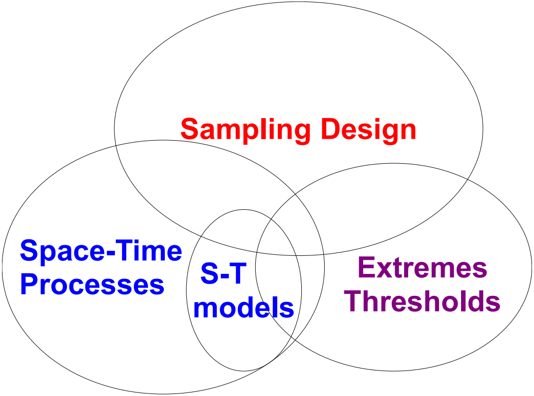
Sampling Design in Space-Time for Threshold Exceedances

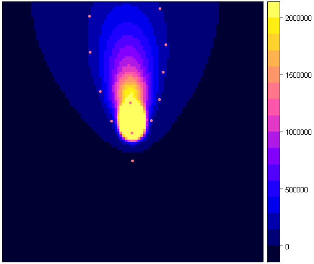
Kristina Helle

12. March 2009

overview


Sampling Design
Space-Time Processes
S-T models
Extremes Thresholds

Usecase: nuclear power plant accident



Where and when to sample?

process

- spread of gas / dust in atmosphere
- no / known chemical / physical reactions
- one (known) point source

aim

- predict areas where threshold is exceeded

Models and Related Sampling

sampling design depends completely on intended use of measurements

$$z(x,t) = y(x,t;\beta) + e(x,t;\alpha)$$

- deterministic trend y (e.g. by differential equations)
  - parameter fitting β
    - samples at **locations** with high information about parameters (uses derivation)
- probabilistic error e (unknown variogram)
  - fitting of covariance parameters α
    - samples at different distances - especially small ones
- probabilistic prediction (with known variogram and trend)
  - interpolation z
    - samples covering space

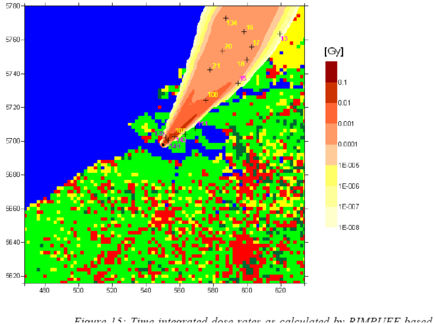
What are good models and designs?

Optimization for interpolation is often performed by minimization of a **global** (cost) function. Model-based sampling desing often uses **local** criteria weighting the importance of each point for sampling.

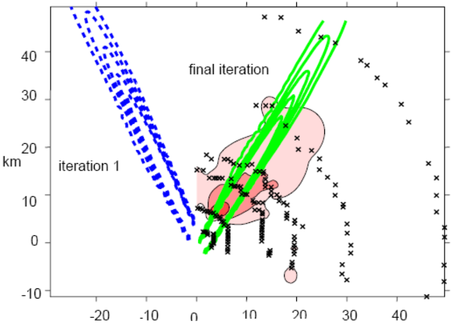
- Q Which model fits the usecase (extreme values)?
- Q How much better are results of complex methods if optimal samples are used?
- Q Which prior knowledge is needed to fit more complex models?
- Q How robust are the models to wrong assumptions?
- Q Can measurements from a sampling design be used for other purposes?

'Data' for Model Comparison

- generated according to the simle models used
- generated by use of different complex models (RIMPUFF...)
- from remote sensing (smoke plumes)(?)
- from experiments - no spatial covering



Palma (2005), p. 29

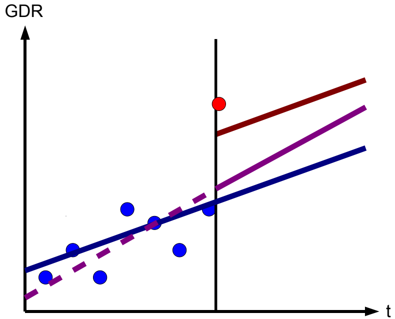


Twenhöfel et al. (2007), p. 28

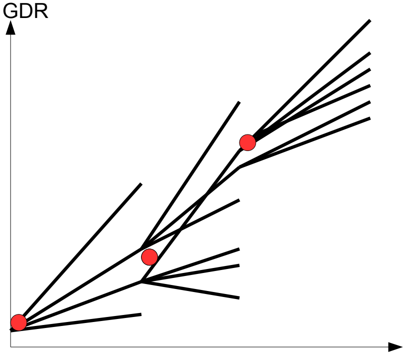
How to use the measured values?

