

Chapter 1

Advanced Mapping of Environmental Data: Introduction

1.1. Introduction

In this introductory chapter we describe general problems of spatial environmental data analysis, modeling, validation and visualization. Many of these problems are considered in detail in the following chapters using geostatistical models, machine learning algorithms (MLA) of neural networks and Support Vector Machines, and the Bayesian Maximum Entropy (BME) approach. The term “mapping” in the book is considered not only as an interpolation in two- or three-dimensional geographical space, but in a more general sense of estimating the desired dependencies from empirical data. The references presented at the end of this chapter cover the range of books and papers important both for beginners and advanced researchers. The list contains both classical textbooks and studies on contemporary cutting-edge research topics in data analysis.

In general, mapping can be considered as: a) a spatiotemporal classification problem such as digital soil mapping and geological unit classification, b) a regression problem such as mapping of pollution and topo-climatic modeling, and c) a problem of probability density modeling, which is not a mapping of values but “mapping” of probability density functions, i.e., the local or joint spatial distributions conditioned on data and available expert knowledge.

2 Advanced Mapping of Environmental Data

As well as some necessary theoretical introductions to the methods, an important part of the book deals with the presentation of case studies. These are both simulated problems used to illustrate the essential concepts and real life applications. These case studies are important complementary parts of the current volume. They cover a wide range of applications: environmental data analysis, pollution mapping, epidemiologic spatiotemporal data analysis, socio-economic data classification and clustering. Several case studies consider multivariate data sets, where variables can be dependent (linearly or nonlinearly correlated) or independent. Common to all case studies is that data are geo-referenced, i.e. they are located at least in a geographical space. In a more general sense the geographical space can be enriched with additional information, giving rise to a high dimensional geo-feature space. Geospatial data can be categorical (classes), continuous (fields) or distributions (probability density functions).

Let us remember that one of the simplest problems – the task of spatial interpolation from discrete measurements to continuous fields – has no single solution. Even with a very simple interpolation method just by changing one or two tuning parameters many different “maps” can be produced. Here we are faced with an extremely important question of model assessment and model selection.

The selection of the method for data analysis, modeling and predictions depends on the quantity and quality of data, the expert knowledge available and the objectives of the study.

In general, two fundamental approaches when working with data are possible: deterministic models, including the analysis of data using physical models and deterministic interpolations, or statistical models which interpret the data as a realization of a random/stochastic process. In both cases models and methods depend on some hypotheses and have some parameters that should be tuned in order to apply the model correctly. In many cases these two groups merge, and deterministic models might have their “statistical” side and vice versa.

Statistical interpretation of spatial environmental data is not trivial because usually only one realization (measurements) of the phenomena under study exists. These cases are, for example, geological data, pollution after an accident, etc. Therefore, some fundamental hypotheses are very important in order to make statistical inferences when only one realization is available: ergodicity, second-order stationarity, intrinsic hypotheses (see Chapter 3 for more detail). While some empirical rules exist, these hypotheses are very difficult to verify rigorously in most cases.

An important aspect of spatial and spatiotemporal data is the anisotropy. This is the dependence of the spatial variability on the direction. This phenomenon can be

detected and characterized with structural analysis such as the variography presented below.

Almost all of the models and algorithms considered in this book (geostatistics, MLA, BME) are based on the statistical interpretation of data.

Another general view on environmental data modeling approaches is to consider two major classes: model-dependent approaches (geostatistical models – Chapter 3 and BME – Chapter 6) and data-driven adaptive models (machine learning algorithms – Chapter 4). Being applied without the proper understanding and lacking interpretability, the data-driven models were often considered as black or gray box models. Obviously, each data modeling approach has its own advantages and drawbacks. In fact, both approaches can be used as complementary tools, resulting in hybrid models that can overcome some of the problems.

From a machine learning point of view the problem of spatiotemporal data analysis can be considered as a problem of pattern recognition, pattern modeling and pattern prediction or pattern completion.

There are several major classes of learning approaches:

- supervised learning. For example, these are the problems of classification and regression in the space of geographical coordinates (inputs) based on the set of available measurements (outputs);
- unsupervised learning. These are the problems with no outputs available, where the task is to find structures and dependencies in the input space: probability density modeling, spatiotemporal clustering, dimensionality reduction, ranking, outlier/novelty detection, etc. When the use of these structures can improve the prediction for a small amount of available measurements, this setting is called semi-supervised learning.

Other directions such as reinforcement learning exist but are rarely used in environmental spatial data analysis and modeling.

1.2. Environmental data analysis: problems and methodology

1.2.1. *Spatial data analysis: typical problems*

First let us consider some typical problems arising when working with spatial data.



