Sensitivity Analysis of Final Repository Models Using Quasi-Random Sampling and a Metamodel Approach

Sabine M. Spiessl and Dirk-A. Becker

Gesellschaft fuer Anlagen- und Reaktorsicherheit (GRS),
Braunschweig, Germany

SAMO 2013

Outline of the Talk

- Final Disposal of Radioactive Waste in a Deep Geological Repository
  - Basic Concept
  - Computational Models for Long-Term Performance Assessment (PA)
  - Test Case
- Approach for Finding Guidelines for Doing Sensitivity Analysis for Complex Models
- Methods
  - Sampling
    - Quasi-Random Sampling / Low Discrepancy Sequences (LDS)
  - Sensitivity Analysis
    - Metamodel (SDP)
    - Simple first-order SI calculation scheme (EASI)
    - Regression-based (SRRC)
    - Graphical (Scatterplot and CSM plot)
- Key Findings
- Summary

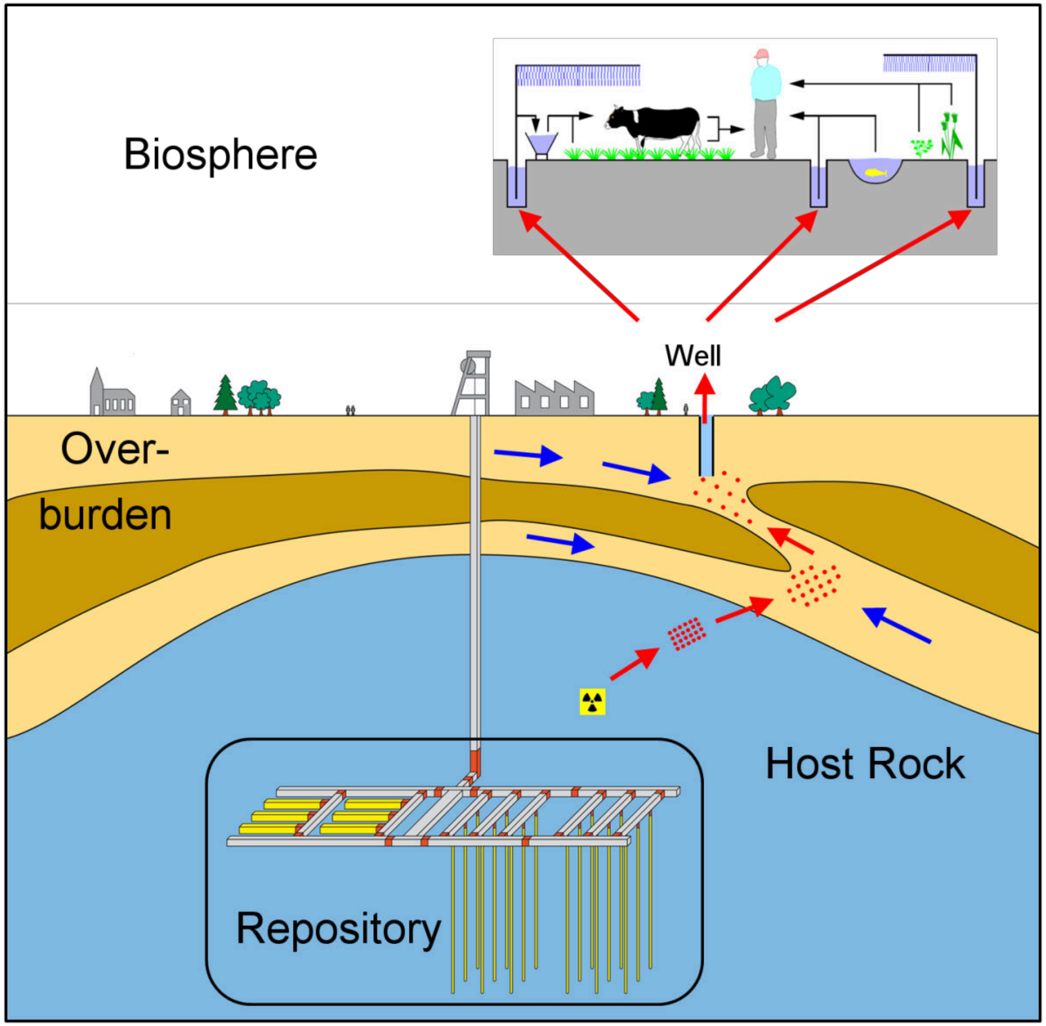
Final Disposal of Radioactive Waste in a Deep Geological Repository

The repository system consists of three parts

- Near field:
  the engineered barrier system and the host rock formation
- Far field:
  the geological layers above the host rock
- Biosphere:
  the biological environment of man

Long-term safety has to be proven

- Release of contaminants to the biosphere
- Effects to human health
- Time frames up to 1 million years
- Use of computer models necessary
- Deterministic and probabilistic approaches



Computational Models for Long-Term Performance Assessment (PA)

Near Field

- Intrusion of water or brine
- Radionuclide mobilization from containers
- Dissolution and precipitation
- Sorption
- Chemical processes
- Mechanical processes
- Gas production and pressure build-up
- Fluid flow
- Transport of radionuclides
- Release to the far field

Far Field

- Dilution by groundwater
- Fluid flow through geological layers
- Precipitation
- Sorption
- Release to the biosphere

