CHARACTERIZATION OF LOCAL AND REGIONAL UNCERTAINTY IN CASE OF MATURE HC RESERVOIRS

Geiger, J., Malvic, T., Horvath, J. Zelenika, K., N.

INTRODUCTION-I.

- Uncertainty ≠ Error
  - Uncertainty expresses our incomplete knowledge → It can not be eliminated
- Types of uncertainties:
  - Local → locations/grid nodes
  - Regional → Above a reservoir
  - Model → Used equation system
- Regionalized variable
  - Multidimensional probability distribution
    - Random variable around a grid point or well location → Local uncertainty
    - There is a function which connects them → Regional uncertainty

INTRODUCTION-II.

SEQUENTIAL SIMULATIONS

- For derivation local PDF's ⟹ Simulation of local uncertainty
- For derivation many realizations of that random function which conects the local PDF's in the space ⟹ Simulation of regional uncertainty

KEY QUESTIONS:

- How many realizations would be enough to get stable description of local uncertainties?
- Does this number depends on grid resolution? How?

GOALS

- **convergence** of the multidimensional conditional distributions derived from increasing subset sizes of 100 realizations which are originated from sequential indicator simulation of well-averaged porosity data.
  → The ***first two moments***:
    - Average/expected value
    - Variance
- how the **local and regional uncertainty** (Goovaerts, 2006) can be formulated by the terminology of **variance components**.
  → ***Variance decomposition***
    - Within group variance (WGV)
    - Between group variance (BGV)
- how criterions for assessing the goodness of simulations given by Deutsch (1997), Goovaerts (1999, 2006) and Emery (2008) can be supplemented on the basis of the theory of variance decomposition.
  → ***Convergences of WGV and BGV***
- an example for the process we suggest
  → ***Turbidity reservoir***

SEQUENCE OF MULTIDIMENSIONAL DISTRIBUTIONS

- Set containing I^{ℓ} stochastic realizations (ℓ=*1,2,...,L*)
- Subsets from I^{ℓ}: $S_{\ell} := \bigcup_{j=1}^{\ell} I_j, \ \ where\ \ell = 1,\ldots,L$
  - ℓ=1 ⟶ *N* nodes
    ℓ=2 ⟶ *2N* nodes
    : :
    ℓ=*L* ⟶ *LN* nodes
  - for each ℓ S_{ℓ} is a realization of the multidimensional distribution generated by the algorithm
  - $S_{\ell} \underset{\ell \to \infty}{\Rightarrow} S$

SEQUENCE OF THE GRID AVERAGES

- E_ℓ^g an average value calculated for S^ℓ $(S^\ell := \bigcup_{j=1}^{\ell} I_j,\ \ell = 1,\ldots,L)$ under g grid resolution.

- $SE_\ell^g := \frac{1}{\ell \cdot N} \sum_{j=1}^{\ell \cdot N} z^{(i)}\left(u_j\right),\ \ell = 1,2,\ldots.$ a sequence of averages for increasing ℓ

  - $$SE_1^g := \{average\ from\ N\ nodes\}$$
    $$SE_2^g := \{average\ from\ 2N\ nodes\}$$
    $$\vdots$$
    $$SE_L^g := \{average\ from\ LN\ nodes\}$$

  - $\lim_{\ell \to \infty} SE_\ell^g = E^g$ according to the strong law of large numbers

