and provide solutions to i) preventive preservation, ii) monitoring of conducted preservation campaigns and iii) stone weathering research under real world conditions *in situ* with the following advantages:

- Covering a greater number of heritage assets over a range of hierarchy of significance due to being low-cost

- Application with greater spatial coverage and more frequently at built heritage to address the 'the whole is greater than the sum of its parts' dilemma and scale issues
- Stone weathering behaviour can be investigated throughout the whole trajectory of stone weathering and not only when mass loss (erosion) takes place, thus contributes to a more holistic understanding of stone weathering behaviour while aiding preventive conservation practices

A truly holistic approach (though partly contradicting the key principle of preservation) would of course be to complement the *in situ* measurements introduced in this thesis with a) more non-destructive methods (both low-cost and sophisticated) and b) take samples from stone heritage with varying weathering-stress patterns and histories to analyse surface and subsurface properties at depth (under laboratory conditions with microscopic investigation, stone property determination etc). Furthermore, these samples could be exposed to long-term controlled laboratory weathering tests to gain deeper insight into resilience and potentially better predict future trajectories of weathering behaviour.

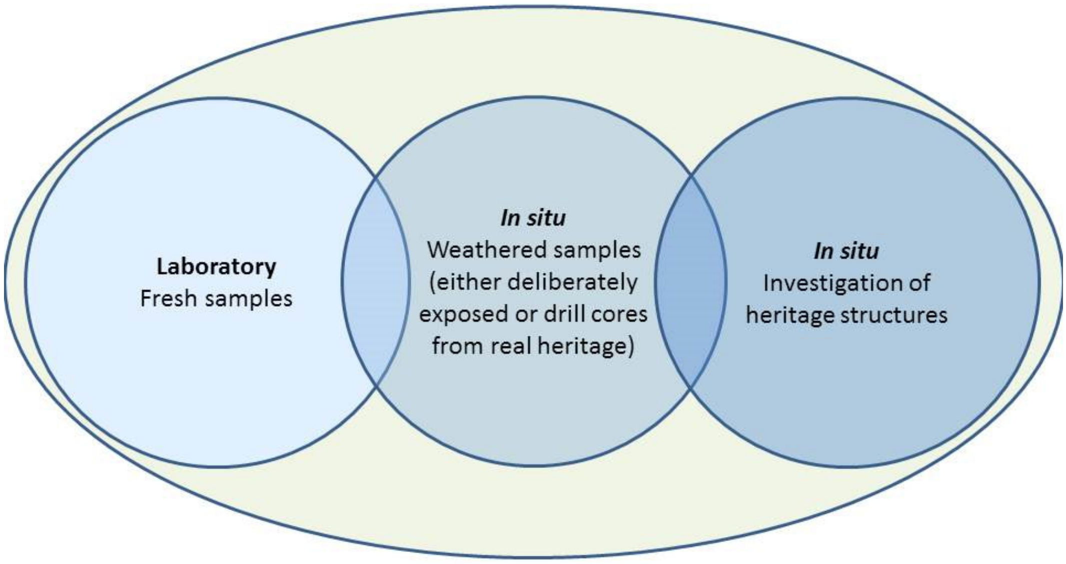


Figure 7.2 Holistic approach (outer circle) to understanding stone weathering behaviour encompasses three single approaches.

The contributions in this thesis are part of the holistic approach shown in figure 7.2 and are located at the intersection of stone weathering research and conservation practice. Thus, they serve a double purpose by advancing theoretical and applied science and inform conservation practice.

7.1.4 Review of answered research questions and main findings

Objective 1

Surface hardness testing

This study found non-normally distributed surface hardness values on porous limestone despite controlled laboratory conditions, fresh limestone and a large sufficient large sample size (120 readings). The four heritage limestones tested (Portland limestone, Bath limestone, Clipsham limestone and Guiting limestone) are all oolitic limestones and show accordingly a natural variability in stone properties. Therefore, the obtained non-normally distributed data (outliers and skewness) was attributed to natural inherent variability of the stones which can be expected for porous limestone as stated by Palmer (2008).

What are the most appropriate statistical methods to handle Equotip data? How should outliers be treated?

This study solved the problem of natural variance of limestone affecting Equotip data output by increasing collected sample size and applying non-parametric statistical data evaluation. The non-parametric statistical approach applied in this study requires no data transformation (e.g. removing outliers and more). This study suggests instead to include outliers as their occurrence is linked to natural stone characteristics – as observed in this study by the presence of fossils and other harder elements of the limestones.

What is an adequate sample size to collect for Equotip applied to porous stone?

Bootstrapping technique was used to determine the width of confidence intervals for the surface hardness median of the individual stone types. Although it was found that appropriate sample sizes of Equotip values depend on the limestone type (varying porosity), for practical reasons for in situ applications on heritage limestone with unknown weathering history this study recommends a general sample size of 45 readings (for a confidence level of 95%). This approach can certainly be transferred to stone and rock with similar porosities and hardness.

It is worth mentioning that calculating sample sizes using a 95% confidence level is a conservative approach. In view of the expected variances for in situ measurements and unknown weathering-stress histories of heritage stone, it might be justified to reduce the confidence level to 90%. This would still provide reliable data output when robust measures are used, but allow for a smaller sample size to be collected.

Is the Equotip appropriate for application on porous limestone?

This study finds that Equotip is suitable for application on porous limestone and thus for generating reliable data for in situ applications, given the advice derived from this research to collect a sufficient sample size and apply non-parametric data evaluation.

How do the Equotip D and DL probes compare?

The DL probe generates higher hardness values compared to the D probe, which was confirmed by Proceq© as being usual (Personal communication 28/11/2013). Therefore, obtained values cannot be compared directly. Both probes are suitable for application to porous limestone, when considering the recommendations on robust statistical data evaluation above. Nevertheless, the DL probe was found to relate better to porosity characteristics than the D probe. Furthermore, with its slim long (82 mm) front section it is a) more suitable for confined spaces and recessed surfaces and b) it also prevents the impact body from transporting particles into the body of the device and offers protection from dust for the Equotip device itself.

How to address affects like surface roughness?

This study adapted the hybrid dynamic hardness (HDH) testing measure first introduced by Yilmaz. This approach combines the two application modes of the Equotip, single (individual) impacts (SIM) and repeated impacts (RIM). The repeated impact measurement (RIM) reflects on the elastic and plastic deformation characteristics. Therefore, using HDH reflects more comprehensively on stone characteristics for the surface and subsurface zone as has been shown in this study. For reasons of reducing the effect of stone variability this study introduced the modified hybrid dynamic hardness (HDH_R), where SIM and RIM are represented by the more robust median.

Handheld electronic moisture meters

This study found that all four tested electrical moisture meters (based on resistivity or capacitance principles) were affected by salt in porous stone. The effect of two realistic salt contamination levels commonly found in built heritage (based on Arendt and Seele 2000) was quantified for two resistivity type moisture meters, the Protimeter and the Resipod, and two capacitance type moisture meters (the Protimeter in resistivity capacitance mode and the CEM).

This study induced dissolution of the salt in the contaminated stone samples by changes of relative humidity only and thus, simulated on-site moisture measurement situations in historic structures where salt deterioration processes are driven by changing relative humidities as found by Colston et al. (2001) and Linnow et al. (2007). In addition, this study isolated the effect of NaCl increasing conductivity (in pore water) from the combined influence of increased conductivity and water retention (at sorption equilibrium). It has been found that the salt contamination effect on the moisture meter data output is more pronounced for higher salt content and when samples have reached sorption equilibrium under 95% RH conditions.

The Protimeter Surveymaster was found to be the most versatile moisture meter of this study and its capacitance mode is useful to complement measurements based on resistivity, in order to help clarify results and discriminate moisture from salt effects. These promising results show the potential for moisture meters to actually detect salt contamination in stone structure. These findings have important implications for situations at built heritage sites where the hygroscopic interaction of salts might be mistaken for rising damp. The latter diagnosis would result in a substantially different conservation intervention compared to desalination procedures or climate control to manage salt contamination problems (Charola, 2000). Thus, the results are promising with regards to a potential quick and simple method to detect salt contamination on site and inform appropriate conservation management decisions. However, further research is needed to verify these findings for heritage materials under real world conditions.

Objective 2

This section focuses on real architectural and archaeological heritage sites with weathering-stress histories of limestone only partly or un-known. The developed non-destructive methods from objective 1 were conjunctly used to establish weathering rates for Portland limestone on a short- and long-term scale (objective 2) and to infer the causes for catastrophic limestone decay at the archaeological excavation site Dülük Baba Tepesi in South Turkey.

Can deterioration rates of Portland limestone monoliths be developed by means of surface hardness changes?

Deterioration rates for Portland limestone gravestones have been established by means of surface hardness changes with low impact surface hardness testing with Equotip in situ. which is a novel application for the Equotip device. In addition this study introduced the novel robust proxy QC_{50} (0.50 quantile regression coefficient of surface hardness QC_{50}) to describe weathering rates. As Equotip is sensitive to minor changes in stone weathering-stress history within the course of a few years, it is suitable to investigate short-term limestone weathering behaviour.

In what ways (linear or not) deteriorate Portland limestone monoliths over a period of 250 years?

Furthermore, for the investigated time-series of gravestones covering a weathering history period of 1 to 248 years non-linear stone weathering behaviour was found. This conforms with findings from other stone weathering studies, for example, by Mosch and Siegesmund, 2007; Palmer, 2008; Hansen et al., 2013; Alberti et al., 2013; Emmanuel, 2015. These studies, however, collect data on stone weathering using very different metrics (usually surface recession rates), thus the current research gives independent confirmation of this behaviour.

The time period was scaled down by segmenting with piecewise regression and a clear distinction was found between gravestones with < 100 years of exposure and those with > 100 to 250 years. Various hypotheses for this finding were proposed, and deserve further testing.

Are there spatial differences in deterioration over time?

Clear differences in spatio-temporal weathering behaviour were observed. The top section of gravestones with c. 100 year exposure history weather at the highest rate of all tested stones with QC_{50} = -2.66, which is more than twice the rate of the bottom section with QC_{50} = -1.19. For gravestones with a longer exposure history considerably lower weathering rates were found with QC_{50} = - 0.65 for top and QC_{50} = - 0.57 for bottom sections and no clear distinction was found between top and bottom.

How do results compare to limestone recession rates derived from former studies?

The QC_{50} weathering rates cannot be directly compared to weathering rates established by former studies based on recession rates or gain/loss in surface height as different weathering-stress parameters are measured. However, following Inkpen et al. (2012b) who point out the advantages in considering both surface recession and the rate of 'surface change', this study adds to the suite of methods to investigate stone weathering rates in situ. The QC_{50} complements other methods like micro erosion meter measurements, water run-off and weight loss. It extends the range of information to be gathered on limestone weathering behaviour by 1) measuring vertical and horizontal stone decay rate (compare to micro erosion measurement, which is mainly applied horizontally), 2) detecting both short- and long-term weathering-stress history changes and 3) when HDH is used, it reflects on surface and subsurface plastic/elastic deformation properties. Thus it is able to detect property changes before actual material loss (i.e. dissolution) takes place and has therefore implications for preventive conservation. It is able to not only detect surface hardness decrease, but also increase, which might indicate ongoing induration and crust forming processes. Therefore, establishing weathering rates using the novel proxy QC_{50} complements existing methodology on stone weathering rate research and provides the basis for a more comprehensive understanding of limestone weathering behaviour.

Objective 3

What caused the catastrophic limestone decay?

The interrelationship between extreme weather events and crust forming behaviour (case hardening) of Hellenistic-Roman limestone remains at the archaeological excavations site Dülük Baba Tepesi (South Turkey) caused catastrophic decay. The influence of the weathering-stress history (including case hardening) on catastrophic decay was reconstructed by applying non-destructive testing methods in situ (surface hardness testing and water uptake) on a time series of blocks excavated in different recent years (2005, 2007 and 2013). Two stone types were identified (Firat and Gaziantep formation) and non-linear stone weathering behaviour observed, with the different stone type showing different crust forming characteristics. The results were further linked to past climate data (precipitation, number of frost days, the number of individual freeze-thaw cycles and Wet-Frost index). The cold period from October 2011 to April 2012, after which the catastrophic decay was observed, had 63 frost days, which is the maximum for the period of interest in this study (2005 – 2013). In addition, unusually high precipitation average and a Wet-Frost index illustrate the damaging nature of the climatic conditions. The results of this study illustrate the increased damage potential of temporal sequences of case hardening processes in limestones and episodic harsh winter condition to catastrophic deterioration.

What are the implications for conservation interventions and future site management?