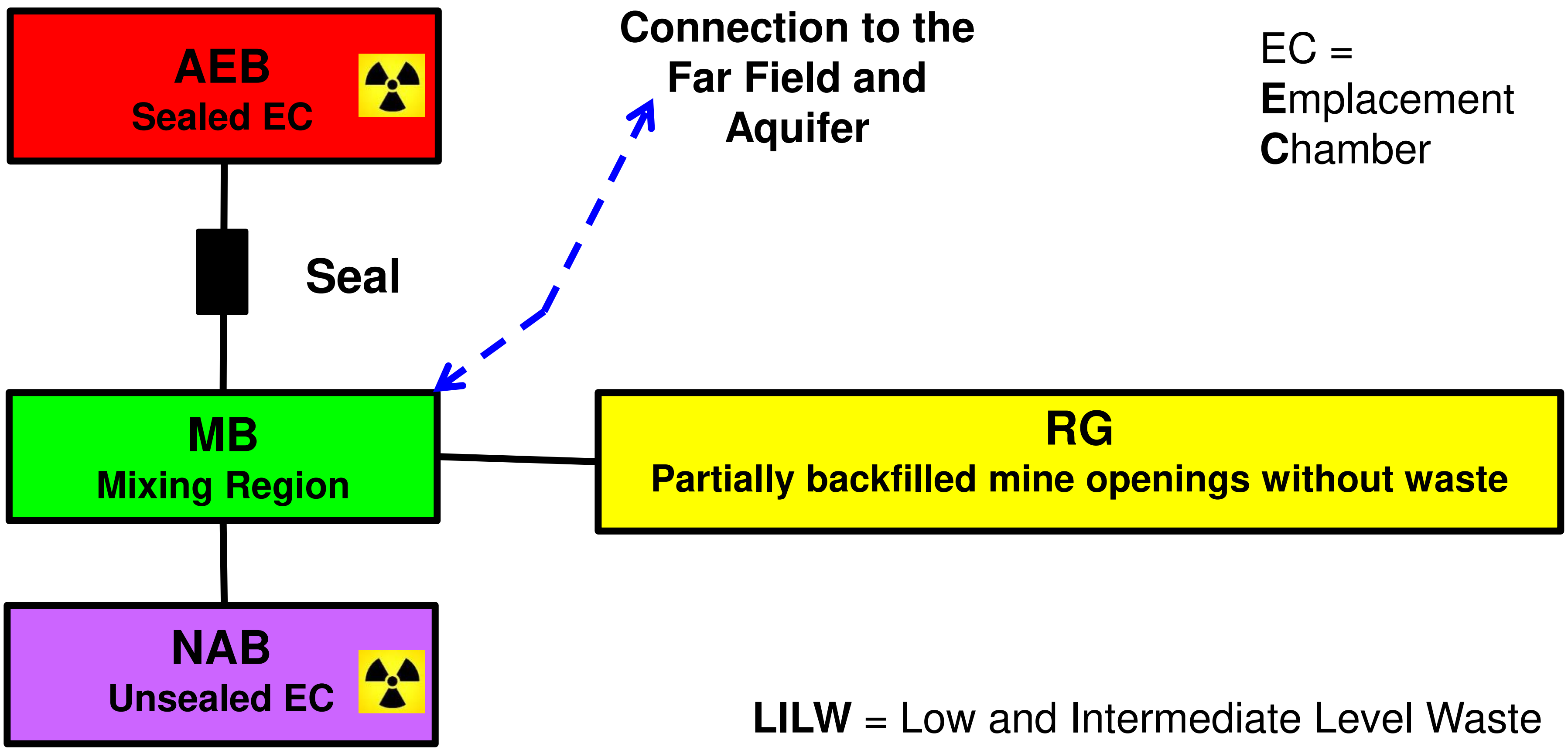
Biosphere

- Concentration of radionuclides in nutrition products
  - Contamination of drinking water
  - Uptake of radionuclides by fish
  - Irrigation of fields
  - Cattle watering
- Direct radiation
- Individual equivalent dose per year

----- Radioactive Decay -----

PA Test Case

<u>Real situation:</u> LILW repository in an abandoned salt mine
<u>Near field model:</u> substantially simplified structure



EC = **E**mplacement **C**hamber

LILW = Low and Intermediate Level Waste

Investigated Model Parameters

	Parameter	Description
1	TBrine	Intrusion time of the brine
2	IniPermSeal	Initial permeability of dissolving seal
3	RefConv	Reference convergence rate
4	GasEntryP	Gas entry pressure
5	BrineMgSat	Volumetric dissolution capacity of the brine
6-9	Factor of local convergence variation	AEBConv – Sealed EC, NABConv - Unsealed EC, MBConv - mixing region, RGConv - RG
10-11	Gas corrosion rate	GasCorrPE - organics, GasCorrFe - metal
12-14	Proportion of the material involved in gas production	AEBGasProd - sealed EC, Unsealed EC - NABGasProd, RGGasProd - RG
15-16	Porosity	RefPor - reference, PorDebris - backfill
17	DiffCoeff	Diffusion coefficient
18	C14Inv	C14 inventory
19	FacDisp	Dispersion length
20	KonvFak	Sheeting factor

