The geogenic radon hazard index – another attempt

Bossew P.[1], Cinelli G.[2], Tollefsen T.[2], DeCort M.[2]

1 German Federal Office for Radiation Protection (BfS), Berlin

2 European Commission, Joint Research Centre (JRC), Directorate for Nuclear Safety and Security, Ispra, Italy

v.3.11.17b

Bundesamt für Strahlenschutz

European Commission

IWEANR 2017

2nd International Workshop on the European Atlas of Natural Radiation

Verbania, Italy, 6 – 9 Nov 2017

Content

- GRHI: rationale & objective
- Concept → definition
- Properties of the GRHI
- Previous attempts
- Case study
- Challenges

The idea of the Geogenic Radon Hazard Index GRHI

A quantity which measures the availability of geogenic Rn at surface level.

Ideally: Geogenic Radon Potential GRP (e.g. Neznal definition); but: available only regionally - CZ, DE, BE, (IT), (ES), (AT), ?

MetroRn
WP 3.2

Other geogenic quantities may be available:

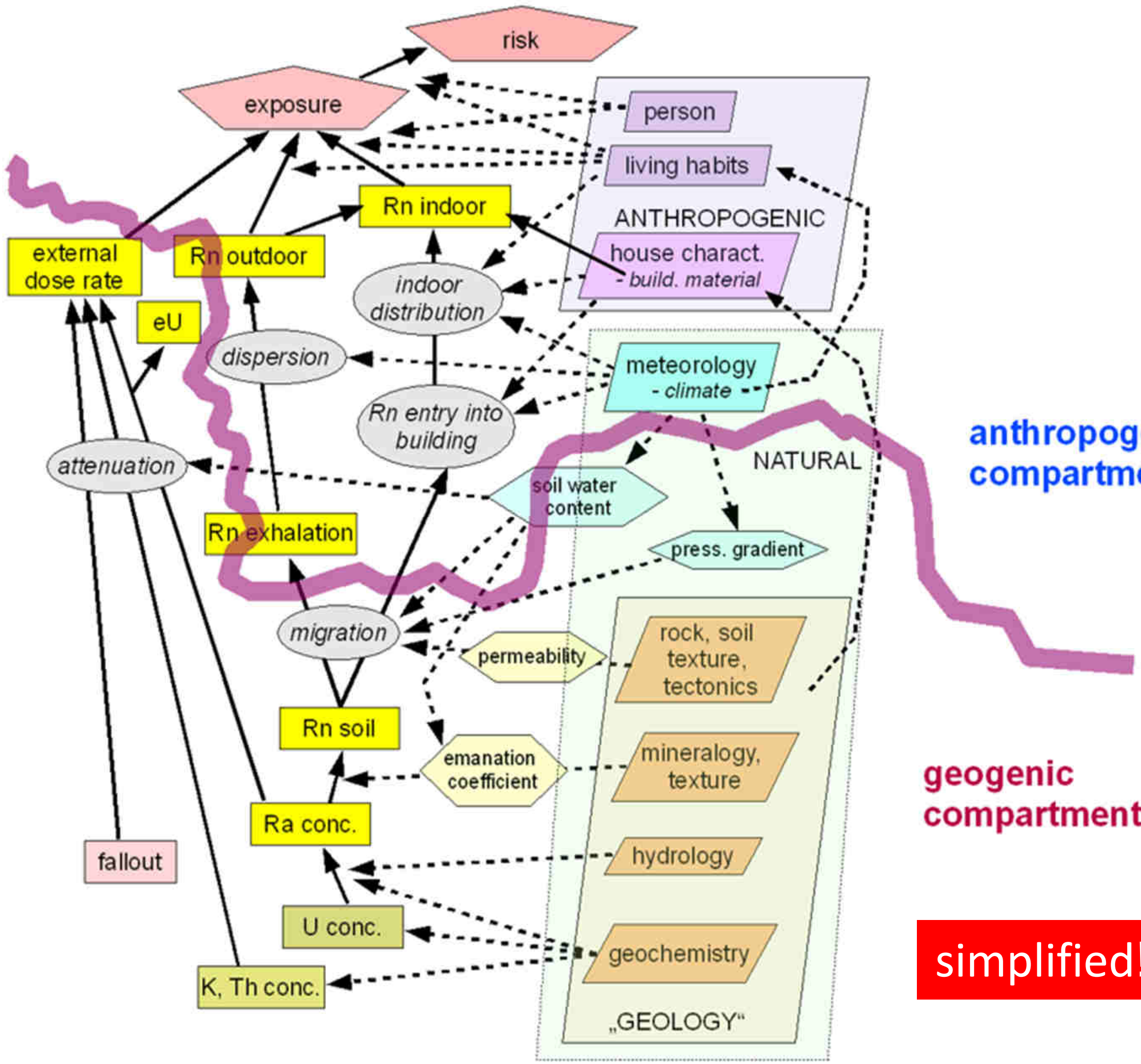
- U concentration,
- ambient dose rate ADR,
- geological units / lithology,
- fault density,
- groundwater recharge coefficient,
- soil properties,
- permeability of the ground, karstification,
- standardized indoor Rn concentration.

GRHI =
measure of "Rn proneness" of an area due to geogenic factors.

Role of MetroRadon

- Development of the GRHI is one of the objectives of MetroRn! (WP 4.3.4)
- Harmonization of geogenic Rn quantification across Europe (~ WP 3.2)
- Possibly harmonized Rn priority areas (delicate subject!) (WP 4.4)

Reminder: Rn - From rock to risk



Radon – a complex system

anthropogenic compartment

geogenic compartment

simplified!

