WORLD METEOROLOGICAL ORGANIZATION WWRP WCRP JWGFVR Environment and Climate Change Canada universität wien

MET and MesoVICT - Tools and Data for the Application and Testing of Established and New Spatial Verification Methods

Part II: MesoVICT
- A review –

Manfred Dorninger
and the MesoVICT community

2020IVMWO
16 Nov. 2020

Outline:

1) Motivation of MesoVICT
2) Aims of MesoVICT
3) Data
4) Strategy and results (selection)
5) Participants
6) Resume

MesoVICT Scientific Committee:

Marion P. Mittermaier (Met Office)
Manfred Dorninger (Univ. Vienna)
Eric Gilleland (NCAR)
Barb G. Brown (NCAR)
Beth E. Ebert (BoM)
Barbara Casati (Env. Canada)
Laurence J. Wilson (former Env. Canada)

1) Motivation

The **Golden Decade** of spatial verification methods development (~2000-2010) peaked in the spatial verification method Intercomparison Project (ICP) to:

- learn about the fundamental behaviour of the methods
- learn about the relation between the methods
- learn about the pro and cons of the methods
- learn about the specific forecast properties described by the methods
- learn about possible unintuitive characteristics of the methods

→ Special collection (some 16 papers in 2009-2010)
https://journals.ametsoc.org/waf/collection/118/Spatial-Forecast-Verification-Methods-Inter

1) Motivation

The Golden Decade

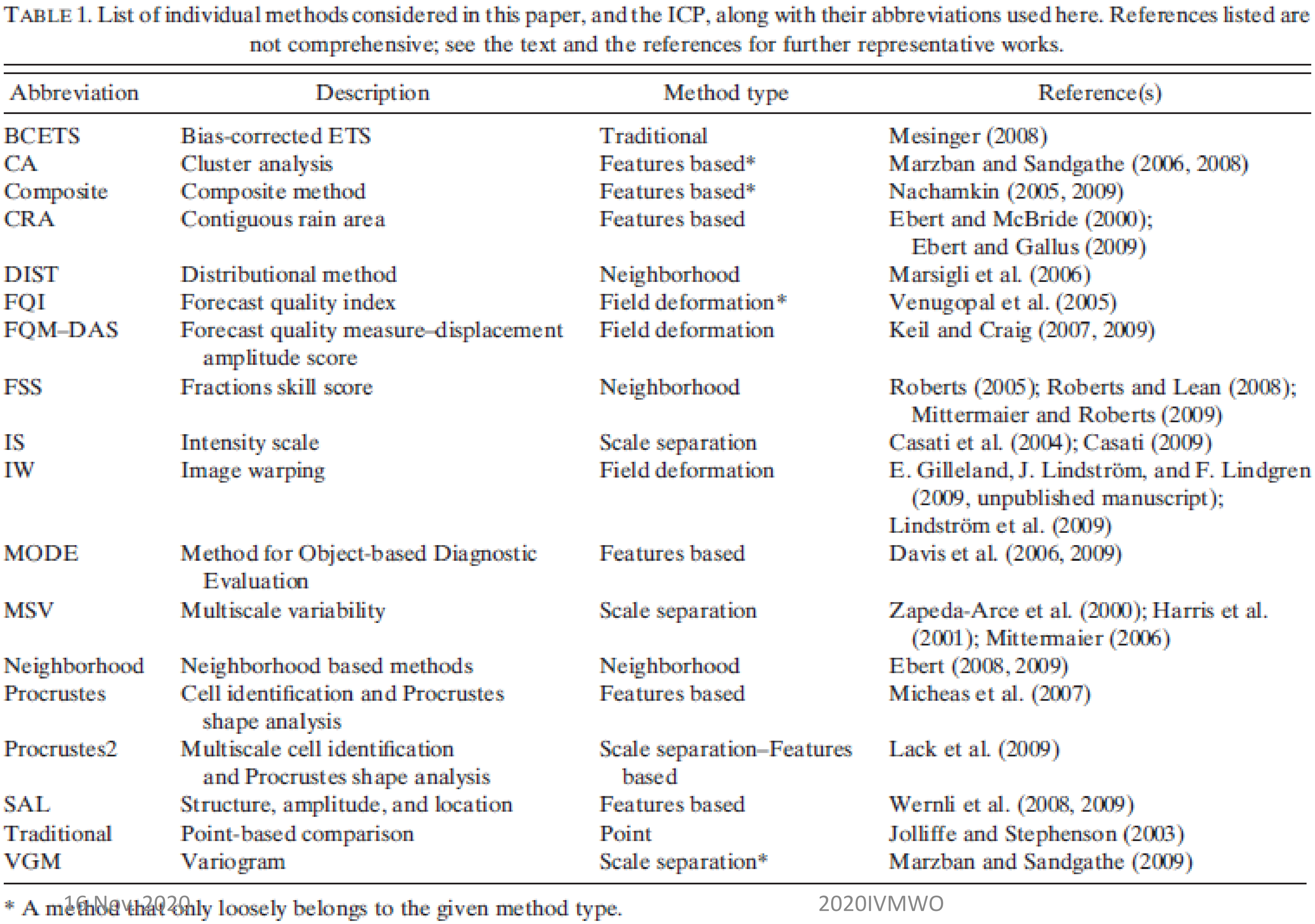


TABLE 1. List of individual methods considered in this paper, and the ICP, along with their abbreviations used here. References listed are not comprehensive; see the text and the references for further representative works.

Abbreviation	Description	Method type	Reference(s)
BCETS	Bias-corrected ETS	Traditional	Mesinger (2008)
CA	Cluster analysis	Features based*	Marzban and Sandgathe (2006, 2008)
Composite	Composite method	Features based*	Nachamkin (2005, 2009)
CRA	Contiguous rain area	Features based	Ebert and McBride (2000); Ebert and Gallus (2009)
DIST	Distributional method	Neighborhood	Marsigli et al. (2006)
FQI	Forecast quality index	Field deformation*	Venugopal et al. (2005)
FQM–DAS	Forecast quality measure–displacement amplitude score	Field deformation	Keil and Craig (2007, 2009)
FSS	Fractions skill score	Neighborhood	Roberts (2005); Roberts and Lean (2008); Mittermaier and Roberts (2009)
IS	Intensity scale	Scale separation	Casati et al. (2004); Casati (2009)
IW	Image warping	Field deformation	E. Gilleland, J. Lindström, and F. Lindgren (2009, unpublished manuscript); Lindström et al. (2009)
MODE	Method for Object-based Diagnostic Evaluation	Features based	Davis et al. (2006, 2009)
MSV	Multiscale variability	Scale separation	Zapeda-Arce et al. (2000); Harris et al. (2001); Mittermaier (2006)
Neighborhood	Neighborhood based methods	Neighborhood	Ebert (2008, 2009)
Procrustes	Cell identification and Procrustes shape analysis	Features based	Micheas et al. (2007)
Procrustes2	Multiscale cell identification and Procrustes shape analysis	Scale separation–Features based	Lack et al. (2009)
SAL	Structure, amplitude, and location	Features based	Wernli et al. (2008, 2009)
Traditional	Point-based comparison	Point	Jolliffe and Stephenson (2003)
VGM	Variogram	Scale separation*	Marzban and Sandgathe (2009)