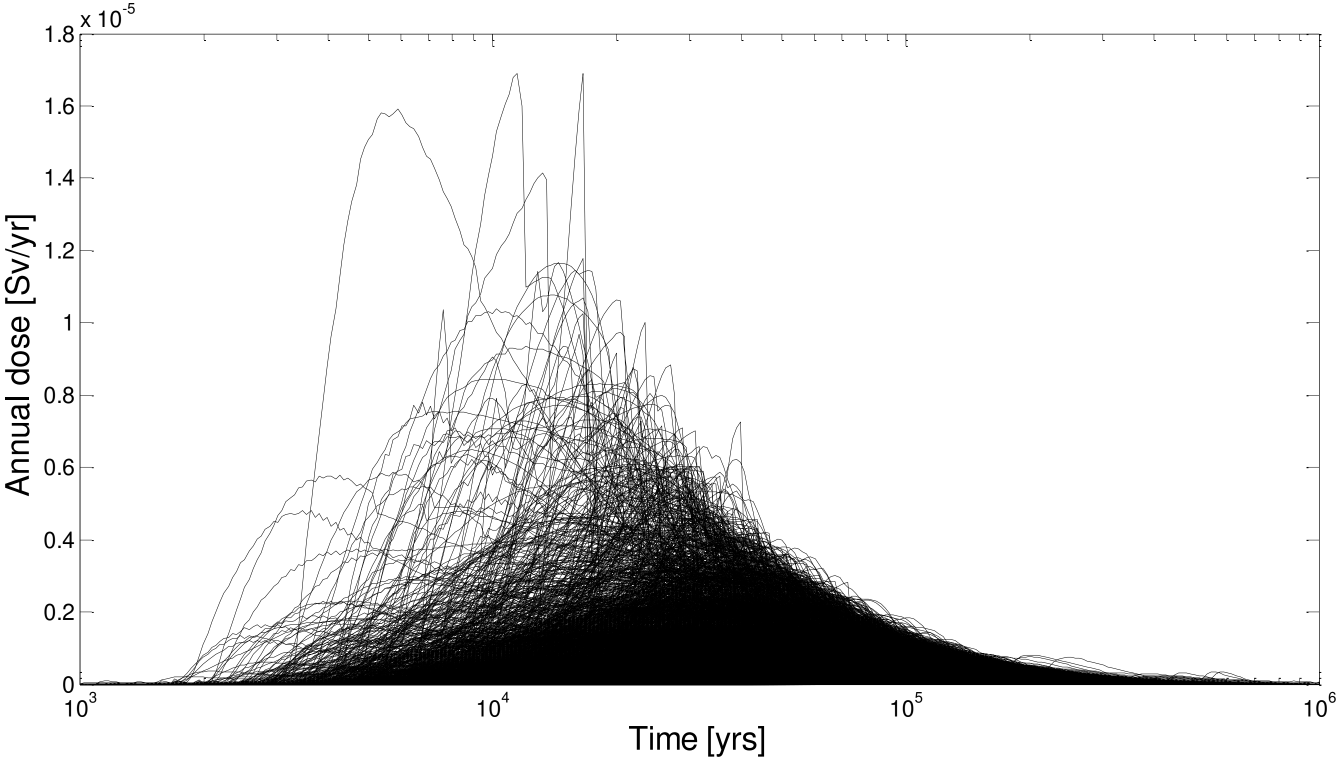
Investigated sets:
6 Parameters
7 Parameters
11 Parameters
20 Parameters

Model Output

Illustration of the annual dose to the biosphere versus time

Skewed distribution of the model output

Few simulations with high model output



Sample with 16384 runs, 11 parameters and LpTau sampling

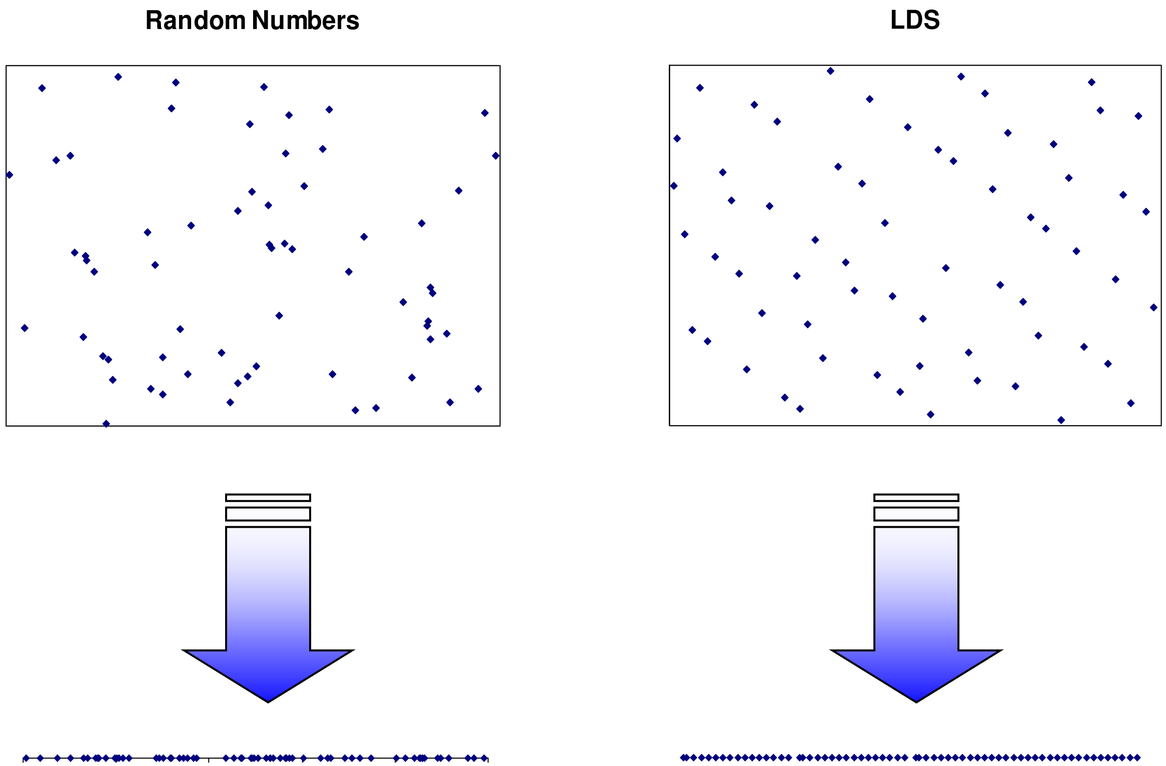
Approach for Finding Guidelines for Doing Sensitivity Analysis for Complex Models