Based on the results a model was developed to describe the observed non-linear stone weathering behaviour. The results indicate that both stone variety (with the typical stone weathering behaviour characteristics) and length of exposure may determine the magnitude of the impact of the severe climatic event. In this study the harder stone (Firat) and the longest exposure period (9 years) are least affected compared to the softer stone (Gaziantep) and a shorter exposure time (7 years). These variations are attributed to crust forming behaviour, which has been found to temporarily affect the long-term stability of built heritage structures as reported by Zehnder (1996) (for wall paintings, as quoted in Charola et al., 2007). The observed variability of stone weathering behaviour at block scale agrees with McCabe et al. (2015), who find both accumulative weathering-stress history and micro-environmental climate results in surface-to-depth heterogeneity at the stone block scale. The results of this study confirm the importance of investigating stone response at a local level in order to understand spatial variability as has been emphasized by both McCabe et al. (2015) and Smith et al. (2011). Furthermore, the observed spatial variability is thought to be key to understanding complex interactions between post-excavation case hardening and extreme winter conditions which can lead to catastrophic deterioration.

7.1.5 Concluding remark

The overall underlying motivation for this research was to contribute to the preventive conservation of vulnerable cultural stone heritage. Thus, the improved methods may assist both 1) understanding heritage stone weathering under real world conditions (without damaging them by sample taking, whilst capturing surface/subsurface changes) and 2) more frequent investigation of the state of preservation/deterioration of stone heritage on-site in order to detect ongoing deterioration at an early stage and thus prevent (or at least mediate) for example catastrophic decay.

8. References

Abdelhafez A, El-Wekeel FM, Ramadan EM, Abed-Allah AA, 2012. Microbial deterioration of archaeological marble: Identification and treatment. *Annals of Agricultural Sciences* **57**: 137–144. DOI: 10.1016/j.aoas.2012.08.007.

Adam, JP, 1999. *Roman building: Materials and Techniques*. London: Routledge.

Adamo P, Violante P, 2000. Weathering of rocks and neogenesis of minerals associated with lichen activity. *Applied Clay Science* **16**: 229–256. DOI: 10.1016/S0169-1317(99)00056-3.

Ahmad A, Pamplona M, Simon S, 2009. Ultrasonic testing for the investigation and characterization of stone - a non-destructive and transportable tool. *Reviews in conservation* **10**: 43–53.

Ahmad A, 2011. Characterization of natural and consolidated stones from Jordan with non-destructive ultrasonic technique and physico-mechanical methods. PhD thesis. University of Dortmund: Dortmund.

Alberti AP, Gomes A, Trenhaile A, Oliveira M, Horacio J, 2013. Correlating river terrace remnants using an Equotip hardness tester: An example from the Miño River, northwestern Iberian Peninsula. *Geomorphology* **192**: 59–70. DOI: 10.1016/j.geomorph.2013.03.017.

Albion's Quarry report 56 (2012), www.albionstone.com (accessed 06/05/2013)

Al-Houdalieh SH, 2009. Political Crisis and Palestine's Cultural Heritage: A Case Study from the Khirbet el-Lauz Site in Area C. *Journal of Field Archaeology* **34**: 338–350. DOI: 10.2307/25608590.

Alexandrowicz Z, Marszałek M, Rzepa G, 2014. Distribution of secondary minerals in crusts developed on sandstone exposures. *Earth Surface Processes and Landforms* **39**: 320–335. DOI: 10.1002/esp.3449.

Aliabdo, AA, Elhakam, EA, Mohamed A, 2012. Reliability of using nondestructive tests to estimate compressive strength of building stones and bricks. *Alexandria Engineering Journal* **51**: 193–203. DOI: 10.1016/j.aej.2012.05.004.

Aliha MRM, Sistaninia M, Smith DJ, Pavier MJ, Ayatollahi MR, 2012. Geometry effects and statistical analysis of mode I fracture in guiting limestone. *International Journal of Rock Mechanics and Mining Sciences* **51**: 128–135. DOI: 10.1016/j.ijrmms.2012.01.017.

Alvarez Grima M, Babuška R, 1999. Fuzzy model for the prediction of unconfined compressive strength of rock samples. *International Journal of Rock Mechanics and Mining Sciences* **36**: 339–349. DOI: 10.1016/S0148-9062(99)00007-8.

Andriani, G. F. and Walsh, N. 2007. The effects of wetting and drying, and marine salt crystallization on calcarenite rocks used as building material in historic monuments. *Geological Society of London Special Publications*, **271**: 179-188.

Arbona A, Bona C, Miñano B, Plastino A, 2014. Statistical complexity measures as telltale of relevant scales in emergent dynamics of spatial systems. *Physica A: Statistical Mechanics and its Applications* **410**: 1–8. DOI: 10.1016/j.physa.2014.05.005.

Aoki H, Matsukura Y, 2007. A new technique for non-destructive field measurement of rock-surface strength: an application of the Equotip hardness tester to weathering studies. *Earth Surface Processes and Landforms* **32**: 1759–1769. DOI: 10.1002/esp.1492.

Aoki H, Matsukura Y, 2008. Estimating the unconfined compressive strength of intact rocks from Equotip hardness. *Bulletin of Engineering Geology and the Environment* **67**: 23–29. DOI: 10.1007/s10064-007-0116-z.

Arendt C, Seele J, 2000. *Feuchte und Salze in Gebäuden.* Leinfelden-Echterdingen: Koch.

Ariño X, Ortega-Calvo JJ, Gomez-Bolea A, Saiz-Jimenez C, 1995. Lichen colonization of the Roman pavement at Baelo Claudia (Cadiz, Spain): biodeterioration vs. bioprotection. *Science of The Total Environment* **167**: 353–363. DOI: 10.1016/0048-9697(95)04595-R.

Arkell WJ, 1947. Oxford Stone. Faber & Faber: London.

Arnold A, 1984. Determination of Mineral Salts from Monuments. *Studies in Conservation* **29**: 129. DOI: 10.2307/1506015.

Ashurst J, Dimes FG, 1998. *Conservation of building and decorative stone.* Butterworth-Heinemann: Oxford, Woburn, Mass.

Auras M, 2011a. Durchführung und Auswertung von Ultraschallmessung. In *Leitfaden Naturstein-Monitoring: Nachkontrolle und Wartung als zukunftsweisende Erhaltungsstrategien*, Auras M (ed). Fraunhofer-IRB-Verl.: Stuttgart; 47–54.

Auras M, 2011b. *Leitfaden Naturstein-Monitoring: Nachkontrolle und* Wartung *als zukunftsweisende Erhaltungsstrategien*. Fraunhofer-IRB-Verl.: Stuttgart.

Aydin A, 2009. ISRM Suggested method for determination of the Schmidt hammer rebound hardness: Revised version. *International Journal of Rock Mechanics and Mining Sciences* **46**: 627–634. DOI: 10.1016/j.ijrmms.2008.01.020.

Aydin A, Basu A, 2005. The Schmidt hammer in rock material characterization. *Engineering Geology* **81**: 1–14. DOI: 10.1016/j.enggeo.2005.06.006.

Banerjee S, Iglewicz B, 2007. A Simple Univariate Outlier Identification Procedure Designed for Large Samples. *Communications in Statistics: Simulation & Computation* **36**: 249–263. DOI: 10.1080/03610910601161264.

Barger S, 1989. Characterization of corrosion products on old protective glass, especially daguerreotype cover glasses. *Journal of Material Science* **24**: 1343–1356.

Bayer K, Köhler W, Schuh H, Wendler E, 2010. Vergleichsuntersuchungen mit unterschiedlichen zerstörungsfreien und zerstörungsarmen Feuchte- und Salzmessmethoden. In *Zerstörungsfreie Oberflächenuntersuchungsmethoden: Stand der Technik, Grenzen und Ausblicke Beiträge des 4. Konservierungswissenschaftlichen Kolloquiums in Berlin-Brandenburg am 29. Oktober in Potsdam*, Arnold B (ed). Werner: Worms: 69–74.

Baykasoglu A, Gullu H, Çanakci H, Ozbakir L, 2008. Prediction of compressive and tensile strength of limestone via genetic programming. *Expert Systems with Applications* **35**: 111–123. DOI: 10.1016/j.eswa.2007.06.006.

Bednarik M, Moshammer B, Heinrich M, Holzer R, Laho M, Rabeder J, Uhlir C, Unterwurzacher M, 2014. Engineering geological properties of Leitha Limestone from historical quarries in Burgenland and Styria, Austria. *Engineering Geology* **176**: 66–78. DOI: 10.1016/j.enggeo.2014.04.005.

Beierkuhnlein C, 2011. *Zwischen Denkmalschutz und Naturschutz: Leitfaden zur naturverträglichen Instandhaltung von Mauerwerk in der Denkmalpflege*. Erich Schmidt: Berlin.

Bell FG, 1993. Durability of carbonate rock as building stone with comments on its preservation. *Environmental Geology* **21**: 187–200. DOI: 10.1007/BF00775905.

Bell FG, Coulthard JM, 1990. Stone preservation with illustrative examples from the United Kingdom. *Environmental Geology and Water Sciences* **16**: 75–81. DOI: 10.1007/BF01702226.

Bellopede R, Manfredotti L, 2006. Ultrasonic sound test on stone: comparison of indirect and direct methods under various test conditions. In *Heritage, weathering and conservation: Proceedings of the International Conference on Heritage, Weathering and Conservation, HWC-2006, 21-24 June 2006, Madrid, Spain*, Fort R (ed). Taylor & Francis: London, New York; 539–546.

Benavente D, García del Cura MA, Fort R, Ordóñez S, 2004. Durability estimation of porous building stones from pore structure and strength. *Engineering Geology* **74**: 113–127. DOI: 10.1016/j.enggeo.2004.03.005.

Benavente D, Cultrone G, Gómez-Heras M, 2008. The combined influence of mineralogical, hygric and thermal properties on the durability of porous building stones. *European Journal of Mineralogy* **20**: 673–685. DOI: 10.1127/0935-1221/2008/0020-1850.

Benedix R, 1999. Chemie für Bauingenieure. Teubner: Stuttgart.

Bernal, Pérez Juan Luis, Bello López MA, 2000. The fractal dimension of stone pore surface as weathering descriptor. *Applied Surface Science* **161**: 47–53. DOI: 10.1016/S0169-4332(00)00031-3.

Bionda D, 2004. Salt deterioration and microclimate in historical buildings. Intermediary report of the project "Sustained care of sensitive historical monuments". Institut für Denkmalpflege, Forschungsstelle Technologie und

Konservierung, ETH Zürich. http://e-collection.library.ethz.ch (accessed: 22/05/15)

Bjelland T, Thorseth IH, 2002. Comparative studies of the lichen–rock interface of four lichens in Vingen, western Norway. *Chemical Geology* **192**: 81–98. DOI: 10.1016/S0009-2541(02)00193-6.

Blakemore P, Morrow A, Liew W, Muruganathan M, Luff R, 2005. Species and treatment corrections for hand-held moisture meters with radiata and slash pine. Victoria: Forest & Wood Products Research & Development Corporation.

Bläuer-Böhm C, Franzen C, Vergès-Belmin V, 2012. Simple field test in stone conservation. In *12th International Congress on the Deterioration and Conservation of Stone*. New York 2012, Wheeler G (ed): New York.

Blömer M, Winter E, 2005. Doliche und das Heiligtum des Iupiter Dolichenus auf dem Dülük Baba Tepesi: 1. Vorbericht (2001-2003). Istanbuler Mitteilungen 55: 187–210.

Blömer M, 2011. Von Kummuh nach Telouch. Archäologische und historische Untersuchungen in Kommagene. In: Winter E, ed. Dolichener und Kommagenische Forschungen IV. Asia Minor Studien 64, Habelt: Bonn.

Blömer M, 2012. Iuppiter Dolichenus zwischen lokalem Kult und reichsweiter Verehrung. In *Iuppiter Dolichenus: Lokalkult und Reichsreligion im Vergleich*, Blömer M and Winter E (eds). Tübingen: Mohr Siebeck, pp. 39-98.

Bonazza A, Sabbioni C, Ghedini N, Hermosin B, Jurado V, Gonzalez JM, Saiz-Jimenez C, 2007. Did Smoke from the Kuwait Oil Well Fires Affect Iranian Archaeological Heritage? *Environmental Science & Technology* **41**: 2378–2386. DOI: 10.1021/es062176s.

Bonazza A, Messina P, Sabbioni C, Grossi CM, Brimblecombe P, 2009. Mapping the impact of climate change on surface recession of carbonate buildings in Europe. Science of *The Total Environment* **407**: 2039–2050. DOI: 10.1016/j.scitotenv.2008.10.067.

Bondell HD, Reich BJ, Wang H, 2010. Non-crossing quantile regression curve estimation. *Biometrika* **97**: 825–838. DOI: 10.1093/biomet/asq048.

