Environment and Climate Change Canada | Environnement et Changement climatique Canada

Main Outcomes form the Spatial Verification Inter-Comparison and MesoVICT Projects

B. Casati, with contributions from M. Dorninger, M. Mittermaier, B.Brown, G.Skok, B. Ebert

Talk outline:

1. Preamble – what was already boiling in the pot?
2. First Spatial Verification Inter-comparison Project (Vx ICP)
3. MesoVICT (aka second Spatial Vx ICP)
4. Conclusions: What have we learned that can be useful for shaping BtG?

Canada

Preamble
what was already boiling in the pot?

The history
of the Joint Working Group in Verification (JWGV),
of the Spatial Verification (Vx) Inter-Comparison Project (ICP),
and of the International Verification Method Workshops (IVMW)
are intimately related

JWGV origins: 1996 WWRP SSC meeting

In the 1996 WWRP SSC meeting (18-21 November, Toulouse, France) Verification was identified as a key research area to be developed (WMO/TD report No. 812).

Section 5.1: "*A number of key technical capacities will be required to support projects proceeding under the auspices of the WWRP. The SSC will continuously review and monitor these key capacities which include : numerical modelling, data assimilation, observational strategies, verification techniques, measurement instruments and their platforms, and integrated forecast systems*"

Section 5.4: "*Attempts should be made to verify the accuracy of the existing forecast system. Any improvements can then be gauged according to this baseline. For many parameters most relevant to the human experience associated with high impact weather events, an international consensus does not exist on verification techniques. An attempt should be made to improve these techniques and to establish international standards.*"

Section 5.6: "*Many new forecast systems are being developed or suggested for periods ranging from nowcasting to the medium range.[…]. All WWRP RDPs will consider integrated forecast system improvements in their design through presently planned workshops and future meetings. High priority issues for the near term future are forecast verification (with emphasis on high impact weather) and nowcasting systems*"

EBERT ET AL.

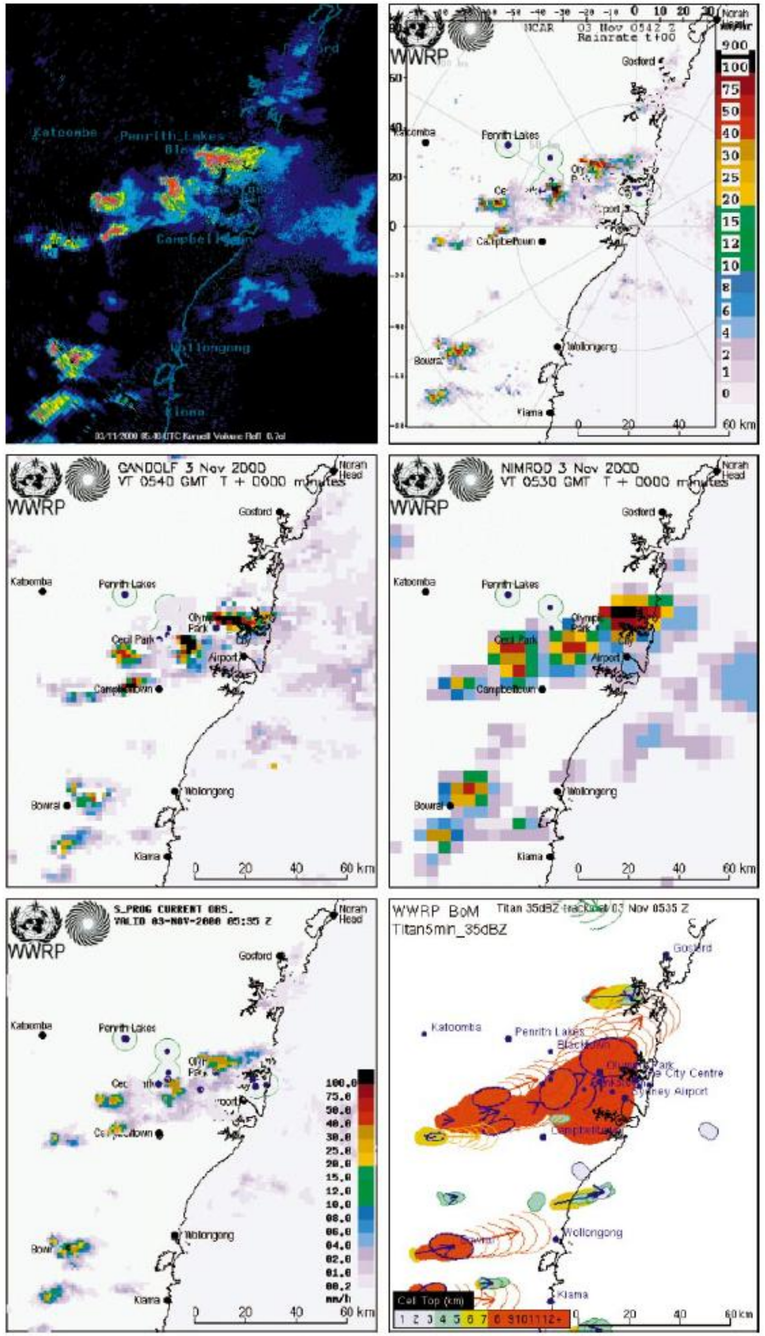
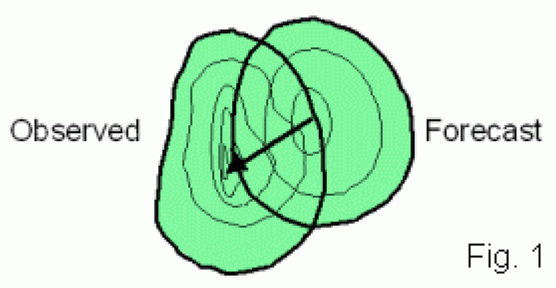


FIG. 1. Radar reflectivity and associated rainfall analyses from five nowcast algorithms at 0530–0540 UTC 3 Nov 2000.

JWGV origins: Sydney 2000 FDP



- Ebert design the Sydney 2000 FDP verification console
- Ebert and McBride (2000) introduce CRA
- First meeting of the to-become JWGV

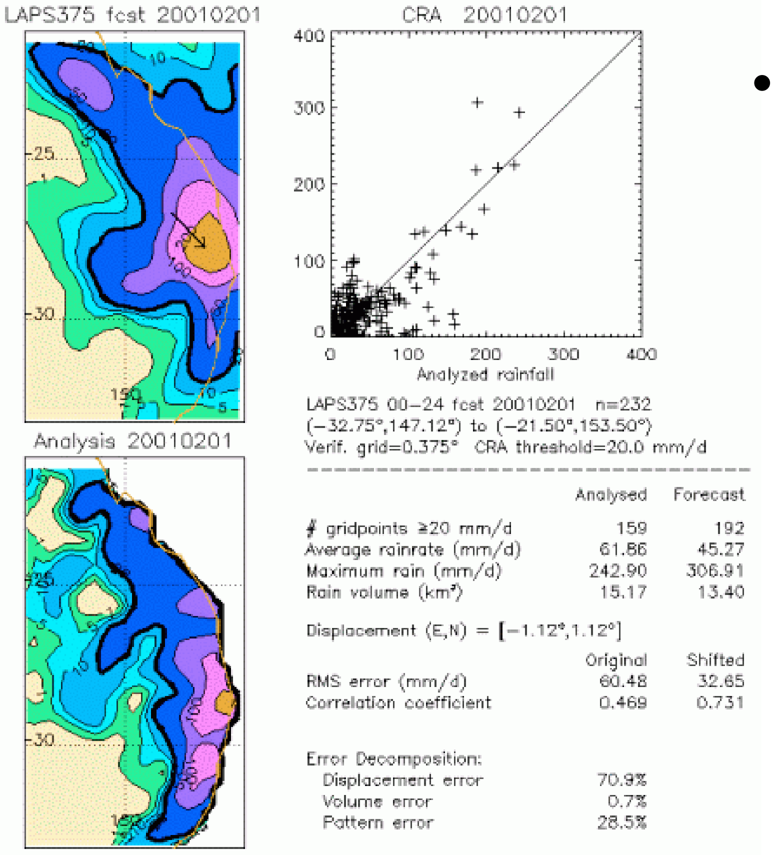
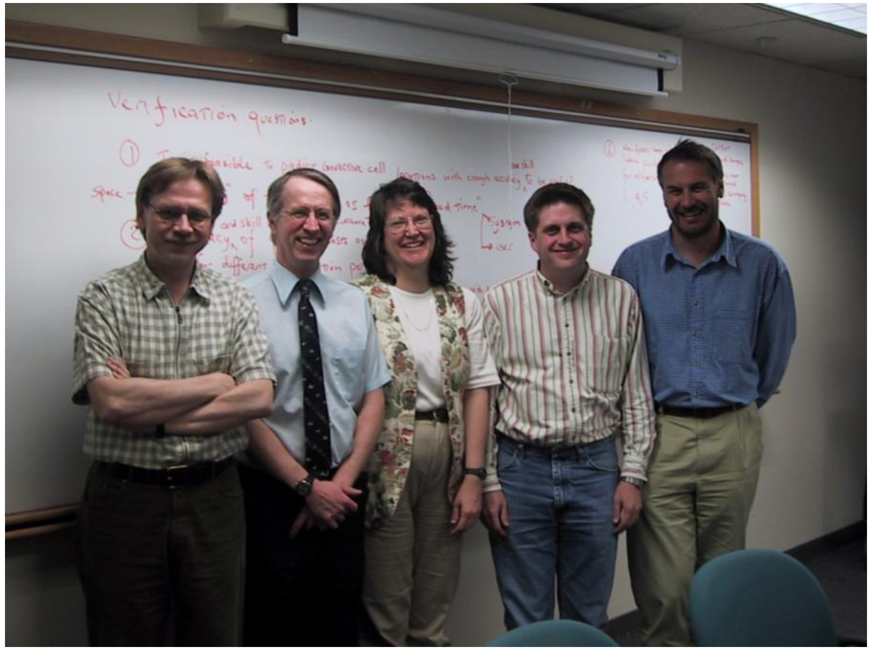


Fig. 2. CRA verification for 24 h rain from the Australian regional model.



2002, JWGV is born

World Weather Research Programme
5th Session of the Science Steering Committee
(Brasilia DF, Brazil, 7-11 November 2002)

Resolution on Verification Working Group

The Committee considered the proposal to form a Working Group on Verification (WGV) under the WWRP. This discussion generated a long debate about the need for guidelines pertaining to the formation of WGs generally, which led to the attached statement of guidelines, many aspects of which your proposal has addressed. The Committee also debated the role of a Verification WG under WWRP considering the traditional home for international verification research has resided under the WGNE.

The Committee is pleased to establish the WWRP WGV with Barbara Brown as its initial Chairwoman. The initial membership shall be as you proposed and the terms of appointment are four years beginning 1 January 2003. The WWRP SSC Liaison to your Group is Rapporteur Radmila Brozkova of the Czech Hydromet Service.

In establishing the WGV, it should be recognized that the WWRP assumes the lead for verification focused on future Forecast Demonstration Projects (FDPs) of the WWRP, oriented to specific groups of "end-users" of high impact weather forecasts. For such efforts to be appropriately targeted, a better specification of what is meant by "user-oriented" verification is needed from the WGV. For example, model developers are a special category of users of verification systems, addressed by the verification development efforts of the WGNE, and having a wide scope.

WWRP FDPs are typically focused on a specific forecast problem involving one or more weather or weather-related hazards at the nowcasting through short ranges of prediction both by non-dynamical and dynamically-based forecast systems. The WGNE will continue to lead verification technique development oriented toward model developers and the broad spectrum of global weather phenomena underlying specific hazards. The purpose of this clarification is to ensure coordination, not to unduly restrict the efforts of either.

While the the WGV is now established solely under the WWRP, serious consideration is being given to a unified approach with the WGNE. In anticipation of a Joint (WWRP-WGNE) Working Group on Verification (JWGV), we ask that consideration be given to broadening the WGV membership over the coming year in such a way as to be inclusive of experts in WGNE verification activities. Coordination with WGNE should be pursued through WGNE Chairman Kamal Puri, BMRC, Australia. The WWRP SSC and the WGNE will review the status of verification activities at the annual meetings in 2003. Two representatives of the WGV will be invited to summarize the current status, and to participate in the discussion related to unification.

ORGANISATION MÉTÉOROLOGIQUE MONDIALE — WORLD METEOROLOGICAL ORGANIZATION

Téléphone: Int'l + 41 (0) 22 730 81 11
Facsimilé: Int'l + 41 (0) 22 730 81 81
Télégrammes: METEOMOND GENEVE
Télex: 41 41 99 OMM CH

SECRÉTARIAT
GENÈVE - SUISSE

7 bis, avenue de la Paix
Case postale No. 2300
CH-1211 Genève 2

Our ref.: 13.220/R/WWRP/SSC-V

Annex: 1

GENEVA, 19 December 2002

Dear Dr Brown,

The Science Steering Committee of the World Weather Research Program held its Fifth Session in Brasilia, Brazil from 7 to 11 November 2002. On behalf of the Committee, I am pleased to communicate a resolution passed by the Committee concerning Guidelines for the Verification Working Group.

If you have any questions or concerns associated with these guidelines, please do not hesitate to contact me or your designated WWRP SSC liaison, Dr Radmila Brozkova. I may be reached via email at carbone@ucar.edu.

I urge you to persevere and to accelerate, wherever possible, your efforts toward improved forecasts, ultimately defined by the enhanced use and value to society.

My best regards to you and your colleagues.

Yours sincerely,

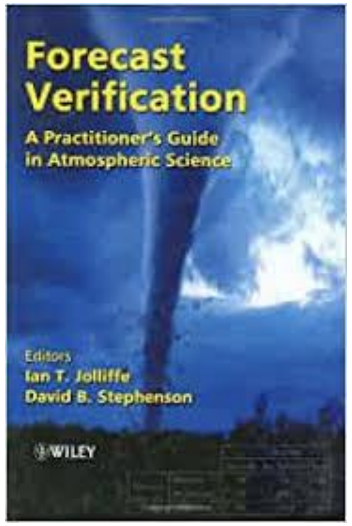
Lei Zhaochong
for (Richard Carbone)
Chairman, WWRP Science Steering Committee

cc: Beth Ebert, Bureau of Meteorology Research Centre, Melbourne, Australia
Laurie Wilson, Research en Prevision Numerique, Dorval, Quebec, Canada
Pertti Nurmi, Finnish Meteorological Institute, Hensinki, Finland
Harold Brooks, National Severe Storms Laboratory, Norman, Oklahoma
Barbara Casati, University of Reading, UK
Chairman, WGNE

How the Spatial Vx ICP started?

- 1996 WWRP SSC recognize the need of a verification group
- Sydney 2000 Olympic games – FDP
- WWRP establishes JWGV - 2002

- 1st IVMW, Boulder 2002
- Jolliffe and Stephenson (2003) book
- 2nd IVMW, Montreal 2004



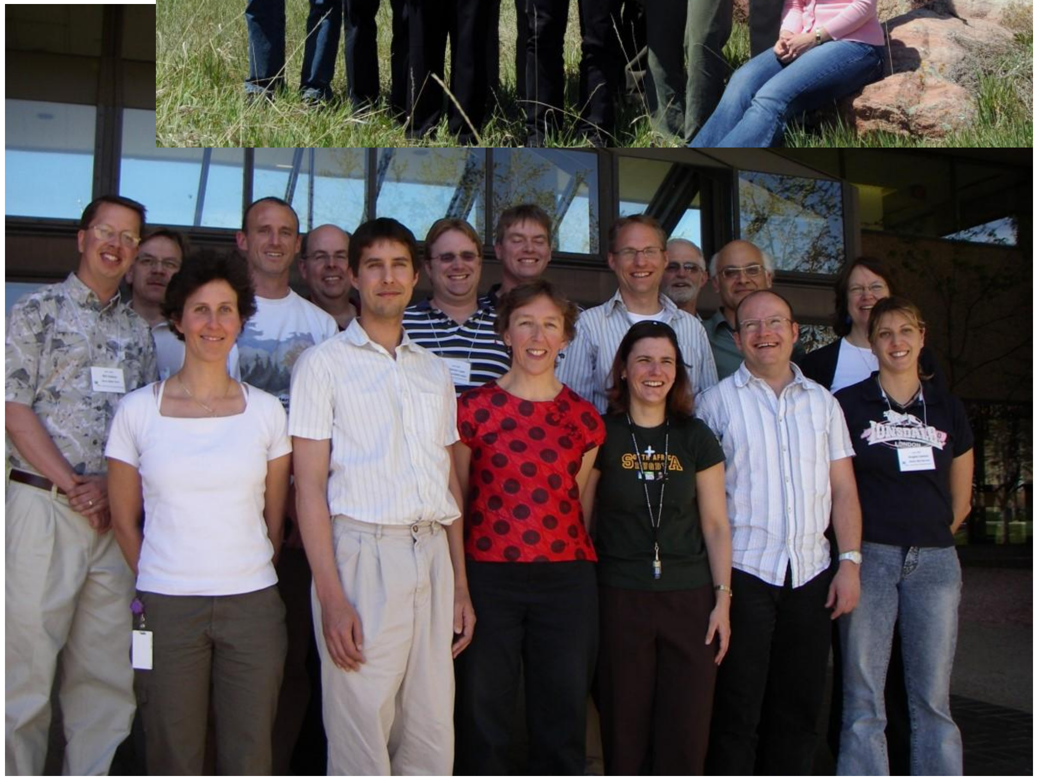
How the Spatial Vx ICP started?

- 1996 WWRP SSC recognize the need of a verification group
- Sydney 2000 Olympic games – FDP
- WWRP establishes JWGV - 2002

- 1st IVMW, Boulder 2002
- Jolliffe and Stephenson (2003) book
- 2nd IVMW, Montreal 2004

- QPF workshop, Boulder 2006
- 3rd IVMW, Reading 2007
- JWGV meeting, Boulder 2008

NCAR: the Spatial Verification (Vx) Inter-Comparison Project (ICP) is established



How the Spatial Vx ICP started? The key elements:

- NWP models began to increase considerably in resolution
- gridpoint-by-gridpoint verification not adequate or sufficient:
  - ➢ Double penalties (e.g., Mass et al. 2002): higher-resolution forecasts exhibit poorer performance than their coarser counterparts, even when considered more realistic in subjective verification.
  - ➢ Not informative in physical terms: displacement? Scale structure? Features? …
- New spatial verification methods were developed very rapidly:
  - ❖ Ebert and McBride, 2000 (CRA); Harris et al. 2001 (wavelets); Baldwin and Lakshmivarahan, 2003 (event-oriented); Casati et al. 2004 (Intensity-Scale); Nachamkin, 2004 (composites); Theis et al. 2005 (pragmatical approach); Marsigli et al. 2008 (distribution); Baldwin and Kain, 2006 (bias-adj TS); Marzban & Sandgathe, 2006 (cluster analysis); Davis et al. 2006 (MODE); Keil and Craig, 2007 (DAS optical flow); Roberts and Lean 2008 (Fraction Skill Score); …

- WMO support: A series of meetings and workshops, leading to the creation of an international verification community
- NCAR leadership – Barb Brown and the RAL group: verification research HUB + IT capacity leading to the development of open-access verification software

The first Spatial Verification Inter-Comparison Project (Vx ICP)

The first Spatial Vx ICP

The first spatial verification methods intercomparison project (ICP; Gilleland et al 2009, Wea&For **24**; Gilleland et al 2010 BAMS) is a meta-verification exercise initiated in 2007 with the aim of better understanding the rapidly increasing literature concerning new spatial verification methods: **analyse the behaviour** and **classify the spatial methods**

