Uncertainty analysis tool as part of safety assessment framework: model-independent or model-tailored?

Valentina Svitelman, Elena Saveleva, Petr Blinov, Dmitrii Valetov
Nuclear Safety Institute of the Russian Academy of Sciences (Moscow)
e-mail: svitelman@ibrae.ac.ru

May 2020
EGU2020: Sharing Geoscience Online
Session HS8.1.6 Handling Uncertainties in Model Concepts, Parameters, Forcings and Forecasts: Diagnostics, Sensitivity, Inversion and Uncertainty Analysis

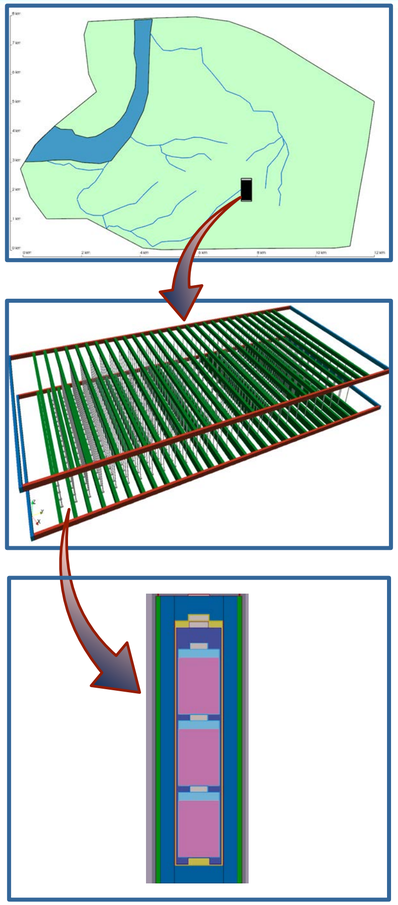
Safety assessment – complex analysis

The focus of our study: to provide uncertainty assessment, sensitivity analysis and calibration tools for the safety assessment framework .

<u>Safety assessment for a radioactive waste disposal facility: consideration of the performance of natural and engineered barriers over long times and assessment of exposures to the environment far in the future.</u>

A lot of activities undertaken in parallel and iteratively:

- characterization of the proposed disposal system (including waste forms, geological setting, repository design);
- identification of relevant factors and arranging them into scenarios and corresponding calculation cases;
- development of the computational models;
- synthesis of the system-level analysis including the assessment of consequences, consistency with regulatory needs and stakeholders interests.

