Verification and Quality Assurance in HIRLAM-C 2018:
Status and Outlook

OUTLINE

- What are the main purposes of NWP verification of the ALADIN-HIRLAM LAM systems ?
- What are the implications of the evolution towards high resolution and the need to forecast shorter time scales ?
- Where do we need to put efforts in the future ?

Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

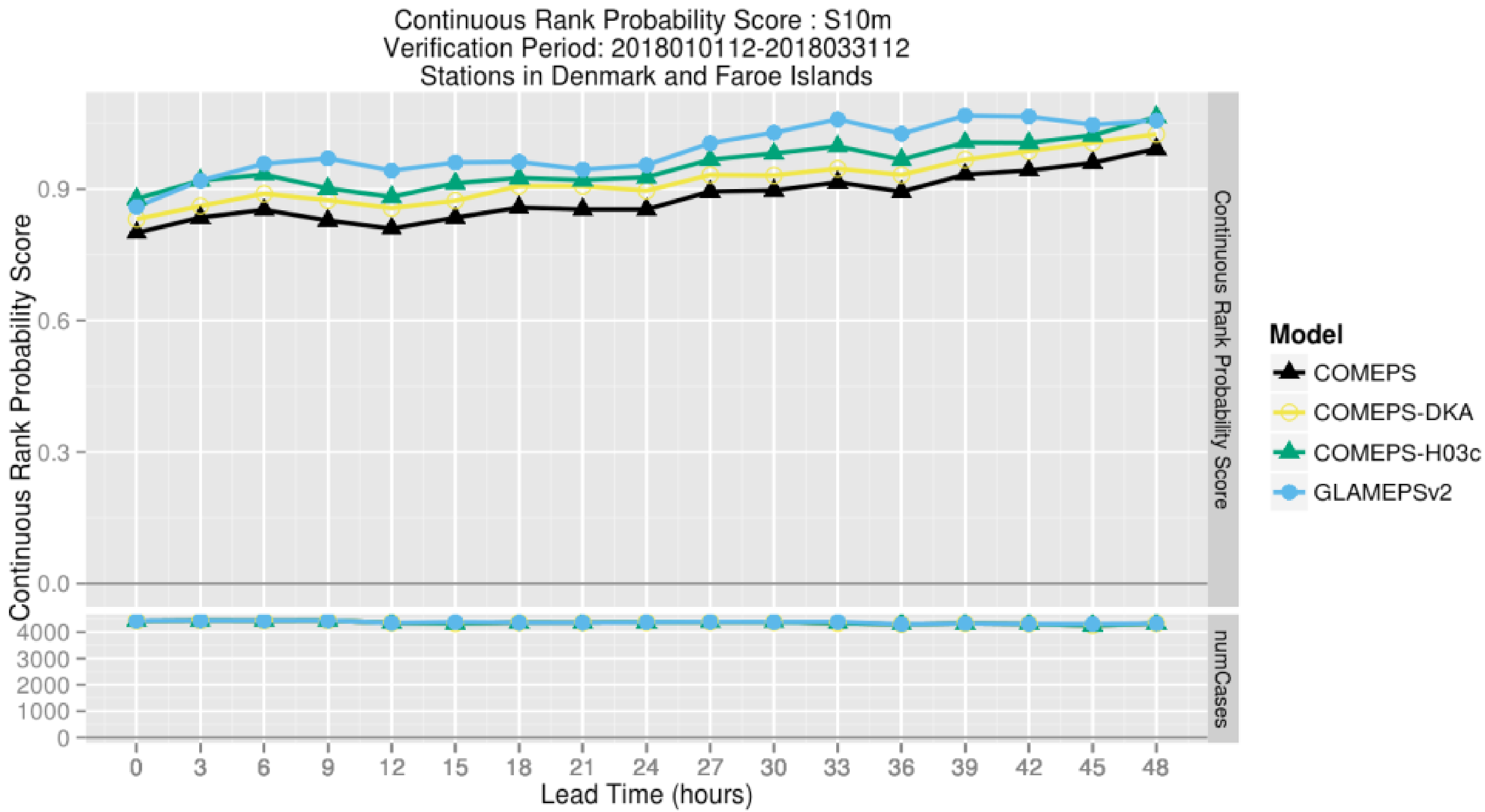
What are the main purposes of NWP verification of ALADIN-HIRLAM LAM systems ?

- **Demonstrate quality improvements in a general sense and that the ALADIN-HIRLAM NWP systems is competitive to host model and alternative model systems available**

- **Create a suitable framework for diagnosing model deficiencies . It needs to be sufficiently advanced to gain insight to relevant aspects of atmospheric modelling**

- **Demonstrate and communicate model skills and limitations to users e.g. forecasters**

