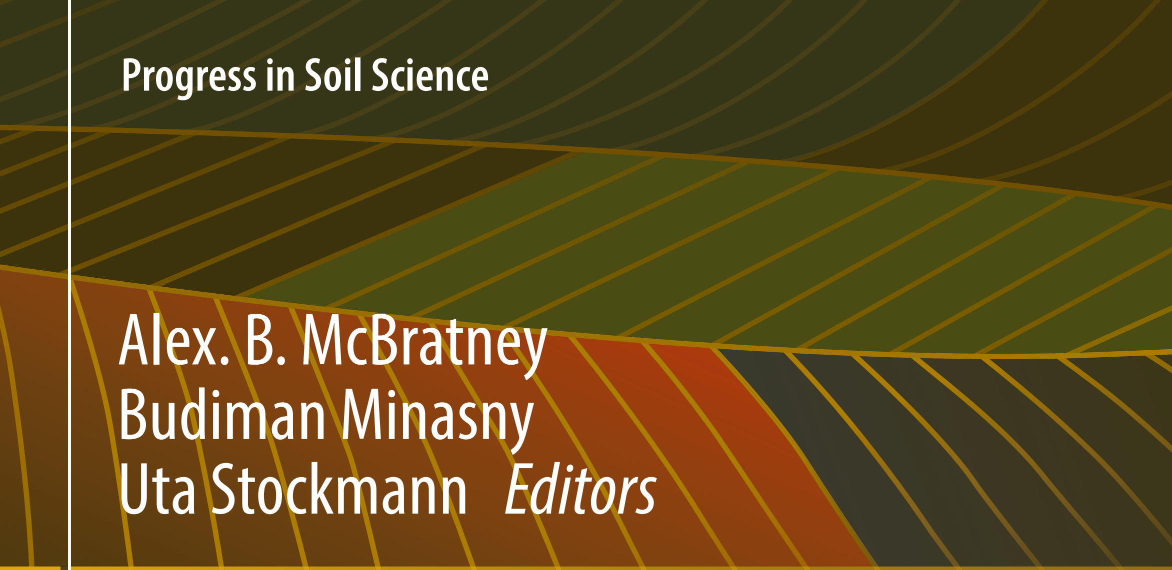
Progress in Soil Science



Alex. B. McBratney
Budiman Minasny
Uta Stockmann *Editors*

Pedometrics

Springer

Progress in Soil Science

Series Editors:
Alfred E. Hartemink, *Department of Soil Science, FD Hole Soils Lab, University of Wisconsin—Madison, USA*
Alex. B. McBratney, *Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Australia*

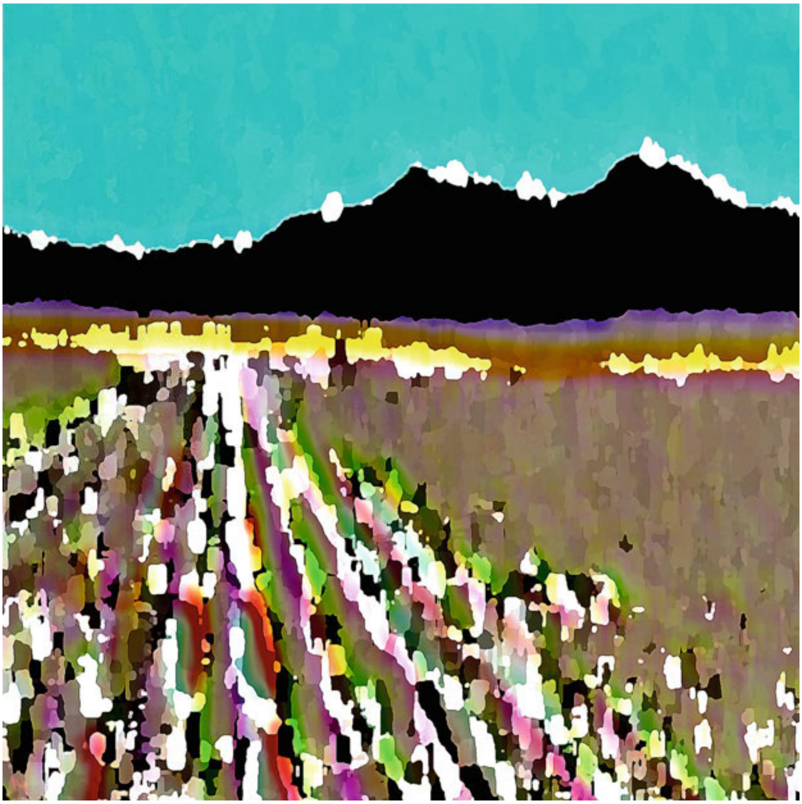
Aims and Scope

Progress in Soil Science series aims to publish books that contain novel approaches in soil science in its broadest sense – books should focus on true progress in a particular area of the soil science discipline. The scope of the series is to publish books that enhance the understanding of the functioning and diversity of soils in all parts of the globe. The series includes multidisciplinary approaches to soil studies and welcomes contributions of all soil science subdisciplines such as: soil genesis, geography and classification, soil chemistry, soil physics, soil biology, soil mineralogy, soil fertility and plant nutrition, soil and water conservation, pedometrics, digital soil mapping, proximal soil sensing, soils and land use change, global soil change, natural resources and the environment.

More information about this series at http://www.springer.com/series/8746

Alex. B. McBratney • Budiman Minasny
Uta Stockmann
Editors

Pedometrics



Springer

Editors
Alex. B. McBratney
Sydney Institute of Agriculture
& School of Life and Environmental
Sciences
The University of Sydney
Sydney, NSW, Australia

Budiman Minasny
Sydney Institute of Agriculture
& School of Life and Environmental
Sciences
The University of Sydney
Sydney, NSW, Australia

Uta Stockmann
Sydney Institute of Agriculture
& School of Life and Environmental
Sciences
The University of Sydney
Sydney, NSW, Australia

ISSN 2352-4774 ISSN 2352-4782 (electronic)
Progress in Soil Science
ISBN 978-3-319-63437-1 ISBN 978-3-319-63439-5 (eBook)
DOI 10.1007/978-3-319-63439-5

Library of Congress Control Number: 2017955385

© Springer International Publishing AG, part of Springer Nature 2018
Chapters 22 and 23 are published with kind permission of © Elsevier B.V. 2018. All Rights Reserved.
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Pedometrics, not in name, but in action, arguably began in the nineteenth century and grew into the twentieth century where we can see early attempts at quantitative expression of spatial and temporal soil variation. The information age which began with the advent of digital computers in the late 1950s saw a second phase of development largely in relation to numerical soil classification and multi-attribute description of soil objects. Geostatistics and information systems brought a large expansion of work in the 1980s and 1990s and also formal recognition and a label – *pedometrics*. The new millennium has seen pedometrics grow from strength to strength expanding the basic science and developing many new areas of application. This book attempts to cover (almost) all the topics that pedometrics comprises in a didactic way. We write in the hope that this text will lead to improved understanding of soil variation and its place in earth system functioning and society.