- Condition
  - Time three years
  - Three test cases to investigate with different challenges (catch characteristic of real system)

- Approach
  - Try to quickly understand basic concept of various methods with different approaches
  - Select methods which look promising
  - Get software
  - Analyse a couple of samples with selected methods
  - Address following questions:
    - Do the results help understanding the system?
    - Do the methods identify important parameters?
    - Do the methods help reducing computational effort?
  - If they are differences in results with the selected methods, try to understand why?
  - Try to understand theory of good working methods in more detail
  - Give recommendation for methods which provide good results

Sampling: Quasi-Random / Low Discrepancy Sequences (LDS)

Discrepancy is a measure of deviation from uniformity



Projections of random and LDS-2D sequences to 1D

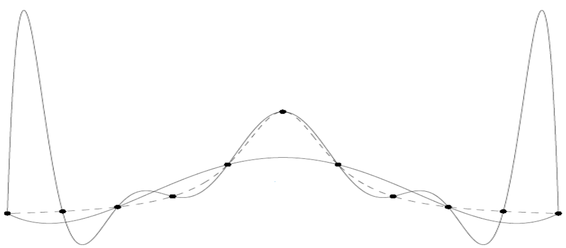
Sergei Kucherenko, 2012

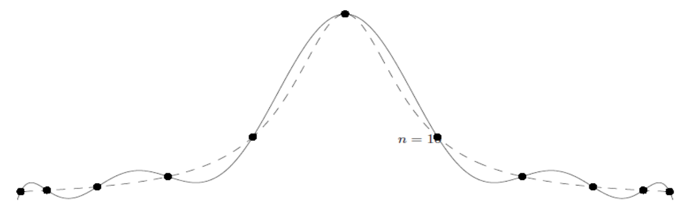
Meta-Modelling

Approximation of the computational model
Use of the approximations to calculate the sensitivity measures

Different methods can be used for the approximation of the computational model

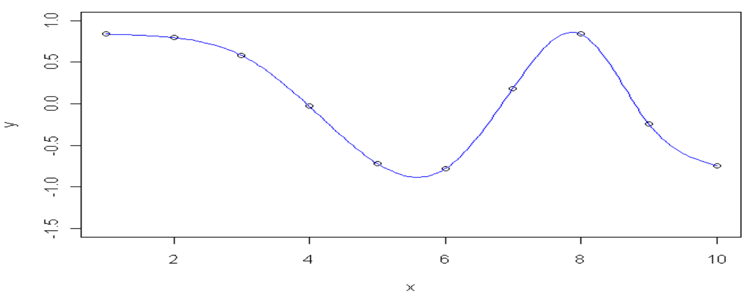
- Linear regression
- Polynomials





- Splines

- Additive models, GAM
- Regression trees
- Neural networks
- Chaos polynomials
- Support Vector Machines
- Kriging – Gaussian process
- State Dependent Parameter (SDP)



Bertrand Iooss, 2012

SDP Meta-Modelling – Insight

High Dimensional Model Representation (HDMR) = decomposition of the computational model *f* into terms of increasing dimensionality $f_i(X_{i,t})$

$$Y_t - f_0 = f_1(X_{1,t}) + f_2(X_{2,t}) + \cdots + f_k(X_{k,t}) + o(XX')$$

1st order

Y_t	Model output	$f_0 = E(Y)$	Expectation of ***Y***
$X_{i,t}$	Input parameter	$o(XX')$	Error term

State Dependent Regression (SDR) model

$$Y_t - f_0 = \boldsymbol{X}_t^T \boldsymbol{p}_t + e_t$$
$$= p_{1,t}X_{1,t} + p_{2,t}X_{2,t} + \cdots + p_{k,t}X_{k,t} + e_t$$

Charactisation of the variability or evolution of the SDP's ($p_{i,t}$)

$p_{i,t}$	SDP	$e_t = N(0, \sigma^2)$	
e_t	Gaussian white noise process with zero mean and variance σ^2	t	Index for sequence of MC runs; 1 ,…, *N*

Comparison of Results

EASI - frequency-based approach (Plischke 2010)

- Usage of existing data possible
- Introduction of frequency to input data by sorting and shuffling the values of the different input parameters
- Arrangement of output according to the input data
- Analyse of arranged data using the power spectrum of the output like it is done in FAST/EFAST

Standardised Rank Regression Coefficients (SRRC)

- Based on the calc. influences of the individual input parameters to a linear regression of the output values

CSM (Contribution to the Sample Mean) plots

- Order the sample according to one parameter
- Calculate relative contribution to sample mean (CSM) for each value
- Plot CSM versus cumulative relative frequency
- Relevant parameters show significant deviance from diagonal

