BIOMATH

Department of Applied Mathematics,
Biometrics and Process Control

Optimal Experimental Design in River Water Quality Modelling

Veronique Vandenberghe
Ann van Griensven
Willy Bauwens
Peter Vanrolleghem

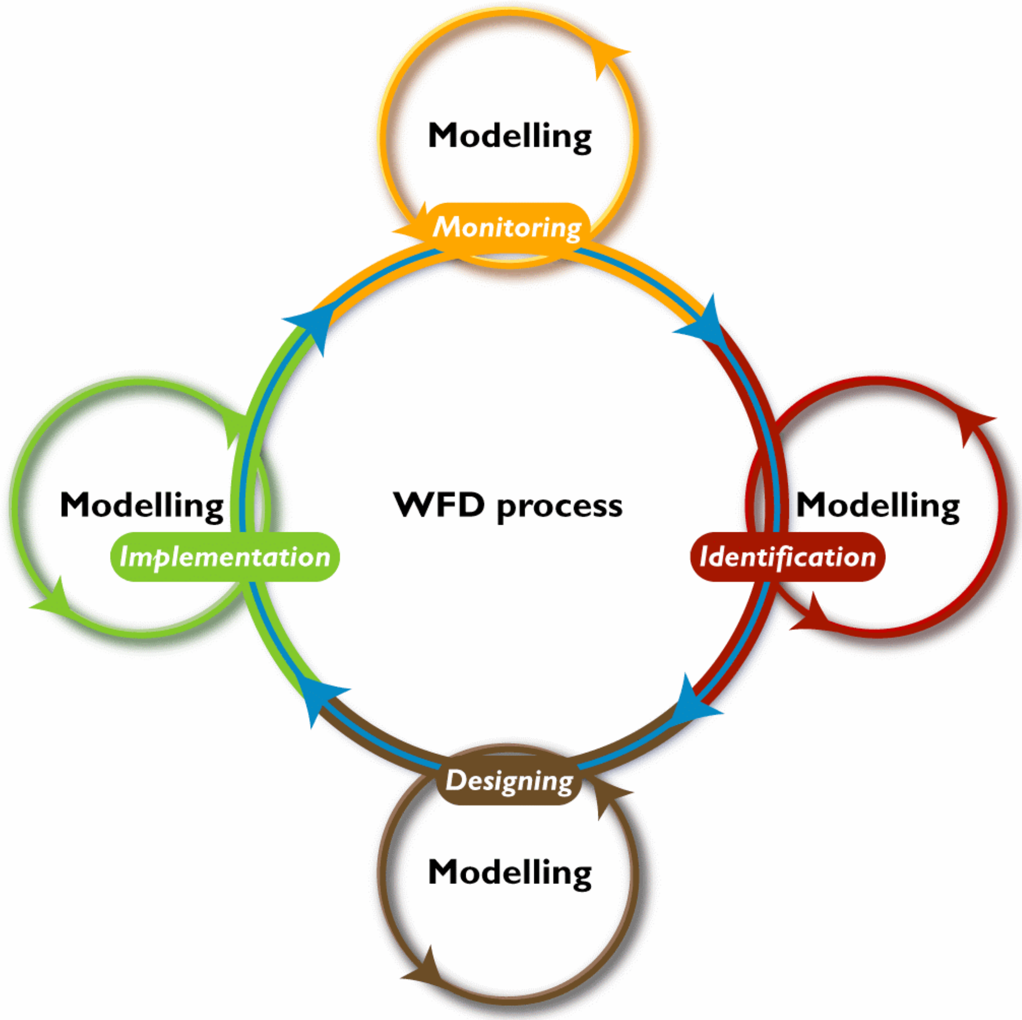
Overview

- Introduction
- Aim
- ESWAT
- Dender case
- Methodology
- Results
- Conclusion

Monitoring

- For the implementation of the WFD is monitoring an important step / for identification/evaluation: mostly ok.

- Models can play a role in every step. So also monitoring needed for building models for evaluation of future pollution abatement scenarios



Monitoring

Problems

- Costly measurements: eg. bod, suspended solids, micro-pollutants
- measurements with wrong frequency, in non-sensitive periods, on wrong places
- lot of effort to maintain large databanks
- Intensive measurement campaigns

510000	12/18	BZV5	=	5,1	mgO2/L
510000	12/18	Cd t	<	0,2	µg/L

510000	11/20	O2 verz	=	79	%
510000	11/20	CZV	=	33,55	mgO2/L

AIM

- **Calibration of water quality models**
  - **identifiability of the model parameters**
  - **reliability of the model**

=> depending on good measurements

Find an optimal set of sampling data for the calibration of a water quality model