Sydney, Australia
March 31, 2017

Alex. B. McBratney
Budiman Minasny
Uta Stockmann

Contents

Part I Introduction: What Is Pedometrics?

Part II Statistical Footings

Part III Soil Measurements and Properties

Part IV Soil Materials, Horizons and Profiles

Part V Soil Variation in Space and Time

Part VI Soil Genesis

Part VII Applications of Pedometrics

Contributors

Dominique Arrouays INRA, US 1106, InfoSol Unit, Ardon, Orléans, France

Pat H. Bellamy Cranfield University, Cranfield, Bedfordshire, UK

Nigel R. A. Bird Research Center for the Management of Agricultural and Environmental Risks (CEIGRAM), Universidad Politécnica de Madrid (UPM), Madrid, Spain

Thomas F. A. Bishop Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Florence Carré INERIS, Parc Technologique Alata, Verneuil en Halatte, France

Estelle J. Dominati Farm Systems and Environment, Land and Environment, AgResearch Grasslands, Palmerston North, New Zealand

Mario Fajardo Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Cécile Gomez LISAH, Univ Montpellier, INRA, IRD, Montpellier SupAgro, Montpellier, France

Juan B. Grau Group of Automation of Signals and Communications, Universidad Politécnica de Madrid (UPM), Spain

Jaap J. de Gruijter Alterra, Wageningen University and Research Centre, Wageningen, The Netherlands

Gerard B. M. Heuvelink Soil Geography and Landscape group, Wageningen University and ISRIC – World Soil Information, Wageningen, The Netherlands

Allan E. Hewitt Manaaki Whenua – Landcare Research, Lincoln, New Zealand

Edward J. Jones Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Mark Kibblewhite Cranfield University, Cranfield, Bedfordshire, UK

Philippe Lagacherie LISAH, Univ Montpellier, INRA, IRD, Montpellier SupAgro, Montpellier, France

Juan C. Losada Complex Systems Group, Universidad Politécnica de Madrid (UPM), Spain

R. Murray Lark British Geological Survey, Keyworth, Nottingham, UK

Brendan P. Malone Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Ben P. Marchant British Geological Survey, Keyworth, Nottingham, UK

Manuel P. Martin INRA, US 1106, InfoSol Unit, Ardon, Orléans, France

Juan J. Martín-Sotoca Research Center for the Management of Agricultural and Environmental Risks (CEIGRAM), Universidad Politécnica de Madrid (UPM), Madrid, Spain

Complex Systems Group, Universidad Politécnica de Madrid (UPM), Madrid, Spain

Alex. B. McBratney Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Jeroen Meersmans Cranfield University, Cranfield, Bedfordshire, UK

Irina Mikheeva Institute of Soil Science and Agrochemistry of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russian Federation

Alice E. Milne Sustainable Agricultural Sciences, Rothamsted Research, Harpenden, UK

Budiman Minasny Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Jason Morris Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Melissa Moyce Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Inakwu O. A. Odeh Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Nathan P. Odgers Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Soils and Landscapes Team, Manaaki Whenua – Landcare Research, Lincoln, New Zealand

Thomas G. Orton University of Queensland, Brisbane, Australia

José Padarian Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Stacey Paterson Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Matthew J. Pringle Department of Science, Information Technology and Innovation, EcoSciences Precinct, Brisbane, QLD, Australia

David G. Rossiter School of Integrative Plant Science, Soil & Crop Sciences Section, Cornell University, Ithaca, NY, USA

ISRIC-World Soil Information, Wageningen, The Netherlands

Antonio Saa-Requejo Research Center for the Management of Agricultural and Environmental Risks (CEIGRAM), Universidad Politécnica de Madrid (UPM), Madrid, Spain

Resource Assessment Group, Universidad Politécnica de Madrid (UPM), Madrid, Spain

Nicolas P. A. Saby INRA, US 1106, InfoSol Unit, Ardon, Orléans, France

Sebastien Salvador-Blanes Laboratoire Géo Hydrosystèmes Continentaux, Faculté des Sciences et Techniques, Parc de Grandmont, Tours, France

Uta Stockmann Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Ana M. Tarquis Research Center for the Management of Agricultural and Environmental Risks (CEIGRAM), Universidad Politécnica de Madrid (UPM), Madrid, Spain

Complex Systems Group, Universidad Politécnica de Madrid (UPM), Madrid, Spain

Iván G. Torre Complex Systems Group, Universidad Politécnica de Madrid (UPM), Madrid, Spain

Department of Applied Mathematics and Statistics, Universidad Politécnica de Madrid (UPM), Madrid, Spain

Tom Vanwalleghem Departamento de Agronomia, Universidad de Córdoba, Córdoba, Spain

Ichsani Wheeler Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia

Brett Whelan Precision Agriculture Laboratory, Sydney Institute of Agriculture, The University of Sydney, Sydney, NSW, Australia

Part I
Introduction: What Is Pedometrics?

"As a science grows, its underlying concepts change..."

Hans Jenny
– Factors of Soil Formation:
A System of Quantitative Pedology (1941), McGraw-Hill,
New York. p. 1

« Les méthodes sont ce qui caractérise l'état de la science à chaque époque et qui détermine le plus ses progrès »

Augustin Pyrame de Candolle (1778–1841)

The first words of Hans Jenny's classic book and those of de Candolle presage the need for this text. Soil science advances, and as a consequence we have to develop and grapple with new concepts and methods. This book is an attempt to present some of these concepts and to hang them in appropriate places on some kind of framework to construct a new body of knowledge. The concepts dealt with here have arisen largely because of a phenomenon that has been going on in the earth and biological sciences since the invention of digital computers just a few years after Jenny's book first appeared, namely that of quantification.

First a short geological excursion to illustrate this. An article entitled *Physicists Invading Geologists' Turf* by James Glanz appeared in the New York Times on November 23rd 1999. Here are some excerpts:

> Dr. William Dietrich has walked, driven and flown over more natural landscapes than he can remember. As a veteran geomorphologist, he has studied how everything from the plop of a raindrop to mighty landslides and geologic uplift have shaped the face of the planet.
>
> So when he admits to the growing influence on his field of an insurgent group of physicists, mathematicians and engineers with all-encompassing mathematical theories but hardly any field experience, the earth almost begins to rumble.
>
> Geology, a field that has always gloried in descriptive detail but has had less luck deriving mathematical generalizations, is changing. Invigorated by satellite maps, supercomputers and fresh ideas from physics, researchers are deriving sweeping theories without ever having put hammer to rock.
>
> The trend is unsettling to some old-school geologists, but even they concede that the work has prompted new research in traditional academic circles. While the new researchers have not yet proved that the shape of every hill and dale can be predicted by simple equations, they have at least raised the question of whether landscapes are sculptured by something other than a reductionistic accumulation of forces.