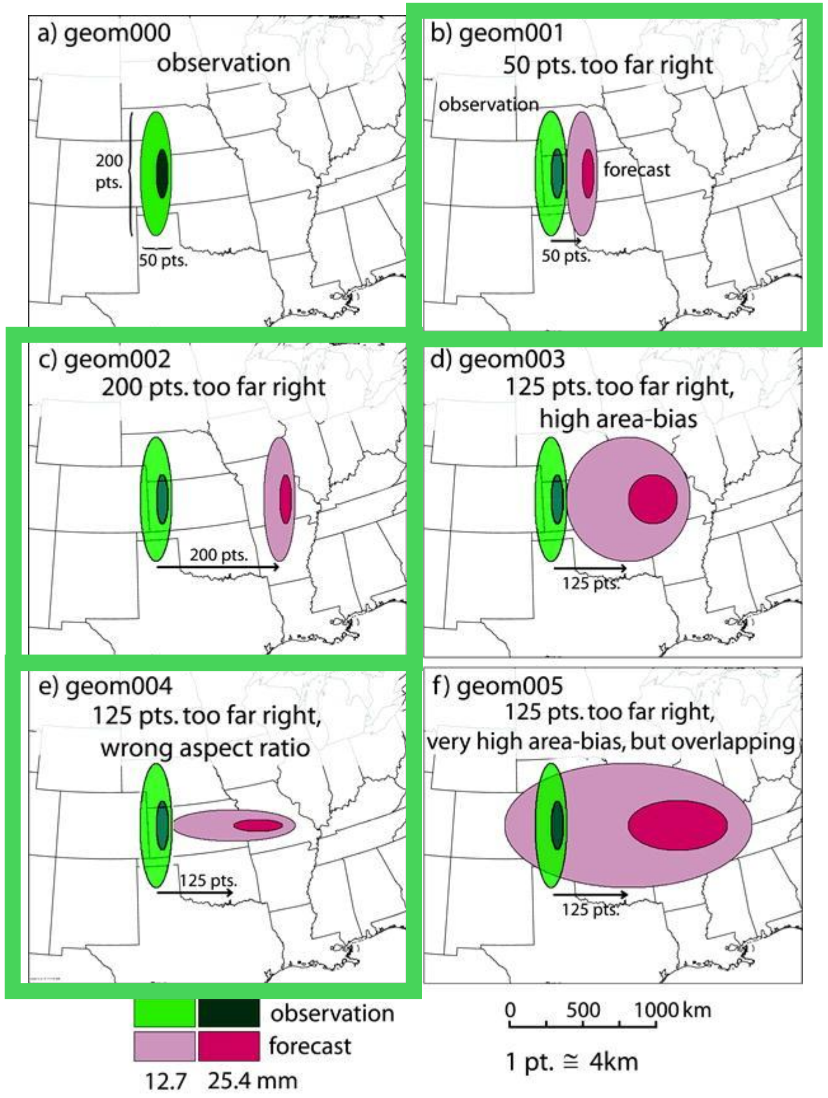
Questions addressed:

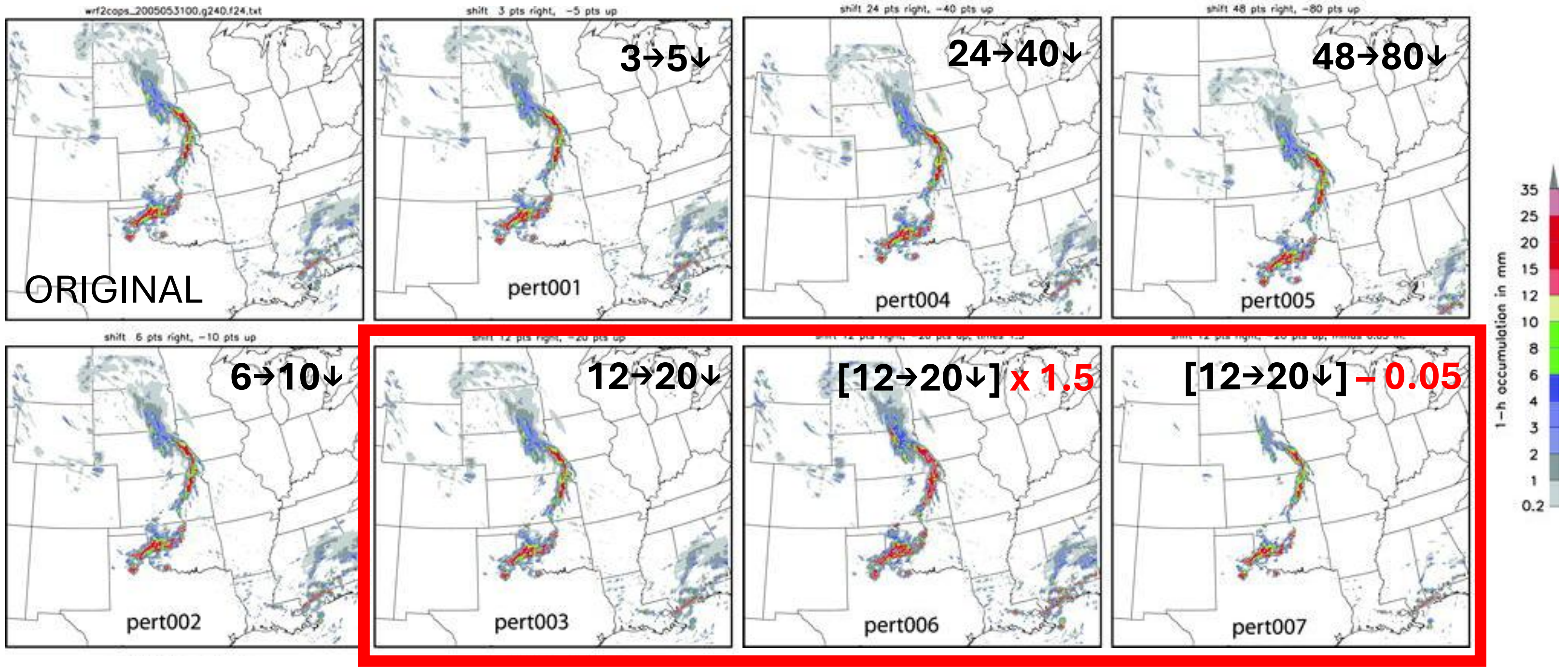
- How does each method inform about forecast performance overall?
- Which aspects of forecast error does each method best identify (e.g., location error, scale dependence of the skill, etc.)?
- Which methods yield identical information to each other, and which methods provide complementary information?

Aim to provide guidance to verification practitioner on which method to use (similarly to WP-MIP)

1st spatial Vx ICP – Common case studies

Ahijevych, D., E. Gilleland, B. Brown, and E. Ebert, 2009: Application of spatial forecast verification methods to gridded precipitation forecasts. *Wea. Forecasting*, **24**, 1485–1497, https://doi.org/10.1175/2009WAF2222298.1.





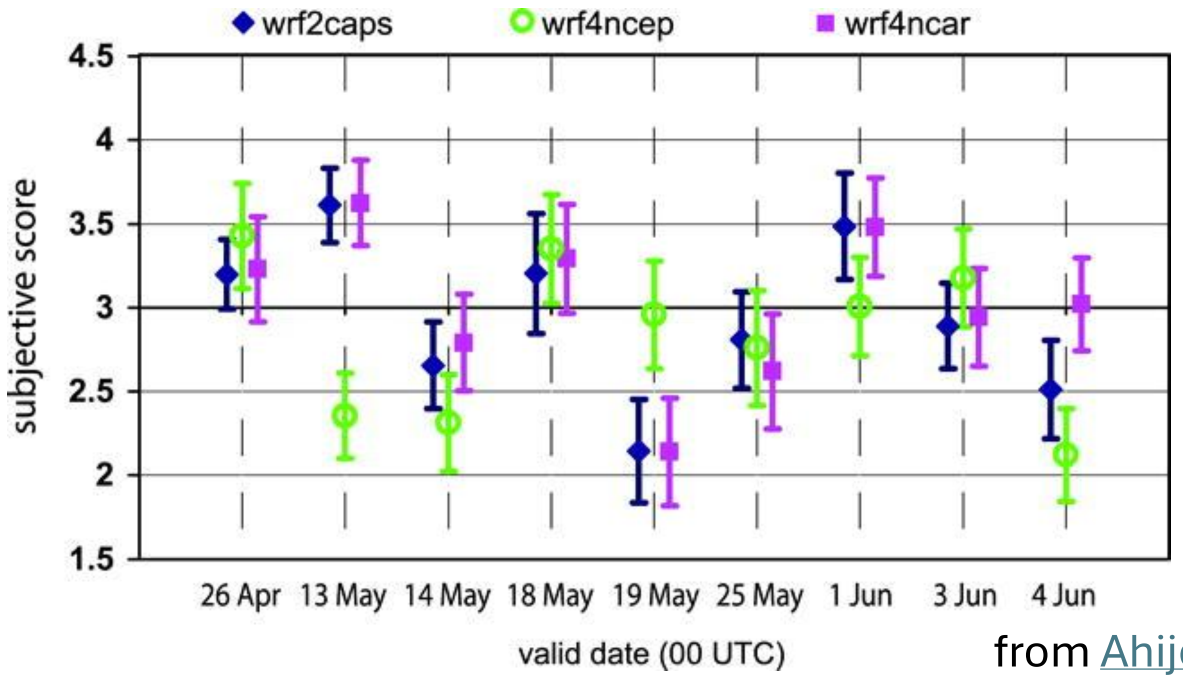
Focus on precipitation:

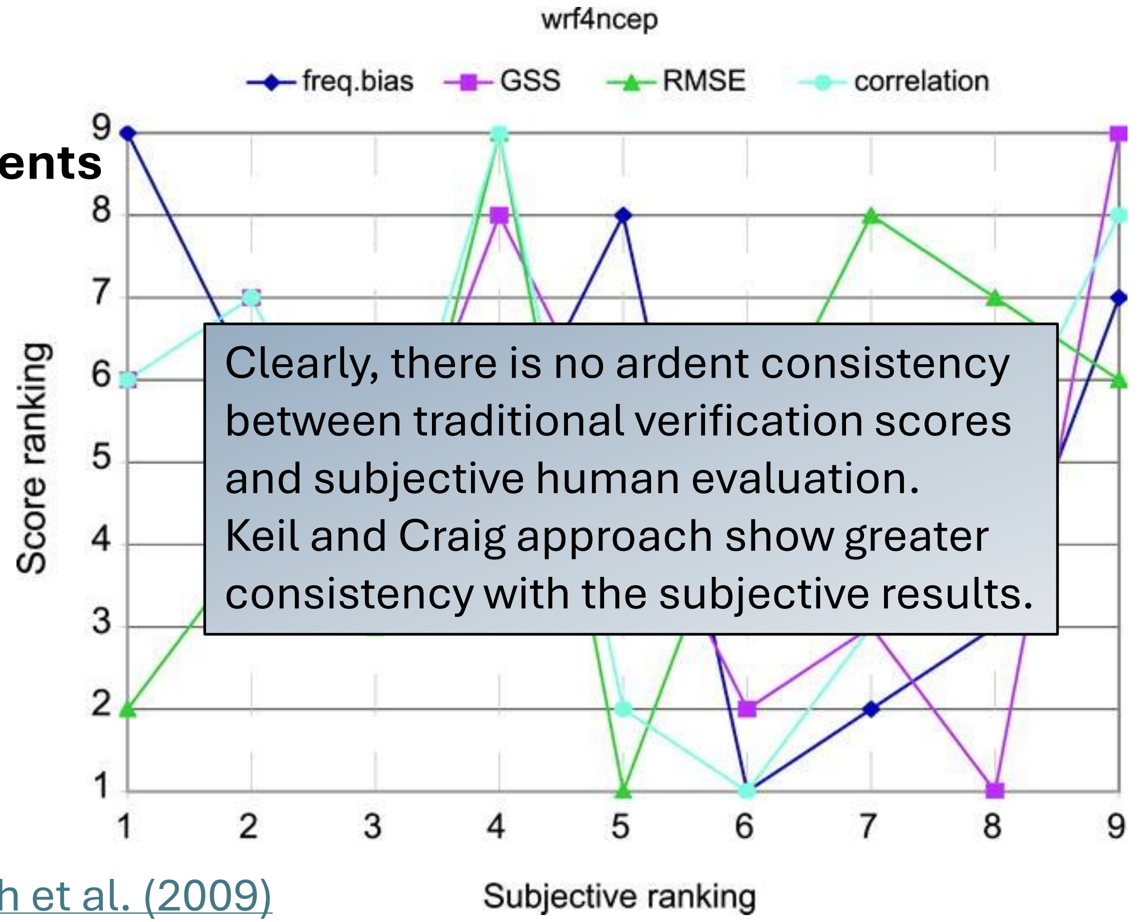
1. Synthetic circular/elliptical **geometric cases**
2. Precipitation forecasts **perturbed with prescribed errors**
3. Nine selected real-data storms extracted from the **SPC/NSSL Spring Program 2005**

SPC/NSSL Spring Program 2005

The spatial verification methods were tested on nine selected case studies (provided by Mike Baldwin) from the 2005 National Severe Storms Laboratory (NSSL) and Storm Prediction Center (SPC) Spring Experiment (Kain et al. 2008, *Wea&For* vol **23**):

- convection-permitting 4–7-km NWP: **wrf4ncar**, **wrf4ncep**, **wrf2caps**
- flat terrain, the Great Plains of the United States
- against Stage II precipitation analyses
- results were compared to **human subjective assessments**

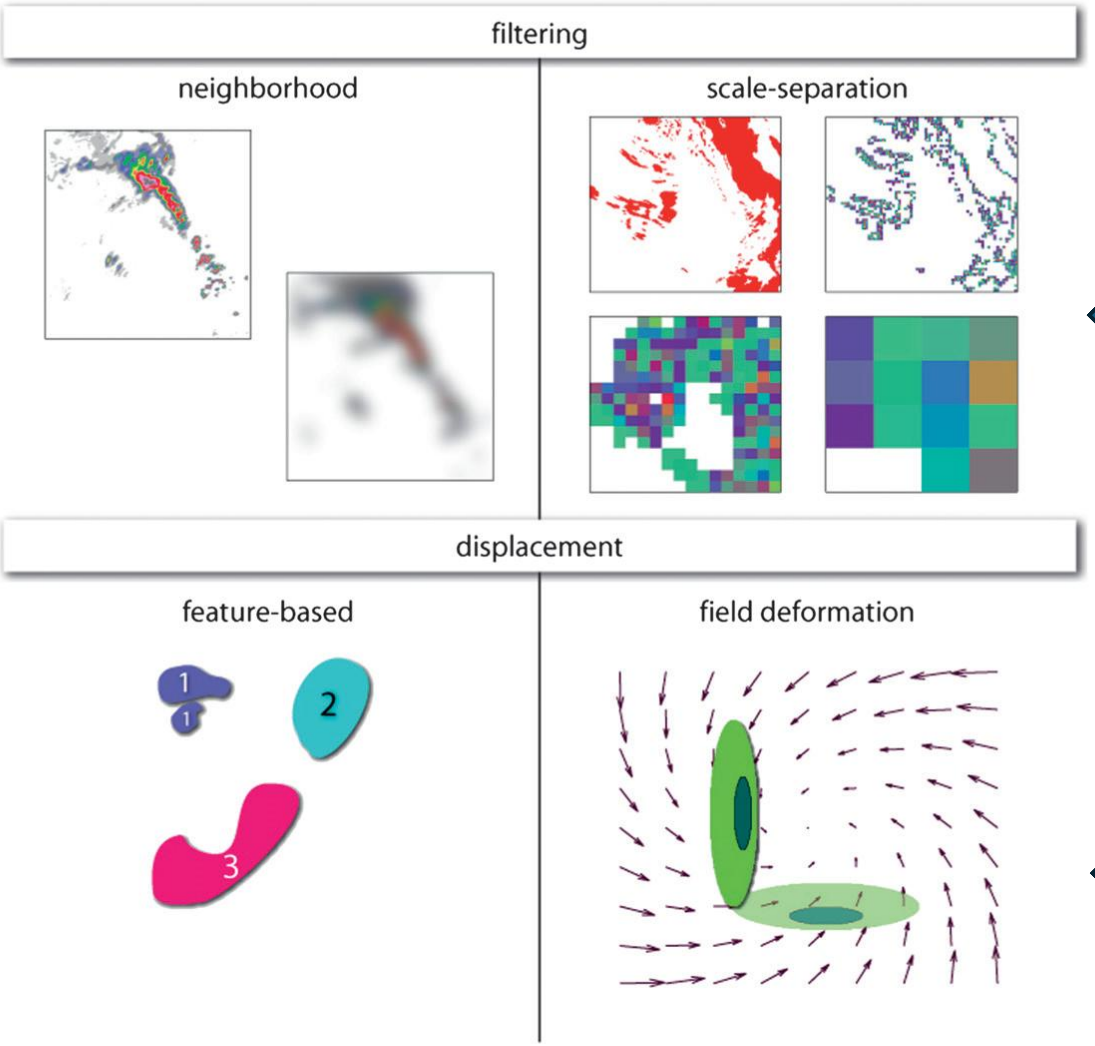


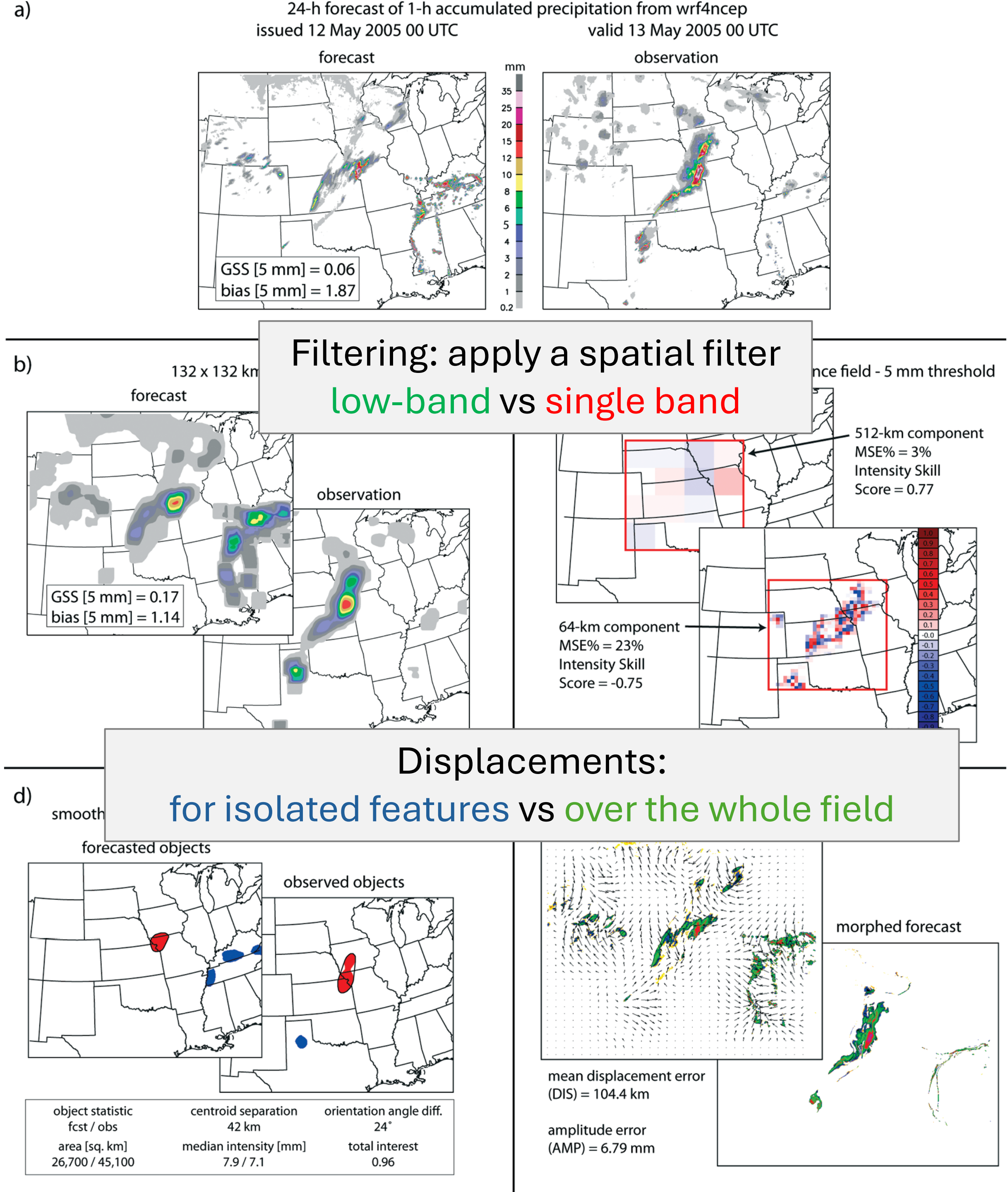


Clearly, there is no ardent consistency between traditional verification scores and subjective human evaluation. Keil and Craig approach show greater consistency with the subjective results.

from Ahijevych et al. (2009)

1st Spatial Vx ICP Classification





Filtering: apply a spatial filter
low-band vs single band

Displacements:
for isolated features vs over the whole field

From Gilleland et al 2009; Gilleland et al 2010 BAMS

Driving questions → guidance to use (and further develop / improve) spatial verification methods …

Questions addressed for each approach (from Gilleland et al 2009, 2010):

1. How do the methods inform about performance at different **scales**?
2. How do the methods provide information on **location** errors?
3. Can the methods be used to identify **timing** errors?
4. Do the methods provide information on **intensity** errors and distributions?
5. Do the methods provide information on **structure** errors?
6. Do the approaches have the ability to provide information about **hits, misses, false alarms, and correct negatives**?