Figure 1.1. *Illustration of the problem of environmental data mapping*

Given measurements of several variables (see Figure 1.1 for the illustration) and a region of the study, typical problems related to environmental data mapping (and beyond, such as risk mapping, decision-oriented mapping, simulations, etc.) can be listed as follows:

- predicting a value at a given point (marked by “?” in Figure 1.1, for example). If it is the only point of interest, perhaps the best way is simply to take a measurement there. If not, a model should be developed. Both deterministic and statistical models can be used;
- building a map using given measurements. In this case a dense grid is usually developed over the region of study taking into account the validity domain (see Chapter 2) and at each grid node predictions are performed finally giving rise to the raster model of spatial predictions. After post-processing of this raster model different presentations are possible – isolines, 3D surfaces, etc. Both deterministic and statistical models can be used;
- taking into account measurement errors. Errors can be either independent or spatially correlated. Statistical treatment of data is necessary;
- estimating the prediction error, i.e. predicting both unknown value and its uncertainty. This is a much more difficult question. Statistical treatment of data is necessary;

- risk mapping, which is concerned with uncertainty quantification for the unknown value. The best approach is to estimate a local probability density function, i.e. mapping densities using data measurements and expert knowledge;

- joint predictions of several variables or prediction of a primary variable using auxiliary data and information. Very often in addition to the main variable there are other data (secondary variables, remote sensing images, digital elevation models, etc.) which can contribute to the analysis of the primary variable. Additional information can be “cheaper” and more comprehensive. There are several geostatistical models of co-predictions (co-kriging, kriging with external drift) and co-simulations (e.g. sequential Gaussian co-simulations). As well as being more complete, secondary information usually has better spatial and dimensional resolutions which can improve the quality of final analysis and recover missing information in the principal monitoring network. This is an interesting topic of future research;

- optimization of the monitoring network (design/redesign). A fundamental question is always where to go and what to measure? How can we optimize the monitoring network in order to improve predictions and reduce uncertainties? At present there are several possible approaches: uncertainty/variance-based, Bayesian approach, space filling, optimization based on support vectors (see references);

- spatial stochastic conditional simulations or modeling of spatial uncertainty and variability. The main idea here is to develop a spatial Monte Carlo model which can produce (generate) many realizations of the phenomena under study (random fields) using available measurements, expert knowledge and well defined criteria. In geostatistics there are several parametric and non-parametric models widely used in real applications (Chapter 3 and references therein). Post-processing of these realizations gives rise to different decision-oriented maps. This is the most comprehensive and the most useful information for an intelligent decision making process;

- integration of data/measurements with physical models. In some cases, in addition to data science-based models – meteorological models, geophysical models, hydrological models, geological models, models of pollution dispersion, etc. are available. How can we integrate/assimilate models and data if we do not want to use data only for the calibration purposes? How can we compare patterns generated from data and models? Are they compatible? How can we improve predictions and models? These fundamental topics can be studied using BME.

1.2.2. Spatial data analysis: methodology

The generic methodology of spatial data analysis and modeling consists of several phases. Let us recall some of the most important.

– **Exploratory spatial data analysis (ESDA).** Visualization of spatial data using different methods of presentation, even with simple deterministic models helps to detect data errors and to understand if there are patterns, their anisotropic structures, etc. An example of sample data visualization using Voronoi polygons and Delaunay triangulation is given in Figure 1.2. The presence of spatial structure and the West-East major axis of anisotropy are evident. Geographical Information Systems (GIS) can also be used as tools both for ESDA and for the presentation of the results. ESDA can also be performed within moving/sliding windows. This regionalized ESDA is a helpful tool for the analysis of complex non-stationary data.

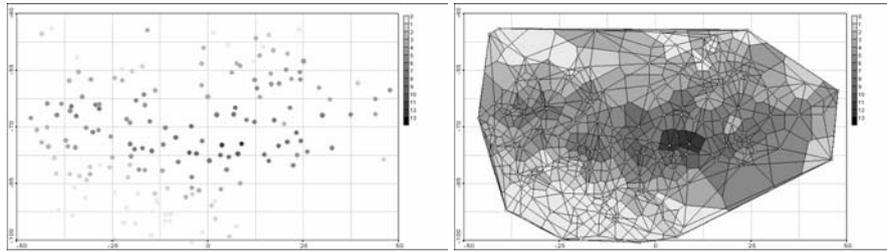


Figure 1.2. Visualization of raw data (left) using Voronoi polygons and Delaunay triangulation (right)

– **Monitoring network analysis and descriptions.** The measuring stations of an environmental monitoring network are usually spatially distributed in an inhomogenous manner. The problem of network homogeneity (clustering and preferential sampling) is closely connected to global estimations, to the theoretical possibility of detecting phenomena with a monitoring network of the given design. Different topological, statistical and fractal measures are used to quantify spatial and dimensional resolutions of the networks (see details in Chapter 2).

