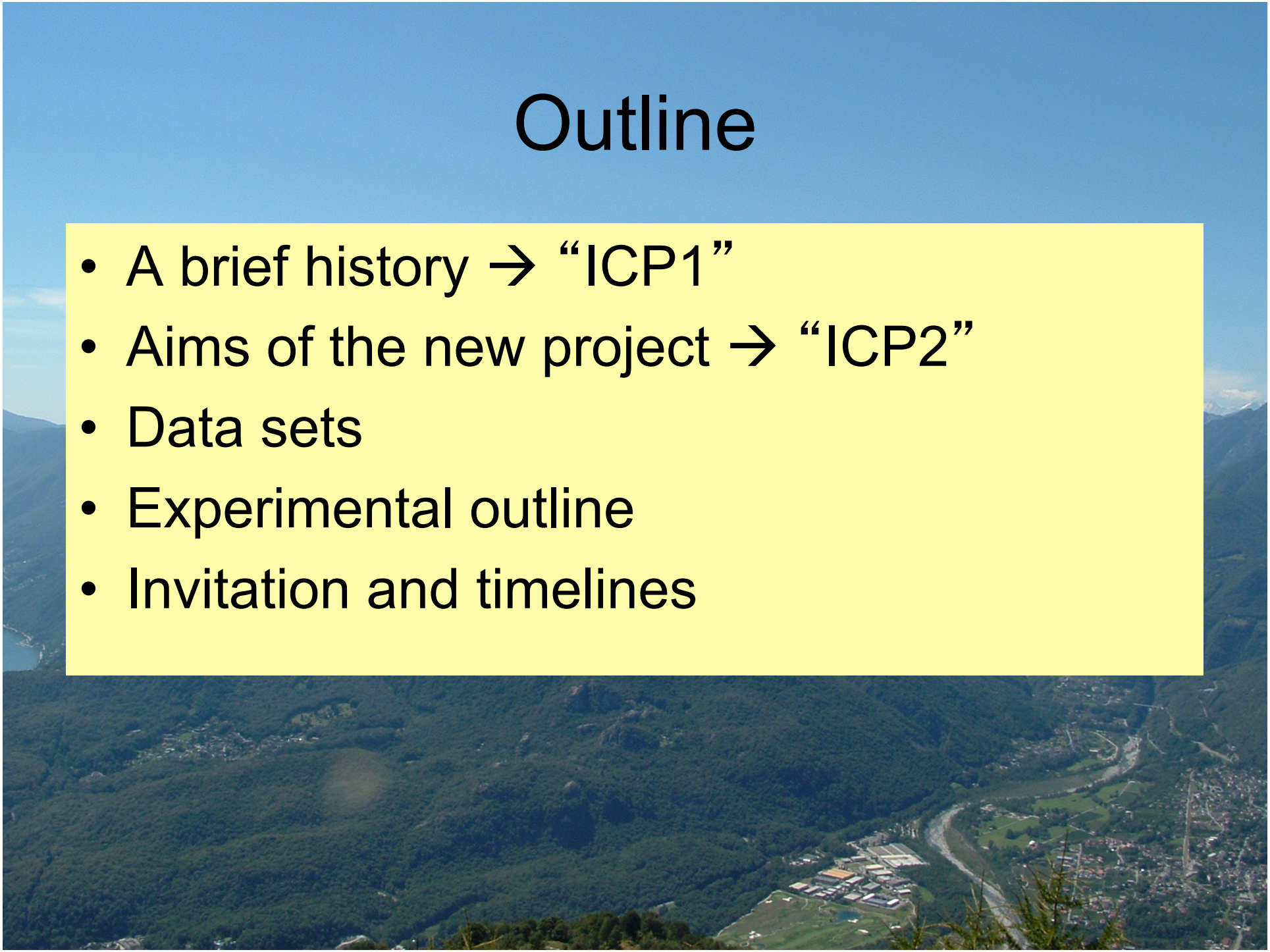
Intercomparison of spatial forecast verification methods: A review and new project launch

Project committee:
Marion Mittermaier, Met Office, UK
Eric Gilleland and Barb Brown, NCAR, USA
Manfred Dorninger, University of Vienna, Austria
Beth Ebert, Bureau of Meteorology, Australia
Laurie Wilson, Environment Canada