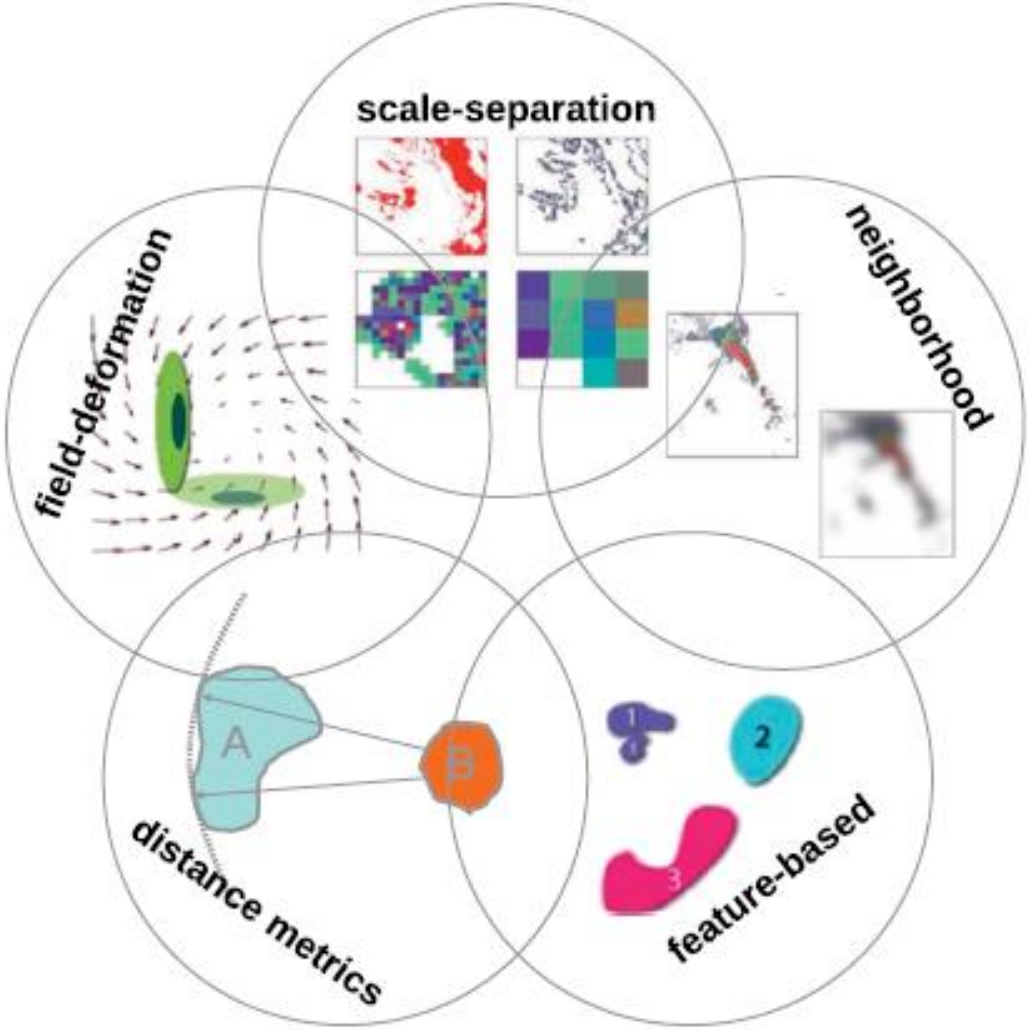
* A method that only loosely belongs to the given method type.

Gilleland, et al., 2009

2) Aims of MesoVICT

MesoVICT in a nutshell:

MesoVICT focuses on the application, capability, and enhancement of **spatial verification methods** as applied to **deterministic and ensemble forecasts** of **precipitation, wind, and temperature over complex terrain** and includes **observation uncertainty** assessment (from Dorninger et al., 2018, BAMS).

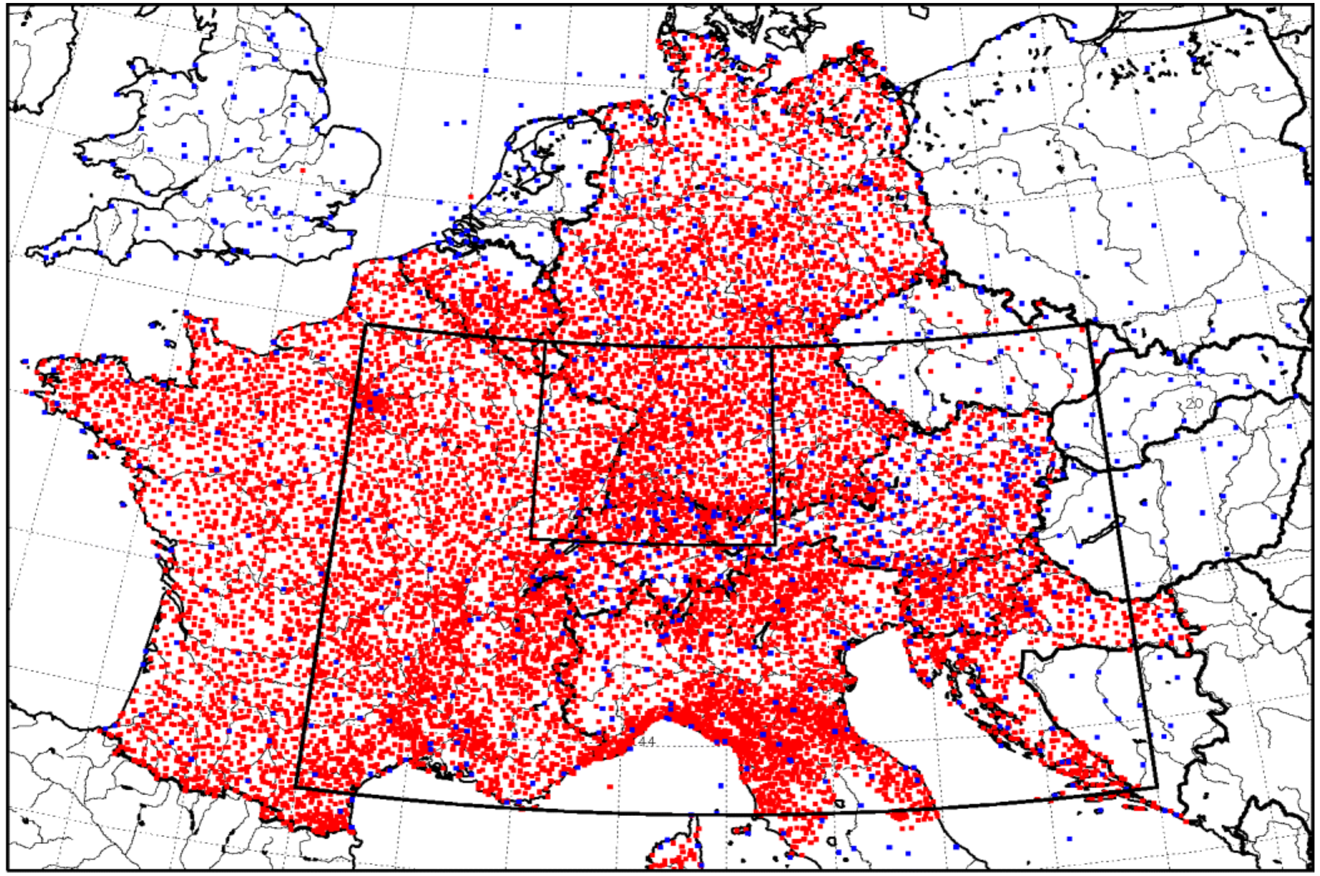


Dorninger et al., 2018, BAMS

MesoVICT as follow-up of the first ICP.
http://www.ral.ucar.edu/projects/icp/index.html