Scatterplot

- Plot output vs. one input parameter

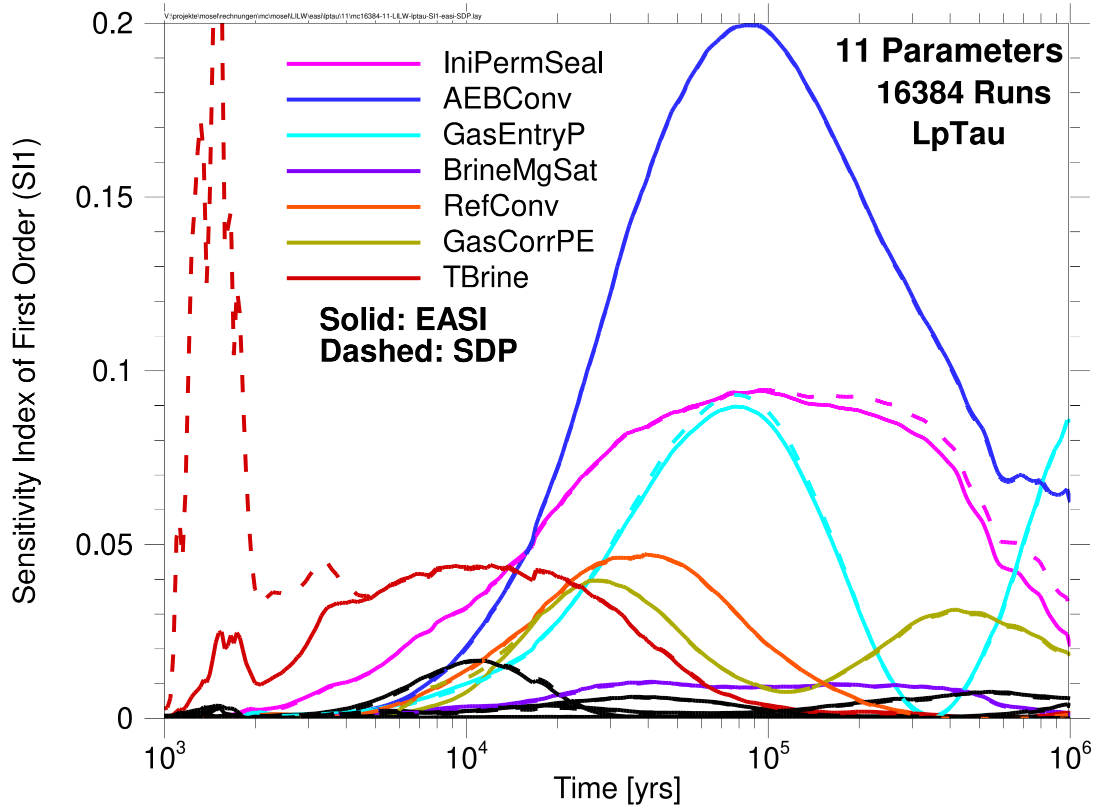
SDP versus EASI

Unexpected, close agreement between SDP and EASI for unknown reason!

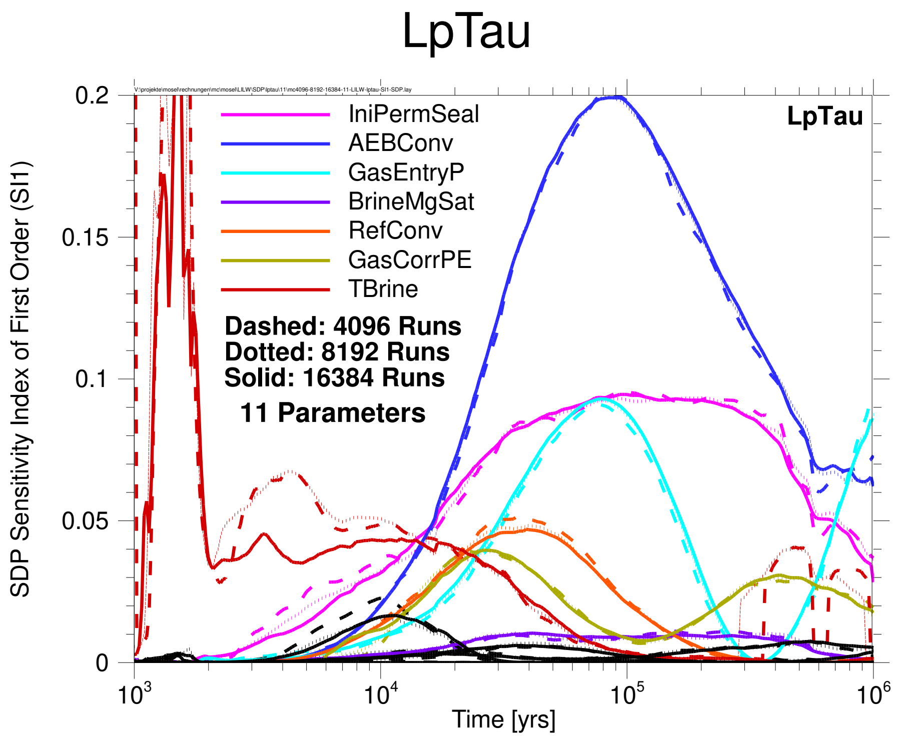
EASI is very fast (a couple of minutes)

