# VALUE
# Defining the Validation Framework

Lisbon, 26/27 March 2013

*«Those among us who are unwilling to expose their ideas to the hazard of refutation do not take part in the scientific game.»*

*Karl Popper, Logic of Scientific Discovery*

# Adaptation to climate change...

*Müritz-Elde-Canal, W. Illner*

# Adaptation to climate change...

in the light of

- competing interests and
- democratic decisions

# Adaptation to climate change...

...requires **robust regional** information

in the light of

- competing interests and
- democratic decisions

# Outline

Summary from Trieste (Revised)

Agenda and Planned Discussions

# Summary from Trieste (Revised)

## Agenda and Planned Discussions

# Aims of a Validation

**We cannot be complete**

- We cannot validate all aspects.
- We have to validate the **relevant** aspects.

**There is no best method (probably)**

- We do not aim to identify a single best method, but all bad methods.
- We aim to identify the (relevant) good and bad aspects of each method.

# Validation Hierarchy

- Level 1: Aspects
  e.g., marginal distribution
- Level 2: Characteristics
  e.g., extremes
- Level 3: Indices
  e.g., 90th percentile
- Level 4: Measures
  e.g., quantile verification score

Evaluation of climatology and long term changes

# Aspects

(of the multi-variate distribution)

## Generic aspects

- marginal distribution
- temporal structure

## Further aspects

- spatial structure
- consistency between variables

Remember: The framework has to work for dynamical and statistical approaches!

# Characteristics

- full marginal distribution
- bulk of the distribution
- wet day probability
- extremes
- annual cycle
- diurnal cycle

...

# Indices

- statistical moments (mean, variance, covariance...)
- 90th percentile
- characteristic time/space scales
- phase/amplitude of annual cycle
- spell length indices

  ...
- measures for spatial patterns?
- measures for long term variability?

# Measures

- skill scores
- root mean squared error
- bias

  ...
- pointwise vs. spatial evaluation?

Summary from Trieste (Revised)

# Agenda and Planned Discussions

# Overview

## First Day

- WG reports;
- presentation of end user questionnaire;
- plenary, breakout and wrap up discussion on end user needs.

## Second Day

- discussion on data and experimental setup;
- plenary, breakout and wrap up discussion on validation framework;
- open discussion.

# Open Discussion

**To focus on VALUE specific topics (see MoU!)**

- save side discussions for the end;
- think now of what you want to discuss;
- who wants to chair?

# Results from End User Questionnaire

**To decide what is required/relevant**

- which variables?
- which aspects?
- which characteristics?
- partly: which indices?

**To asses the validation results**

- at which accurracy?

# Discussions on End User Needs

Integrate the results from the questionnaire into our hierarchy

- which variables?
- which aspects?
- which characteristics?

A decision has to be drawn!

# Data and Experimental Setup Discussion

**Observational data**

- which regions?
- which station/gridded data sets?
- how much do we cover with public data?
- time period?

**RCMs and pseudo reality**

- ERA40 vs. ERA interim?
- which resolution for pseudo reality?
- which GCM(s) for pseudo reality?

# Validation Discussion

## Filling the levels

- Level 3: which indices?
- Level 4: which measures?

A decision on level 3 has to be drawn!

## Lessons from forecast verification

- We need to assess predictive power for all methods.
  $\rightarrow$ event-wise validation!
- To what extent can we use forecast verification methods in a distribution-wise validation? (decomposition of scores)