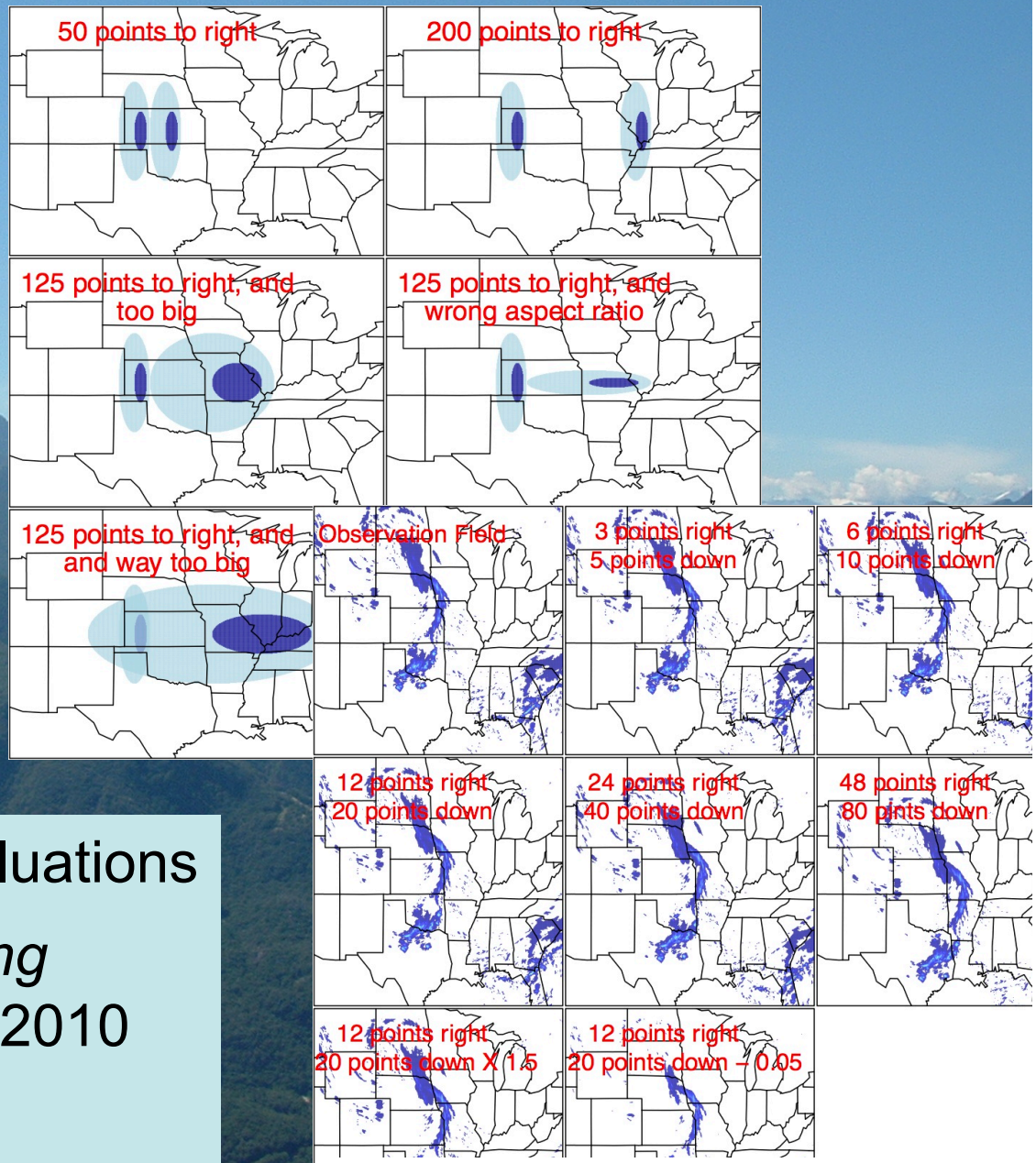
Outline

- A brief history → “ICP1”
- Aims of the new project → “ICP2”
- Data sets
- Experimental outline
- Invitation and timelines