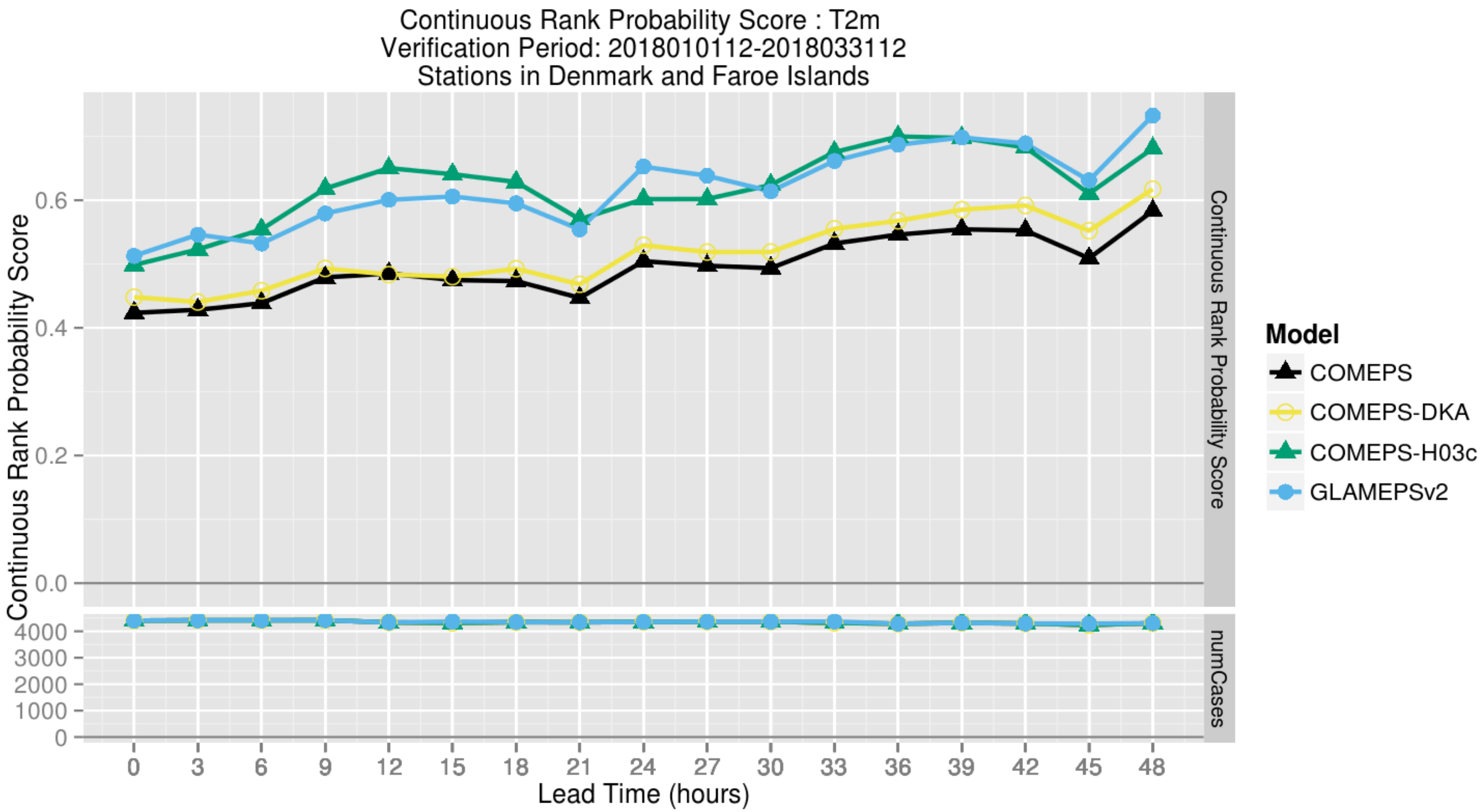
Verification of ensembles (HARP)
Recent results (V10m)



Models run at DMI: COMEPS (full system) COMEPS-DKA is without HIRLAM members, COMEPS-H03c is COMEPS with only HIRLAM.
GLAMEPSv2 is run at lower model resolution at ECMWF

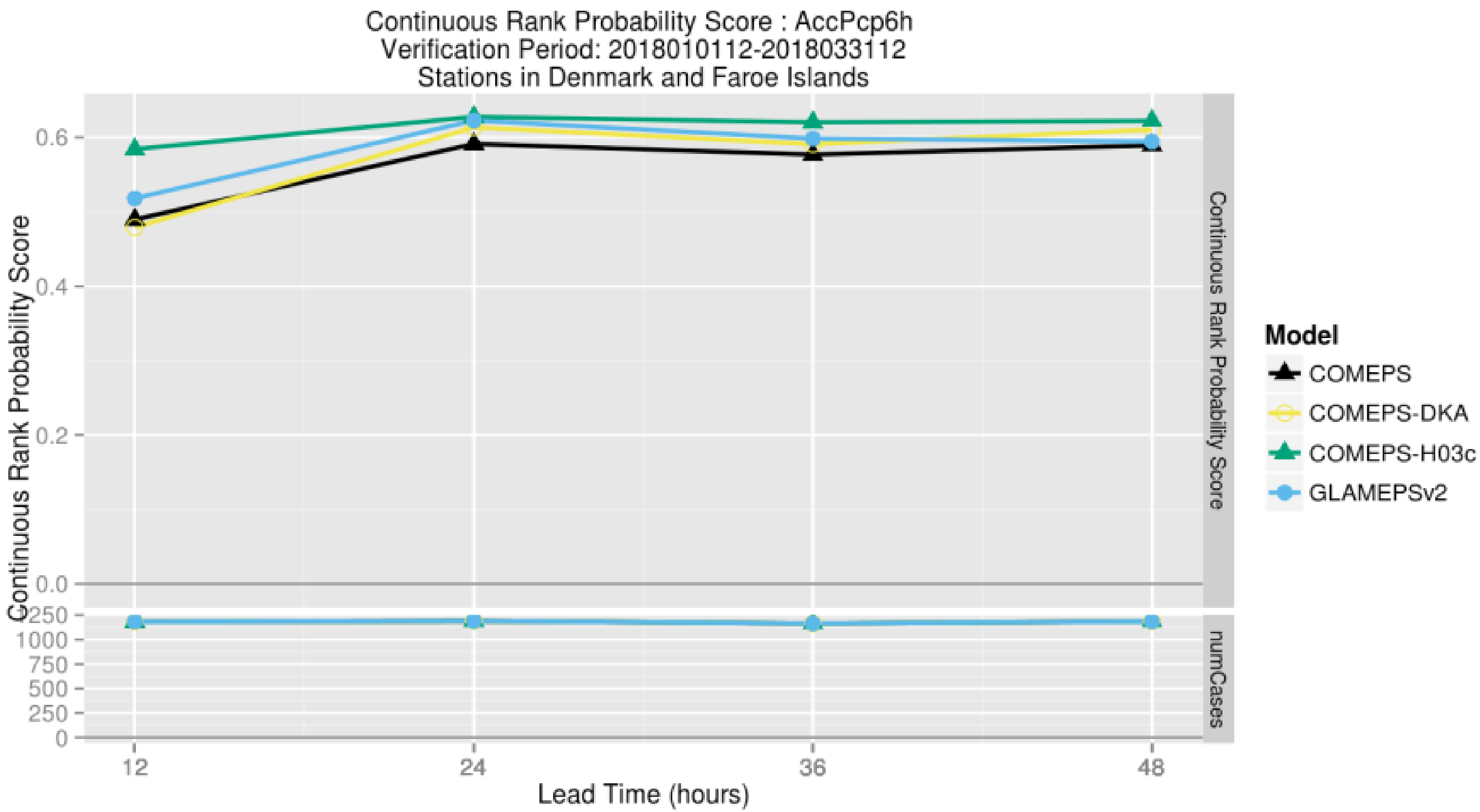
Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

Verification of ensembles (HARP)
Recent results (T2m)



Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

Verification of ensembles



Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

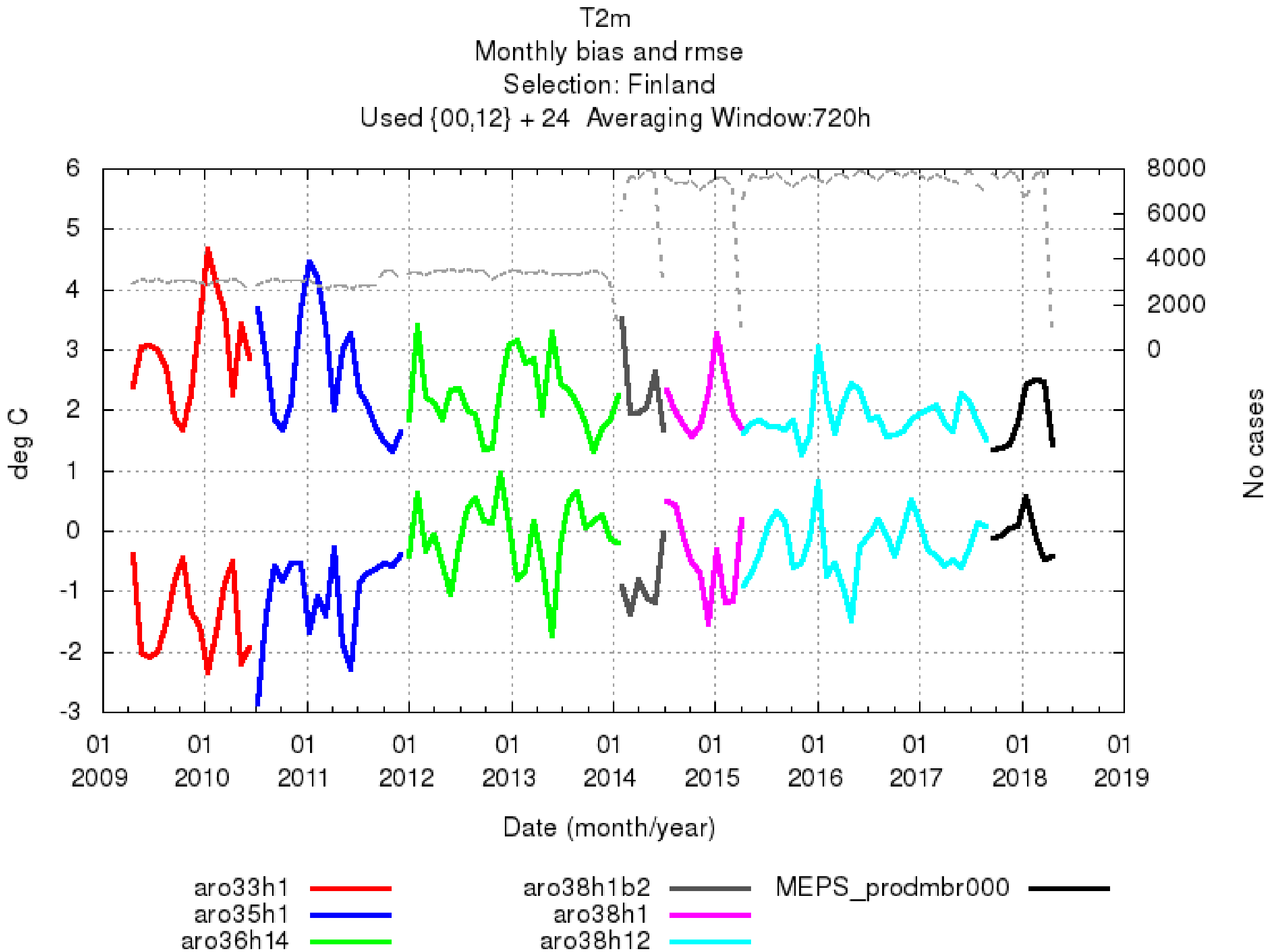
➢ Use of monitor, update of settings in future common experiments

MONITOR is still the most commonly used verification tool, but this should change over the next years. Main limitation of MONITOR: It is based on *point verification* .

Modified common settings involves by the end of 2017:

1) **Time series with a 1 hour resolution:**
2) **Min/Max temp included**
3) **Gust verification supported (max gust within last 1 hour)**
4) **Visibility from Harmonie (list of visibility thresholds chosen**
5) **Choice of verification scores : As a minimum ETS will be used for precipitation verification + frequency bias + KSS + Symmetric extremal dependency index**

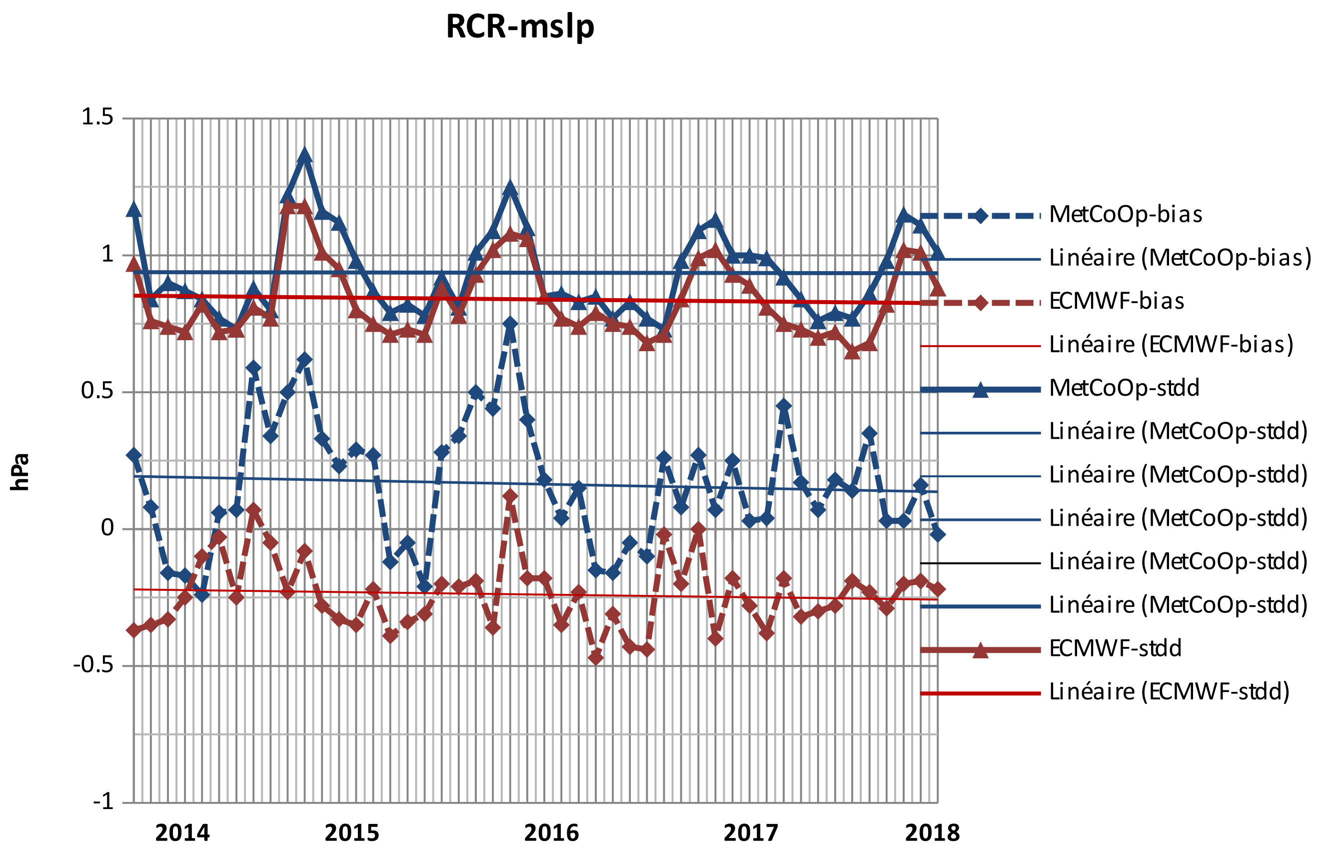
➢ Point verifications: has Harmonie-Arome improved over the years
Figures from FMI since 2009



