Uncertainty in Spatial Models
Geostatistics and Machine Learning

Vasily Demyanov

Institute of Petroleum Engineering, Heriot-Watt University, Edinburgh, UK
vasily.demyanov@pet.hw.ac.uk

Presented at

"Analysing and Visualising Spatial Uncertainty" workshop

Environmental Statistics Section of the Royal Statistical Society

and The Food and Environment Research Agency

27 January 2010

Sources of Uncertainty

Natural system modelling is subject to:

- observation errors (calibration, positioning, etc)
- unknown free model parameters
- computational discretisation – solution errors
- conceptual uncertainty – model inadequacy

Uncertainty Modelling Questions

- How accurate is the prediction?
- What is the risk of taking a decision on the prediction?
- How variable and uncertain are spatial predictions?
- How the model uncertainty is propagated into further predictions?
- Where to obtain further measurements to improve the prediction quality?

Modelling Approaches

- Geostatistics
- Machine Learning
- Bayesian Maximum Entropy
- Uncertainty Quantification in Inverse Problems

Spatial Modelling Approaches

- Deterministic
  - rely on analytical assumptions about model dependencies
  - uncertainty quantification is limited to parameter selection
- Geostatistics
  - stochastic nature of data $Z(x) = m(x) + S(x)$
  - spatial correlation (covariance) model
  - family of kriging models (regression)
  - stochastic simulations (multiple realisations)
- Machine learning
  - data driven approach
  - model choice is based on the learning principles
  - suffer from poor quality and insufficient amount of data

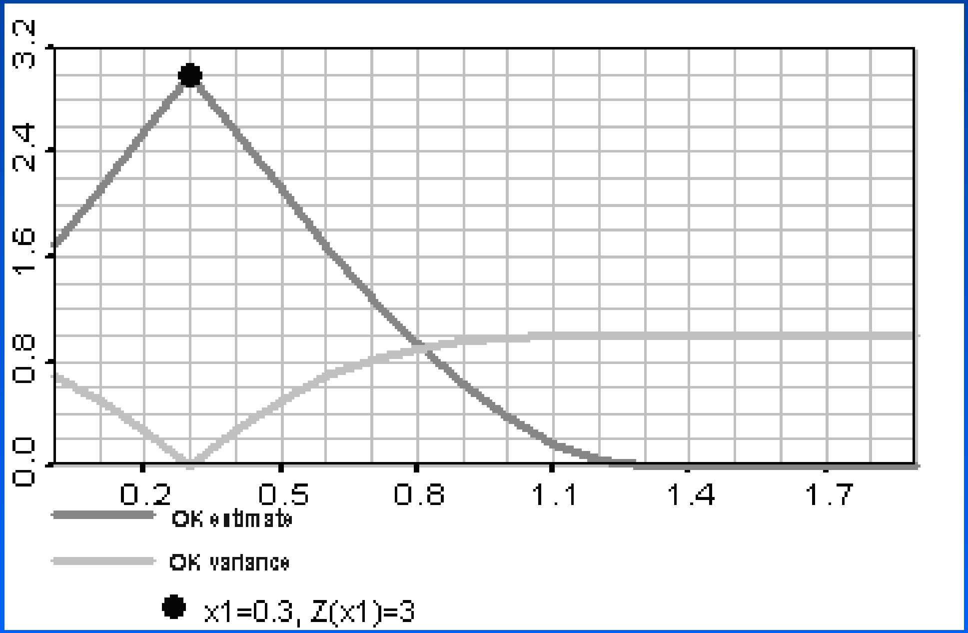
Uncertainty Modelling Questions

- How accurate is the prediction?
- What is the risk of taking a decision on the prediction?
- How variable and uncertain are spatial predictions?
- How the model uncertainty is propagated to the further predictions?
- Where to obtain further measurements to improve the prediction quality?

Geostatistical Predictors

Kriging:

- Unbiased linear estimator
- “Best” in terms of minimum variance
- Honours the data



Ordinary kriging estimate:

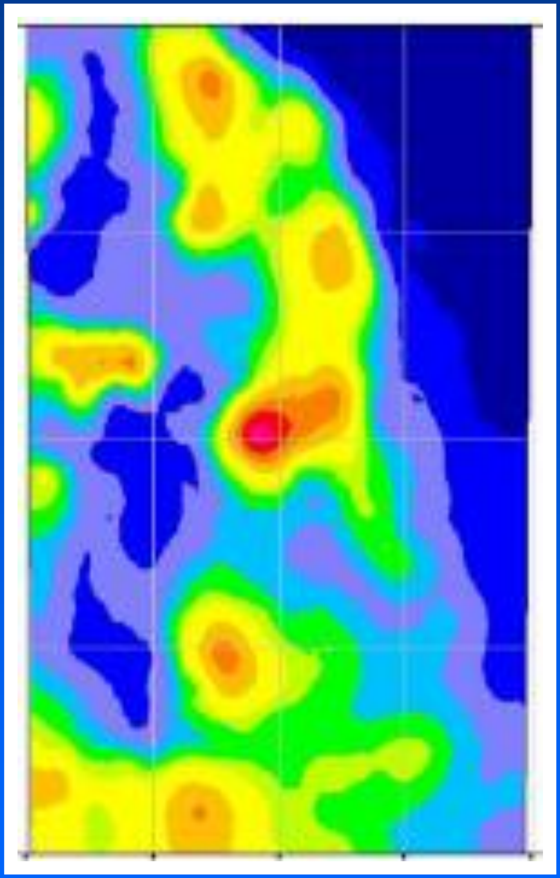
$$Z^*(x_0) = \sum_{j=1}^{n} w_j Z(x_j)$$

Ordinary kriging variance:

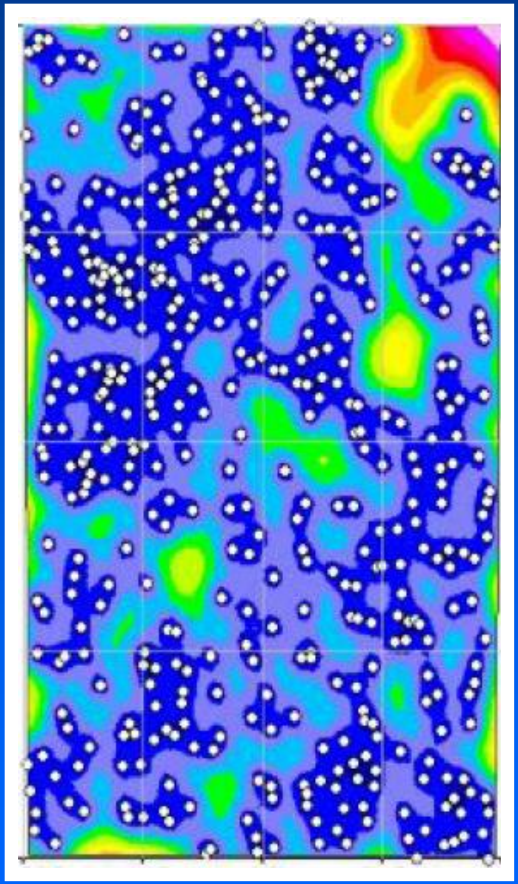
$$\sigma_{OK}^2(x) = \sigma_Z^2 - \sum_{j=1}^{n(x)} w_j(x) C_{j0} + \mu(x)$$

Kriging Predictions

Kriging estimate



Kriging variance



Kriging variance is unconditional:

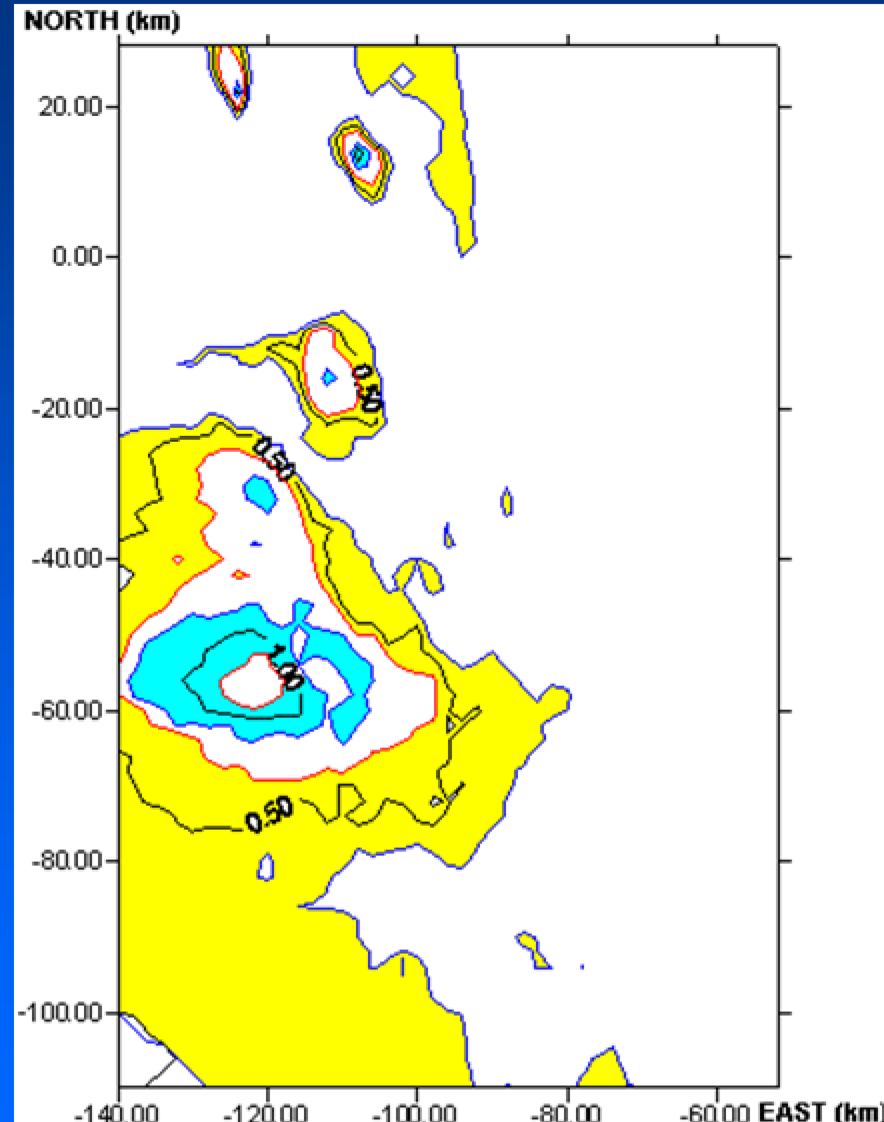
- depends on the data density
- does not reflect function value

Chernobyl radioactive soil contamination

Uncertainty Visualisation: Thick Contours

Kriging variance determines the thickness of the uncertainty interval around the kriging estimate contour

Chernobyl radioactive soil contamination



Uncertanity contours for isoline at 0.5 Ci/sq.km

Uncertanity contours for isoline at 1.0 Ci/sq.km

M. Kanevski, et.al. Mapping of Radioactively Contaminated Territories with Geostatistics and Artificial Neural Networks. Contaminated Forests , Kluwer 1999

Learning from Data

- Conventional model driven approach
  - Develop a model with known functional dependencies
  - Fit the model parameters to the available data
- Data driven approach
  - Model dependencies are not explicitly defined in functional form due to their complexity or lack of knowledge
  - Model dependencies are extracted from data via training
- Artificial Neural Networks can be trained to
  - store, recognise, and associatively retrieve patterns;
  - filter noise from measurement data;
  - estimate functions of unknown analytical form.