– **Structural analysis (variography).** Variography is an extremely important part of the study. Variograms and other functions describing spatial continuity (rodogram, madogram, generalized relative variograms, etc.) can be used in order to characterize the existence of spatial patterns (from a two-point statistical point of view) and to quantify the quality of machine learning modeling using variography of the residuals. The theoretical formula for the variogram calculation of the random variable $Z(\mathbf{x})$ under the intrinsic hypotheses is given by:

$$\gamma(\mathbf{x}, \mathbf{h}) = \frac{1}{2} \text{Var} \{ Z(\mathbf{x}) - Z(\mathbf{x} + \mathbf{h}) \} = \frac{1}{2} E \left\{ (Z(\mathbf{x}) - Z(\mathbf{x} + \mathbf{h}))^2 \right\} = \gamma(\mathbf{h})$$

where \mathbf{h} is a vector separating two points in space. The corresponding empirical estimate of the variogram is given by the following formula

$$\gamma_{ij}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} (Z_i(\mathbf{x}) - Z_i(\mathbf{x} + \mathbf{h}))^2$$

where $N(\mathbf{h})$ is a number of pairs separated by vector \mathbf{h} .

The variogram has the same importance for spatial data analysis and modeling as the auto-covariance function for time series. Variography should be an integral part of any spatial data analysis independent of the modeling approach applied (geostatistics or machine learning). In Figure 1.3 the experimental variogram rose for the data shown in Figure 1.2 is presented. A variogram rose is a variogram calculated in several directions and at many lag distances. A variogram rose is a very useful tool for detecting spatial patterns and their correlation structures. The anisotropy can be clearly seen in Figure 1.3.

– **Spatiotemporal predictions/simulations**, modeling of spatial variability and uncertainty, risk mapping. The following methods are considered in this book:

- *Geostatistics* (Chapter 3). Geostatistics is a well known approach developed for spatial and spatiotemporal data. It was established in the middle of the 20th century and has a long successful history of theoretical developments and applications in different fields. Geostatistics treats data as realizations of random functions. The geostatistical family of kriging models provides linear and nonlinear modeling tools for spatial data mapping. Special models (e.g. indicator kriging) were developed to “map” local probability density functions, i.e. modeling of uncertainties around unknown values. Geostatistical conditional stochastic simulations are a type of spatial Monte Carlo generator which can produce many equally probable realizations of the phenomena under study based on well defined criteria.

- *Machine Learning Algorithms* (Chapter 4). Machine Learning Algorithms (MLA) offer several useful information processing capabilities such as nonlinearity, universal input-output mapping and adaptivity to data. MLA are nonlinear universal tools for obtaining and modeling data. They are excellent exploratory tools. Correct application of MLA demands profound expert knowledge and experience. In this book several architectures widely used for different applications are presented: neural networks: multilayer perceptron (MLP), probabilistic neural network (PNN), general regression neural network (GRNN), self-organizing (Kohonen) maps (SOM), and from statistical learning theory: Support Vector Machines (SVM), Support Vector Regression (SVR), and other kernel-based methods. At present, the conditional stochastic simulations using machine learning is an open question.

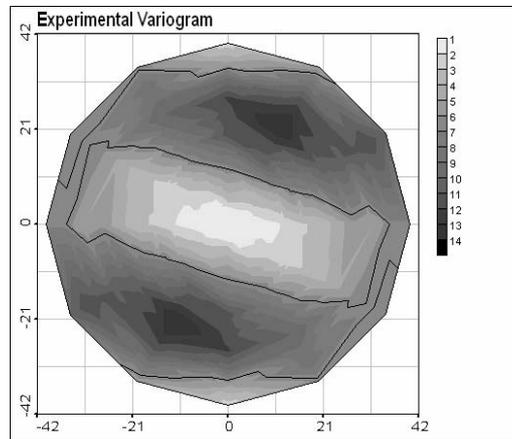


Figure 1.3. *Experimental variogram rose for the data from Figure 1.2*

- *Bayesian Maximum Entropy* (Chapter 6). Bayesian Maximum Entropy (BME) is based on recent developments in spatiotemporal data modeling. BME is extremely efficient in the integration of general expert knowledge and specific information (e.g. measurements) for the spatiotemporal data analysis, modeling and mapping. Under some conditions BME models are reduced to geostatistical models.