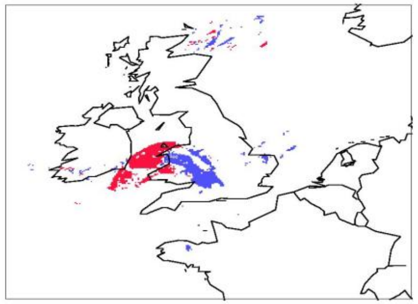
Abbreviation	Scales?	Location errors?	Intensity errors?	Structure errors?	Hits, misses, false alarms, and correct negatives?
BCETS	No*	Indirectly	Yes	No	Yes
CA					
Procrustes					
Procrustes2					
Composite					
CRA					
DIST					
FQI					
FQM–DAS					
FSS					
IS					
IW					
MSV					
MODE	Yes (see text)	Yes	Yes	Yes	Yes (see text)
Neighborhood	Yes (see text)	Indirectly	Yes	No	Yes
SAL	No*	Yes	Yes	No	No
Traditional	No*	No	Yes	No	Yes
VGM	Yes	No (see text)	No (see text)	Yes*	No

TABLE 1. Summary of the types of information that the general categories of methods can provide. Entries should be interpreted as being a general characteristic of the method type, and it should be understood that there are exceptions among individual methods; see the text for more information.

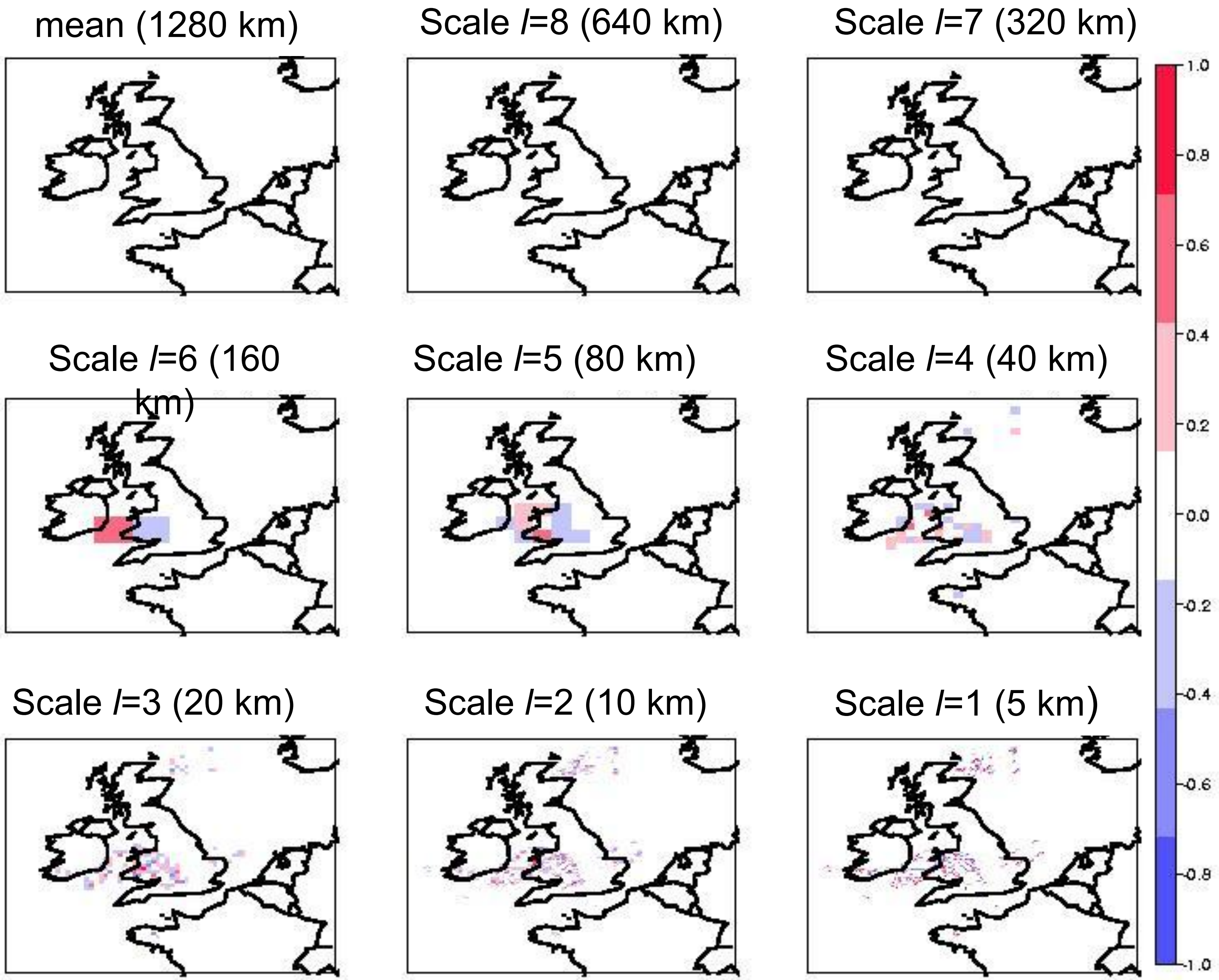
Category	Scales with skill	Location errors	Intensity errors	Structure errors	Occurrence (hits, misses, and false alarms)
Neighborhood	Yes	No	Yes	No	Yes
Scale separation	Yes	No	Yes	No	Yes
Features based	No	Yes	Yes	Yes	Yes
Deformation	No	Yes	Yes	No	No

7. Do the methods do anything that is **counterintuitive**?
8. Do the methods have **selectable parameters** and how sensitive are the results to parameter choice?
9. Can the results be easily **aggregated** across multiple cases?
10. Can **confidence intervals** or hypothesis tests be readily computed for the method?

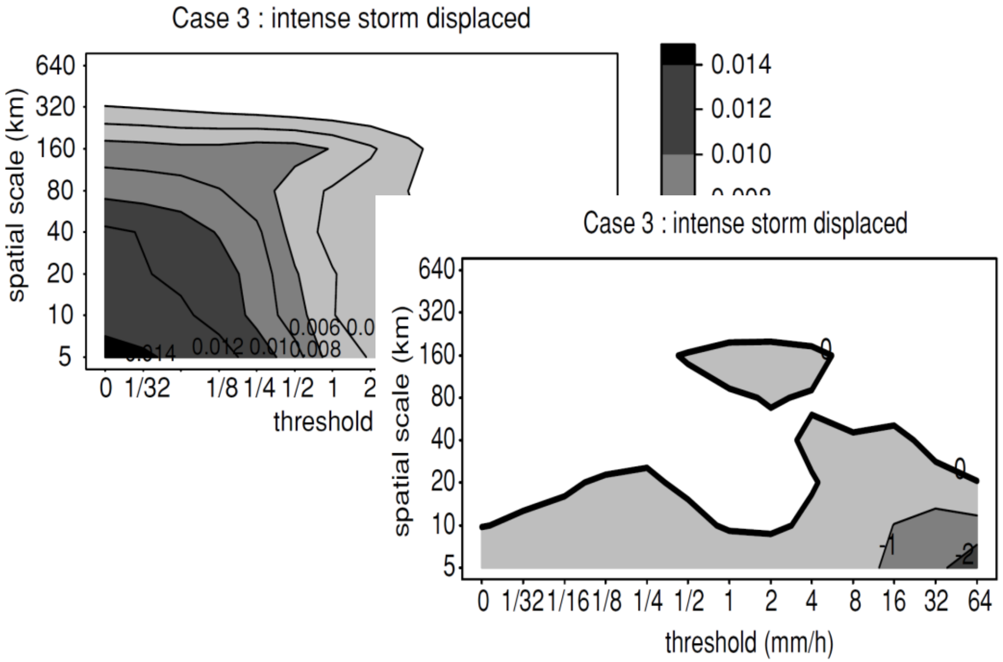
Before 1st SpVx ICP: Intensity-scale error and skill score (Casati et al. 2004)



= ∑

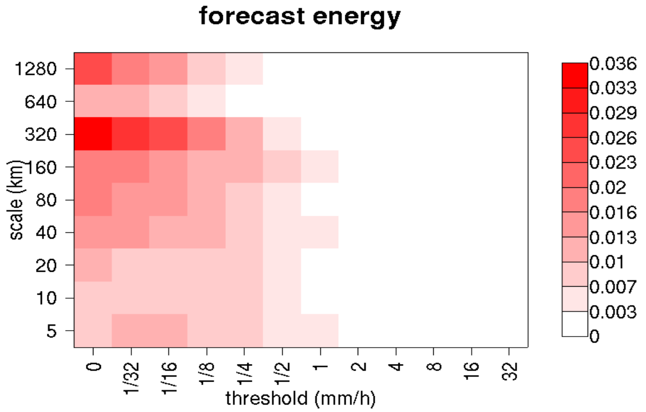


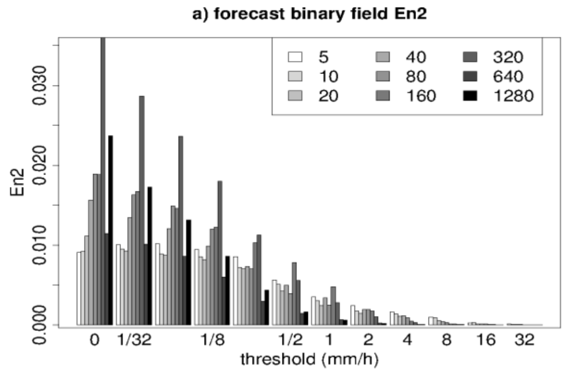
- **intensity**: precipitation thresholding
- **scale**: Haar wavelet decomposition
- **skill score** ↔ categorical approach

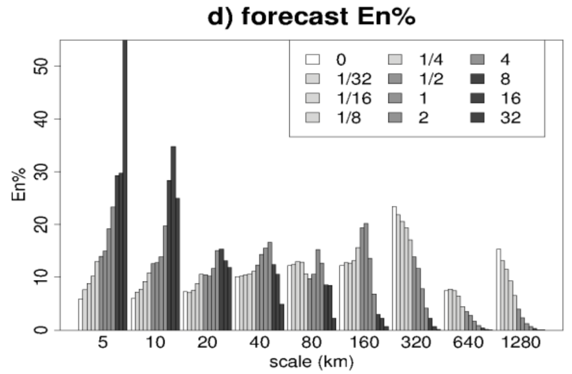


2D Haar wavelet decomposition of binary error image

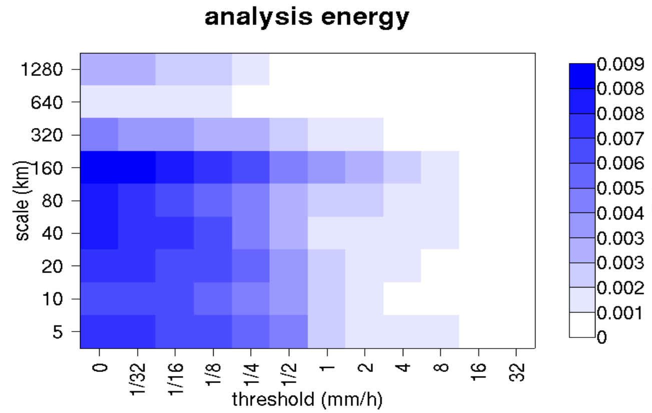
1stSpVx ICP innovation: Intensity-scale energy: bias and scale structure (Casati, 2010)

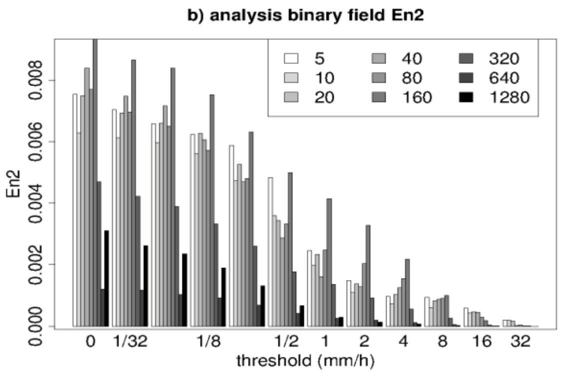


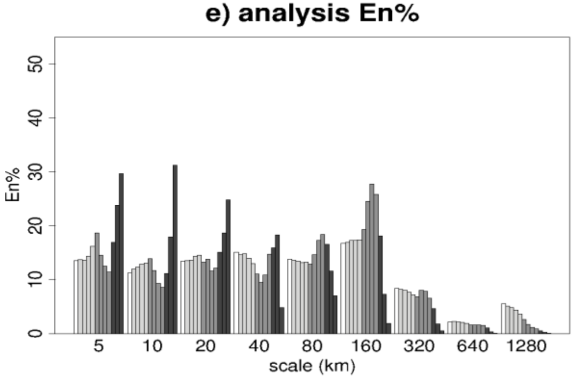




The energy *(En2)* or a field is the average of the squared values of the field grid-points.

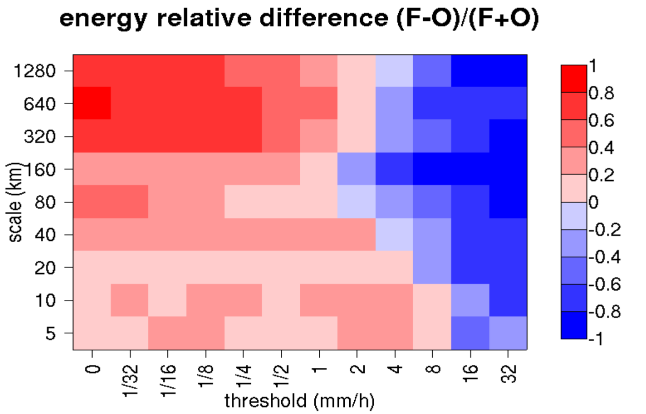


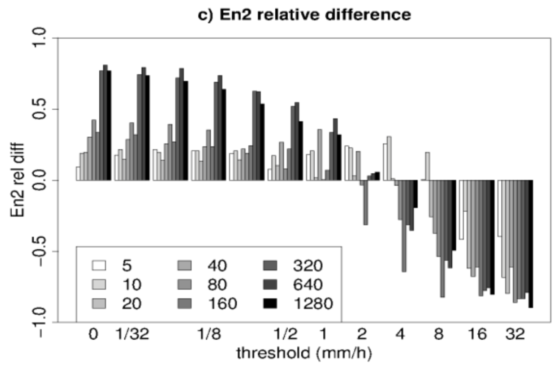




$$En_u^2 = \sum_{l=1}^{L} En_{u,l}^2$$

$$En\%_{u,l}^2 = {En_{u,l}^2}/{En_u^2}$$





<u>For each threshold and scale:</u>

- **energy** informs on the amount of events;
- energy relative difference measures the **bias**;
- energy % assess the **scale structure**.

1st Spatial Vx ICP publication outcomes

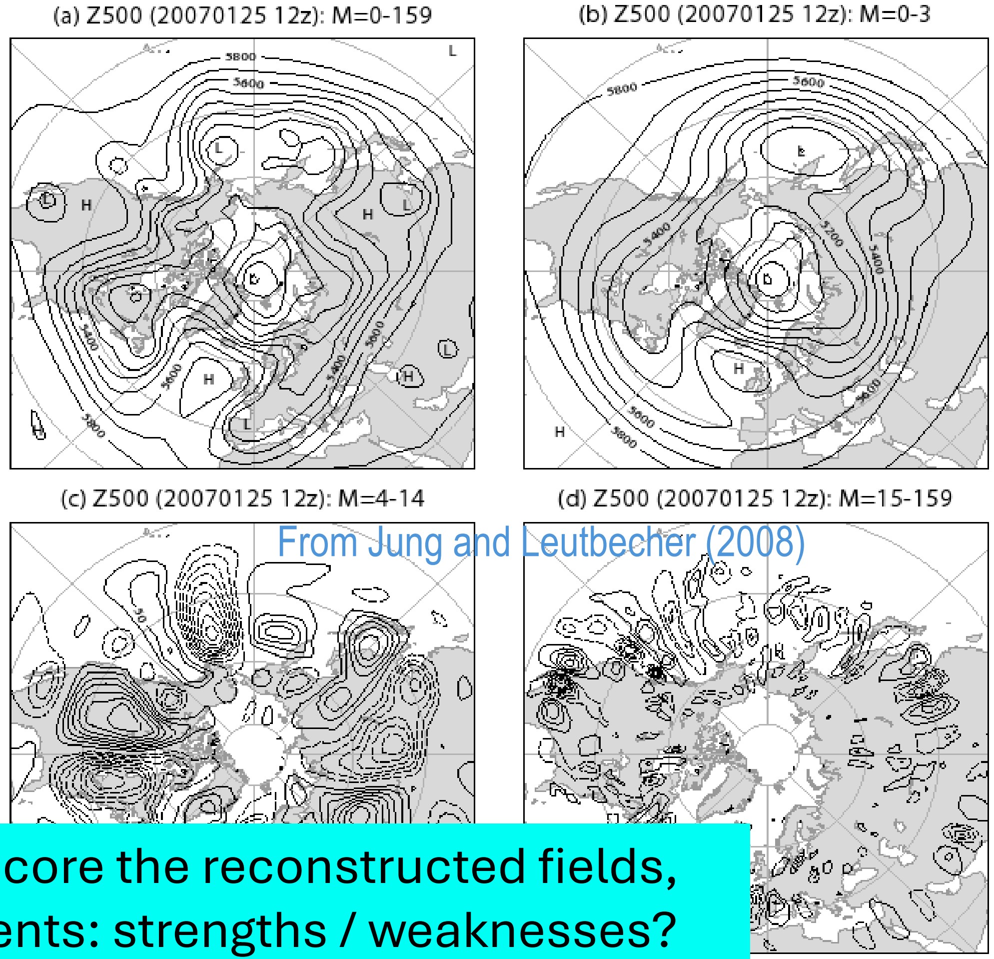
Results for specific methods from the first intercomparison of spatial verification methods can be found in the *Weather and Forecasting (vol.24, 2009)* Special Collection: Spatial Forecast Verification Methods Inter-Comparison Project (ICP)

- Wernli (2008) (SAL) → Wernli et al. 2009;
- Davis et al. 2006 (MODE, timing and aggregation) → Davis et al. 2009 (sensitivity to tuning parameters);
- Nachamkin 2004 (composites) → Nachamkin 2009;
- Casati et al. 2004 (Intensity-Scale) → Casati 2010 (En2 bias and scale structure);
- Micheas et al 2007, Lack 2007 (procrustes) → Lack et al. 2010 (multi-scale procrustes);
- Ebert and McBride 2000 (CRA) → Grams et al. 2006 (MSE = corr)→ Ebert and Gallus 2009;
- Ebert 2008 (neighbourhood framework) → Ebert 2009 (the double definition of scale)
- Roberts and Lean 2008 (Fraction Skill Score) → Mittermaier and Roberts, 2010
- Keil and Craig 2007 (DAS optical flow) → Keil and Craig 2009;
- Gilleland et al. 2010 (image warping).
- Marzban & Sandgathe, 2006 (cluster analysis) → Marzban et al. 2009 (cluster analysis + variograms + optical flow)
  → Marzban and Sandgathe 2009 (variograms) Marzban and Sandgathe 2010 (optical flow)
- Laksmanan and Kain 2010 (Gaussian Mixture Model – cluster method)
- Gilleland et al. 2008 (Baddeley for MODE) → Gilleland, 2011 (Baddeley)
- Baldwin and Kain (2006) → Schwedler and Baldwin 2011 (bin. img. Distance metrics)
- Brill 2009 (analytical method), Brill and Mesinger 2009 (bias-adjusted TS);

1. Scale-separation verification techniques

Rationale: weather phenomena of different spatial scales (e.g. fronts vs convective cells) are governed by different physical processes. Scale separation verification method can help identifying which processes need improvement.

1. Decompose forecast and obs. fields on different scales (wavelets, discrete cos transforms, Fourier, spherical harmonics)
2. Perform verification on different scale components, separately (continuous and categorical scores, probabilistic scores)



From Jung and Leutbecher (2008)

- assess bias, error and skill on different spatial scales
- assess forecast vs observed scale structure
- scale dependency of forecast predi
  skill transition scale)

Note: some techniques score the reconstructed fields, others score the coefficients: strengths / weaknesses?

Some selected References: Stamus et al (1992); **Briggs & Levine (1997)** wavelets+denoising; **Casati et al (2004),Casati (2010)** wavelets+ categorcal; **Casati & Wilson (2007)** wavelets+probability; **Casati et al (2023)** wavelets+continuous; **Denis et al (2002, 2003), De Elia et al (2002),** cosine discrete transform+continuous; **Zepeda -Arce et al (2000), Harris et al (2001), Tustison et al (2003)** wavelets+scale structure; **Jung and Leutbecher (2008)** sph-harmonics+ensembles; **Livina et al (2008)** wavelet+continuous; **Weniger (2017), Buschow and Friederichs (2020, 2021)** SAD: scale, anisotropy and direction, **Buschow (2022)** displacement, complex wavelets+continuous; **Husain et al (2025)**, sph-harmonics + continuous. **Rodwell (2025)** sph-harm ens

2. Neighbourhood verification

Ebert (2008) "*Fuzzy verification of high resolution gridded forecasts: **a review and proposed framework***", Met Applications, vol **15**, 51-64 pp

Some selected References:
Atger (2001), spatial multi-event ROC curve; **Yates (2006),** upscaling, cont&cat scores; **Rezacova and Sokol (2005)**, rank RMSE. **Theis et al (2005)**; pragmatical approach; **Marsigli et al (2006, 2008)** probabilistic approach.
Roberts and Lean (2008) Fraction Skill Score;
Ebert (2008,2009) framework and comparison**;**
Duc et al (2013) ensembles+time;
BenBoualegue and Theis (2014), ensembles;
Nachamkin and Schmidt (2014) quilt-plots;
Mittermaier (2014), Mittermaier and Csima (2017)
Schwartz (2017), categorical; **Schwartz and Sobash (2017),** ensembles; **Stein and Stoop (2019),** categorical contingency tables; **Stein and Stoop (2022)** CRPS; **Stein and Stoop (2024)** Brier Divergence Skill Score. **Necker et al (2024)** ensembles.