Learning Approaches

- Supervised learning
  - Inputs and the corresponding known target output values are available form the training set;
  - ANN output is computed for each set of inputs presented to ANN;
  - ANN output is compared with the target value;
  - ANN weights are updated to minimise error measure between the ANN output and the target output.

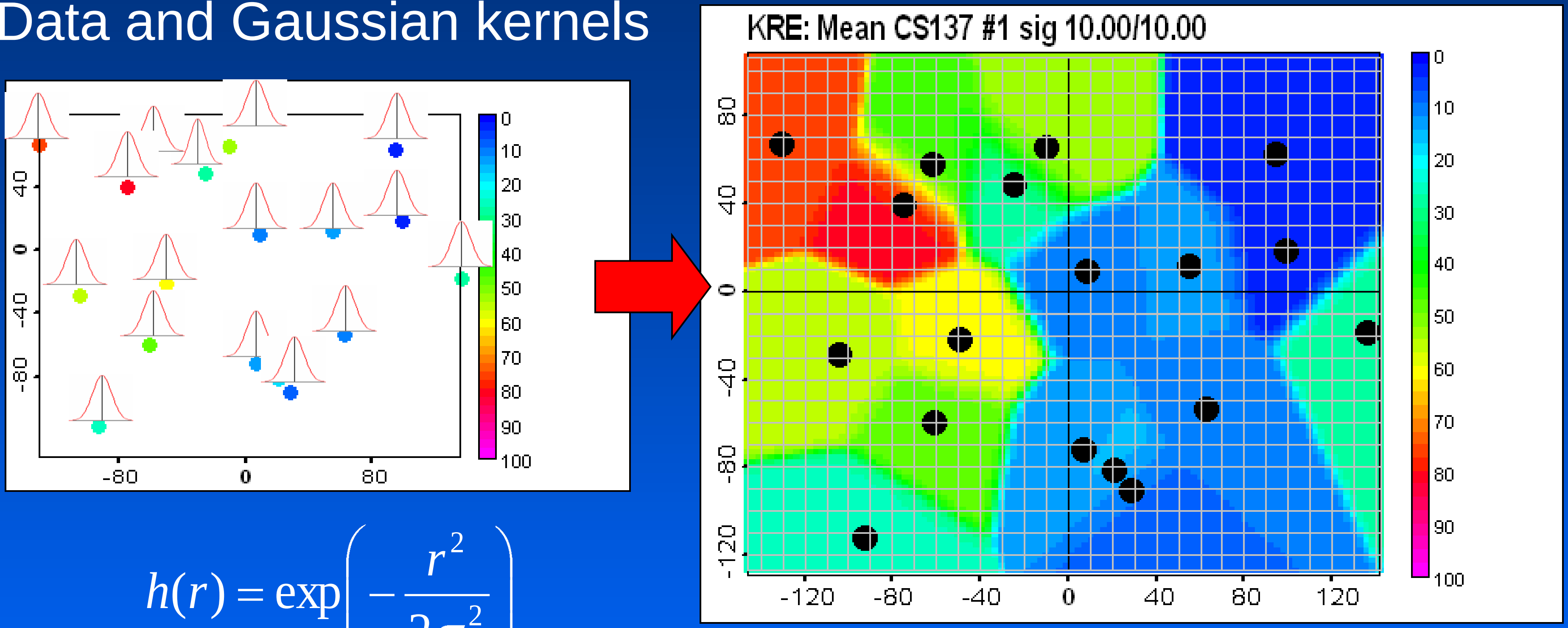
- Unsupervised learning
  - A set of inputs is presented to the network with no target outputs
  - Inputs are assumed to belong to several classes and the ANN output is an identification of the class to which its input belongs.
  - Competitive learning rule may be used: a “winner-takes-all”

Learning Approahes

- Semi-supervised learning
  - Use labelled and unlabeled data for training and prediction
- Active Learning
  - Interaction between User and learning machine
  - Update the training data set with particular new samples
- Reinforcement Learning
  - Learn how to act given an observation
- Transduction
  - Predict new outputs based on training data and new inputs

General Regression Neural Network (GRNN)

Data and Gaussian kernels

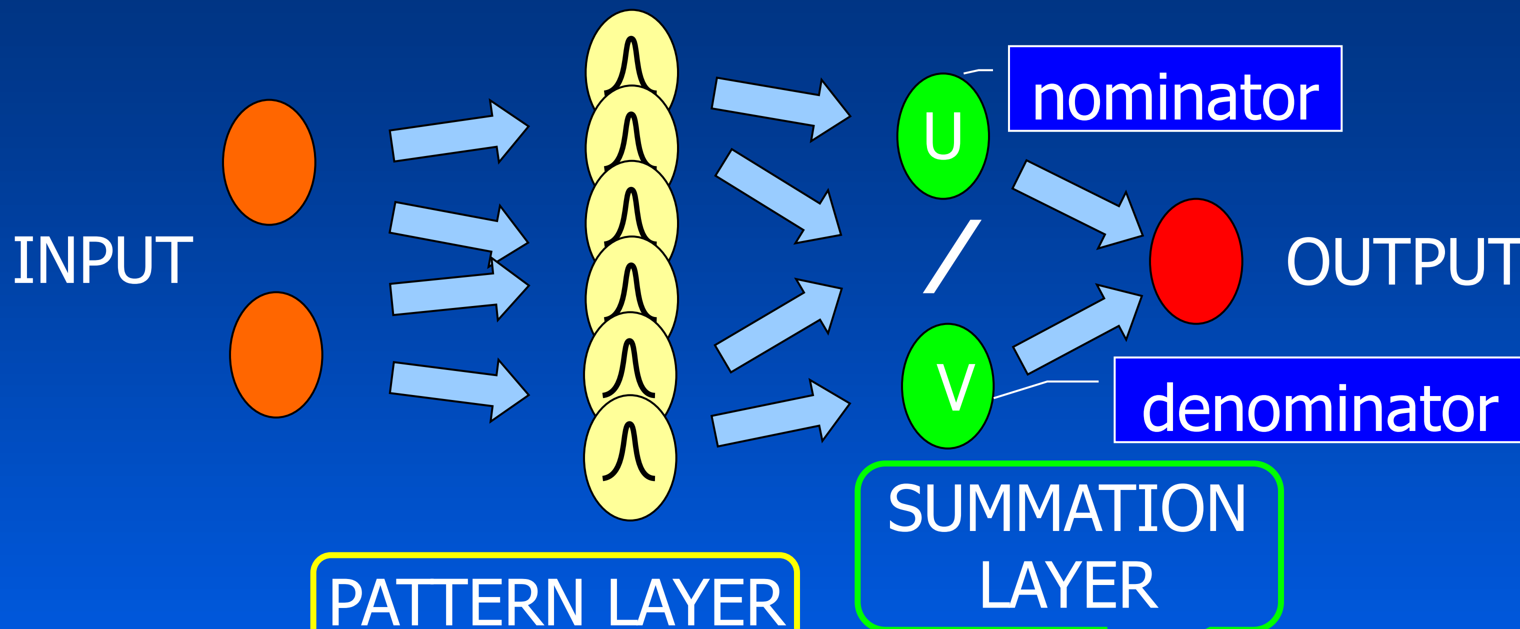


$$h(r) = \exp\left(-\frac{r^2}{2\sigma^2} \right)$$

Kernel smoothing parameter: $\overleftrightarrow{\sigma}=10$

How to choose the kernel width σ? – Cross-validation

General Regression Neural Network (GRNN)



Pattern layer consists of N kernels – one for every available data point

Signal from pattern units:

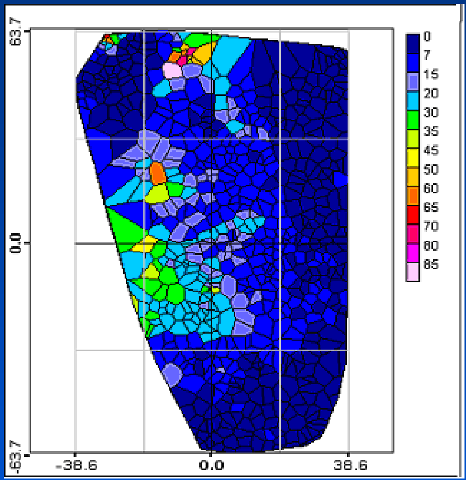
$$U = \sum_{i=1}^{N} Z_i \exp\left(-r_i^2 / 2\sigma^2\right)$$

Weights:

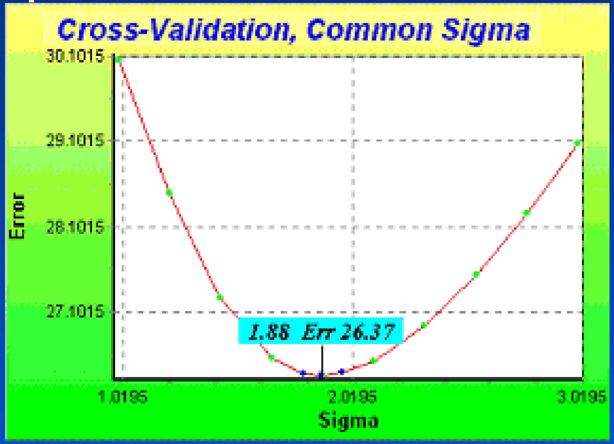
$$V = \sum_{i=1}^{N} \exp\left(-r_i^2 / 2\sigma^2\right)$$

