Minutes: VALUE General Assembly Athens

Renate Wilcke

24.09.2014

Douglas:
Outline:

- Overview CORDEX ESD (Rasmus, Douglas, José)
- Overview NCPP Downscaling Metadata Project
- Augsburg Meeting
- Rome Meeting (Rasmus, validation controlled vocabulary, downscaling glossary)
- VALUE framework paper

Scope of meeting:

- to be approved:
  - selection of data
  - experimental framework including the protocol and terms and conditions
  - the portal functions
- planing further meetings
- to discuss the outreach, role and future of VALUE

WG2 report (José)

- observation stations 786, about 5 per country
- !TODO: Sweden - 5 obs stations to be suggested 1979-2008 (Thomas, Fredrik and Renate) comments about represented climate of the stations (freely available) (trying to cover all relevant variables)
- Poland, France, Norway, Italy, Greece

VALUE Terms and Conditions for validation portal (Martin)

uploading:

- non-public mode:
  - version control, each upload will be logged. Logs are visible to the user and the version control managing group. (Who will that be?)
  - Jens: logging is extremely important for if people are trying to sabortage research (like uploading fake data)
- public mode:
  - restricted exclusively to VALUE members till 201X (2 years? needs decision)
  - if validation not done, data is flagged

downloading (looking at):

- no responsibility if errors occur
- acknowledgements

version control:

- how do we define versions, new version? expert judgement?
- commercial and non-commercial use...?
- what information on models needs to be given?
- citation of data VALUE

comments:

- Jens: being careful about commercial use, not ignore the issue
- Heike: make all data free. a lot of work with version control and logging - who is doing it and for how long?

Portal (José)

- functions are implemented
- validation indices can be computed
- webportal created and maintained by external company (spin-off of Santander Group)

Controlled vocabulary (Rasmus)

- precise definitions → improved communications.
- validation is a new niche.
- all else is on metadata, common standards and coordination.
- common methods (Data Reference Syntax DRS)
- transparency, visibility of validation
- replicable and traceability
- community-based vocabulary, reflect common use, technical terms, explanations/definitions. Sometimes different meanings, which can be mapped, referenced, examples.
- should help a lot with our portal and version control, metadata describing the model. Helps in data search and selection engage the community (VALUE, EUPORIAS, CORDEX, ClipC...)
- within VALUE: read and comment, discuss comments, poll, collect comments from users
- e.g. R-package 'esd' (check zoo objects)

Discussions

framework decision

(Bruce) SD doesn't get teleconnections right. → Should we add a fifth aspect to capture that?
The fifth aspect would consider large scale, but also small scale (monthly fluxes), physical consistency, teleconnections, scale consistency.
We might replace inter-variable by physical consistency and divide that with inter-variable and large-scale. not decided yet.

indices

"indices" is a not good name, as it is also a one-dimensional represenatative of a multidimensional thing. "statistics" has the same "problem". "diagonstics" not good as well as it is not including the validation? → beer-discussion.

data issue

stations

50 **stations** and 50 gridboxes "over" those stations. No high resolution gridded data available. How do we get the 50 gridboxes? Intersection of high resolution gridded data and 50 stations. NO EOBS! → 5 regions *rightarrow* 5 stations each → 25 boxes

nested stations

Nested station data: high station density in southern Sweden for a catchment. (Thomas, Fredrik, Renate) Decide on stations as soon as possible. (for multi-site weather generators)
Discussion:

Alrun do we need that anyway, as comparison with obs is done anyway by everyone self, as we can't apply pseudo reality.

Martin 5 stations, and then stick to 5 and change size of region, or only change size with increasing number of stations? do we need to nail that down? *leave it open.*

Sven if it is only a small number of people/groups working on this, should we skip it in the framework, even though doing it in VALUE?

Douglas huge community on applied small scale statistical downscaling (detached to climate community), no skipping. *(decided by majority)*

Jens it is useful to figure out in which context which method is applicable.

Andreas F. prepared such a experiment with Ole R. in Switzerland (but tyvärr not publicly available data)

subdaily

(Douglas) suggest to use station data, not gridded

experiment 3: gcm errors

to test the whole regional simulation including GCM errors.

Jens biggest issue is to actually characterize the gcm error. if like IPCC (only max and min) we miss a lot.

Douglas on what scale should we validate? only with gridded data.

cross-validation

kfold most likely 5 fold
data upload:

- 1 series for deterministic mthods
- 100 series for stochastic
- train model 5 times (5 period slices)
- in VLAUE → systematic sampling → 5 fold cross-validation of 5 consecutive blocks
- beyond VALUE → random sampling

(douglas) downscaled rcm precip and obs correlation order 0.9 for seasonal. (study of martin and douglas) pairwise comparison on larger time scales. do we want that? problem of miss interpreting things (sven). issue to discuss.

validation options in portal

Validation for specific variables for specific location → spatial validation.

Douglas is that implemented in the portal or easy to do?

José yes, not a problem.

Douglas how do we validate inter-variable relations?