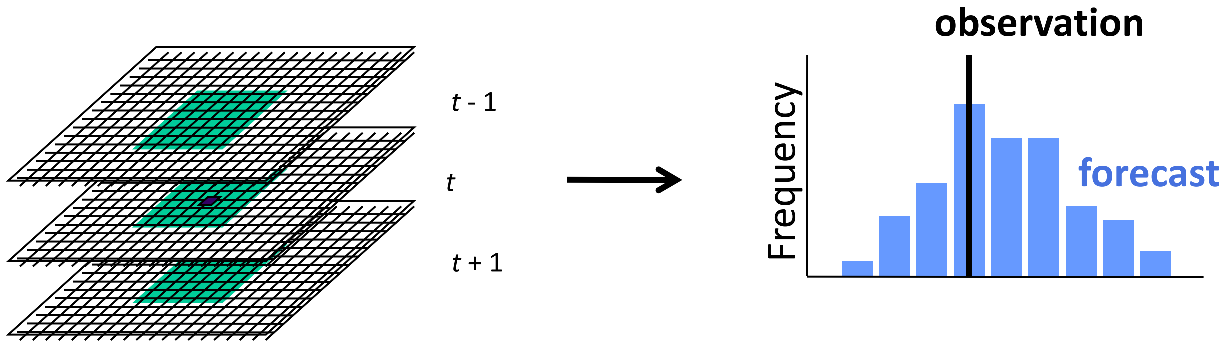
Motivation: enhanced resolution NWP systems produce precipitation fields enriched of realistic small-scale features. Traditional (point-to-point) verification approaches fail to detect the added value of enhanced resolution, possibly due to the higher variability and small space-time displacements (double penalties).

2. Neighbourhood verification: Ebert 2008 framework

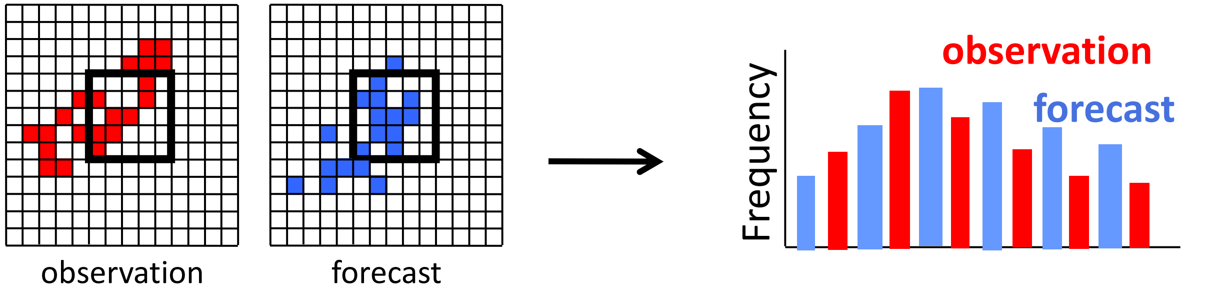
Step 1. Define neighbourhood of grid-points:

Relax requirements for exact positioning in space and time within the neighborhood, while matching forecast and observations (reward "close" forecast, mitigating double penalty: suitable for high resolution models); account for forecast and observation small time-space uncertainty (unpredictable scales).

SO-NF
Single Obs - Neighbourhood of Forecast



NO-NF
Neighbourhood Obs - Neighbourhood For



Step 2. Perform verification over neighbourhoods of different sizes

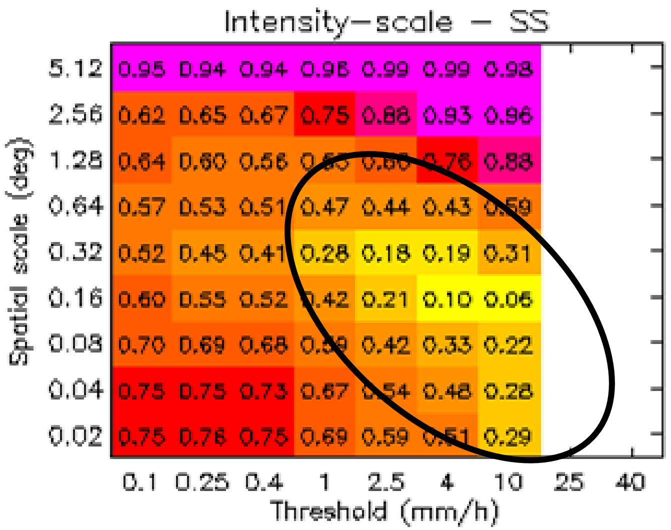
- up-scaling / smoothing; verify using traditional continuous & categorical scores within neighbourhoods;
- verify event occurrence / frequencies within the neighbourhood; verify distribution attributes (e.g. quantiles)
- **verify a deterministic forecast with a probabilistic approach.**

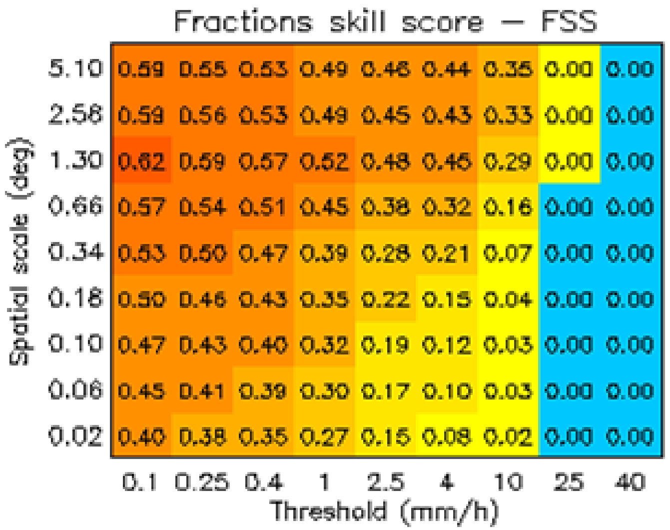
The double definition of "scale"

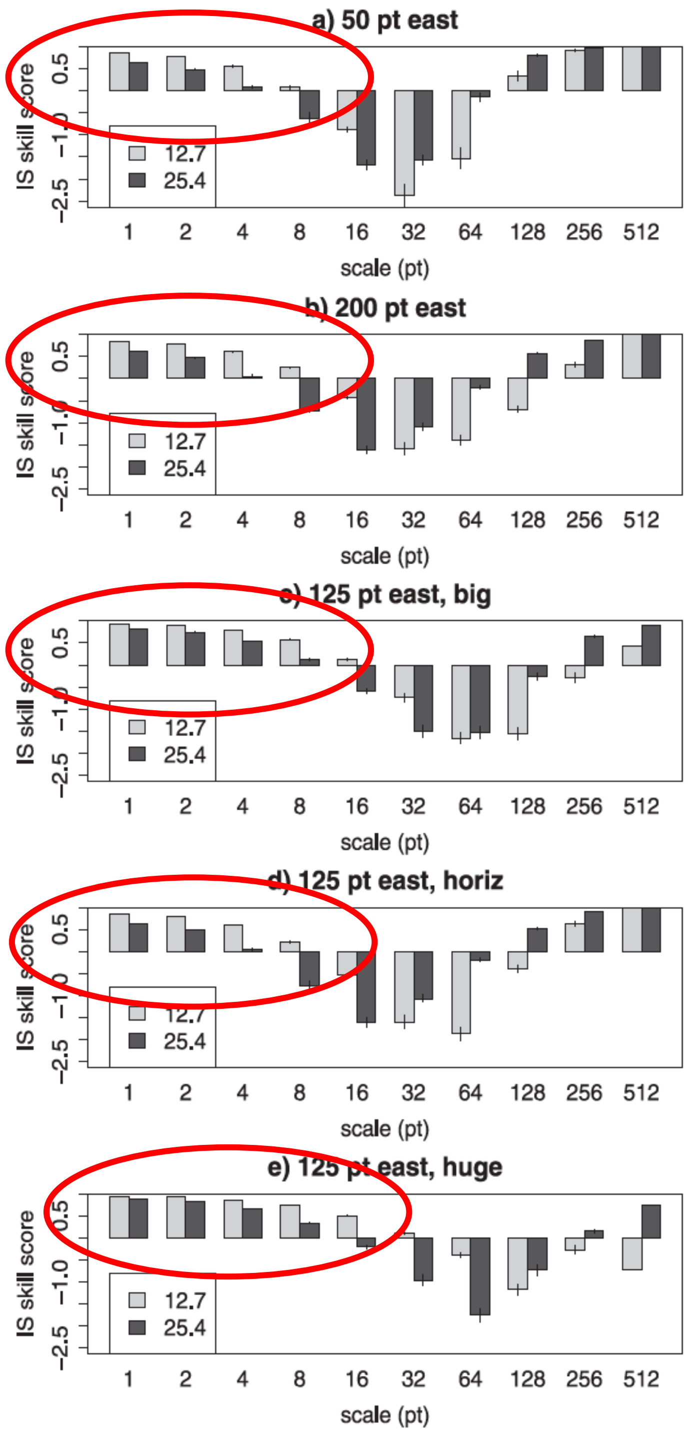
Neighbourhood verification: the scale is the neighbourhood size (low band pass filter): as the scale increases the exact positioning requirements are more and more relaxed (the performance is bound to improve)

Scale-decomposition verification: the scales are separated by a single-band spatial filter - inform on performance on individual separate scales

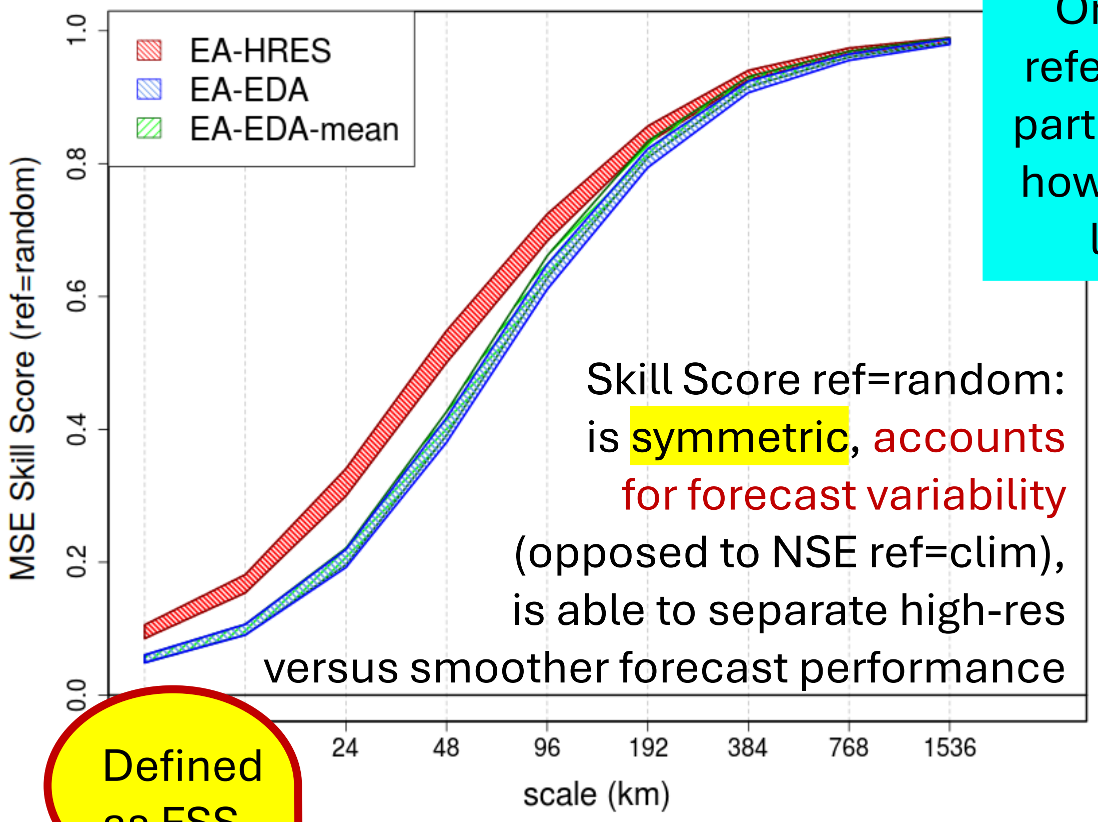
In general, forecasts exhibit negative skill on small scales, and positive skill on large scales (predictability), and both approaches are expected to provide (similar) no-skill to skill transition scale. However ...





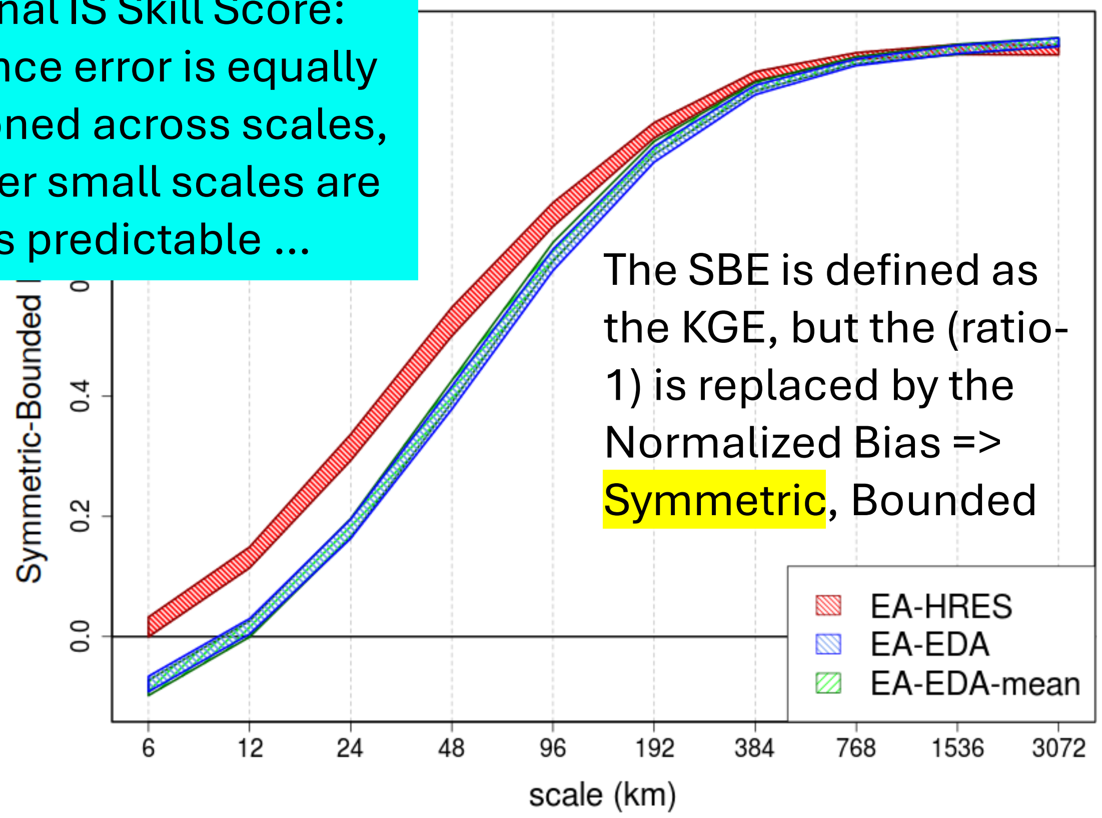


Post MesoVICT innovation: ~~Intensity~~ **Scale separation diagnostics and the Symmetric Bounded Efficiency** ... (Casati, Lussana, Crespi, 2023, *Int.J.Clim*) introduce scores more suitable for the inter-comparison of gridded products with different resolutions



Skill Score ref=random: is symmetric, accounts for forecast variability (opposed to NSE ref=clim), is able to separate high-res versus smoother forecast performance

Original IS Skill Score: reference error is equally partitioned across scales, however small scales are less predictable ...



The SBE is defined as the KGE, but the (ratio-1) is replaced by the Normalized Bias => Symmetric, Bounded

Defined as FSS

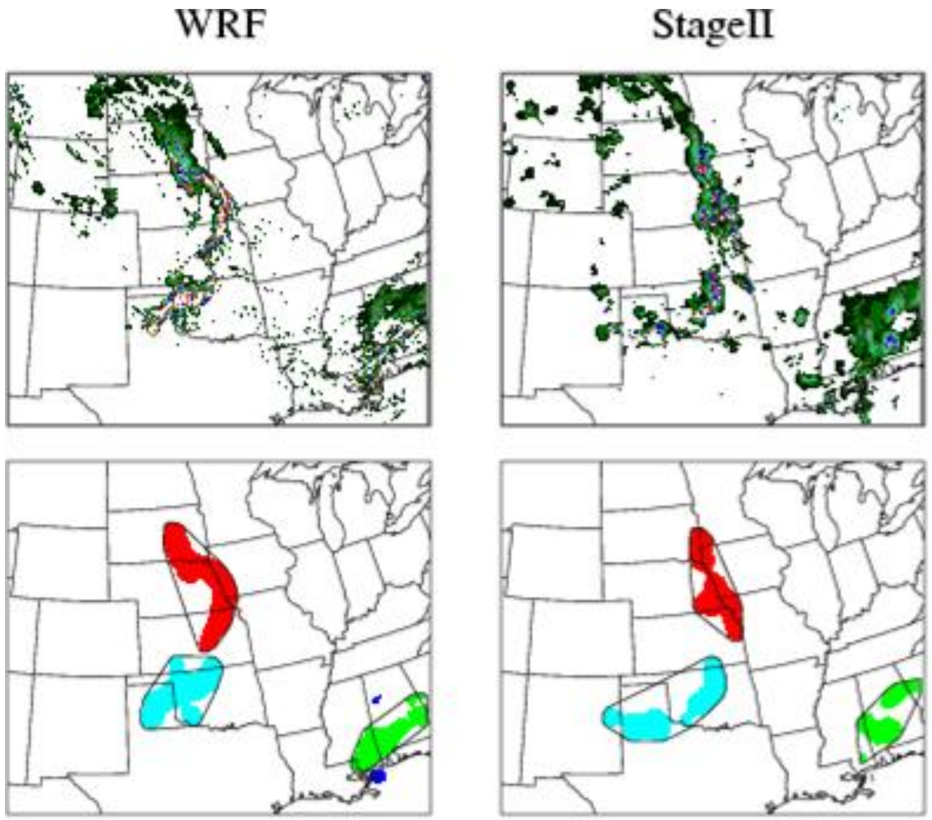
$$SSSSj = \boxed{1 - \frac{MSEj}{\sigma_{Yj}^2 + \sigma_{Xj}^2}} = \frac{2\sigma_{Yj}\sigma_{Xj}r_{Yj,Xj}}{\sigma_{Yj}^2 + \sigma_{Xj}^2}$$

$$SBE = 1 - \sqrt{\underbrace{(r-1)^2}_{\text{Corr}} + \underbrace{\left(\frac{\sigma_Y - \sigma_X}{\sigma_Y + \sigma_X}\right)^2}_{\text{Variability}} + \underbrace{\left(\frac{\mu_Y - \mu_X}{|\mu_Y| + |\mu_X|}\right)^2}_{\text{Bias}}}$$

Reference = MSE random = variances of both obs and forecast

3. Feature-based techniques

1. Identify and isolate (precipitation) **features** in forecast and observation fields (thresholding, composites, cluster analysis, binary image metrics, …)
2. assess displacement and amount (extent and intensity) error for each pairs of obs and forecast features:
    - Compare attributes of object pairs (e.g. intensity, area, centroid location);
    - evaluate distance-based contingency tables and categorical scores;
    - perform verification as function of feature size (scale);
    - add time dimension for the assessment of the timing error of precipitation systems



Some selected References:

Ebert and McBride (2000), Grams et al (2006), Ebert and Gallus (2009): CRA

Davis, Brown, Bullok (2006) I and II, Davis et al (2009): MODE

Wernli, Paulat, Frei (2008): SAL score

Nachamkin (2004, 2005): composites

Marzban and Sandgathe (2006): cluster

Lack et al (2010): Procrustes

Skok (2023): precipitation attribution distance

4. Field-deformation methods

1.Use a vector (wind) field to deform the forecast field towards the obs field
2.Use an amplitude field to correct intensities of (deformed) forecast field to those of the obs field

- Originated from nowcasting and data assimilation needs, can be employed to correct the position of meteorological features
- Vector and amplitude fields provide physically meaningful information: e.g. can diagnose systematic errors in the advection (wind, position of mass fields).
- Error decomposition is performed on different scales components: inform on small scales uncertainty versus large scale errors.