3) Data

Observation data: surface data: WWRP COPS (RDP, Wulfmeyer, et al., 2008, BAMS) and WWRP D-PHASE (FDP, Rotach, et al., 2009, BAMS),



- data covering whole year 2007
- 32 data providers
- GTS-Stations: 1232
- NGTS-Stations: > 13000
- Mean station distance: GTS: ~ 36km
  GTS+Non-GTS: ~ 12km

Frames: D-PHASE (large)
& COPS (small) areas

Dorninger et al., 2013, NCAR TN-505+STR

3) Data

Analysis data: VERA (Vienna Enhanced Resolution Analysis) including analysis ensemble

Data quality control scheme
+
Thin-Plate-Spline algorithm
+
Downscaling via the „Fingerprint“ method
↓
Not dependent on first guess fields – „model independent“

Wind

Potential Temperature

MSL - pressure

Equivalent – Pot. Temperature

Precipitation: Accumulated to 1h, 3h, 6h, 12h, 24h

Post processing:
- Mixing Ratio
- Moisture Flux Divergence

Further reading:

Steinacker, et al. 2000 (MWR),
Steinacker, et al. 2006 (MWR),
Steinacker, et al. 2011(MWR)

3) Data

Input data: JDC data-set

VERA analysis:

- VERA has been run for the whole year 2007
- Time resolution: hourly
- Horizontal resolution: 8 km

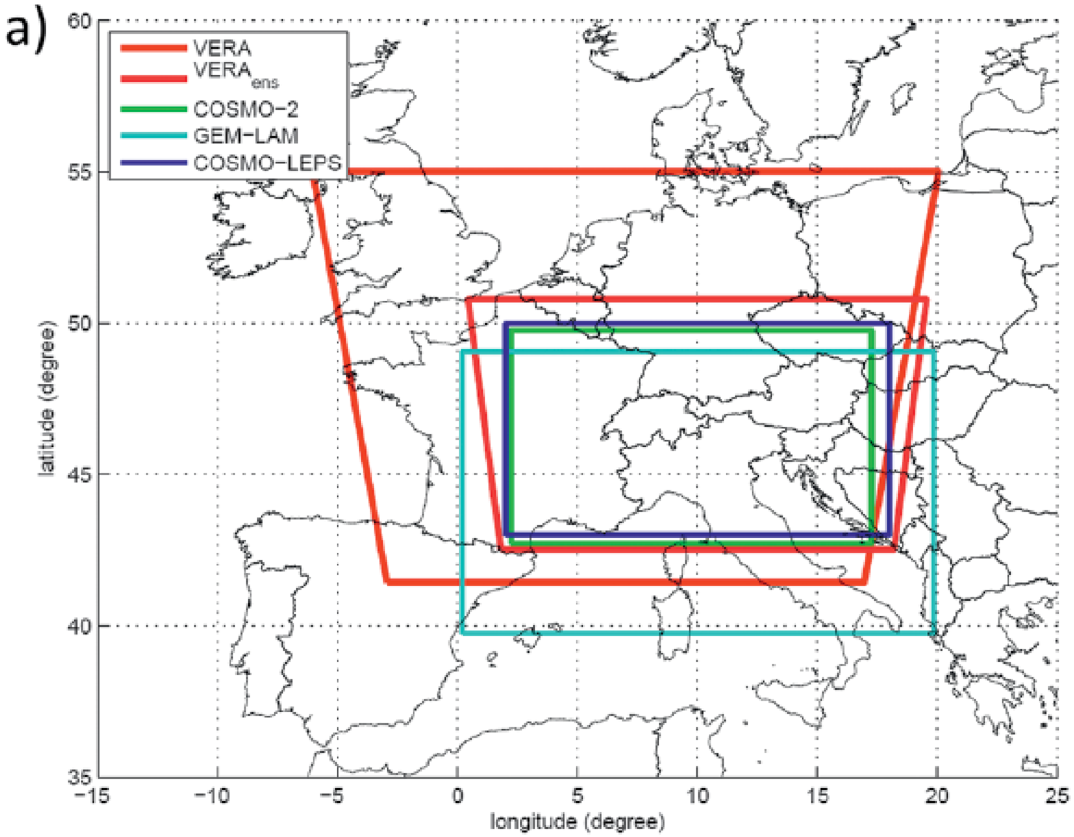
VERA analysis ensemble:

- ensemble generation:

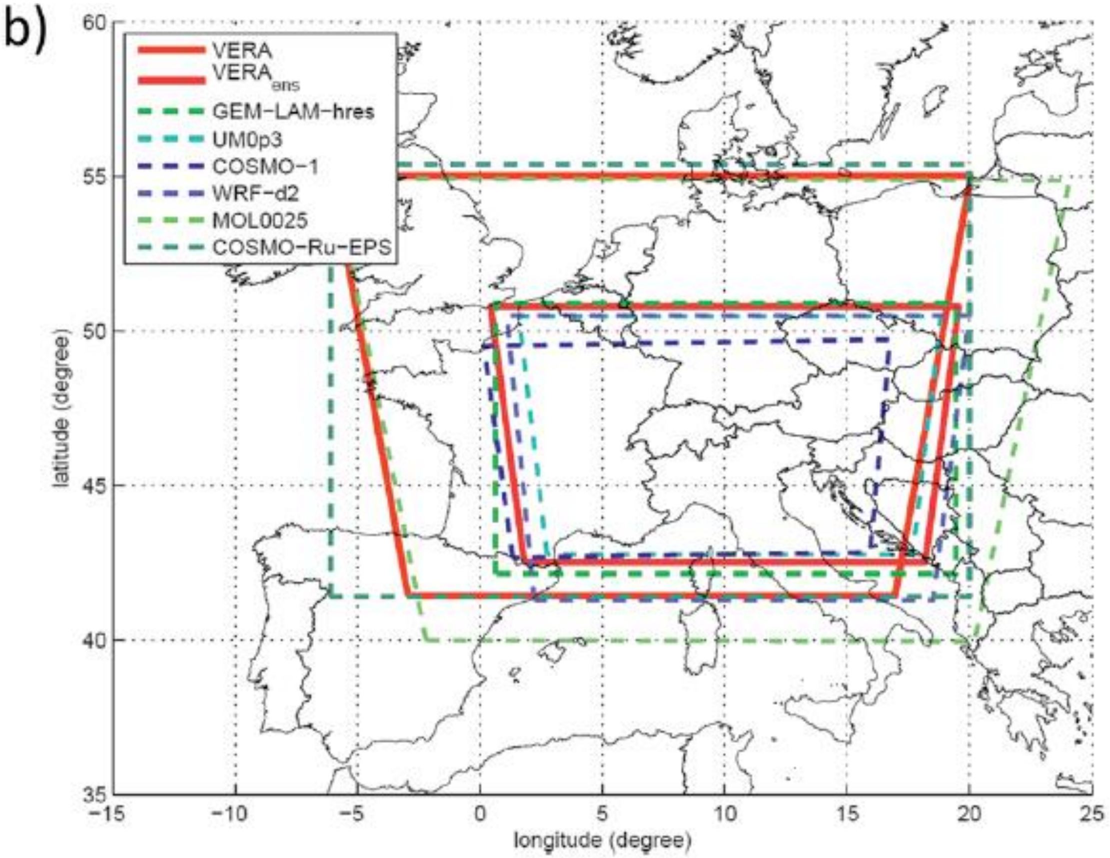
  Gorgas T., and M. Dorninger, 2012: Concepts for a pattern-oriented analysis ensemble based on observational uncertainties. *Q. J. R. Meteorol. Soc.*, **138**, 769-784.
- consists of 50 analyses, all VERA Parameters, 8 km grid
- created for the core case (20-22 June 2007)