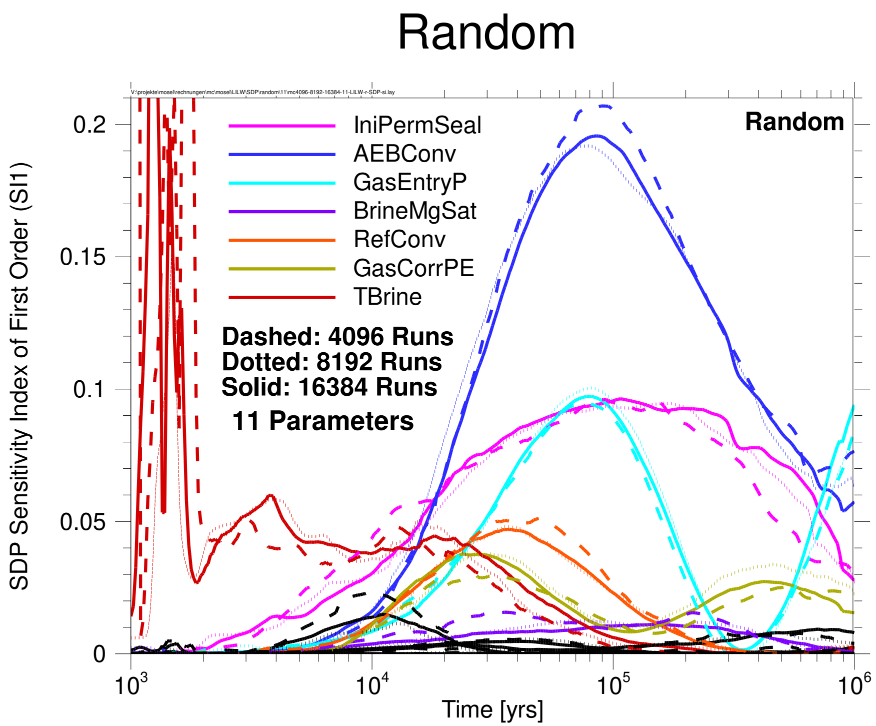
For samples with high number of simulations and number of parameters, SDP needs a long time for the estimation of the SI1 index (order of days)

SDP has more difficulties of handling the small output of the few runs at the beginning of the simulation



Results – SDP Meta-Modelling – LpTau versus Random - 11 Parameters



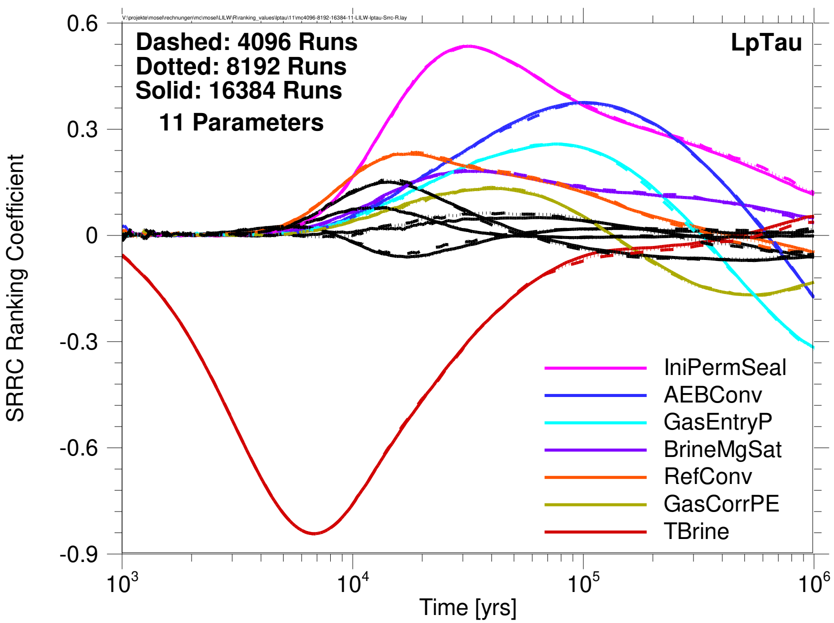


Parameter ranking using random sampling agrees with the one using LpTau sampling

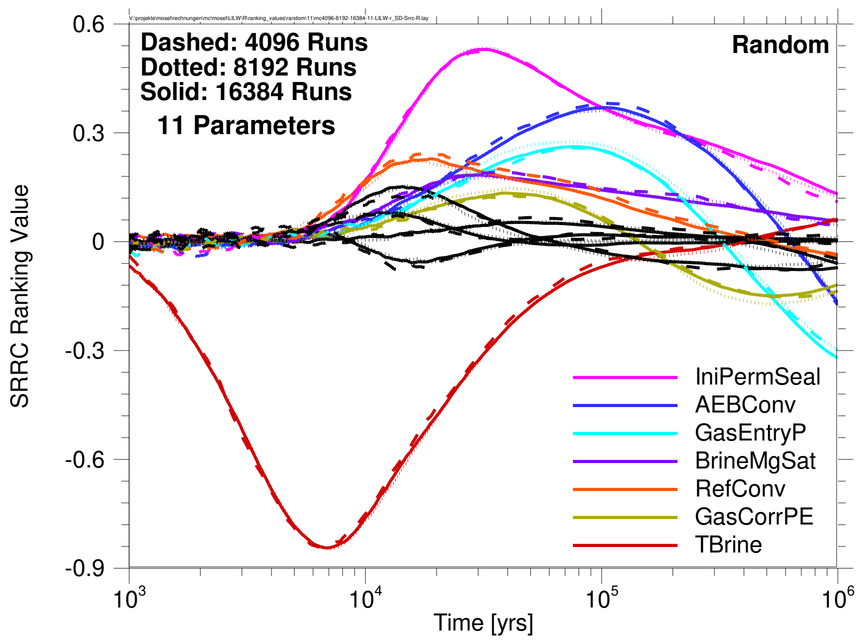
For random sampling, the deviations between the different sets with same number of runs are bigger compared to the ones using LpTau sampling