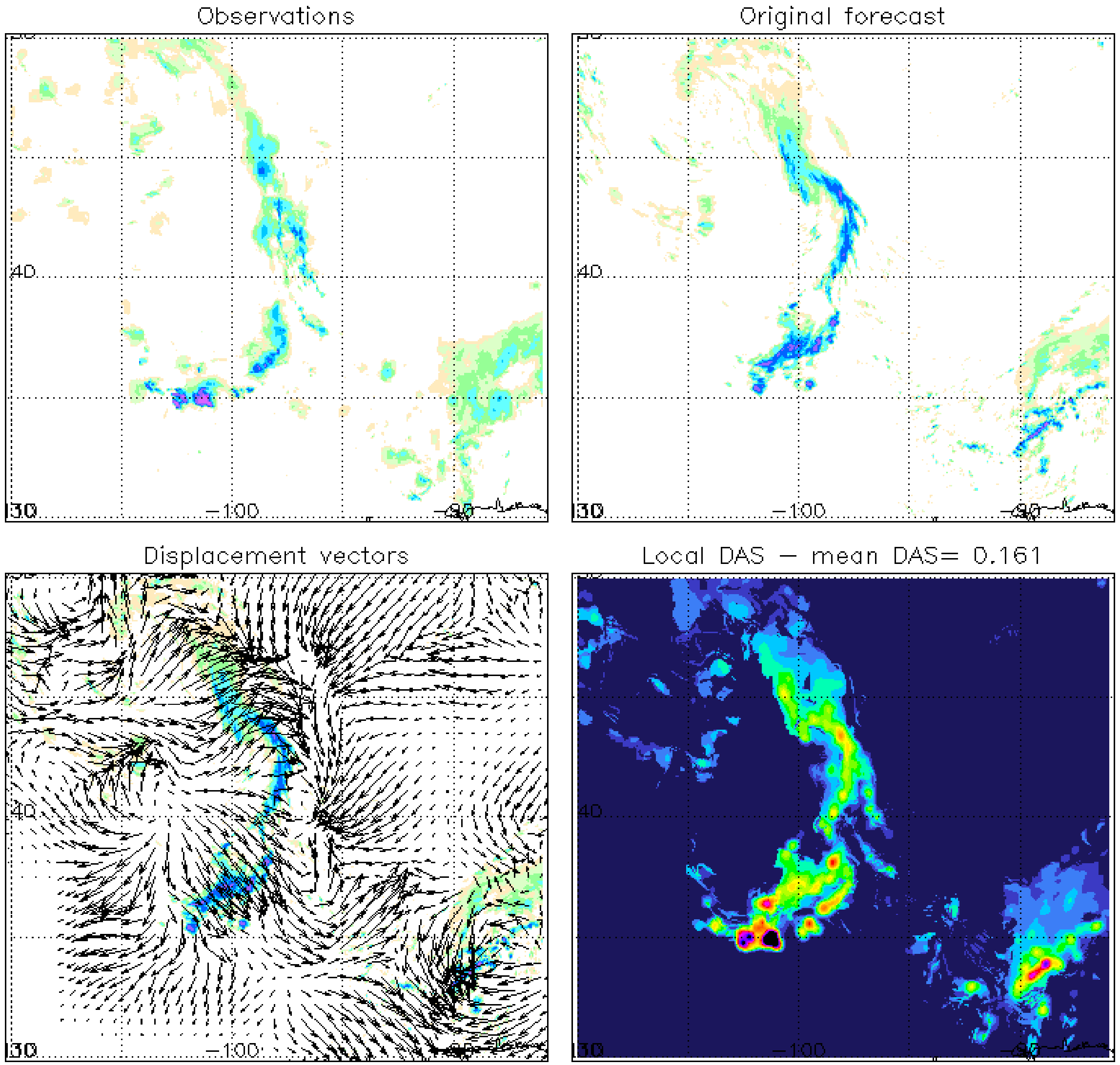
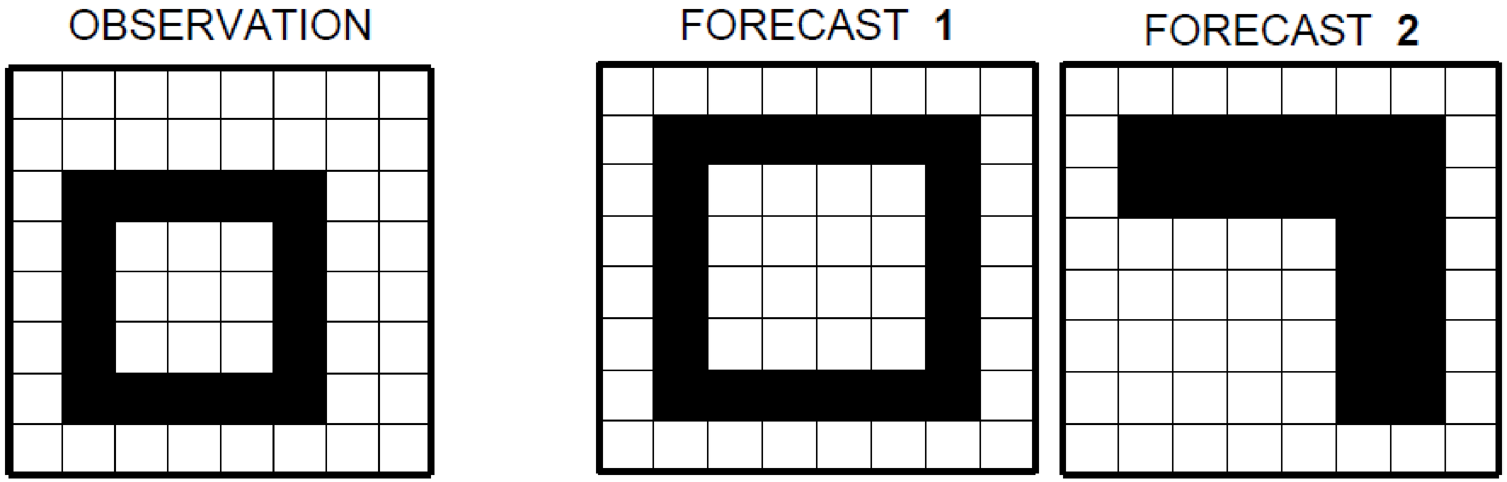


Image from Keil and Craig (2009), Wea&For

Some selected References: Hoffmann et al (1995); Hoffman and Grassotti (1996), Nehrkorn et al. (2003); **Brill (2002); Germann and Zawadzki (2002, 2004); Keil and Craig (2007, 2009)** DAS; **Marzbar and Sandgathe (2010)** optical flow; **Alexander et al (1999), Gilleland et al (2010)** image warping; **Francis et al (2025)** optimal transport

5. Distance metrics for binary images

- Average distance
- K-mean
- Fréchet distance
- **Hausdorff metric**
  - **Modified Hausdorff**
  - **Partial Hausdorff**
  - **Baddeley metric**
- Pratts' figure of merit



$\Delta = 0.5625$	$\Delta = 0.96875$
hits = 9; false alarms = 11; misses = 7; nils = 37	

- Evaluate distances for all grid-points
- Account for distance between objects, and similarity in shapes.
- Compare binary images: alternative metrics to be used along with traditional categorical scores

Image Processing: **Baddeley (1992); Dubuisson and Jain (1994).** Precipitation: **Gilleland et al.(2008, 2011)**, Schwedler & Baldwin (2011), Venugopal et al. (2005); Zhu et al (2010), Aghakouchak et al (2011). **Brunet and Sills (2015).** Sea-ice: Heinrichs et al (2006); **Dukhovskoy et al (2015)**, Hebert et al (2015).

Open-access software

The R Project for Statistical Computing

[Home]

Download

CRAN

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To **download R**, please choose your preferred CRAN mirror.

Gilleland, E. (2022). Comparing spatial fields with SpatialVx: Spatial forecast verification in R. doi:10.5065/4px3-5a05

Brown et al (2021) “The Model Evaluation Tools (MET): More than a Decade of Community-Supported Forecast Verification” BAMS

METplus: https://dtcenter.org/community-code/metplus

- Community based software, is a **tool for research-to-operation tech transfer.** Operational at UKMO, BoM, NOAA, ..
- Modular, python: Data Base, scoring, visualization.
- Includes verification tools for
  - Stations and gridded obs
  - Spatial verification
  - Inference (Confidence Intervals)
- Documentation, user support, online tutorials and workshops

DTC Developmental Testbed Center

ABOUT TESTING + EVALUATION COMMUNITY CODE VISITOR PROGRAM

NEWS EVENTS

Welcome

Welcome to the users page for the enhanced Model Evaluation Tools (METplus) verification system. METplus was developed by the Developmental Testbed Center (DTC) through the generous support of the 557th Weather Wing of the United States Air Force, the National Oceanic and Atmospheric Administration (NOAA), and the National Center for Atmospheric Research (NCAR).

Description

METplus is a verification framework that spans a wide range of temporal (warn-on-forecast to climate) and spatial (storm to global) scales. It is intended to be extensible through additional capability developed by the community. The core components of the framework include MET, the associated database and display systems called METviewer and METexpress, and a suite of Python wrappers to provide low-level automation and examples, also called use-cases. METplus will be a component of NOAA's Unified Forecast System (UFS) cross-cutting infrastructure as well as NCAR's System for Integrated Modeling of the Atmosphere (SIMA).

METplus is being actively developed by NCAR/Research Applications Laboratory (RAL), NOAA/Earth Systems Research Laboratories (ESRL), NOAA/Environmental Modeling Center (EMC), and is open to community contributions.

METplus Components

Links to the code repository and documentation for each METplus component are provided below:

- METplus Wrappers (repo, docs)
- MET (repo, docs)
- METviewer (repo, docs)
- METexpress (repo, docs)
- METplotpy (repo, docs)
- METcalcpy (repo, docs)
- METdatadb (repo, docs)

METplus Sponsors

- National Center for Atmospheric Research (NCAR)
- National Oceanic and Atmospheric Administration (NOAA)
- United States Air Force (USAF)

NCAR

METPLUS COMPONENTS

METPLUS

- Home
- System Architecture
- Download +
- Documentation
- User Support +

LATEST RELEASE

METexpress Version 4.5.3
Released: 2022-10-18

METplus Version 4.1.4
Released: 2022-09-09

MET Version 10.1.2
Released: 2022-05-16

Coordinated METplus Version 4.1
Released: 2022-03-15

METdatadb Version 1.1.0
Released: 2022-03-11

METcalcpy Version 1.1.0
Released: 2022-03-11

METplotpy Version 1.1.0
Released: 2022-03-11

METviewer Version 4.1.0
Released: 2022-03-11

UPCOMING EVENTS

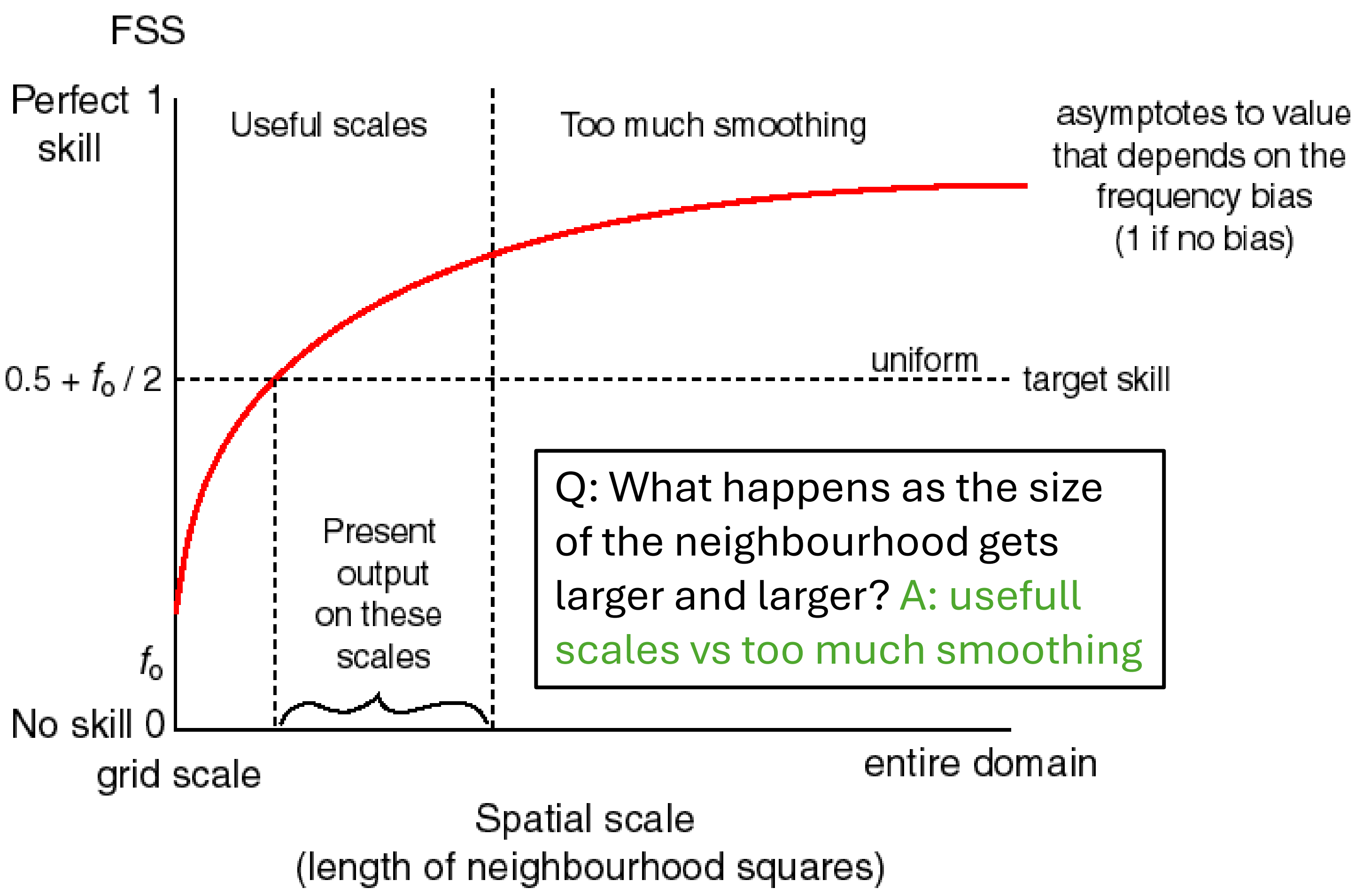
No upcoming events

PAST EVENTS

No recent events

Intermezzo: the most debated score

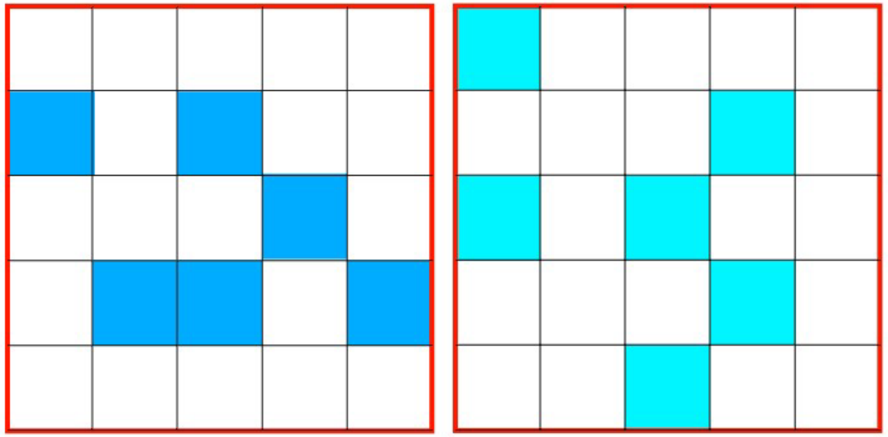
The **Fraction Skill Score**, introduced by Roberts and Lean (2008), *MonWeaRev* **136**: simple to compute!



Q: What happens as the size of the neighbourhood gets larger and larger? A: usefull scales vs too much smoothing

P_{fcst} is the fraction of forecast gpt > threshold

P_{obs} is the fraction of obs gpt > threshold



Fraction = 6/25 = 0.24 Fraction = 6/25 = 0.24

$$\mathrm{FSS} = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(P_{fcst} - P_{obs})^2}{\frac{1}{N}\sum_{i=1}^{N}P_{fcst}^{\ 2} + \frac{1}{N}\sum_{i=1}^{N}P_{obs}^{\ 2}}$$

Q: What is a skillfull FSS?
A: defining the reference target skill

The FSS ~~challenged~~ history

- Skok (2016) Analysis of Fraction Skill Score properties for a displaced rainy grid point in a rectangular domain AtmRes
- Skok and Roberts (2016) Analysis of Fractions Skill Score properties for random precipitation fields and ECMWF forecasts QJRMS
- Skok and Roberts (2018) Estimating the displacement in precipitation forecasts using the Fractions Skill Score QJRMS

- Skok (2023) Precipitation attribution distance AtmRes
- Skok and Lledo (2025) Spatial verification of global precipitation forecasts QJRMS

- Mittermaier (2018) How interpolation and resolution can affect verification scores: A study based on the Fractions Skill Score MetZeit
- Mittermaier (2021) A “Meta” Analysis of the Fractions Skill Score: The Limiting Case and Implications for Aggregation MWR
- Mittermaier (2025) localized FSS QJRMS

- Antonio and Aitchison (2025) How to derive skill from the FSS MWR

Neighbourhood verification for ensembles

Atger (2001) multi-event contingency tables

Marsigli et al. 2008: compare forecasts and obs distribution parameters over neighbourhoods

Theis et al (2005) Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach → Ben Bouallegue & Theis (2014) Spatial techniques applied to precipitation ensemble forecasts: from verification results to probabilistic products

Mittermaier (2014) W&F and Mittermaier and Csima (2017): compared deterministic and ensembles forecasts on the kilometer scale.