Samples can be taken at many points in space and time. How to integrate the measured values in space-time models?

Q How to include data from different time steps?

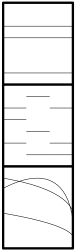


Q How to compute several scenarios?

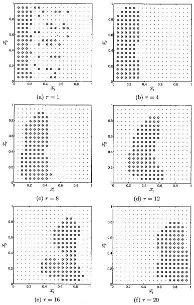


Where and when can we measure?

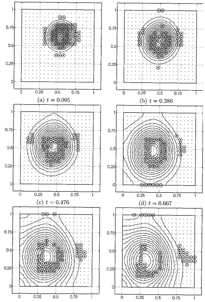
- Q What are reasonable designs?
- Q How to search the infinite dimensional space of designs?



stationary
scanning
moving



Uciński 2005, p.119
Fedorov: determine sensor density



Uciński 2005, p.224
keep sensors stationary for time intervals

Additional constraints to movements

no clustering

- Hide redundant information by noise (Müller).
- Bound sensor density (Fedorov).

tabu areas

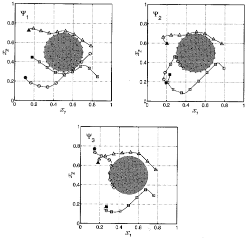
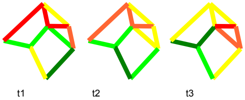


FIGURE 5.1
Optimal sensor trajectories for different criteria. The lightly shaded circle constitutes the forbidden region $\mathcal{X}$ where spatial prediction is required.

Uciński 2005, p.158

street networks



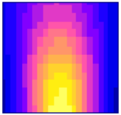
Q How to choose a / several ways?

Q How to select speed / times to change direction?

measurement → target value

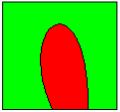
The knowledge we really aim at is not necessary identical with the values we measure.

- Spatio-temporal resolution of the aim
  - global / local on regular grid /local on administrative units /pointwise
  - temporal average / average on intervals / continuous
- The target variable may depend (non)linear on the measurements.
  - Q Error propagation from measured value to target value.
  - Q If there are several possibilities to use measurements for prediction, how to balance sampling design.
  - Q If covariates used, how to determine combined sampling designs?



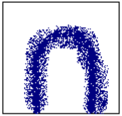
Threshold exceedances can be addressed by different kinds of models and related criteria:
global criteria

- area classified correctly
- risk that very high values remain undetected



local criteria

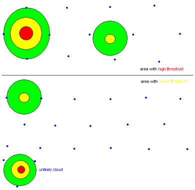
- uncertainty of classification
- cost of misclassification depending on amount of exceedance



Q Which global / local criteria are useful for threshold exceedance?

Q Which optimization algorithm is best for which of those criteria?

Is there a typical shape of plumes?



- Q Which density of a network is needed to detect plumes?
- Q How to build "fences" of sensors to protect areas?
- Q Are there simple geometric aproximations to plume shapes for fast detection of boundaries?

Basic Literature on Sampling Design

de Gruijter, J., Brus, D.J., Bierkens, M.F.P. & Knotters, M. (2006). Sampling for Natural Resource Monitoring. Springer, Berlin.
(mainly probability-based sampling design, spatio-temporal correlation patterns)

Müller, W. (2007). Collecting Spatial Data. Optimum Design of Experiments of Random Fields. 3rd edition, Springer, Berlin.
(s.d. for geostatistics (model-based): parameter fitting, spatial correlation)

Uciński, D. (2005). Optimal Measurement Methods for Distributed Parameter Identification. CRC Press, Boca Raton.
(s.d. for PDE, spatio-temporal)

Further References

Palma, C. (2005): Data Assimilation for Off site Nuclear Emergency Management. DAONEM FIKR-CT-2000-00025

Twenhöfel, Ch.; van Troost, M. & Bader, S. (2007): Uncertainty analysis and parameter optimisation in early phase nuclear emergency management – A case study using the NPK-PUFF dispersion model. RIVM Report 861004001/2007

Thank you for your attention!

Questions?

Discuss!