The influence of observation protocols on reported values of Rn in soil air

Peter Bossew
German Federal Office for Radiation Protection,
Köpenicker Allee 120-130, 10318 Berlin

11th INTERNATIONAL WORKSHOP on the GEOLOGICAL ASPECTS OF RADON RISK MAPPING

September 18th – 20th(22nd), 2012, Prague, Czech Republic

Content

1. Motivation
2. Soil Rn sampling
3. Numerical experiment
4. conclusions

Motivation, 1

- Often a quantity of interest cannot be observed directly.
  Observation = sampling design, sampling, measuring, evaluation.
- It must be estimated from quantities which can be observed.
- Examples from the Rn world:
  1) "long-term indoor Rn concentration":
  observed: concentration over a period, conditions ± controlled
  2) "Rn risk", defined as prob(indoor $C>c_0$):
  probability cannot be observed, but only estimated from data.
  3) "Rn concentration in soil air":
  quantity is under-defined in real situations.

Motivation, 2

- What do we do in practice?
- Apply assumptions:
  e.g. indoor-measurement over 1 year equals long-term concentration
- Apply models:
  - e.g. from indoor-measurement over 1 month, apply seasonal factor → estimate of long-term.
  - e.g. for risk: assume log-normal distribution of concentrations, estimate parameters from data and calculate prob from LN-model.
- Define protocol, which yields an *operational quantity*, which substitutes the ideal one.
  - e.g.: soil Rn: concentration in 1 m depth, sampled with a certain device, evaluated with certain rule.

Motivation, 3

- This presentation:
  observation of Rn concentration in soil air

- Appears particularly delicate:
  - most soils are not homogeneous horizontally and vertically
  ⇒ *what does "concentration in soil" mean?*

  - temporal variability due to meteorological and hydrological influence
  ⇒ *how to produce representative values?*

  - sampling procedure influences the quantity which is being sampled
  ⇒ *can one control this effect?*

Sampling soil air

- Basically 2 classes of methods:
- Active:
  borehole(s), probe, insulate against surface air, extract air (grab or continuous), measure, evaluate.
- Passive:
  borehole, bury passive detector (e.g. TE), close hole, expose for some period, recover, count tracks, evaluate.
- Here: discussion of two active methods

2 active methods

- for short: “DE” (Kemski et al.) and “CZ” (Neznal et al.) protocols
- why these? - very popular
- not discussed here:
  - temporal aspect: it has been shown that under most conditions (certain ones excluded) yield reproducible values for a location, i.e. “noise” introduced by temporal variability does not obscure the “signal”;
  - influence of method on quantity: ± understood qualitatively.
- In any case:
  The observation protocol actually defines the quantity in detail, which it intends to observe.