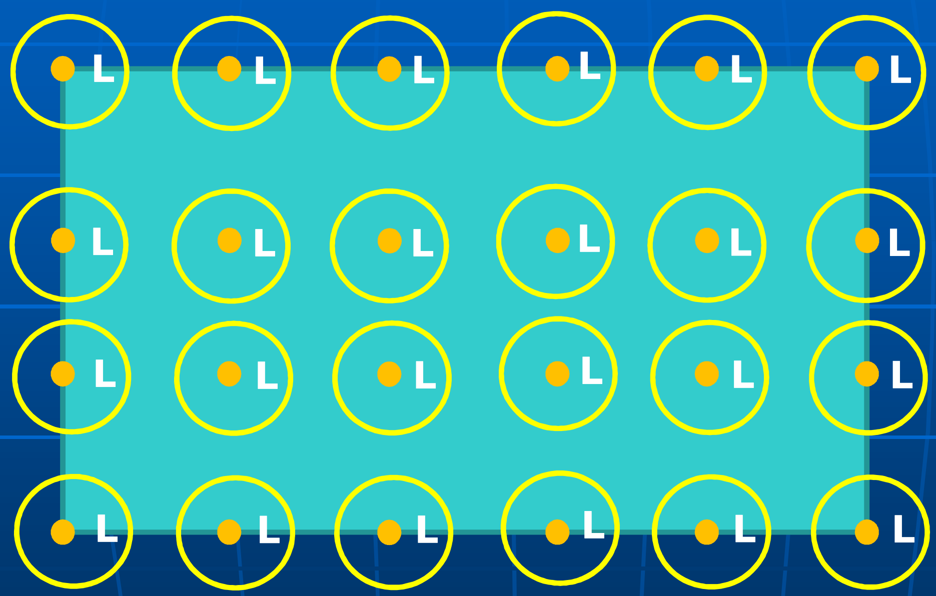
ABOUT THE CONVERGENCE OF GRID AVERAGES

$$\lim_{\ell \to \infty} SE_{\ell}^{g} = E^{g}$$

- Valid only for the underlying *g* grid geometry since
  - stationary is the property of the multivariate random function and not that of the underlying geological process;
  - this random function may change with the change of scale.
- It is conditioned to
  - the original data set
  - spatial configuration.
- A kind of ‘conditional convergence’.

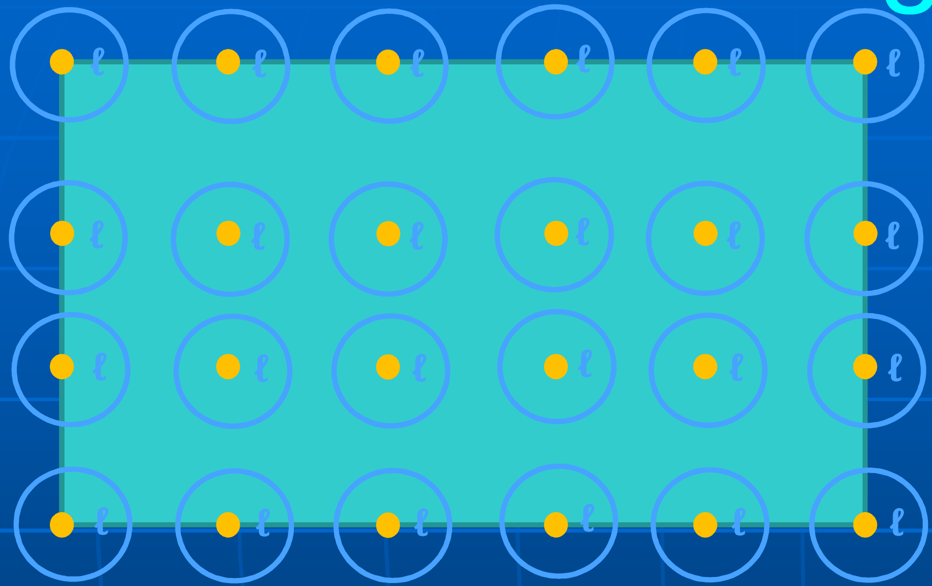
VARIANCE DECOMPOSITION-I

- There are *L* stochastic images above *N* grid nodes



- One-way ANOVA: **there is a set of *LN* values which are nested within *N* nodes**