Results – SRRC – LpTau versus Random – 11 Parameters

LpTau

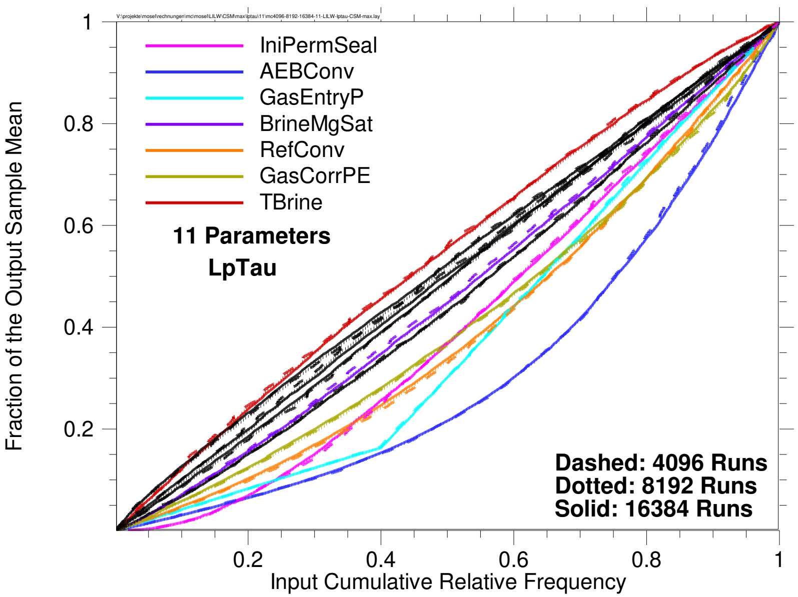


Random

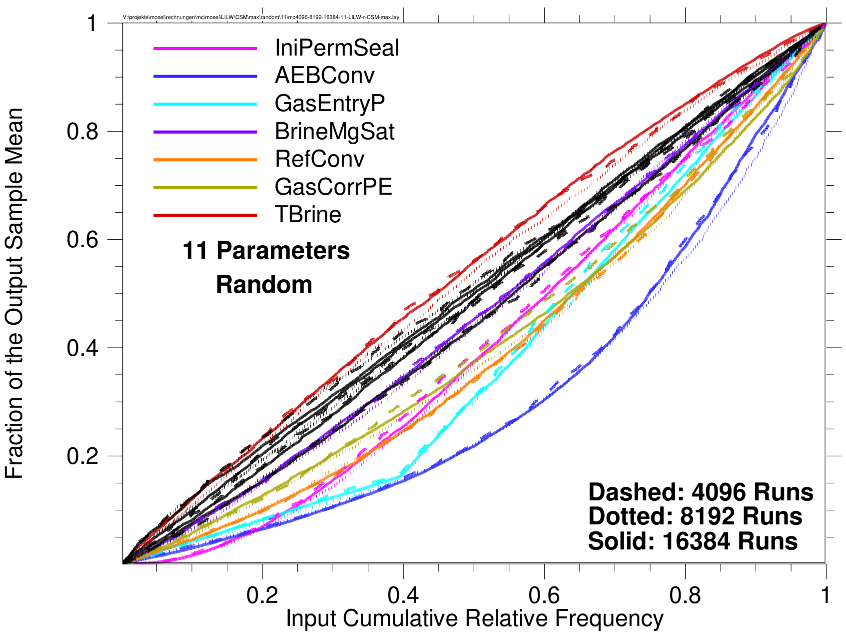


Results – CSM Plot – LpTau vs Random – 11 Parameters – Peak Dose

LpTau



Random



Parameter Ranking

Parameter	SDP/EASI	SRRC	CSM
AEBConv	1	3	1
IniPermSeal	2	2	3
GasEntryP	3	4	2
TBrine	4	1	4
GasCorrPE	5	4	3
RefConv	5	4	3
BrineMgSat	6	5	5

Major differences !!!

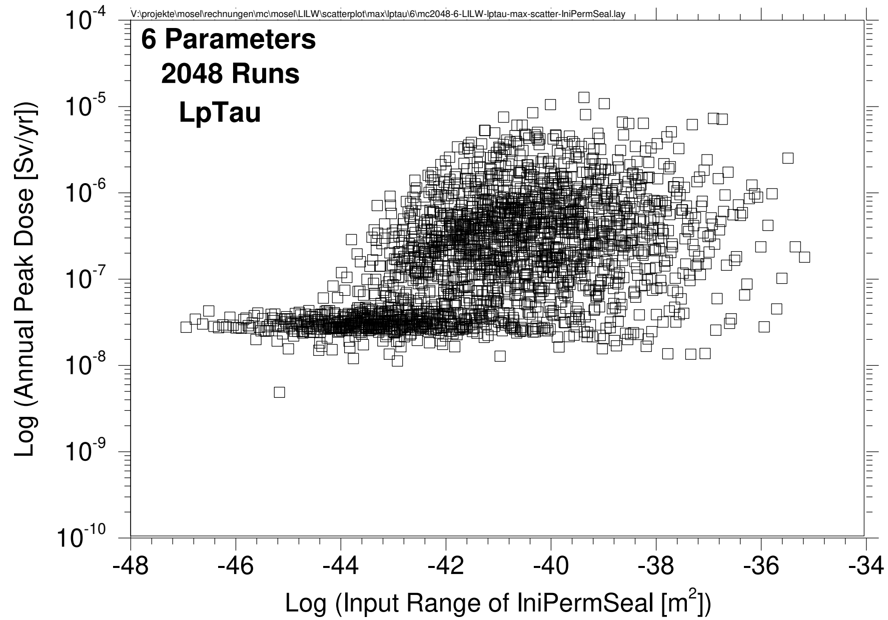
Thus, a unique list of parameter ranking cannot be provided