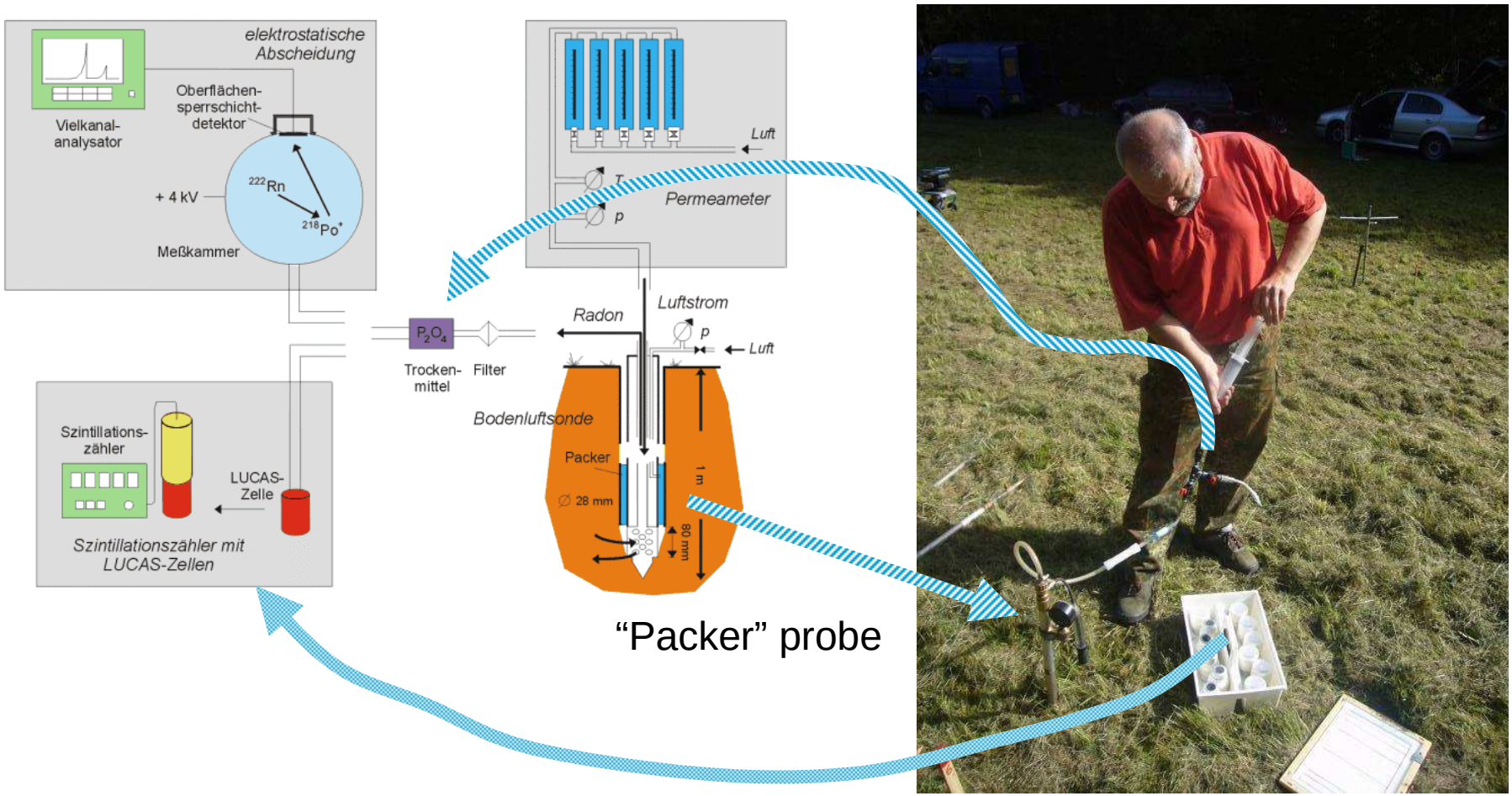
consequence

- just to repeat because it is so important:
  The observation protocol actually defines the quantity which it intends to observe.
- ⇒ Different protocols yield, in general, systematically (i.e. apart from observation uncertainty) different values of the nominal quantity “soil Rn”, under the same objective conditions.
- **⇒ Need to investigate these differences in order to be able to compare reported values of “soil Rn”.**

“DE” vs. “CZ” protocols

	“DE”	**“CZ”**
horizonal design	3 boreholes, equilateral triangle, ~5 m side.	≥ 15 holes, distributed over the investigation area (e.g. building site)
vertical design	standard 1 m deep, if <1 m: correction factor	standard 0.7-0.8 m deep
borehole	“Packer” probe, inserted into drilled hole	“lost tip” probe makes the hole
air sampling	2 samples per hole	1 sample
measurement	Lucas cell; m = 2 - 3 measurements per sample; AM of 2·m measurements per hole	Lucas cell
evaluation	“soil Rn” = maximum of the AM’s of the 3 holes.	“soil Rn” = 75%-quantile of the holes. Values <1 kBq/m³ excluded.