FMI long time series of Harmonie verifications as used in HIRLAM consortium (2m temperatures in Finland) : Improvement clear over time !

Demonstrate that the ALADIN-HIRLAM NWP systems is competitive to host model and alternative model systems available

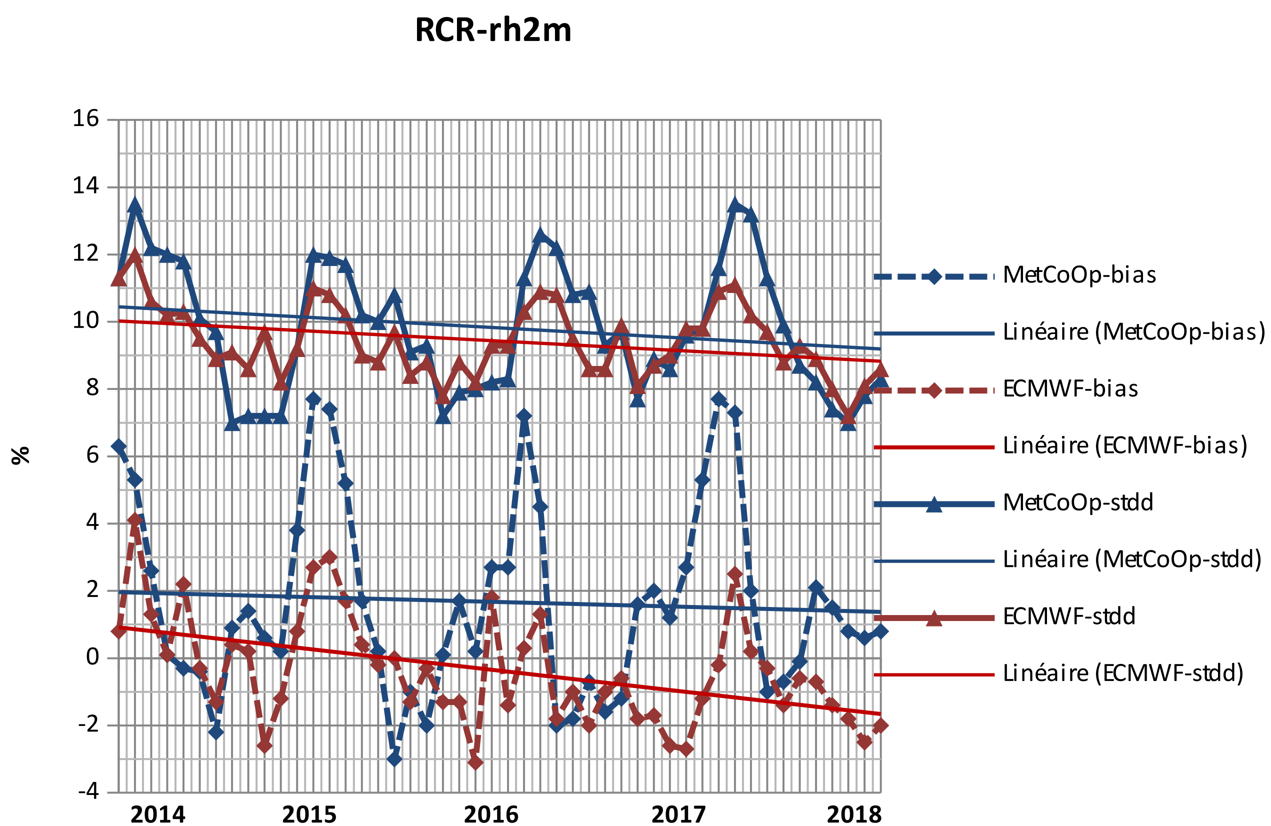
RCR HARMONIE-AROME MSLP verification in MetCoOp (24h) (station list : Scandinavia)



Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

➢ Demonstrate that the ALADIN-HIRLAM NWP systems is competitive to host model and alternative model systems available