> "In some ways, they irritate us," said Dr. Dietrich, a professor at the University of California in Berkeley, speaking for university and government geologists.
>
> But Dr. Dietrich added that the new, physics-inspired theories of landscape formation, which have found success by ignoring the detail-oriented approach of traditional geology and focusing instead on the earth's overall patterns, are "a way of discussing some sense of regularity in what otherwise is a very messy world."
>
> "I've come to appreciate the perspective," he said.
>
> Inspired by ideas long familiar in physics, and fueled by the recent availability of high-quality satellite maps of much of the planet's surface, the new approach turns the earth sciences on their head, asserting that the most prominent structures on the surface of the planet are shaped not by just local factors, as had been thought, but by the most general properties of physics and mathematics.
>
> For all the insights of the new research, geologists are far from abandoning their worldview. In fact, they too are using satellite data — to perform detailed computer calculations of how rain and wind erode one particular landscape into another over time. "But even there", Dr. Dietrich said, "the physics-oriented researchers have inspired discoveries of order amid the reductionistic detail."
>
> "The places where they accomplish that," Dr. Dietrich said, "will, I think, have a lasting effect on how we think about the planet".

This example illustrates the disciplinary and human relationships[1] between traditional and mathematical geologists, which is completely akin to that between conventional pedologists and pedometricians. Pedologists study pedology, and pedometricans develop pedometrics. (Each group has its contribution to make.) Pedology and pedometrics are closely related. Flippantly, pedology is about augering; pedometrics is about auguring.

Pedology (from the Greek *pedon*, 'ground'), a term first coined in Saxony in 1862 (Simonson 1999), is the scientific study of the soil. More specifically it is the study of soil as part of the natural environment. It is concerned with soil description, spatial distribution, genesis and sustainable use (inter alia Joffe 1949; Buol et al. 1997). Traditionally, it has had a descriptive and field focus (Basher 1997).

Pedometrics generally addresses the same issues as pedology but focuses on specific kinds of problems, those that can be formulated quantitatively and can be solved with quantitative mathematical and statistical techniques. The coining of the term pedometrics and its first definition is by McBratney (1986):

> The use of quantitative methods for the study of soil distribution and genesis and as a sustainable resource.

Pedometrics is a neologism derived from the Greek roots:

πεδον PEDON, the ground, earth, soil
μετρον METRON, measurement

Webster (1994) reminds us that pedometrics is used analogously to other words such as biometrics, psychometrics, econometrics, chemometrics, and the oldest of all geometrics. Etymologically, the word covers two main ideas. First the 'pedo' part corresponds roughly to that branch of soil science we call pedology and the soil,

---

[1] Science is a human construct and as such shows all the glories and imperfections of human nature.

and the metric part has been restricted to quantitative mathematical and statistical methods. If we borrow from the wide definition of biometrics, we get:

> The development and application of statistical and mathematical methods applicable to data analysis problems in soil science.

So it is essentially the application of probability and statistics to soil. Webster (1994), in addition to McBratney's earlier definition, suggested an alternative problem-oriented meaning, which he paraphrased as

> Soil science under uncertainty.

In this sense, pedometrics deals with uncertainty in soil models that are due to deterministic or stochastic variation, vagueness and lack of knowledge of soil properties and processes. Thus, mathematical, statistical and numerical methods could be applied to resolve the uncertainty and complexity inherent in the soil system, including numerical approaches to classification, which deals with supposedly deterministic variation.

Pedometrics is not new, although it was first formally recognised as a different branch of soil science to traditional pedology at the end of the 1980s. Mathematical and statistical methods have been applied to soil studies generally since at least the 1960s with the availability of digital computers and software. The thread stretches back much further to precomputer days, however. It appears to have its origins in agronomy and soil survey, rather than strictly pedology, in the early part of the twentieth century. Harris' (1915) study of soil spatial variation in experimental fields and Robinson and Lloyd's (1915) concern over soil survey error are early examples. We suggest that Forbes' (1846) study of the temporal variation of soil temperature modelled by Fourier series is indeed a pedometric study, and no doubt there are even earlier ones.

For several decades of the twentieth century, pedometrics (although unrecognised and undefined) was a tool for designing experiments and surveys and in advisory work. In the 1960s, pedometricians were concerned with the difficult problem of soil classification and applied the methods of numerical taxonomy. In the late 1970s, pedometricians began to treat soil properties as spatially correlated random processes and to utilise geostatistics for analysis and prediction. Indeed, in a recent collection of the most important papers in soil science historically (Hartemink et al. 2009), this earlier work from pedometrics was recognised in four papers: Youden and Mehlich (1937) on efficient soil sampling, Rayner (1966) on numerical soil classification, Beckett and Webster on soil spatial variation and Burgess and Webster (1980) on soil spatial prediction. More recently, in addition to these earlier themes, pedometrics has begun to attempt to elucidate pedogenesis by quantifying relations between individual soil properties and controlling factors (e.g. Minasny et al. 2008).

Some might argue that Jenny (1941) was the seminal text, especially because of the title of his book *Factors of Soil Formation. A System of Quantitative Pedology*. The text of Webster (1977) and revised a decade or so later by Webster and Oliver (1990) were clearly milestones. Other important texts such as Burrough (1986), Goovaerts (1997), Webster and Oliver (2001), Nielsen and Wendroth (2004),

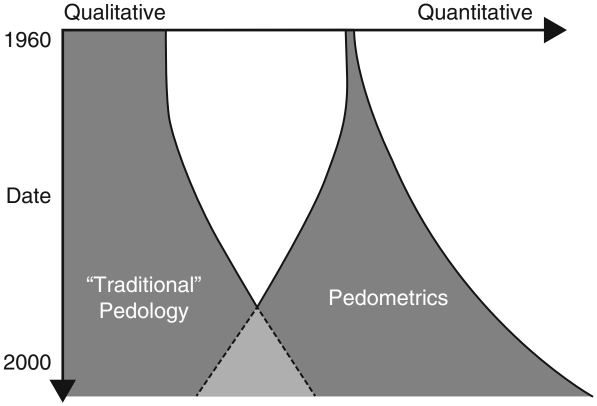


Fig. 1 A timeline of the growth of pedology and pedometrics (After McBratney et al. 2000, Fig. 1)

Grunwald (2005), de Gruijter et al. (2006) and Hengl and Reuter (2009) have a basis in pedometrics and contribute significantly to its canon. Pedometrics has contributed to, and gained from, applied statistics, geostatistics, GIS science, environmental sampling and geomorphometry.