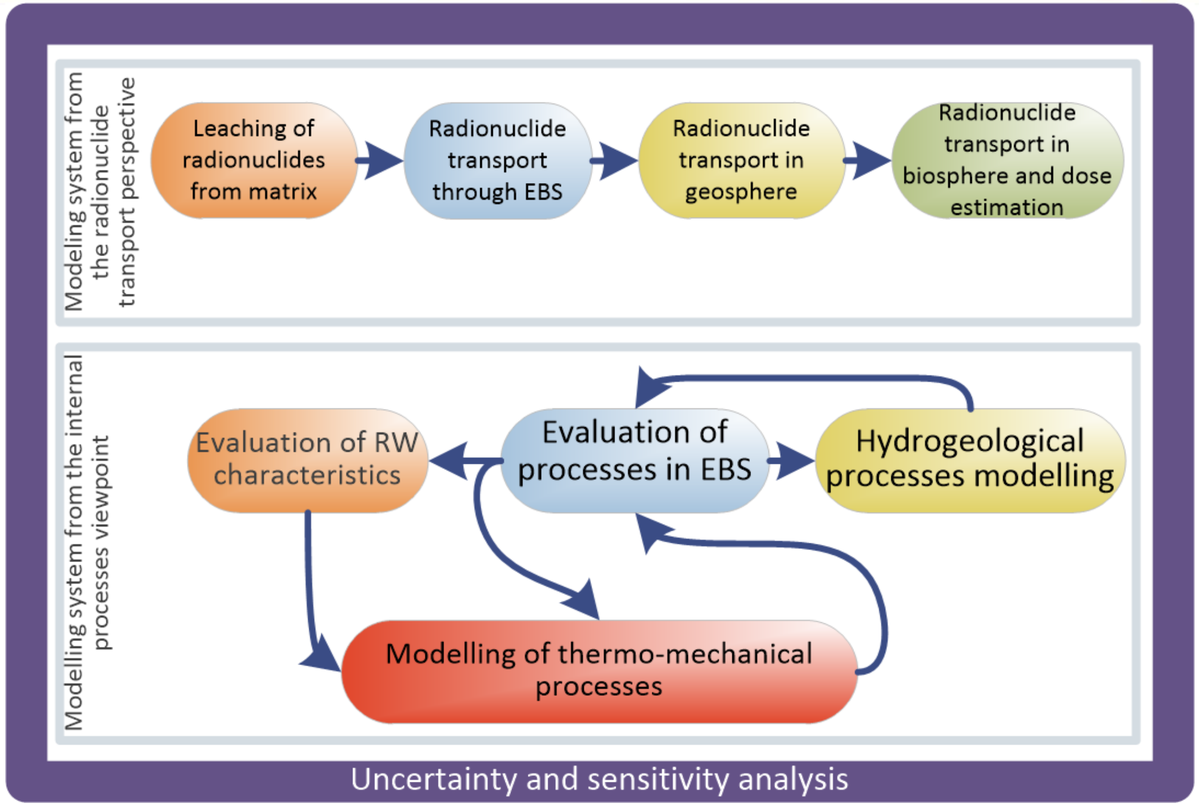


Different scales of models in safety assessment

[1] Dorofeev AN, Bolshov LA, Linge II, Utkin SS, Saveleva EA. Strategic master-plan for research demonstrating the safety of construction, operation and closure of a deep geological repository for radioactive waste. Radioactive Waste 2017;1:19–26.

System of numerical models for safety assessment

- Numerous models impemented in various software:
  - 3D flow and transport – **GeRa**, **Modflow+MT3DMS**
  - Near field processes – **DESTRUCT**, **AMBER, PhreeqC**
  - 3D thermomechanics – **FENIA**
  - RW properties – **CORIDA**
- Uncertainty analysis is needed for each of them and for their combinations → our group is developing **MoUSE** software package.



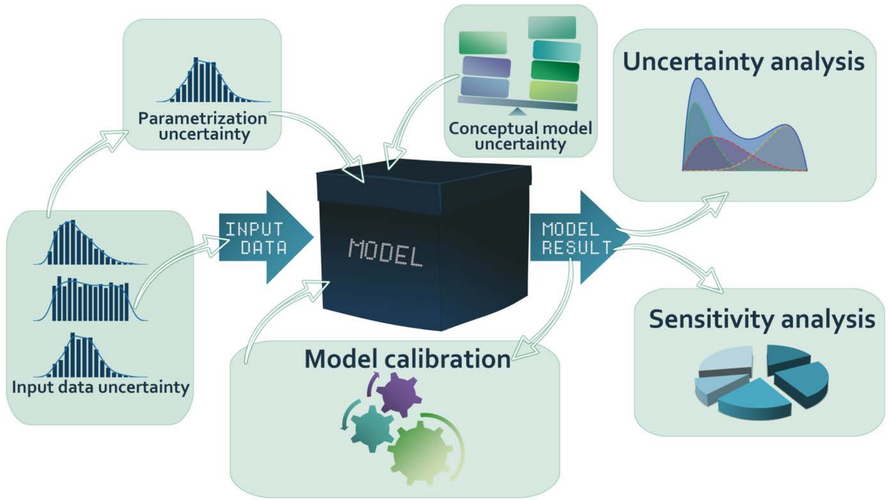
* Both **in-house developed software** and **external software** is used for model development

[2] Valetov D, Neuvazhaev G, Svitelman V, Saveleva E. Hybrid Cuckoo Search and Harmony Search Algorithm and Its Modifications for the Calibration of Groundwater Flow Models: Proceedings of the 11th International Joint Conference on Computational Intelligence, Vienna, Austria: SCITEPRESS – Science and Technology Publications; 2019, p. 221–228. doi: 10.5220/0008345502210228.
[3] Romanchuk A, Larina A, Semenkova A, Svitelman V, Blinov P, Kalmykov S. Sorption of radionuclides onto minerals surfaces: new approach to the modelling // 17th International Conference on Chemistry and Migration Behaviour of Actinides and Fission Products in the Geosphere. Kyoto, Japan: 2019.
[4] Svitelman V, Saveleva E, Gorelov M, Moiseenko E, Drobyshevsky N. The numerical model of the planned URF thermo-mechanical experiment: sensitivity analysis. DECOVALEX 2019 Symposium Abstracts, DECOVALEX 2019.