Integrated modelling tool

SWAT98

- Catchment hydrology
- Agricultural pollution
- Constant point sources

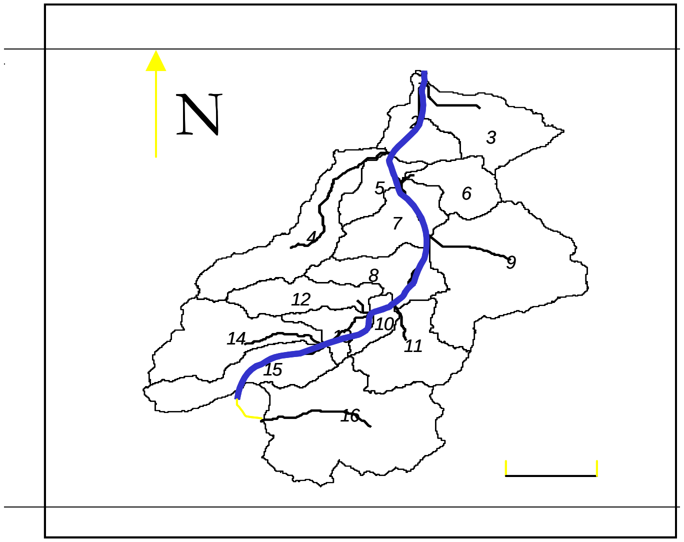
ESWAT (van Griensven)

- Hourly time step (land and river hydrology)
- River water quality processes
- Dynamic point sources
- Urban drainage system

Dender basin model

MODEL: 700 km²

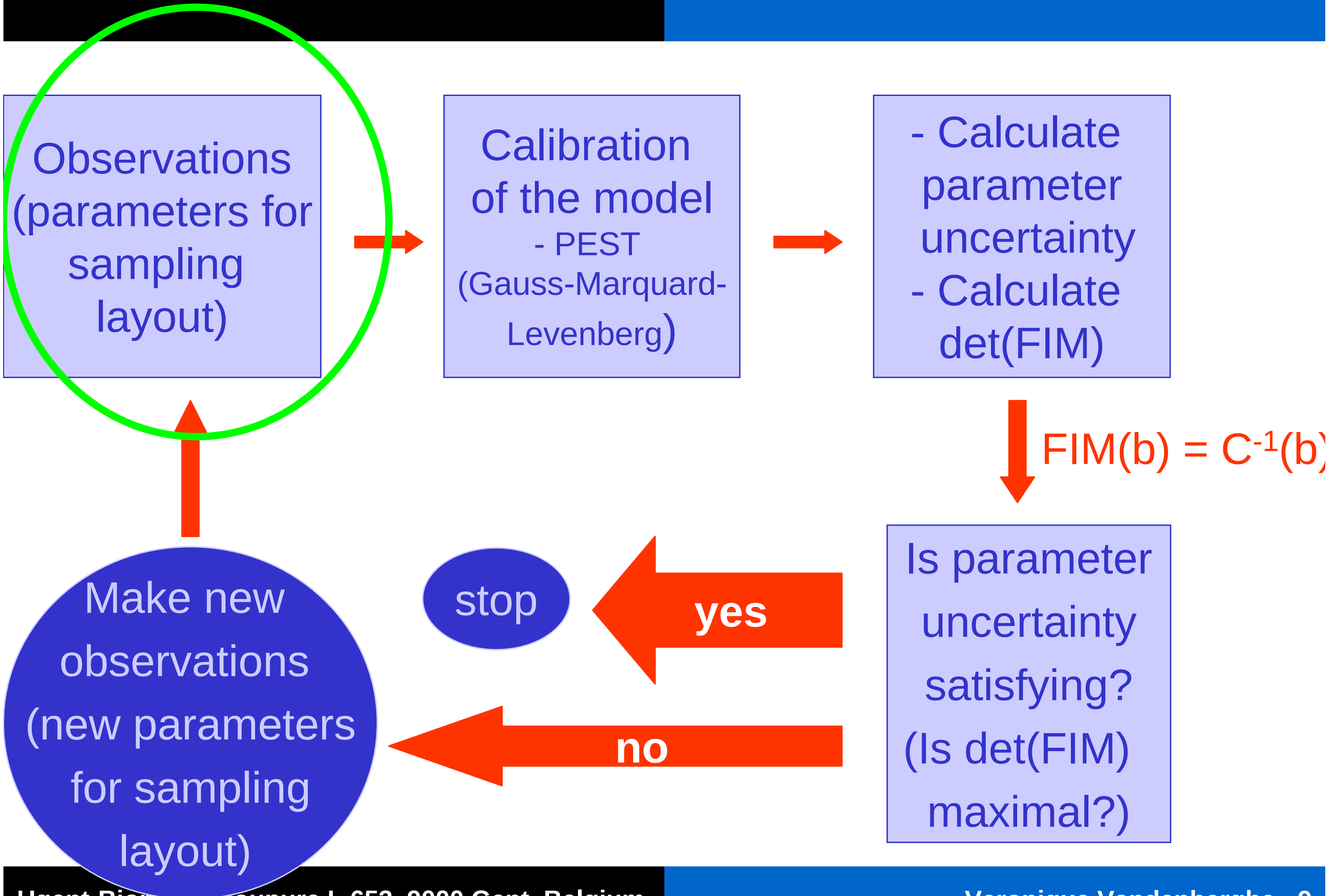
- 15 subbasins / 8 tributaries
- 80 HRU's (=combination land use and soil type)
- 10 point source locations
- 8 sluices



OPTIMAL EXPERIMENTAL DESIGN

- Start with a calibrated model but with high uncertainty bounds caused by a lack of good measurements for a better calibration.
- Generate time series with the model for water quality variables
- See where, when, how many measurements/samples are practically feasible and define parameters of sampling layout