Over time, the use of computers has increased in both pedometrics and in traditional pedology, and the difference between the two has decreased and in some cases overlapped (as shown in Fig. 1). Traditional pedology has, of necessity, become more quantitative through the increased use of computerised soil information systems and field-based measuring devices. Pedometrics has developed quantitative methods, which attempt to account for conceptual pedological models of soil variation. Now there is a strong and growing overlap and synthesis between traditional pedology and pedometrics.

In a bibliometric study of the composition of papers in a leading soil science journal from its inception in 1967 until 2001, Hartemink et al. (2001) showed that papers on pedometrics have risen from less than 3% in 1967 to around 18% of all papers in 2000. It seems, as shown in Fig. 2, that more qualitative soil genesis and morphological studies decreased to make way for the increase in more quantitative studies. By 2016, the proportion of papers in Geoderma on pedometrics had risen to ~27%.

From a pedological point of view, Mermut and Eswaran (2001) saw pedometrics as a research tool with the potential to complement conventional soil surveys and a crucial technique in precision agriculture. By 2016, Brevik et al. (2016) recognised the role of pedometrics in soil mapping, classification and pedological modelling. A great potential has been demonstrated in applications such as digital soil mapping (Lagacherie et al. 2007; Hartemink et al. 2008; Boettinger et al. 2010) with a sixfold linear increase in published papers between 2001 and 2015 and proximal soil sensing (Viscarra Rossel et al. 2010).

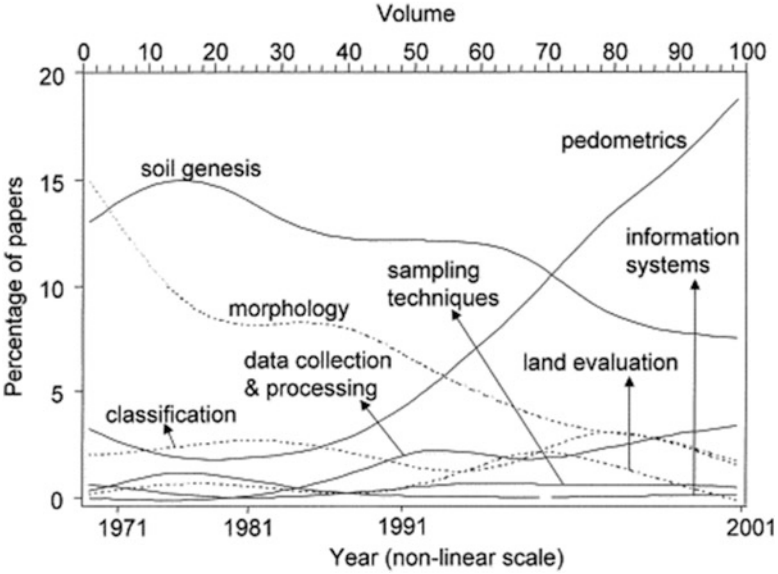


Fig. 2 Trends in pedological subdisciplines and subjects reported in Geoderma between 1967 and 2001. The ordinate is the percentage of papers in Geoderma (After Hartemink et al. 2001, Fig. 16)

References

Basher LR (1997) Is pedology dead and buried? Aust J Soil Res 35:979–994

Boettinger JL, Howell DW, Moore AC, Hartemink AE, Kienast-Brown S (eds) (2010) Digital soil mapping: bridging research, environmental application, and operation. Springer, Dordrecht

Brevik EC, Calzolari C, Miller BA, Pereira P, Kabala C, Baumgarten A, Jordán A (2016) Soil mapping, classification, and pedologic modeling: history and future directions. Geoderma 264B:256–274

Buol SW, Hole FD, McCracken RJ, Southard RJ (1997) Soil genesis and classification. Iowa State University Press, Ames

Burgess TM, Webster R (1980) Optimal interpolation and isarithmic mapping of soil properties. 1. The semi-variogram and punctual kriging. J Soil Sci 31:315–331

Burrough PA (1986) Principles of geographical information systems. Oxford University Press, Oxford

De Gruijter JJ, Brus D, Bierkens M, Knotters M (2006) Sampling for natural resources monitoring. Springer, Heidelberg

Forbes JD (1846) Account of some experiments on the temperature of the earth at different depths and in different soils near Edinburgh. Trans R Soc Edinb 16:189–236

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York

Grunwald S (ed) (2005) Environmental soil-landscape modeling: geographic information technologies and pedometrics. CRC Press, Boca Raton

Harris JA (1915) On a criterion of substratum homogeneity (or heterogeneity) in field experiments. Am Nat 49:453

Hartemink AE, McBratney AB, Cattle JA (2001) Developments and trends in soil science: 100 volumes of Geoderma (1967–2001). Geoderma 100:217–268

Hartemink AE, McBratney AB, Mendonça Santos ML (eds) (2008) Digital soil mapping with limited data. Springer, Heidelberg

Hartemink AE, McBratney AB, White RE (eds) (2009) Soil science vol. I soil in space and time. Earthscan, London

Hengl T, Reuter HI (eds) (2008) Geomorphometry: concepts, software, applications. Elsevier, Amsterdam

Hengl T, Reuter HI (2009) Geomorphometry: concepts, software, applications. Elsevier, Amsterdam/Boston/Heidelberg/London/New York/Oxford/Paris/San Diego/San Francisco/Singapore/Sydney/Tokyo

Jenny H (1941) Factors of soil formation. A system of quantitative pedology. McGraw Hill, New York

Joffe JS (1949) Pedology, 2nd edn. Pedology Publications, New Brunswick

Lagacherie P, McBratney AB, Voltz M (eds) (2007) Digital soil mapping: an introductory perspective, Developments in Soil Science 31. Elsevier, Amsterdam

McBratney AB (1986) Introduction to pedometrics: a course of lectures, CSIRO Division of Soils Technical Memorandum 53/1986. (mimeo). CSIRO Division of Soils, Glen Osmond. 150 pp

McBratney AB, Odeh IOA, Bishop TFA, Dunbar MS, Shatar TM (2000) An overview of pedometric techniques for use in soil survey. Geoderma 97:293–327

Mermut AR, Eswaran H (2001) Some major developments in soil science since the mid-1960s. Geoderma 100:403–426

Minasny B, McBratney AB, Salvador-Blanes S (2008) Quantitative models for pedogenesis – a review. Geoderma 144:140

Nielsen DR, Wendroth O (2004) Spatial and temporal statistics. Sampling field soils and their vegetation. Catena Verlag, Reiskirchen

Rayner JH (1966) Classification of soils by numerical methods. J Soil Sci 17:79–92. 20