- **Model assessments/model validation.** This is a final phase of the study. The “best” models are selected and justified. Their generalization capabilities are estimated using a validation data set – a completely independent data set never used to develop and to select a model.

- **Decision-oriented mapping.** Geomatics tools such as Geographical Information Systems (GIS) can be used to efficiently visualize the prediction results. The resulting maps may include not only the results of data modeling but other thematic layers important for the decision making process.

- **Conclusions, recommendations, reports, communication of the results.**

1.2.3. Model assessment and model selection

Now let us return to the question of data modeling. As has already been mentioned, in general, there is no single solution to this problem. Therefore, an extremely important question deals with model selection and model assessment procedures. First we have to choose the “best” model and then estimate its generalization abilities, i.e. its predictions on a validation data set which has never been used for model development.

Model selection and model assessment have two distinct goals [HAS 01]:

- *Model selection*: estimating the performance of different models in order to choose the best one: the most appropriate, the most adapted to data, best matching some prior knowledge, etc.

- *Model assessment*: having chosen a model, model assessment deals with estimating its prediction error on new independent data (generalization error).

In practice these problems are solved either using different statistical techniques or empirically by splitting the data into three subsets (Figure 1.4): training data, testing data and validation data. Let us note that in this book the traditional definition used in environmental modeling is used. The machine learning community splits data in the following order: training/validation/testing.

The training data subset is used to train the selected model (not necessarily the optimal or best model); the testing data subset is used to tune hyper-parameters and/or for the model selection, and the validation data subset is used to assess the ability of the selected model to predict new data. The validation data subset is not used during the training and model selection procedure. It can be considered as a completely independent data set or as additional measurements.

The distribution of percentages between data subsets is quite free. What is important is that all subsets characterize the phenomena under study in a similar way. For environmental spatial data it can be the clustering structure, the global distributions and variograms which should be similar for all subsets.

Model selection and model assessment procedures are extremely important especially for data-driven machine learning algorithms, which mainly depend on data quality and quantity and less on expert knowledge and modeling assumptions.

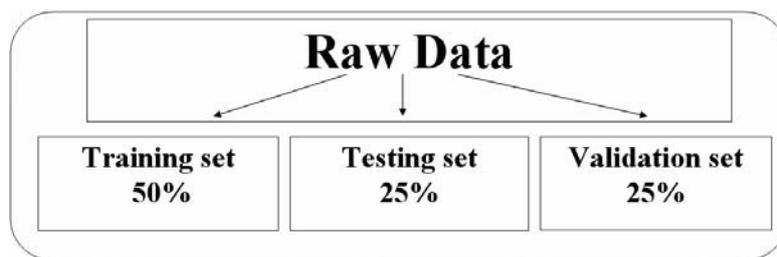


Figure 1.4. Splitting of raw data

A scheme of the generic methodology of using machine learning algorithms for spatial environmental data modeling is given in Figure 1.5. The methodology is similar to any other statistical analysis of data. The first step is to extract useful information (which should be quantified, e.g. as information described by spatial correlations) from noisy data. Then, the quality of modeling has to be controlled by analyzing the residuals. The residuals of training, testing and validation data should be uncorrelated white noise. Unfortunately in many applied publications this important step of the residual analysis is neglected.

Another important aspect of environmental decisions both during environmental modeling or environmental data analysis and forecasting deals with uncertainties of the corresponding modeling results. Uncertainties have great importance in intelligent decisions; sometimes they can be even more important than the particular prediction values. In statistical models (geostatistics, BME) this procedure is inherent and under some hypotheses confidence intervals can be derived. With MLA this is a slightly more difficult problem, but many theoretical and operational solutions have already been proposed.