Model independent uncertainty analysis software – how it works

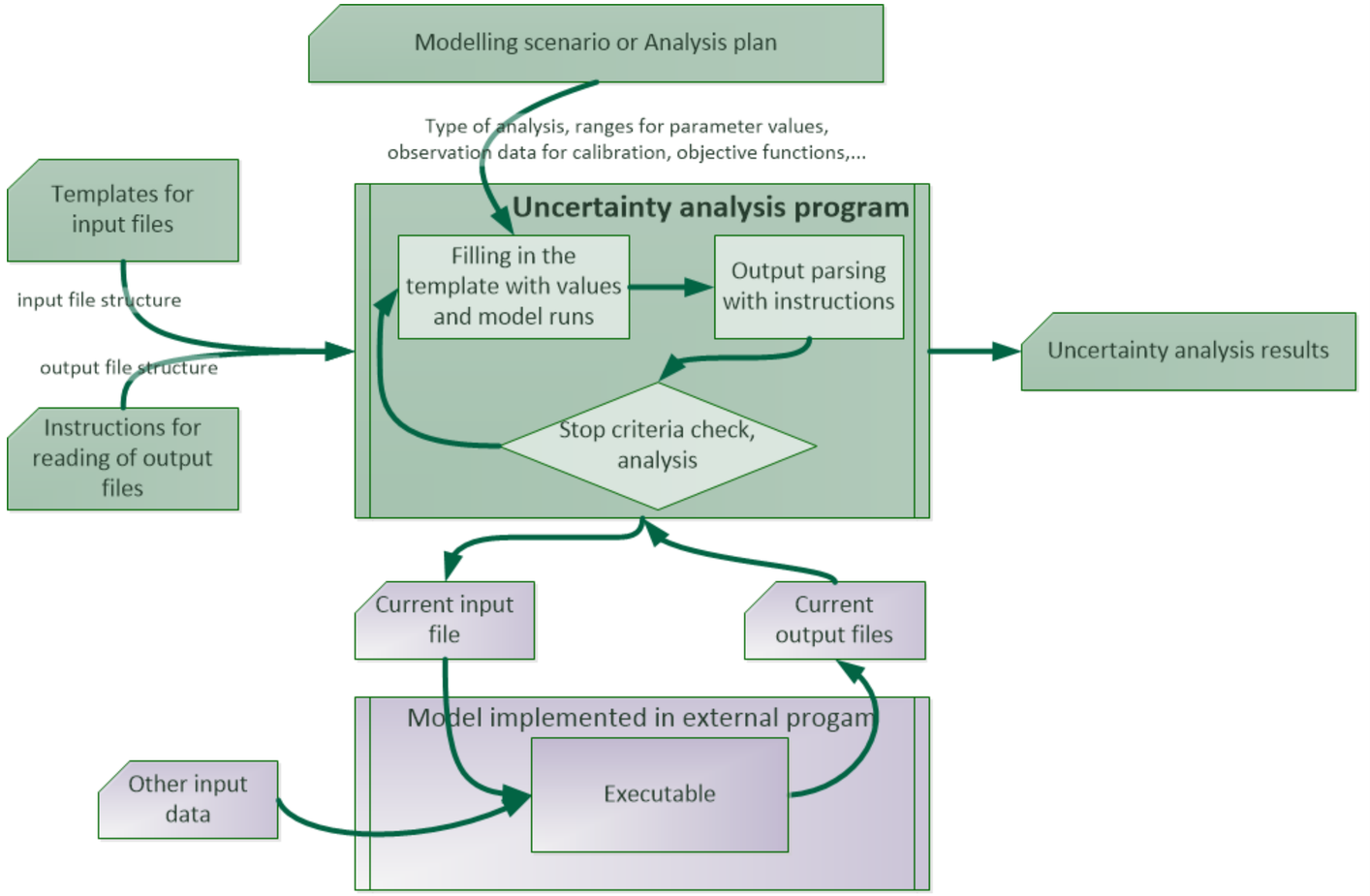
We follow generally accepted practice:

- Black-box executables – could run any program with command line interface
- Templates for input files: similar to the native input files except “gaps” for parameters that are varied for analysis
- Instructions for output files: rules for extraction of specific segments of output data



Analysis options

Interaction with model



Real world differences: inputs

Different input formats examples

PhreeqC

DATABASE	*database for the simulations*
SOLUTION .... END	*composition of an aqueous solution*
REACTION ... END	*irreversible reactions*
EXCHANGE ... END	*exchange assemblage composition*
... ...	*Various keyword-specified options: EQUILIBRIUM_PHASES for a combination of minerals and/or gases which react reversibly to a prescribed equilibrium; KINETICS for chemicals which react depending on time and composition of the solution; REACTION_TEMPERATURE for changing the temperature*
USER PUNCH ... END	*Print user-defined quantities to the selected-output file*