Duc et al (2013) ensemble FSS, including time

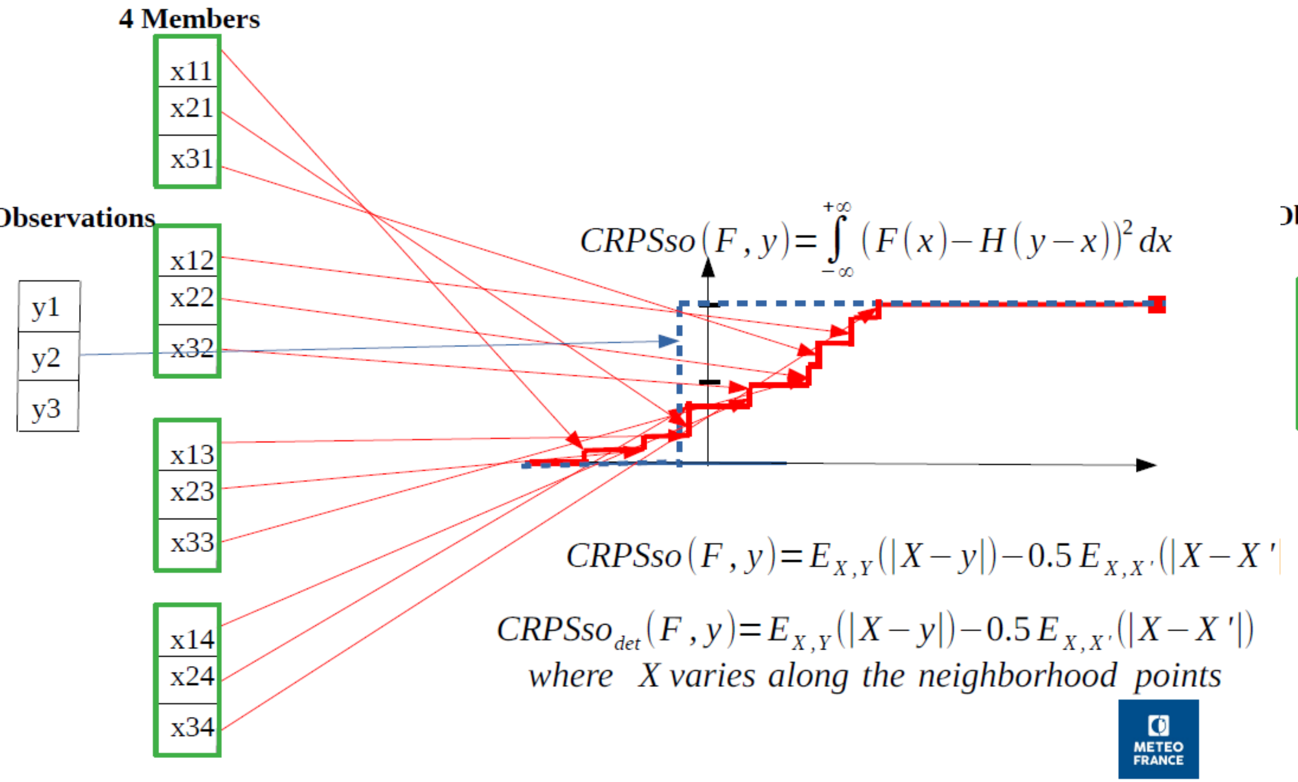
- Stein and Stoop (2022); Neighborhood-Based Ensemble Evaluation Using the CRPS
- Stein and Stoop (2024) Evaluation of Probabilistic Forecasts of Binary Events with the Neighborhood Brier Divergence Skill Score Brier Divergence Skill Score
- Stein (2025) Diagonal Scores and Neighborhood: Definitions and Application to Idealized Cases

Necker et al (2024) The fractions skill score for ensemble forecast verification QJRMS

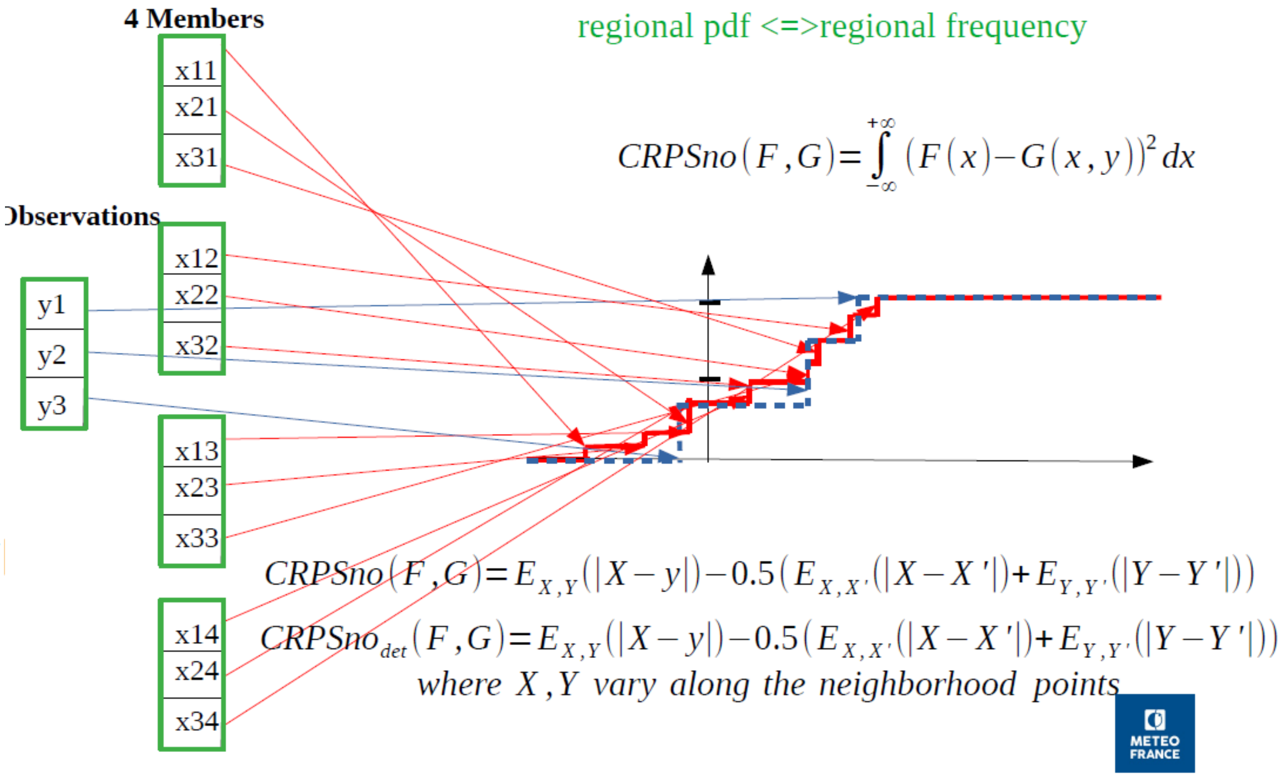
Spatial verification technique built around a **proper scoring rules,** prevent models from artificially gaming the score through strategic over-smoothing ?

Example: CRPS calculated in the neighbourhood, for both SO-NF and NO-NF (from Stein and Stoop, 2022)

CRPSso : increase of the number of Members



CRPSno : comparison at the neighborhood scale



images courtesy of Joel Stein, from the 2020-IVMW presentation

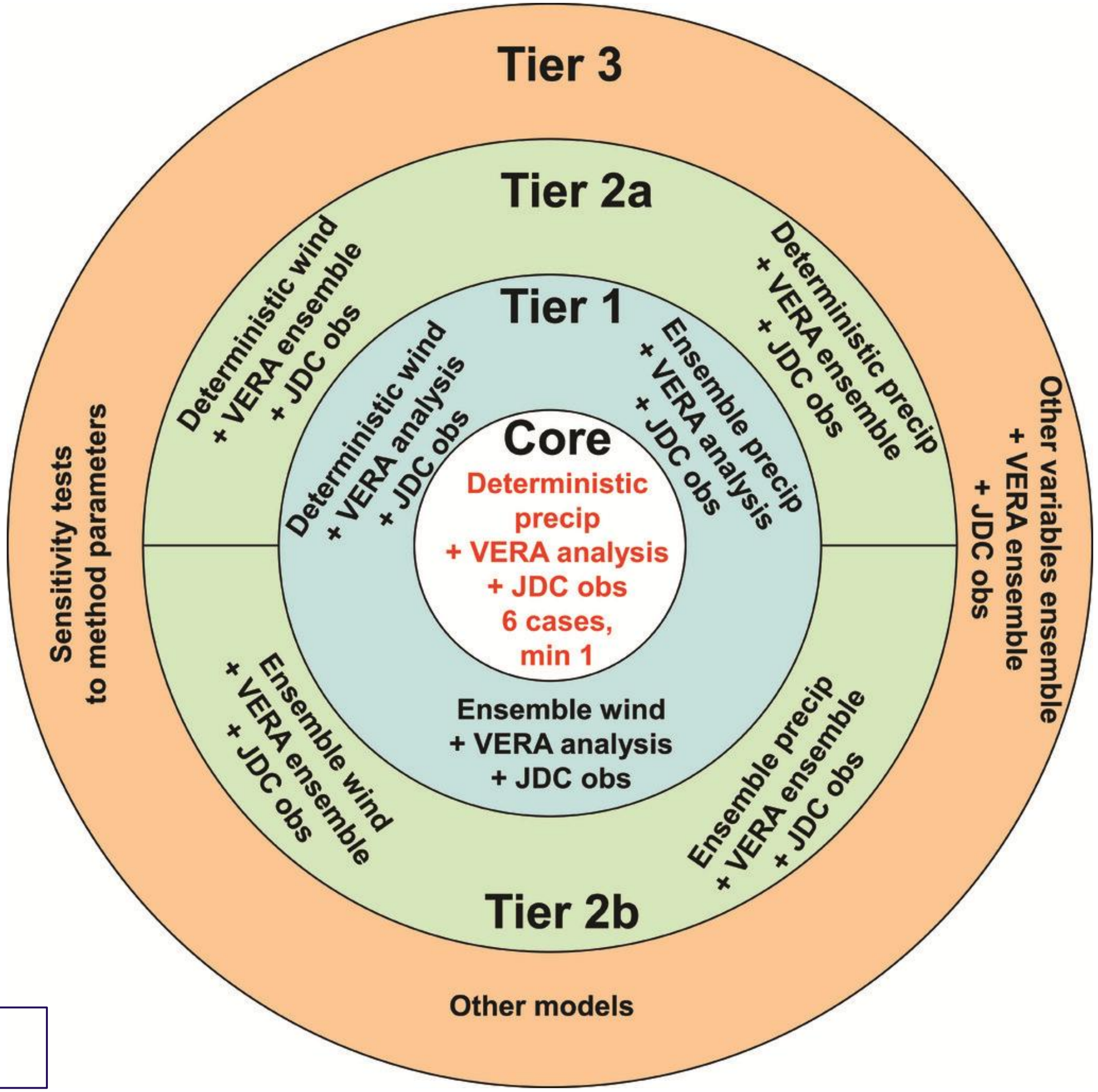
MesoVICT
The Mesoscale Verification Inter-Comparison over Complex Terrain project

MesoVICT innovative goals and experimental design

Dorninger et al (2018): The set-up of the Mesoscale Verification Inter-Comparison over Complex Terrain project. *BAMS*. 99 (9):1887-1906, doi: 10.1175/BAMS-D-17-0164.1.

How do spatial verification methods perform for:

1. Additional variables beyond precipitation, such as wind and temperature.
2. Deterministic **and ensemble** forecasts.
3. Challenges of **complex terrain.**
4. Account for **observation uncertainty:** VERA ensemble analysis
5. Sensitivity to own tuning parameters, domain size, interpolation and re-gridding?
6. Can compare models with different resolutions?

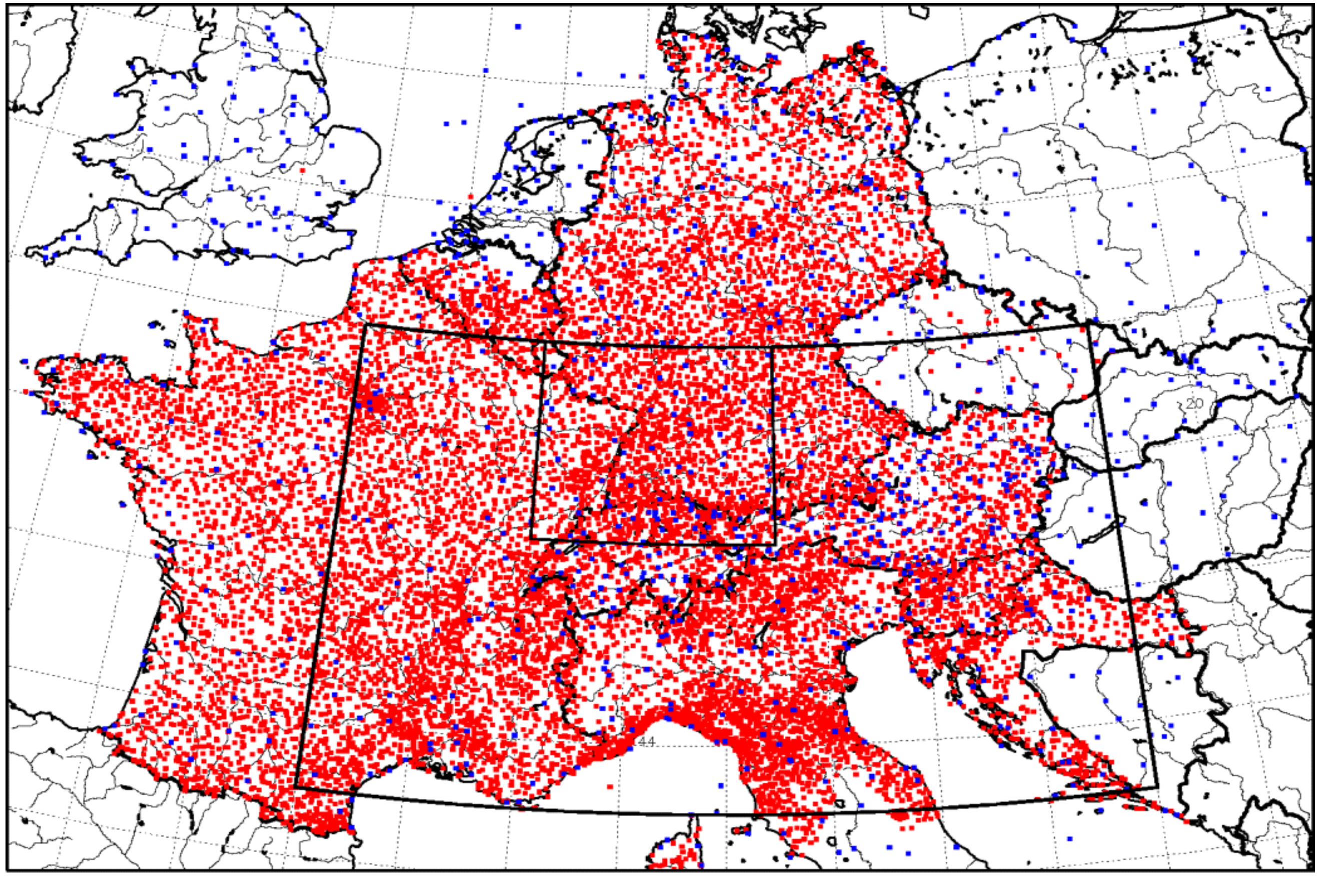


Provide a community testbed for testing and comparison

MesoVICT dataset: verification reference

32 data providers
- GTS stations: 1232 (blue)
- non GTS Stations: > 13000

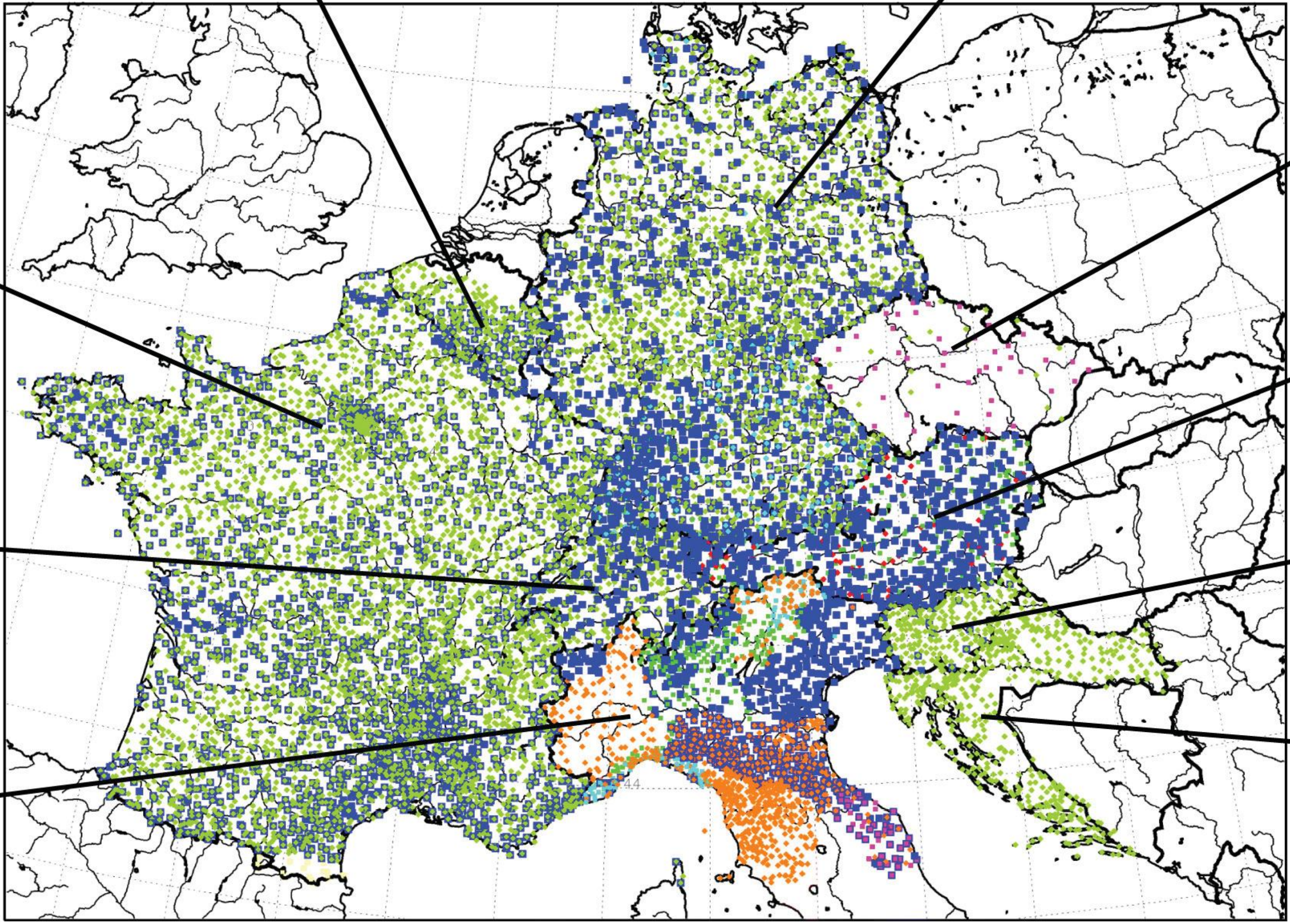
Mean station distance:
- GTS: ~ 36km
- GTS+Non-GTS: ~ 12km



Observational data:
Joint D-PHASE-COPS (JDC) data set

Gridded model-independent analysis:
Vienna Enhanced Resolution Analyses (VERA)
- Quality control (Steinacker et al. 2011)
- Obs gridding by thin-plate spline
- Downscaling by fingerprint method (Steinacker et al. 2006; Bica et al. 2007): thermal and dynamical effects
- 8km resolution, hourly fields
- Model interpolation to VERA grid by inverse distance method (Cressman 1959)
- The VERA ensemble is generated from observation uncertainties - Error estimates are perturbed assuming a gaussian distribution (Gorgas & Dorninger, 2012: Concepts for a pattern-oriented analysis ensemble based on observational uncertainties. *Q. J. R. Meteorol. Soc.*, **138**, 769-784).

Challenges: harmonizing precipitation accumulations



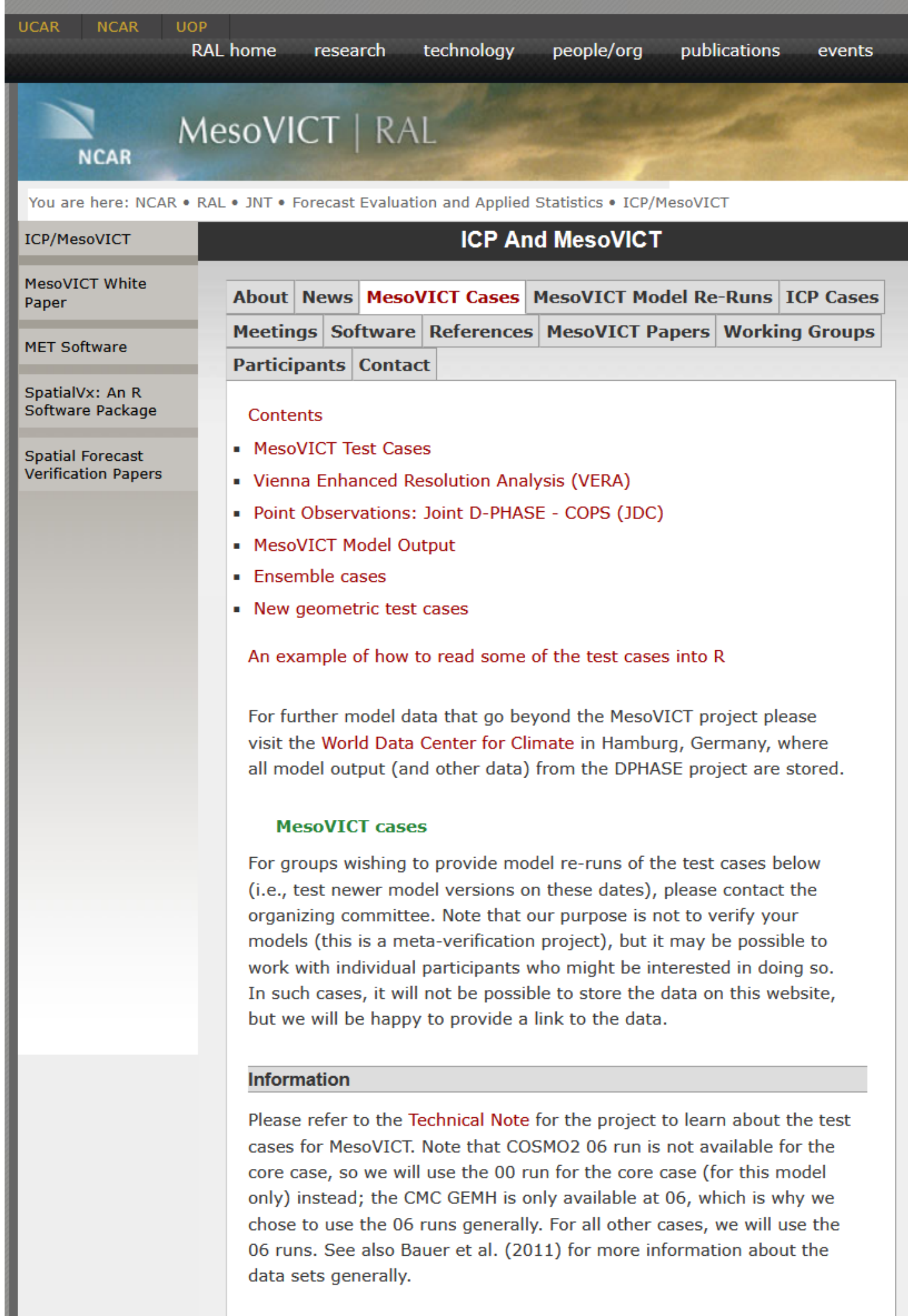
MesoVICT case studies

Forecasts from two WWRP Forecast Demonstration Projects:

- Mesoscale Alpine Program (MAP)–Demonstration of Probabilistic Hydrological and Atmospheric Simulation of Flood Events in the Alpine Region (D-PHASE) project (Rotach et al. 2009)
- Convective and Orographically-Induced Precipitation Study (COPS; Wulfmeyer et al. 2008)

Three selected models:
Swiss COSMO-2, the Canadian GEM-LAM, and the Limited Area ensemble COSMO-LEPS (ARPA Emilia-Romagna, Italy). Several institutions provided re-runs.