Methodology

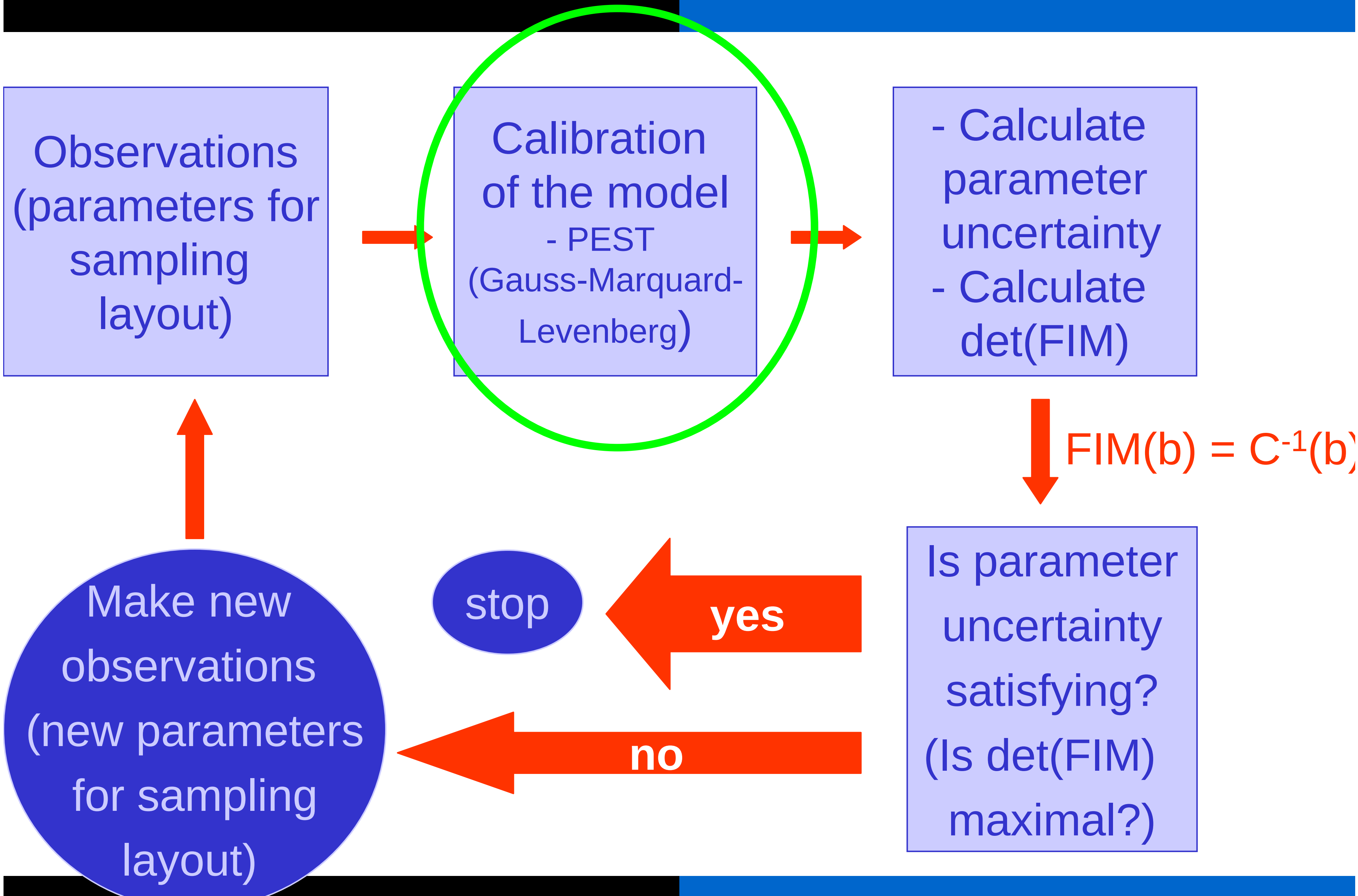


observations

- **No historic series of high frequency water quality data series available**
- **=> Synthetic observation series generated by ESWAT + pseudo-random noise terms**
- **Noise terms: consistent with the accuracy of the measurement devices**
- **Parameters for sampling design: frequency, measurement place, period of the year**