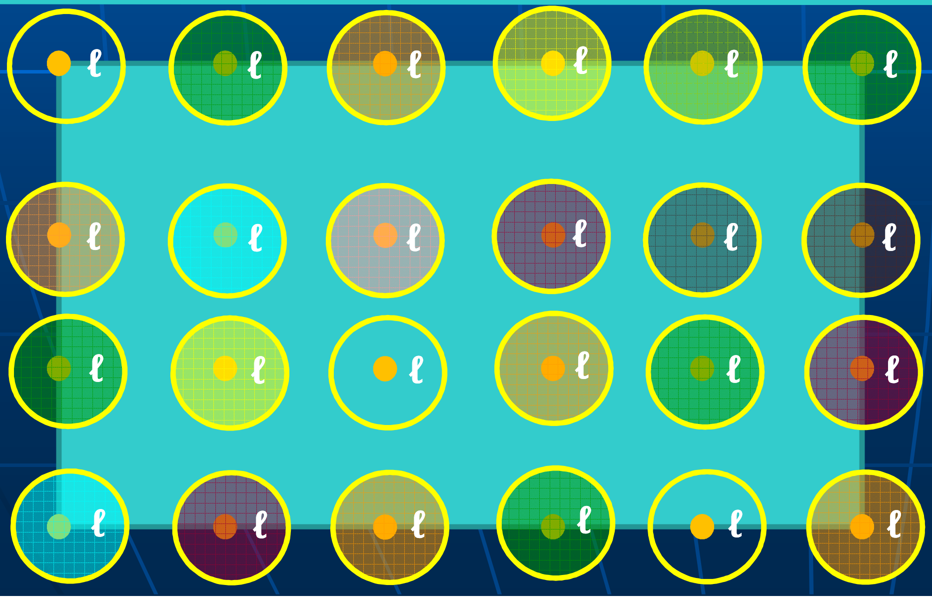
VARIANCE DECOMPOSITION-II
Averages



AVERAGE FROM *N*ℓ VALUE

ONE AVERAGE FOR THE SET OF ℓ REALIZATIONS

GRID-AVERAGE

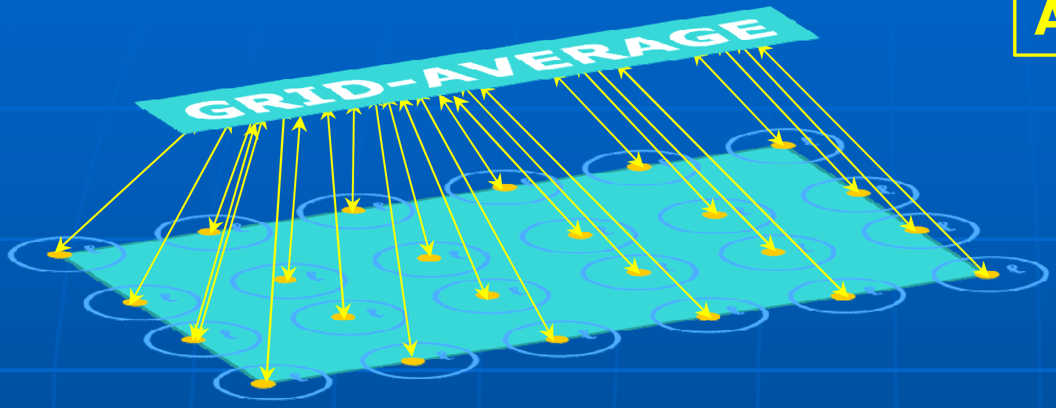


AVERAGES FOR EACH NODES FROM THE ℓ SIMULATED VALUES

E-TYPE ESTIMATION

NODE-AVERAGES

VARIANCE DECOMPOSITION-III
Within-group and between-group variances



AROUND THE GRID AVERAGE

It measures variance of the node-averages (group averages) around the grid average (main-average).