Bourges A, 2006. Holistic correlation of physical and mechanical properties of selected natural stones for assessing and weathering in the natural environment. PhD thesis. Ludwig-Maximilians-Universität: Munich.

Breysse D, 2012. Non destructive evaluation of concrete strength: An historical review and a new perspective by combining NDT methods. *Construction and Building Materials* **33**: 139–163. DOI: 10.1016/j.conbuildmat.2011.12.103.

Brimblecombe P, 2010. Environment and architectural stone. In *Stone in Architecture: Properties, Durability*, Siegesmund S, Snethlage R, Winkler E (eds). Springer: Berlin; 317–346.

Brimblecombe P, Grossi CM, 2009. Millennium-long damage to building materials in London. *Science of The Total Environment* **407**: 1354–1361. DOI: 10.1016/j.scitotenv.2008.09.037.

British Standards Institute (BSI), 2006a. Natural stone test methods – Determination of uniaxial compressive strength. BSI: London. (BS EN 1926:2006).

British Standards Institute (BSI), 2006b. Natural stone test methods— Determination of real density and apparent density, and of total and open porosity. BSI: London. (BS EN 1936:2006).

British Standards Institute (BSI), 2006c. Natural stone test methods— Determination of water absorption at atmospheric pressure. BSI: London. (BS EN 13755:2008).

British Standards Institute (BSI), 2013. Conservation of cultural heritage - Test methods - Measurement of water absorption by pipe method. (BS EN 16302:2013 (EN 16302:2013 (E)). BSI Standards Limited 2013: London.

Bruno F, Bianco G, Muzzupappa M, Barone S, Razionale AV, 2011. Experimentation of structured light and stereo vision for underwater 3D reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing* **66**: 508–518. DOI: 10.1016/j.isprsjprs.2011.02.009.

Bruno G, Vessia G, Bobbo L, 2013. Statistical Method for Assessing the Uniaxial Compressive Strength of Carbonate Rock by Schmidt Hammer Tests Performed on Core Samples. *Rock Mechanics and Rock Engineering* **46**: 199–206. DOI: 10.1007/s00603-012-0230-5.

Brunsden D, 2001. A critical assessment of the sensitivity concept in geomorphology. *CATENA* **42**: 99–123. DOI: 10.1016/S0341-8162(00)00134-X.

Building Research Establishment (BRE), 1997a. Portland Base Bed Limestone (Bowers Quarry): Technical Data Sheet. The BRE / British Stone Stone List (Accessed 06/05/2013).

Building Research Establishment (BRE), 1997b. Portland Whit Bed Limestone (Bowers Quarry): Technical Data Sheet. The BRE / British Stone Stone List. Accessed 06/05/2013).

Building Research Establishment (BRE), 1997c. Cotswold Hill Quarry Cream(Bed 2 – White Guiting): Technical Data Sheet. The BRE / British Stone Stone List (Accessed 06/05/2013).

Burkinshaw R, 2002. What is the moisture meter trying to tell us? *Structural Survey* **20**: 162–172. DOI: 10.1108/02630800210456814.

Burkinshaw R, 2004. Which instruments should surveyors use to monitor moisture condition? *Structural Survey*, **22**: 7–19.

Burkinshaw R, Parrett M, 2003. *Diagnosing damp*. Coventry: RICS Books.

Butlin RN, Coote AT, Devenish M, Hughes I, Hutchens CM, Irwin JG, Lloyd GO, Massey SW, Webb AH, Yates T, 1992. Preliminary results from the analysis of stone tablets from the National Materials Exposure Programme (NMEP). *Atmospheric Environment. Part B. Urban Atmosphere* **26**: 189–198. DOI: 10.1016/0957-1272(92)90022-K.

Cade BS, Noon BR, 2003. A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment* **1**: 412–420. DOI: 10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2.

Cámara B, Álvarez de Buergo, M., Fort R, Ascaso C, Roís, A. de los, Gomez-Heras M, 2014. Another source of soluble salts in urban environments due to recent social behaviour pattern in historical centres. In *Science, technology and cultural heritage: Proceedings of the second International Congress on Science and Technology for the Conservation of Cultural Heritage, Sevilla, Spain, 24-27 June 2014*, Rogerio-Candelera MÁ (ed). CRC Press/Balkema: London; 89–94.

Camuffo D, 1995. Physical weathering of stones. *Science of The Total Environment* **167**: 1–14. DOI: 10.1016/0048-9697(95)04565-I.

Camuffo D, 2014. *Microclimate for cultural heritage: Conservation and restoration of indoor and outdoor monuments*. Elsevier: Amsterdam, New York.

Camuffo D, Sturaro G, 2001. The climate of Rome and its action on monument decay. *Climate Research* **16**: 145–155.

Çanakci H, 2007. Collapse of caves at shallow depth in Gaziantep city center, Turkey: a case study. *Environmental Geology* **53**: 915–922. DOI: 10.1007/s00254-007-0802-y.

Çanakci H, Demirboğa R, Burhan Karakoç M, Şirin O, 2007. Thermal conductivity of limestone from Gaziantep (Turkey). *Building and Environment* **42**: 1777–1782. DOI: 10.1016/j.buildenv.2006.01.011.

Caneva G, Nugari MP, Salvadori O et al., 2008. *Plant biology for cultural heritage: Biodeterioration and conservation*. Getty Conservation Institute: Los Angeles.

Carter N, Viles H, 2005. Bioprotection explored: the story of a little known earth surface process. *Geomorphology* **67**: 273–281. DOI: 10.1016/j.geomorph.2004.10.004.

Cassar J, Winter MG, Marker BR, Bromhead EN, Smith, J. W. N., Toll DG, Walton, N. R. G., Entwisle DC, Dijkstra TA, 2013. Introduction to the Stone Cycle and the Conservation of Historic Buildings. *Quarterly Journal of Engineering Geology and Hydrogeology* **46**: 363–366. DOI: 10.1144/qjegh2013-052.

Cataldo R, Leucci G, Siviero S, Pagiotti R, Angelini P, 2009. Diagnostic of the conservation state in the crypt of the Abbey of Montecorona: biological, microclimatic and geophysical evaluations. *Journal of Geophysics and Engineering* **6**: 205–220. DOI: 10.1088/1742-2132/6/3/001.

Charola AE, 2000. Salts in the Deterioration of Porous Materials: An Overview. *Journal of the American Institute for Conservation* **39**: 327–343. DOI: 10.2307/3179977.

Charola AE, Ware R, 2002. Acid deposition and the deterioration of stone: A brief review of a broad topic. *Geological Society, London, Special Publications 205*: 393–406. DOI: 10.1144/GSL.SP.2002.205.01.28.

Charola AE, Pühringer J, Steiger M, 2007. Gypsum: a review of its role in the deterioration of building materials. *Environmental Geology* **52**: 339–352. DOI: 10.1007/s00254-006-0566-9.

Chen J, Blume H, Beyer L, 2000. Weathering of rocks induced by lichen colonization — a review. *CATENA* **39**: 121–146. DOI: 10.1016/S0341-8162(99)00085-5.

Choquette PW, Pray LC, 1970. Geologic nomenclature and classification of porosity in sedimentary carbonates. *American Association of Petroleum Geologists Bulletin* **54**: 207–250.

Church, M., 1980. Records of recent geomorphological events. *In Timescales in Geomorphology*. Cullingford, RA, Davidson DA, Lewin J (eds). Wiley.

Churches Conservation Trust (CCT) www.visitChurches.org.uk (accessed 07/2015)

Colston BJ, Watt DS, Munro HL, 2001. Environmentally-induced stone decay: the cumulative effects of crystallization–hydration cycles on a Lincolnshire oopelsparite limestone. *Journal of Cultural Heritage* **2**: 297–307. DOI: 10.1016/S1296-2074(01)01129-3.

Colvin H, 2008. *A biographical dictionary of British architects, 1600-1840*. Yale University Press: New Haven.

Commonwealth War Graves Commission, www.cwgc.org. (Accessed 01/09/2015).

Cooke RU, Doornkamp JC, 1990. *Geomorphology in environmental management: A new introduction*. Clarendon Press; Oxford University Press: Oxford, New York.

Cooke RU, Gibbs GB, 1994. Crumbling heritage? Studies of stone weathering in polluted atmospheres. *Atmospheric Environment* **28**: 1355–1356. DOI: 10.1016/1352-2310(94)90284-4.

Cooke RU, Inkpen RJ, Wiggs, G. F. S., 1995. Using gravestones to assess changing rates of weathering in the United Kingdom. *Earth Surface Processes and Landforms* **20**: 531–546. DOI: 10.1002/esp.3290200605.

Coombes MA, Feal-Pérez A, Naylor LA, Wilhelm K, 2013. A non-destructive tool for detecting changes in the hardness of engineering materials: Application of the Equotip durometer in the coastal zone. *Engineering Geology* **167**: 14–19. DOI: 10.1016/j.enggeo.2013.10.003.

Cooper TP, O'Brien PF, Jeffrey DW, 1992. Rates of Deterioration of Portland Limestone in an Urban Environment. *Studies in Conservation* **37**: 228. DOI: 10.2307/1506352.

Cosentino PL, Capizzi P, Martorana R, Messina P, Schiavone S, 2011. From Geophysics to Microgeophysics for Engineering and Cultural Heritage. *International Journal of Geophysics* **2011**. DOI: 10.1155/2011/428412.

Coskun B, 2000. The Dead Sea Fault and related subsurface structures, Gaziantep Basin, southeast Turkey. *Geological Magazine* **137**: 175–192. DOI: 10.1017/S0016756800003770.

Costanzo A, Minasi M, Casula G, Musacchio M, Buongiorno MF, 2015. Combined use of terrestrial laser scanning and IR thermography applied to a historical building. *Sensors* **15**: 194–213. DOI: 10.3390/s150100194.

Council of Europe, 1992. *European Convention on the Protection of the Archaeological Heritage (Revised)* (also known as Malta Convention or Valetta Treaty). Council of Europe: Strasbourg.

Crawley MJ, 2005. Statistics: *An introduction using R*. J. Wiley: Chichester, West Sussex, England.

Crawley MJ, 2007. *The R book*. Wiley: Chichester, England, Hoboken, N.J.

Cutler NA, Viles HA, Ahmad S, McCabe S, Smith BJ, 2013. Algal ‘greening’ and the conservation of stone heritage structures. Science of The Total Environment **442**: 152–164. DOI: 10.1016/j.scitotenv.2012.10.050.

Dagistan H, Simsek S, 2005. Geological and Hydrogeological Investigation of Kozluk-Taslidere (Batman) Geothermal Field. In *Proceedings World Geothermal Congress*. (WGC 2005), Horne R, Okandan E (eds). Turkey, Antalya.

Dancey CP, Reidy J, 2004. *Statistics without maths for psychology: Using SPSS for Windows*. Prentice Hall: New York.

Dette H, Volgushev S, 2008. Non-crossing non-parametric estimates of quantile curves. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **70**: 609–627. DOI: 10.1111/j.1467-9868.2008.00651.x.

D'ham G, Meinhardt J, Niemeyer R, 2011. Bestimmung der kapillaren Wasseraufnahme mti Messröhrchen nach Karsten und Mirowski. In *Leitfaden Naturstein-Monitoring: Nachkontrolle und Wartung als zukunftsweisende Erhaltungsstrategien*, Auras M (ed). Fraunhofer-IRB-Verl.: Stuttgart; 82–92.

Diaz Gonçalves T, Delgado Rodrigues J, 2006. Evaluating the salt content of salt-contaminated samples on the basis of their hygroscopic behavior. Part I. Fundamentals, scope and accuracy of the method. *Journal of cultural heritage* **7**: 79-84

Dinçer I, Acar A, Çobanoğlu I, Uras Y, 2004. Correlation between Schmidt hardness, uniaxial compressive strength and Young’s modulus for andesites, basalts and tuffs. *Bulletin of Engineering Geology and the Environment* **63**: 141–148. DOI: 10.1007/s10064-004-0230-0.

Doehne EF, Price CA, 2010. *Stone conservation: An overview of current research.* Getty Conservation Institute: Los Angeles.

Drdácký M, Hasníková H, Le`sak, Pavel Z, 2012. Innovated water uptake measurements on historic stone surfaces. In *12th International Congress on the Deterioration and Conservation of Stone. New York 2012*, Wheeler G (ed). New York.

Drdácký M, Slížková Z, 2013. Enhanced affordable methods for assessing material characteristics and consolidation effects on stone and mortar. *Journal of Geophysics and Engineering* **10**: 064005-064011. DOI: 10.1088/1742-2132/10/6/064005.