Meteorological events evolving over time, over few days.
Description in Dorninger et al (2013) NCAR TechNote 505.
Data available at: https://projects.ral.ucar.edu/icp/


UCAR NCAR UOP
RAL home research technology people/org publications events

MesoVICT | RAL

You are here: NCAR • RAL • JNT • Forecast Evaluation and Applied Statistics • ICP/MesoVICT

ICP/MesoVICT
MesoVICT White Paper
MET Software
SpatialVx: An R Software Package
Spatial Forecast Verification Papers

ICP And MesoVICT

About | News | MesoVICT Cases | MesoVICT Model Re-Runs | ICP Cases | Meetings | Software | References | MesoVICT Papers | Working Groups | Participants | Contact

Contents
- MesoVICT Test Cases
- Vienna Enhanced Resolution Analysis (VERA)
- Point Observations: Joint D-PHASE - COPS (JDC)
- MesoVICT Model Output
- Ensemble cases
- New geometric test cases

An example of how to read some of the test cases into R

For further model data that go beyond the MesoVICT project please visit the World Data Center for Climate in Hamburg, Germany, where all model output (and other data) from the DPHASE project are stored.

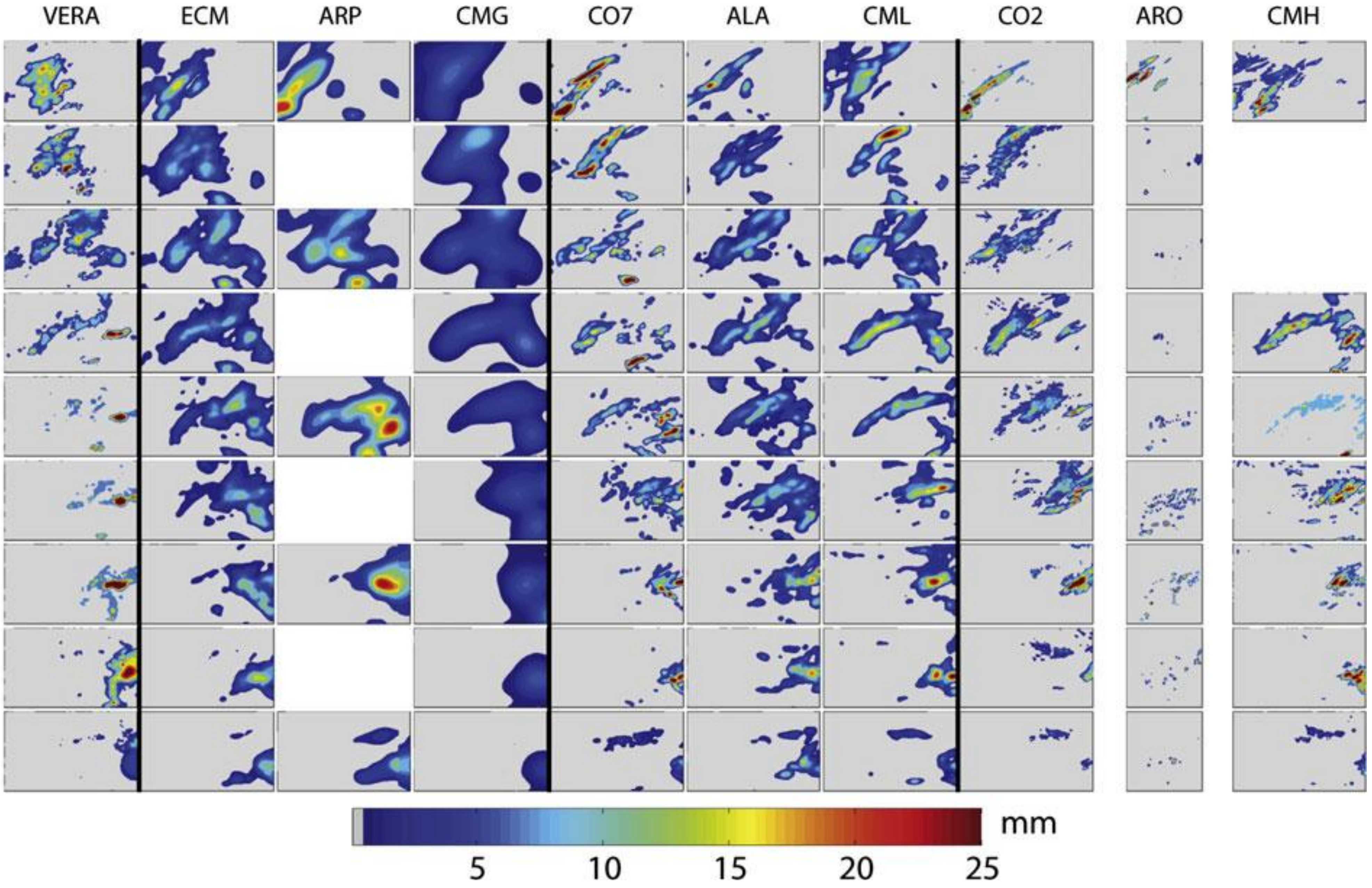
MesoVICT cases

For groups wishing to provide model re-runs of the test cases below (i.e., test newer model versions on these dates), please contact the organizing committee. Note that our purpose is not to verify your models (this is a meta-verification project), but it may be possible to work with individual participants who might be interested in doing so. In such cases, it will not be possible to store the data on this website, but we will be happy to provide a link to the data.

Information

Please refer to the Technical Note for the project to learn about the test cases for MesoVICT. Note that COSMO2 06 run is not available for the core case, so we will use the 00 run for the core case (for this model only) instead; the CMC GEMH is only available at 06, which is why we chose to use the 06 runs generally. For all other cases, we will use the 06 runs. See also Bauer et al. (2011) for more information about the data sets generally.

18 Sept 2007: fast moving front from west to east impinging the Alps



The MesoVICT new penta-classification

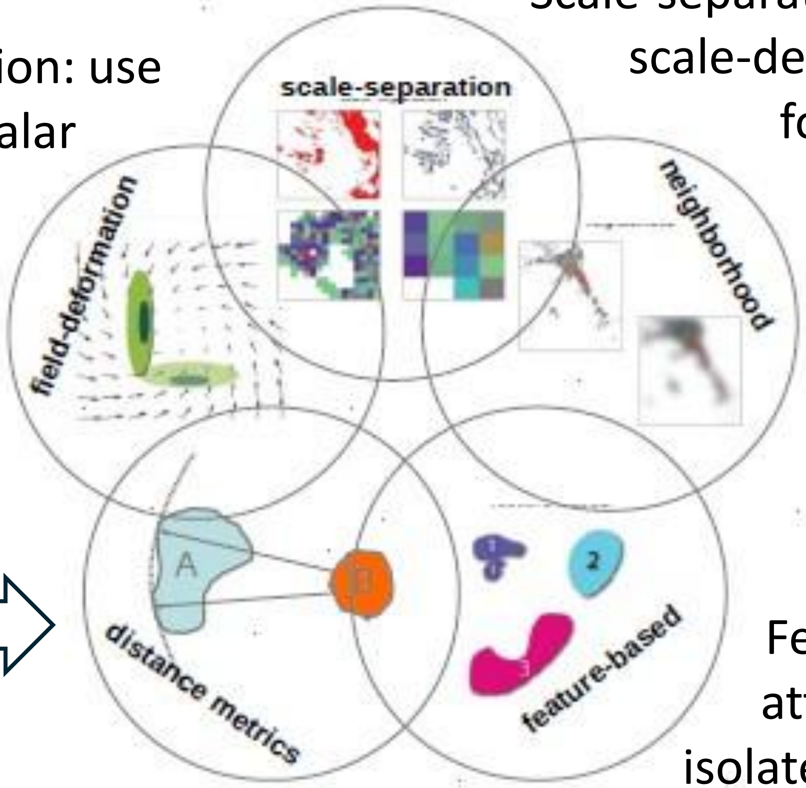
- Account for **coherent spatial structure** and the presence of **features**
- Avoid **double-penalty** issue, relax exact matching allowing **small time-space discrepancies**
- Provide information on **error in physical terms: e.g.** assess **scale structure** and **displacement error** (separately from **intensity error**) – meaningful + informative diagnostics

Spatial Verification classes overlap, as some methods belong to more than one class: proximity indicates methods which are related / connected.

Addition of a class of methods: distance measures for binary images, evaluate distances for selected pixels / binary features over the whole field – bridges feature based and morphing methods

Field-deformation: use a vector and scalar field to morph forecast into obs field

Scale-separation: analyse scale-dependency of forecast error



Neighborhood: relax requirement of exact space-time matching

Feature-based: evaluate attributes of individually isolated and paired features

From Dorninger et al (2018) *BAMS*

MesoVICT Meetings and Workshops

13th EMS/11th ECAM	Reading, UK	9-13 Sep 2013	initiation of the project
1st MesoVICT workshop (kick-off)	Vienna, Austria	2-3 Oct 2014	develop the program
15th EMS/12th ECAM	Sofia, Bulgaria	7-11 Sep 2015	MesoVICT session
16th EMS/11th ECAC	Trieste, Italy	12-16 Sep 2016	MesoVICT session
2nd MesoVICT workshop	Bologna, Italy	21-23 Sep 2016	intermediate workshop
7th International Verification Methods Workshop	Berlin, Germany	3-11 May 2017	MesoVICT exercises & MesoVICT talks
EMS annual meeting	Budapest, Hungary	3-7 Sep 2018	MesoVICT talks
MesoVICT final workshop	Vienna, Austria	8-9 July 2019	resume

Final MesoVICT meeting: https://mesovict.univie.ac.at/
MesoVICT meeting and publications: https://projects.ral.ucar.edu/icp/

Exploratory studies and some grey literature

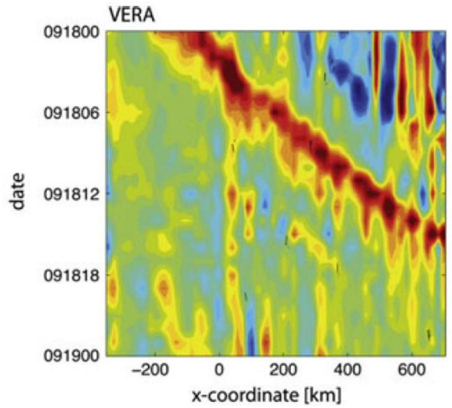
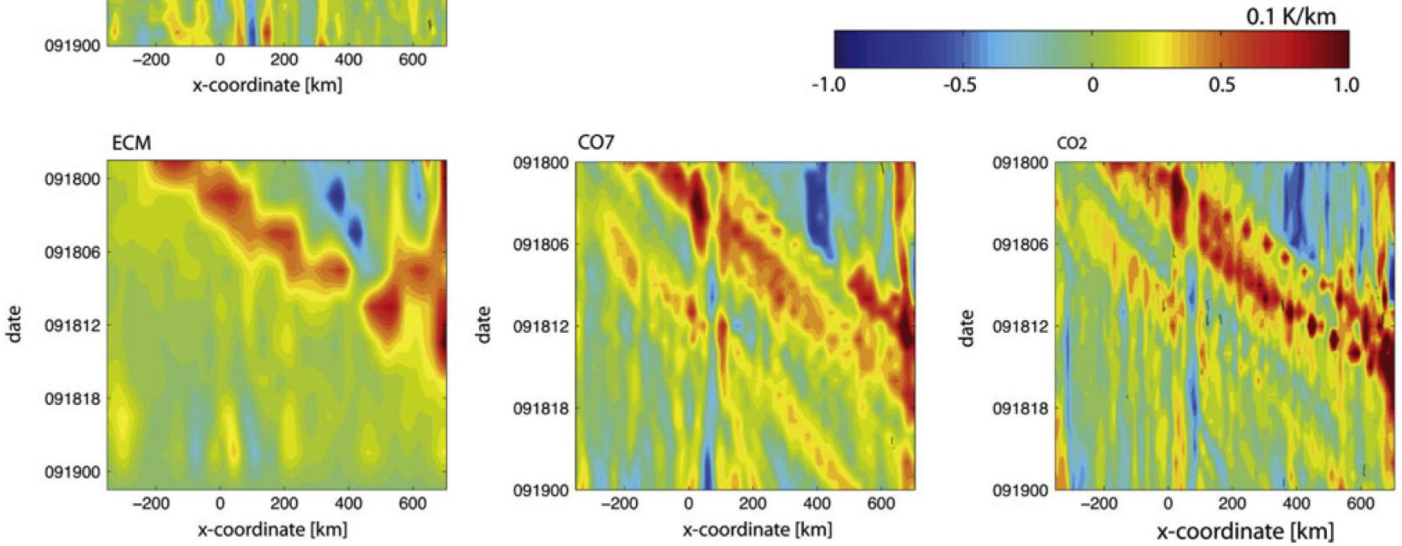


Figure 9: Hovmoller diagram for the meridionally averaged gradient of equivalent potential temperature for VERA and the model chain ECM-CO7-CO2 for 18 September 2007, 00 UTC to 19 September 2007, 00 UTC. Red colours: separate warmer air mass to the east from colder to the west, blue colours: separate colder air to the east from warmer to the west;



Dorninger and Gorga (2013): Comparison of NWP-model chains by using novel verification Methods. *MetZeit*. [10.1127/0941-2948/2013/0488](10.1127/0941-2948/2013/0488) – explore SAL, IS, Bias Corrected RMSE …
exploit existing (NCAR) code

Geiß, S., 2015: Comparison of spatial verification methods. Bachelor thesis. Ludwig-Maximilians-University Munich Meteorological Institute Munich, Supervisor: Christian Keil, 55 pp. (pdf)
Kloiber, S., 2017: Verification in Complex Terrain with Ensemble-Analysis. MSc thesis, Dep. of Meteorology and Geophysics, University of Vienna, 74 pp.
COSMO Priority Project INSPECT (COSMO Tech Report No. 37, 2019): comprehensive testing of spatial verification methods to provide criteria for the best suited method for a specific application.

Some MesoVICT outcomes / publications

Gilleland 2017: A new characterization false alarms, misses, and overall patterns10.1175/WAF-D-16-0134.1.
Gilleland 2020: novel set of test cases (pathological, circular, elliptic and scattered fields) DOI: https://doi.org/10.1175/MWR-D-19-0256.1
Gilleland 2021: Novel measures (G and G_beta):10.5194/ascmo-7-13-2021.

Mariani and Casaioli, 2018: effects of domain size and resolution on CRA. **10.1127/metz/2018/0897**.
Radanovics et al. 2018: SAL for ensembles. **10.1175/WAF-D-17-0162.1.**
Skok and Hladnik, 2018: FSS adapted to wind. **https://doi.org/10.1175/MWR-D-16-0471.1.**
Buschow and Friederichs (2020, 2021) SAD: scale, anisotropy and direction (account for orography)
Buschow (2022) displacement, complex wavelets+continuous;

Han, F. and Szunyogh, I.2016. A morphing-based technique for the verification of precipitation forecasts. *Mon. Wea. Rev.* **144**, 295–313.https://doi.org/10.1175/MWR-D-15-0172.1
Han, F. and Szunyogh, I.2017. *A technique for the verification of precipitation forecasts and its application to a problem of predictability*. *Mon. Wea. Rev.* **146**, 1303-1318. DOI: https://doi.org/10.1175/MWR-D-17-0040.1 introduce the amplitude and structure similarity index metric (ASSIM), related to Wang & Bovik (2002,2009), Wang et al. (2004)

… and possibly other studies …

There are also several studies which have been showcased at conferences but have never been published!

What have we learned that can be useful for steering / shaping WP-MIP and BtG ?

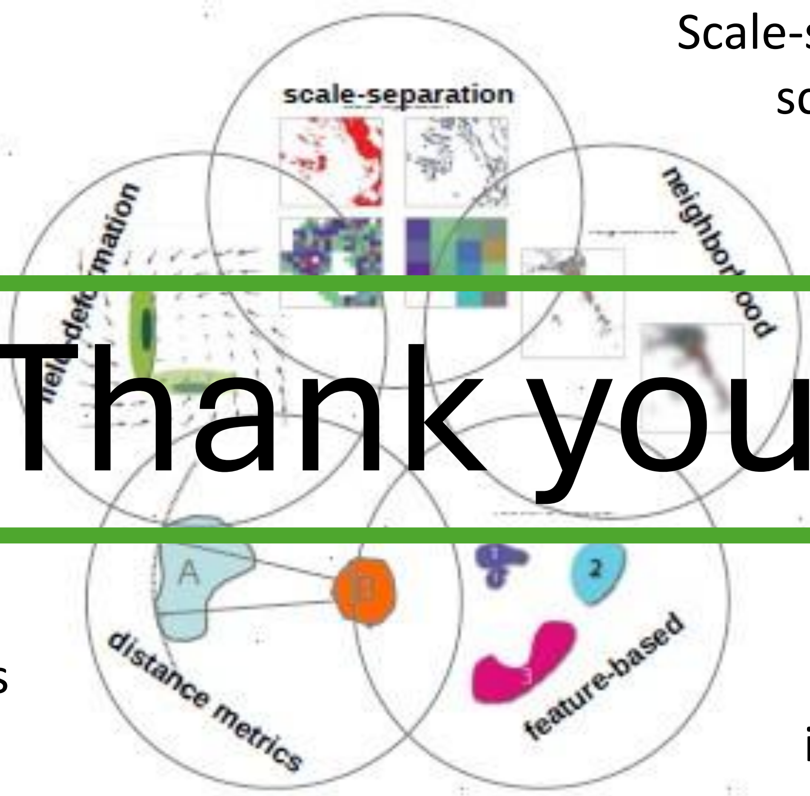
1st spatial Vx ICP	MesoVICT	WP-MIP	B-i-i-G
Pressing scientific gap/need (high-res models + double penalty + more informative verification)	… less pressing / innovative goals …	AI, pressing+ innovative	EPS, innovative
Critical mass of already published techniques → classification	Techniques already published (twice), classification just needed refining …	Already published to test + develop physical coherence	review
Guidance to verif practitioners	Guidance to verif practitioners	WIPPS guidance	
WWRP + workshops → verif community engagement	WWRP projects / double-tasks → less engagement from 1st SpVx ICP players	WWRP+WIPPS & verif comm.	WWRP verif comm.
NCAR driving: verif hub + IT capacity	→ exploiting existing code	BoM scores	
Few clearly outlined research questions	More challenging research questions	… a zoo of research questions …	
Well thought case studies	Realistic multi-dimensional case studies (time + more variables)	One full year, many models, many variables: overwhelming?	
Gridded verification reference	Very well thought verification references	No reference	

Spatial verification approaches

- Account for **coherent spatial structure** and the presence of **features**
- Avoid **double-penalty** issue, relax exact matching allowing **small time-space discrepancies**
- Provide information on **error in physical terms: e.g.** assess **scale structure** and **displacement error** (separately from **intensity error**) – meaningful + informative diagnostics

Scale-separation: analyse scale-dependency of forecast error

Field-deformation: use a vector and scalar field to morph forecast into obs



Thank you!