Robinson GW, Lloyd WE (1915) On the probable error of sampling in soil survey. J Agric Sci (Camb) 7:144–153

Simonson RW (1999) Origin and acceptance of the term pedology. Soil Sci Soc Am J 63:4–10

Viscarra Rossel RA, McBratney AB, Minasny B (eds) (2010) Proximal soil sensing. Springer, Heidelberg

Webster R (1994) The development of pedometrics. Geoderma 62(1):1–15

Webster R (1977) Quantitative and numerical methods in soil classification and survey. Clarendon Press, The University of Michigan

Webster R, Oliver MA (1990) Statistical methods in land and resource survey. Oxford University Press, Oxford

Webster R, Oliver MA (2001) Geostatistics for environmental scientists. Wiley, New York

Youden WJ, Mehlich A (1937) Selection of efficient methods for soil sampling. Contrib Boyce Thompson Inst Plant Res 9:59–70

Chapter 1
Scope of Pedometrics

Alex. B. McBratney and R. Murray Lark

> *“Those are my principles, and if you dont’t like them well, I have others.”*
>
> Groucho Marx

1.1 The Agenda of Pedometrics

Why do we need pedometrics? What is its agenda? Pedometrics addresses certain key soil-related questions from a quantitative point of view. The need for the quantitative approach arises from a general demand for quantitative soil information for improved economic production and environmental management. Pedometrics addresses four main areas which are akin to the problems of conventional pedology:

1. Understanding the pattern of soil distribution in character space – soil classification
2. Understanding soil spatial and temporal variation
3. Evaluating the utility and quality of soil
4. Understanding the genesis of soil

A.B. McBratney (✉)
Sydney Institute of Agriculture & School of Life and Environmental Sciences,
The University of Sydney, Sydney, NSW 2006, Australia
e-mail: alex.mcbratney@sydney.edu.au

R.M. Lark
British Geological Survey, Keyworth, Nottingham NG12 5GG, UK
e-mail: mlark@bgs.ac.uk

A.B. McBratney et al. (eds.), *Pedometrics*, Progress in Soil Science,
DOI 10.1007/978-3-319-63439-5_1

1.1.1 *Unravelling the Structure of the Soil Character Space: Soil Classification*

We all search for organising generalisations in the face of complexity. Many soil properties need to be observed to characterise a soil. The organisation of this soil complexity through classification has been problematic, probably because classification is an innate human faculty. Leeper (1956, p. 59) noted "*When scientists discuss methods of analysing a solution for phosphate, they are practical, reasonable, and unemotional. When the same men discuss the classification of soils, these virtues are likely to evaporate*". Pedometrics attempts *to resolve some of the polemic of soil classification by a* search for insight into the structure of soil character space. Multivariate and numerical taxonomic methods (Webster 1977a) attempt to unravel this structure, with a view to better prediction and understanding. Chapters 4, 8 and 9 explore these issues in more detail.

1.1.2 *Understanding Soil Spatial and Temporal Variation*

The soil skins the land, and its attributes vary spatially and temporally in a sometimes continuous, sometimes discrete and sometimes haphazard fashion. Using current technologies, we can still only measure most attributes of the soil at a finite number of places and times on relatively small volumes, and therefore statements concerning the soil at other places or times involve estimation and prediction and an inevitable uncertainty. Such prediction and estimation are required for inventory, assessment and monitoring.

One of the tasks of pedometrics is to quantify this inexactitude in order that it can be known and managed accordingly (Heuvelink and Webster 2001). Pedometrics also seeks insight into such spatio-temporal patterns using geostatistical and other spatial and temporal description and prediction tools. Chapters 10, 11 and 12 explore these issues in more detail.

1.1.3 *Evaluating the Utility and Quality of Soil*

The importance of knowledge and awareness of soil resources ranging from individual fields to the global scale is axiomatic. Over and above the problems of spatial and temporal variation, we must put objective value judgments on the utility of soil for specified purposes. This may involve encapsulating practical experience and developing objective rules for management. Conventionally, this has been called land evaluation, and quantification has been underway since Storie (1933) or earlier and has been developed further by inter alia Rossiter (1990).

More recently, the focus has moved to the environmental performance of soil (Adhikari and Hartemink 2016). The somewhat problematic concept of soil quality includes assessment of soil properties and processes as they relate to the ability of soil to function effectively as a component of a healthy ecosystem (Doran and Parkin 1994). The quality of soil as part of an ecosystem may depend on its pedodiversity (Ibáñez et al. 1995). Soil can provide a number of valuable ecosystem services (Dominati et al. 2010) that require to be quantified. This is also reflected in the newly emerging multidimensional concept of soil security where soil plays an integral part in the global environmental sustainability challenges (McBratney et al. 2014). In recent years, pedometrics has contributed more and more to these areas, with the use of pedometric products for land suitability analysis (e.g. Kidd et al. 2015) and assessment of the change in the soil resource in space and time (e.g. Stockmann et al. 2015), for example. These and related issues are explored in Chap. 17.

1.1.4 Understanding the Genesis of Soil

Ultimately, pedometrics would attempt to provide quantitative models of soil formation. This is achieved by encapsulating knowledge of pedological processes in mathematical forms as an alternative to purely statistical approaches. The success of such a modelling approach depends, however, on pedological knowledge and the non-linearity of processes (Phillips 1998). The advantages of a successful modelling of soil formation are substantial and manifold. The three previously discussed agenda items of pedometrics, namely, soil classification, spatial and temporal variation and soil utility and quality, would be predictable from such a model. Therefore, attempts at such an approach have emerged in recent years. Hoosbeek and Bryant (1992) perhaps first outlined the problem. Minasny and McBratney (1999, 2001), Cohen et al. (2010) and Vanwalleghem et al. (2013) have provided the first substantive, but still rudimentary models, followed by more sophisticated models on the soil profile scale and soil landscape scale. This intriguing approach is discussed further in Chaps. 18 and 19.

1.2 Types of Models and Their Evaluation

Models are abstractions of reality. Harvey (1969) gives an early general discussion of models and quantification. Dijkerman (1974) discusses the kinds of models used in soil science generally. Because the soil is a complex and variable system, we represent it with a simpler or more abstract model. Originally, these models were descriptive mental models, but over time they have become more quantitative (Dijkerman 1974). In the context of pedometrics, with its quantitative impulse, a model generally corresponds to one of Webster"s dictionary definitions, i.e. *a system of postulates, data, and inferences presented as a mathematical description of an entity or state of affairs*. Underlying this general statement, there is a variety of model types which we shall now consider and then go on to discuss how to judge when models are performing adequately.

1.2.1 Types of Models

Models can be described in various ways by dividing them into a number of binary categories.