Dubelaar CW, Engering S, Hees R, Koch R, Lorenz H, 2003. Lithofacies and petrophysical properties of Portland Base Bed and Portland Whit Bed limestone as related to durability. *Heron* **48**: 221–229.

Dunham R J, 1962. Classification of carbonate rocks according to depositional texture. In *Classification of carbonate rocks: American Association of Petroleum Geologists Memoir,* Ham WE (ed.); 108-121.

Duval M, Smith B, 2013. Rock art tourism in the uKhahlamba/Drakensberg World Heritage Site: Obstacles to the development of sustainable tourism. *Journal of Sustainable Tourism* **21**: 134–153. DOI: 10.1080/09669582.2012.699060.

Edmunds FH, Schaffer RJ, 1932. Portland Stone; its geology and properties as a building stone. *Proceedings of the Geologists' Association* **43**: 225–IN4. DOI: 10.1016/S0016-7878(32)80019-2.

Efron B, 1987. Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association* **82**: 171–185. DOI: 10.2307/2289144.

Eklund JA, Zhang H, Viles HA, Curteis T, 2013. Using Handheld Moisture Meters on Limestone: Factors Affecting Performance and Guidelines for Best Practice. *International Journal of Architectural Heritage* **7**: 207–224. DOI: 10.1080/15583058.2011.626491.

Embry AF, Klovan JE, 1971, A Late Devonian reef tract on Northeastern Banks Island, *NWT: Canadian Petroleum Geology Bulletin* **19**: 730-781.

Emmanuel S, Levenson Y, 2014. Limestone weathering rates accelerated by micron-scale grain detachment. *Geology* **42**: 751–754. DOI: 10.1130/G35815.1.

Emmanuel S, 2015. Evidence for non-Gaussian distribution of rock weathering rates. *Earth Surface Dynamics* **3**: 441-445. DOI: 10.5194/esurf-3-441-2015.

Erceg-Hurn DM, Mirosevich VM, 2008. Modern robust statistical methods: an easy way to maximize the accuracy and power of your research. *The American psychologist* **63**: 591–601. DOI: 10.1037/0003-066X.63.7.591.

Erkal A, D'Ayala D, Stephenson V, 2013. Evaluation of environmental impact on historical stone masonry through on-site monitoring appraisal. *Quarterly Journal of Engineering Geology and Hydrogeology* **46**: 449–458.

European Union, 2013. *Final Report - CRIC (Identity and conflict. Cultural heritage and the re-construction of identities after conflict.): Project reference: 217411. Funded under FP7-SSH.* © European Union: Luxembourg. http://cordis.europa.eu/publication/rcn/14683_en.html. (Accessed 31 August 2015).

Everett DH, 1961. The thermodynamics of frost damage to porous solids. *Transactions of the Faraday Society* **57:** 1541–1551.

Feal-Pérez A, Blanco-Chao R, 2012. Characterization of abrasion surfaces in rock shore environments of NW Spain. *Geo-Marine Letters* **33**: 1–9. DOI: 10.1007/s00367-012-0300-4.

Ferreira Pinto AP, Delgado Rodrigues J, 2008. Consolidation of carbonate stones: Influence of treatment procedures on the strengthening action of consolidants. *Journal of cultural heritage* **9**: 38–53. DOI: 10.1016/j.culher.2011.07.003.

Field AP, 2009. *Discovering statistics using SPSS*. SAGE: Los Angeles, [Calif.], London.

Filzmoser P, Todorov V, 2013. Robust tools for the imperfect world. *Statistics with Imperfect Data* **245**: 4–20. DOI: 10.1016/j.ins.2012.10.017.

Fitzner B, Basten D, 1994. Gesteinporosität — Klassifizierung, messtechnische Erfassung und Bewetung ihrer Vervitterungsrelevanz. In *Jahresberichte Steinzerfall — Steinkonzervierung 1992,* Snethlage R (ed). Ernst & Sohn, Berlin; 19–32.

Fitzner B, Heinrichs K, Kownatzki R, 1997. Weathering forms at natural stone monuments: Classification, mapping and evaluation. *Restoration of buildings and monuments = Bauinstandsetzen und Baudenkmalpflege* **3**: 105–124.

Fitzner B, 2002. Damage diagnosis on stone monuments - in situ investigation and laboratory studies. In *Proceedings of the International Symposium of the Conservation of the Bangudae Petroglyph 15.07.2002, Ulsan City / Korea*, Stone Conservation Laboratory SNUSK (ed): Korea; 29–71.

Fitzner B, Heinrichs K, 2002. Damage diagnosis on stone monuments - weathering forms, damage categories and damage indices. In *Understanding and managing of stone decay: (SWAPNET 2001) [proceedings of the international conference stone weathering and atmospheric pollution network]*, Viles HA, Přikryl R (eds). Karolinum: Prague; 11–56.

Flint R, Jackson P, McCann D, 1999. Geophysical imaging inside masonry structures. *NDT & E International* **32**: 469–479.

Folk RL, 1959, Practical petrographic classification of limestones: *American Association of Petroleum Geologists Bulletin* **43**: 1-38.

Forsén H, Tarvainen V, 2000. *Accuracy and functionality of hand held wood moisture content meters*. Espoo: Technical Research Centre of Finland.

Fort R, Alvarez de Buergo M, Perez-Monserrat EM et al., 2013. Non-destructive testing for the assessment of granite decay in heritage structures compared to quarry stone. *International Journal of Rock Mechanics and Mining Sciences* **61**: 296–305. DOI:10.1016/j.ijrmms.2012.12.048

Fowler J, Cohen L, Jarvis P, 1998. *Practical statistics for field biology*. John Wiley: Chichester.

Franzen C, Mirwald PW, 2004. Moisture content of natural stone: static and dynamic equilibrium with atmospheric humidity. *Environmental Geology* **46**: 391–401. DOI: 10.1007/s00254-004-1040-1.

Franzen, C. and Mirwald, P. W. 2009. Moisture sorption behaviour of salt mixtures in porous stone. *Chemie der Erde - Geochemistry* **69**: 91–98.

Fuente D, Vega JM, Viejo F, Díaz I, Morcillo M, 2013. Mapping air pollution effects on atmospheric degradation of cultural heritage. *Journal of Cultural Heritage* **14**: 138–145. DOI: 10.1016/j.culher.2012.05.002.

Futagami Y, 2010. Joint research project for the conservation of stone monuments at the Ta Nei Temple: A 2008–2009 study on the effect of microorganisms on the stone surface. National Research Institute for Cultural Properties, Tokyo. In *16th Plenary Session - Activity reports for distribution and discussion during the - ICC Angkor: Siem Reap - 15 December 2009. English version (Original: French)*, ICC-UNESCO (ed): Phnom Penh; 50–53.

Futagami Y, Kuchitsu N, Uno T, 2008. Joint research project on conservation of stone at Ta Nei Temple in 2007. Joint research project on conservation of stone at Ta Nei Temple in 2007. In *14th Plenary Session - Activity reports for distribution and discussion during the - ICC Angkor: Siem Reap - 28 November 2007. English version (Original: French)*, ICC-UNESCO (ed): Phnom Penh; 62–63.

Garcia-Vallès M, Topal T, Vendrell-Saz M, 2003. Lichenic growth as a factor in the physical deterioration or protection of Cappadocian monuments. *Environmental Geology* **43**: 776–781.

GE Sensing - EMEA. 2006. *Calibration tables for wood: GE Protimeter*. Co. Clare: GE Sensing - EMEA.

General Electric Company. 2011. *Measure and control solutions - Surveymaster Protimeter dual-function moisture meter*. Billerica, US: General Electric Company.

General Electric Sensing. 2009. *Surveymaster, Protimeter dual-function moisture meter – Instruction manual, INS5360*, Rev.B. Billerica, US: General Electric Company.

Ghobadi MH, Torabi-Kaveh M, 2014. Assessing the potential for deterioration of limestones forming Taq-e Bostan monuments under freeze–thaw weathering and karst development. *Environmental Earth Sciences* **72**: 5035–5047. DOI: 10.1007/s12665-014-3373-8.

Giesen MJ, Mazel AD, Graham DW, Warke PA, 2011. Care and Management of Ancient Stone Monuments During Environmental Change. *International Journal of Heritage and Sustainable Development* **1**: 60–71.

Godden M, 2012. *Portland's quarries and its stone*. http://www.dorsetgeologistsassociation.com/Portland-Stone/Portland_Stone_Document_-_7_June_12.pdf (Accessed 07/08/2015).

Gomez-Heras M, Fort R, 2007. Patterns of halite (NaCl) crystallisation in building stone conditioned by laboratory heating regimes. *Environmental Geology* **52**: 259-267.

Gomez-Heras M, Smith BJ, Viles HA, 2008. Laboratory modelling of gypsum crust growth on limestone related to soot pollution and gaseous sulphur: implications of "cleaner" environments for stone decay. In *11th International Congress on Deterioration and Conservation of Stone, 15-20 September 2008, Torun, Poland: Proceedings*, Łukaszewicz JW, Niemcewicz P (eds). Uniwersytetu Mikołaja Kopernika: Toruń; 105–112.

Gomez-Heras M, Smith BJ, Viles HA, 2010. Oxford stone revisited: causes and consequences of diversity in building limestone used in the historic centre of Oxford, England. *Geological Society, London, Special Publications* **333**: 101–110. DOI: 10.1144/SP333.10.

Gomez-Heras M, McCabe S, 2015. Weathering of stone-built heritage: A lens through which to read the Anthropocene. *Anthropocene* **11**: 1–13. DOI: 10.1016/j.ancene.2015.12.003.

Gómez-Laserna O, Prieto-Taboada N, Ibarrondo I, Martínez-Arkarazo I, Olazabal MA, Madariaga JM, 2012. Raman spectroscopic characterization of brick and mortars: The advantages of the non- destructive and in situ analysis. In *Brick and mortar research*, Rivera SM, Pena Diaz, Antonio L (eds). Nova Science Publishers: Hauppauge, N.Y; 195–214.

Good PI, Hardin JW, 2009. *Common errors in statistics (and how to avoid them)*. Wiley: Hoboken, N.J.

Goudie AS, 1999. A comparison of the relative resistance of limestones to frost and salt weathering. *Permafrost and Periglacial Processes* **10**: 309–316. DOI: 10.1002/(SICI)1099-1530(199910/12)10:4<309:AID-PPP330>3.0.CO;2-C.

Goudie AS, 2006. The Schmidt Hammer in geomorphological research. *Progress in Physical Geography* **30**: 703–718. DOI: 10.1177/0309133306071954.

Granger F (Vitruvius Pollio, Marcus), 1970. *On architecture*. Heinemann W; Harvard University Press: London, Cambridge Mass.

Gray W, 1861-1862. On the Geology of the Isle of Portland: Ordinary meeting, Monday, November 4th, 1861. *Proceedings of the Geologists' Association* **1**: 128.

Grimm W, 2010. Verwendung, Verwitterung und Konservierung von Denkmalgesteinen auf alten Friedhöfen, auch mit Beispielen aus Stuttgart. In *Natursteinsanierung Stuttgart 2010: Neue Natursteinsanierungsergebnisse und messtechnische Erfassungen sowie Sanierungsbeispiele ; Tagung am 26. März 2010 in Stuttgart*, Grassegger-Schön G (ed). Fraunhofer IRB-Verlag: Stuttgart; 153–193.

Grimmer, A. E., ed. 1984. *A Glossary of Historic Masonry Deterioration Problems andPreservation Treatments*. Washington, DC: Dept. of the Interior, National ParkService, Preservation Assistance Division (US Government Printing Office).

Grossi CM, Brimblecombe P, Harris I, 2007. Predicting long term freeze–thaw risks on Europe built heritage and archaeological sites in a changing climate. *Science of The Total Environment* **377**: 273–281. DOI: 10.1016/j.scitotenv.2007.02.014.

Güney A, 2005. Evaluation of the Relationships between Schmidt Hardness Rebound Number and Other (Engineering) Properties of Rocks. In *The 19th International Mining Congress and Fair of Turkey, IMCET2005, İzmir, Turkey, June 09-12, 2005*.

Hachinohe S, Hiraki N, Suzuki T, 2000. Rates of weathering and temporal changes in strength of bedrock of marine terraces in Boso Peninsula, Japan. *Engineering Geology* **55**: 29–43. DOI: 10.1016/S0013-7952(99)00104-0.

Hack H, Hingira J, Verwaal W, 1993. Determination of discontinuity wall strength by Equotip and ball rebound tests. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* **30**: 151–155. DOI: 10.1016/0148-9062(93)90707-K.

Hack R, Huismann M, 2002. Estimating the intact rock strength of a rock mass by simple means. In *Engineering Geology for Developing Countries — Proceedings of 9th Congress of the International Association for Engineering Geology and the* Environment. Durban, South Africa (2002), Rooy JL, Jermy CA (eds). 1971–1977.