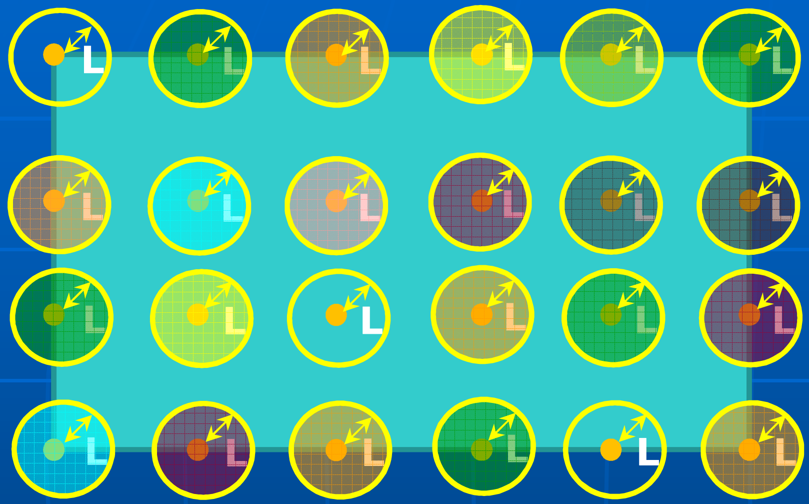
$$BGV = \frac{1}{SUM}\sum_{i=1}^{m} n_i \cdot \left(\bar{x}_i - \bar{X}\right)$$

AROUND THE GRID-NODE AVERAGES

It measures how individual values (simulated node values) deviate from their particular node average (group average).

$$WGV = \frac{1}{SUM}\sum_{i=1}^{m} n_i \cdot \sigma_i^2$$

INTERPRETATION OF WGV



$$WGV = \frac{1}{SUM} \sum_{i=1}^{m} n_i \cdot \sigma_i^2$$

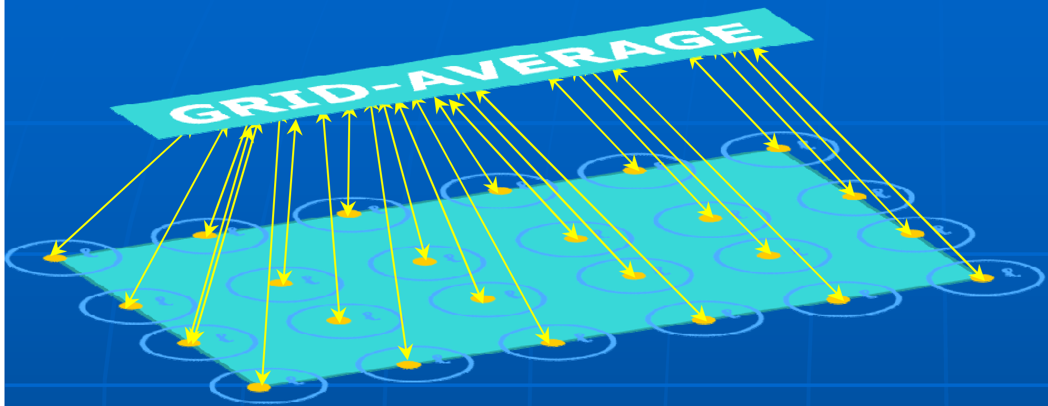
It measures how individual values deviate from the group average in the whole sample.

The smaller the WGV the more stable the estimations can be derived at grid nodes , in average.

This component measures the average stability of our estimation at grid nodes in the ℓ realizations.

Its convergence means that the estimation stability approaches such a value which can not be decreased so far within the simulation method and the *g* grid resolution. The limit is the **accesible maximum estimation stability**

INTERPRETATION OF BGV



$$BGV = \frac{1}{SUM} \sum_{i=1}^{m} n_i \cdot \left(\bar{x}_i - \bar{X} \right)$$

It measures variance of the group averages (node averages) around the main average (grid-average).

Node averages show a direct relation to the underlying geological process

⇓

The more heterogeneous this geological background the larger the variance being measured between node averages.

⇓

It can reflect the degree of lateral heterogeneity of that geological process which is transformed to the grid.

In the limit, the E-type grid shows that lateral heterogeneity which will not change so far under the used simulations algorithm and grid resolution.

⇓

The limit of the sequence of between-group variances is exactly that variance which in this situation measures the variability between grid nodes.