3) Data

Model data (D-PHASE):

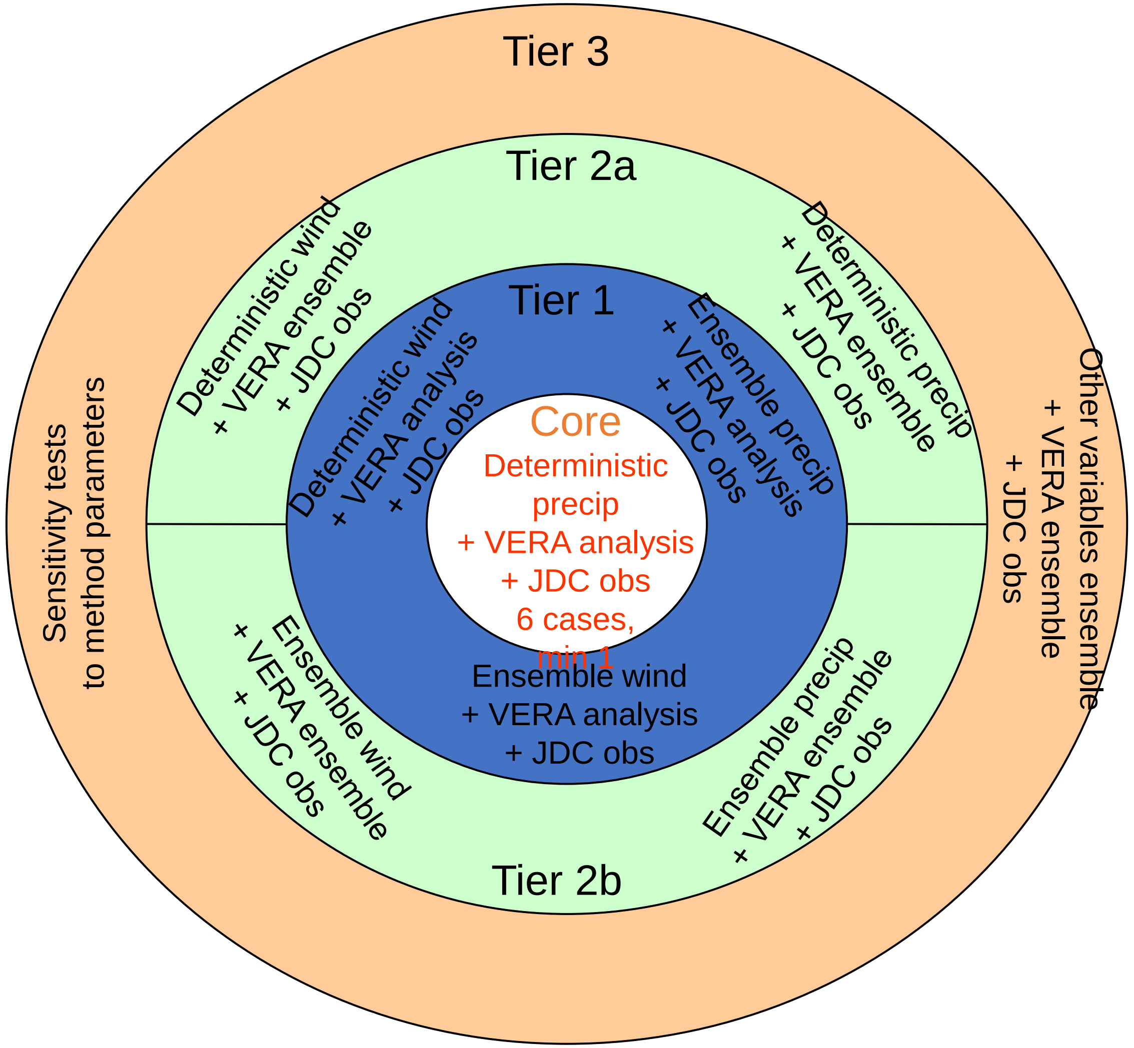


Model data (model re-runs):



All model data are interpolated on the VERA grid, same parameters as VERA, same file format (ASCII)

4) Strategy and results



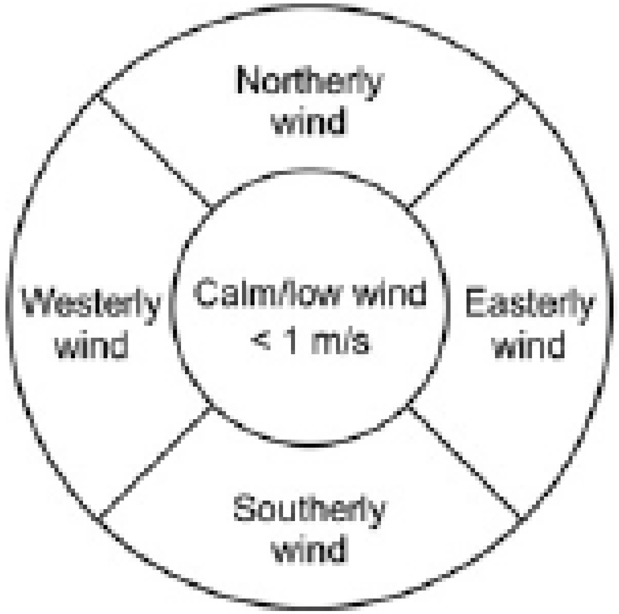
Case selection citerion:
Cases should cover a broad range of meteorological phenomenon in and around the Alps

For a more in depth synoptic description see:
Dorninger, et al., 2013: *NCAR-TN* 505.