GRNN Mapping

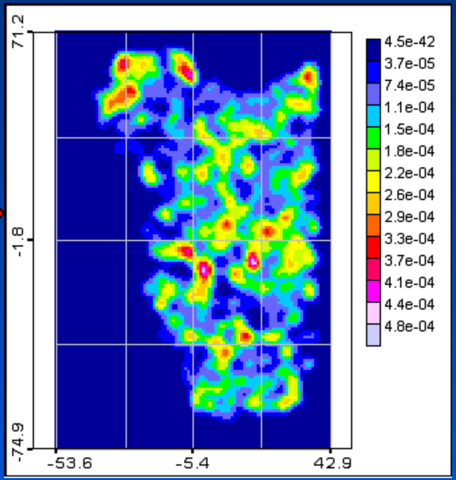
Data



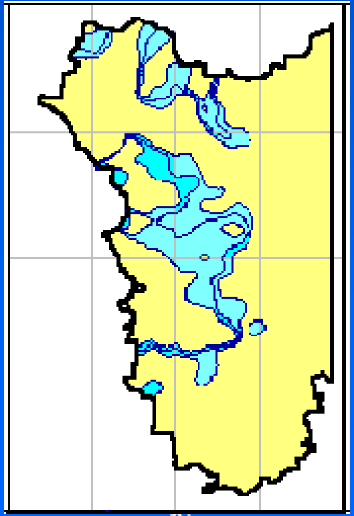
Optimise kernel width σ



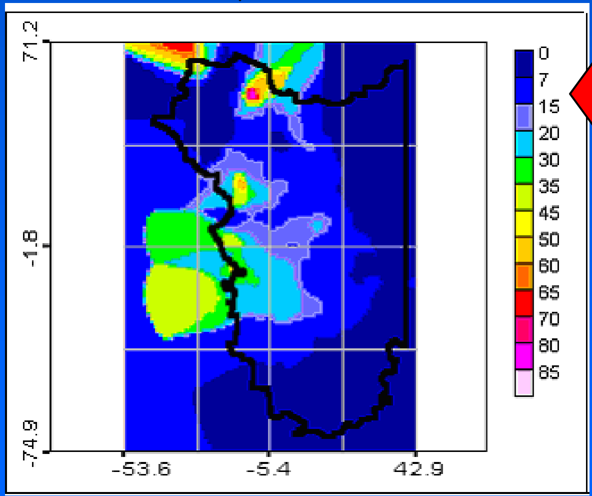
Kernel density



residual estimation



Thick contours



GRNN estimate

Uncertainty Quantification

Chernobyl radioactive soil contamination

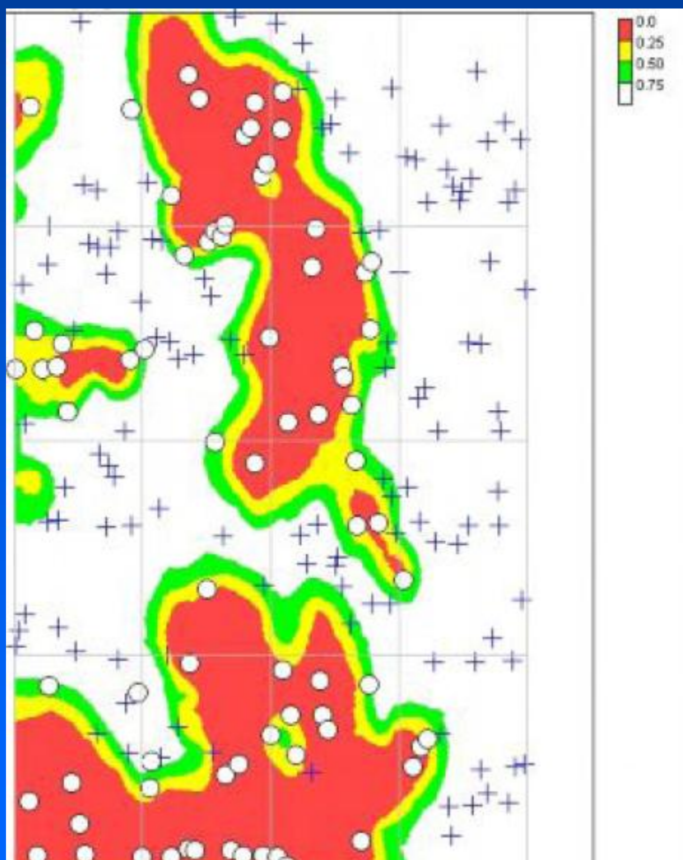
M.Kanevski, V.Timonin

Uncertainty Modelling Questions

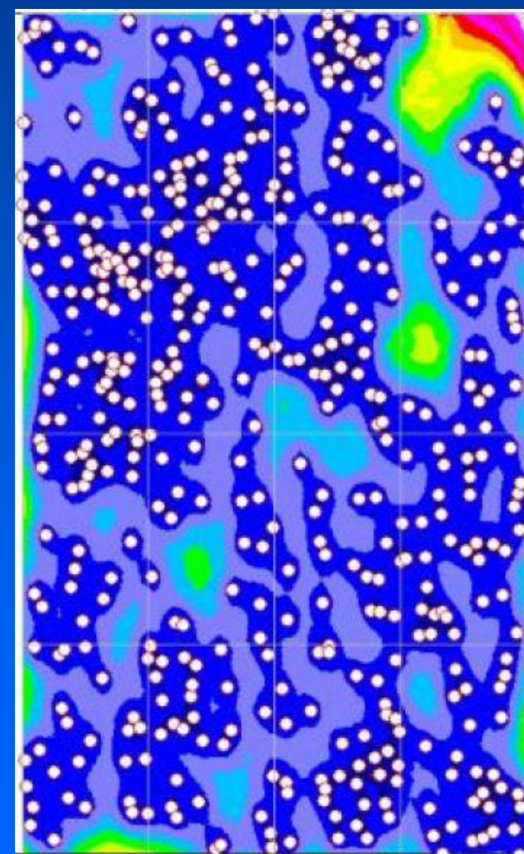
- How accurate is the prediction?
- What is the risk of taking a decision on the prediction?
- How variable and uncertain are spatial predictions?
- How the model uncertainty is propagated to the further predictions?
- Where to obtain further measurements to improve the prediction quality?

Probabilistic Prediction with Indicator Kriging

Probability to be below a level based on local cdf



Indicator kriging variance



○ data locations

Kriging variance is unconditional:

- depends on data density
- does not reflect function value

Chernobyl radioactive soil contamination

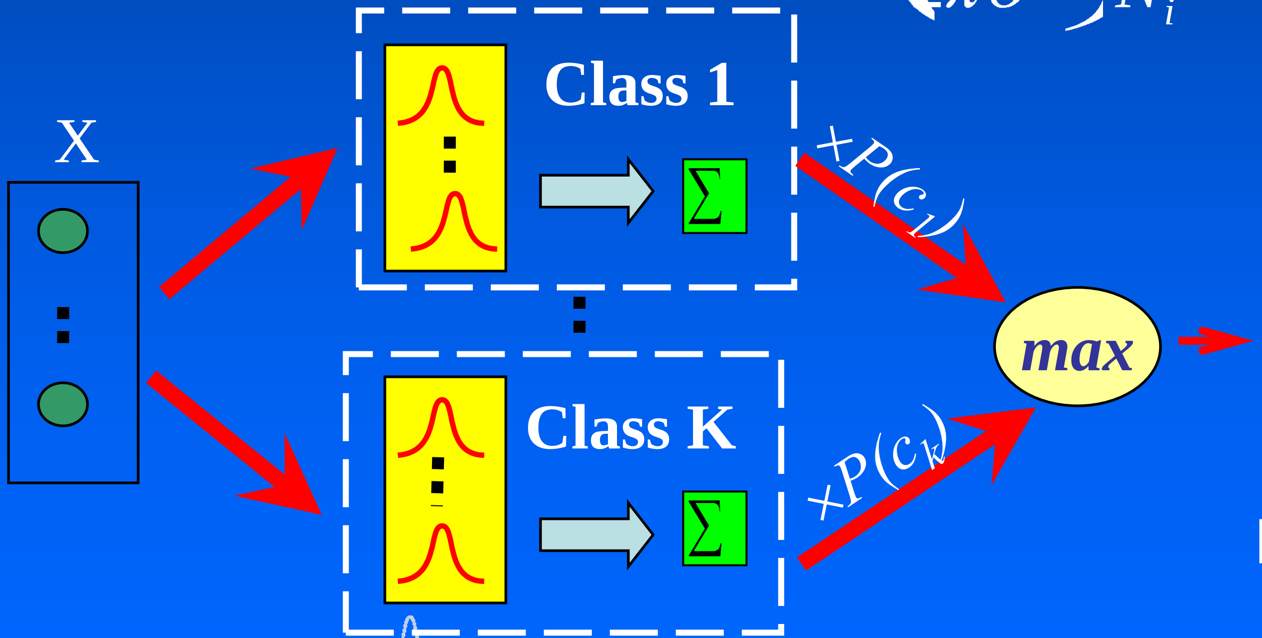
Probabilistic Neural Network (PNN)

Classification model based on maximum posterior decision rule:

$$C(x) = \{c_1, \ldots, c_K\} \arg\max_{c_i} P(c_i)\,p(x|c_i)$$

Class probability density kernel estimator:

$$p(x|c_i) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}} N_i} \sum_{n=1}^{N_i} \exp\left(\frac{-\left\| x - x_i^{(n)} \right\|^2}{2\sigma^2} \right)$$

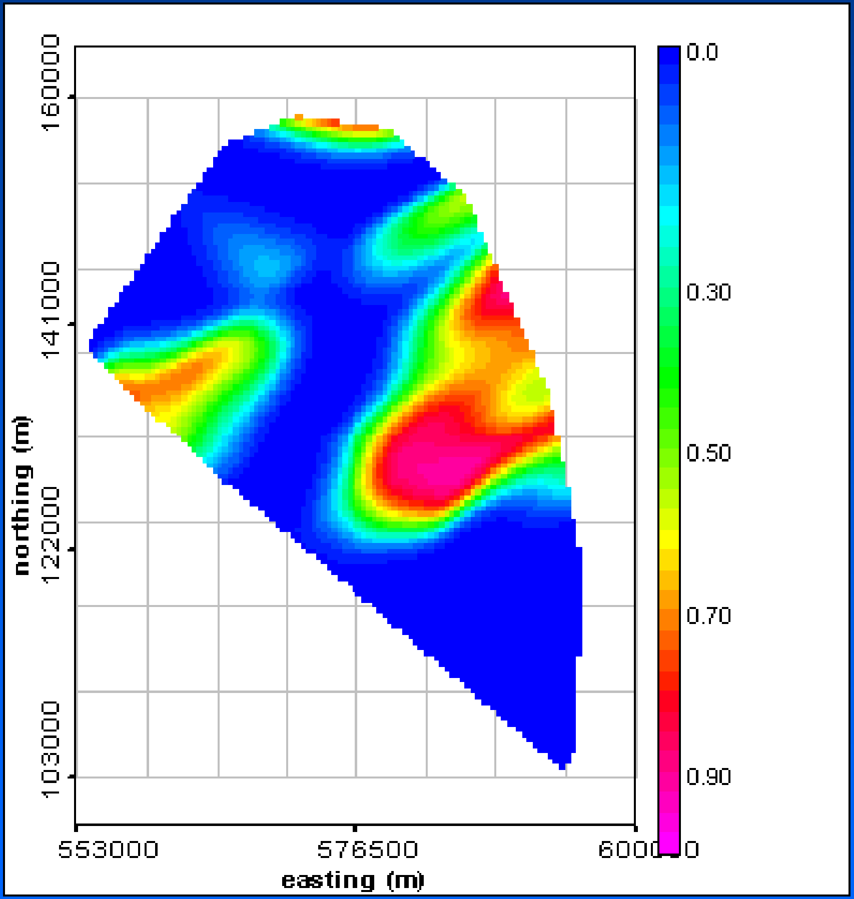


$$P(x|c_i) = \frac{P(c_i)\,p(x|c_i)}{\sum_{k-1}^{K} P(c_k)\,p(x|c_k)}$$

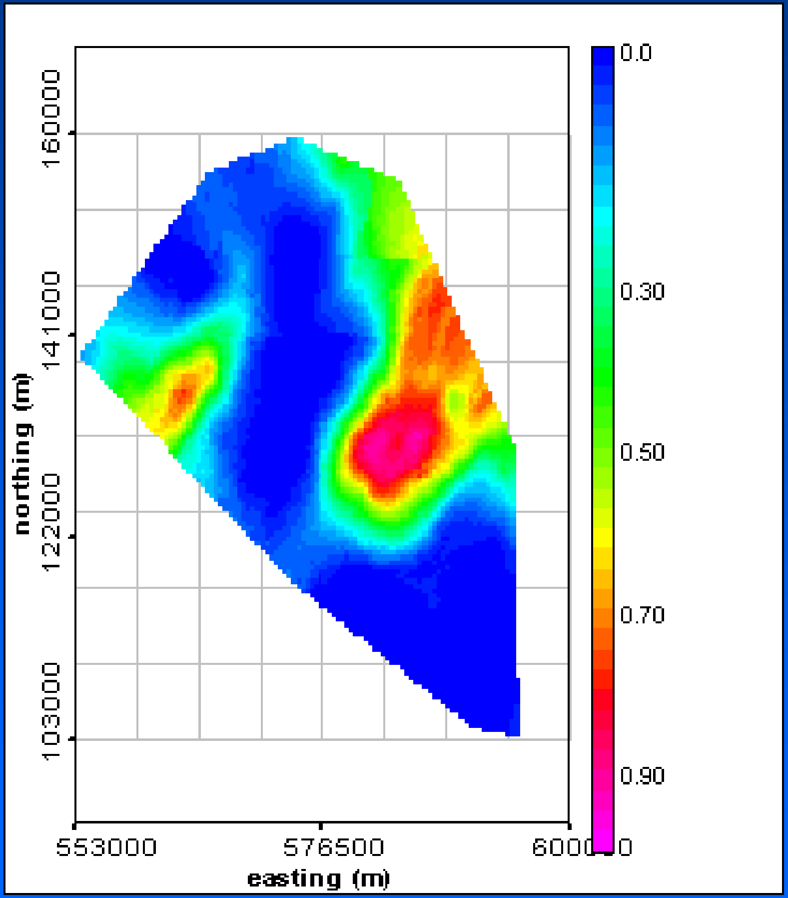
Posterior probability to belong to class *i*:

Probability Class Predictions

PNN prediction



Indicator kriging prediction



Probability of Ringold lower mud presence

Kansvski et.al, Characterization of Hydrogeologic Systems with Machine Learning Algorithms and Geostatistical Models, 2001

Uncertainty Modelling Questions

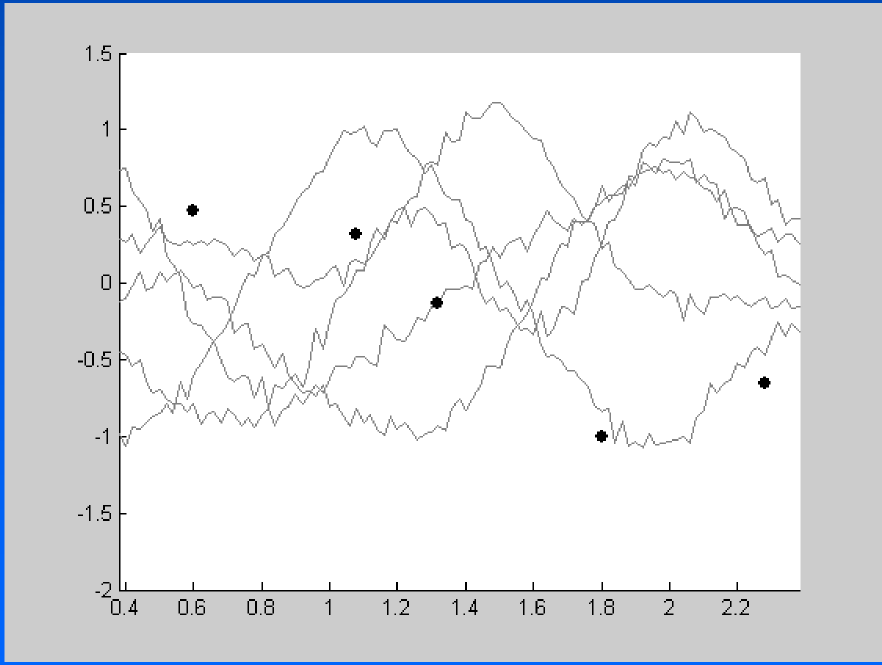
- How accurate is the prediction?
- What is the risk of taking a decision on the prediction?
- How variable and uncertain are spatial predictions?
- How the model uncertainty is propagated to the further predictions?
- Where to obtain further measurements to improve the prediction quality?

Geostatistical Simulations

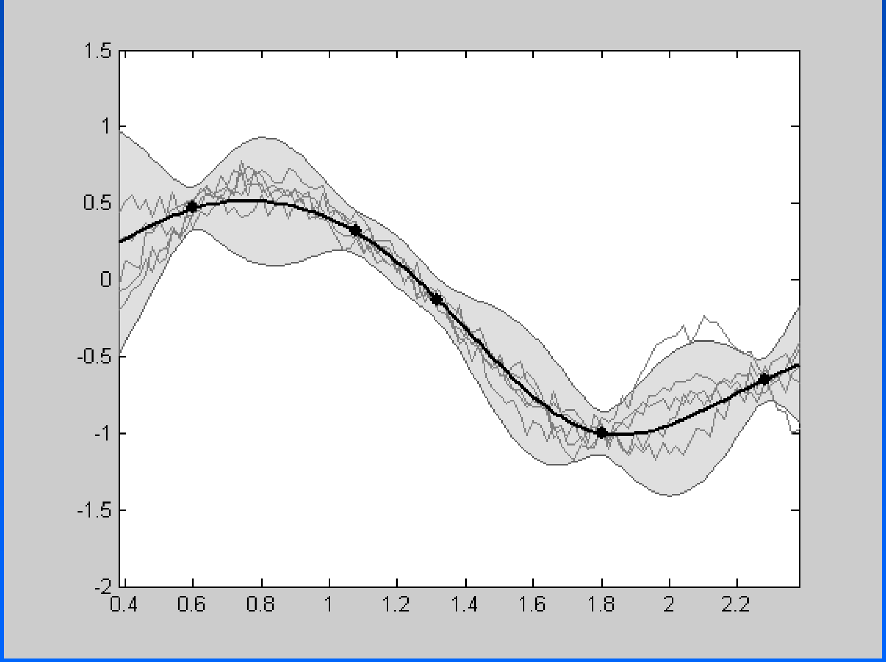
Kriging estimates vs. stochastic simulations:

- Stochastic realisations describe variability and local uncertainty

Unconditional

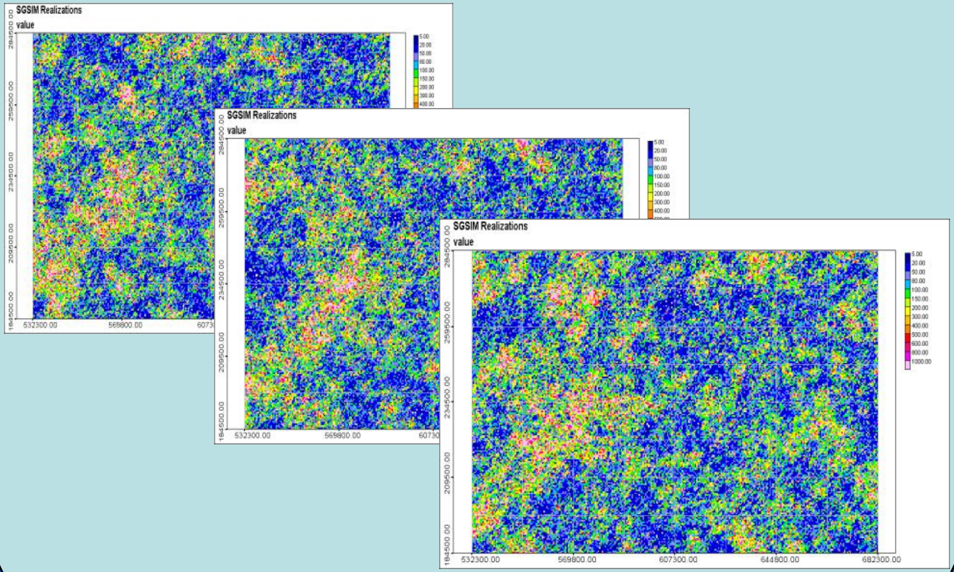


Conditional



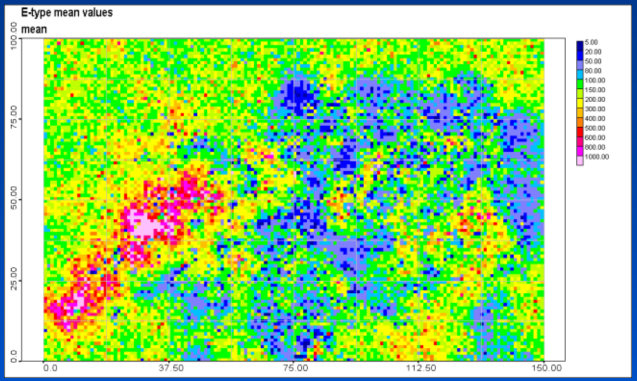
Stochastic Realisations

Multiple realisations

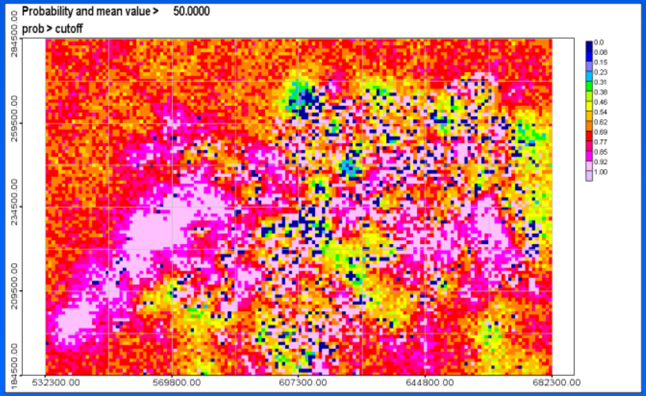


averaging

E-type estimate



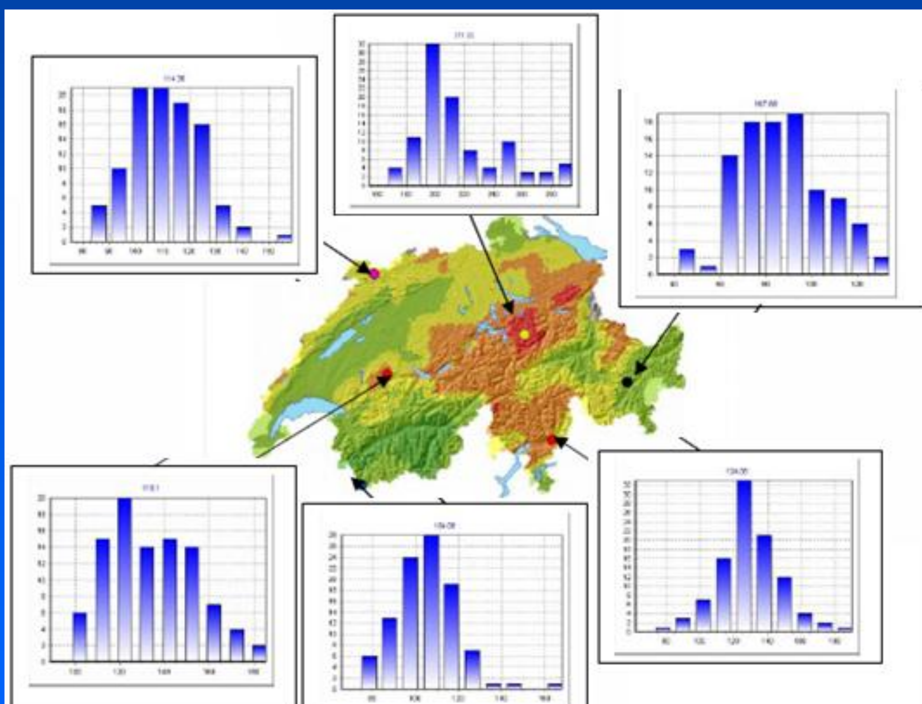
Probability above a level



Demyanov et al., Indoor Radon Risk Assessment with Geostatistics and Artificial Neural Networks, Geostatistics 2000