  e.g. every 12 hours, at the mouth + 5km more upstream, from 01/04 till 31/08

Methodology



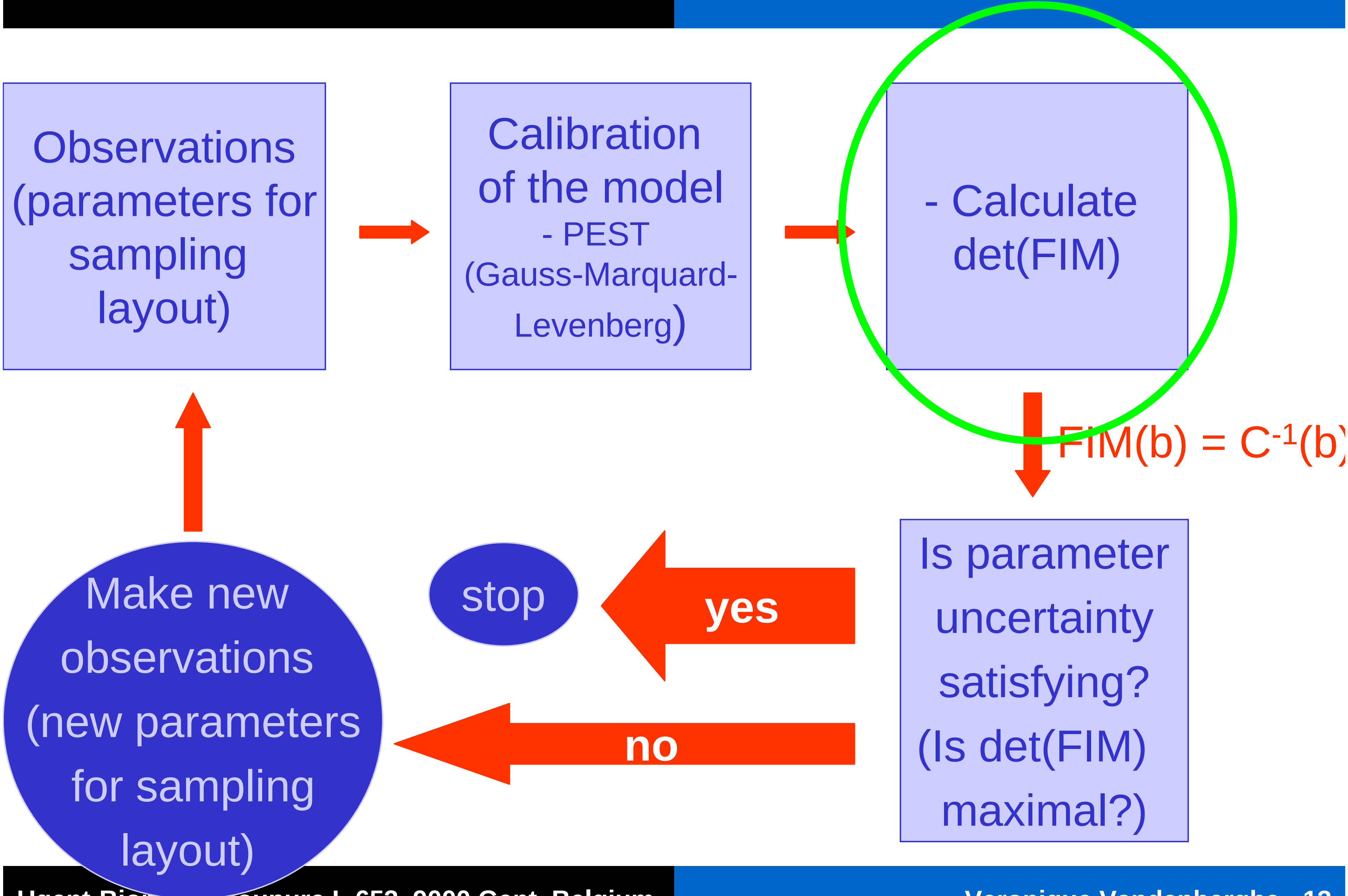
CALIBRATION

- Start calibration at a point ‘close’ to the former calibration (to avoid local minima)

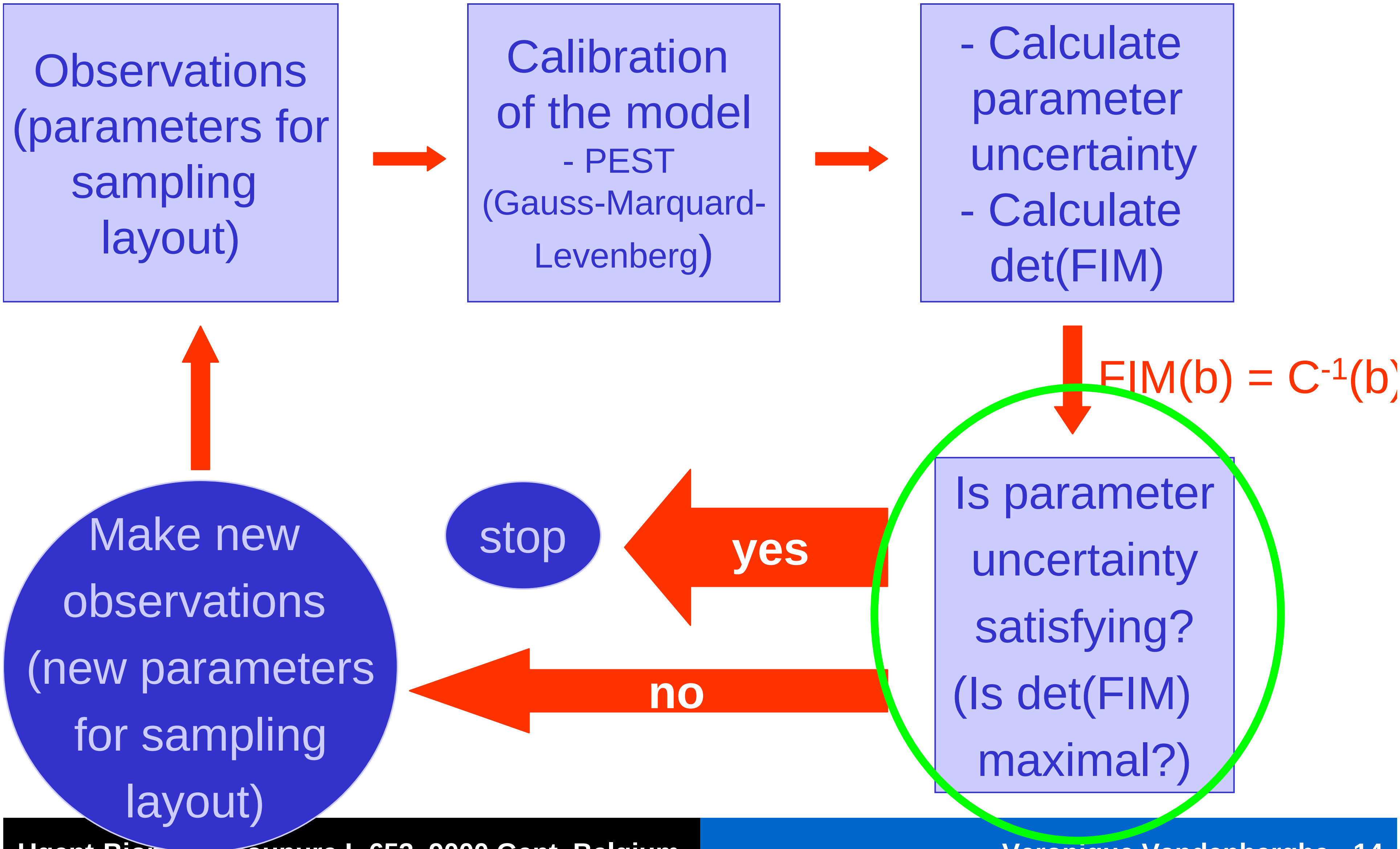
⇨ **During the calibration the covariance matrix and the uncertainty bounds on the parameters are calculated**