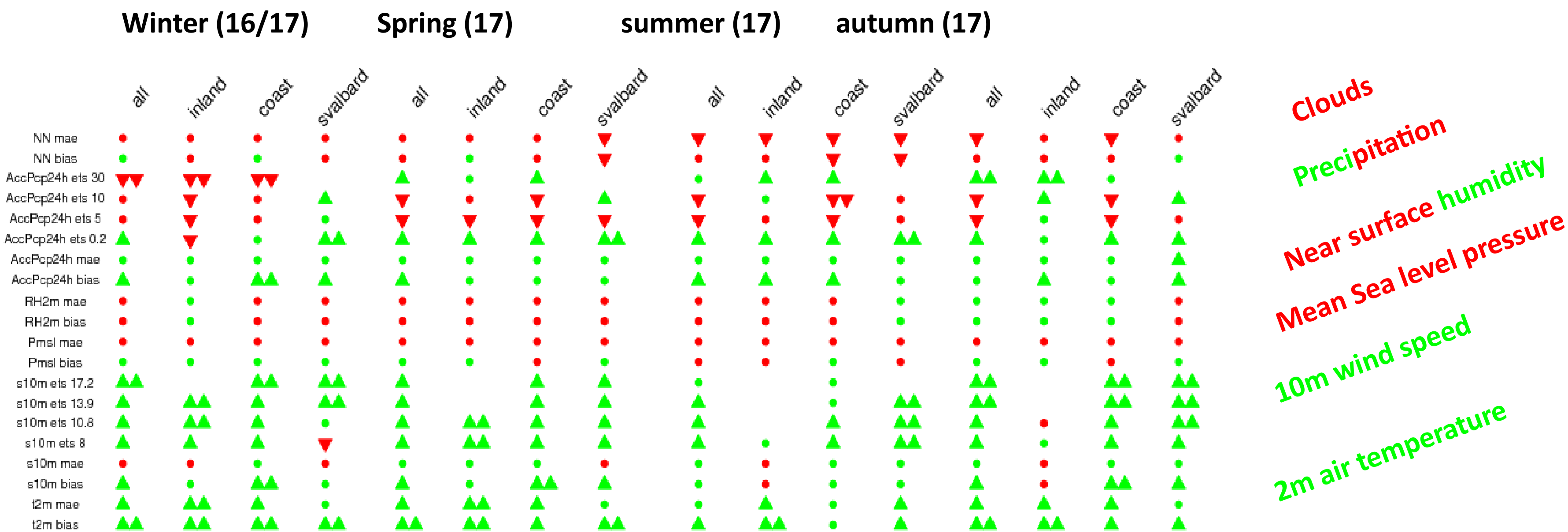
RCR HARMONIE-AROME RH2m verification in MetCoOp (24h) (station list : Scandinavia)



Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

AROME Arctic is very competitive compared with global models

Scorecard AROME Arctic vs IFS HRES



Green AROME Arctic performs better than IFS HRES

Red AROME Arctic performs worse than IFS HRES

(circle small differences, triangles bigger differences)

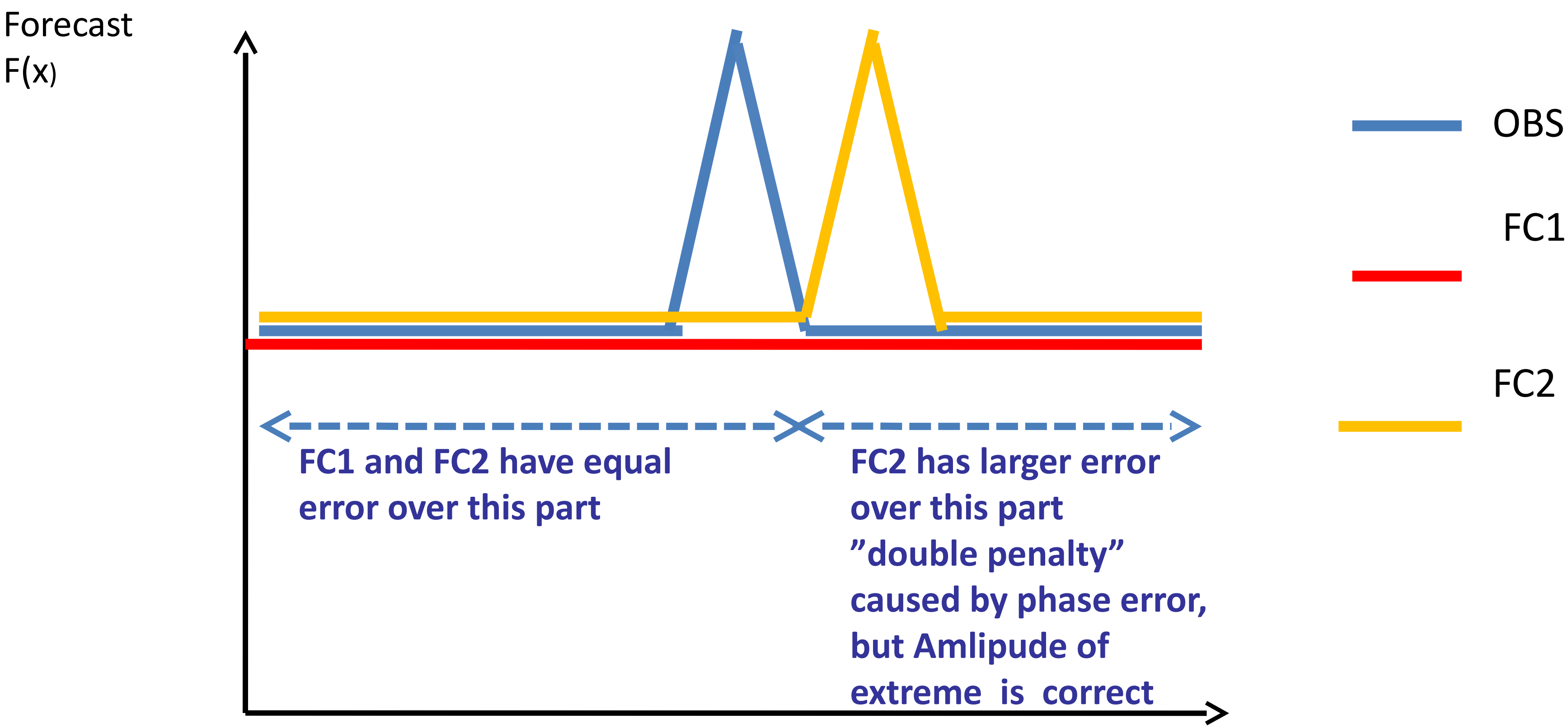
From Morten A. Ødegaard, Met.Norway

Challenges using point verifications
(MONITOR package in HIRLAM-C)

1) Limitations due to ”Double penalty”, in space (and time)