Stochastic Variability

Local distributions based on multiple realisations



Variance of a set of stochastic realisations is conditional to the functional value unlike kriging variance

Monthly precipitation, Switzerland

Bayesian Maximum Entropy

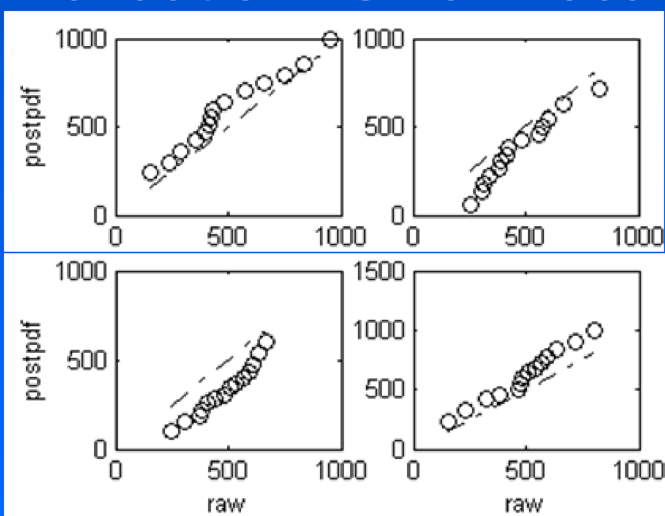
Epistemic principle – information is maximised subject to available general knowledge

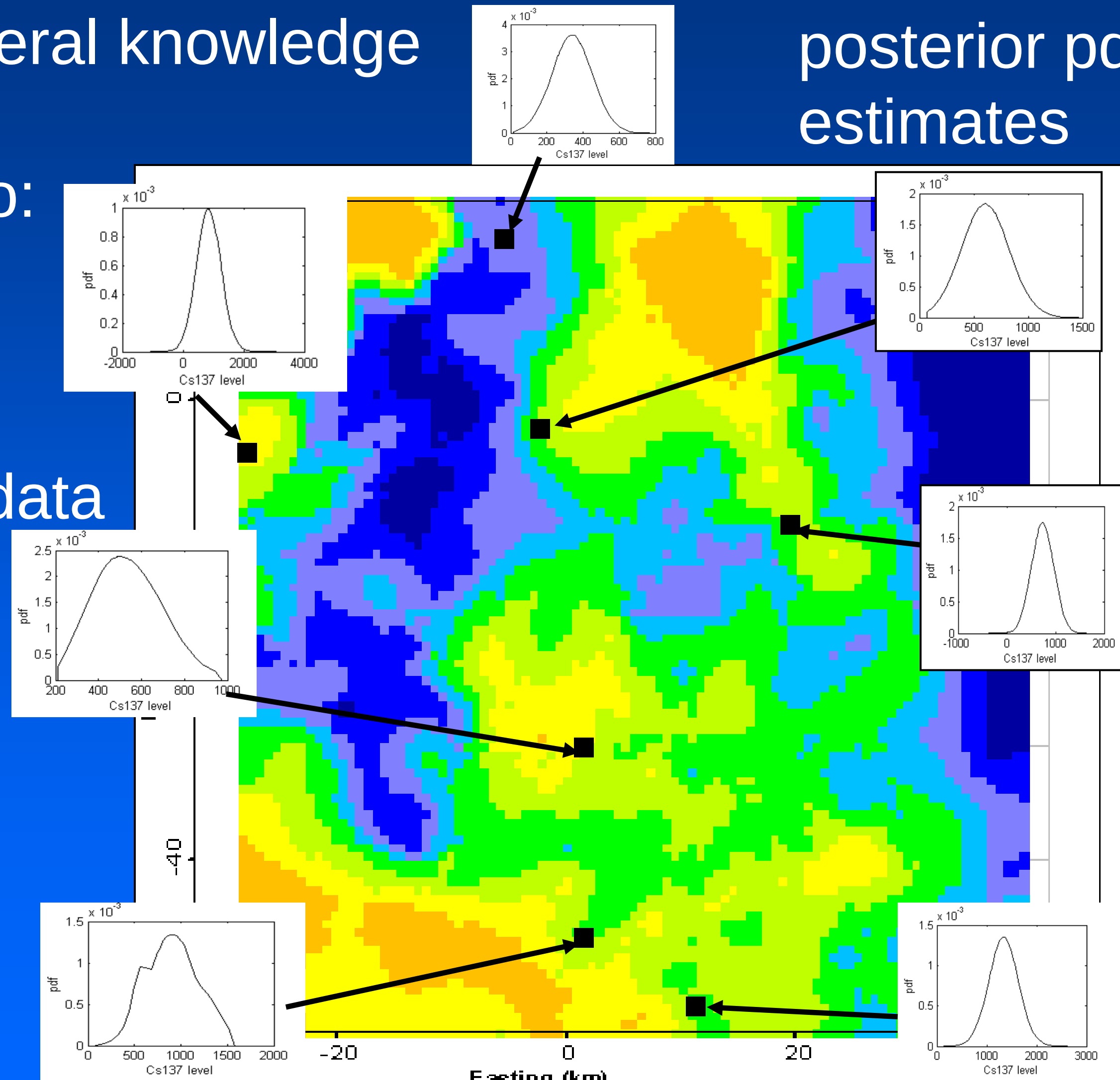
posterior pdf estimates

Bayesian conditioning to:

- Hard data
- Soft data

Validation vs. raw local data



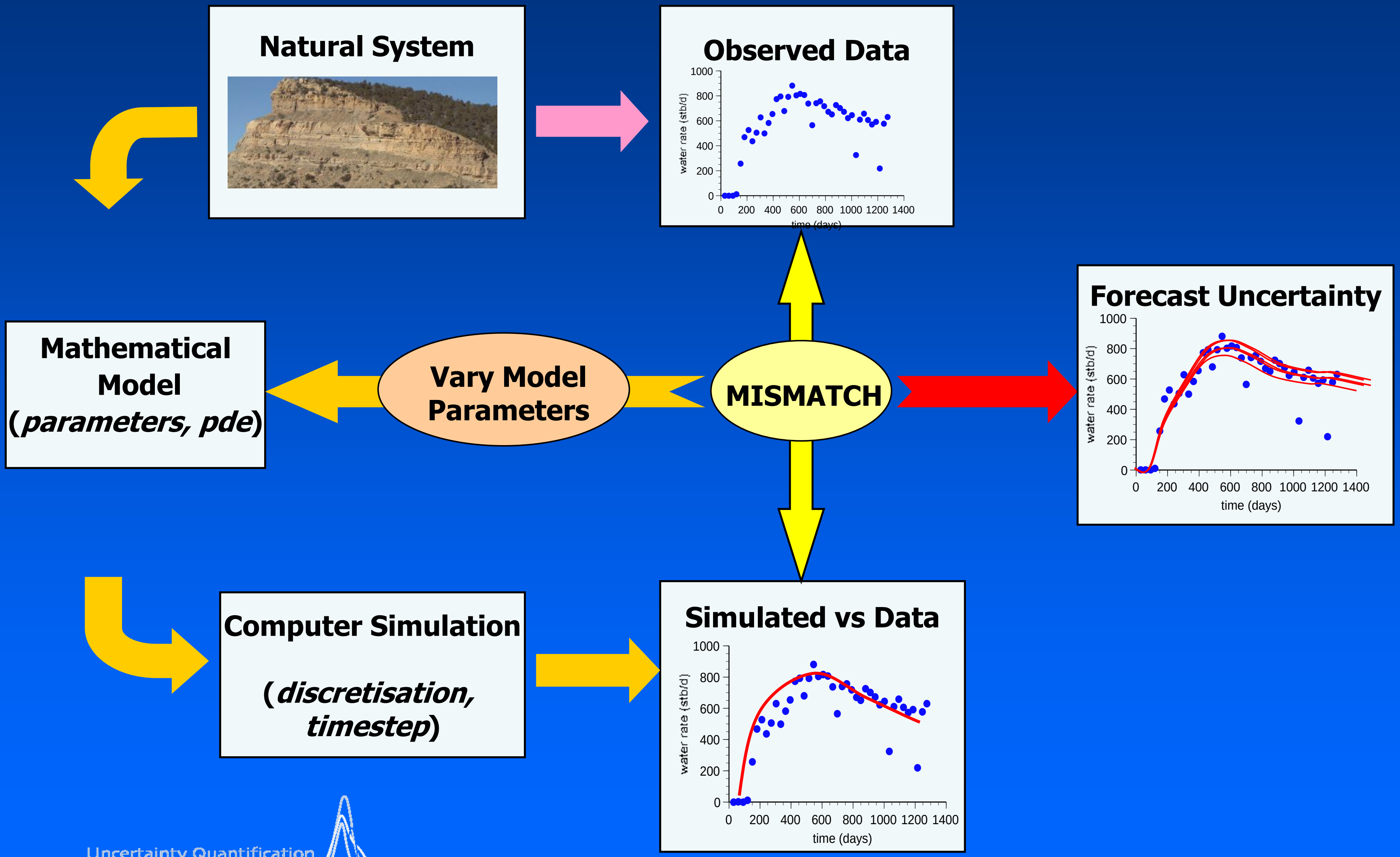


E.Savelieva, et.al. BME-based uncertainty assessment of Chernobyl fallout, Geoderma,2005

Uncertainty Modelling Questions

- How accurate is the prediction?
- What is the risk of taking a decision on the prediction?
- How variable and uncertain are spatial predictions?
- How the model uncertainty is propagated to the further predictions?
- Where to obtain further measurements to improve the prediction quality?