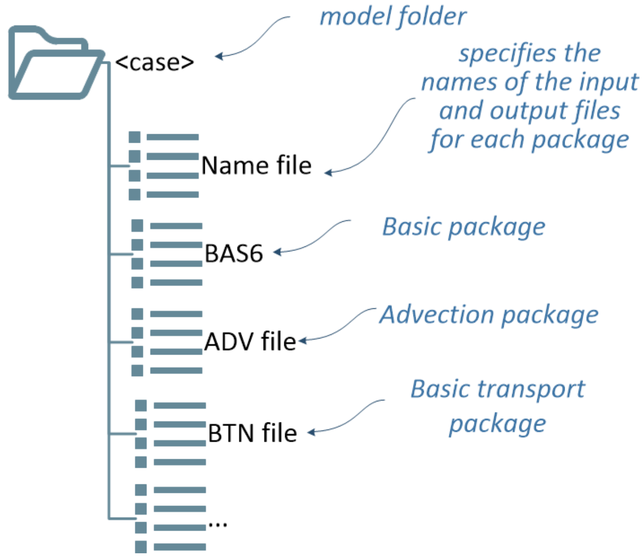
Single files with keywords

Amber

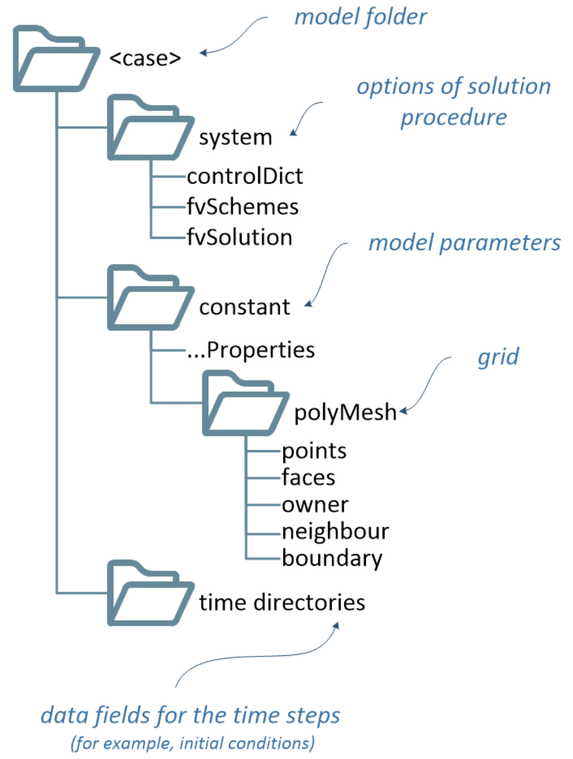
"case name" "case description" AMBER_VERSION: version № GENERATED_BY: who DATE: text entry for date TIME: text entry for time CONTAMINANT name	*Case Information*
atomic_mass "description" DECAY parent daughter rate	*Contaminants*
name "description" COMPARTMENT name "description"	*Decays*
...	*Source Terms, Transfers, Submodels, Calculation Options, User Units, Export Files, etc.*

Folders with multiple specific files

Modflow & MT3DMS



OpenFOAM



...and a lot of less known program-specific formats

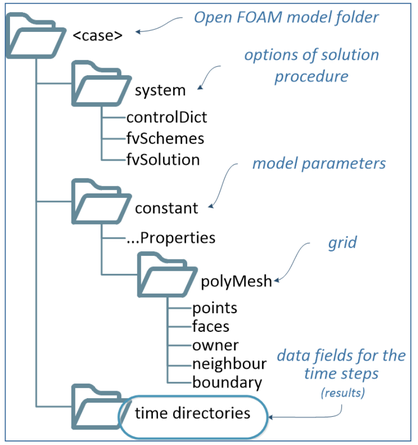
Real world differences: outputs

- Dimensionality:
  - One or several separate values
  - Time series
  - 2D or 3D field
  - Multiple different dependencies (for example heads and concentrations in groundwater flow and transport models, isotherms and ph-dependencies in sorption models)
- Different output formats:
  - Plain-text vs Binary vs Database
  - One file vs multiple files
  - Tables with header in the beginning of the file vs long output listings, where values could be found in the n^{th} line after keyword.