Models may be *qualitative* or *quantitative*. In traditional pedology models tend to be qualitative, whereas in pedometrics the tendency is for quantification. The main advantage of quantitative models seems to be that of reproducibility. Some would argue that quantitative models are more objective, but this is a difficult area from a philosophical point of view. So qualitative models may be considered inferior; however, they do contain knowledge and that knowledge may be captured in some objective way. Qualitative models can be made quantitative. So a qualitative model may be a good starting point for subsequent quantitative investigation, e.g. qualitative models expressing the expert knowledge of a group of experienced scientists and practitioners may be formally written down and used as hypothesis for quantitative testing.

Models may be *static* or *dynamic*. Pedometrics tends to focus on static models, whereas the models of soil physics are dynamic. In the future, pedometric models probably will be increasingly dynamic.

Models may be *empirical* or *mechanistic*. By empirical models we simply describe a phenomenon with as few parameters as possible without necessarily seeking to describe the mechanism underlying the phenomenon, whereas mechanistic models attempt to describe the mechanism (at some scale). For example, we could model the changes in moisture content at some location as a function of time using a purely empirical time series or transfer function model. Alternatively, we could describe the same observations using Richards' equation (Pachepsky et al. 2003) and various other soil physical mechanisms. The latter requires knowledge of basic soil properties, whereas the former does not. Whether one uses a mechanistic or an empirical model will depend on our level of process knowledge and the relative predictability of the models. Another way of describing empirical versus mechanistic models is to term them functional and physical models (Addiscott and Wagenet 1985). Functional is largely synonymous with empirical. Sometimes the empirical and mechanistic categorisation is confused with static and dynamic. Mechanistic models are usually dynamic but as an example given above, empirical models can be dynamic as well as static.

A very important distinction is to consider *deterministic* or *stochastic* models. Underlying Laplace's (1749–1827) statement, "*All the effects of Nature are only the mathematical consequences of a small number of immutable laws*", is the origin of determinism described by Laplace and quoted by Addiscott and Mirza (1998), "*An intellect which at any given moment knew all the forces which animate nature and the mutual positions of all the beings that comprise it, if this intellect was vast enough to submit its data to analysis, could condense into a single formula the movement of the greatest bodies of the universe and that of the lightest atom: for such an intellect nothing could be uncertain, and the future just like the past would be present before his eyes*". Stochastic models concern some phenomenon that has randomness innate to its structure, whereas deterministic models have not. Many earth scientists do not readily accept that such randomness is inherent to nature. This unease recalls Einstein's famous saying "*The Lord does not play with dice*". No one has suggested that any soil process is constitutionally random unless we consider them at the quantum level. The ensemble averages are considered to be deterministic. So, for example, at the so-called Darcy scale, soil water phenomena can be described by a deterministic equation. However, given the inevitable lack of understanding and particularly lack of knowledge of soil processes, then randomness in models is a way of incorporating our ignorance. So in this sense, stochastic models may be seen as pragmatic. Up until the 1980s, stochastic models were rarely recognised or used in soil science. Dexter (1976) was one of the first examples. In hydrology on the other hand, stochastic models are abound (Yevjevich 1987). Pedometrics will probably make an ever-increasing use of stochastic models.

More recently, scientists have realised that deterministic models may have outcomes that look for all intents and purposes like stochastic ones – the outcomes appear random. These models describe processes which have a sensitive dependence on initial conditions. These are the so-called non-linear dynamic models, the basis of chaos theory (Gleick 1987). Such chaotic models blur the distinction between determinism and stochasticity (Addiscott and Mirza 1998). Phillips (1993, 1998) has suggested this kind of determinism as the basis for soil variability itself, i.e. soil formation is a chaotic process. Non-wetting phenomena in soil physics (Persson et al. 2001) are probably chaotic processes that cause soil variation. Minasny and McBratney (2001) give an example of a chaotic model and suggest how a chaotic deterministic model may be distinguished from a random one (Sugihara and May 1990).

As described above, models may be (1) qualitative or quantitative, (2) static or dynamic, (3) empirical or mechanistic (functional or physical) and (4) deterministic or stochastic. If we consider these four distinctions, then there are 16 combinations. Not all of them are sensible ones. In pedometrics, it is certainly worth considering the quantitative combinations and particularly those relating to stochastic and deterministic and empirical and mechanistic as shown in Table 1.1. So these four kinds of models are rather important in soil science. Another category has been

Table 1.1 Examples of quantitative models

	Deterministic		
	Certain	Uncertain	Stochastic
Empirical	Functional leaching model	Pedotransfer functions Fuzzy models	Regionalised variables geostatistics Markov chain models
Mechanistic	Richards' equation	Non-wetting phenomena Soil-landscape model	Molecular-scale diffusion Soil mechanics

added to the table, and we distinguish between certain and uncertain deterministic models. These models lie between purely deterministic and purely stochastic models. Examples of soil phenomena modelled in the various categories are given.

Pedometrics deals particularly with uncertain deterministic models and stochastic ones, i.e. models where randomness has to be introduced to deal with uncertainty (pedotransfer functions, Wösten et al. 2001) and/or randomness is seen as part of nature itself (soil as a realisation of random field – geostatistics). As such many pedometric models have a statistical basis. The uncertain deterministic, mechanistic and the stochastic, mechanistic models (e.g. soil mechanics, Manolis 2002) are not well developed in soil science or pedometrics.

One class of model that does not readily fit in to this scheme are the fuzzy models. They can be possibly thought of as uncertain deterministic, empirical models. They are probably best thought of as another categorisation, that of continuous or discrete states. Bárdossy et al. (1995) modelled soil water movement using the fuzzy approach as described in McBratney and Odeh (1997). Dou et al. (1999) modelled solute movement in a similar way. Pedotransfer functions (Wösten et al. 2001) which are largely linear or non-linear regression equations are empirical deterministic uncertain models, which will be discussed further in Chap. 7 of this book.

1.2.2 Critical Evaluation of Models and Their Parameters

Critical evaluatory procedures for models are needed to maintain the integrity of modelling and to ensure that the increasingly widespread use of models does not result in the propagation of misleading information. Generally speaking, models particularly quantitative pedometric ones, should be good predictors, while at the same time, they should not have too many parameters. The quality of prediction has to be tested with independent data sets by comparing predicted and observed values (Van Kuilenburg et al. 1982) or less satisfactorily with some kind of cross-validation procedure (Solow 1990).

The Franciscan friar, William of Ockham, was concerned about the number of parameters; he wrote in 1332, "*Pluralitas non est ponenda sine necessitate*", which roughly translates as "*Don't make things more complicated than you have to*". This is called Ockham's razor. The two ideas of quality prediction and minimising the number of parameters quantitatively are reflected in measures such as the Akaike information criterion (Webster and McBratney 1989).