ICP

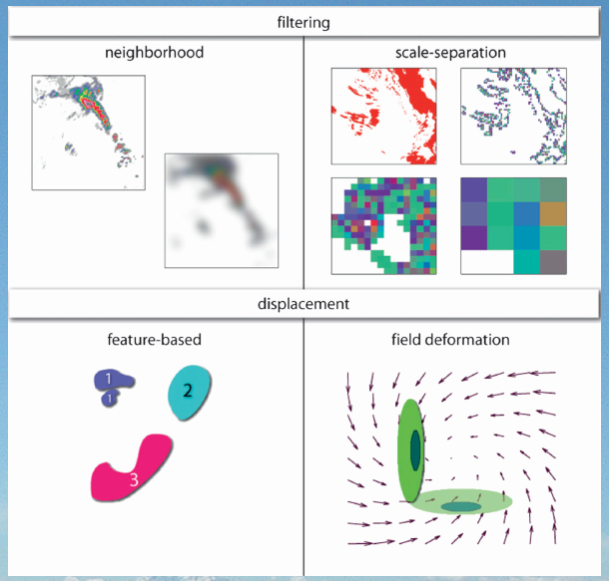
- 2008-2010
- Focus on precipitation
- Methods applied by researchers to same datasets (real forecasts; perturbed cases; idealized cases)
- Subjective forecast evaluations
- *Weather and Forecasting* special collection 2009-2010
- Code available online



http://www.ral.ucar.edu/projects/icp/

Weather and Forecasting special collection 2009-201

Categorisation of methods



Category	Scales with skill	Location errors	Intensity errors	Structure errors	Occurrence (hits, misses, false alarms)
Traditional (gridpoint)	×	×	✓	×	✓
Neighbourhood	✓	×	✓	×	✓
Scale separation	✓	×	✓	×	✓
Features based	×	✓	✓	✓	✓
Deformation	×	✓	✓	×	×

Gilleland *et al.*, *Bulletin of the American Meteorological Society*, 2010

Different methods have different aims

Scale separation and neighbourhood methods

- Focus on accuracy quantification
- What is the forecast accuracy at small scales? Large scales? Low / high intensities?
- What spatial scales and intensities have reasonable accuracy?
- Different methods emphasize different aspects of accuracy

Feature-based methods

- Focus on describing the error
- What is the error in *this forecast*?
- What is the cause of this error (wrong location, wrong size, wrong intensity, etc.)?

Field deformation methods (morphing)

- Focus on describing phase errors
- Does the shape/placement of the forecast resemble the observations?

Usage

Scale separation and neighbourhood methods

- Whenever high density observations are available over a reasonable domain
- When knowing scale- and intensity-dependent skill is important
- When comparing forecasts at different resolutions

Feature-based methods

- When features are well defined (organized systems, longer rain accumulations)
- When it is important to measure how well the forecast predicts the properties of systems
- When size of domain >> size of systems

Field deformation (morphing)

- When forecasts have a fair resemblance to the observations
- Absence of a feature in the observations or forecast leads to peculiar behaviour of distortion vectors
- When knowing phase errors of the field is important

Aims of ICP2

How do/can spatial methods:

- Transfer to other regions with *complex terrain*, and other parameters: *wind* (speed and direction) and *rain*?
- Work with *ensembles*?
- Incorporate *observations uncertainty*?