Hall K, 2004. Evidence for freeze–thaw events and their implications for rock weathering in northern Canada. *Earth Surface Processes and Landforms* **29**: 43–57. DOI: 10.1002/esp.1012.

Hall M, 2011. *Towards world heritage: International origins of the preservation movement, 1870-1930*. Burlington, VT; Ashgate: Farnham, Surrey.

Haneef SJ, Dickinson C, Johnson JB, Thompson GE, Wood GC, 1992. Simulation of the Degradation of Coupled Stones by Artificial Acid Rain. *Studies in Conservation* **37**: 105–112. DOI: 10.2307/1506402.

Hansen CD, Meiklejohn, K. Ian, Nel W, Loubser MJ, Van Der Merwe, Barend J., 2013. Aspect-controlled Weathering Observed on a Blockfield in Dronning Maud Land, Antarctica. Geografiska Annaler: Series A, *Physical Geography* **95**: 305–313. DOI: 10.1111/geoa.12025.

Haubeck K, Prinz T, 2013. A UAV-based low-cost stereo camera system for archaeological surveys - experiences from Doliche (Turkey). *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XL-1/W2**: 195–200. DOI: 10.5194/isprsarchives-XL-1-W2-195-2013.

Hauke J, Kossowski T, 2011. Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data. *Quaestiones Geographicae* **30**: 87–93. DOI: 10.2478/v10117-011-0021-1.

Hendrickx R, 2013. Using the Karsten tube to estimate water transport parameters of porous building materials. *Materials and Structures* **46**: 1309-1320. DOI: 10.1617/s11527-012-9975-2.

Hodge VJ, Austin J, 2004. A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review* **22**: 85–126. DOI: 10.1007/s10462-004-4304-y.

Hodges, JL. Jr., Lehmann, EL., 1963. Estimates of Location Based on Rank Tests. *The Annals of Mathematical Statistics* **34**: 598–611. DOI: 10.2307/2238406.

Hoke GD, Turcotte DL, 2004. The weathering of stones due to dissolution. *Environmental Geology* **46**:305–310**.** DOI: 10.1007/s00254-004-1033-0.

Honeybourne, DB, Harris PB 1958. The structure of porous building stone and its relation to weathering behaviour. In *Proceedings Tenth Symposium of the Colston,* Research Society; 343-363.

Howard AJ, 2013. Managing global heritage in the face of future climate change: The importance of understanding geological and geomorphological processes and hazards. *International Journal of Heritage Studies* **19**: 632–658. DOI: 10.1080/13527258.2012.681680.

Howell J, 1995. *Moisture measurement in masonry: Guidance for surveyors. COBRA 1995*. London: Royal Institution of Chartered Surveyors Research.

Huang C, 2012. Jurassic. In *The Geologic Time Scale 2012 2-Volume Set*, Ogg, James G. (James George), 1952-, Schmitz M, Ogg G (eds). Elsevier Science; 731–791.

Huber PJ, 1981. *Robust statistics*. Wiley: New York.

ICOMOS (International Council on Monuments and Sites), 1964. *Venice-Charter. International charter for the conservation and restoration of monuments and sites: Second international congress of architects and technicians of historic monuments*. ICOMOS: Venice.

Iglewicz B, Hoaglin DC, 1993. *How to detect and handle outliers*. ASQC Quality Press: Milwaukee, Wis.

Ingham JP, 2005. Predicting the frost resistance of building stone. *Quarterly Journal of Engineering Geology and Hydrogeology* **38**: 387–399. DOI: 10.1144/1470-9236/04-068.

Inkpen RJ, Jackson J, 2000. Contrasting weathering rates in coastal, urban and rural areas in southern Britain: preliminary investigations using gravestones. *Earth Surface Processes and Landforms* **25**: 229–238. DOI: 10.1002/(SICI)1096-9837(200003)25:3<229:AID-ESP52>3.0.CO;2-Y.

Inkpen RJ, Viles H, Moses C, Baily B, 2012a. Modelling the impact of changing atmospheric pollution levels on limestone erosion rates in central London, 1980–2010. *Atmospheric Environment* **61**: 476–481. DOI: 10.1016/j.atmosenv.2012.07.042.

Inkpen RJ, Viles H, Moses C, Baily B, Collier P, Trudgill ST, Cooke RU, 2012b. Thirty years of erosion and declining atmospheric pollution at St Paul's Cathedral, London. *Atmospheric Environment* **62**: 521–529. DOI: 10.1016/j.atmosenv.2012.08.055.

Jackson M, Marra F, Hay RL, Cawood C, Winkler EM, 2005. The judicious selection and preserveration of tuff and travertine building stone in ancient Rome. *Archaeometry* 47: 485–510. DOI: 10.1111/j.1475-4754.2005.00215.x.

Jaynes SM, Cooke RU, 1987. Stone weathering in Southeast England. *Atmospheric Environment* 21: 1601–1622. DOI: 10.1016/0004-6981(87)90321-0.

Jha AK, Duyne JE, 2010. *Safer homes, stronger communities: A handbook for reconstructing after natural disasters.* World Bank: Washington, DC.

Kahraman S, Gunaydin O, 2007. Empirical methods to predict the abrasion resistance of rock aggregates. *Bulletin of Engineering Geology and the Environment* **66**: 449–455. DOI: 10.1007/s10064-007-0093-2.

Kamh GME, Koltuk S, 2015. Micro-topographic and Geotechnical Investigations of sandstone Wall on Weathering Progress, Aachen City, Germany, case study. *Arabian Journal for Science and Engineering* **41**: 2285-2294. DOI: 10.1007/s13369-015-1887-3.

Kaymakci N, Inceoz M, Ertepinar P, Koc A, 2010. Late Cretaceous to Recent kinematics of SE Anatolia (Turkey). *Geological Society, London, Special Publications* **340**: 409–435. DOI: 10.1144/SP340.18.

Kelley K, 2005. The effects of nonnormal distributions on confidence intervals around the standardized mean difference. *Educational and Psychological Measurement* **65**: 51–69. DOI: 10.1177/0013164404264850.

Kemp M, 2006. *Leonardo da Vinci – The marvellous works of nature and man.* Oxford University Press: Oxford; New York.

Keselmann H, 1998. Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA and ANCOVA analysis. *Review of Educational Research* **68**: 350–386.

Kilic G, 2015. Using advanced NDT for historic buildings: Towards an integrated multidisciplinary health assessment strategy. *Journal of Cultural Heritage* **16**: 526–535. DOI: 10.1016/j.culher.2014.09.010.

Klopfer H, 1985. *Lehrbuch der Bauphysik*. Teubner: Stuttgart.

Klopfer, H., 1979. *Das Kapillarverhalten poröser Baustoffe, Technische Akademie Esslingen,Lehrgang Nr. 4280/79.88*, Klopfer H, Bauphysikalische Aspekte-Anforderungen an Sanierungen, TechnischeAkademie Esslingen, Lehrgang Nr. 4406/79.88, (1980).

Klopfer H, Homman M, 2008. 'Feuchte', in *Lehrbuch der Bauphysik*, 6th ed.,Vieweg+Teubner Verlag, Wiesbaden. 338-490.

Koenker R, 2005. *Quantile regression*. Cambridge University Press: Cambridge, New York.

Kraska S, 1998. *Zerstörungsfreie Erfassung und Bewertung von Salzgehalten in Baustoffen mit Hilfe der Wenner Methode. Diplomarbeit* (eq. Master thesis). Weimar: Bauhaus-Universität Weimar. Fakultät Bauingenieurwesen. Professur Bauchemie.

Kucera V, 2002. *Dose-response functions as a basis for assessment of acceptable levels*. Schwedish Corrosion Institute: Stockholm.

Leary E, 1983. *The building limestones of the British Isles. Dept. of The Environment*, Building Research Establishment; H.M.S.O.: Garston, Watford, London.

Lednicá M, Kaláb Z, 2012. Evaluation of granite weathering in the Jeroným mine using non-destructive methods. *Acta Geodyn. Geomater* **9**: 211–220.

Leucci G, Masini N, Persico R, 2012. Time–frequency analysis of GPR data to investigate the damage of monumental buildings. *Journal of Geophysics and Engineering* **9**: 81-91. DOI: 10.1088/1742-2132/9/4/S81.

Lewry AJ, Asiedu-Dompreh J, Bigland DJ, Butlin RN, 1994. The effect of humidity on the dry deposition of sulfur dioxide onto calcareous stones. *Construction and Building Materials* **8**: 97–100. DOI: 10.1016/S0950-0618(09)90018-6.

Leys C, Ley C, Klein O, Bernard P, Licata L, 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology* **49**: 764–766. DOI: 10.1016/j.jesp.2013.03.013.

Linnow K, Juling H, Steiger M, 2007. Investigation of NaCl deliquescence in porous substrates using RH-XRD. *Environmental Geology* **52**: 317-327.

Lipfert FW, 1989. Atmospheric damage to calcareous stones: Comparison and reconciliation of recent experimental findings. *Atmospheric Environment* **23**: 415–429. DOI: 10.1016/0004-6981(89)90587-8.

Lix, Lisa M., Keselman Joanne C., and Keselman H. J. "Consequences of Assumption Violations Revisited: A Quantitative Review of Alternatives to the One-Way Analysis of Variance "F" Test." *Review of Educational Research* **66**: 579-619. http://www.jstor.org/stable/1170654.

Logan JM, Hastedt M, Lehnert D, Denton M, 1993. Variations of rock properties within a quarry. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* **30**: 1527–1530. DOI: 10.1016/0148-9062(93)90151-3.

Loke MH, 1999. *Electrical imaging surveys for environmental and engineering studies—a practical guide to 2-D and 3-D surveys.* Penang, Malaysia: Loke M.H.

Lyalikova NN, Petushkova YP, 1991. Role of microorganisms in the weathering of minerals in building stone of historical buildings. *Geomicrobiology Journal* **9**: 91–101. DOI: 10.1080/01490459109385991.

Maierhofer C, 2010. *Monitoring - Zuverlässige Quantifizierung und Bewertung von Schädigungsprozessen an Bauteiloberflächen und -grenzflächen mit Hilfe optischer und thermografischer zerstörungsfreier Prüfverfahren // Monitoring - zuverlässige Quantifizierung und Bewertung von Schädigungsprozessen an Bauteiloberflächen und -grenzflächen mit Hilfe optischer und thermografischer zerstörungsfreier Prüfverfahren*. Fraunhofer-IRB-Verlag: Stuttgart.

Maierhofer C, Wöstmann J, 1998. Investigation of dielectric properties of brick materials as a function of moisture and salt content using a microwave impulse technique at very high frequencies. *NDT & E International* **31**: 259–263.

Martinez A, Byrnes A, 2001. Modeling dielectric-constant values of geologic materials: An aid to ground-penetrating radar data collection and interpretation. *Current Research in Earth Sciences, Bulletin* **247**(Part1): 1–16.

Martínez-Martínez J, Benavente D, Gomez-Heras M, Marco-Castaño L, García-del-Cura MÁ, 2013. Non-linear decay of building stones during freeze–thaw weathering processes: 25th Anniversary Session for ACI 228 – Building on the Past for the Future of NDT of Concrete. *Construction and Building Materials* **38**: 443–454. DOI: 10.1016/j.conbuildmat.2012.07.059.

Massey SW, 1999. The effects of ozone and NOx on the deterioration of calcareous stone. *Science of The Total Environment* **227**: 109–121. DOI: 10.1016/S0048-9697(98)00409-4.

Matsui T, Sawada M, Inoue S, Ebisawa T, Kawasaki E, Atomi Y, Nakagawa T, 2014. *A Study on Conservative Material for the Bas-Reliefs of the Bayon Temple in Angkor Monuments. In Proceedings of the International Conference on Conservation of Stone and Earthen Architectural Heritage: July 2014 Gongju, Republic of Korea*, Lee CH, Kim J, Kim RH (eds). ICOMOS-ISCS; Kongju National University: Gongju; 119–126.

Matthews JA, Geraint O, 2008. Endolithic lichens, rapid biological weathering and Schmidt Hammer R-values on recently exposed rock surfaces: Storbreen

Glacier Foreland, Jotunheimen, Norway. Geografiska Annaler: Series A, *Physical Geography* **90**: 287–297. DOI: 10.1111/j.1468-0459.2008.00346.x.

May A, 1998. Die Verwitterung von karbonatischen Naturbaustein in Mitteleuropa. *Geoökodynamik* **19**: 71–88.