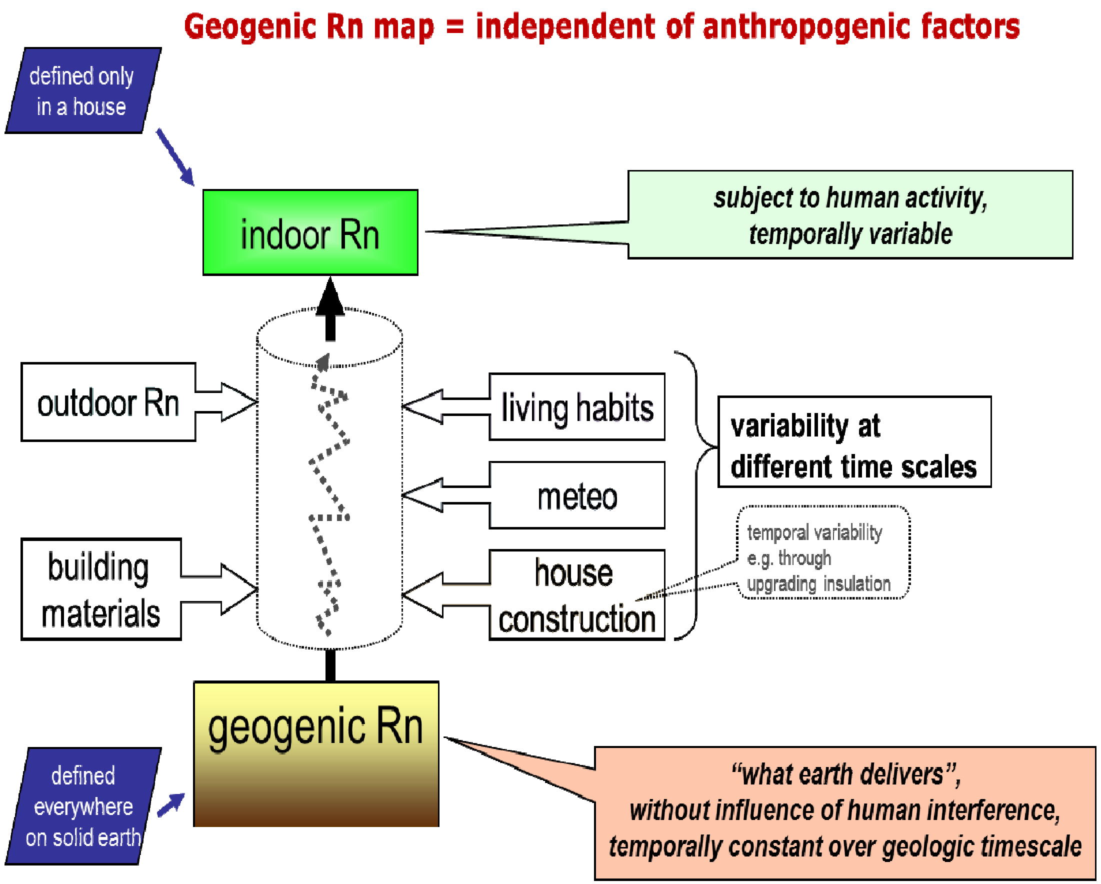
Often factors are
- by themselves heterogeneous
- interact in complicated way, sometimes not well known

Result: complicated dependence of Rn quantities.

Further,
often factors are
- fuzzy or ill defined;
- not well known;

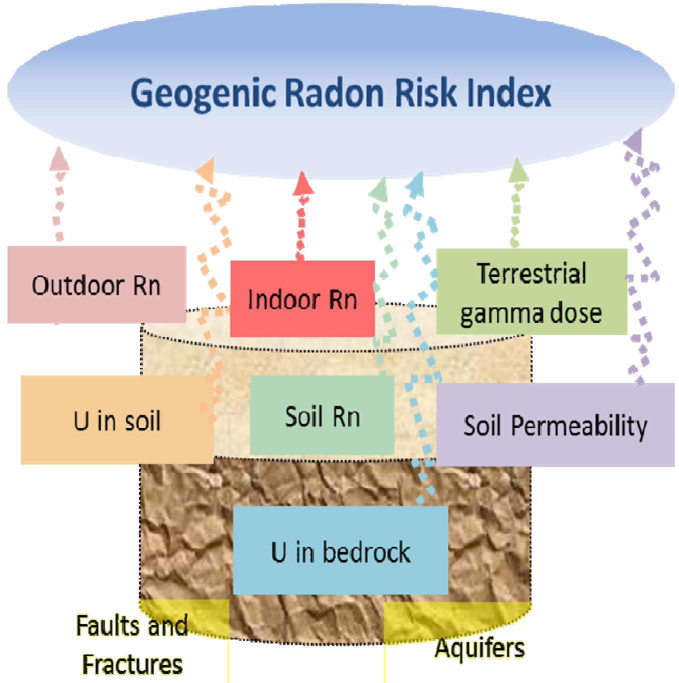
Result: difficult to understand the source of variability

The geogenic radon potential

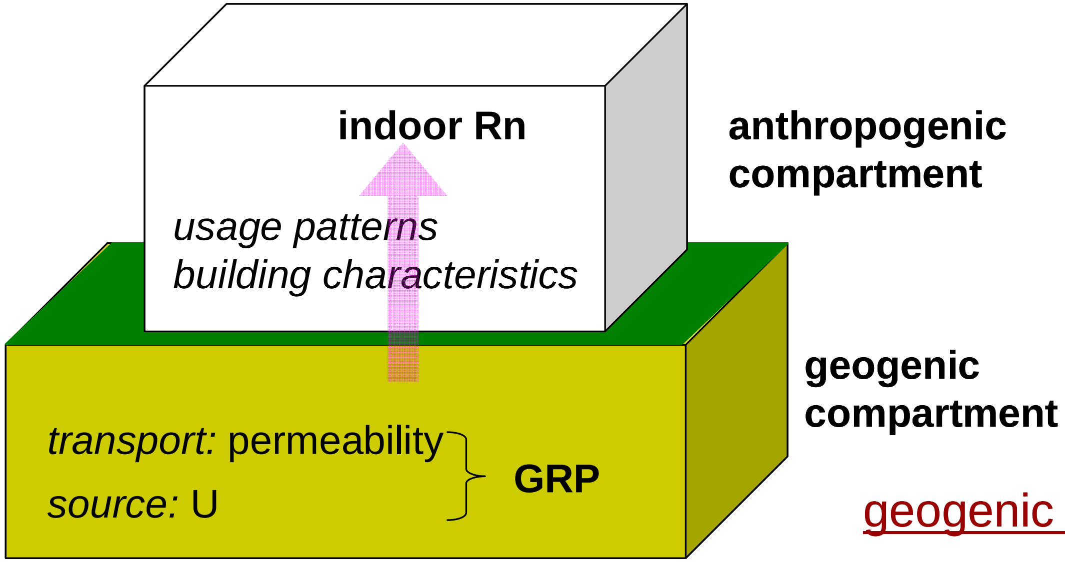


Wanted:

Multivariate definition of Geogenic Radon Risk Index

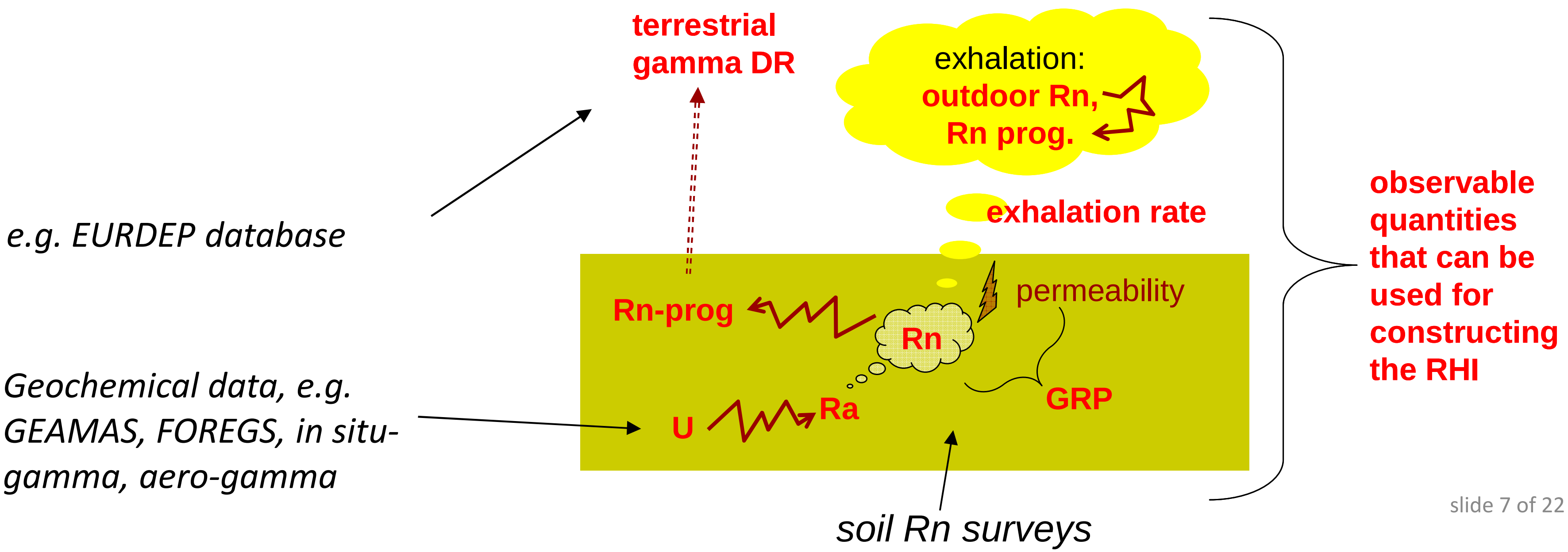


geogenic and anthropogenic compartments

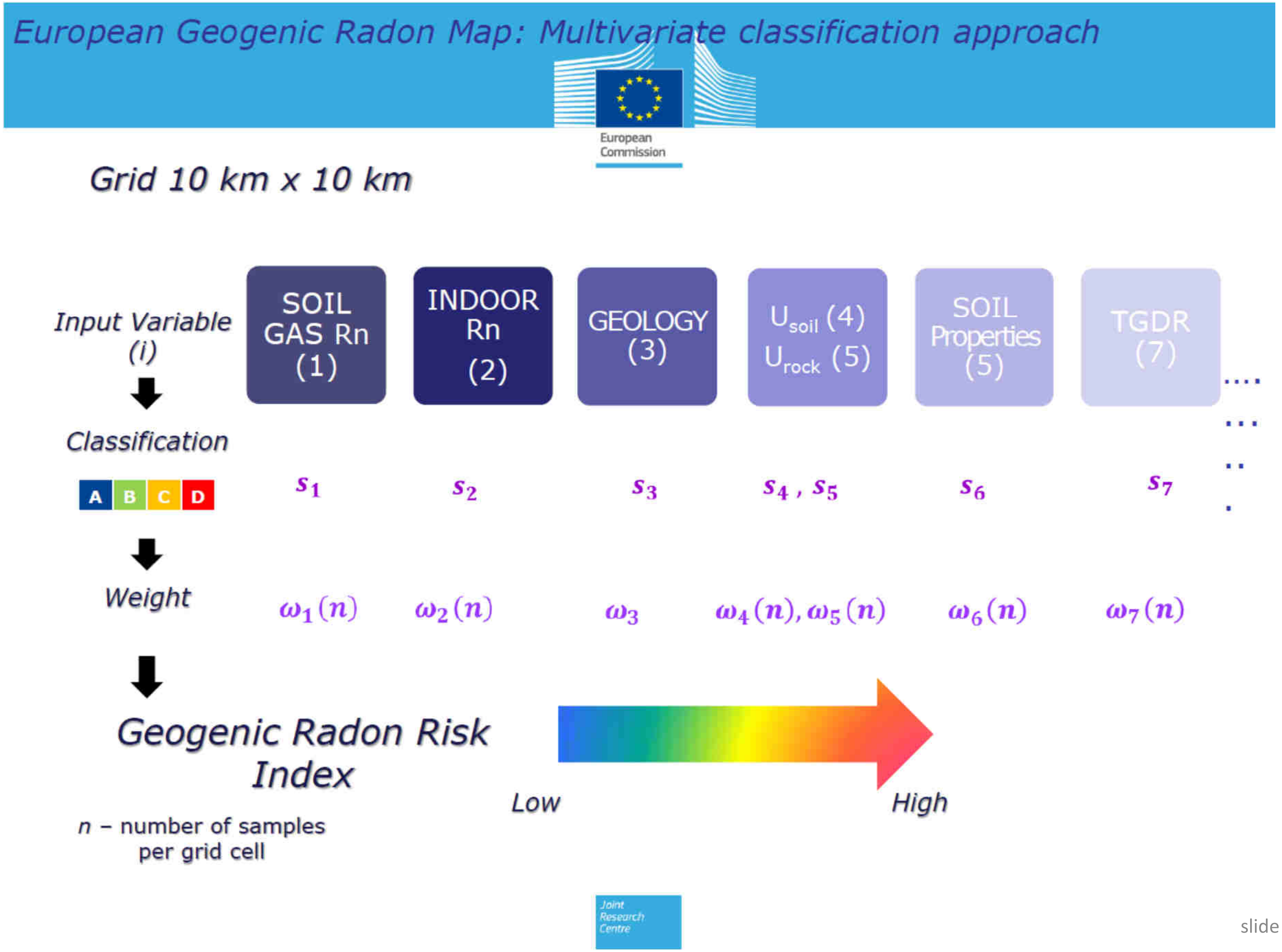


- The GRP quantifies availability of Rn for infiltration
- Anthropogenic factors determine, to which extent available geogenic Rn leads to indoor Rn concentration… “infiltration and accumulation potential”

geogenic quantities



Initial idea (Cinelli et al. 2015)