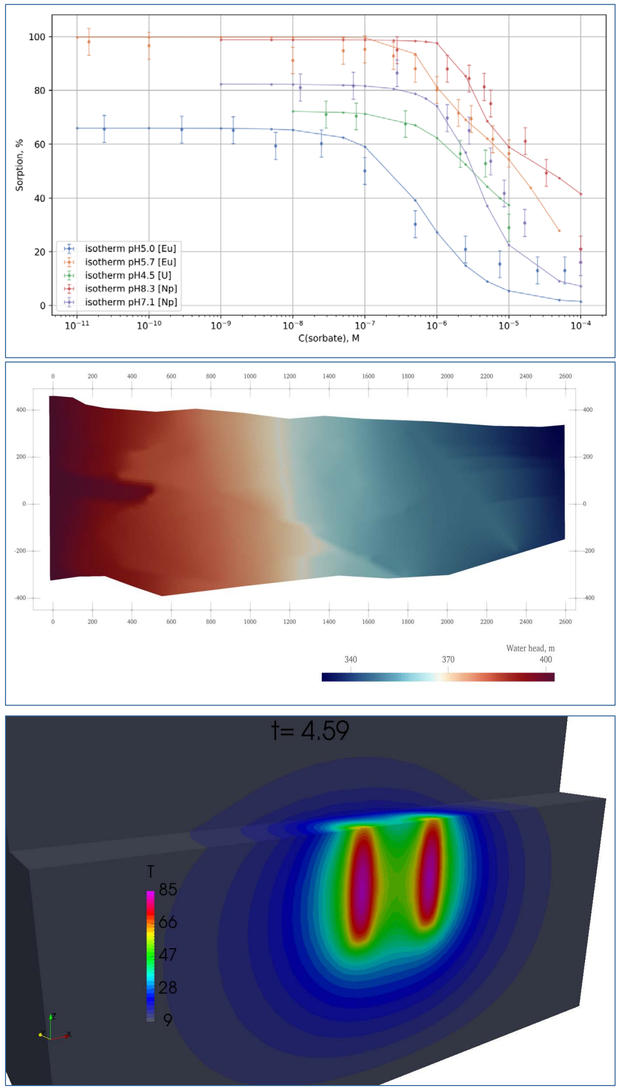
Examples of output formats

#vtk DataFile	*Header*
Data	*Title (max. 256 characters)*
ASCI | BINARY	*Data type(ASCII or binary)*
DATASET type …	*Geometry/Topology. One of:* *STRUCTURED_POINTS STRUCTURED_GRID FIELD UNSTRUCTURED_GRID POLYDATA RECTILINEAR_GRID*
POINT_DATA n	*Dataset attributes*
…	*Point dataset*
CELL_DATA n	*Dataset attributes*
…	*Cell dataset*

#GEO-EAS DataFile	*Header*
n	*Number of data columns*
Variable_1 … Variable_n (n)	*Column names*
data_value … data_value data_value … data_value data_value … data_value data_value … data_value data_value … data_value data_value … data_value (n)	*Data*



Different dimensions of outputs



Real world differences: analysis needs and properties

- Normalization of inputs: if one parameter changes in linear scale and another in logarithmic scale, should we harmonize them before analysis? Do we know this a priori?
- What do we do with time-dependent or spatial outputs: analyze each point, select significant points, analyze integral characteristics?
- Different objective functions for model calibration:

$$MSE = \frac{\sum_{i=1}^{N}(obs_i - sim_i)^2}{N}$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^{N}(obs_i - sim_i)^2}{N}}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(obs_i - sim_i)^2}{\sum_{i=1}^{N}(\overline{obs} - obs_i)^2}, \quad \overline{obs} = \sum_{i=1}^{N} obs_i$$