Neighborhood: relax requirement of exact space-time matching

Distance metric for binary images: evaluate distances for features and fields

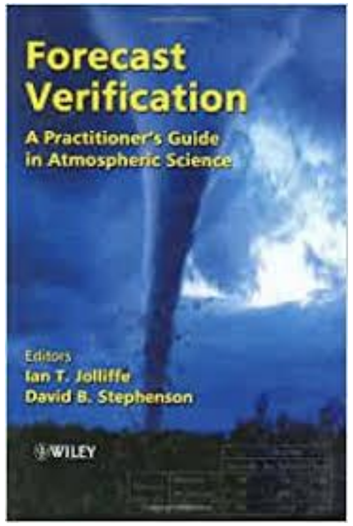
Feature-based: evaluate attributes of individually isolated and paired features

Key references: Gilleland et al (2009, 2010), Dorninger et al (2018)

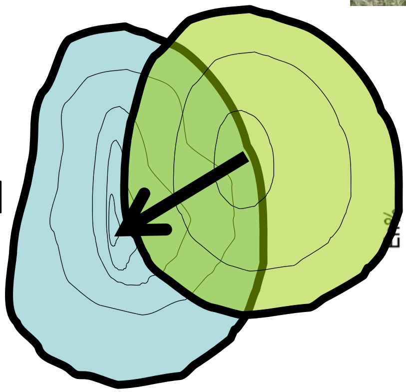
Extras

JWGFVR History

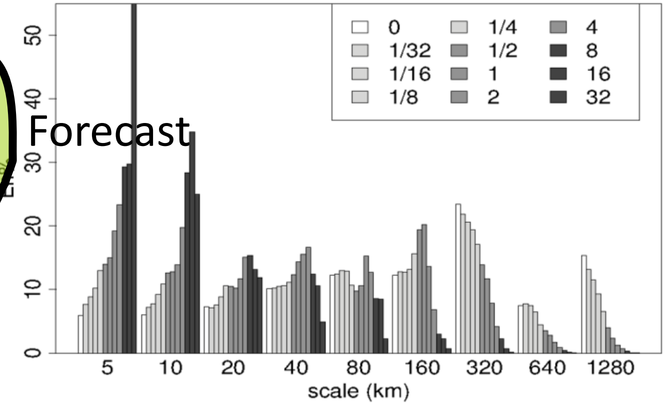
- Sydney 2000 real time verification (Beth Ebert): a small group of verification experts is formed
- First two verification method workshops: 2002, Barb Brown (NCAR); 2004 Laurie Wilson (ECCC);
- 2007 ECMWF first workshop + tutorials.
- Web-based tools: EUMETCAL verification modules (L. Wilson and P. Nurmi)
- Forecast Verification FAQ webpage (Beth Ebert);
- Joliffe and Stephenson (2003): book on Forecast Verificaton
- Initial publications on spatial verification methods:
  Ebert and Mc.Bride (2000) CRA; Casati et al (2004) Intensity Scale technique; Roberts and Lean (2008) Fraction Skill Score; Davis et al (2006) MODE; Keil and Craig (2008) DAS
- The first spatial verification method intercomparison project (Gilleland et al 2010)
- The second ICP, MesoVICT (Dorninger et al 2018)
- R verification software, Met, METplus (NCAR)

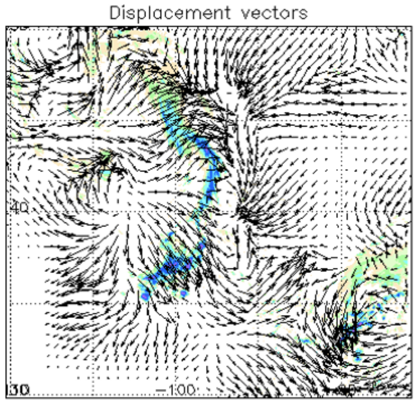


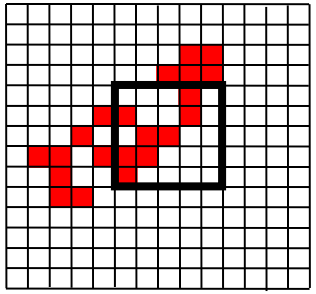
Observed



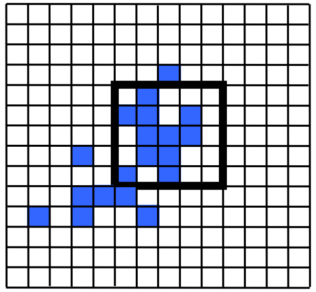
Forecast







observation



forecast

Spatial Verification Inter-Comparison Projects (Spatial Vx ICP):

In the past two decades, the JWGFVR has coordinated the intercomparison of spatial verification methods in two international projects. The intercomparisons were coordinated over a set of common idealized and real case-studies, and have enabled advancing our understanding on the diagnostic capabilities of spatial verification methods. These two projects have involved more than 100 worldwide scientists, meeting regularly at the International Verification Method Workshops, and lead to several peer-reviewed publications.

Information on the Spatial VX ICP, as well as a list of peer-review publications related to spatial verification methods, can be found on the NCAR web page: https://projects.ral.ucar.edu/icp/ .

A concise description of the spatial verification inter-comparison projects is outlined in the following two key references:

- **1st Spatial VX ICP:** Gilleland, E., D. Ahijevych, B.G. Brown, B. Casati, and E.E. Ebert, 2009: Intercomparison of Spatial Forecast Verification Methods. *Weather Forecast.*, **24** (5), 1416 - 1430, doi: 10.1175/2009WAF2222269.1.
- **2nd Spatial VX ICP (MesoVICT):** Dorninger, M., E. Gilleland, B. Casati, M. P. Mittermaier, E. E. Ebert, B. G. Brown, and L. Wilson, 2018: The set-up of the Mesoscale Verification Inter-Comparison over Complex Terrain project. *Bull. Amer. Meteorol. Soc.*, 99 (9), 1887 - 1906, doi: 10.1175/BAMS-D-17-0164.1.

Joint Working Group in Forecast Verification Research

https://community.wmo.int/activity-areas/wwrp/wwrp-working-groups/wwrp-forecast-verification-research

WORLD METEOROLOGICAL ORGANIZATION
Weather · Climate · Water

Mission: The JWGFVR aims to advance the development and application of improved diagnostics and verification methods to assess the quality and enable improvement of weather and environmental predictions, for time scales encompassing from weather forecasts to sub-seasonal and seasonal predictions, to decadal and climate projections.

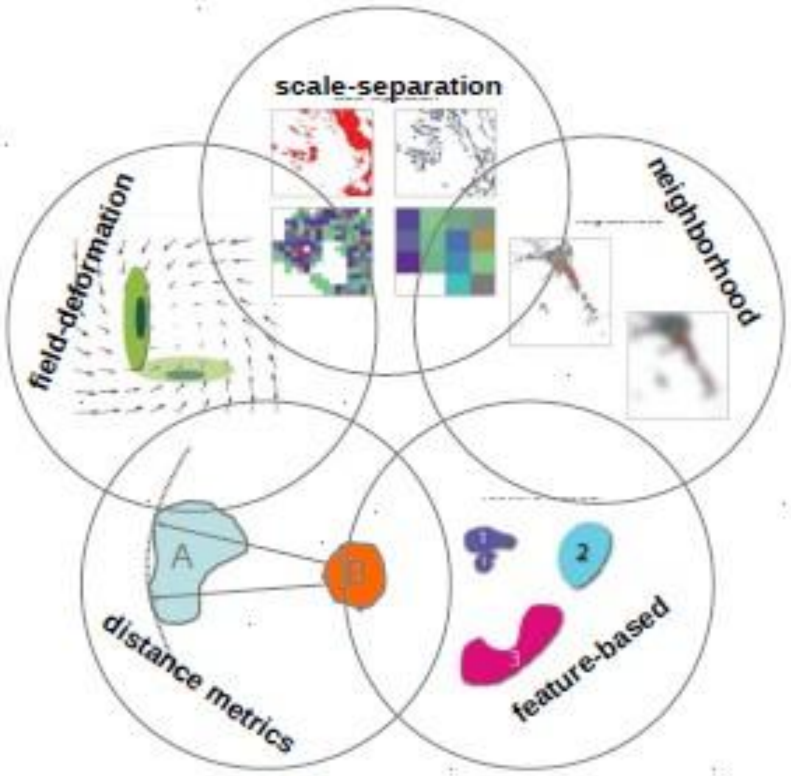
Promote good verification practices :
- Verification web-page
- Verification tutorials
- Verification software
- WMO recommendation reports and verification standards for operational centers -> INFCOM

Advance verification research:
- Spatial verification method intercomparisons
- International verification methods workshops
- Verification challenges
- Special issues & publications

Support verification activities in **WWRP and WGNE/WCRP**
- South Am. PP
- AvRDP2
- HIW
- SAGE
- PCAPS
- InPRHA
- TC-PFP
- URBAN
- PEOPLE
- ADVANCE
- WGNE WP-MIP
- WGSIP
- WGCM

SPATIAL VERIFICATION METHODS

- Account for **coherent spatial structure** and the presence of **features**
- Aim to provide information on **error in physical terms (meaningful verification): e.g.** assess **scale structure** and **displacement error** (separately from **intensity error**)
- Account for **small time-space uncertainties** (avoid **double-penalty** issue)



- Spatial Verification Inter-Comparison Project (ICP) Gilleland et al (2010), BAMS
- Mesoscale Verification Intercomparison in complex Terrain (MesoVICT) Dorninger et al (2018), BAMS

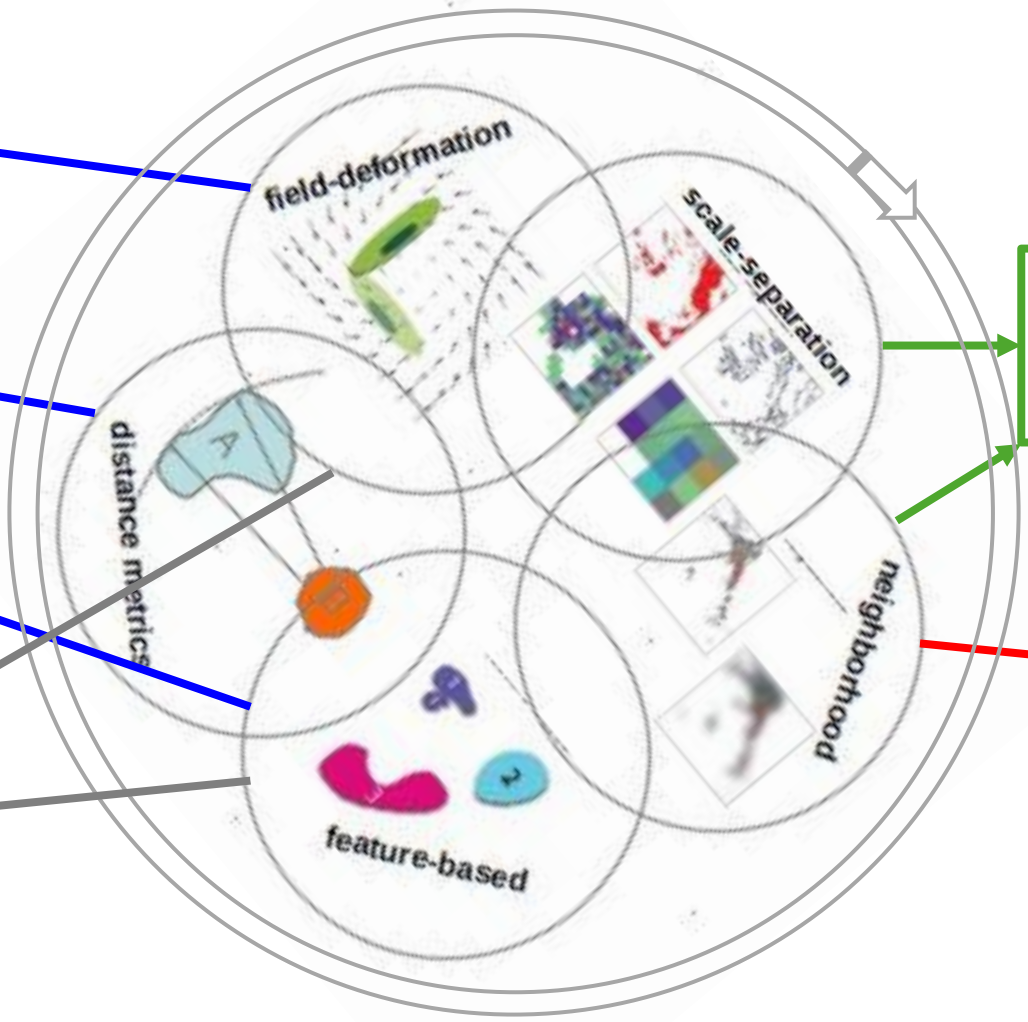
Description, case-studies and references at: https://projects.ral.ucar.edu/icp/

Open source community verification tools:

- SpatialVx R verification package (E.Gilleland)
- MET and METplus (NCAR): https://dtcenter.org/community-code/metplus

Spatial verification approaches – blending classes

Assess the **displacement error**:
- over the whole field => field-morphing techniques
- for binary images => distance metrics
- for individually isolated features => feature-based techniques



Analyze the **scale** dependency of forecast skill (single-band vs low-pass)

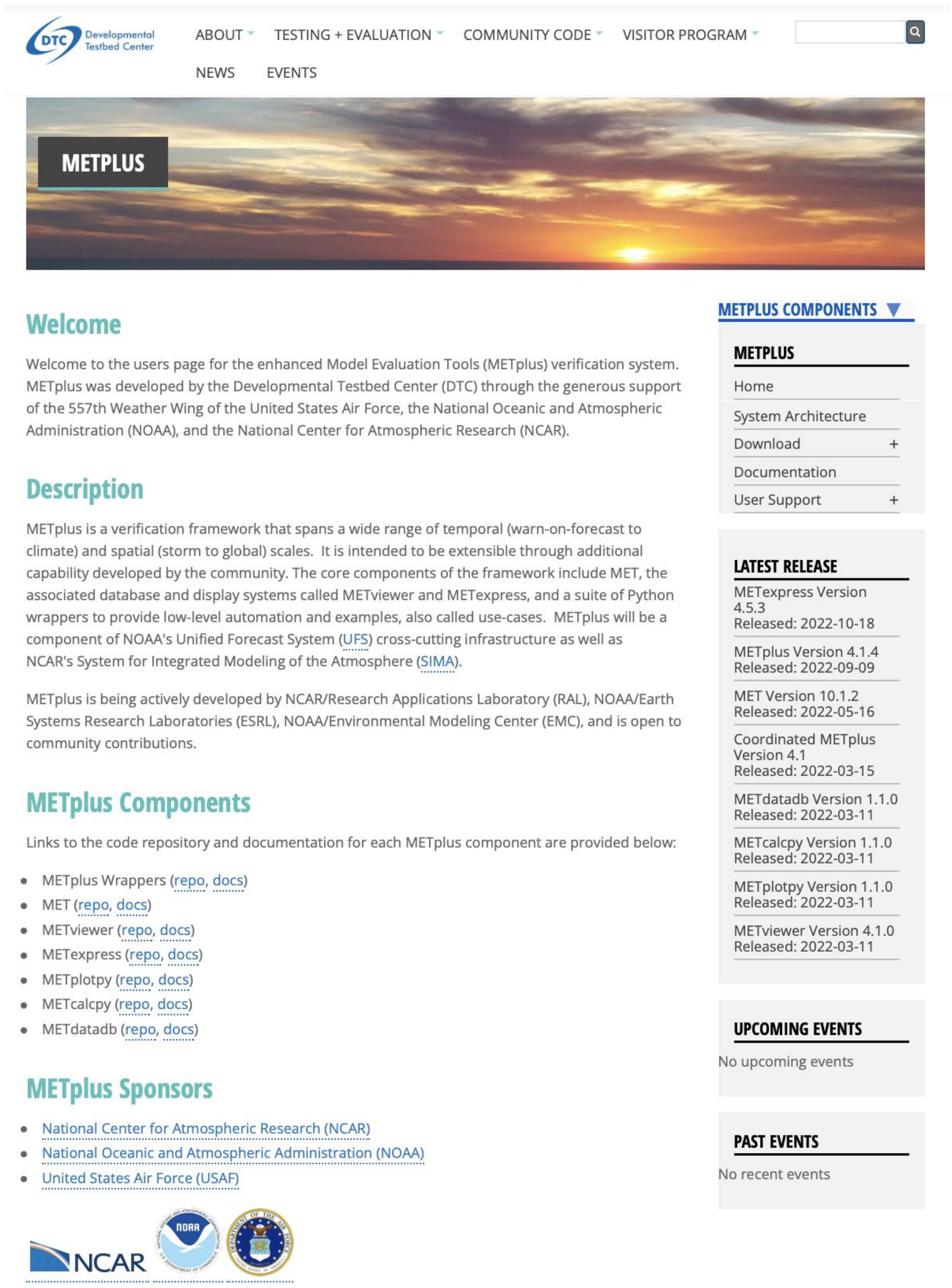
Directly address the **double penalty** issue

Decomposition of the forecast error = displacement error + amplitude error + residual error

As the spatial verification techniques evolve the classes tend to blend: scale-separation and neighbourhood include the displacement error; morphing techniques leverage on spatial filters; ...

Verification Workshops and Tutorials

- 30 July – 1 Aug **2002, Boulder**: "Making Verification More Meaningful" (Barb Brown).
- 15-17 Sept **2004, Montreal**: 2nd International Verification Workshop (Laurie Wilson)
- 31 Jan – 2 Feb **2007, Reading**: 3rd International Verification Workshop & Tutorials (Anna Ghelli) https://www.ecmwf.int/en/learning/workshops-and-seminars/past-workshops/2007-international-verification-methods
  Ebert and Ghelli (2008) ed. Met Apps Special Issue; Casati et al (2008) review article
- 4 -10 June **2009, Helsinki**: 4th International Verification Workshop & Tutorials (Pertti Nurmi) https://space.fmi.fi/Verification2009/
- 1-7 Dec **2011, Melbourne**: 5th International Verification Workshop & Tutorials (Beth Ebert)
  Ebert et al (2013) review article
- 13-19 March **2014, New Delhi**: 6th International Verification Workshop & Tutorials (Raghu Ashrit)
- 3-11 May **2017, Berlin**: 7th International Verification Workshop & Tutorials (Martin Goeber)
  https://www.7thverificationworkshop.de
  Dorninger et al (2018) ed. Met Zet special issue; Dorninger et al (2020) ed. Met Apps special issue.
- 9-20 November **2020, online,** 8th International Verification Method Workshop (Barbara Casati & Manfred Dorninger) https://jwgfvr.univie.ac.at Casati et al (2021) BAMS workshop summary
- 21-25 June **2021, online:** MPE-CDT + JWGFVR verification summer school
  https://mpecdt.ac.uk/mpe-cdt-jwgfvr-forecast-verification-summer-school
- 15-22 May 2024, Cape Town: 9th International Verification Method Workshop (B.Casati) and Tutorial (C.Marsigli), SAWS host (S.Landman): https://ivmw2024.weathersa.co.za MetApps Special Issues (B.Casati C.Coelho)



DTC Developmental Testbed Center

ABOUT · TESTING + EVALUATION · COMMUNITY CODE · VISITOR PROGRAM

NEWS · EVENTS

METPLUS

Welcome

Welcome to the users page for the enhanced Model Evaluation Tools (METplus) verification system. METplus was developed by the Developmental Testbed Center (DTC) through the generous support of the 557th Weather Wing of the United States Air Force, the National Oceanic and Atmospheric Administration (NOAA), and the National Center for Atmospheric Research (NCAR).

Description

METplus is a verification framework that spans a wide range of temporal (warn-on-forecast to climate) and spatial (storm to global) scales. It is intended to be extensible through additional capability developed by the community. The core components of the framework include MET, the associated database and display systems called METviewer and METexpress, and a suite of Python wrappers to provide low-level automation and examples, also called use-cases. METplus will be a component of NOAA's Unified Forecast System (UFS) cross-cutting infrastructure as well as NCAR's System for Integrated Modeling of the Atmosphere (SIMA).

METplus is being actively developed by NCAR/Research Applications Laboratory (RAL), NOAA/Earth Systems Research Laboratories (ESRL), NOAA/Environmental Modeling Center (EMC), and is open to community contributions.

METplus Components

Links to the code repository and documentation for each METplus component are provided below:

- METplus Wrappers (repo, docs)
- MET (repo, docs)
- METviewer (repo, docs)
- METexpress (repo, docs)
- METplotpy (repo, docs)
- METcalcpy (repo, docs)
- METdatadb (repo, docs)

METplus Sponsors

- National Center for Atmospheric Research (NCAR)
- National Oceanic and Atmospheric Administration (NOAA)
- United States Air Force (USAF)

NCAR

METPLUS COMPONENTS

METPLUS

- Home
- System Architecture
- Download +
- Documentation
- User Support +

LATEST RELEASE

METexpress Version 4.5.3
Released: 2022-10-18

METplus Version 4.1.4
Released: 2022-09-09

MET Version 10.1.2
Released: 2022-05-16

Coordinated METplus Version 4.1
Released: 2022-03-15

METdatadb Version 1.1.0
Released: 2022-03-11

METcalcpy Version 1.1.0
Released: 2022-03-11

METplotpy Version 1.1.0
Released: 2022-03-11

METviewer Version 4.1.0
Released: 2022-03-11

UPCOMING EVENTS

No upcoming events

PAST EVENTS

No recent events

METplus (lead: Tara Jensen, NCAR)

https://dtcenter.org/community-code/metplus

- Community based software
- Modular, python: DB, scoring, visualization, ...
- Documentation and User support
  - Sept 2021- April 2022 online tutorials,
  - June 2022 user workshop
- Includes verification tools for
  - Stations and gridded obs
  - Spatial verification
  - Inference (Confidence Intervals)
- Operational at UKmetO, BoM, NOAA, ..

METplus: an optimal tool for research-to-operation tech transfer

Brown et al (2021) “The Model Evaluation Tools (MET): More than a Decade of Community-Supported Forecast Verification” BAMS