4) Strategy and results (selection)

I. What is the ability of the method to verify forecasts of variables other than precipitation (e.g., wind)?

Skok, and Hladnik, 2018, (MWR) present an adapted FSS for wind verification



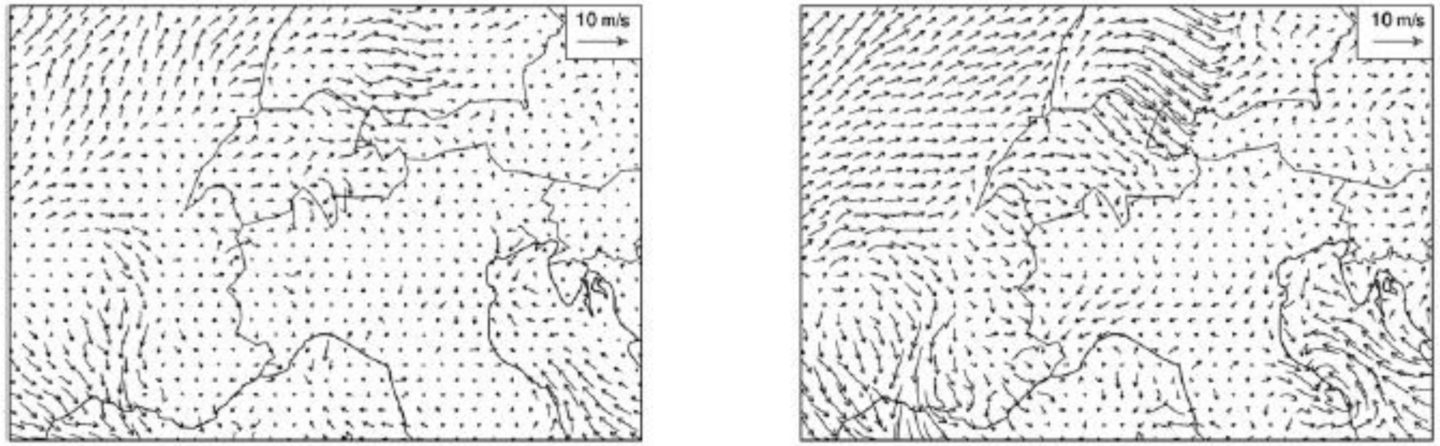
Basic wind class definition

$$\mathrm{FSS}_{\mathrm{wind}} = 1 - \frac{\sum_{k}\sum_{i,j}\left[O_k(i,j) - M_k(i,j)\right]^2}{\sum_{k}\sum_{i,j} O_k(i,j)^2 + \sum_{k}\sum_{i,j} M_k(i,j)^2},$$

O, M	fraction value for observation O and forecast M
i,j	Location
k	wind class at location (i,j)

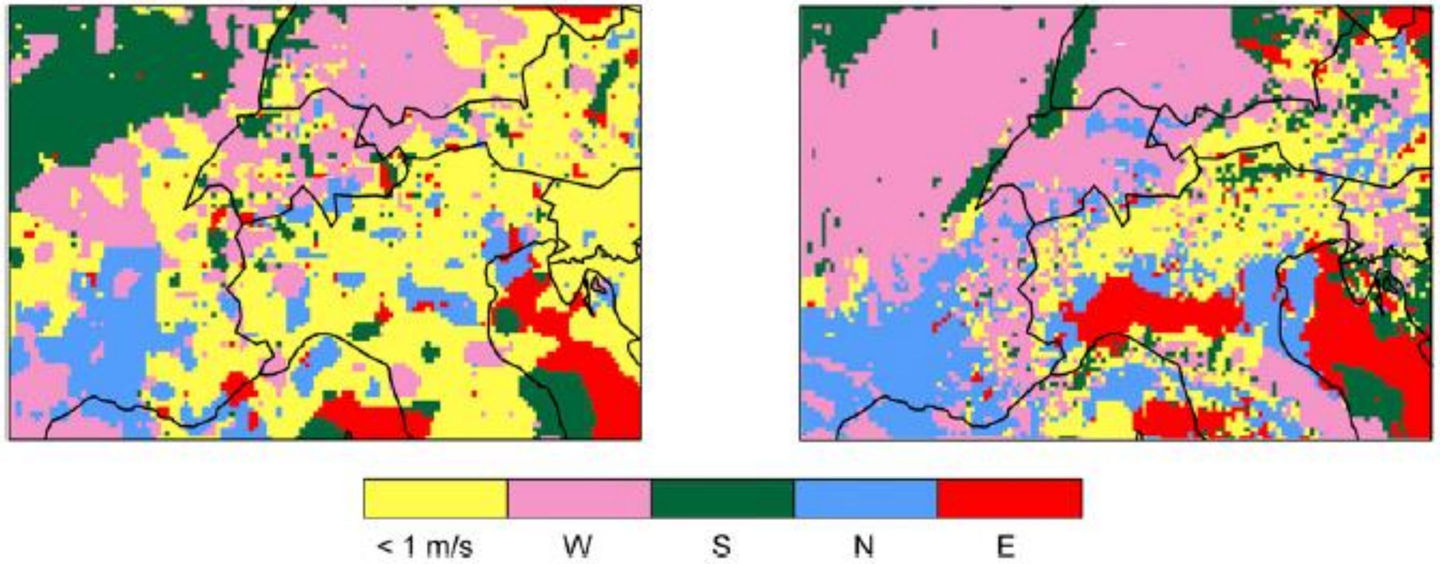
4) Strategy and results (selection)

Skok, G. and V. Hladnik, 2018, (MWR)