Properties of the GRHI

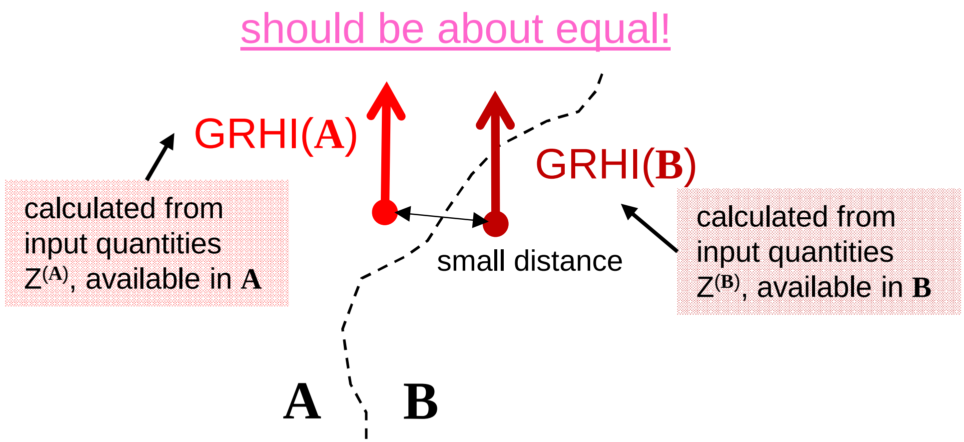
- Consistency: see next slide
- should include as much information as possible
- should be flexible, i.e. to be applied to as many different situations as possible
- should be simple to calculate!

consistency, 1

Its value at a location must be independent on which quantities it has been estimated from.

I.e., GRHI calculated from U concentration in soil should have approximately the same value as if calculated from dose rate or GRP, etc.

This follows from the requirement to be consistent across borders, or regions in which different input quantities are available.



consistency, 2

Given input quantities (U, DR, geol. class).
Then should be:
GRHI(U,.,.) ≅ GRHI(.,DR,.) ≅ GRHI(U,.,Geo) ≅ GRHI(U,DR,Geo) etc.
≅ means "up to deviations which are due to the imperfect correlation between geogenic quantities & statistical uncertainty"

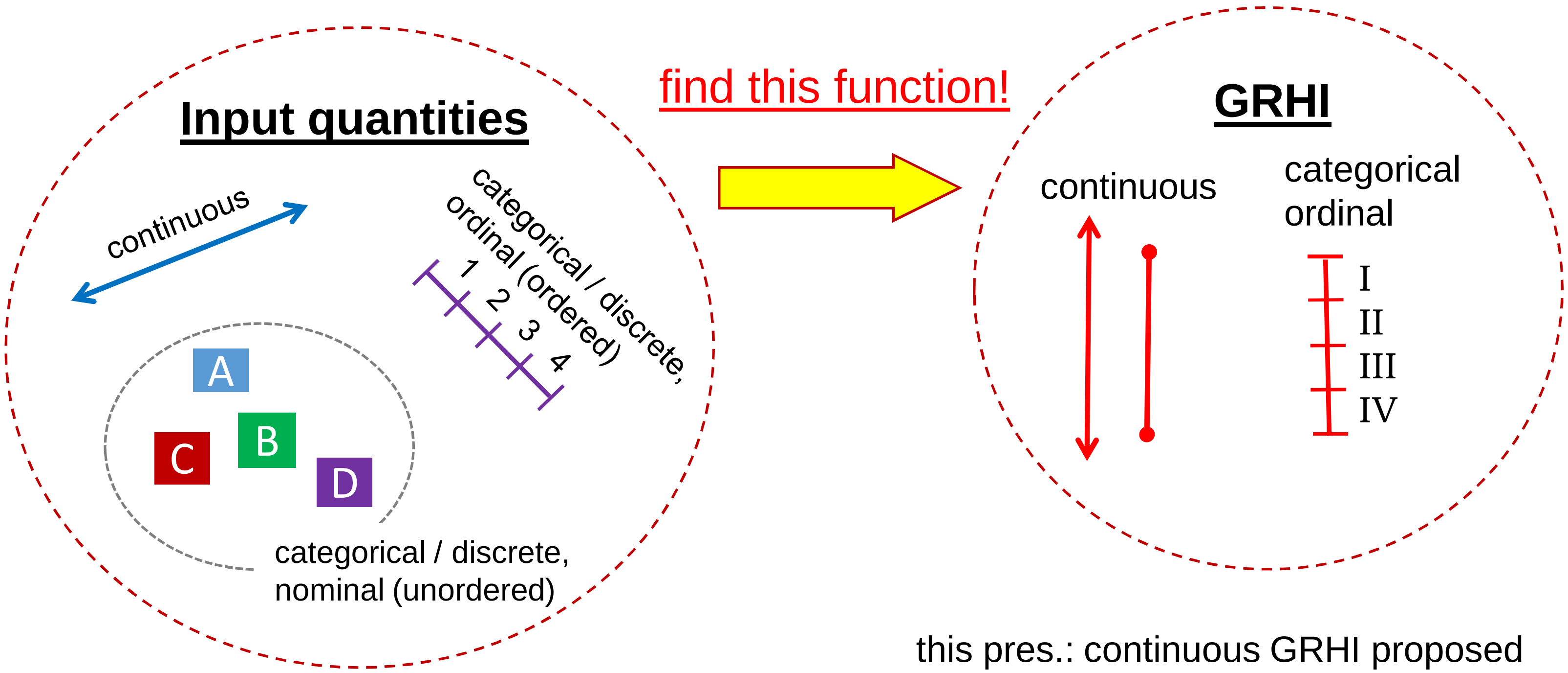
or: $E[GRHI_1 - GRHI_2]=0$

Why?
Because it shall be applicable independent of the input quantities in a region.
This is the most difficult condition!