$$R^2_{adj} = (1 - R^2)\frac{N-1}{N-k}$$

obs_i – observed values

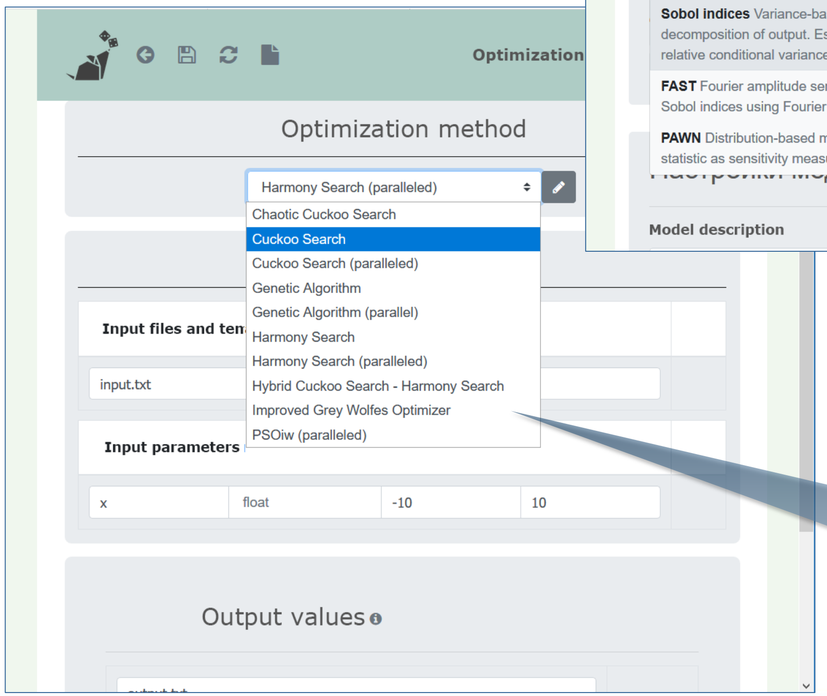
sim_i – simulated values

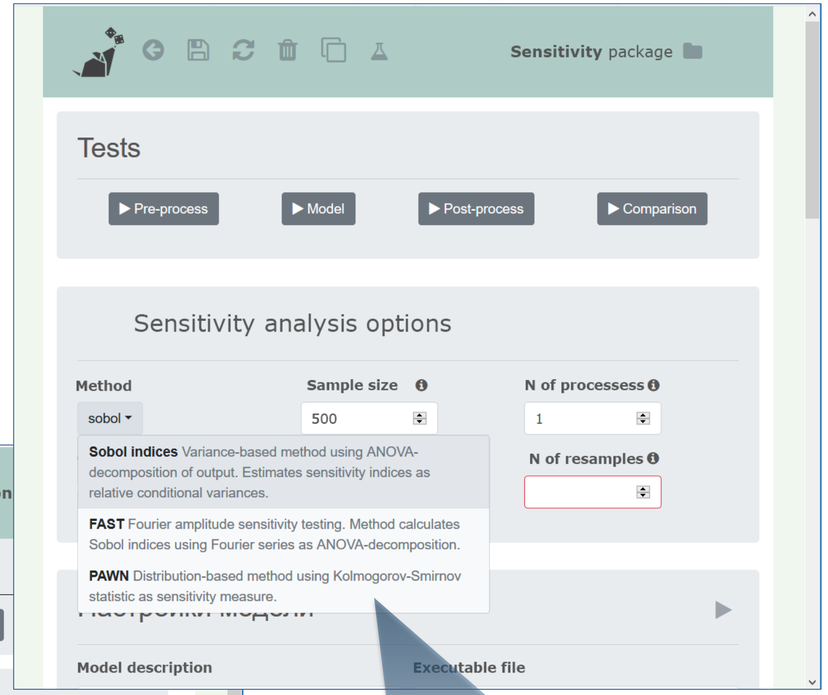
N – number of observations,

k – number of parameters

Method choice considerations

- Choice of the method for sensitivity analysis is conditioned by a lot of factors:
  - model properties if known (linearity, monotonicity, multimodality, asymmetry, and so on)
  - unfortunately – computational cost (a lot)
- Method choice for model calibration is basically empirical:
  - one heuristic algorithm to outperform another is to adjust to the structure of the specific problem (“no free lunch theorem”).
- Sampling:
  - A lot of sensitivity analysis methods require specific sampling strategies, could we use these samples also for output uncertainty analysis?
  - How we foresee the necessity to extend sample?