a) VERA analyses wind

b) CO2_00 model wind

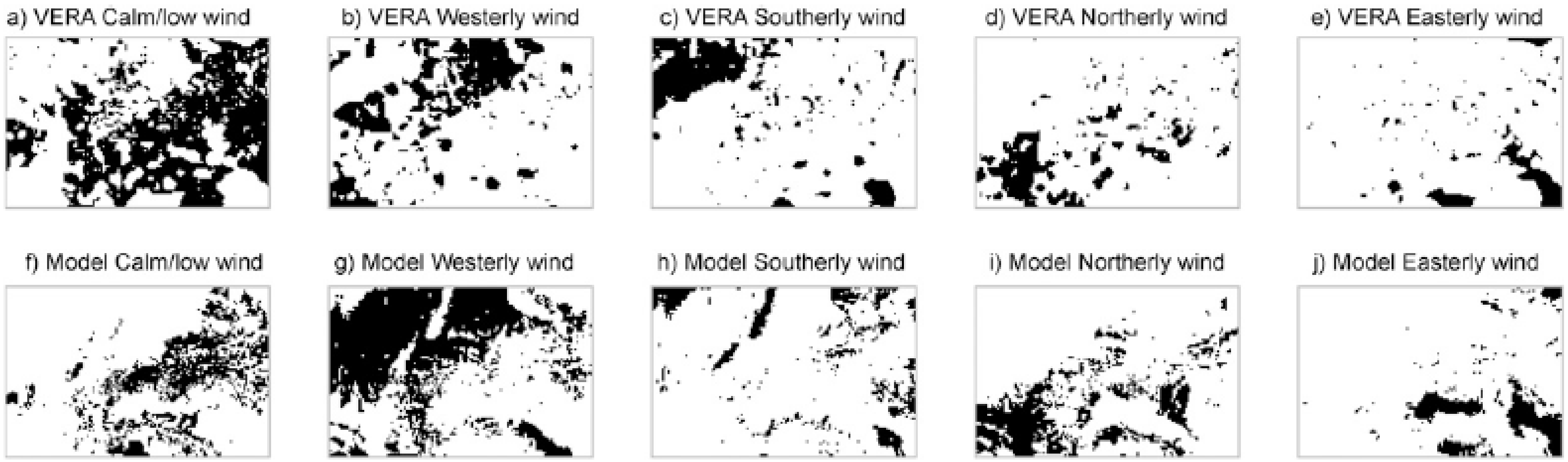


c) VERA analyses wind classes

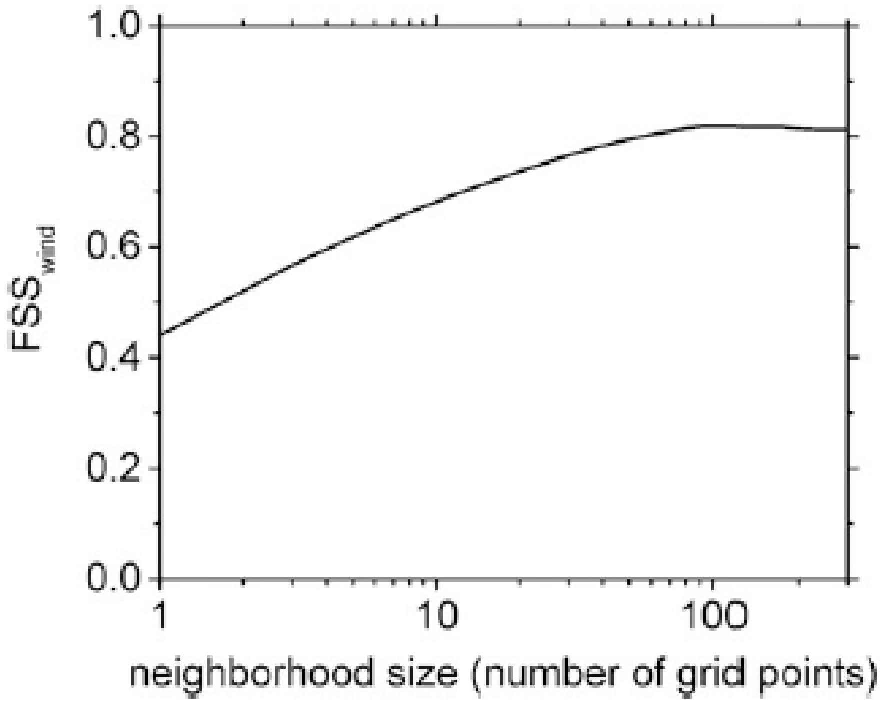
d) CO2_00 model wind classes

4) Strategy and results (selection)

Skok, G. and V. Hladnik, 2018, (MWR)



Binary wind class fields used to calculate the fractions



Method can be applied to other parameters as well (e.g., precipitation classes)

asymptotic value of ~0,8 indicates forecast bias

4) Strategy and results

II. How can the method be adapted to evaluate ensemble forecasts?

Radanovics, et al., 2018 (WAF) introduce an adapted SAL for ensemble forecasts

Original SAL definition

$$S = \frac{V_{\mathrm{mod}} - V_{\mathrm{obs}}}{0.5\left(V_{\mathrm{mod}} + V_{\mathrm{obs}}\right)}, \quad A = \frac{\overline{\mathrm{rr}}_{\mathrm{mod}} - \overline{\mathrm{rr}}_{\mathrm{obs}}}{0.5\left(\overline{\mathrm{rr}}_{\mathrm{mod}} + \overline{\mathrm{rr}}_{\mathrm{obs}}\right)},$$

$$L_1 = \frac{\left|\mathbf{x}\left(\mathrm{rr}_{\mathrm{mod}}\right) - \mathbf{x}\left(\mathrm{rr}_{\mathrm{obs}}\right)\right|}{d},$$

$$L_2 = 2\left[\frac{\left|\left(\frac{\sum_i \mathrm{rr}_i \left|\mathbf{x}_i - \mathbf{x}\right|}{\sum_i \mathrm{rr}_i}\right)_{\mathrm{mod}} - \left(\frac{\sum_i \mathrm{rr}_i \left|\mathbf{x}_i - \mathbf{x}\right|}{\sum_i \mathrm{rr}_i}\right)_{\mathrm{obs}}\right|}{d}\right],$$

$$L = L_1 + L_2.$$

eSAL definition

$$eS = \frac{\left\langle V_{\mathrm{mod}}\right\rangle - \left\langle V_{\mathrm{obs}}\right\rangle}{0.5\left(\left\langle V_{\mathrm{mod}}\right\rangle + \left\langle V_{\mathrm{obs}}\right\rangle\right)}, \quad eA = \frac{\left\langle \overline{\mathrm{rr}}_{\mathrm{mod}}\right\rangle - \left\langle \overline{\mathrm{rr}}_{\mathrm{obs}}\right\rangle}{0.5\left(\left\langle \overline{\mathrm{rr}}_{\mathrm{mod}}\right\rangle + \left\langle \overline{\mathrm{rr}}_{\mathrm{obs}}\right\rangle\right)},$$

$$eL_1 = \frac{\left|\left\langle \mathbf{x}\left(\mathrm{rr}_{\mathrm{mod}}\right)\right\rangle - \left\langle \mathbf{x}\left(\mathrm{rr}_{\mathrm{obs}}\right)\right\rangle\right|}{d},$$

$$eL_2 = 2 \times \mathrm{CRPS}\left[P\left(\frac{r_{\mathrm{mod}}}{d}\right), P\left(\frac{r_{\mathrm{obs}}}{d}\right)\right],$$

$$eL = eL_1 + eL_2.$$

4) Strategy and results

II. How can the method be adapted to evaluate ensemble forecasts?

Radanovics, et al., 2018 (WAF)