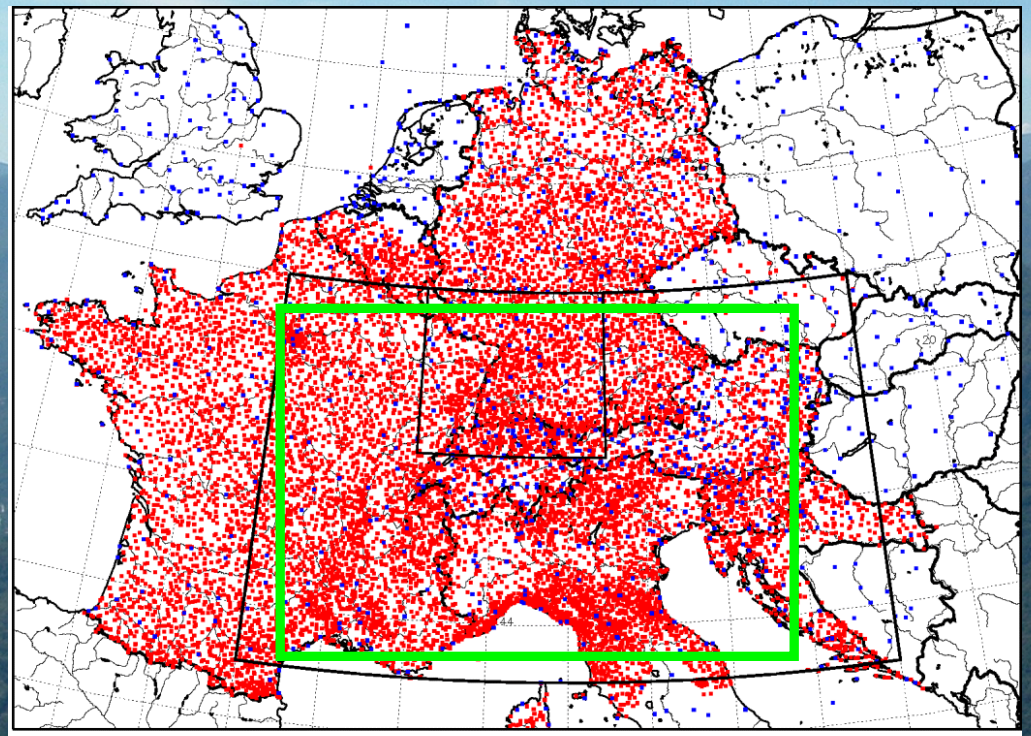
Models

- From MAP D-PHASE COPS archive
  - Deterministic 2 km COSMO-2 Init-time:
    - Initialised 06 UTC FC-range: 24h
  - Ensemble 10 km CMC-GEM-H Init-time:
    - Initialised 06 UTC FC-range: 18h

- Invitation for modelling centres to produce re-runs of cases with more up-to-date model configurations (Tier 3), but core experiments to be done using COSMO-2 and CMC-GEM-H.

Observations data set

JDC-data: D-PHASE (FDP, Rotach, et al., 2009, BAMS) and WWRP COPS (RDP, Wulfmeyer, et al., 2008, BAMS), data available: (http://cera-www.dkrz.de/WDCC/ui/Index.jsp)



- 32 data providers
- GTS-Stations: 1232
- NGTS-Stations: > 13000
- Mean station distance: GTS: ~ 36km
  GTS+Non-GTS: ~ 12km

Frames: D-PHASE (black, large)
COPS (black, small)
this study (green)

Red: Non-GTS stations
Blue: GTS stations

VERA analysis scheme

(Vienna Enhanced Resolution Analysis)