RELATIONS BETWEEN WGV AND BGV-I

If (WGV < BGV) and ($\ell < \infty$)

- Under the given ℓ realizations, the heterogeneity of the geological process is higher than the average estimation stability obtained at grid nodes
- Under the given grid resolution, the total variance of the **E-type grid** can rather be drawn back **to the underlying geology** than the instability of the estimation method applied.
- The pooled ℓ realizations on the given grid resolution **can be accepted**, because it reflects mostly the underlying geology
- Since ℓ is finite, there is **a possibility for changing this relation** if the number realization is increased.

If (WGV < BGV) in the limit

- there is **no way for changing** this relation without altering either the simulation algorithm or the grid resolution.
- **in general** the actual grid resolution and simulation algorithm are **adequate** to the given problem. At least in the sense that if ℓ is large enough, the E-type estimation of the pooled realizations will be affected mainly by the geological situation and not the instability of the node-estimation processes.

RELATIONS BETWEEN WGV AND BGV-II

If (WGV > BGV) and ($\ell < \infty$)

- under the given ℓ realizations the revealed lateral heterogeneity of the underlying geological process is smaller than the average estimation stability obtained at grid nodes.
- under the given grid resolution, the total variance of the **E-type grid is mainly controlled by the instability of estimation** method applied.
- the pooled ℓ realizations on the given grid resolution **cannot be accepted**, because it does not reflect the geology.
- since ℓ is finite, there is **a possibility for changing this relation** if the number realizations is increased.

If (WGV > BGV) in the limit

- there is **no way for changing this relation** without altering either the simulation algorithm or the grid resolution.
- **in general** the actual grid resolution and simulation algorithm are **not adequate** to the given problem. At least in the sense that if ℓ is large enough, the E-type estimation of the pooled realizations will be affected mainly by instability of the local estimations.

APPROXIMATION OF THE LIMIT

- the limit must be existing; ⟹ - *Unquestionably*
- the “tendency” of the members of sequences can be evaluated from the finite number of members ⟹ - *Rather ‘hope’…*