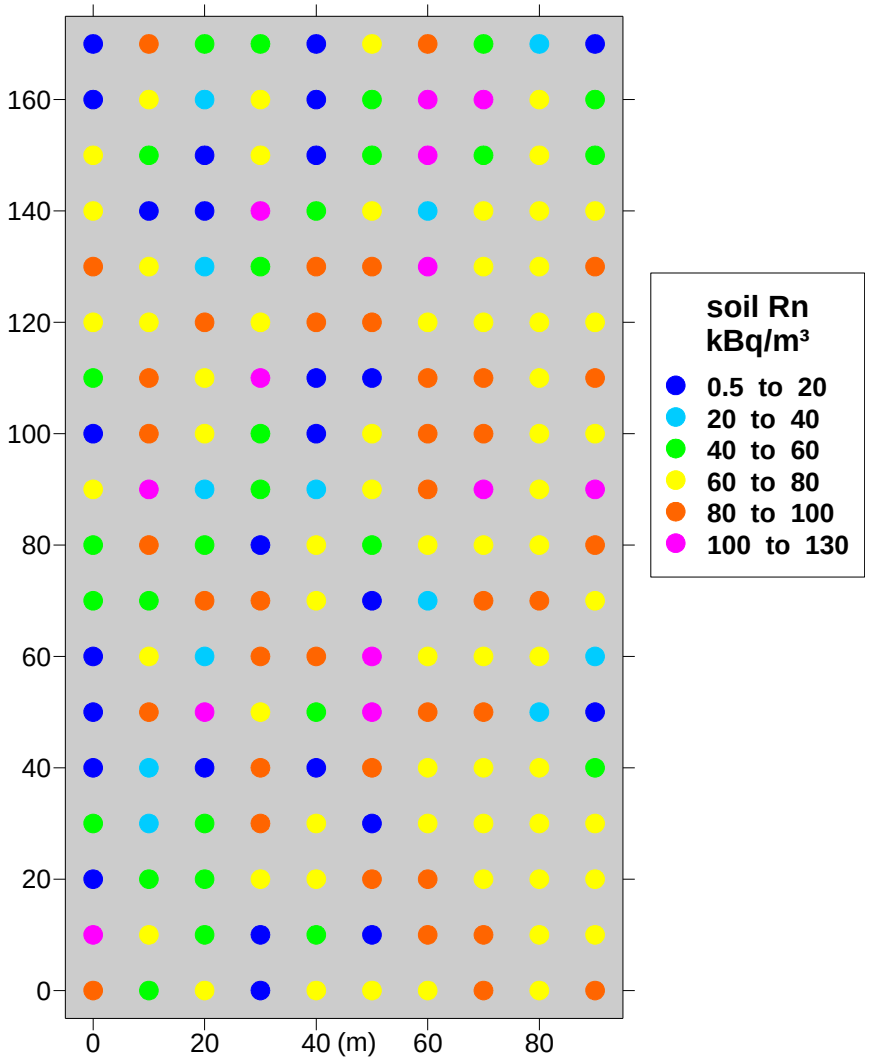
how it looks in practice



“Packer” probe

German (DE) protocol

simulation study: 1) data

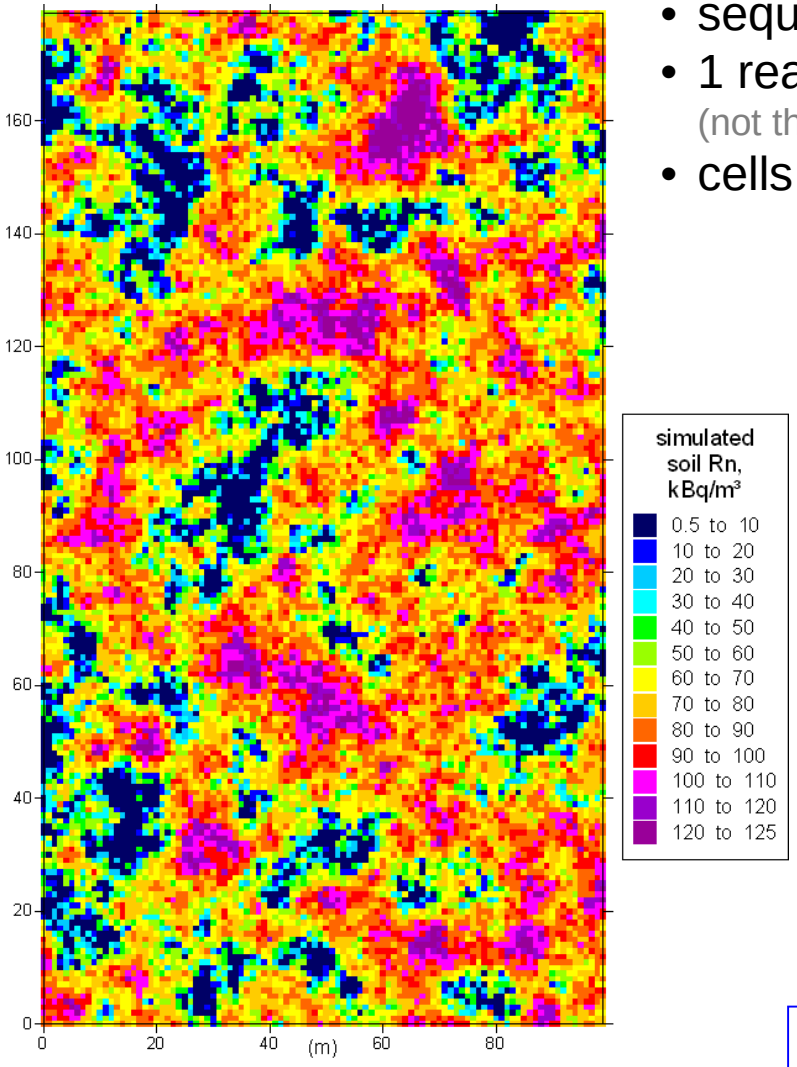


- data: real values, provided by M. & M. Neznal;
- Observation: “CZ” protocol;

- These data taken as “true” values of the quantity “Rn conc. in soil”