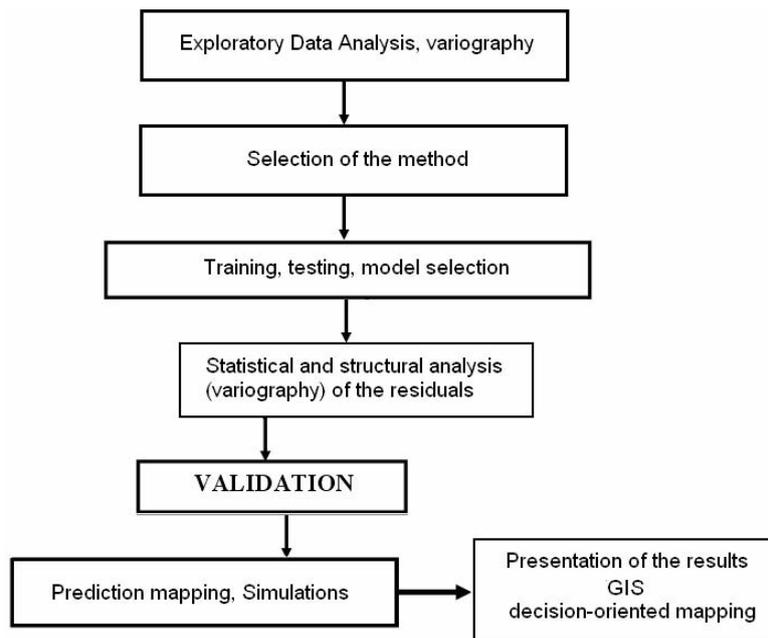


Figure 1.5. Methodology of MLA application for spatial data analysis

Concerning mapping and visualization of the results one possibility to summarize both predictions and uncertainties is to use “thick isolines” which characterize the uncertainty of spatial predictions (see Figure 1.6). For example, under some hypotheses which depend on the applied model, the interpretation is that with a probability of 95% an isoline of the predefined decision level can be found in the thick zone. Correct visualization is important in communicating the results to decision makers. It can be used also for monitoring network optimization procedures by demonstrating regions with high or unacceptable uncertainties. Let us note that such a visualization of predictions and uncertainties is quite common in time series analysis.

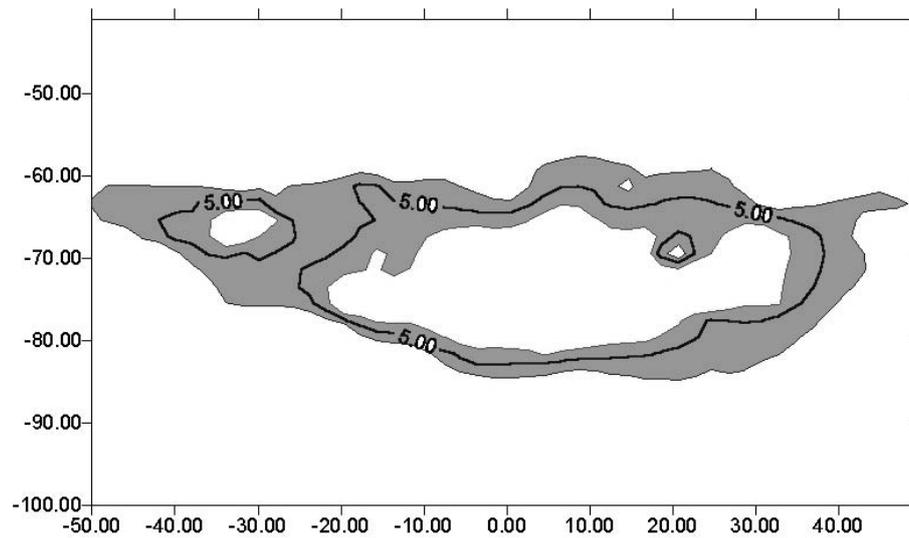


Figure 1.6. *Combining predictions with uncertainties: “thick isolines”*

In this section some basic problems of spatial data analysis, modeling and visualization were presented. Model-based (geostatistics, BME) methods and data-driven algorithms (MLA) were mentioned as possible modeling approaches to these tasks. Correct application of both of them demands profound expert knowledge of data, models, algorithms and their applicability. Taking into account the complexity of spatiotemporal data analysis, the availability of good literature (books, tutorials, papers) and software modules/programs with user-friendly interfaces are important for learning and applications.

In the following section some of the available resources such as books and software tools are given. The list is very short and far from being complete for this very dynamic research discipline, sometimes called environmental data mining.

1.3. Resources

Some general information, including references to the conferences, tutorials and software for the methods considered in this book can be found on the Internet, in particular on the following sites:

- web resources on geostatistics and spatial statistics can be found at <http://www.ai-geostats.org>;

- on machine learning: <http://www.kernel-machines.org/>, <http://www.support-vector.net/>; <http://mloss.org/about/> – machine learning open source software; <http://www.cs.iastate.edu/~honavar/Courses/cs673/machine-learning-courses.html> – index of ML courses, <http://www.patternrecognition.co.za/tutorials.html> – machine learning tutorials; very good tutorials on statistical data mining can be found on-line at <http://www.autonlab.org/tutorials/list.html>;

- Bayesian maximum entropy: some resources related to Bayesian maximum entropy (BME) methods. For a more complete list or references see Chapter 6; see also the BMELab site at <http://www.unc.edu/depts/case/BMELab>.