Different concepts

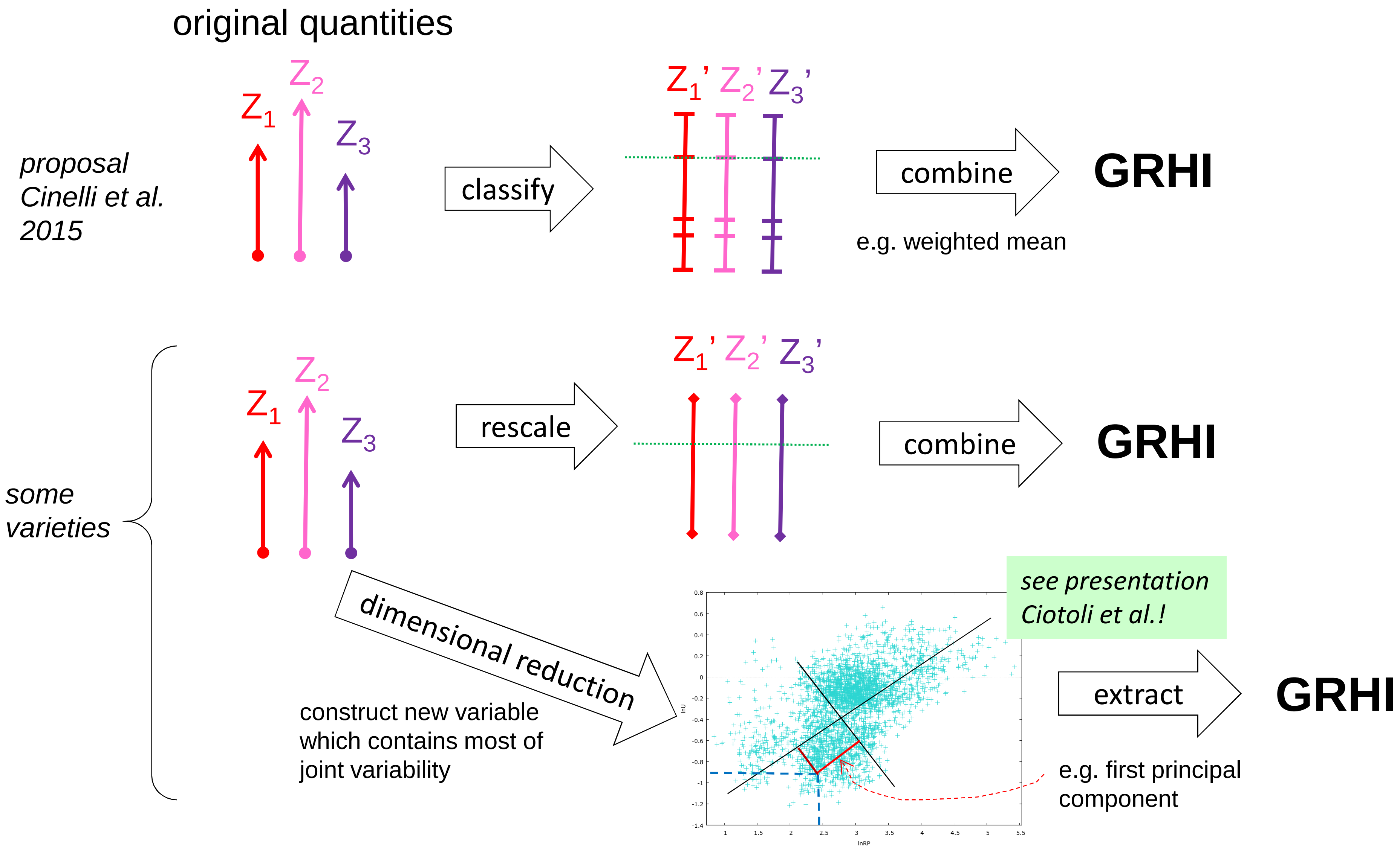
Geogenic Rn hazard index GRHI can be:

- <u>continuous</u> index, e.g. ∈ [0,1] or (-∞,∞) etc.
- <u>discrete</u> index or score, e.g. ∈ {I,II,III,IV} or {low, medium, high} etc.



this pres.: continuous GRHI proposed

some options



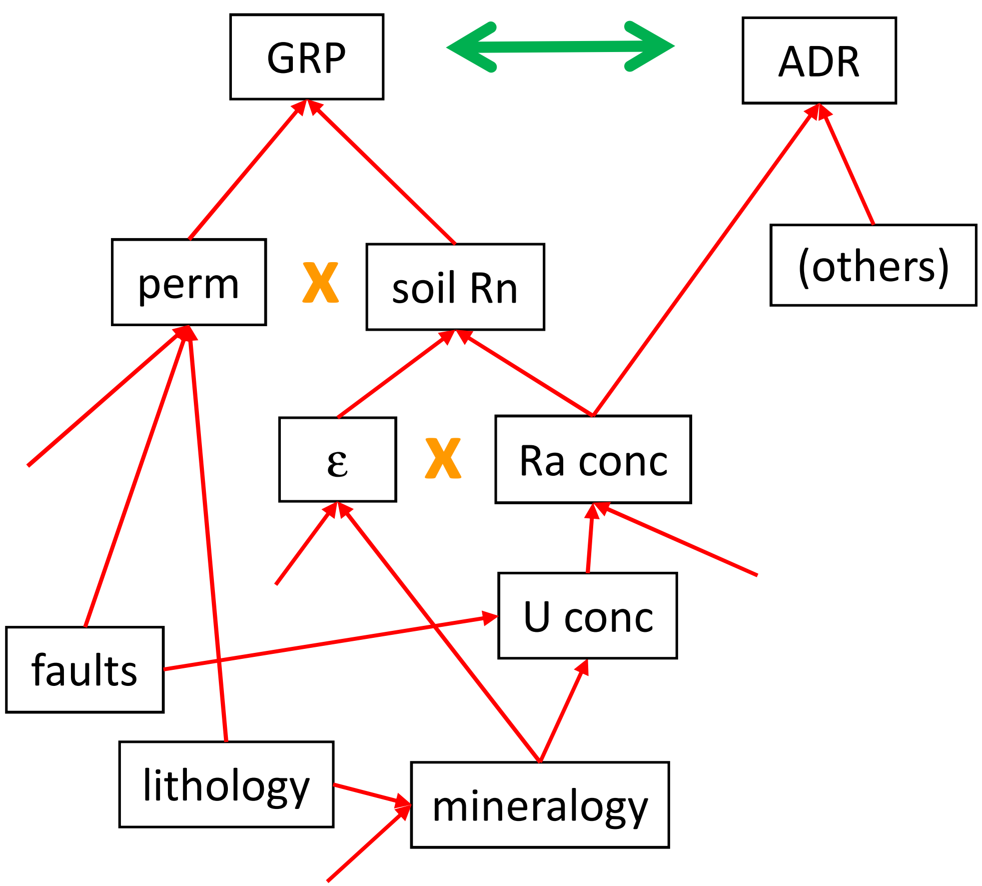
Previous attempts

- TREICEP-5, Veszprém 2016:
  - transformed variables
  - options: GRHI constructed such that
    (a) covariates considered as proxies or predictors of GRP; or
    (b) covariates should best predict indoor Rn
  - weights:
    (1) through correlations between variables;
    (2) loadings of 1. principal component
  - performance of GRHI assessed as RPA predictor, DE data