- Construct “true” Rn field;

- Simulate “DE” and “CZ” protocol on this “true” field;
- Compare results.

2) simulated “true” field



small-scale structure strongly depends on variogram model!

- sequential simulation based on variogram;
- 1 realization chosen
  (not the mean, in order to preserve the strong variability);
- cells 1m x 1m.

basic statistics:

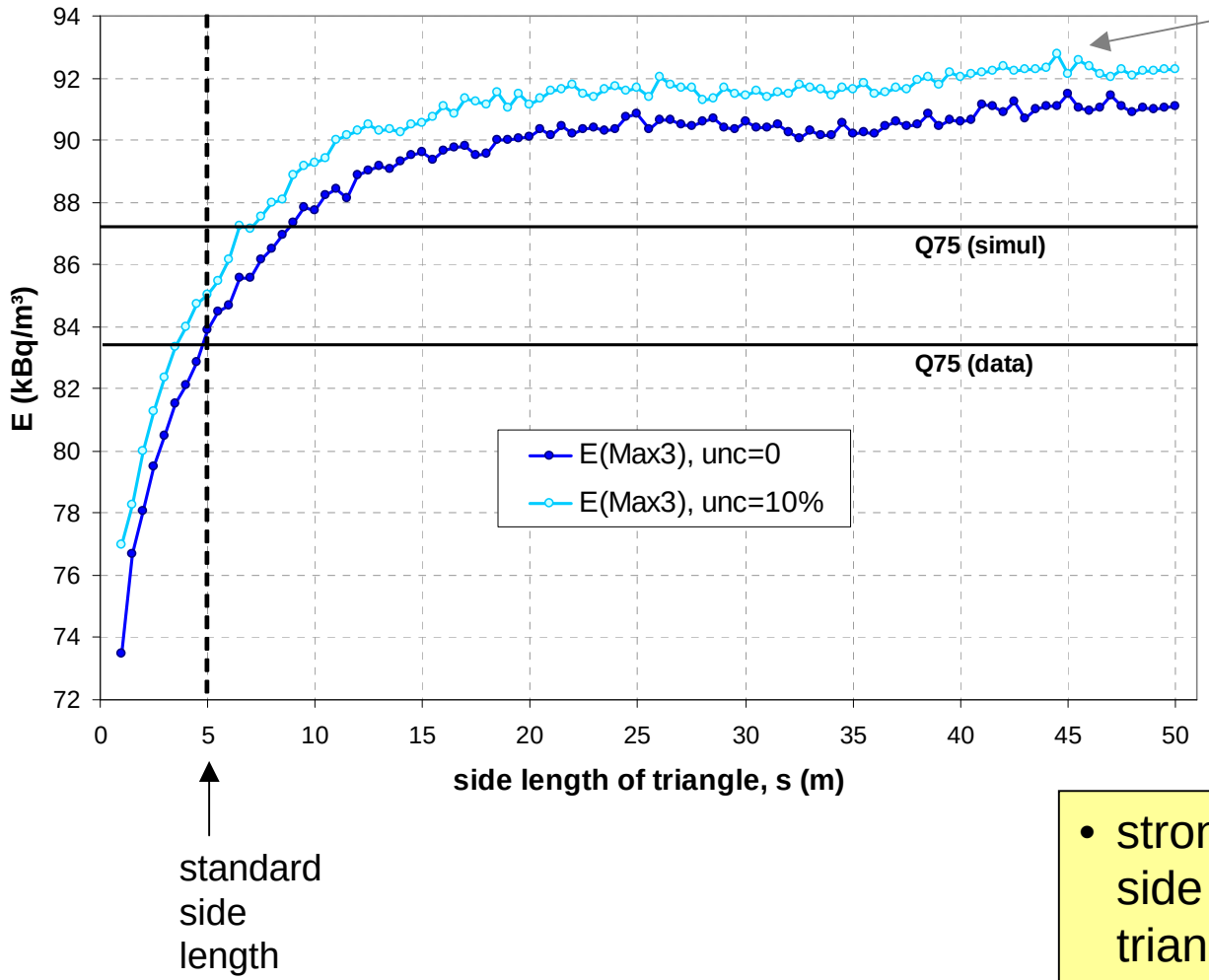
statistic	original data	simulation
n	180	18000
AM (kBq/m³)	62.26	66.64
SD (kBq/m³)	29.80	30.02
CV	48%	45%
Min (kBq/m³)	0.5	0.5
Max (kBq/m³)	122.1	122.1
Med (kBq/m³)	69.1	73.4
Q75 (kBq/m³)	83.2	86.5
Q75 (Z>1 only) (kBq/m³)	83.4	87.2
GM (kBq/m³)	44.4	48.1
GSD	3.26	3.40