- PEST: uses Gauss-Marquard-Levenberg

Methodology



Methodology



D-optimal design technique

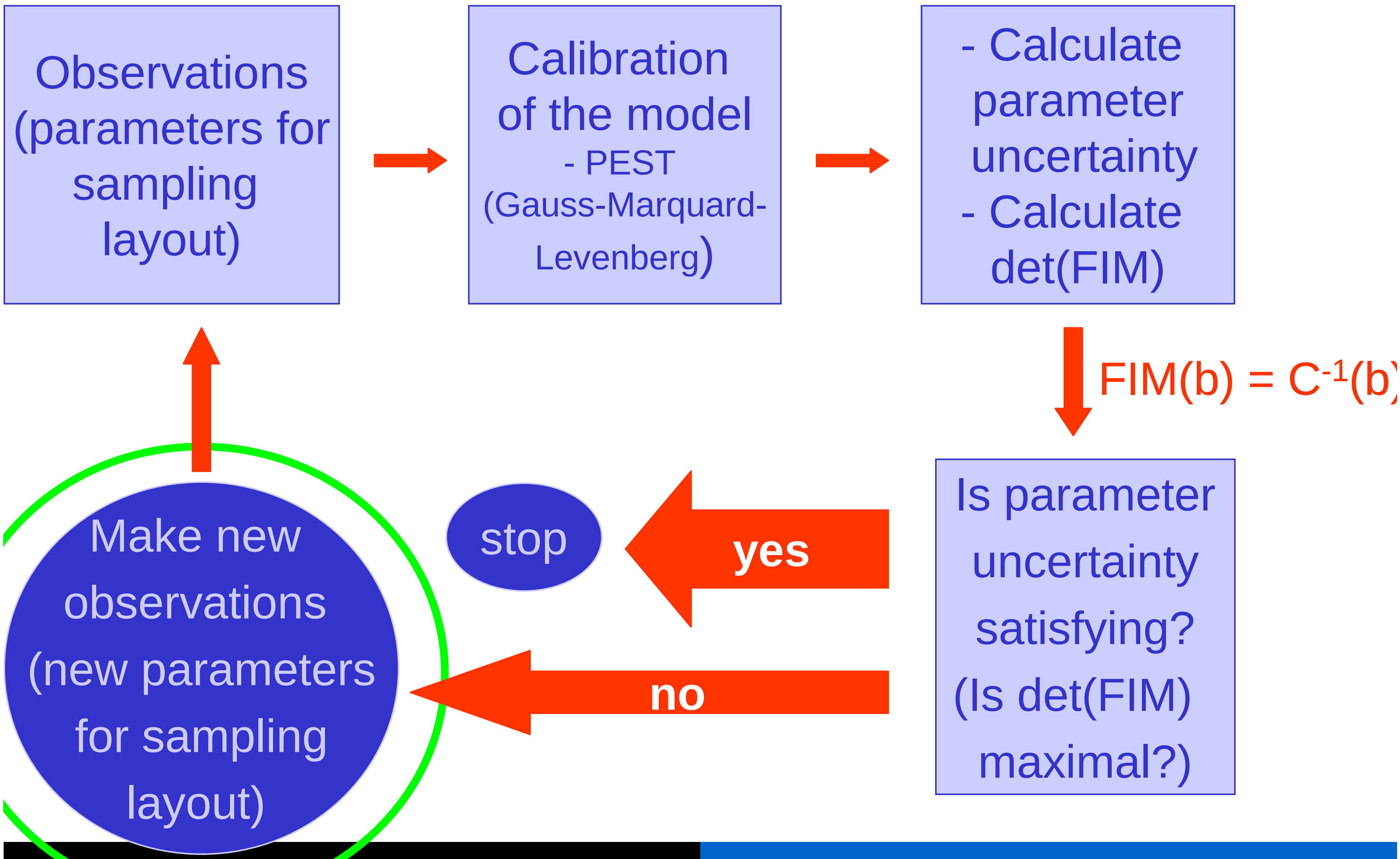
- new measurement layout => with D-optimal design technique (= maximisation of Det(FIM))