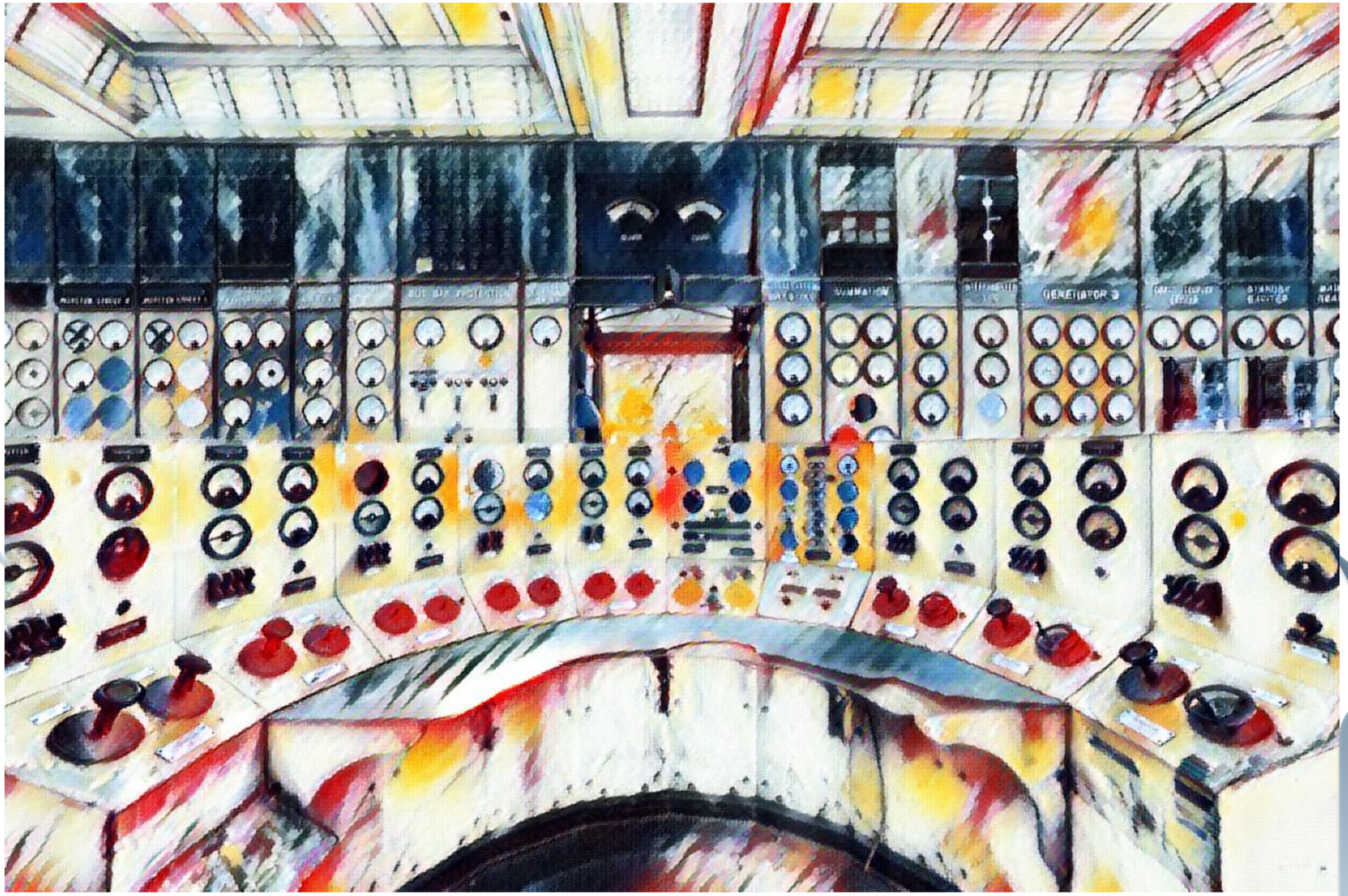


User 1: How do I choose between them?

User 2: Hmmm…That’s a lot of options, just tell me what’s best?

Hypothetical absolute model-independent uncertainty analysis software

User 1: Could you please help me with this just for the first time?