- GARRM-13, Prague 2016:
  - 3 “families” of methods:
    ‘F’: GRHI=mean of distribution functions of covariates;
    ‘R’: GRHI=mean of GRP predicted by covariates through regression;
    ‘P’: 1.PC, as above.
  - performance of RHI assessed as predictor of indoor Rn exceedance probability, DE data;
    no convincing advantage of any method

- TEERAS, Sofia 2017:
  - Case study Cantabria:
    covariates: soil Rn, GDR, fault density, U in soil, lithology, permeability, karstification
  - weights: correlation with indoor Rn; GDR and U excluded
  - 3 “hazard classes”: if prob(C>300), estimated from GRHI, >0.1 → high;
    if prob(C>100)<0.1 → low; otherwise medium.
  - Performance through underestimation rate (2.kind error): 7%

Predictors and proxies or surrogates

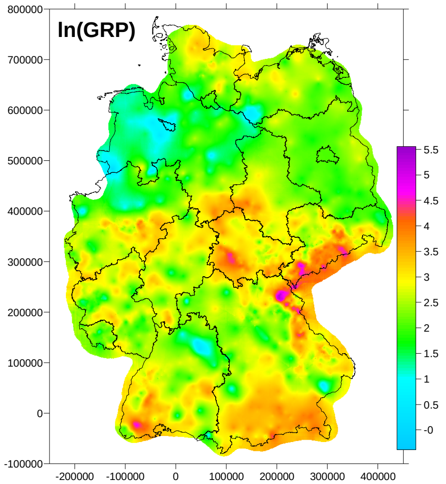


- ADR is proxy to GRP: no physical causal, but **statistical** relationship
- Red arrows: **physical** causality: predictors or controls; direct or indirect
- X : no identifiable relationship, perhaps because other controlling factors are dominant

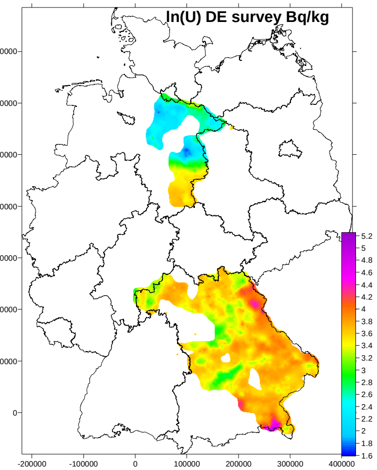
- GRHI candidate covariates are predictors or proxies to the GRP;
- The stronger the statistical relationship, the better!