Czech classification: **Rn class 2** (Q75 between 30 and 100 kBq/m³ for low permeable soil).

3) numerical Rn sampling

- **“DE” protocol:**
  - 10000 equilateral triangles, side=s, located randomly over the field;
  - Gaussian noise 0%; 10% assumed;
  - E(max3) = expectation acc. DE protocol;
  - prob(wrong) = probability of wrong classification (true class = 2);
  - prob(class 1, 3) = probability that the protocol wrongly classifies as class 1 or class 3.
- **“CZ” protocol:**
  - n≥15 random samples taken, 2000 realizations;
  - within cluster of radius r.
  - Gaussian noise 0%; 10% assumed;
  - E(Q75) = expectation acc. CZ protocol;
  - prob(wrong) etc. as above;
  - dev := 100·(E(Q75)-Q75true)/Q75true, deviation from the “true” Q75 = the one of the simulated field = 87.2.
- Effects of different sampling depth and of different extracted soil air volumes not considered!

result, DE protocol, 1

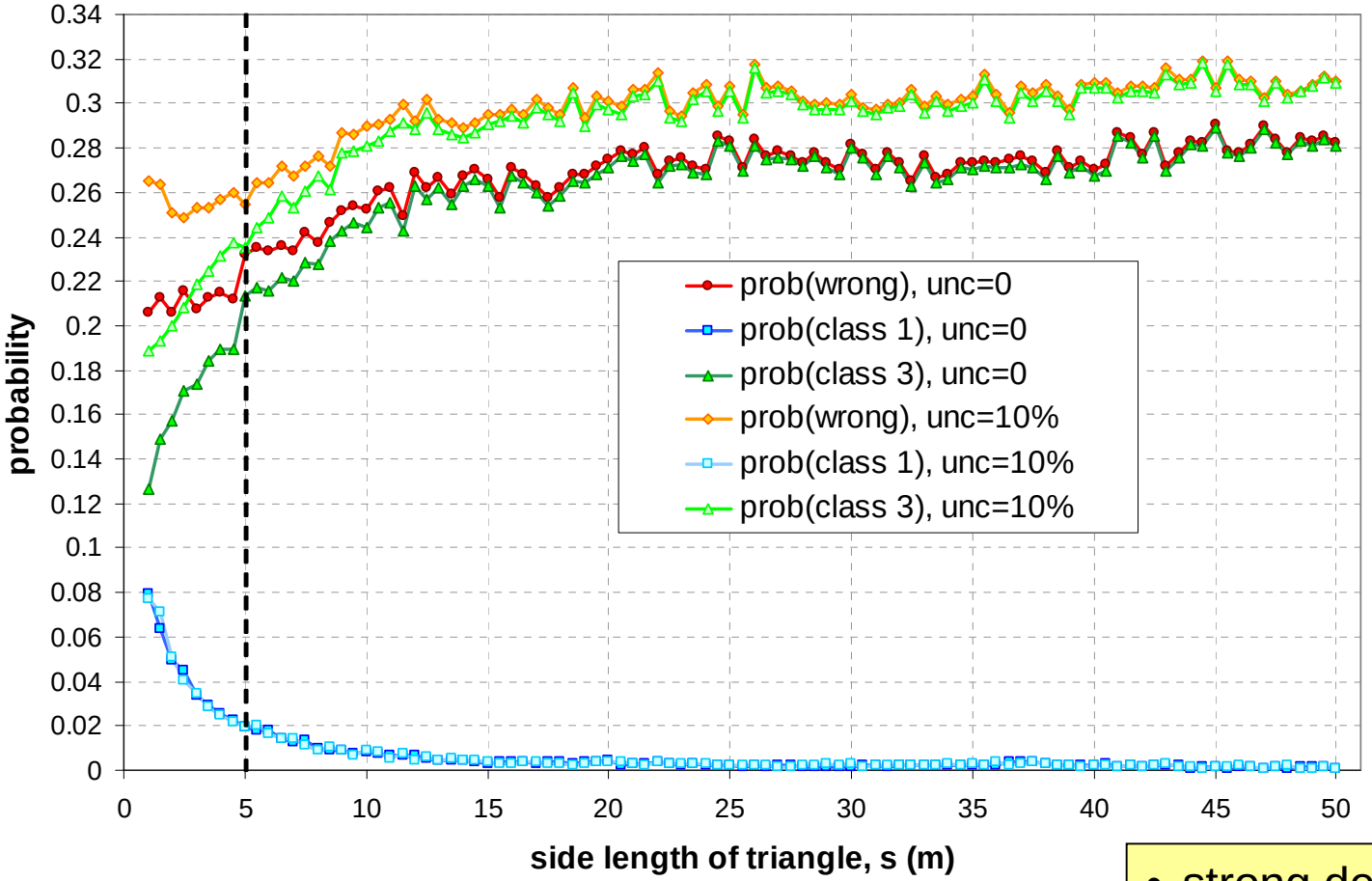


fluctuations due to stochastic character of simulation

E(Max3) for s=5 m between Q75(true) and Q75(data)

- strong dependence on side length (s) of the triangle !
- dependence on assumed uncertainty.