Addiscott et al. (1995) pointed out that no model can be validated in the sense that it has been unequivocally justified. All that can be achieved is to show how small the probability is that the model has been refuted. Whether this probability is acceptable remains a subjective decision. In general, the further the data used for parameterisation are removed from the data to be simulated, the better. Problems can arise in both parameterisation and validation if the model is non-linear with respect to its parameters, and the latter have appreciable variances. Parameterisation and validation become more difficult as the complexity of the model or the scale at which it is used increases.

1.3 Methods × Problems

The gamut of pedometrics can also be thought of as a two-way table. The rows are the problems that pedometrics can possibly address, and the columns are the potential mathematical and statistical methods that could be applied to such problems. Some cells of the table will be richly endowed, whereas others will require thought and invention to come up with new approaches. This book, which is essentially a contribution to soil science, and thereby earth, environmental and ecological sciences, is arranged with a problem focus rather than on the methods themselves which would be more appropriate to an applied statistics text.

1.3.1 The Problems of Pedometrics

> *"Sir Walter Blunt, new lighted from his horse.*
> *Stain'd with the variation of each soil*
> *Betwixt that Holmedon and this seat of ours;"*
>
> William Shakespeare, Henry IV Part 1 Act 1 Scene 1.

Pedometrics arises principally from this common observation that the soil is spatially variable. The soil varies laterally and with depth, and this variation has implications for its use and management. This is not a new insight. The rabbinical biblical commentators on the book of Genesis discussed the question of how much water is needed to sustain plant growth:

> *"How much rain must descend that it may suffice for fructification? As much as will fill a vessel of three handbreadths. This is Rabbi Meir's opinion. Rabbi Judah said: in hard soil, one handbreadth; in average soil, two; in humid soil three."*
>
> Midrash Genesis Rabbah XIII, 13.

Later in history, Thomas Tusser, in his *One Hundred Points of Good Husbandry* published in England in the sixteenth century, noted pithily that *each divers soil hath divers toile*. More recently, the development of agronomy, soil management and engineering has shown that the spatial variation of the soil means that its suitability for different purposes will vary in space. This is important for the planning of infrastructure, the government agencies making planning decisions about land use and, in recent years, the individual farmer trying to improve the efficiency of cropping systems by managing inputs in response to fluctuations in crop requirements at within-field scales.

We can think of other problems where the spatial variation of the soil will be of practical significance. Throughout the industrialised world, it is recognised that the soil is at risk of pollution from unregulated inputs and emissions. If we are to monitor this problem and focus protection and remediation where the problem is most severe, or the soil most vulnerable, then we must be able to detect changes in pollutants, which arise from complex processes of deposition and transport, against a background of complex intrinsic variation of the soil. A similar problem arises when we set out to monitor changes in the carbon content of the soil to evaluate its significance as a source and sink for greenhouse gases.

Soil variation is linked to some pressing problems, but it is also pertinent to other questions. Faced with variation in the soil cover of any terrain, the natural response is to ask how and why it arose. Many soil scientists would not have become interested in their subject if the soil was more or less uniform in space. To understand the variation of the soil at scales from the aggregate to the continent and the relationship of this variation to that of the vegetation which it supports is a basic scientific challenge. The better we understand these basic questions, the better equipped we will be to address the practical ones. Pedometrics is concerned both with meeting the practical requirements of soil management and also with generating insight into how and why the soil material varies as it does.

These, in summary, are the problems of pedometrics. In the remainder of this section, we want to set them in a general framework, with examples, as a prelude to introducing the methods which are used in their solution.

1.3.1.1 Prediction

The first type of problem is prediction. The basic question requiring a prediction has the form:

1. What soil conditions pertain at position $\mathbf{x}$?

Vector $\mathbf{x}$ may contain two or three Cartesian coordinates which define a location in space and possibly a further number which defines a time. This very general question may be refined in different ways:

1.a What is the value of soil variable s at $\mathbf{x}$, $s(\mathbf{x})$, *given a set of observations of the variable at other locations?*

The variable $s(\mathbf{x})$ might be the concentration of available potassium in the topsoil at a location in space $\{x,y\}$. Note that some finite volume of the soil is implicitly of interest. We call this volume the geometric support of $s(\mathbf{x})$.

This problem arises because very few if any soil properties of direct practical interest will have been measured exhaustively across a region. The soil is a continuum, and the constraints of costs and time mean that the soil may only be sampled and measured at a few sites. Information will inevitably be required about soil which has not been measured directly.

Reflection on this problem makes it clear that the best answer we can obtain to our question, short of actually sampling the soil at $\mathbf{x}$, will be an estimate of $s(\mathbf{x})$,

$\widehat{s}(\mathbf{x})$, a number with attendant uncertainty. The way in which this uncertainty is quantified and controlled is a critical issue in pedometrics, to which we will return repeatedly.

A second refinement of the question is possible:

1.b. What is the value of, $y(\mathbf{x})$, *a mathematical function of soil property s at* $\mathbf{x}$*?*

This may seem at first sight like a pedantic variant of question *1.a.*, but it raises an important issue, and a practical one. We may be able, for example, to obtain a reasonable estimate of the clay content of the soil at $\mathbf{x}$, but are really interested in its available water capacity. A reasonable estimate of this latter quantity might be obtained as a function of the clay content. If y is a linear function of s, then the transformation of the estimate $\widehat{s}(\mathbf{x})$ to an estimate $\widehat{y}(\mathbf{x})$ is simple. If the relationship is not linear, then the uncertainty of the estimate $\widehat{s}(\mathbf{x})$ must be accounted for.

Other variants on the simple prediction problem are possible. For example:

1.c. What is the difference between $s\{x,y,t_i\}$ *and* $s\{x,y,t_j\}$*?*

where the third term in the vector denotes a time. This question will arise in environmental monitoring. If $s\{x,y,t_i\}$ is the concentration of a pollutant at x at time t_i, then the difference may be a measure of the success of a soil remediation campaign or the environmental impact of a change in regulations.

This problem will vary in form. A critical question is whether it is possible in principle to measure both variables – given the support of $s\{x,y,t_i\}$, does the disturbance of the initial sampling prevent a meaningful measurement from being made at $\{x,y,t_j\}$? This problem presents pedometricians with interesting challenges. Since politicians are increasingly interested in reliable estimates of the change of organic carbon in the soil, in response to the Kyoto protocol, the problem is also timely and topical.

So far, we have considered the soil over a small volume about a notional two- or three-dimensional location. In practice, we may be concerned more often with predictions about larger parcels or blocks of land, regularly or irregularly shaped. Such parcels may constitute management units, for example. The general problem, then, is

1.d. What is the value of $s(\mathbf{X}) = \int_{x \subset X} s(x)\mathrm{d}x$?