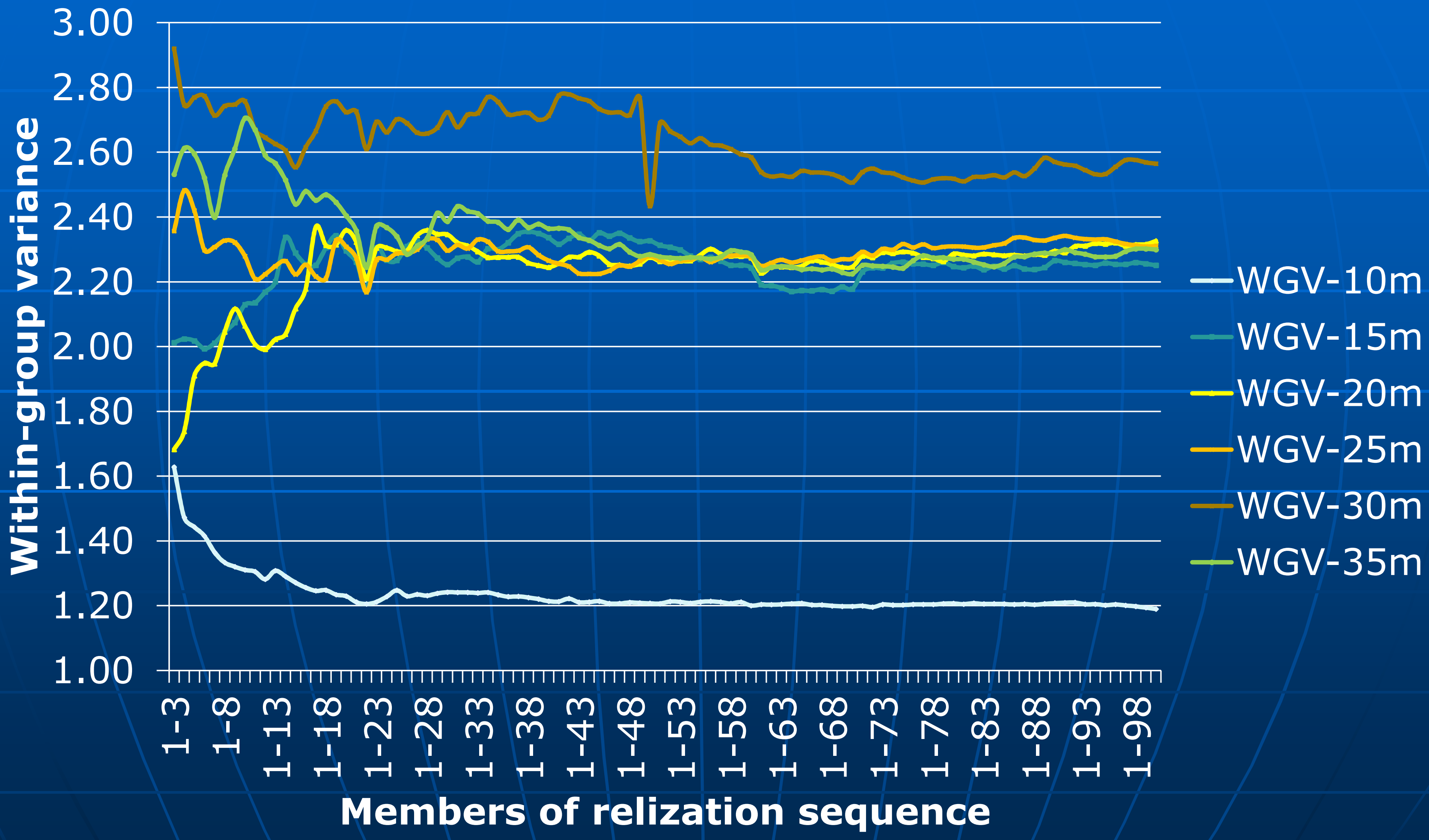
The analytical form of the sequences can be approximated by a bounded, monoton, non-linear regression, if its domain is restricted to the set of integer numbers

If ℓ_{min} be that particular serial number from which this approximation can be done, then