José experiments can be joined. But probelm for metadata. We need multidimensional metadata

- needs to be discussed, but not here in the plenary → will be smaller group.
- spatial val → small working group on deciding and implementing that. (in principle possible)

experitment 1

Was approved.

experiment 2

pseudo reality (do our methods fail under climate change applications)

1. perfect predictor: RCM downscaling GCM ← pseudo reality. SD downscaling GCM. Comparison (check Frias et al for ENSEMBLES)

Jens shouldn't we do a pilot study with more than one RCM (pseudo reality)?

Douglas More GCMS? Like 2 GCMS, 2 RCM?

Sven like we said in Augsburg?

Christian and Alrun... more RCMs than GCMs. Like 1 GCM, and multiple RCMs. How many RCMs?

Rasmus What are the questions we want to answer?

- Decide that tomorrow after checking Augsburg agenda.

2. imperfect predictor (RCM bias correction): pseudo reality a GCM, a RCM : validation is a GCM, b,c,d...RCM (multiple RCMs different from a)

3. imperfect predicor (GCM bias correction): pseudo reality a GCM, a RCM: validation is b,c,d,... GCM, a RCM

covering non-stationarity - pseudo reality

Alrun Do we test if the blocks are statistically different?

Douglas No, as this is not the reason for this experiment. But it is to avoid the artificial skill of overfitting.

- Discussion about cross-validation fundamentals. Different methods (Joanna, Julie), like sequential and (something I missed) ... leave-one-out.
- We define cross-validation as José suggested (*approved*).
- More flexible implementation regarding folds and variations of cross-validation

Terms and conditions (Martin

The document will be on webpage.
Collection of questions/topics:

Christian raw downscaled data re-distribution

José no distribution to third parties, to keep VALUE server the up-to-date one, to keep track of our data, to avoid mirroring. José will send a sentence from UERRA. Default answer is no, if request for collaboration comes, we discuss again.

Douglas fake data upload

Jens people who are interested in destroying research/databases. We would need to do a "police force" to check the data and reserve the rights to remove the data. Abuse of names, and upload fake data with wrong name. Thus we should include a sentence that we will warn, discuss and remove data when suspecting fake data upload. (action might be taken if uploaded data jeopardise the functionality of the portal. Rerserve the right to remove data after warning, block accounts.)

Portal

How and who work is done

- during VALUE?

Heike we need a webmaster to control the data transfer... for a certain time.

Douglas Portal will be maintain by Predictia in Santander (needs to be discussed)

- what do we have to check?

Heike look for "stupid" output, and report it.

Heike webmaster checks what disrupts functionality

José you need to fill in a lot of metadata, which can help against fake data.

José technical attacks are covered already.

Douglas community is quite small, we know each other. It is difficult to check the data properly. So we will find out only when doing the comparison and writing a paper. Maybe that is enough?

José Douglas a sentence that we will help with uploading results.

webmaster, who?

(Douglas) someone outside Predictia. Decide later.

after VALUE?

post VALUE discussion tomorrow

version control

José we need to trust people, as there are to many ways to fake.

heike you will have to give a history about your model anyway, there is nothing to hide.

Douglas what kind of metadata is saved?

José user needs to label what he is uploading (we need guideline). having a free-text field to comment on version changes.

Martin user is obliged to provide a complete set of metadate, we reserve to force people to label their upload differently. We will have a definition of what is a version and what is a model. Discussion needs to be continued related to NPCC, but in a smaller subgroup.

- abuse
- statement of scientific conduct and responsibility add a sentence on good scientific practice.
- commercial use of the results

minimum number of stations

Duglas suggesting incomplete upload should be allowed, even 1 station.

Renate repeating augsburg

Christion agreed to Douglas, as e.g. weather-typing methods are (workwise, timewise) expensive when they need to be recalibrated to each region.

José experiments are predefined, so they can be incomplete but not changed.

Heike go back to teh definition of metadata, including which stations.

José for the intercomparison 1 station wouldn't be enough as everyone might upload different stations, no intercomparison will be possible.

Douglas papers are a different story, pilot papers, of course, there everyone should do as much as possible.

Heike everyone should be allowed to upload and validate whatever stations (within the ones we provide), also to encourage people to do more,

José experiment 1 paper, as open as possible

Heike & Douglas if a method is not designed for Spain, why then evaluating it there? that makes no sense.

Douglas minimum could be to pick one of those 10 station, to have a common station for each region.

Acknowledgements

COST, data provider, framework paper. put acknowledgements in a central place, so that everyone can cite it easily independent on what data was used (stations, grids...)

whether you should be allowed to upload results selectively

José
- Metadata based selection.
- Defined by webportal group.
- User can't decide about which measure is calculated, but can see it clearly which is calculated.
- It is easier when the user can't decide on validation, but if we change the framework, we will inform the user on the changes.

Douglas sentence be more like, this method is not designed for this meassure...

Heike comment field for numbers the modeller is not "happy" with

Renate making predefined validation fields transparent is important.

Renate can we change the framework later, and calculate measures after teh validation of a method has been done already?

Douglas yes, as data is publicly available anyway. so everyone could do this extra validation.

José keep tracks on comments

Christian can one look at his data before publishing

José yes, you don't have to publish.

Douglas is complete withdrawel allowed? → default **no**, but if yes inform "us" and "we" will withdraw your method.

Douglas after publication, no withdrawel, stays in public space.

what if someone drops out of VALUE (post VALUE legacy)?