1.3.1. Books, tutorials

The list of books, given in the reference section below, is not complete but gives good references on introductory and advanced topics presented in the book. Some of these are more theoretical, while some concentrate more on the applications and case studies. In any case, most of them can be used as text books for the educational purposes as well as references for research.

1.3.2. Software

All contemporary data analysis and modeling approaches are not feasible without powerful computers and good software tools. This book does not include a CD with software modules (unfortunately). Therefore, below we would like to recommend some cheap and “easy to find” software with short descriptions.

- *GSLIB*: a geostatistical library with Fortran routines [DEU 1997]. The *GSLIB* library, which first appeared in 1992, was an important step in geostatistics applications and stimulated new developments. It gave many researchers and

students the possibility of starting with geostatistical models and learning corresponding algorithms having access to the codes. Description: the GSLIB modeling library covers both geostatistical predictions (family of kriging models) and conditional geostatistical simulations. There is a version of GSLIB with user interfaces which can be found at <http://www.statios.com/WinGslib>.

– *S-GeMS* is a piece of software for 3D geostatistical modeling. Description: it implements many of the classical geostatistics algorithms, as well as new developments made at the SCRF lab, Stanford University. It includes a selection of traditional and the most recent geostatistical models: kriging, co-kriging, sequential Gaussian simulation, sequential indicator simulation, multi-variate sequential Gaussian and indicator simulation, multiple-point statistics simulation, as well as standard data analysis tools (histogram, QQ-plots, variograms) and interactive 3D visualization. Open source code is available at <http://sgems.sourceforge.net>.

– *Geostat Office* (GSO). An educational version of GSO comes with a book [KAN 04]. The GSO package includes geostatistical tools and models (variography, spatial predictions and simulations) and neural networks (multilayer perceptron, general regression neural networks and probabilistic neural networks).

– *Machine Learning Office* (MLO) is a collection of machine learning software modules: multilayer perceptron, radial basis functions, general regression and probabilistic neural networks, support vector machines, self-organizing maps. MLO is a set of software tools accompanying the book [KAN 08].

– *R* (<http://www.r-project.org>). R is a free software environment for statistical computing and graphics. It is a GNU project which is similar to the S language and environment. There are several contributed modules dedicated to geostatistical models and to machine learning algorithms.

– *Netlab* [NAB 01]. This consists of a toolbox of Matlab® functions and scripts based on the approach and techniques described in “Neural Networks for Pattern Recognition” by Christopher M. Bishop, (Oxford University Press, 1995), but also including more recent developments in the field. <http://www.ncrg.aston.ac.uk/netlab>.

– *LibSVM*: <http://www.csie.ntu.edu.tw/~cjlin/libsvm> is quite a popular library for Support Vector Machines.

– *TORCH* machine learning library (<http://www.torch.ch>). The tutorial on the library, <http://www.torch.ch/matos/tutorial.pdf>, presents TORCH as a machine learning library, written in C++, and distributed under a BSD license. The ultimate objective of the library is to include all of the state-of-the-art machine learning algorithms, for both static and dynamic problems. Currently, it contains all sorts of artificial neural networks (including convolutional networks and time-delay neural networks), support vector machines for regression and classification, Gaussian mixture models, hidden Markov models, k-means, k-nearest neighbors and Parzen

windows. It can also be used to train a connected word speech recognizer. And last but not least, bagging and adaboost are ready to use.

– *Weka*: <http://www.cs.waikato.ac.nz/~ml/weka>. Weka is a collection of machine learning algorithms for data-mining tasks. The algorithms can either be applied directly to a dataset or taken from your own Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules and visualization. It is also well-suited for developing new machine learning schemes.

– *Machine Learning Open Source Software (MLOSS)*: <http://mloss.org/about>. The objective of this new interesting project is to support a community creating a comprehensive open source machine learning environment.

– *SEKS-GUI* (Spatiotemporal Epistematics Knowledge Synthesis software library and Graphic User Interface). Description: advanced techniques for modeling and mapping spatiotemporal systems and their attributes based on theoretical modes, concepts and methods of evolutionary epistemology and modern cognition technology. The interactive software library of SEKS-GUI explores heterogenous space-time patterns of natural systems (physical, biological, health, social, financial, etc.); accounts for multi-sourced system uncertainties; expresses the system structure using space-time dependence models (ordinary and generalized); synthesizes core knowledge bases, site-specific information, empirical evidence and uncertain data; and generates meaningful problem solutions that allow an informative representation of the real-world system using space-time varying probability functions and the associated maps (predicted attribute distributions, heterogeneity patterns, accuracy indexes, system risk assessment, etc.). <http://geography.sdsu.edu/Research/Projects/SEKS-GUI/SEKS-GUI.html>. Manual: Kolovos, A., H-L Yu, and Christakos, G., 2006. SEKS-GUI v.0.6 User Manual. Dept. of Geography, San Diego State University, San Diego, CA.