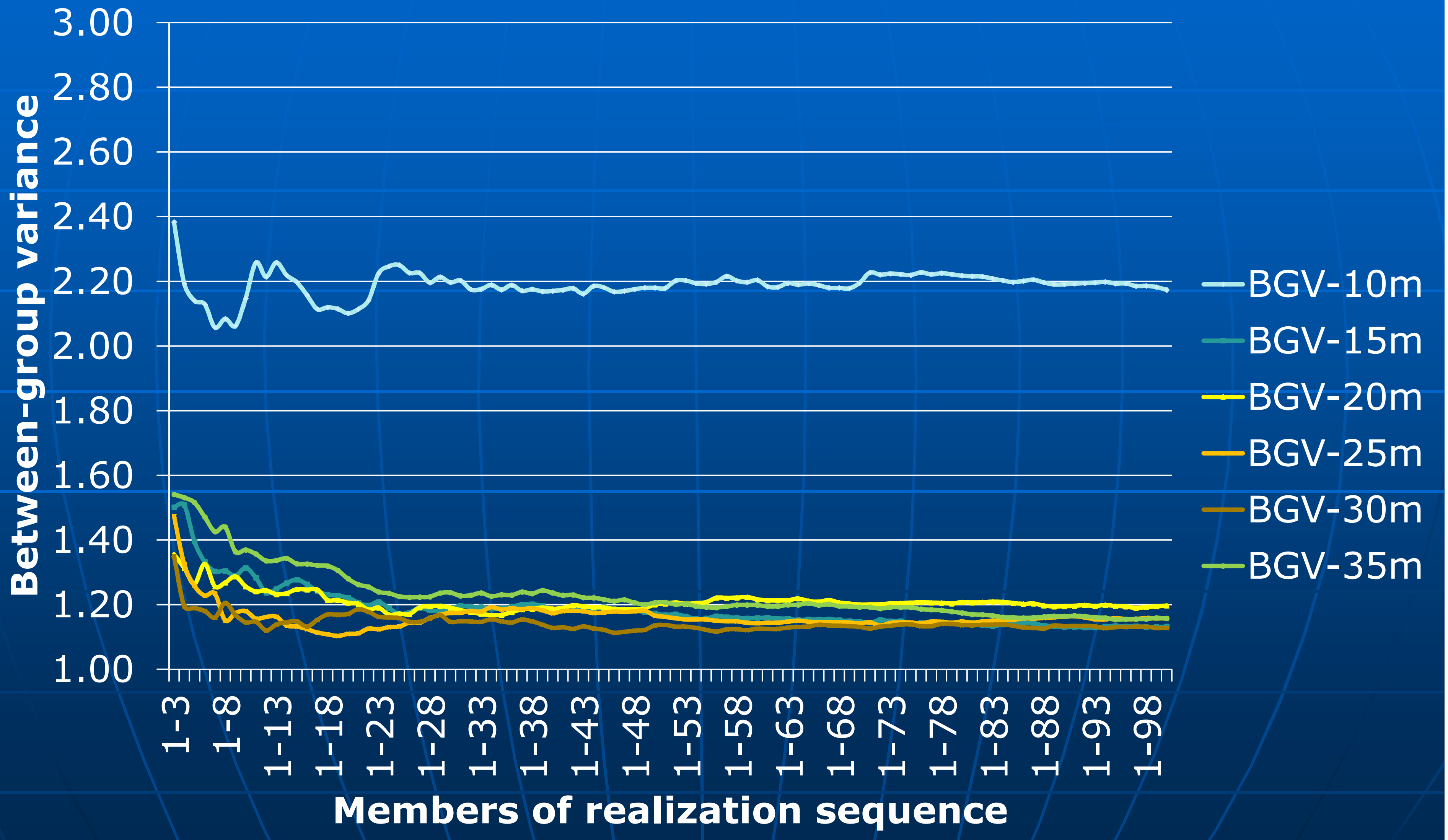
since the convergence does not depend on omitting finite number of members,

$\underline{\ell_{min}}$ gives the minimum number of realizations which is <u>necessary</u> <u>but not sufficient</u> for assessing the regional uncertainty.