Data quality control scheme

+

Thin-Plate-Spline algorithm

+

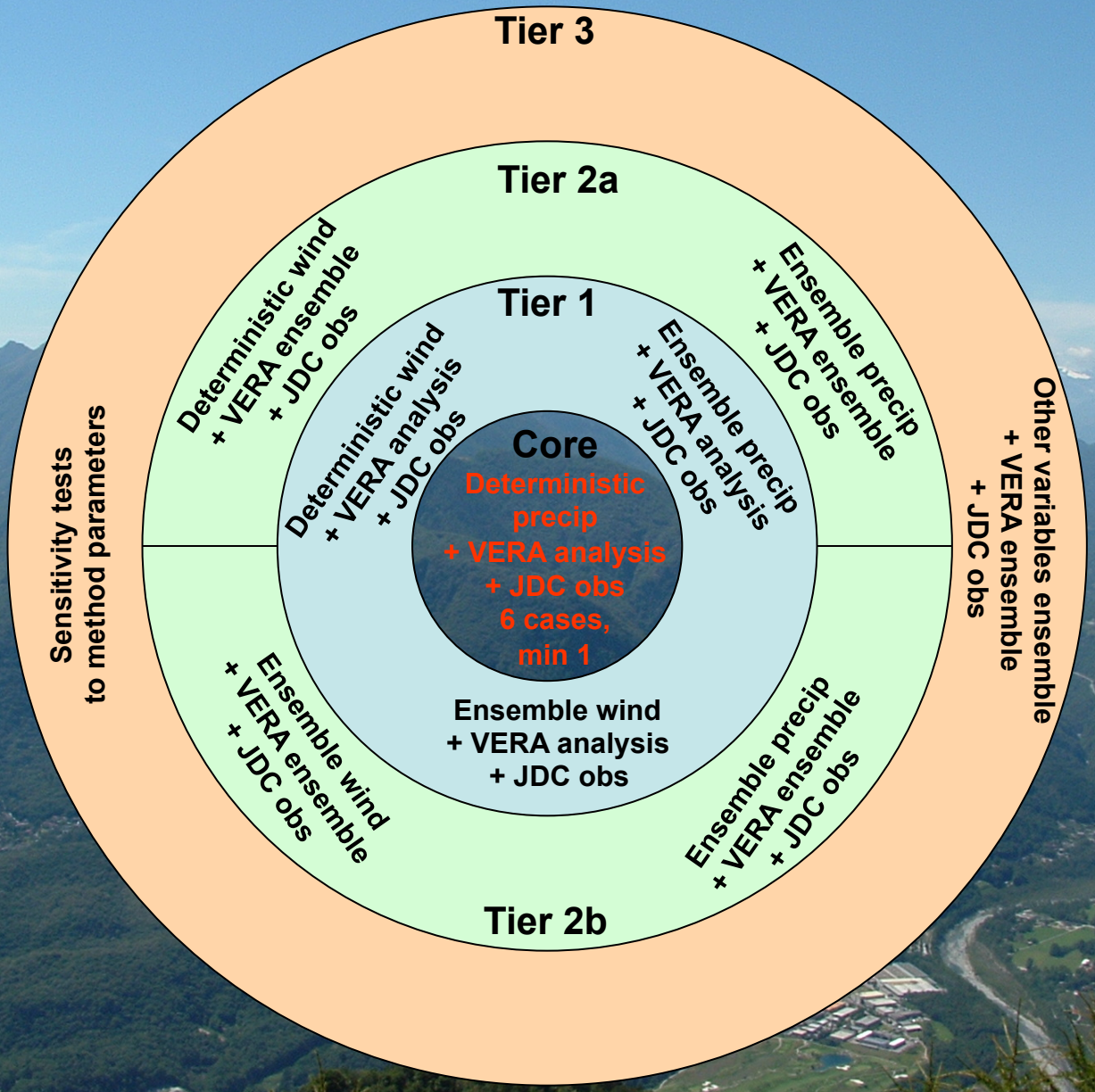
Downscaling via the „Fingerprint“ method

Not dependent on first guess fields – „model independent“

Wind	Potential Temperature	Precipitation: Accumulated to 1h, 3h, 6h, 12h, 24h	Post processing: - Mixing Ratio - Moisture Flux Divergence
MSL - pressure	Equivalent – Pot. Temperature		

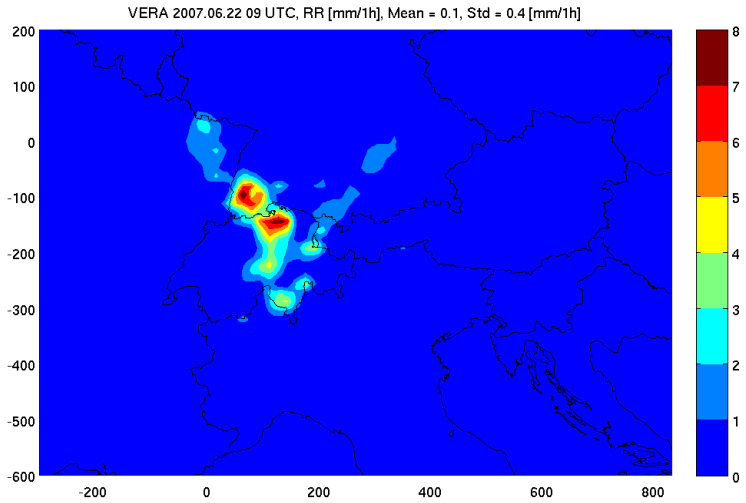
Further reading: Steinacker, et al. 2000 (MWR), Steinacker, et al. 2006 (MWR), Steinacker, et al. 2011(MWR)