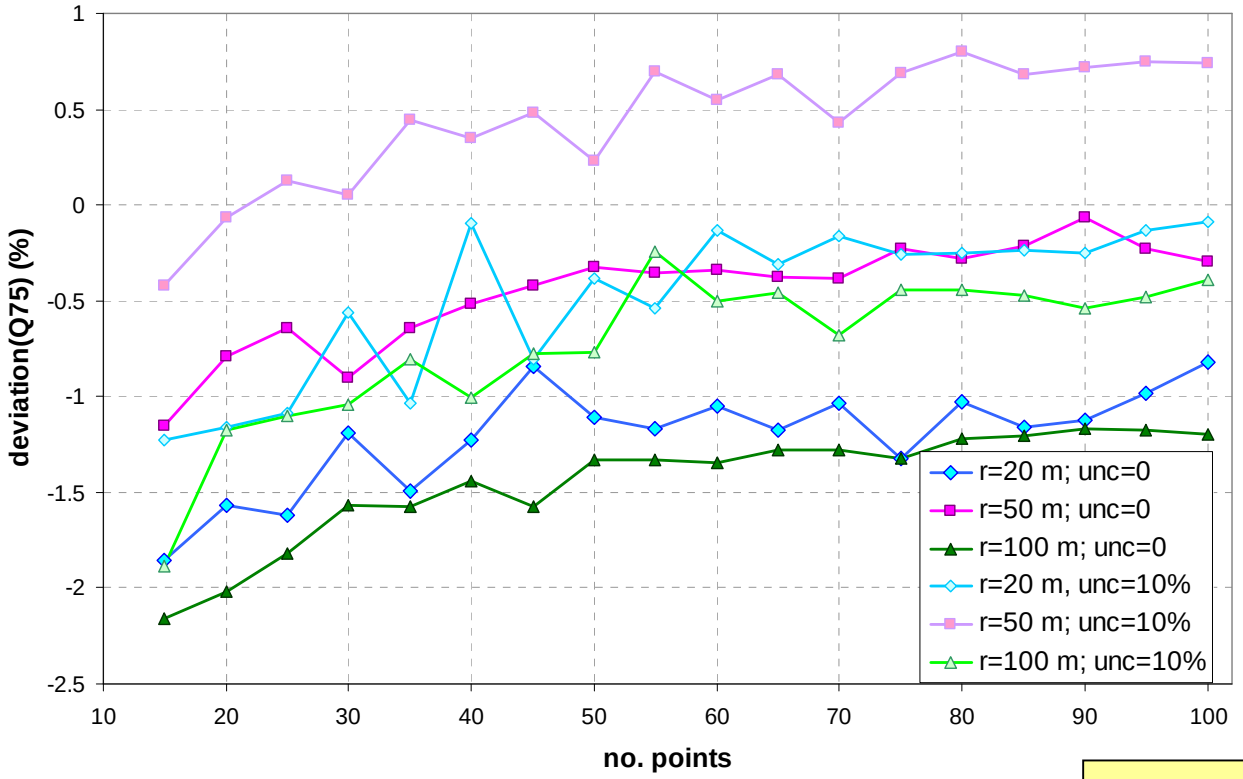
result, DE protocol, 2



class 2 definition: 30 < Q75 < 100
since Q75(true) = 87.2 = closer to 100 than to 30:
mis-classification into class 3 more likely than into class 1.

- strong dependence on side length (s) of the triangle !
- dependence on assumed uncertainty.

result, CZ protocol



very low under-estimation of true Q75

- dependence on number of samples (n)
- dependence on size of the cluster (r)
- dependence on assumed uncertainty.

comparison of the methods

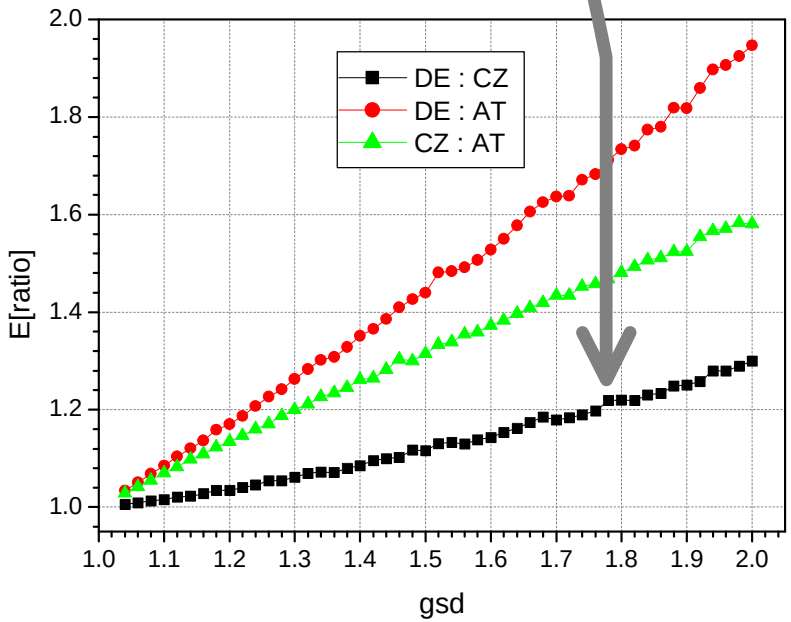
protocol	**value (kBq/m³)**	
true values: AM	**66.6** ± 30.0	
true values: Q75(Z>1)	**87.2**	
measurement unc.:	unc=0	unc=10%
DE (s= 5m), E(Max3)	**83.9** ± 20.4, p(wrong)=23%	**85.0** ± 22.0, p(wrong)=25%
CZ; 15 points, r=20m, E(Q75)	**85.6** ± 8.6, p(wrong)=4.5%	**86.1** ± 9.4, p(wrong)=5.9%
CZ; 15 points, r=50 m	**86.2** ± 6.5, p(wrong)=2.6%	**86.8** ± 7.3, p(wrong)=4.0%
CZ, 15 points. r=100 m	**85.3** ± 6.2, p(wrong)=1.2%	**85.6** ± 7.3, p(wrong)=2.3%
CZ, 100 points, r=100 m	**86.2** ± 2.4, p(wrong)<0.05%	**86.9** ± 2.9, p(wrong)<0.05%
CZ, original sample: r=total domain, 180 points	83.4	