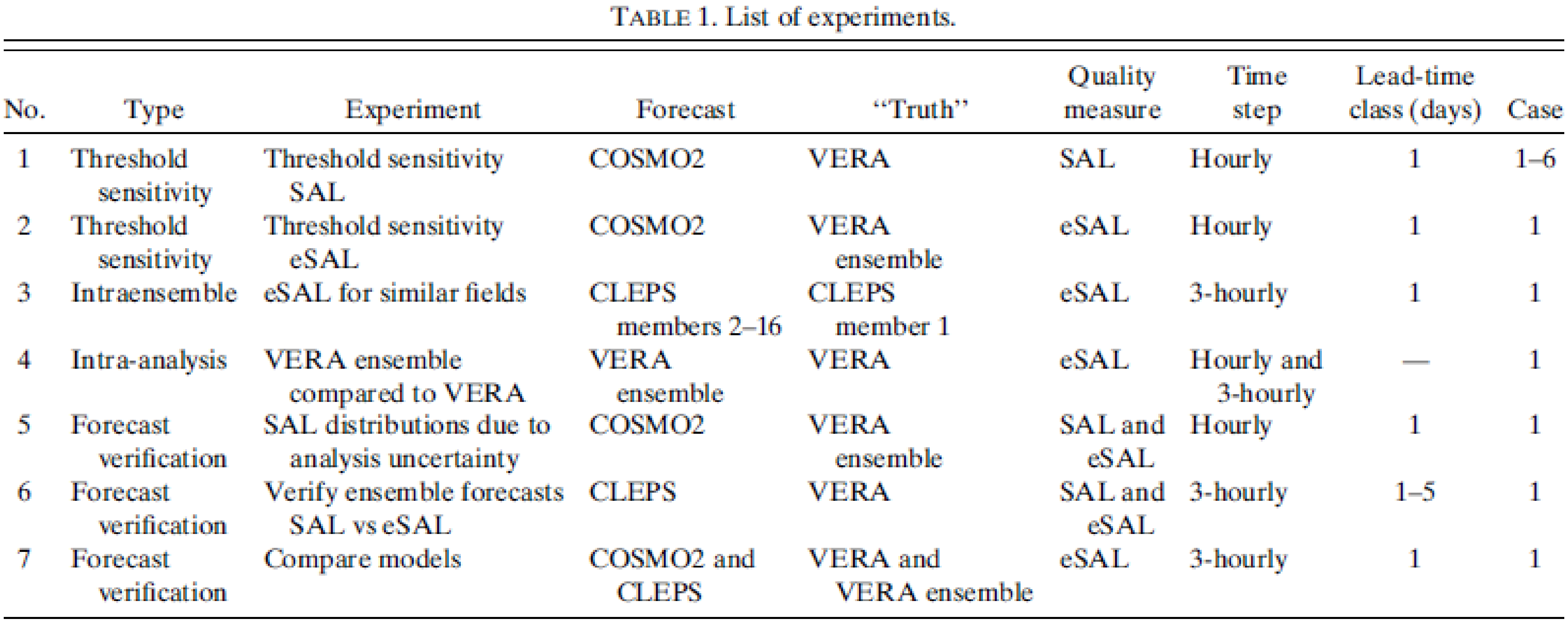


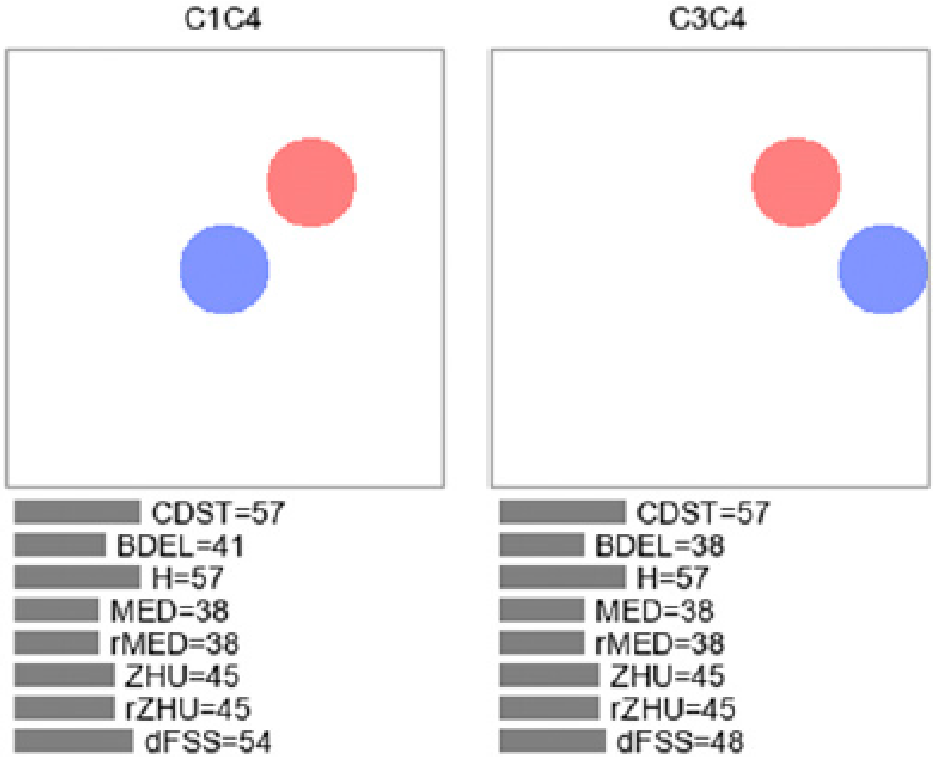
TABLE 1. List of experiments.

No.	Type	Experiment	Forecast	"Truth"	Quality measure	Time step	Lead-time class (days)	Case
1	Threshold sensitivity	Threshold sensitivity SAL	COSMO2	VERA	SAL	Hourly	1	1–6
2	Threshold sensitivity	Threshold sensitivity eSAL	COSMO2	VERA ensemble	eSAL	Hourly	1	1
3	Intraensemble	eSAL for similar fields	CLEPS members 2–16	CLEPS member 1	eSAL	3-hourly	1	1
4	Intra-analysis	VERA ensemble compared to VERA	VERA ensemble	VERA	eSAL	Hourly and 3-hourly	—	1
5	Forecast verification	SAL distributions due to analysis uncertainty	COSMO2	VERA ensemble	SAL and eSAL	Hourly	1	1
6	Forecast verification	Verify ensemble forecasts SAL vs eSAL	CLEPS	VERA	SAL and eSAL	3-hourly	1–5	1
7	Forecast verification	Compare models	COSMO2 and CLEPS	VERA and VERA ensemble	eSAL	3-hourly	1	1

4) Strategy and results

III. How can the method deal with specific geometric cases?

Gilleland, et al., 2020 (MWR) used a novel set of verification test fields (pathological, circular, elliptic and scattered fields) and investigates the performance of distance measures under these different scenarios



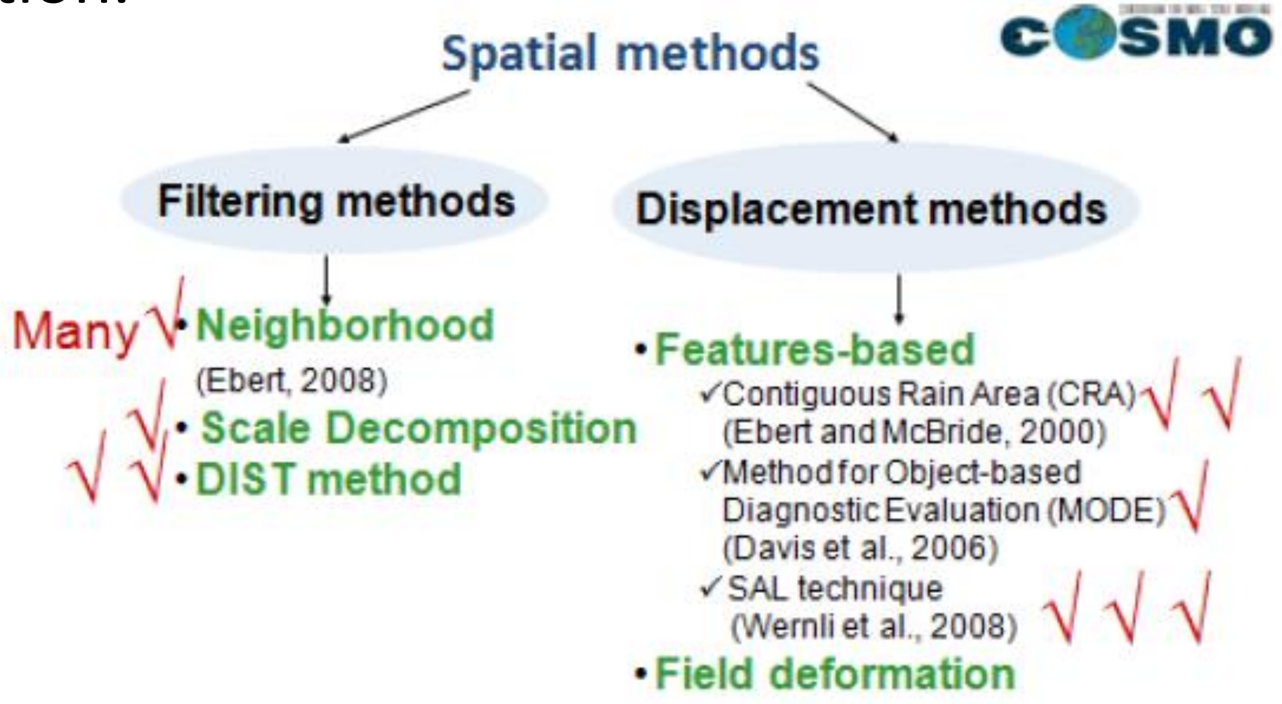
7 pathological cases
12 circular cases
20 elliptical cases
3 scattered cases

The new geometric test cases are included on the website (https://ral.ucar.edu/projects/icp/#NewGeom)

4) Strategy and results

COSMO Priority Project INSPECT includes the MesoVICT-project (COSMO Technical Report No. 37, 2019)

Comprehensive testing of spatial verification methods to provide criteria for the best suited method for a specific application.