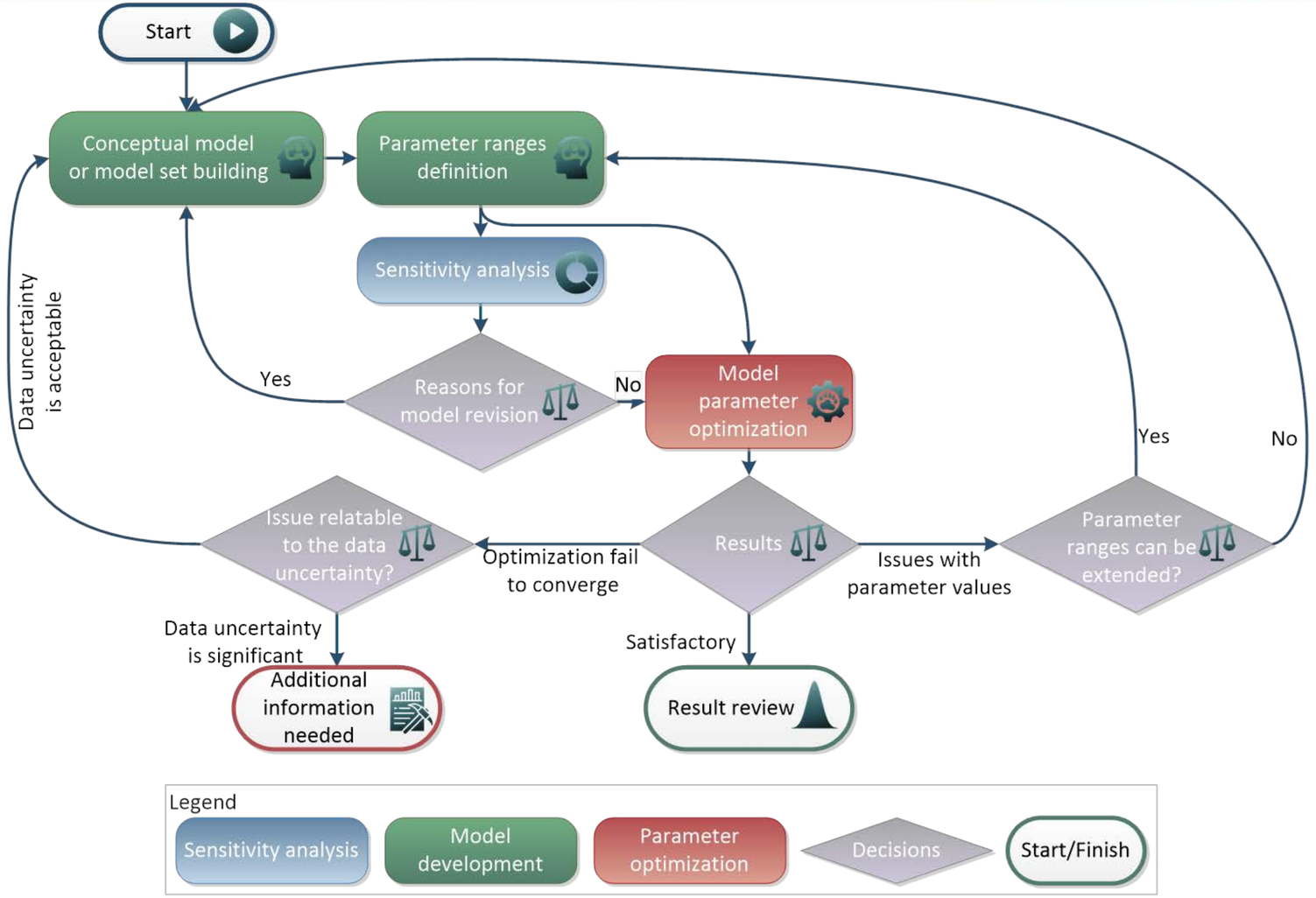


User 2: Hmmm…It's too complicated, maybe I could live with manual model calibration

Addition: model evolution due to sensitivity analysis and calibration

Another argument for “user-friendly” uncertainty analysis tools:

- Very often sensitivity analysis and calibration procedures are considered only as final one-time analysis steps.
- In our experience, if these results are interpreted and communicated then they lead to the new enhanced versions of the models



[5] Saveleva E, Svitelman V, Blinov P, Valetov D, Neuvazhaev G. Coupling of sensitivity analysis and model calibration in radioactive waste disposal safety assessment. Ninth International Conference on Sensitivity Analysis of Model Output, Barcelona: 2019.

Discussion: how to balance

- Uncertainty analysis software alternatives:
  - Model-independent package with lots of options, choices, most possible variants
    → definitely not user-friendly!
    → requires not only field expert (physicist, geologist, chemist, etc.), but «uncertainty analysis expert»
    → rarely takes into account modern tendencies in methods development (e.g. built-in sensitivity analysis in GoldSim)
  - Solutions with predefined options for specific group of tasks.
    - For example, GUI tool for calibration of geochemical models developed using PhreeqC
      → could not use it to calibrate MODFLOW groundwater flow model.
- How to balance?
  - Universal (model independent) library (Python, Matlab, R) + Model-tailored executable on demand

Conclusion

- Uncertainty and sensitivity analysis and calibration are often considered as final step out of scope of model development process.
- Uncertainty management for the safety assessment requires close co-operation of «model developers» and «uncertainty analysts».
- Unfortunately it is impossible to implement fully model-independent uncertainty analysis software – need to balance.