McCabe S, Brimblecombe P, Smith BJ, McAllister D, Srinivasan S, Basheer PAM, 2013. The use and meanings of 'time of wetness' in understanding building stone decay. *Quarterly Journal of Engineering Geology and Hydrogeology* **46**: 469–476. DOI: 10.1144/qjegh2012-048.

McCabe S, McAllister D, Warke PA, Gomez-Heras M, 2015. Building sandstone surface modification by biofilm and iron precipitation: emerging block-scale heterogeneity and system response. *Earth Surface Processes and Landforms* **40**: 112–122. DOI: 10.1002/esp.3665.

McCann D, Forde M, 2001. Review of NDT methods in the assessment of concrete and masonry structures. *NDT & E International* **34**: 71–84. DOI: 10.1016/S0963-8695(00)00032-3.

McCarroll D, 1991. The Schmidt Hammer, weathering and rock surface roughness. *Earth Surface Processes and Landforms* **16**: 477-480. DOI: 10.1002/esp.3290160510.

McGreevy JP, Warke PA, Smith BJ, 2000. Controls on Stone Temperatures and the Benefits of Interdisciplinary Exchange. *Journal of the American Institute for Conservation* **39**: 259. DOI: 10.2307/3180095.

McIlroy de la Rosa JP, Warke PA, Smith B, 2014. The effects of lichen cover upon the rate of solutional weathering of limestone. *Geomorphology* **220**: 81–92. DOI: 10.1016/j.geomorph.2014.05.030.

Meierding TC, 1993a. Marble Tombstone Weathering and Air Pollution in North America. *Annals of the Association of American Geographers* **83**: 568–588. DOI: 10.1111/j.1467-8306.1993.tb01954.x.

Meierding TC, 1993b. Inscription legibility method for estimating rock weathering rates. *Geomorphology* **6**: 273–286. DOI: 10.1016/0169-555X(93)90051-3.

Meinhardt-Degen J, 2005. *Geologisch mineralogische und materialtechnische Untersuchungen zur Risikoabschätzung von Folgekonservierungen bei Sandsteinen am Bespiel von Regensburger Grünssandstein und Grünem Mainsandstein*. PhD thesis. Technical University: München.

Meinhardt-Degen J, 2011. Leitfaden Naturstein-Monitoring - Methoden zur Bewertung von Konservierungsmaßnahmen. In *Einfache, zerstörungsfreie Prüfverfahren: Beiträge des 5. Konservierungswissenschaftlichen Kolloquiums in Berlin / Brandenburg am 18. November 2011 in Potsdam*, Karg D (ed). Imhof: Petersberg; 68–79.

Meinhardt-Degen J, Snethlage R, Auras M, 2012. Natural stone monitoring - Investigation methods for a reliable evaluation of the effectiveness of conservation measures. In *12th International Congress on the Deterioration and Conservation of Stone. New York 2012*, Wheeler G (ed): New York.

Meneely J, Smith BJ, Curran J et al., 2009. *Developing a 'Non-destructive scientific toolkit' to monitor monuments and sites. ICOMOS Scientific Symposium: Changing World, Changing Views of Heritage*. http://www.international.icomos.org/adcom/malta2009/pdf/ADCOM_200910_SYMP_1_Documentation_John_Meneely.pdf.

Met Office. www.metoffice.gov.uk/ (accessed 10/08/2015)

Missouri Department of Natural Resources, Division of Geology and Land Survey, 2011. Definition and Classification of Limestone. http://dnr.mo.gov/geology/geosrv/imac/limestone.htm. (Accessed 25/09/2015).

Mitchell DJ, Searle DE, 2004. *Stone deterioration in polluted urban environments*. Science Publishers: Enfield, NH.

Mol L, Viles HA, 2012. The role of rock surface hardness and internal moisture in tafoni development in sandstone. *Earth Surface Processes and Landforms* **37**: 301–314. DOI: 10.1002/esp.2252.

Montgomery DR, 1999. Process domains and the river continuum. *Journal of the American Water Resources Association* **35**: 397–410. DOI: 10.1111/j.1752-1688.1999.tb03598.x.

Mooney CZ, Duval RD, 1993. *Bootstrapping: A nonparametric approach to statistical inference*. Sage Publications: Newbury Park, Calif.

Moore L, Byers JaBD, 1989. *Statistical test of reproducibility and operator variance in thin-section modal analysis of textures and phenocrysts in the Topopah Spring Member, drill hole USW VH-2, Crater Flat, Nye County, Nevada (LA-11452-MS)*. (Report). Los Alamos National Laboratory: Los Alamos.

Moos Av, Quervain Fd, 1948. *Technische Gesteinskunde*. Birkhäuser: Basel.

Moroni B, Pitzurra L, 2008. Biodegradation of atmospheric pollutants by fungi: A crucial point in the corrosion of carbonate building stone. *International Biodeterioration & Biodegradation* **62**: 391–396. DOI: 10.1016/j.ibiod.2008.03.011.

Mosch S, Siegesmund S, 2007. Petrophysical and technical properties of dimensional stones: a statistical approach. *Zeitschrift der Deutschen Gesellschaft für Geowissenschaften* **158**: 821–868. DOI: 10.1127/1860-1804/2007/0158-0821.

Moses C, Robinson D, Barlow J, 2014. Methods for measuring rock surface weathering and erosion: A critical review. *Earth-Science Reviews* **135**: 141–161. DOI: 10.1016/j.earscirev.2014.04.006.

Mottershead D, Gorbushina A, Lucas G, Wright J, 2003. The influence of marine salts, aspect and microbes in the weathering of sandstone in two historic structures. *Building and Environment* **38**: 1193–1204. DOI: 10.1016/S0360-1323(03)00071-4.

Muggeo VMR, 2003. Estimating regression models with unknown break-points. *Statistics in medicine* **22**: 3055–3071. DOI: 10.1002/sim.1545.

Muggeo VMR, 2008. Segmented: an R package to fit regression models with broken-line relationships. *R News* **8**: 20–25.

Myrin MMK, 2006. A case study on the evaluation of consolidation treatments of Gotland sandstone by use of ultrasound pulse velocity measurements. In *Heritage, weathering and conservation: Proceedings of the International Conference on Heritage, Weathering and Conservation, HWC-2006, 21-24 June 2006, Madrid, Spain*, Fort R (ed). Taylor & Francis: London, New York.

Nava S, Becherini F, Bernardi A, Bonazza A, Chiari M, García-Orellana I, Lucarelli F, Ludwig N, Migliori A, Sabbioni C, Udisti R, Valli G, Vecchi R, 2010. An integrated approach to assess air pollution threats to cultural heritage in a semi-confined environment: The case study of Michelozzo's Courtyard in Florence (Italy). *Science of The Total Environment* **408**: 1403–1413. DOI: 10.1016/j.scitotenv.2009.07.030.

Neisel JD, *Rasterelektronenmikroskopische Untersuchung der Gefügemerkmale von ungeschützten und imprägnierten Sandsteinen nach Verwitterungsbeanspruchung*, PhD thesis, Rheinisch-Westfälische Technische Hochschule Aachen (1995).

Nicholson DT, 2001. Pore properties as indicators of breakdown mechanisms in experimentally weathered limestones. *Earth Surface Processes and Landforms* **26**: 819–838. DOI: 10.1002/esp.228.

Niedzielski T, Migoń P, Placek A, 2009. A minimum sample size required from Schmidt hammer measurements. *Earth Surface Processes and Landforms* **34**: 1713–1725. DOI: 10.1002/esp.1851.

Niemeyer R, 2013. *Bestimmung der kapillaren Wasseraufnahme mit dem Prüfröhrchen nach Karsten*. Niedersächsisches Landesamt für Denkmalpflege.

Noiriel C, Renard F, Doan M, Gratier J, 2010. Intense fracturing and fracture sealing induced by mineral growth in porous rocks. *Chemical Geology* **269**: 197–209. DOI: 10.1016/j.chemgeo.2009.09.018.

O'Brien P, Bell E, Orr T, Cooper TP, 1995. Stone loss rates at sites around Europe. *The Science of the Total Environment* **167**: 111–121.

Osborn G, McCarthy D, LaBrie A, Burke R, 2015. Lichenometric dating: Science or pseudo-science? *Quaternary Research* **83**: 1–12. DOI: 10.1016/j.yqres.2014.09.006.

Özvan A, Dinçer İ, Akın M, Oyan V, Tapan M, 2015. Experimental studies on ignimbrite and the effect of lichens and capillarity on the deterioration of Seljuk Gravestones. *Engineering Geology* **185**: 81–95. DOI: 10.1016/j.enggeo.2014.12.001.

Palmer T, 2008. Limestone petrography and durability in English Jurassic Freestones. In *England's heritage in stone: Proceedings of a conference, Tempest Anderson Hall, York, 15-17 March, 2005*, Doyle P (ed). English Stone Forum: Folkestone, Kent; 66–78.

Pamplona Mea, 2008. Drilling resistance; the state of art and future developments. In *11th International Congress on Deterioration and Conservation of Stone, 15-20 September 2008, Torun, Poland: Proceedings*, Łukaszewicz JW, Niemcewicz P (eds). Uniwersytetu Mikołaja Kopernika: Toruń.

Paradise TR, 1998. Limestone weathering and rate variability, Great Temple of Amman, Jordan. *Physical Geography* **19**: 133–146.

Park HM, 2008. *Univariate Analysis and Normality Test Using SAS, Stata, and SPSS. Technical Working Paper*. The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University.

Petzet M, 2010. *International principles of preservation*. Bäßler: Berlin.

Pierce M, Yu D, Richards-Kortum R, 2011. High-resolution Fiber-optic microendoscopy for in situ cellular imaging. *Journal of Visualized Experiments* **47**:2306-2311. DOI: 10.3791/2306.

Pinchin, SE,2008. Techniques for monitoring moisture in walls. *Reviews in conservation* **9**:33–45.

Pollok H, Mennicke RT, 2010. Using Equotip hardness test blocks. www.proceq.com. Accessed 09/2015.

Pope GA, 2000. Weathering of petroglyphs: direct assessment and implication for dating methods. *Antiquity* **74**: 833–843.

Pope GA, Meierding TC, Paradise TR, 2002. Geomorphology's role in the study of weathering of cultural stone. *Geomorphology* **47**: 211–225. DOI: 10.1016/S0169-555X(02)00098-3.

Posas PJ, 2011. Exploring climate change criteria for strategic environmental assessments. *Progress in Planning* **75**: 109–154. DOI: 10.1016/j.progress.2011.05.001.

Poschlod K, 1990. *Das Wasser im Porenraum kristalliner Naturwerksteine und sein Einfluss auf die Verwitterung*. F. Pfeil: München.

Powell J, 1996. Estimation of semiparametric models. In Handbook of econometrics. North-Holland: Amsterdam; 2444–2514.

Price CA, 2000. An expert chemical model for determining the environmental conditions needed to prevent salt damage in porous materials: Technologies to protect and rehabilitate *European culture heritage final report*. Archetype: London.

Price CA, 2007. Predicting environmental conditions to minimise salt damage at the Tower of London: a comparison of two approaches. *Environmental Geology* **52**: 369–374. DOI: 10.1007/s00254-006-0477-9.

Prieto-Taboada N, Ibarrondo I, Gómez-Laserna O, Martinez-Arkarazo I, Olazabal MA, Madariaga JM, 2013. Buildings as repositories of hazardous pollutants of anthropogenic origin. *Journal of Hazardous Materials* **248**-249: 451–460. DOI: 10.1016/j.jhazmat.2013.01.008.

Prikryl R, 2013. Durability assessment of natural stone. *Quarterly Journal of Engineering Geology and Hydrogeology* **46**: 377–390. DOI: 10.1144/qjegh2012-052.

Proceq© SA, 2006. Concrete Test Hammer – Operating instructions. www.proceq.com. Accessed 07/2015.

Proceq© SA, 2010. Operating instructions portable metal hardness tester. www.proceq.com. Accessed 1. September 2012.

Quervain F, 1967. *Technische Gesteinskunde*. Birkhäuser Basel; Imprint: Birkhäuser: Basel.

Razali NM, Wah YB, 2011. Power comparisons of Shapiro-Wilk, Kolgomorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics* **2**: 21–33.

Reimann C, 2008. *Statistical data analysis explained: Applied environmental statistics with R*. John Wiley & Sons: Chichester, England, Hoboken, NJ.

Reunion Internationale des Laboratoires D'Essais et de Recherches sur les Materiaux et les Constructions (RILEM), 1980. Recommandations provisoires de la commission 25-PEM Protection et érosion des monuments. Essais recommandés pour mesurer l'altération des pierre et évaluer l'efficacité des méthodes de traitment. *RILEM: Paris*. (vol. **13** nr. 75).