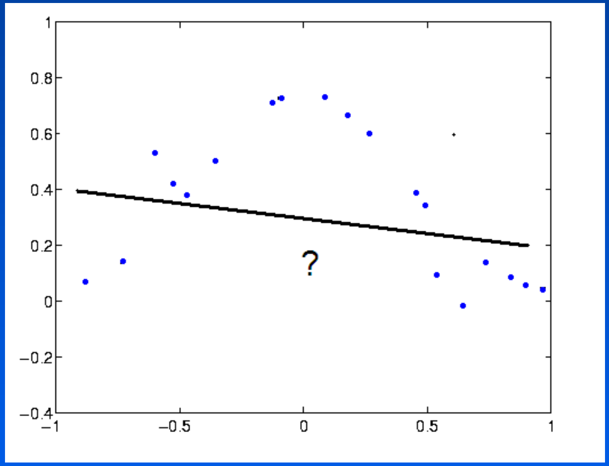
Uncertainty Quantification Framework

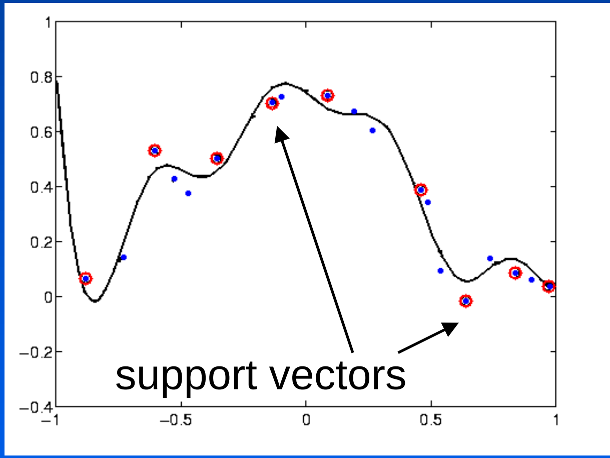


Support Vector Regression (SVR)

Linear regression in hyperspace

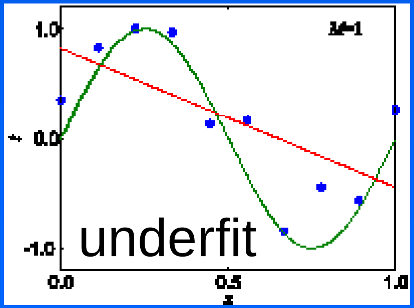
Kernel trick projects data into sufficiently high dimensional space: $f(x) = wx + b \longrightarrow f(x) = \sum_{i=1}^{L} y_i \alpha_i K(x, x_i) + b$

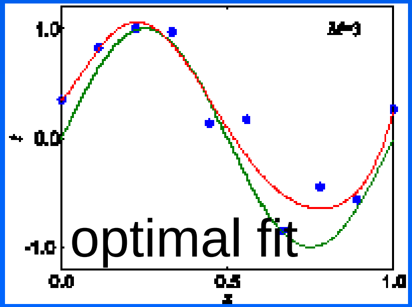


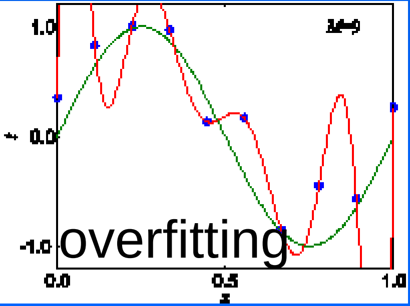


Complexity control with training errors

- data
- model
- truth

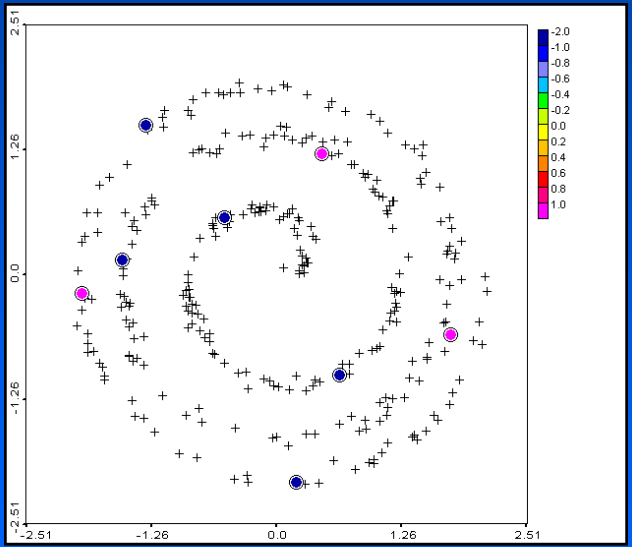




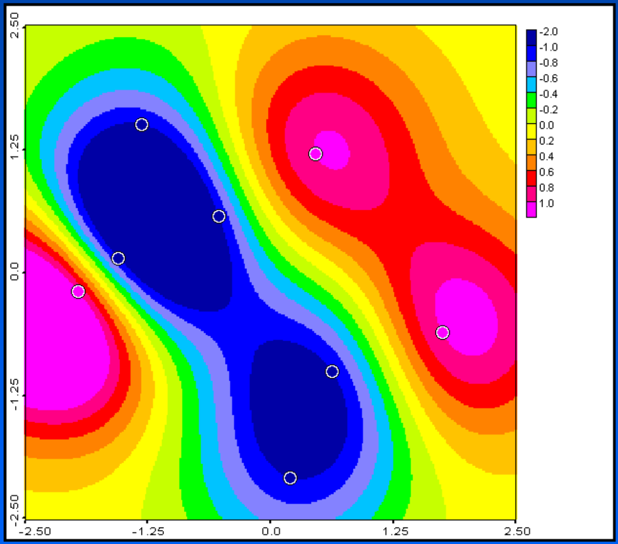


Semi-supervised SVR Model

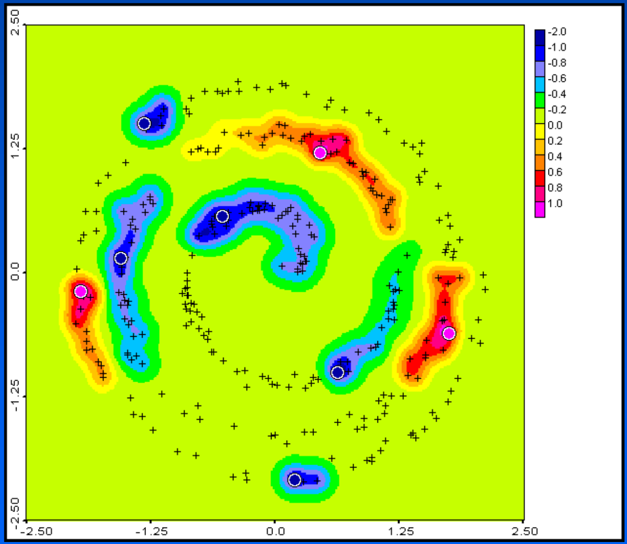
- Incorporate prior knowledge as graphs in input space
- Kernel function enforces continuity along the graph model – manifold – obtained from the prior information



Spiral manifold represented by unlabelled points (+)

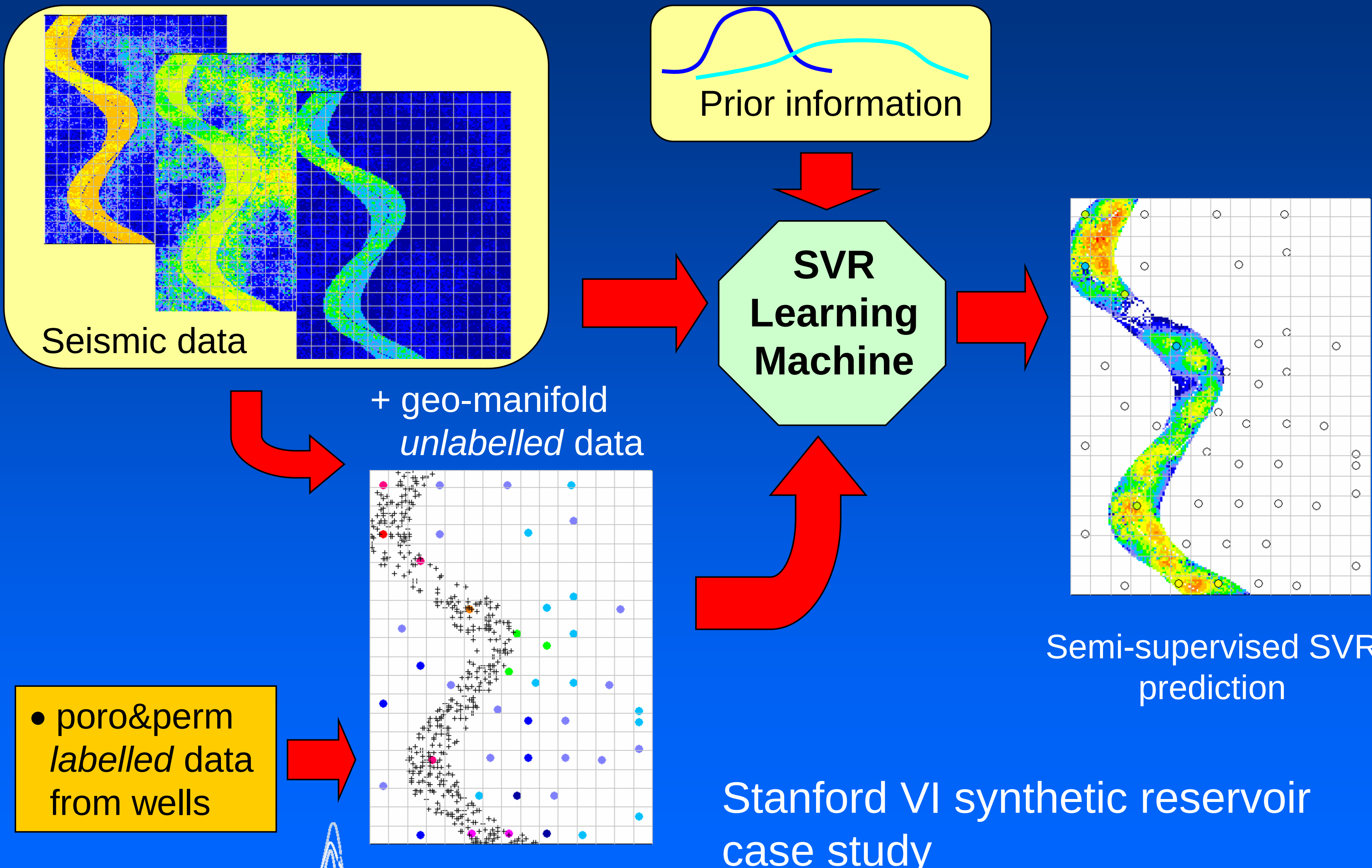


Conventional regression estimate based on labelled data only (●)