$$C(b) = \partial^2 (J^t Q J)$$

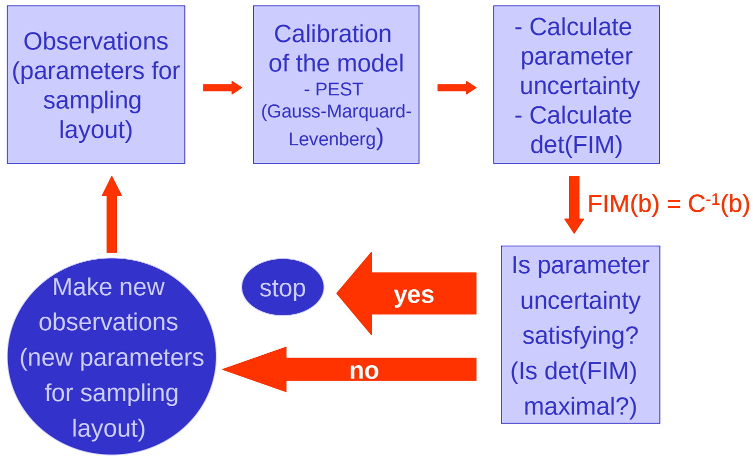
$$Fisher\ Information\ Matrix = C^{-1}(b)$$

- maximisation with shuffled complex technique (SCE-UA method) by changing parameters of sampling layout (frequency, place, period,…)

Methodology



Methodology



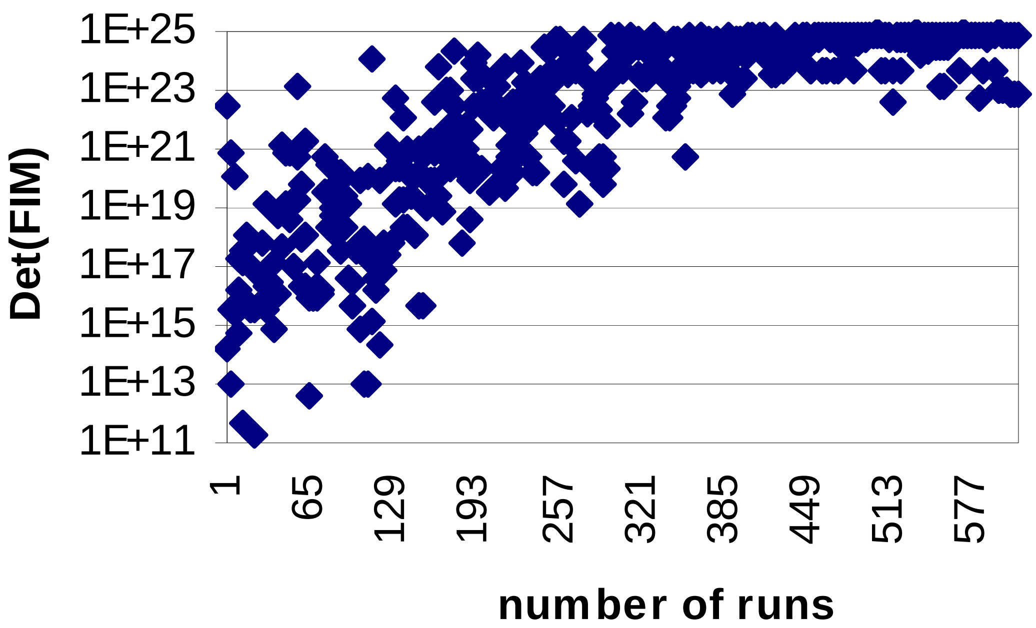
- **When 'STOP' = make new measurements following the obtained sampling parameters**
- **Then start the procedure all over again=>iterative method**
- **The sampling procedure can change year after year, following natural trends e.g. climate change or human interactions e.g. a dam, a new discharge point**

Results

Five parameters of the sample layout variable:

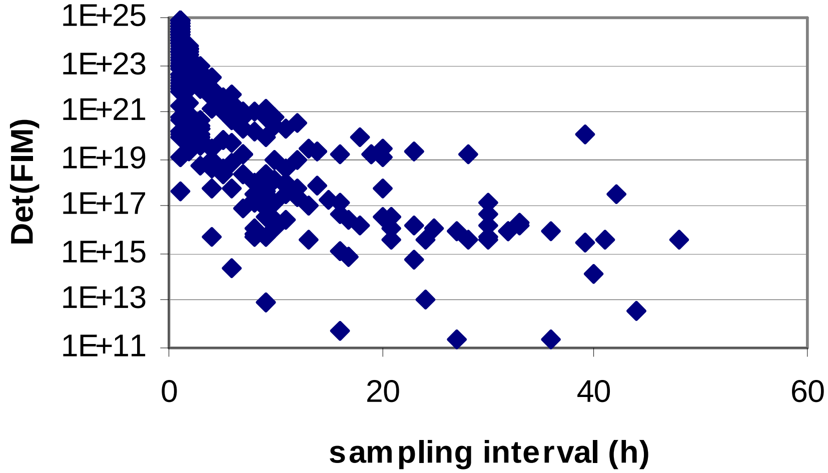
- **frequency (every hour - every two days)**
- **period (summer, winter, mixed summer-winter)**
- **total amount of samples (365*24)**
- **only DO or combined DO-NO_3, DO-NO_3-BOD or DO-NO_3-BOD-NH_4)**
- **the sample location (4 possible combinations of 3 possible locations: upstream, halfway, downstream)**