Robertson A, Boulton SJ, Taslı K, Yıldırım N, İnan N, Yıldız A, Parlak O, 2015. Late Cretaceous-Miocene sedimentary development of the Arabian continental margin in SE Turkey (Adıyaman Region): implications for regional palaeogeography and the closure history of Southern Neotethys. *Journal of Asian Earth Sciences* (in press). DOI: 10.1016/j.jseaes.2015.01.025.

Rodriguez-Navarro C, Doehne E, Sebastian E, 2000. How does sodium sulfate crystallize? Implications for the decay and testing of building materials. *Cement and Concrete Research* **30**: 1527–1534. DOI: 10.1016/S0008-8846(00)00381-1.

Rose, D.A., ‘Water movement in unsaturated porous materials’, *RILEM Bulletin* 29 (1965) 119-124.

Rosner B, 1983. Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics* **25**: 165–172. DOI: 10.2307/1268549.

Ross KD, Butlin RN, 1989. Durability tests for building stone. *Building Research Establishment*: Garston.

Ross KD, Hart D, Butlin RN, 1991. Durability testsfor natural building stone. *In: Baker* JM, Nixon PJ,Majudmar J, Davies H, (eds). *Durability of BuildingMaterials and Components, Proceedings of the Fifth International Conference.* Chapman and Hall, London, 97–111.

Ruddiman WF, 2010. *Plows, plagues, and petroleum: How humans took control of climate*. Princeton University Press: Princeton, NJ.

Sabbioni C, Brimblecombe P, Cassar M et al., 2010. *The atlas of climate change impact on European cultural heritage: Scientific analysis and management strategies*. Anthem: London, New York.

Sass O, Viles H, 2010. 2D resistivity surveys of the moisture contents of historic limestone walls in Oxford UK Implications for understanding catastrophic stone deterioration. In *Limestone in the built environment: Present-day challenges for the preservation of the past*, Smith BJ (ed). Geological Society: London.

Schön J, 2011. *Physical properties of rocks: A workbook*. Elsevier: Oxford.

Schön J, 2015. *Physical properties of rocks: Fundamentals and principles of petrophysics*. Elsevier: Amsterdam, Netherlands.

Schuh H, Grüner F, Köhler W, Wendler E, 2011. Vergleichsuntersuchungen mit unterschiedlichen zerstörungsfreien und zerstörungsarmen Feuchte- und Salzmessmethoden In *Einfache, zerstörungsfreie Prüfverfahren: Beiträge des 5. Konservierungswissenschaftlichen Kolloquiums in Berlin / Brandenburg am 18. November 2011 in Potsdam*, Karg D (ed). Petersberg: Imhof.

Schwarz H, Was sind Salze? http://193.175.110.91/salzwiki/index.php/Was_sind_Salze%3F. (Accessed 20 September 2015).

Searle DE, Mitchell DJ, 2006. The effect of coal and diesel particulates on the weathering loss of Portland Limestone in an urban environment. *Science of The Total Environment* **370**: 207–223. DOI:10.1016/j.scitotenv.2006.07.005.

Seaward M, 2001. The role of lichens in the biodeterioration of ancient monuments with particular reference to central Italy. *International Biodeterioration & Biodegradation* **48**: 202–208. DOI: 10.1016/S0964-8305(01)00082-8.

Sharp AD, Trudgill ST, Cooke RU, Price CA, Crabtree RW, Pickles AM, Smith DI, 1982. Weathering of the balustrade on St. Paul's Cathedral, London. *Earth Surface Processes and Landforms* **7**: 387–389. DOI: 10.1002/esp.3290070410.

Siedel H, Plehwe-Leisen E von, Leisen H, 2008. Salt load and deterioration of sandstone at the temple of Angkor Wat, Cambodia. In *11th International Congress on Deterioration and Conservation of Stone, 15-20 September 2008, Torun, Poland: Proceedings*, Łukaszewicz JW, Niemcewicz P (eds). Uniwersytetu Mikołaja Kopernika: Toruń; 267–274.

Siedel H, Siegesmund S, 2010. Characterisation of stone deterioration on buildings. In *Stone in Architecture: Properties, Durability,* Siegesmund S, Snethlage R, Winkler E (eds). Springer: Berlin; 347–410.

Siegesmund S, Kirchner D, 2003. Frost-Tau-Verhalten ausgewählter Naturwerksteine. *Naturstein* **2**: 52–56.

Siegesmund S, Török A, Hüpers A, Müller C, Klemm W, 2007. Mineralogical, geochemical and microfabric evidences of gypsum crusts: a case study from Budapest. *Environmental Geology* **52**: 385–397. DOI:10.1007/s00254-006-0588-3.

Siegesmund S, Dürrast H, 2010. Physical and mechanical properties of rocks. In *Stone in Architecture: Properties, Durability*, Siegesmund S, Snethlage R, Winkler E (eds). Springer: Berlin; 97–225.

Siegesmund S, Snethlage R, Winkler E, 2010. *Stone in Architecture: Properties, Durability*. Springer: Berlin.

Simon S, 2001. *Zur Verwitterung und Konservierung kristallinen Marmors: Untersuchungen zu physiko-mechanischen Gesteinskennwerten, zur Oberflächenchemie von Calcit und zur Anpassung und Überprüfung von Gesteinsschutzmitteln.* PhD thesis. Ludwig-Maximilians-Universität: München.

Smith BJ, Török A, McAlister JJ, Megarry Y, 2003. Observations on the factors influencing stability of building stones following contour scaling: a case study of oolitic limestones from Budapest, Hungary. *Building Stone Decay: Observations, Experiments and Modeling* **38**: 1173–1183. DOI: 10.1016/S0360-1323(03)00076-3.

Smith BJ, Viles H, 2006. Rapid, catastrophic stone decay of building limestones: thoughts and causes, effects and consequences. In *Proceedings of the International Conference on Heritage, Weathering and Conservation, HWC-2006, 21-24 June 2006*, Madrid, Spain. Taylor & Francis: London; New York.

Smith BJ, Prikryl R, 2007. Diagnosing decay: the value of medical analogy in understanding the weathering of building stones. *Geological Society, London, Special Publications* **271**: 1–8. DOI: 10.1144/GSL.SP.2007.271.01.01.

Smith BJ, Gomez-Heras M, McCabe S, 2008. Understanding the decay of stone-built cultural heritage. *Progress in Physical Geography* **32**: 439–461. DOI: 10.1177/0309133308098119.

Smith BJ, Gomez-Heras M, Viles HA, 2010. Underlying issues on the selection, use and conservation of building limestone. *Geological Society, London, Special Publications* **331**: 1–11. DOI: 10.1144/SP331.1.

Smith BJ, McCabe S, McAllister D, Adamson C, Viles HA, Curran JM, 2011a. A commentary on climate change, stone decay dynamics and the 'greening' of natural stone buildings: new perspectives on 'deep wetting'. *Environmental Earth Sciences* **63**: 1691–1700. DOI: 10.1007/s12665-010-0766-1.

Smith BJ, Srinivasan S, Gomez-Heras M, Basheer PAM, Viles HA, 2011b. Near-surface temperature cycling of stone and its implications for scales of surface deterioration: Scale Issues in Geomorphology. *Geomorphology* **130**: 76–82. DOI: 10.1016/j.geomorph.2010.10.005.

Snethlage R, 1984. *Steinkonservierung - Forschungsprogram des Zentrallabors für Denkmalpflege 1979-1983: Bericht für die Stiftung des Volkswagenwerk*. Bayerisches Landesamt für Denkmalpflege: Munich.

Snethlage R, Wendler E, 1997. *Moisture cycles and sandstone degradation. In Saving our architectural heritage: The conservation of historic stone structures report of the Dahlem Workshop on Saving our Architectural Heritage-the Conservation of Historic Stone Structures, Berlin, March 3-8, 1996*, Baer NS, Snethlage R (eds). J. Wiley: Chichester, New York.

St. Clair L, Seaward MRD, 2004. *Biodeterioration of stone surfaces: Lichens and biofilms as weathering agents of rocks and cultural heritage*. Kluwer Academic Publishers: Dordrecht, London.

Stahl T, Winkler S, Quigley M, Bebbington M, Duffy B, Duke D, 2013. Schmidt hammer exposure-age dating (SHD) of late Quaternary fluvial terraces in New Zealand. *Earth Surface Processes and Landforms* **38**: 1838–1850. DOI: 10.1002/esp.3427.

Steiger M, 2004. Influence of salts on the freezing temperature of water: implications on frost damage to porous materials. In *Proceedings of the 10th International Congress on Deterioration and Conservation of Stone, Stockholm June 27 - July 2, 2004*, Kwiatkowski D, Löfvendahl R (eds). ICOMOS, Sweden: Stockholm; 179–186.

Steiger M, Linnow K, Juling H, Gülker G, Jarad AE, Brüggerhoff S, Kirchner D, 2008. Hydration of MgSO 4 ·H 2 O and Generation of Stress in Porous Materials. *Crystal Growth & Design* **8**: 336–343. DOI: 10.1021/cg060688c.

Steiger M, Charola AE, Sterflinger K, 2010. Weathering and deterioration In *Stone in Architecture: Properties, Durability*. Edited by S. Siegesmund, R. Snethlage, and E. Winkler. Berlin: Springer.

Stephenson WJ, Kirk RM, Hemmingsen SA, Hemmingsen MA, 2010. Decadal scale micro erosion rates on shore platforms. *Geomorphology* **114**: 22–29. DOI: 10.1016/j.geomorph.2008.10.013.

Stone Federation of Great Britain, 1991. *Natural Stone Glossary*. Linked Advertising & Marketing: London.

Svahn H, 2006. *Non-destructive Field Tests in Stone Conservation*. Riksantikvarieämbetet: Stockholm.

Taber S, 1929. Frost heaving. *The Journal of Geology* **37:** 428–461.

Taber S, 1930. The mechanics of frost heaving. *The Journal of Geology* **38:** 303–317.

Tapete D, Fanti R, Cecchi R, Petrangeli P, Casagli N, 2012. Satellite radar interferometry for monitoring and early-stage warning of structural instability in archaeological sites. *Journal of Geophysics and Engineering* **9**: 10. DOI: 10.1088/1742-2132/9/4/S10.

Tegethoff FW, 2001. *Calcium Carbonate*. Birkhäuser Basel: Basel.

The Churches Conservation Trust (CCT), www.visitChurches.org.uk (accessed 07/2015)

Thomas R, Rawlinson G, 1997. *The histories / Herodotus*; With an introduction by Rosalind Thomas; Transl. by George Rawlinson. Everyman's library: London.

Tiryaki B, 2008. Predicting intact rock strength for mechanical excavation using multivariate statistics, artificial neural networks, and regression trees. *Engineering Geology* **99**: 51–60. DOI: 10.1016/j.enggeo.2008.02.003.

Török Á, 2003. Surface strength and mineralogy of weathering crusts on limestone buildings in Budapest. *Building Stone Decay: Observations, Experiments and Modeling* **38**: 1185–1192. DOI: 10.1016/S0360-1323(03)00072-6.

Török Á, 2007. Characteristics and morphology of weathering crusts on porous limestone, the role of climate and air pollution. In *Preservation of Natural Stone and Rock Weathering*, Olalla C, Estaire J, Sola P (eds). Taylor & Francis; 61–66.

Török Á, 2008. Black crusts on travertine: factors controlling development and stability. *Environmental Geology* **56**: 583–594. DOI: 10.1007/s00254-008-1297-x.

Townson WG, 1975. Lithostratigraphy and deposition of the type Portlandian. *Journal of the Geological Society* **131**: 619–638. DOI: 10.1144/gsjgs.131.6.0619.

Trotman PM, Sanders C, Harrison HW, 2004. *Understanding dampness*. Watford: BRE Bookshop.

Trudgill ST, Viles HA, Inkpen RJ, Cooke RU, 1989. Remeasurement of weathering rates, St. Paul's Cathedral, London. *Earth Surface Processes and Landforms* **14**: 175–196. DOI: 10.1002/esp.3290140302.

Trudgill ST, Viles HA, Cooke RU, Inkpen RJ, Heathwaite AL, Houston J, 1991. Trends in stone weathering and atmospheric pollution at St Paul's Cathedral, London 1980–1990. *Atmospheric Environment. Part A. General Topics* **25**: 2851–2853. DOI: 10.1016/0960-1686(91)90210-X.

Trudgill ST, Viles HA, 1998. Field and laboratory approaches to limestone weathering. *Quarterly Journal of Engineering Geology and Hydrogeology* **31**: 333–341. DOI: 10.1144/GSL.QJEG.1998.031.P4.06.

Trudgill ST, Ferguson RI, Ball J, Crabtree RW, Gent R, 1994. Ten year remeasurement of chemical denudation on a magnesian limestone hillslope. *Earth Surface Processes and Landforms* **19**: 109–114. DOI: 10.1002/esp.3290190203.