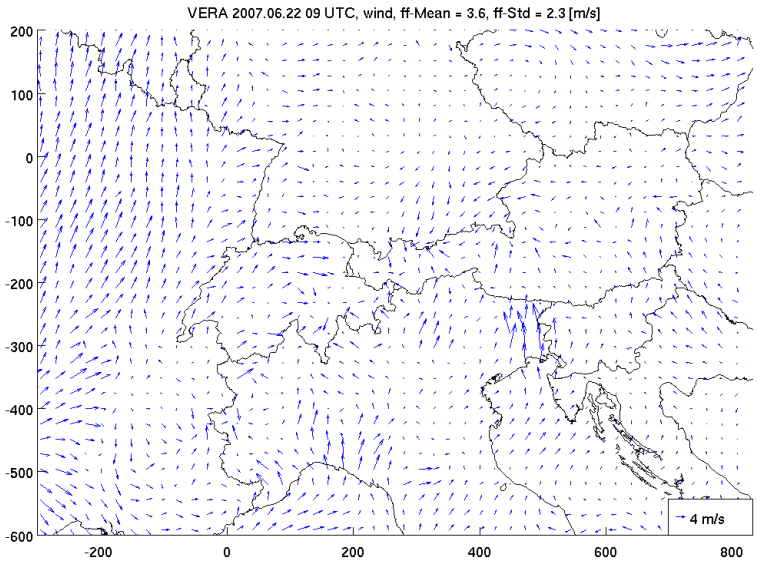
Experimental design

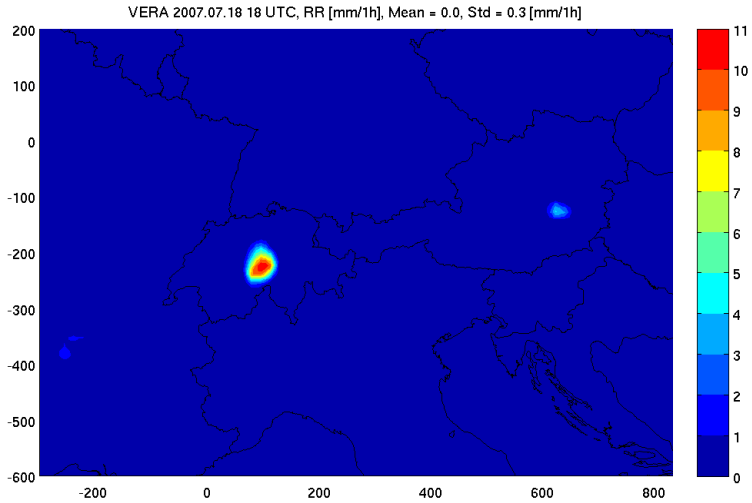


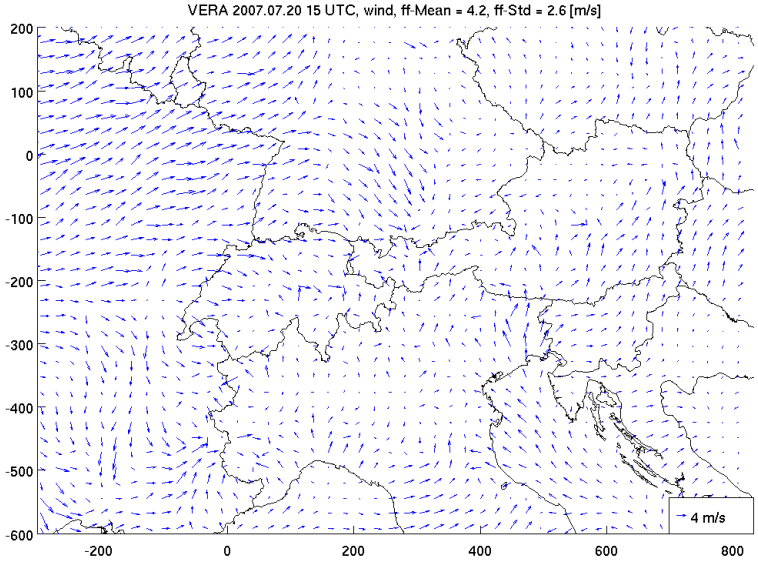
Outcomes

- Participants must complete the core experiment for at least case 1 to formally be classed participants.
- This requires the provision of hourly verification statistics (following the forecast evolution) for
  - Hourly precipitation (and 6h precipitation)
  - Hourly wind speed and direction
- Participation in subsequent tiers 1-3 is at the discretion of participants, but output should follow the same rules as above.
- We extend an invitation to modelling centres to produce re-runs of cases with more up-to-date model configurations under Tier 3, but core experiments must be done using COSMO-2 and CMC-GEM-H.