Case study: covariates

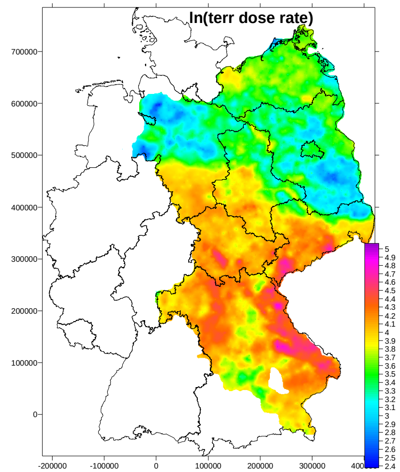
German data



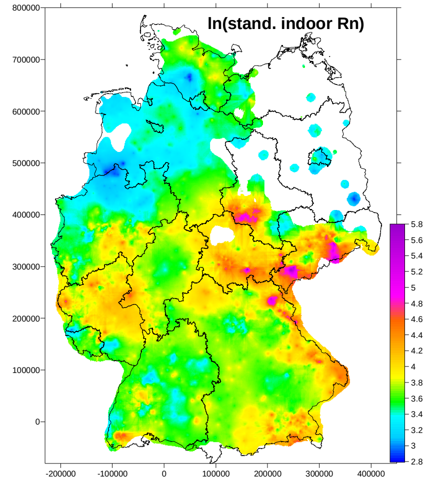
GRP, n=4728



U (Bq/kg), survey, n=2194

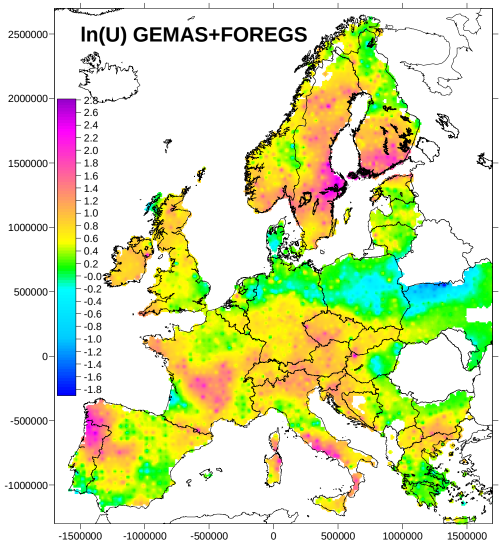


terr. DR (nSv/h), survey, n=10,931

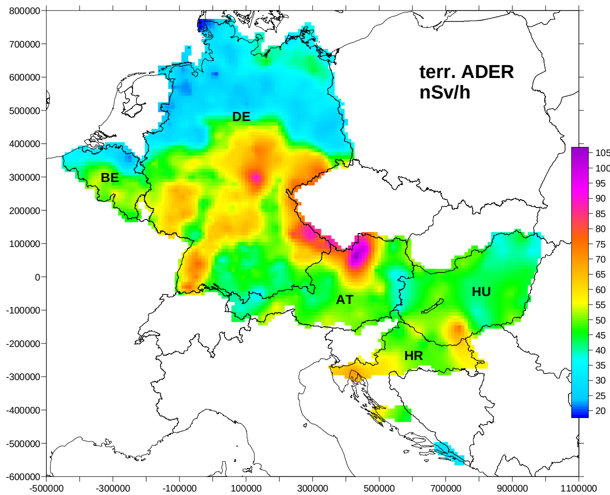


stand. indoor Rn (Bq/m³), n=39,809

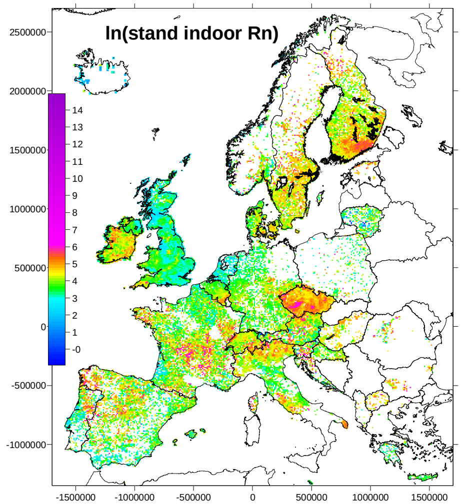
Atlas data



U (ppm), GEMAS-FOREGS, n=4970



terr. DR (nSv/h), from EURDEP stations, n=1343



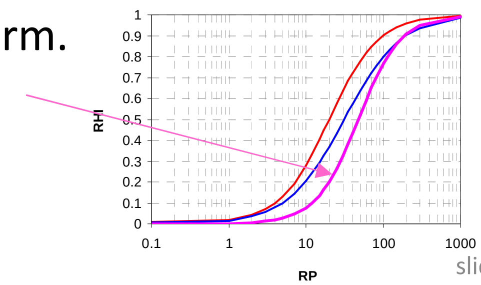
stand. indoor Rn (Bq/m³), n=24,269

Approach

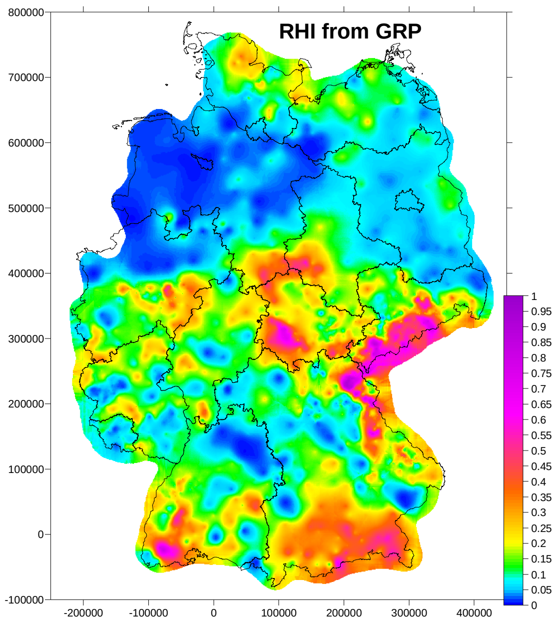
- Understand the GRP as “best” realization of the GRHI at a location.
- For all covariates Y^i (e.g. DR, U, stand. indoor, geology,…): establish all possible functional dependencies
  $GRP = f(Y^i)$, $GRP = f(Y^i, Y^j)$, … (“transfer models”)
  method: estimate Y^i at locations of GRP, in regions where GRP and Y^i are available. Where possible, the f should be regionally determined, otherwise generic.

MetroRn
WP 4.2.1

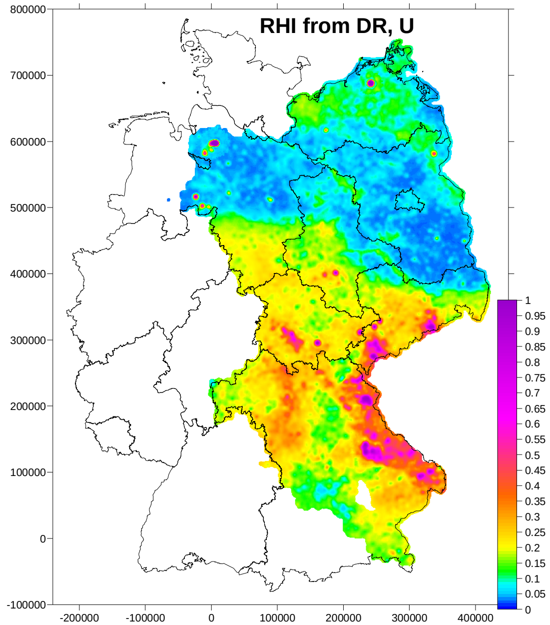
- At locations x where Y^i, Y^j,… are available (data $y^i(x)$, $y^j(x)$,..):
  Calculate $GRP^*(x)=f(y^i(x))$, $f(y^i(x), y^j(x))$,…
- Merge datasets of GRP and GRP*, whichever available, and use for mapping.
- Technicality: Transform GRP to GRHI $\in$ [0,1), by tgh transform.
  Here: so that GRHI(GRP=20)=0.2 and GRHI(GRP=300)=0.95



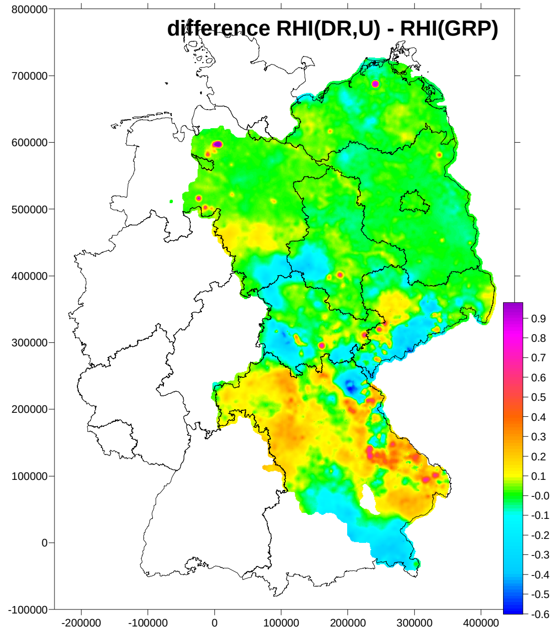
Example 1



“truth”



estimate



error

AM[diff]=0.015

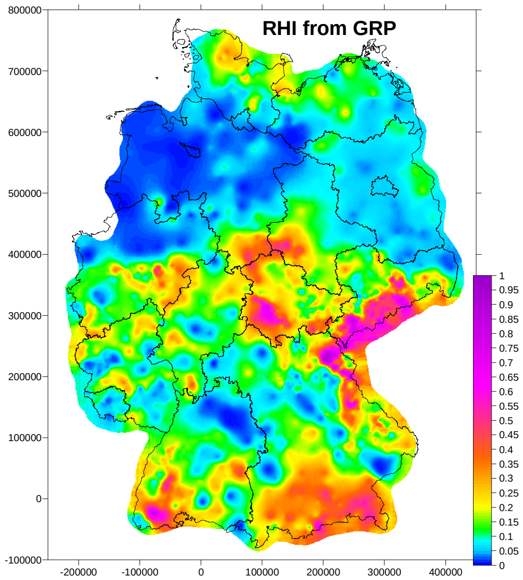
ln(GRP)=poly(ln(DR))

ln(GRP)=poly(ln(U),ln(DR))