– *BMELib Matlab library* (Matlab®) and its applications can be found on <http://www.unc.edu/depts/case/BMElab/>.

1.4. Conclusion

The problem of spatial and spatiotemporal data analysis is becoming more and more important: many monitoring stations around the world are collecting high frequency data on-line, satellites produce a huge amount of information about Earth on a daily basis, an immense amount of data is available within GIS.

Environmental data are multivariate and noisy; highly variable at many geographical scales – from local variability in hot spots to regional trends; many of them are unique (only one realization of the phenomena under study); usually environmental data are spatially non-stationary.

The problem of the reconstruction of random fields using discrete data measurements has no single solution. Several important, and difficult to verify, hypotheses have to be accepted and tuning of the model-dependent parameters has to be carried out before arriving at a “unique and in some sense the best” solution.

In general, different data analysis approaches – both model-based and data-driven can be considered as complementary. For example, MLA can be efficiently used already at the phase of exploratory data analysis or for de-trending in a hybrid scheme. Moreover, there are links between these two groups of methods, such that, under some conditions, kriging (as a Gaussian process) can be considered as a particular neural network and vice versa.

Therefore, in this book, three currently different approaches are presented as possible solutions to the same problem of analysis and mapping of spatial data. Each of them has its own advantages and drawbacks in comparison with the others. In some cases they have quite unique properties to solve more specific tasks: geostatistical simulations and BME to model joint probability density functions (random fields), MLA when working with high dimensional and multivariate data. Hybrid models based on both approaches can overcome some difficulties and produce better results.

We propose to apply different methods and tools in order to produce alternative and complementary results which can improve decision-making processes.

1.5. References

- [ABE 05] ABE S., *Support Vector Machines for Pattern Classification*, Springer, 2005.
- [BIS 07] BISHOP C.M., *Pattern Recognition and Machine Learning*, Springer, 2007.
- [CHE 98] CHERKASSKY V. and MULIER F., *Learning from Data*, John Wiley & Sons, 1998.
- [CHI 99] CHILES J.-P., DELFINER P., *Geostatistics: Modelling Spatial Uncertainty*, Wiley series in probability and statistics, John Wiley and Sons, 1999.
- [CHR 92] CHRISTAKOS G., *Random Field Models in Earth Sciences*, Academic Press, San Diego, CA, 1992.
- [CHR 98] CHRISTAKOS G. and HRISTOPULOS D.T., *Spatiotemporal Environmental Health Modelling*, Kluwer Academic Publ., Boston, MA, 1998.
- [CHR 00a] CHRISTAKOS G., *Modern Spatiotemporal Geostatistics*, Oxford University Press, New York, 2000.
- [CHR 02c] CHRISTAKOS G., BOGAERT P. and SERRE M.L., *Temporal GIS*, Springer-Verlag, New York, NY, with CD-ROM, 2002.

- [CHR 05] CHRISTAKOS G., OLEA R.A., SERRE M.L., YU H.L. and WANG L-L., *Interdisciplinary Public Health Reasoning and Epidemic Modelling: The Case of Black Death*, Springer-Verlag, New York, NY, 2005.
- [CRE 93] CRESSIE, N., *Statistics for Spatial Data*, John Wiley and Sons, NY, 1993.
- [CRI 00] CRISTIANINI N. and SHAWE-TAYLOR J., *Support Vector Machines*, Cambridge University Press, 2000.
- [DAV 88] DAVID M., *Handbook of Applied Advanced Geostatistical Ore Reserve Estimation*, Elsevier Science Publishers, Amsterdam B.V., 216 p., 1988.
- [DEU 97] DEUTSCH C.V. and JOURNEL A.G., *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, 1997.
- [DOB 07] DOBESCH H., DUMOLARD P., and DYRAS I (eds.), *Spatial Interpolation for Climate Data: The Use of GIS in Climatology and Meteorology*, Geographical Information Systems series, ISTE, 2007.
- [DUB 03] DUBOIS G., MALCZEWSKI J., and DE CORT M. (eds.), *Mapping Radioactivity in the Environment, Spatial Interpolation Comparison 97*, European Commission, JRC Ispra, EUR 20667, 2003.
- [DUB 05] DUBOIS G. (ed.), *Automatic Mapping Algorithms for Routine and Emergency Data*, European Commission, JRC Ispra, EUR 21595, 2005.
- [DUD 01] DUDA R., HART P. and STORK D., *Pattern Classification*, 2nd edition, John Wiley & Sons, 2001.
- [GAN 63] GANDIN L.S., *Objective Analysis of Meteorological Fields*, Israel program for scientific translations, 1963, Jerusalem.
- [GOO 97] GOOVAERST P., *Geostatistics for Natural Resources Evaluation*, Oxford University Press, 1997.
- [GRU 06] DE GRUIJTER J., BRUS D., BIERKENS M.F.P. and KNOTTERS M., *Sampling for Natural Resource Monitoring*, Springer, 2006.
- [GUY 06] GUYON I., GUNN S., NIKRAVESH M., and ZADEH L. (eds.), *Feature Extraction: Foundations and Applications*, Springer, 2006.
- [HAS 01] HASTIE T., TIBSHIRANI R., and FRIEDMAN J., *The Elements of Statistical Learning*, Springer, 2001.
- [HAY 98] HAYKIN S., *Neural Networks: a Comprehensive Foundation*, Pearson Higher Education, 2nd edition, 842 p., 1999.
- [HIG 03] HIGGINS N. A. and JONES J. A., *Methods for Interpreting Monitoring Data Following an Accident in Wet Conditions*, National Radiological Protection Board, Chilton, Didcot, 2003.
- [HYV 01] HYVARINEN A., KARHUNEN J., OJA E., *Independent Component Analysis*, Wiley Interscience, 2001.
- [ISA 89] ISAACS E., SHRIVASTAVA M., *Applied Geostatistics*, Oxford University Press, 1989.