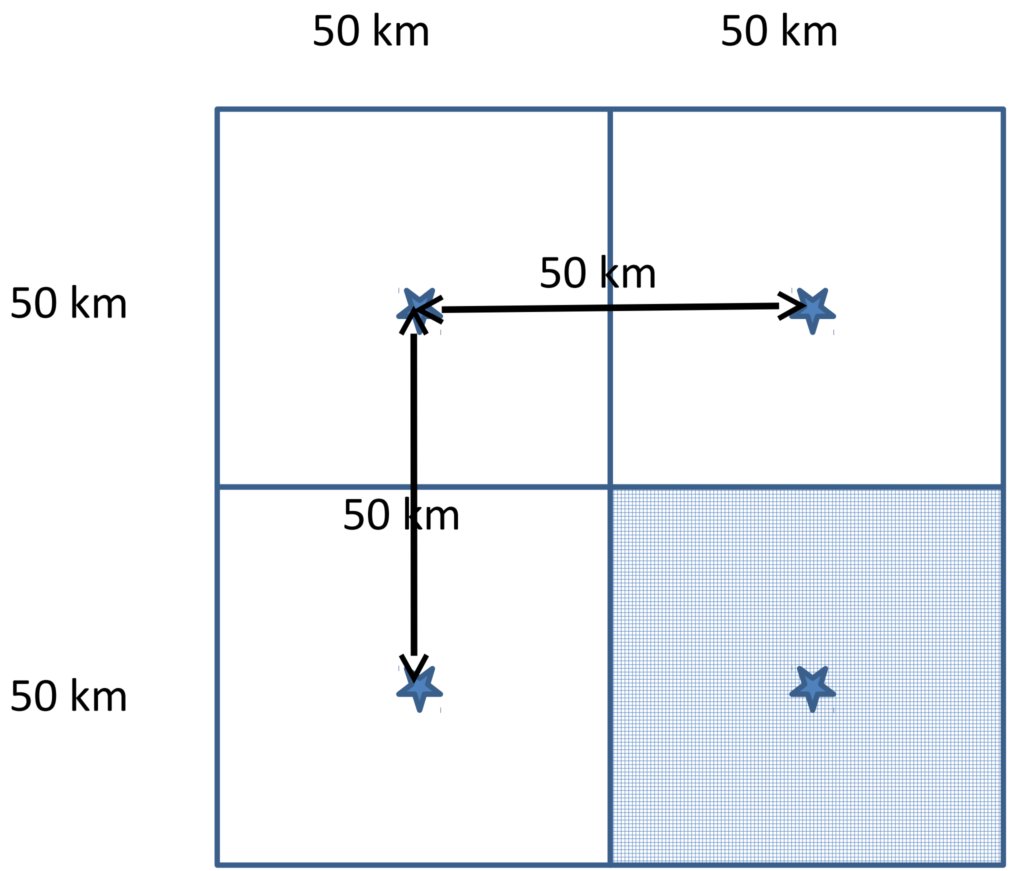
2) Previous (1980s -1990s) reasonable balance between number of observation points and model grid points has been offset by recent big model resolution increases !

Illustration of `double penalty´ issue



Independent space variabel X could be replaced by time t,
FC2 gets no ‘reward ‘ for predicting the amplitude of the fluctuation correctly !

EWGLAM station list typical distances



Present problems :

1) Very low resolution observations compared with present day model resolution with up to 10000 grid points pr. 50*50 km !

2) No guarantee that observations pick up the occurring extremes of the real world

Verification challenges of Greenland

There are strong indications from local reports that the relevant observations (~85 km apart in the picture) do not describe adequately what people in the area experience. Apparently a high resolution Harmonie-Arome 0.75 km grid (Fig1b) produces more realistic wind forecasts than the operational 2.5 km grid model (Fig1a). This represents a verification challenge since in Greenland it is not easy to implement a high density observation network