The integral implies that the new variable is effectively the average of all the notional point values, $s(\mathbf{x})$, in the region. Thus, we might be asked, 'what is the mean concentration of lead in the soil at this former factory site' or 'what is the mean concentration of available phosphorous in the soil of this field'? Since our prediction will be based on observations which are effectively point samples on volumes of soil, very small by comparison to the block $\mathbf{X}$, the problem involves generalisation from one spatial scale to a coarser scale. This is sometimes called 'upscaling' or 'aggregation'.

The change of scale can cause problems. Consider the following:

1.e. What is the value of $y(X)$ *where y is a mathematical function of soil variables w and z at the scale of point observations* $s(x)$ *and* $w(x)$*?*

Again, y might be the available water capacity of the soil and w and s bulk the density and clay content, respectively. If w and s are measured at all locations, then we might evaluate y at all these locations then aggregate these to the coarser unit. If the two variables have been measured independently, then we can combine two aggregated values, $s(X)$ and $w(X)$, but only if the functional relationship is linear. Otherwise, it is necessary to make some inference about the joint variation of variables s and w at the original point scale; this is sometimes called 'downscaling' or 'disaggregation'.

We have raised the issue of uncertainty already in the context of simple prediction, but uncertainty may sometimes be addressed directly in a pedometric problem. Consider the following:

1.f. What is the risk that $s(\mathbf{x})$ *exceeds at threshold t?*

The variable s might be the concentration of a pollutant and t a regulatory threshold. For example, the limit for concentration of lead in the soil set by the ANZECC guidelines is 300 mg/kg.

A type 1a question requires a simple prediction, $\widehat{s}(\mathbf{x})$, but in the context of the present problem, this is generally not adequate. If we act on the prediction, there remains the risk that contaminated soil is left untreated (because $\widehat{s}(\mathbf{x}) < t$ but $s(\mathbf{x}) > t$) or that expensive remediation is applied to land unnecessarily (because $\widehat{s}(\mathbf{x}) > t$ but $s(\mathbf{x}) < t$). If our predictions are unbiased, i.e. on average $\widehat{s}(\mathbf{x}) = s(\mathbf{x})$, and the costs of an error in one direction are more or less equal to the costs of an error in the other direction, then all we can do is try to reduce the uncertainty of our predictions as far as possible. Often, however, the costs of an error in one direction are much steeper than the other (e.g. the fines for leaving land unremediated may be large compared to the costs of remediation). In these circumstances, the best decision must account for the risk that $s(\mathbf{x}) > t$, given all available knowledge implicit in the prediction $\widehat{s}(\mathbf{x})$.

This latter problem was expressed in terms of a near-point support at **x**. Answering the final prediction-type problem:

1.g. What is the risk that the mean of s over region **X** *exceeds threshold t?*

raises further problems for the pedometrician. Since management decisions will generally be made about a parcel of land, practical questions will often be framed this way.

1.3.1.2 Inventory and Allied Problems

A second category of pedometric problems can be recognised. If the problems in category 1 are variants of the question 'what conditions pertain at **x**', then class 2 consists of variants on:

2. Over what subregion does conditions pertain?

At its simplest, the problem may be one of inventory – enumeration of the sites over which certain conditions are found. Thus, for example:

2.a. Over what region, X', does the value of soil variable s fall within the range $s_1 < s < s_2$?

As with *questions of* type 1.a., there is uncertainty attendant on any answer to this question since the soil properties are not measured exhaustively.

This problem might be posed by the land manager who wants to know where the depth of the soil over the underlying rock is large enough to permit the growth of a tree crop. It might also be asked by the farmer who wants to know where soil pH is likely to limit certain crops. In reality, a more complex problem might be posed:

2.b. Over what region, X', does some function of soil variablesy $= f(s, w)$ fall within the range $s_1 < y < s_2$?

If we want to identify areas where the likely erosion losses exceed some threshold, then we may have to compute some non-linear function of soil properties like the universal soil loss equation (Wischmeir and Smith 1978). This poses similar problems to the analogous prediction problem 1.b.

Temporal monitoring of the soil may also generate problems with an inventory flavour such as:

2.c. Over what region, X', does the change in soil property s from time t_1 to time t_2 fall within the range $s_1 < s < s_2$?

This problem arises when the policy maker wants to know over what proportion of a landscape the concentrations of pollutants in the soil are diminishing or where the organic carbon content is increasing.

1.3.1.3 Decision Support

Prediction and inventory are tools. Managers require that the answers to these basic questions are integrated in a way which aids decision-making directly. This is the sphere of decision-support systems, and pedometric problems arise. There are two general types of problems:

3.a. What is the optimum management strategy over region X' to achieve goals A subject to constraints B?

So, for example, what is the optimum nitrogen rate to prescribe for a particular parcel of a field to maximise the economic return to the producer subject to the constraints that emissions to the environment through leaching and denitrification do not exceed some threshold? The problem requires process models of an appropriate level of sophistication to describe the whole system under different scenarios. Information on soil properties within the parcel of concern will also be required – subproblems of type 1.c., for example. The optimisation subject to a constraint may be done numerically – i.e. by computing several runs of the system model in an ordered way to find the scenario which best meets the goals subject to the constraints.

In practice, decisions have to be made under uncertainty about many critical conditions, e.g. weather in the case of crop management. We then have a problem;

> *3.b. What is the optimum management strategy over region X' to achieve goals A subject to constraints B and given that the conditions C have a particular statistical distribution?*

1.3.1.4 In Pursuit of Insight

The problems so far have had a strongly practical flavour, but pedometric problems include basic scientific questions where a quantitative account of the spatial variation of the soil is required. We can identify such a problem:

> *4.a. Which factors appear to determine the lateral and horizontal spatial variation of soil property s?*

This question implies a spatial scale – are we concerned with variations at the scale of microbial activity or geomorphic processes or some range between? This general problem may be addressed using analytical methods, but these always rest on assumptions which may not be realistic. An approach to the problem can never be driven purely by data analysis, however. Our investigation will be most fruitful if it is structured around a hypothesis. The approach may be statistical:

> *4.b. Is hypothesis* H *about the causes of soil variation supported by the given observations of properties s...?*

or structured around a mechanistic model:

> *4.c. Given a process model linking input variables s, w and output y, can this particular set of observations of the variables be held to validate the model?*

1.3.2 The Methods Pedometrics Uses

These questions are addressed by pedometrics. We now offer an overview of the pedometric methods that are treated in more detail in later chapters.

1.3.2.1 Statistical Prediction and Modelling

Random Variables

Ideally, soil scientists would like to base pedometric methods on quantitative understanding of soil processes. The ideal way of predicting the value of a soil property at location x would be to enumerate factors of soil formation which