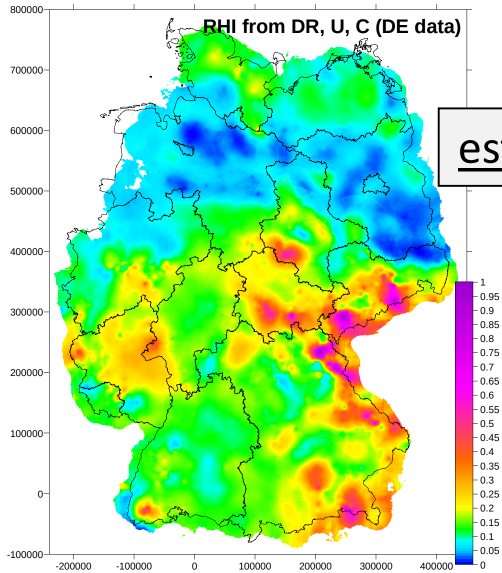
- coefficients found by multiple regression and backward selection
- no physical base of the model!

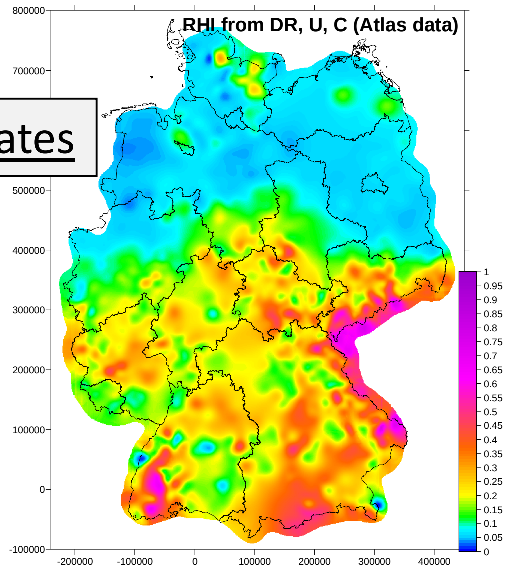
evidently errors are not random, but have regional trend. Why…?
Violates consistency requirement!

Example 2

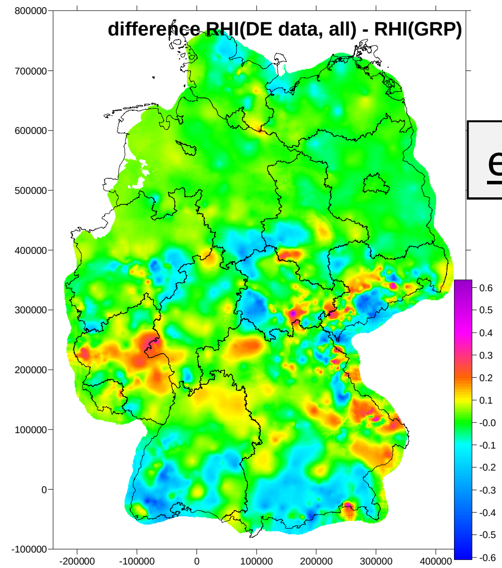


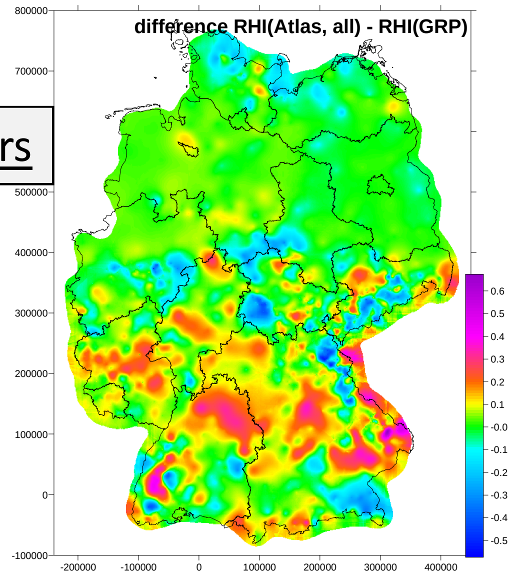
“truth”





estimates





errors

AM[diff]=0.0017

AM[diff]=0.034

why?

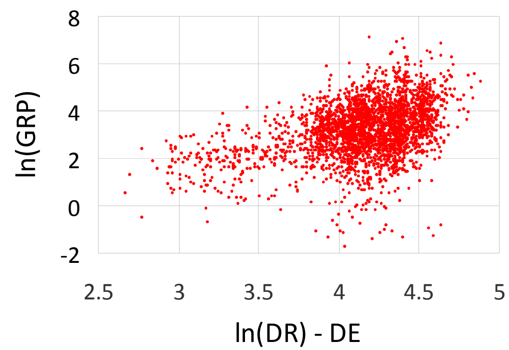
<u>Observation:</u>

- AM[diff] should be =0; in reality ≠0, but quite low → no high bias.
- Most unpleasant: spatial trends of the errors!

<u>Possible sources of the errors</u>

1. Data (value and location) uncertainty: would lead to randomly distributed errors.
2. Predictors & proxies do not allow perfect reproduction of the GRP because important control factors are missing. (See “rock to risk”!) I.e., models are incomplete.
   if these missing factors are regionally differently important ⇒ error has geographical trend.
3. Transfer models (by regression) are uncertain:
   - a) unc. of model structure,
   - b) unc. of estimated parameters;
   - c) residual error.

   (a+c) partly related to 2.



	lnDR-DE	lnU-DE	lnC-DE	lnDR-Atlas	lnU-Atlas	lnC-Atlas
lnGRP	0.37	0.46	0.42	0.35	0.36	0.38
lnDR-DE		0.68	0.47	0.73	0.79	0.34
lnU-DE			0.66	0.50	0.54	0.64
lnC-DE				0.60	0.50	0.82
lnDR-Atlas					0.76	0.55
lnU-Atlas						0.46

Spearman-r

Conclusions & *to-do*

- Idea of GRHI is relatively simple
- Different ways of defining it from predictors or proxies
- Main problem:
  poor correlation between GRP and candidates for covariates
- Dependence structure (and correlation) is regionally variable;
  how to parametrize this while staying simple?
- Here: GRP predicted from covariates, model determined by regression
- Works moderately well, local errors to be expected!
- GRHI classes (see Cantabria study, TEERAS 2017):
  how to define class limits; classification errors?
- *To do: exercises with regional datasets;*
  include more predictors and proxies!

Thank you!