- [ISA 90] ISAAKS E. H. and SRIVASTAVA R. M., *An Introduction to Applied Geostatistics*, Oxford University Press, 1990.
- [JEB 04] JEBARA T., *Machine Learning: Discriminative and Generative*, Kluwer Academic Publ., 2004.
- [JOU 78] JOURNAL A.G. and HUIJBREGTS C.J., *Mining Geostatistics*, Academic Press, 600 p., London, 1978.
- [KAN 04] KANEVSKI and MAIGNAN, M., *Analysis and Modelling of Spatial Environmental Data*, EPFL Press, 2004.
- [KAN 08] KANEVSKI M., POZDNOUKHOV A. and TIMONIN V., *Machine Learning Algorithms for Environmental Spatial Data. Theory, Applications and Software*, EPFL Press, Lausanne, 2008.
- [KOH 00] KOHONEN T., *Self-Organising Maps*, Springer, NY, 2000.
- [LEE 07] LEE J and VERLEYSEN M., *Nonlinear Dimensionality Reduction*, Springer, NY, 2007.
- [LEN 06] LE N.D. and ZIDEK J.V., *Statistical Analysis of Environmental Space-Time Processes*, Springer, NY, 2006.
- [LLO 06] LLOYD C.D., *Local Models for Spatial Analysis*, CRC Press, 2006.
- [MAT 63] MATHERON G., *Principles of Geostatistics Economic Geology*, vol. 58, December 1963, p. 1246-1266.
- [MUL 07] MULLER W.G., *Collecting Spatial Data. Optimum Design of Experiments for Random Fields*, 3rd edition, Springer, NY, 2007.
- [NAB 01] NABNEY I., *Netlab: Algorithms for Pattern Recognition*, Springer, 2001.
- [RAS 06] RASMUSSEN C.E. and WILLIAMS C.K.I., *Gaussian Processes for Machine Learning*, MIT Press, 2006.
- [SCH 05] SCHABENBERGER O. and GOTWAY C., *Statistical Methods for Spatial Data Analysis*, Chapman and Hall/CRC, 2005.
- [SCH 06] SCHÖLKOPF B. *et al.* (eds.), *Semi-Supervised Learning*, Springer, 2006.
- [SCH 98] SCHÖLKOPF B., SMOLA A., and MÜLLER K., "Nonlinear Component Analysis as a Kernel Eigenvalue Problem", *Neural Computation*, vol. 10, 1998, p. 1299-1319.
- [SHA 04] SHAW-TAYLOR J. and CRISTIANINI N., *Kernel Methods for Pattern Analysis*, Cambridge University Press, 2004.
- [VAP 06] VAPNIK V., *Estimation of Dependences Based on Empirical Data* (2nd Edition), Springer, 2006.
- [VAP 95] VAPNIK V., *The Nature of Statistical Learning Theory*, Springer, 1995.
- [VAP 98] VAPNIK V., *Statistical Learning Theory*, Wiley, 1998.
- [WAC 95] WACKERNAGEL H., *Multivariate Geostatistics*, 3rd edition, Springer-Verlag, 387 p., Berlin, 2003.