VERA 2007.06.22 09 UTC, RR [mm/1h], Mean = 0.1, Std = 0.4 [mm/1h]
200
100
0
-100
-200
-300
-400
-500
-600
-200
0
200
400
600
800
8
7
6
5
4
3
2
1
0


VERA 2007.06.22 09 UTC, wind, ff-Mean = 3.6, ff-Std = 2.3 [m/s]
200
100
0
-100
-200
-300
-400
-500
-600
-200
0
200
400
600
800
4 m/s


VERA 2007.07.18 18 UTC, RR [mm/1h], Mean = 0.0, Std = 0.3 [mm/1h]
200
100
0
-100
-200
-300
-400
-500
-600
-200
0
200
400
600
800
11
10
9
8
7
6
5
4
3
2
1
0


VERA 2007.07.20 15 UTC, wind, ff-Mean = 4.2, ff-Std = 2.6 [m/s]
200
100
0
-100
-200
-300
-400
-500
-600
-200
0
200
400
600
800
4 m/s

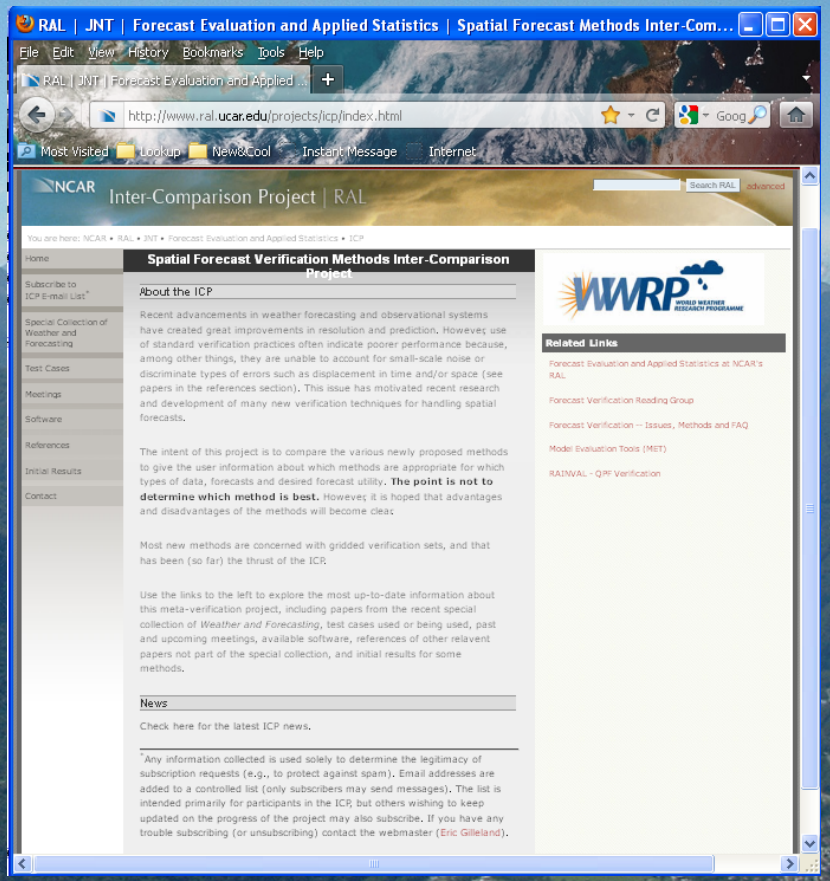
Invitation and timeline

- ***Now***: Recruiting participants
- ***Jan-Feb***: Data sets available from NCAR ftp; collecting datasets; work commences
- ***Mar 2014:*** Initial results session at

6th international verification methods workshop

17-19 March 2014

New Delhi



To express interest:

Send email to Eric Gilleland (ericg@ucar.edu)

http://www.ral.ucar.edu/projects/icp/index.html