Results – Scatterplots – Peak Dose

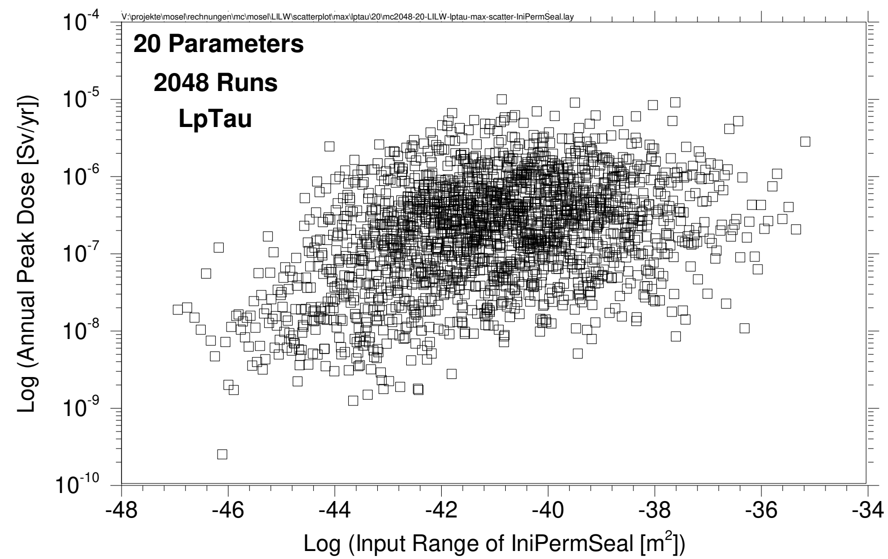
Scatterplot confirms important parameters found by the SDP, EASI and SRRC methods as well as the CSM plot

Increased pattern in samples with smaller number of parameters

6 Parameters



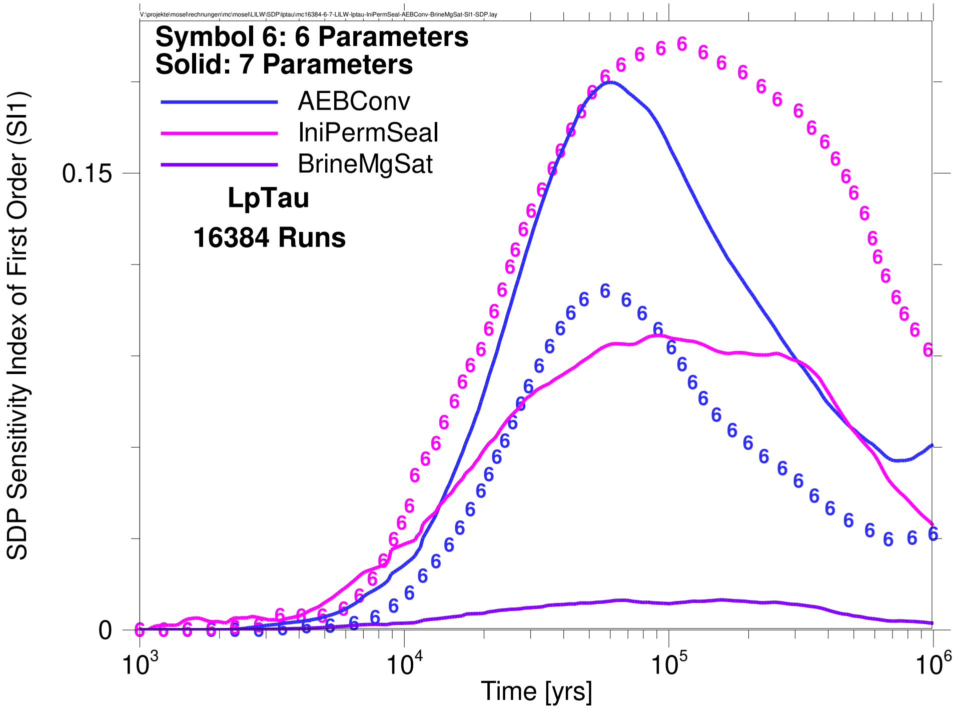
20 Parameters



Parameter: IniPermSeal (initial permeability of dissolving seal)

2048 Simulations – LpTau

Undetected Significant Parameter



- Parameter **BrineMgSat** (**volumetric dissolution capacity of the brine)** not identified as important by any method
- If unconsidered, other parameter curves show significantly different shapes
- Apparently, the parameter has an indirect influence

Summary

- A metamodel (SDP) approach in combination with different sampling methods (random and LpTau) was investigated
- SDP results were compared to a simple first-order SI calculation scheme (EASI), a regression-based method (SRRC) and two graphical methods (Scatterplot and CSM plot)
- Major differences in parameter ranking obtained from the CSM plot, SRRC and SDP methods
- Therefore, a unique list of parameter ranking cannot be provided
- However, all methods identified same parameters as important
- For random sampling, the deviations between the different sets with same number of runs are bigger compared to the ones using LpTau sampling
- EASI is quicker than SDP
- None of the investigated methods could find all significant parameters
    - One parameter identified as nearly non-important has significant impact upon the two important parameters, e.g. ranking is turned around

THANK you for your attention

Study is sponsored by the German Federal Ministry of Economics and Technology (BMWi)