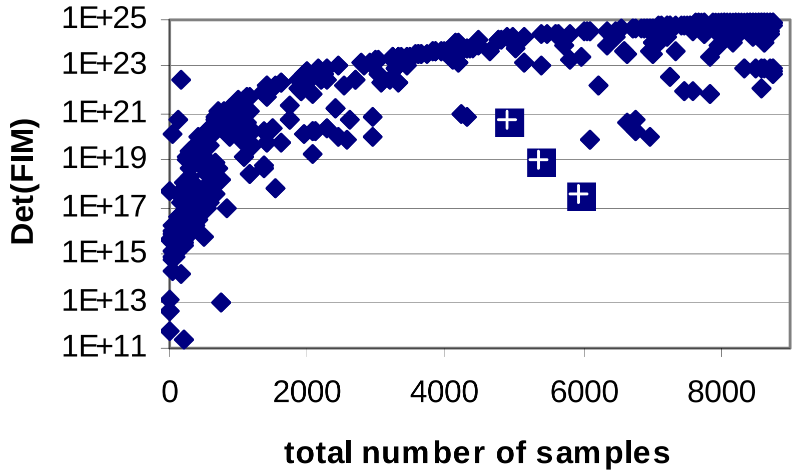
Results



- 635 iterations needed

Results





- **Best : hourly time base, nearly the whole year, on three locations and with the four variables**
- **Other sampling schemes possible that provide a quasi simular accuracy**
- **Some sampling schemes non-optimal**

non-optimal sampling schemes

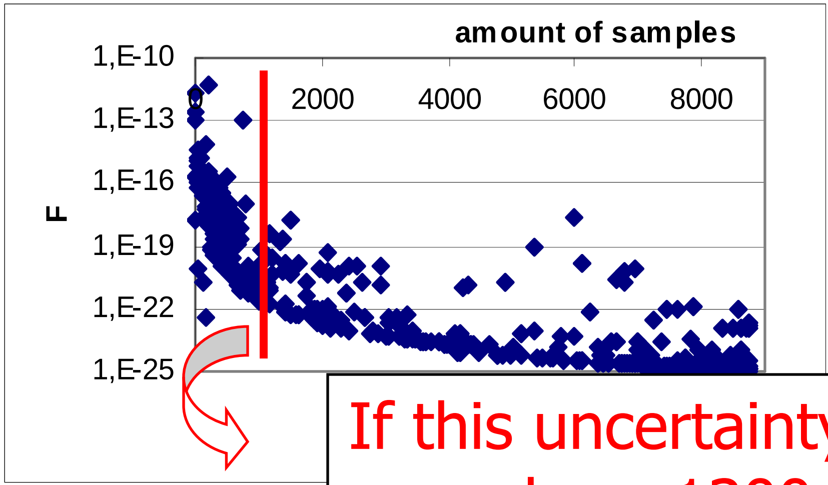
Table 1. Non-optimal sampling designs

Sampling interval (h)	Number of samples	Period	Location	Observed variables	Det(FIM)
1	5972	16 Apr.-31 Dec.	Geraardsbergen	DO-NO_3	4,08E+17
1	5340	22 May-15 Nov.	Geraardsbergen	DO-NO_3-BOD	1,19E+19
1	4902	11 May-31 Dec.	Geraardsbergen	DO-NO_3-BOD	5,92E+20

- **Factors that are influencing:**
  - **place (upstream)**
  - **period (no measurements in spring)**
  - **no NH_4**