Trudgill ST, Viles HA, Inkpen R, Moses C, Gosling W, Yates T, Collier P, Smith DI, Cooke RU, 2001. Twenty-year weathering remeasurements at St Paul's Cathedral, London. *Earth Surface Processes and Landforms* **26**: 1129–1142. DOI: 10.1002/esp.260.

Tucker ME, Wright VP, 1990. Carbonate Sedimentology. Blackwell Publishing Ltd: Oxford, UK.

Tuğrul A, 2004. The effect of weathering on pore geometry and compressive strength of selected rock types from Turkey. *Engineering Geology* **75**: 215–227. DOI: 10.1016/j.enggeo.2004.05.008.

Tukey JW, 1977. *Exploratory data analysis*. Addison-Wesley Pub. Co.: Reading (Mass.).

Turkington AV, Paradise TR, 2005. Sandstone weathering: a century of research and innovation. *Geomorphology* **67**: 229–253. DOI: 10.1016/j.geomorph.2004.09.028.

Türkkan A, 2011. Yavuzelý - Araban (Gazýantep) dolayinin stratýgrafýsý ve Fırat formasyonu resýfal kýrectaplarinin dodal yapi malzemesý olarak kullanilabýlýrlýdýnýn arabtirilmasi. MSc thesis. Çukurova Üniversitesi Fen Býlýmlerý Enstýtüsü (Çukurova university) (Department of Geological Engineering) Adana. http://library.cu.edu.tr/tezler/8383.pdf

UNI (Ente nazionale italiano di unificazione/Italian Organization for Standardization). 2006. UNI 11182 *Beni culturali: Materiali lapidei naturali ed artificiali; Descrizione della forma di alterazione; Termini e definizioni = Cultural heritage: Natural and artificial stone; Description of the alteration; Terminology and definition.* Milan and Rome: UNI (Ente nazionale italiano di unificazione).

Upreti DK, Divakar PK, Shukla V et al., 2015. *Recent Advances in Lichenology: Modern Methods and Approaches in Biomonitoring and Bioprospection, Volume 1*. Springer: New Dehli, Heidelberg, New York, Dordrecht, London.

Uraibi HS, Midi H, Talib BA, Yousif JH, 2009. Linear Regression Model Selection Based on Robust Bootstrapping Technique. *American Journal of Applied Sciences* **6**: 1191–1198. DOI: 10.3844/ajassp.2009.1191.1198.

Urosevic M, Sebastián E, Cardell C, 2013. An experimental study on the influence of surface finishing on the weathering of a building low-porous limestone in coastal environments. *Engineering Geology* **154**: 131–141. DOI: 10.1016/j.enggeo.2012.12.013.

Vallero DA, 2008. *Fundamentals of air pollution*. Elsevier: Amsterdam, Boston.

Van de Wall ARG, Ajalu JS, 1997. Characterization of the geotechnical properties of rock material for construction purposes. *International Journal of Rock Mechanics and Mining Sciences* **34**: 319.e1 - 319.e12. DOI: 10.1016/S1365-1609(97)00208-6.

Vandevoorde D, Cnudde V, Dewanckele J, Brabant L, Bouw M de, Meynen V, Verhaeven E, 2012. Validation of in situ applicable measuring techniques for analysis of the water absorption by stone. *Procedia Chemistry* **8**: 317–327.

Vandevoorde D, Pamplona M, Schalm O, Vanhellemont Y, Cnudde V, Verhaeven E, 2009. Contact sponge method: Performance of a promising tool for measuring the initial water absorption. *Journal of Cultural Heritage* **10**: 41–47. DOI: 10.1016/j.culher.2008.10.002.

Vasconcelos G, Lourenço PB, Alves CA, Pamplona J. 2007. Prediction of the mechanical properties of granites by ultrasonic pulse velocity and Schmidt hammer hardness. In *Proceedings of the Tenth North American Masonry Conference*, The Masonry Society (ed). St. Louis MO. http://hdl.handle.net/1822/9166. (Accessed 20 September 2015).

Vasconcelos G, Lourenço PB, Alves CAS, Pamplona J, 2008. Ultrasonic evaluation of the physical and mechanical properties of granites. *Ultrasonics* **48**: 453–466. DOI: 10.1016/j.ultras.2008.03.008.

Vergès-Belmin V (ed), 2008. *Illustrated Glossary on Stone Deterioration Patterns =Glossaire illustré sur les formes d'altération de la pierre*. English-French ed. Monuments & Sites 15. Paris: ICOMOS (International Council on Monuments and Sites) and ISCS (International Scientific Committee for Stone). http://international.icomos.org/publications/monuments_and_sites/15/pdf/Monuments_and_Sites_15_ISCS_Glossary_Stone.pdf.

Verwaal W, Mulder A, 1993. Estimating rock strength with the Equotip hardness tester. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* **30**: 659–662. DOI: 10.1016/0148-9062(93)91226-9.

Viles H, 1997. Group report: What is the state of our knowledge of the mechanisms of deterioration and how good are our estimates of rates of deterioration? In *Saving our architectural heritage: The conservation of historic stone structures report of the Dahlem Workshop on Saving our Architectural*

Heritage--the Conservation of Historic Stone Structures, Berlin, March 3-8, 1996, Baer NS, Snethlage R (eds). J. Wiley: Chichester, New York; 95–112.

Viles H, 2001. Scale issues in weathering studies. *Weathering and Geomorphology* **41**: 63–72. DOI: 10.1016/S0169-555X(01)00104-0.

Viles H, 2002a. Implications of future climate change for stone deterioration. Geological Society-Special Publications 205:407-418. DOI: 10.1144/GSL.SP.2002.205.01.29.

Viles H, 2002b. Soiling and decay of N.M.E.P. limestone tablets. *The Science of the Total Environment* **292**: 215–229. DOI: 10.1016/S0048-9697(01)01124-X.

Viles HA, 2003. Conceptual modeling of the impacts of climate change on karst geomorphology in the UK and Ireland. *Journal for Nature Conservation* **11**: 59–66. DOI: 10.1078/1617-1381-00035.

Viles H, 2005. Microclimate and weathering in the central Namib Desert, Namibia. *Geomorphology* **67**: 189–209. DOI: 10.1016/j.geomorph.2004.04.006.

Viles H, 2013. Durability and conservation of stone: coping with complexity. *Quarterly Journal of Engineering Geology and Hydrogeology* **46**: 367–375. DOI: 10.1144/qjegh2012-053.

Viles H, Cutler N, 2012. Global environmental change and the biology heritage structures. *Global Change Biology* **18**: 2406–2418.

Viles H, Goudie A, Grab S, Lalley J, 2011. The use of the Schmidt Hammer and Equotip for rock hardness assessment in geomorphology and heritage science: a comparative analysis. *Earth Surface Processes and Landforms* **36**: 320–333. DOI: 10.1002/esp.2040.

Vleugels G, Dewolfs R, van Grieken R, 1993. On the memory effect of limestone for air pollution. *Atmospheric Environment. Part A. General Topics* **27**: 1931–1934. DOI: 10.1016/0960-1686(93)90298-D.

Wang C, Jones R, Perry M, Johnson C, Clark P, 2013. Using an ultrahigh-resolution regional climate model to predict local climatology. *Quarterly Journal of the Royal Meteorological Society* **139**: 1964–1976. DOI: 10.1002/qj.2081.

Warke PA, Curran JM, Turkington AV, Smith BJ, 2003. Condition assessment for building stone conservation: a staging system approach: Building Stone Decay: Observations, Experiments and Modeling. *Building and Environment* **38**: 1113–1123. DOI: 10.1016/S0360-1323(03)00085-4.

Warke PA, Curran JM, Smith BJ, Gardiner M, Foley C, 2010. Post-excavation deterioration of the Copney Bronze Age Stone Circle Complex: A geomorphological perspective. *Geoarchaeology* **25**: 541–571. DOI: 10.1002/gea.20324.

Warscheid T, Braams J, 2000. Biodeterioration of stone: a review. *International Biodeterioration & Biodegradation* **46**: 343–368. DOI: 10.1016/S0964-8305(00)00109-8.

Weber H, 1988. *Mauerfeuchtigkeit*. Ehningen bei Böblingen: Expert.

Wedekind W, Rüdrich J, Siegesmund S, Rieffel Y, 2014. Long-term Evaluation of the Conservation State of Marble Statues. In *Proceedings of the International Conference on Conservation of Stone and Earthen Architectural Heritage: July 2014 Gongju, Republic of Korea*, Lee CH, Kim J, Kim RH (eds). ICOMOS-ISCS; Kongju National University: Gongju; 17–24.

Wexler AS, Seinfeld JH, 1991. Second-generation inorganic aerosol model. *Atmospheric Environment. Part A. General Topics* **25**(12):2731–2748.

Wilcox RR, 2003. *Applying contemporary statistical techniques*. Academic Press: Amsterdam, Boston.

Wilcox RR, 2005. *Introduction to robust estimation and hypothesis testing*. Elsevier/Academic Press: Amsterdam, Boston.

Wilcox RR, 2012. *Introduction to robust estimation and hypothesis testing*. Academic Press: Amsterdam, Boston.

Wilhelm K, Viles H, Burke Ò, 2016a. Low impact surface hardness testing (equotip) on porous surfaces - advances in methodology with implications for rock weathering and stone deterioration research. *Earth Surface Processes and Landforms* (in press). DOI: 10.1002/esp.3882.

Wilhelm K, Viles H, Burke Ò, 2016b. The influence of salt on handheld electrical moisture meters: Can they be used to detect salt problems in porous stone?: *International Journal of Architectural Heritage* (in press). DOI: 10.1080/15583058.2015.1109733.

Wilhelm K, Viles H, Burke O, Mayaud J, 2016c. Surface hardness as a proxy for weathering behaviour of limestone heritage: A case study on dated headstones on the Isle of Portland, UK. *Environmental Earth Sciences* **75**. DOI: 10.1007/s12665-016-5661-y.

Williams RBG, Robinson DA, 2000. Effects of aspect on weathering: anomalous behaviour of sandstone gravestones in southeast England. *Earth Surface Processes and Landforms* **25**: 135–144. DOI: 10.1002/(SICI)1096-9837(200002)25:2<135:AID-ESP50>3.0.CO;2-9.

Winchester V, 1988. An assessment of lichenometry as a method for dating recent stone movements in two stone circles in Cumbria and Oxfordshire. *Botanical Journal of the Linnean Society* **96**: 57–68. DOI: 10.1111/j.1095-8339.1988.tb00627.x.

Winkler EM, Wilhelm EJ, 1970. Salt Burst by Hydration Pressures in Architectural Stone in Urban Atmosphere. *Geological Society of America Bulletin* **81**: 567. DOI: 10.1130/0016-7606(1970)81[567:SBBHPI]2.0.CO;2.

Winkler, E.M., 1994. *Stone in architecture: Properties, durability*. Springer-Verlag: Berlin, New York.

Winter, E., 2014. Kult und Herrschaft am Euphrat In: *Winter, E., ed. Dolichener und Kommagenische Forschungen VI. Asia Minor Studien* 73, Habelt: Bonn.

Yavuz H, Altindag S, Sarac S, Ugur I, Sengun N, 2006. Estimating the index properties of deteriorated carbonate rocks due to freeze-thaw and thermal shock weathering. International *Journal of Rock Mechanics and Mining Science* **43**: 767–775.

Yilmaz NG, 2013. The Influence of Testing Procedures on Uniaxial Compressive Strength Prediction of Carbonate Rocks from Equotip Hardness Tester (EHT) and Proposal of a New Testing Methodology: Hybrid Dynamic Hardness (HDH). Rock Mechanics and *Rock Engineering* **46**: 95–106. DOI: 10.1007/s00603-012-0261-y.

Yu S, Oguchi CT, 2009. Complex relationships between salt type and rock properties in a durability experiment of multiple salt–rock treatments. *Earth Surface Processes and Landforms* **34**: 2096–2110.

Zalasiewicz J, Williams M, Haywood A, Ellis M, 2011. The Anthropocene: a new epoch of geological time? Philosophical transactions. Series A, Mathematical, Physical, and *Engineering Sciences* **369**: 835–841. DOI: 10.1098/rsta.2010.0339.

Zehnder K, 1996. Gyspsum efflorescences in the zone of rising damp. Monitoring of slow decay processes caused by crystallizing salts on wall paintings. In: Riederer J (ed) Proceedings 8th International congress on deterioration and

Zehnder K, 2007. Long-term monitoring of wall paintings affected by soluble salts. *Environmental Geology* **52**: 353–367. DOI: 10.1007/s00254-006-0463-2.