Semi-supervised regression estimation follows the smoothness along the graph

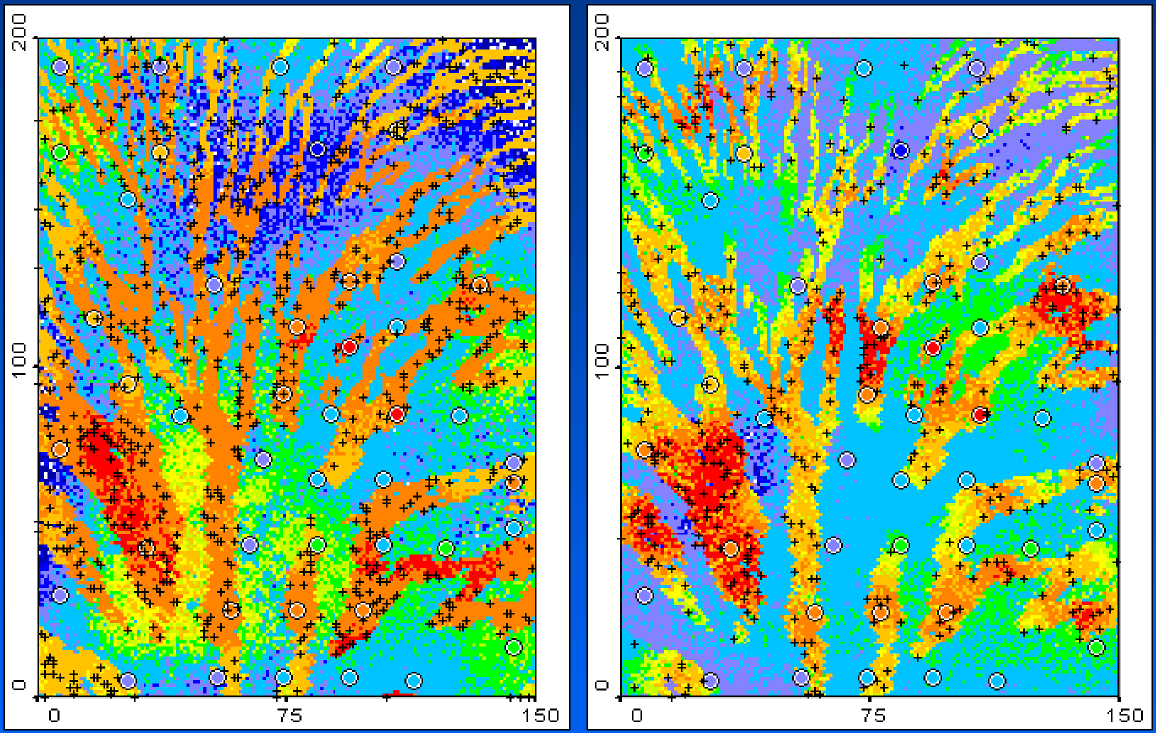
Semi-supervised SVR Reservoir Geomodel



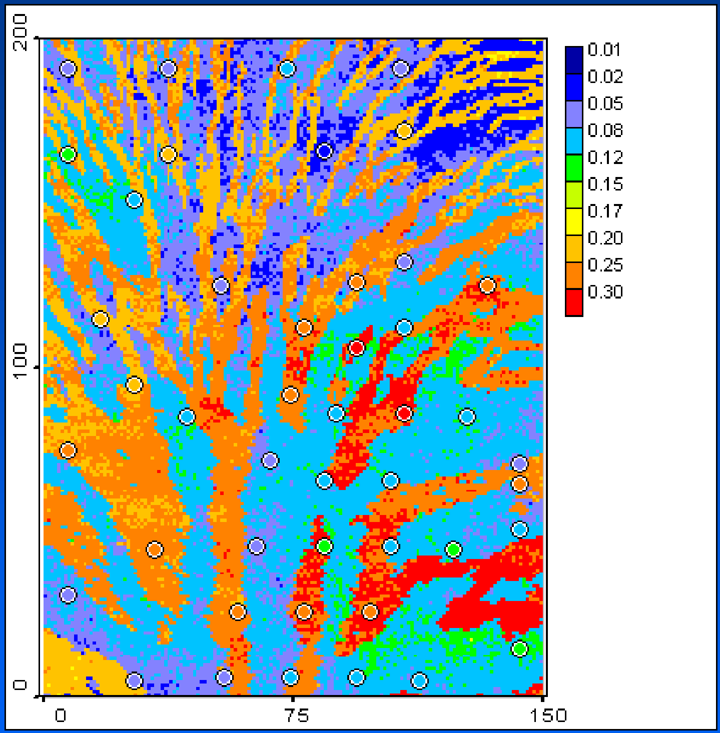
V. Demyanov, et.al. Geomodelling of a Fluvial System with Semi-Supervised Support Vector Regression, Geostatistics Congress, 2008

Multiple Realisations vs Truth case

Multiple good fitting porosity models



Truth case porosity



labelled (hard) data

\+ unlabelled data

- The river delta front structure is preserved
- Data conditioning
- Local spatial variability

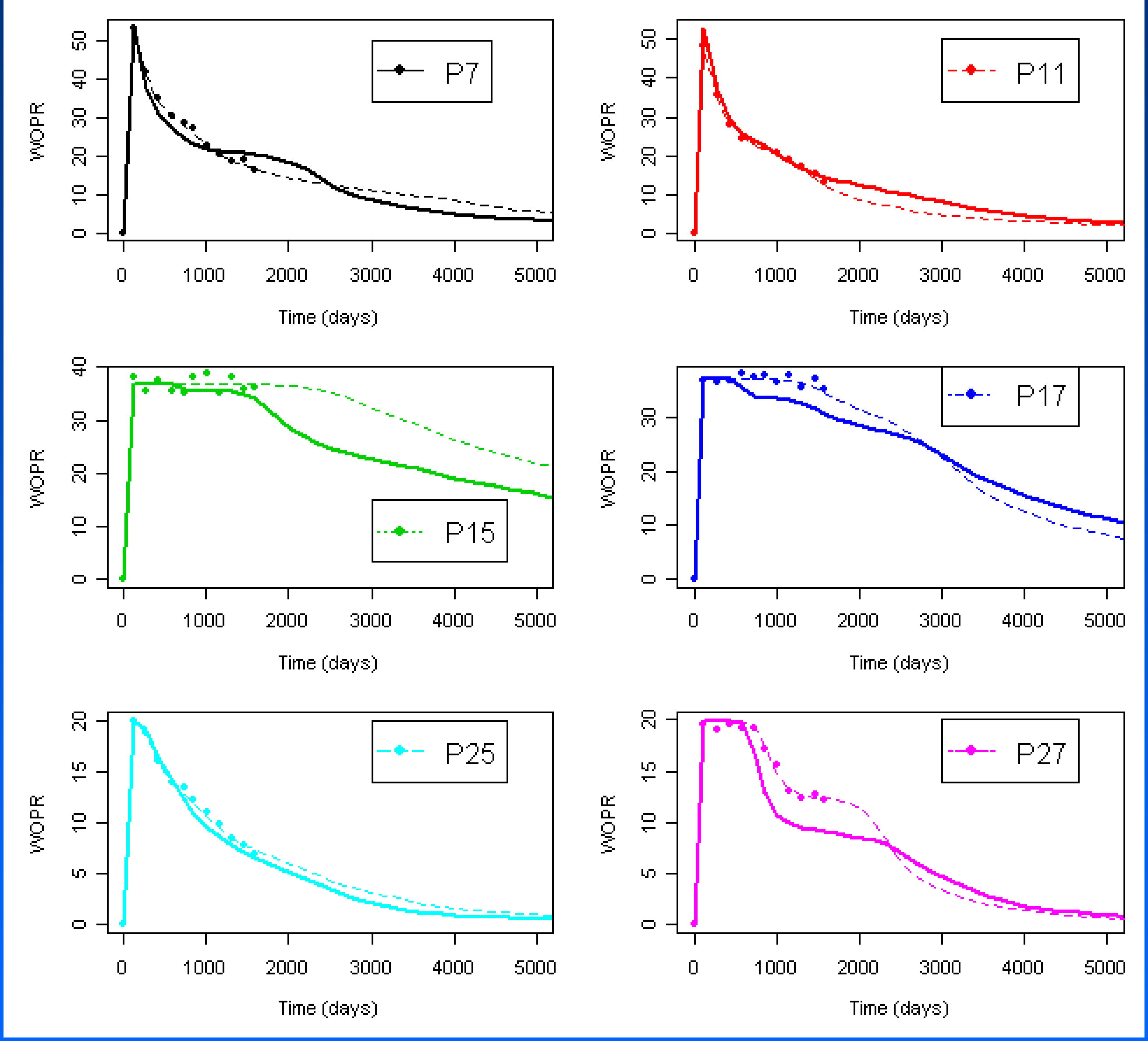
Model Forecast: Production Profiles

Oil production from 6 largest producing wells:

Past history data (truth case + noise)

— Fitted model

- - Truth case



Forecast with Uncertainty

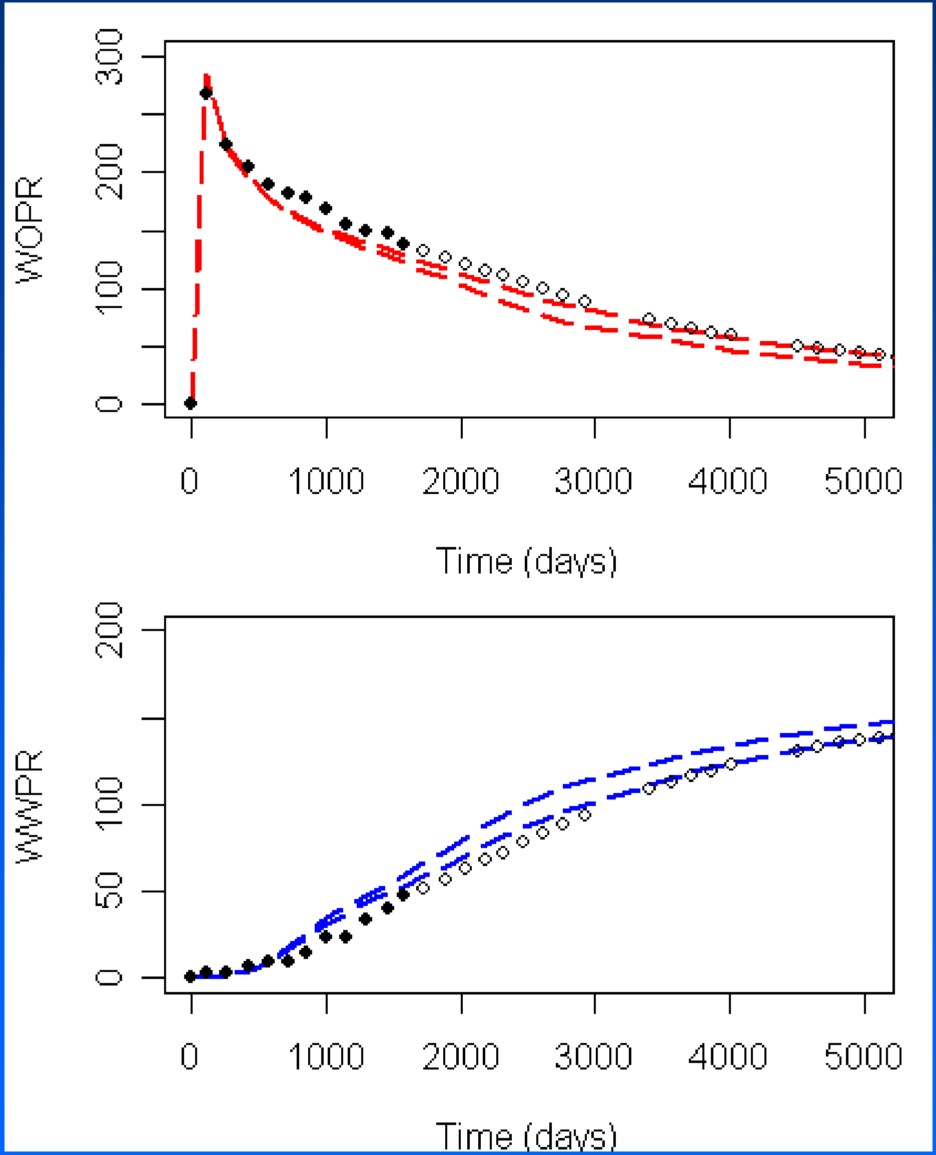
Confidence P10/P90 interval for production forecast based on multiple models

Total oil and water production profiles:

Past history data (truth case + noise)