OED with practical considerations

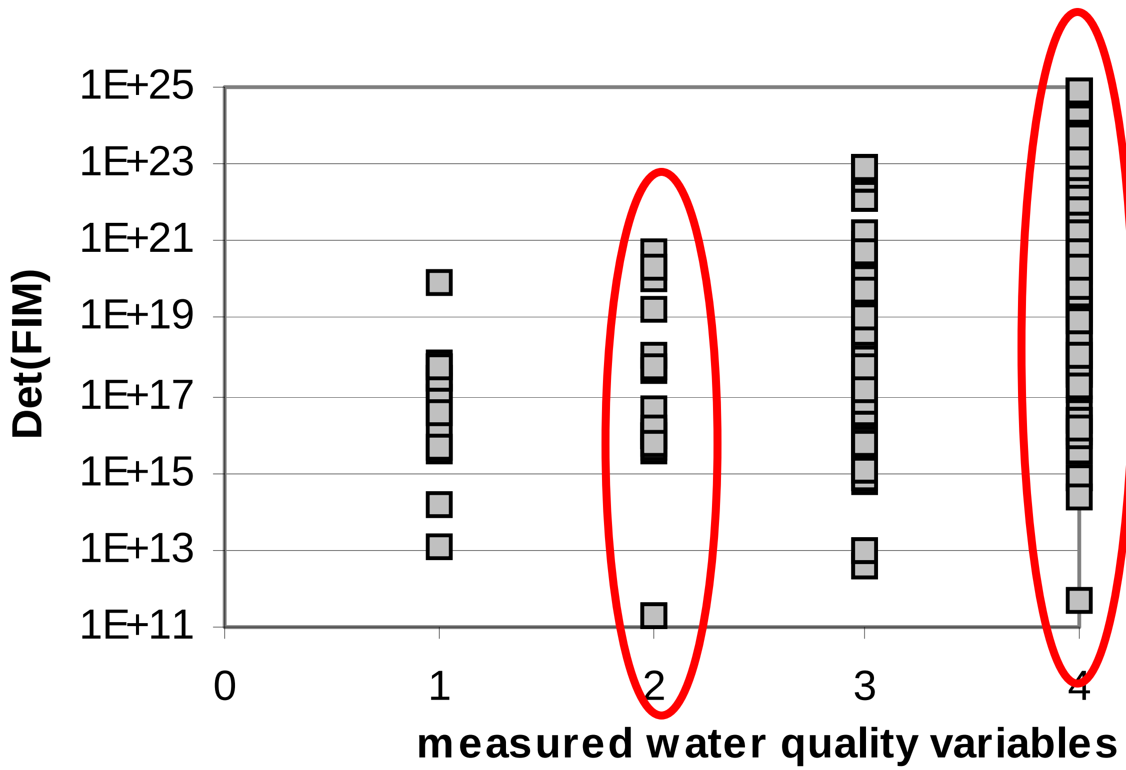
- Final uncertainty on model results is important => can be related to cost and practical implications



If this uncertainty is good enough=> 1300 samples instead of 85OO

Results

OED with practical considerations

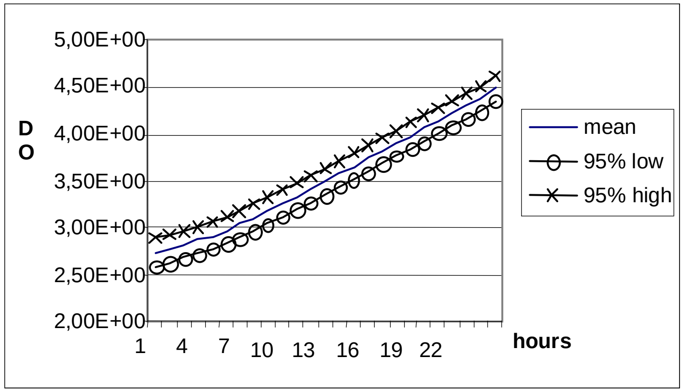


Replacement of expensive measurements with cheaper ones?

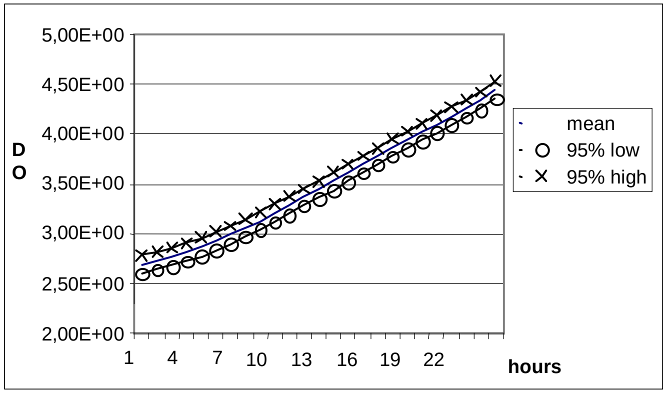
DO + NO_3 or DO + NO_3 + BOD + NH_4

Results

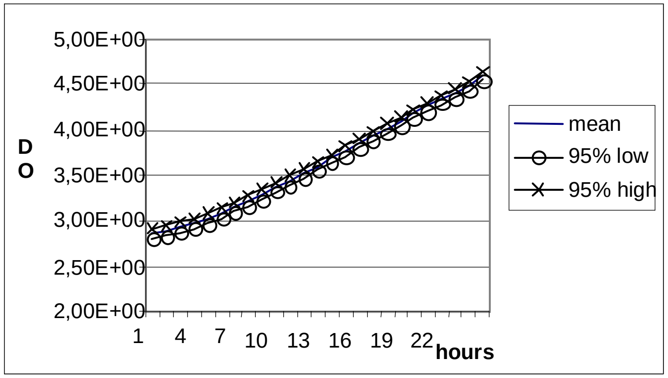
Comparison of three sampling schemes



1



2



3

- Average width of the confidence interval on the model results:

- reduction of 45% (2-1)

- reduction of 60% (3-1)

Conclusion

- **OED for calibration of water quality models => measurement strategy**
- **Dender: optimal sampling strategy with highest amount of samples, highest sampling frequency, max. number of locations and max. number of variables measured.**
- **Usefulness of the method: evaluation of sub-optimal sampling strategies, in view of limitations, costs and practical considerations.**