- all very similar to “true” Q75 !
- all > AM(true) … due to definition of Max3 and Q75 → conservative!

p(wrong)(CZ) < p(wrong)(DE) but DE-protocol does not pretend to estimate Q75 and class acc. CZ scheme.

conclusion & caveat & to-do, 1

- Max3 (DE-protocol) and Q75 (CZ-protocol): remarkably similar results!
- depth correction: DE value x 0.93 (low permeable loamy soil, acc. Kemski 2002)
- previous simpler numerical study (EGRM-report v.1, sec. 4.5.2.4): samples taken from marginally LN population, spatial structure not considered: difference DE - CZ protocols depends strongly on GSD of true field.

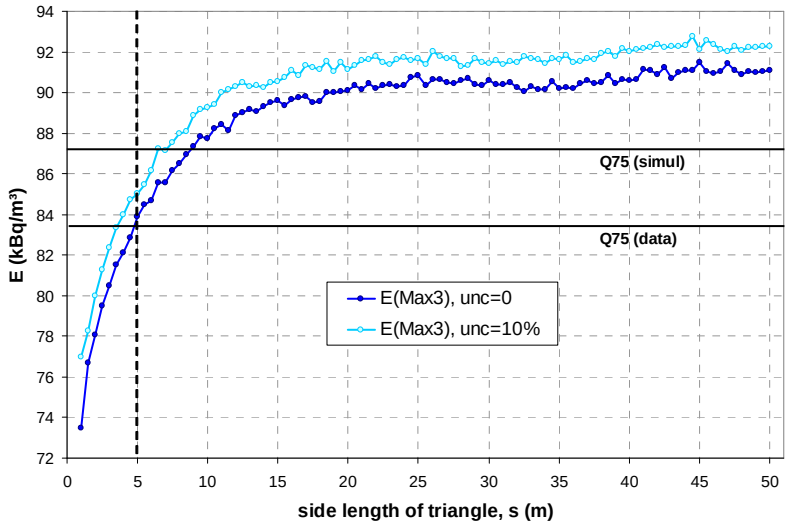


empirical study, Barnet et al. 2010: DE/CZ ≈ 1.14

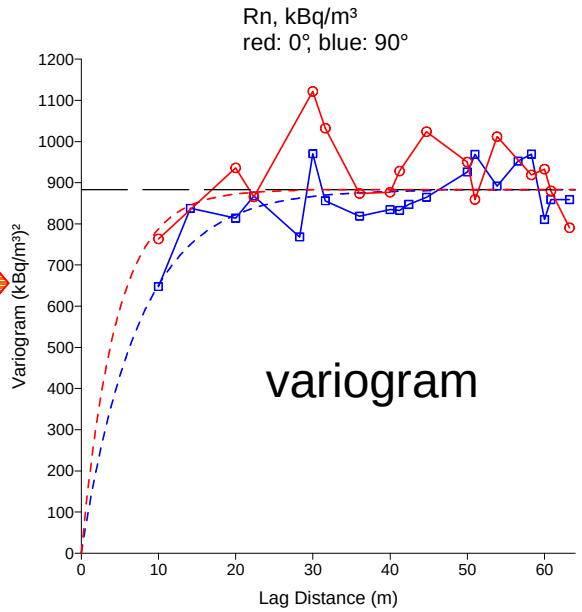
conclusion & caveat & to-do, 2

very suspicious: result appears to depend on spatial structure !!

Maybe result typical for the assumed true Rn field ??



!!!



⇒ to do:

investigate dependence on spatial structure of the true field
by assuming different variogram models → many simulations
→ quite heavy numerical work – was not possible within this study!

Thank you !