Validation truth case forecast data

– – P10/P90 production forecast confidence bounds



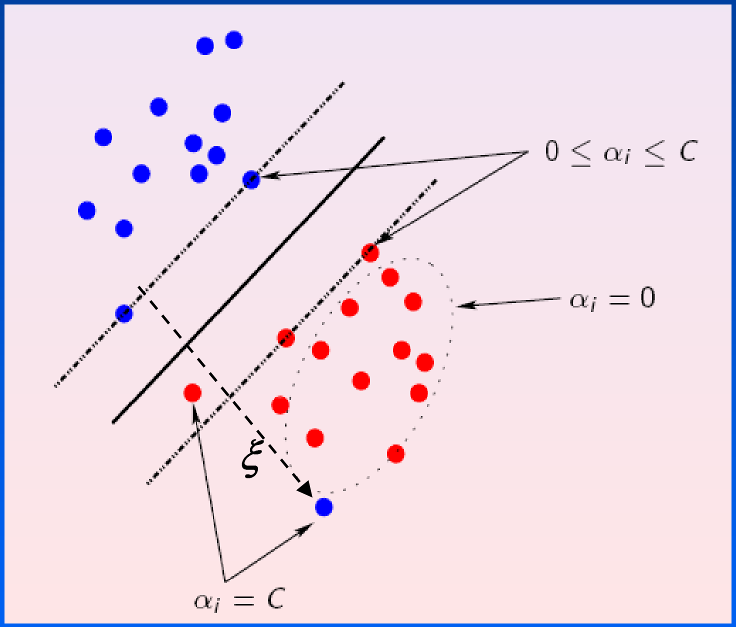
Uncertainty Quantification

Uncertainty Modelling Questions

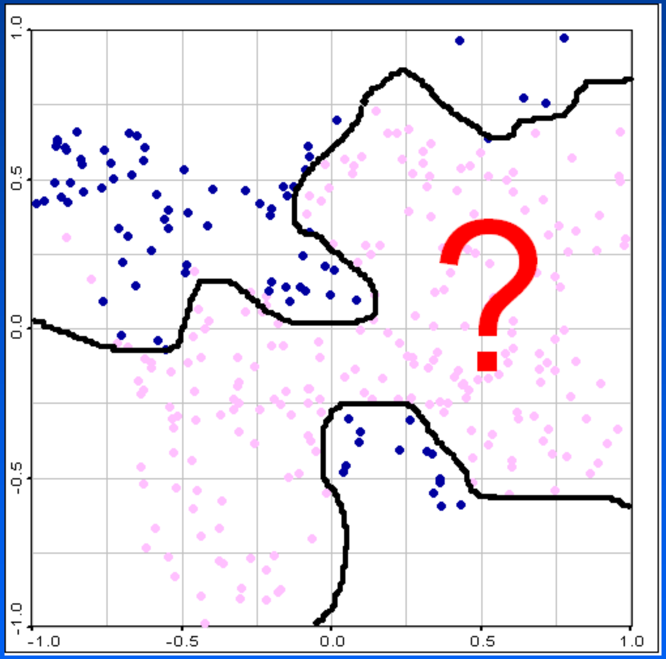
- How accurate is the prediction?
- What is the risk of taking a decision on the prediction?
- How variable and uncertain are spatial predictions?
- How the model uncertainty is propagated to the further predictions?
- Where to obtain further measurements to improve the prediction quality?

Monitoring Optimisation

- Monitoring Network
- Classification problem
- SVM classification model



- Find locations for additional measurements to refine the current model



Support Vectors (SVs): $0 < \alpha_i < C$

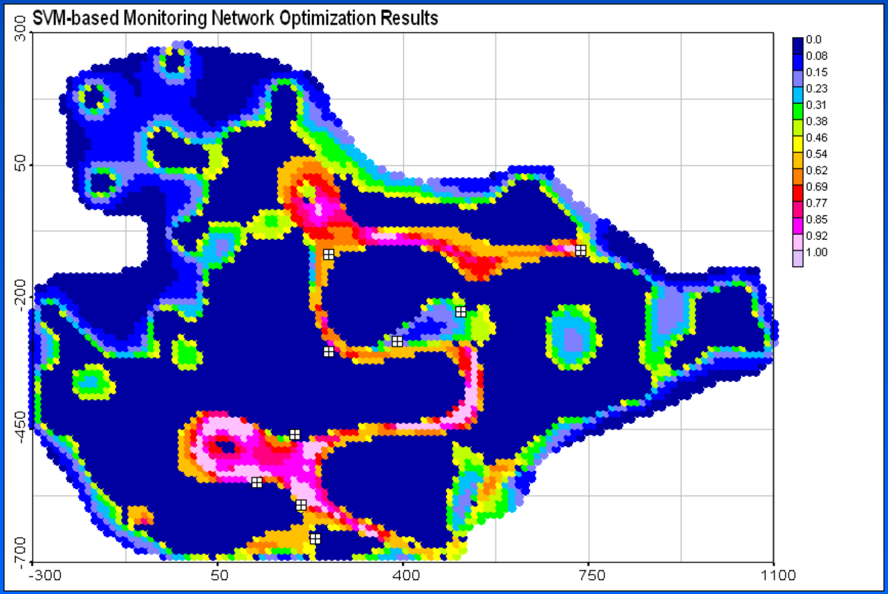
- only SVs contribute to maximum margin solution.
- SVs are the closest samples to the decision boundary

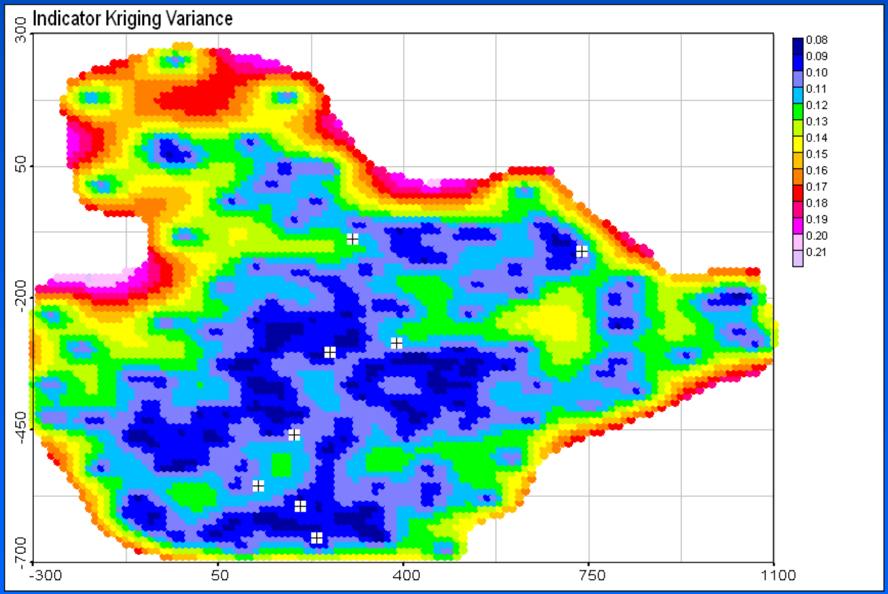
Active Learning with Support Vectors

- Sample decision boundary to label most uncertain locations
- Minimize cross-validation error for misclassifications

SV-based importance measure
- task-oriented result
- follows classification model

Kriging variance
- improves network topology only





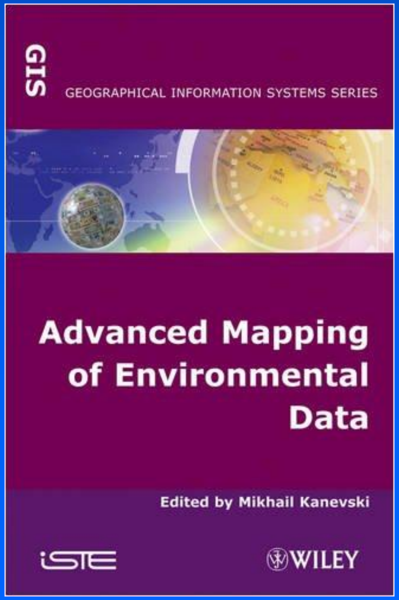
A. Pozdnoukhov, M. Kanevski, Monitoring Network Optimisation for Spatial Data Classification Using Support Vector Machines. International Journal of Environment and Pollution, 2005

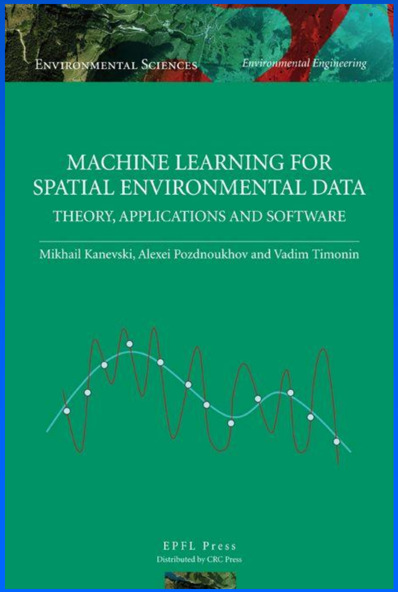
Summary

- Approaches for spatial uncertainty modelling
  - Geostatistics
  - Machine Learning
  - Combination of both
- Need for stochastic models for adequate uncertainty description
- Bayesian approach handles uncertainty of the model definitions and data uncertainty
- Uncertainty modelling for sampling optimisation

Acknowledgement

- M. Christie, Heriot-Watt University
- M. Kanevski, V. Timonin, University of Lausanne
- A. Pozdnoukhov, University of Ireland Mynooth
- G. Christakos, San-Diego State University
- E. Savelieva, Nuclear Safety Institute, Moscow





References

1. Advanced Mapping of Environmental Data: Geostatistics, Machine Learning and Bayesian Maximum Entropy (2008) ed. M. Kanevski, ISTE & Wiley, 328 p.
2. V. Demyanov, A. Pozdnoukhov, M. Christie, M. Kanevski (2009) *Uncertainty Quantification of a Semi-supervised Support Vector Regression Reservoir Model*, International Association for on Mathematical Geosciences (IAMG) conference, Stanford.
3. Demyanov V., Pozdnoukhov A., Kanevski M., Christie M. (2008) *Geomodelling of a Fluvial System with Semi-Supervised Support Vector Regression*, Proceedings of the VII International Geostatistics Congress, pp. 627-636, GECAMIN, Chile.
4. V.Demyanov, M. Kanevski, M. Maignan, E. Savelieva, V. Timonin, S. Chernov, G. Piller (2000*) Indoor Radon Risk Assessment with Geostatistics and Artificial Neural Networks*, International Geostatistics Congress, Cape Town.
5. M.Kanevski, A.Pozdnoukhov, V.Timonin (2009) *Machine Learning for Environmental Data*, PPUR, 368 p.
6. M. Kanevski , L. Bolshov, E. Savelieva, A. Pozdnukhov, V. Timonin, S. Chernov (2001) *Characterization of Hydrogeologic Systems with Machine Learning Algorithms and Geostatistical Models*. International Containment & Remediation Technology Conference, Orlando, USA.
7. M. Kanevski, R. Arutyunyan, L. Bolshov, V. Demyanov, S. Chernov, E.Savelieva, V. Timonin, M. Maignan, M.F. Maignan (1999*) Mapping of Radioactively Contaminated Territories with Geostatistics and Artificial Neural Networks*. In Contaminated Forests I. Linkov and W.R. Schell (eds.), pp. 249-256, Kluwer Academic Publishers.
8. A. Pozdnoukhov, M. Kanevski (2005) *Monitoring Network Optimisation for Spatial Data Classification Using Support Vector Machines*. International Journal of Environment and Pollution.
9. E. Savelieva, V. Demyanov, M. Kanevski, M. Serre, G. Christakos.(2005) *BME-Based Uncertainty Assessment of the Chernobyl Fallouts.* Geoderma , vol. 128, Issues 3-4, pp. 312-324.