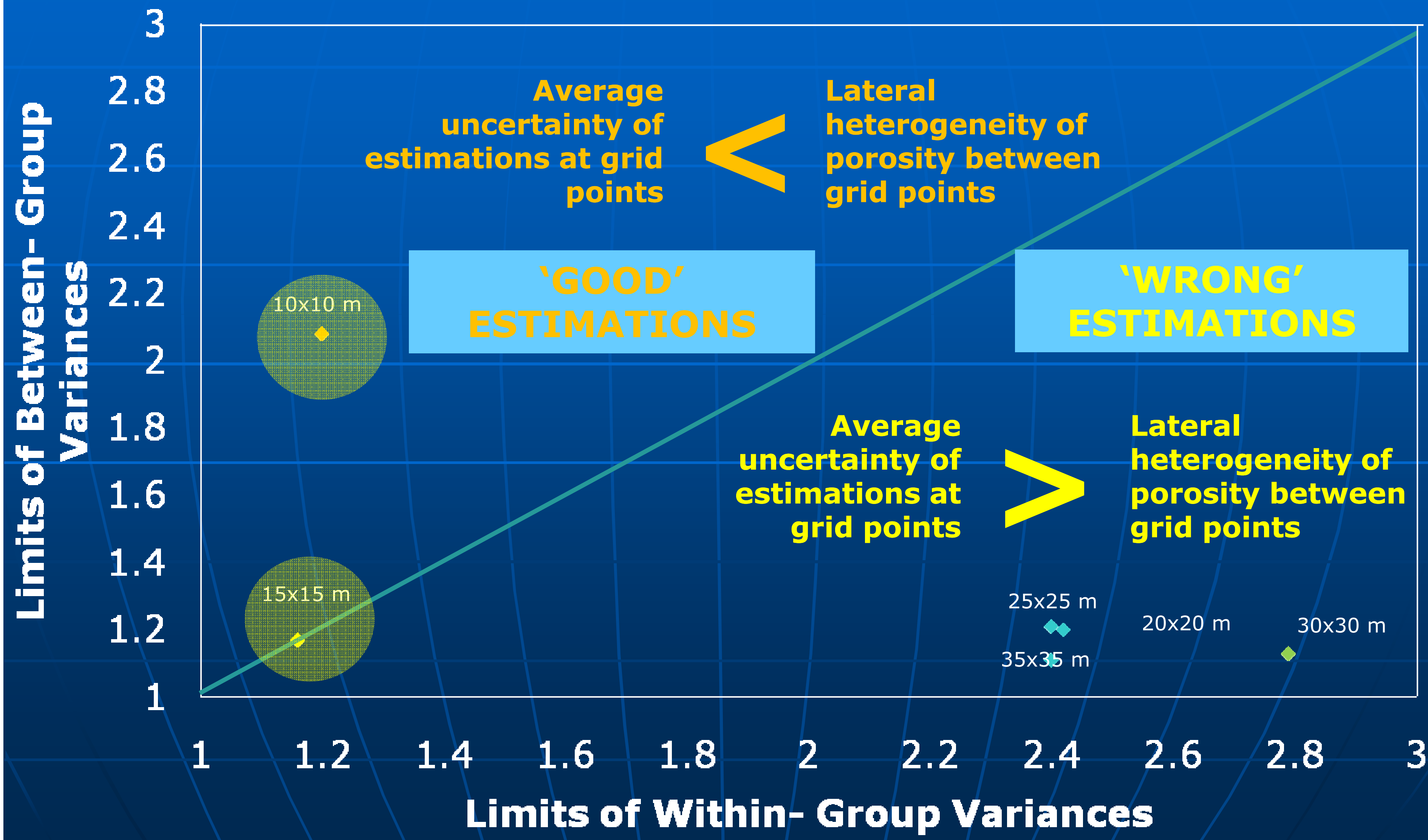
CONVERGENCE OF WITHIN GROUP VARIANCES (average estimation stability)



CONVERGENCE OF BETWEEN GROUP VARIANCES
(revealed average geological heterogeneity)



CRITERIA FOR 'GOOD' AND 'WRONG' ESTIMATIONS



CONSEQUENCE

- "How many realizations are enough for reaching a 'good' spatial average?"
  - Note: "if the limit has to be approached within a particular (ε_0) deviation"
  - this question can be drawn back to the Heine's definition of convergence

$$\left|SE_\ell^g - E^g\right| < \varepsilon_0,\ \left(\ell_0 < \ell\right)$$

$$\left(SE_\ell^g - \varepsilon_0\right) < E^g < \left(SE_\ell^g + \varepsilon_0\right) \quad if\ \ \ell_0 < \ell$$

the infinum of the set of $\left\{\ell \middle| \ell > \ell_0\right\}$ is the requested solution.

SUMMARY

- Relations
  - WGV $\longleftrightarrow$ Estimation stability/heterogeneity
  - BGV $\longleftrightarrow$ Regional/Geological heterogeneity
- Another criterion for assessing the goodness of simulation

$$WGV_\ell^g \underset{\ell\to\infty}{\Rightarrow} WGV^g < BGV_\ell^g \underset{\ell\to\infty}{\Rightarrow} BGV^g$$

- Criterion for the necessary and sufficient number of realizations

$$\left|SE_\ell^g - E^g\right| < \varepsilon_0,\ \left(\ell_0 < \ell\right)$$

$$\left(SE_\ell^g - \varepsilon_0\right) < E^g < \left(SE_\ell^g + \varepsilon_0\right) \quad if\ \ell_0 < \ell$$

MANY THANKS FOR YOUR ATTENTION