Spatial methods

Filtering methods

Displacement methods

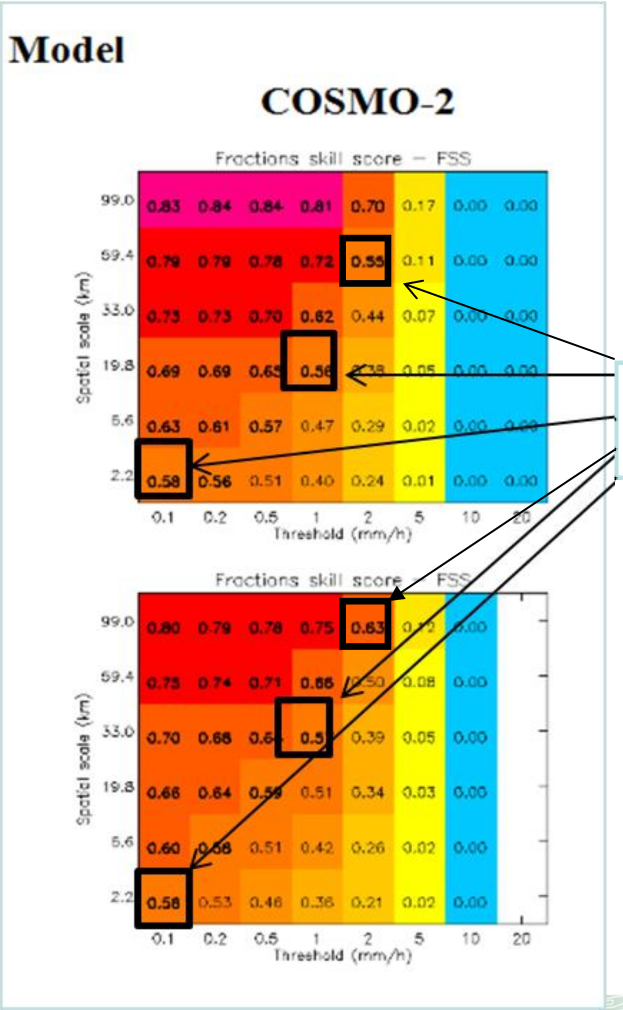
- Many √ Neighborhood (Ebert, 2008)
- √ Scale Decomposition
- √ √ DIST method

- Features-based
  - ✓Contiguous Rain Area (CRA) (Ebert and McBride, 2000) √ √
  - ✓Method for Object-based Diagnostic Evaluation (MODE) (Davis et al., 2006) √
  - ✓SAL technique (Wernli et al., 2008) √ √ √
- Field deformation

4) Strategy and results

Lead time

MAM2015



Model

COSMO-2

Fractions skill score – FSS

Spatial scale (km)	0.1	0.2	0.5	1	2	5	10	20
99.0	0.83	0.84	0.84	0.81	0.70	0.17	0.00	0.00
59.4	0.79	0.79	0.78	0.72	0.55	0.11	0.00	0.00
33.0	0.73	0.73	0.70	0.62	0.44	0.07	0.00	0.00
19.8	0.69	0.69	0.65	0.56	0.38	0.05	0.00	0.00
5.6	0.63	0.61	0.57	0.47	0.29	0.02	0.00	0.00
2.2	0.58	0.56	0.51	0.40	0.24	0.01	0.00	0.00

Threshold (mm/h)

01-12

Fractions skill score – FSS

Spatial scale (km)	0.1	0.2	0.5	1	2	5	10	20
99.0	0.80	0.79	0.78	0.75	0.63	0.12	0.00	
59.4	0.75	0.74	0.71	0.66	0.50	0.08	0.00	
33.0	0.70	0.68	0.6	0.5	0.39	0.05	0.00	
19.8	0.66	0.64	0.59	0.51	0.34	0.03	0.00	
5.6	0.60	0.58	0.51	0.42	0.26	0.02	0.00	
2.2	0.58	0.53	0.46	0.38	0.21	0.02	0.00	

Threshold (mm/h)

13-24

Useful scale as a function of lead time

FSS: Fractions Skill Score

P is the event fraction in the neighborhood.

$$FSS = 1 - \frac{\frac{1}{N}\sum_{N}(P_f - P_o)^2}{\frac{1}{N}\left[\sum_{N} P_f^2 + \sum_{N} P_o^2\right]}$$

Threshold **Lead time**	**0.1 mm/h**	**1mm/h**	**2 mm/h**
01-12	2.2 km	19.8 km	59.4 km
13-24	2.2 km	33.0 km	99.0 km

Thus, the main benefit of INSPECT *"that the wide range of spatial verification methods will become commonly used within the COSMO community"* is achieved.

5) Participants

NCAR, USA, E Gilleland et al.: testing and refining the software SpatialVx; website, geometric case studies.

University of Vienna, M. Dorninger, Simon Kloiber: VERA and JDC data, VERA ensemble, observation uncertainty

COSMO Priority project: INSPECT, A . Bundel, F. Gofa et al.

MeteoSwiss, high resolution model re-runs including forecast ensemble (COSMO-1, COSMO-2 and COSMO-E)

University of Ljubljana, G. Skok: FSS adapted for wind

UK MetOffice, UK, M. Mittermaier et al.: model re-runs, FSS

University of Bonn, Germany, P. Friederichs et al.: probabilistic forecasts and observation uncertainty, image warping and wavelet analysis

ISPRA, Italy, Stefano Mariani et al.: BOLAM, MOLOCH re-runs, CRA

Environment Canada, Canada, B. Casati: model re-runs,

CETEMPS, Italy, R. Ferretti et al.: WRF-CETEMPS re-runs

6) Resume

We could make a step forward but there is still a (long) way to go concerning:

- Adding effects of complex terrain to spatial verification methods (E. Gilleland, et al., 2020)
- Apply and adapt the methods for other parameters (Skok and Hladnik, 2018)
- Apply and adapt the methods to ensemble forecasts (Radanovics, et al., 2018)
- Sensitivity of the methods on model domain and horizontal grid (Mariani and Casaioli, 2018)
- Testing and establishing spatial verification methods at NWS (COSMO, INSPECT)

BUT the best is yet to come:

All data (observations, VERA, VERAens, D-PHASE models and model re-runs, geometric cases) are still available at: https://ral.ucar.edu/projects/icp/ (including SpatialVx)