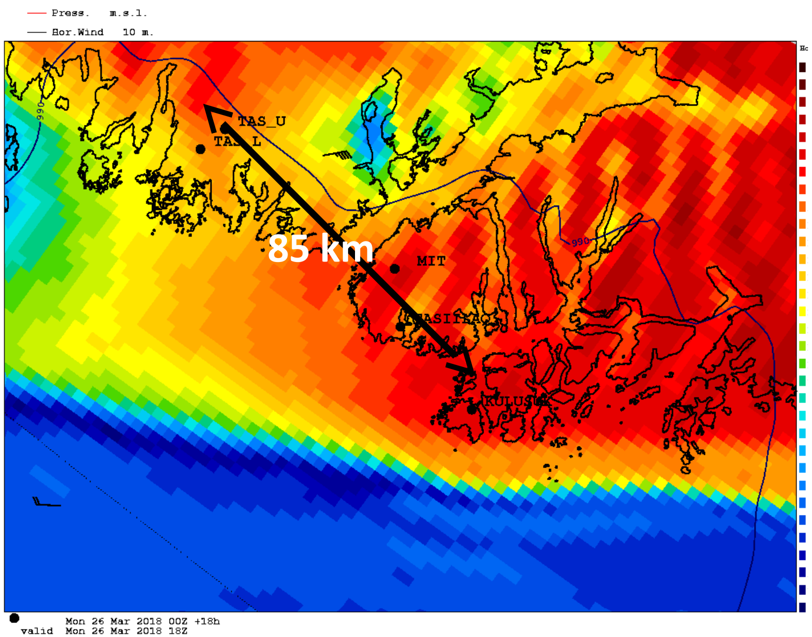


Fig 1a: IGB 2.5 km mesh
10m wind 26/3 00utc+18h

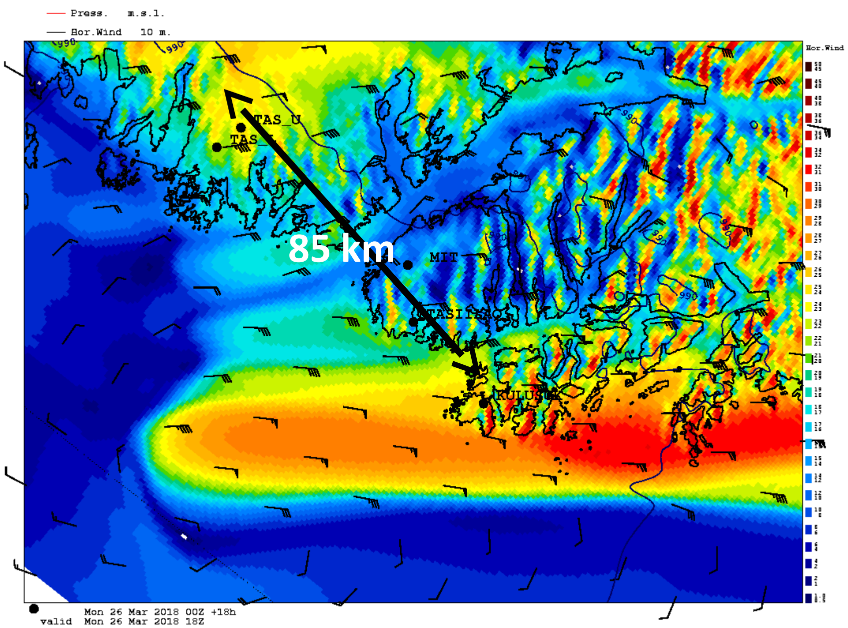
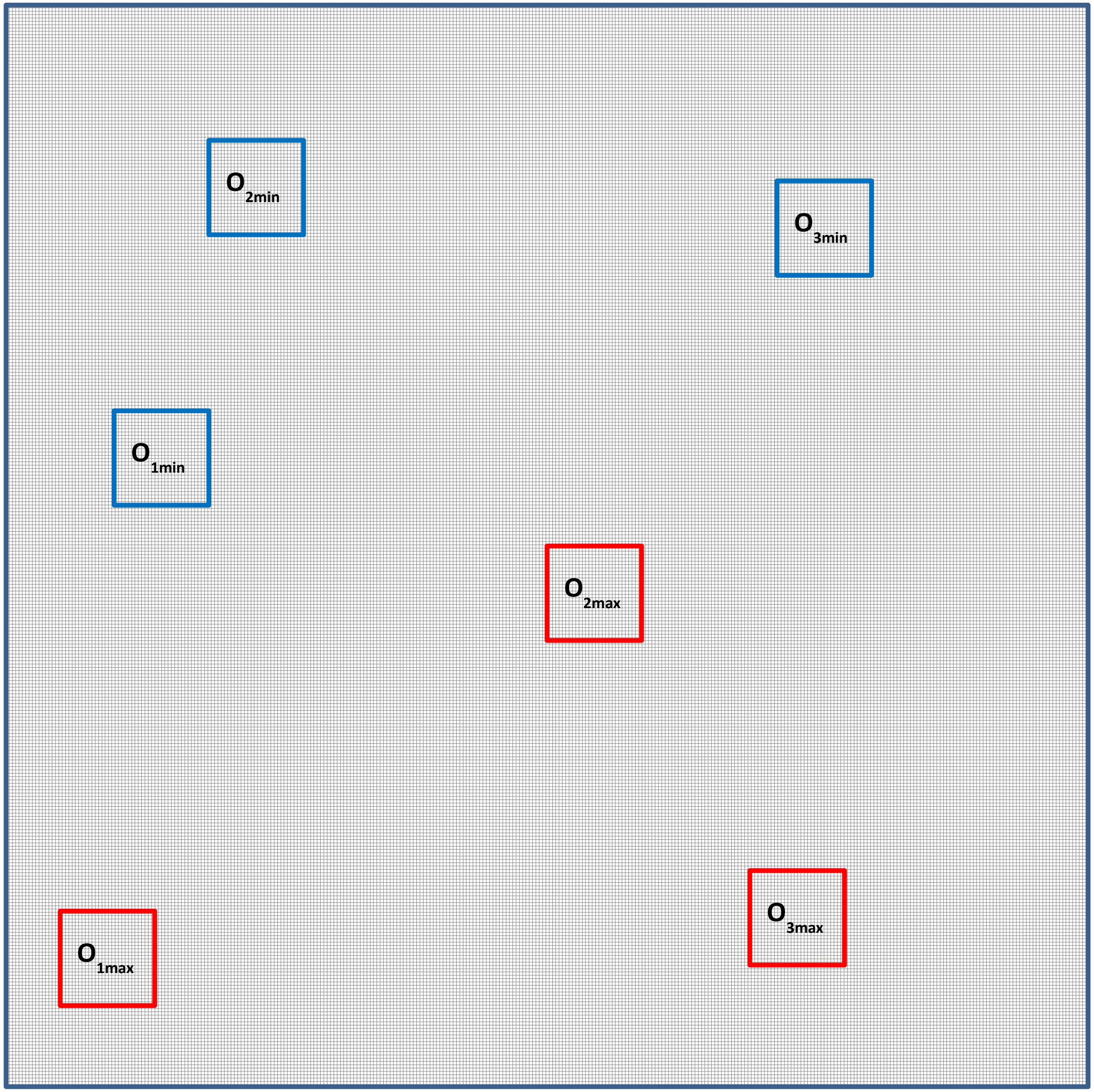


Fig. 1b: TAS 0.75 km mesh
10m wind 26/3 00UTC+18h

Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

CONCLUSION:

- **For present day high resolution forecasting we would like to verify a high resolution forecast field with a dense observation net or an analysis field in the same grid as the model to characterize model quality. <u>The latter is strongly desirable and is done in HARP spatial verification</u> !**



Example of simple spatial verification scheme looking at <u>observed extremes</u> and the ability of the spatial forecast to match the extremes

(SWS scheme by Sass and Yang 2012)

The figure illustrates the 3 highest and the 3 lowest observations in the area and the associated spatial environments considered

Significant weather score SWS
Sass, B.H., and Yang, X, 2012:

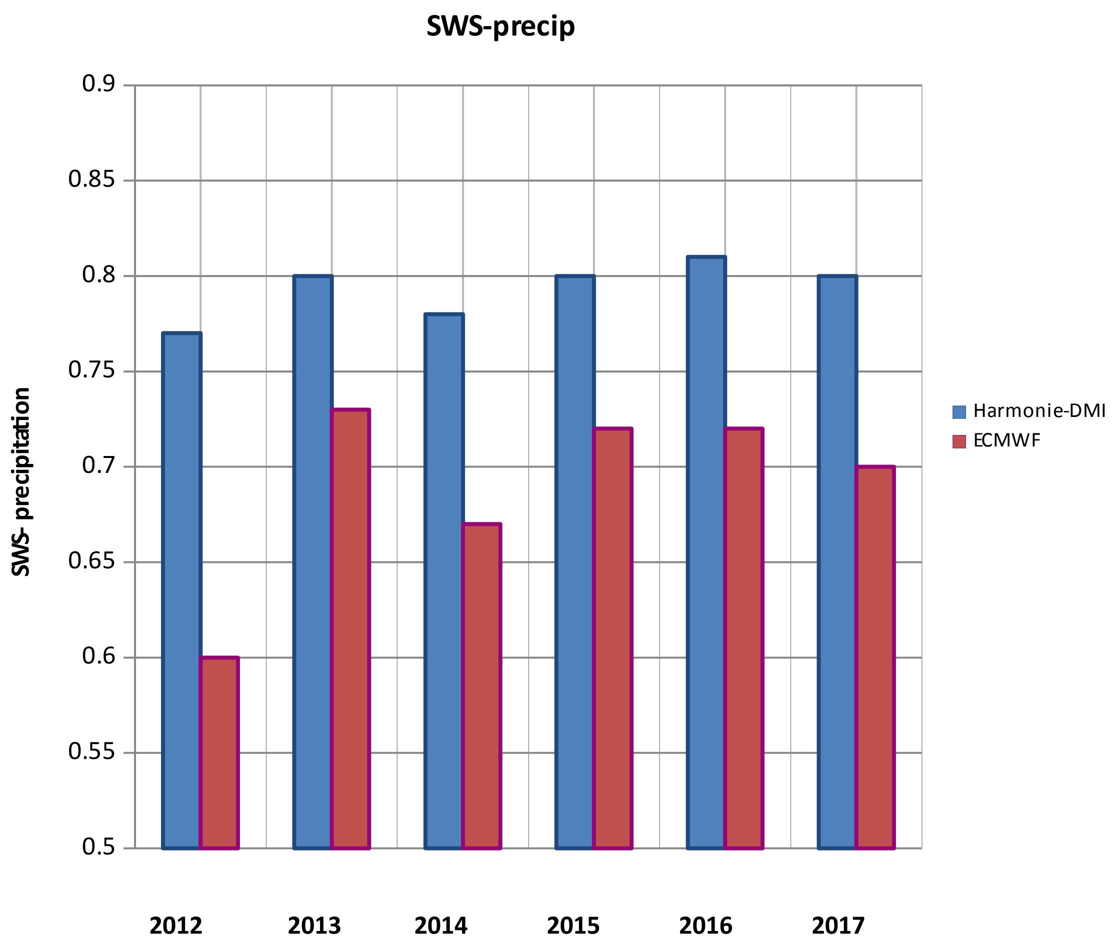
Current implementation:

$SWS_f = (\Sigma F_{meso})/ N$, is the average success of forecasting extreme events in an environment around a local extreme <u>observed</u> in the model domain. The number of observed extremes is a choice of the actual setup.

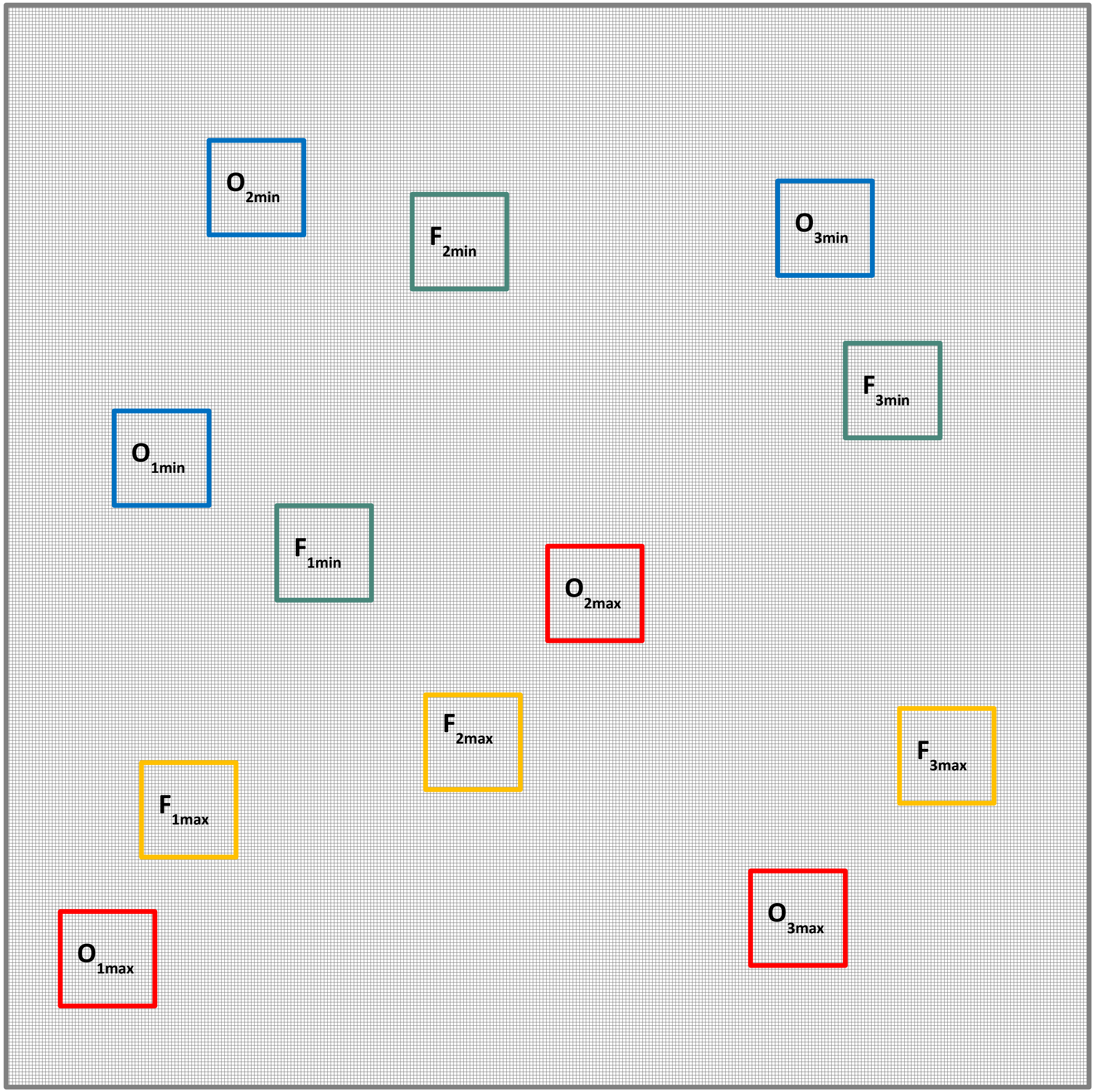
$0 < F_{meso} < 1$, - For each extreme observation event we define a number F_{meso} between 0 and 1 , with 1 being a perfect forecast, 0 a failure and values between 0 and 1 as partially successful. N is the number events (observations) considered in total

Annual SWS index verification at DMI comparing HARMONIE against ECMWF , 2012-2017 (Precipitation , Danish station list)

A benefit of Harmonie –Arome relative to ECMWF using a 15 km box size is seen



Bent Hansen Sass
ASM 16-20 April 2018
Météo-France



Easy extension in HARP:

Extension of SWS scheme by including forecasted local extremes

The figure shows the 3 highest and the 3 lowest observations as well as forecasts in the area.

Properties:

- ***Treatment of extremes (good for warnings)***
- ***Systemtic model bias will be revealed due to the min & max methodology***

Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

Challenge :

- Can we produce reliable high enough resolution analyses

- There will be a close link to data-assimilation and quality control of observations

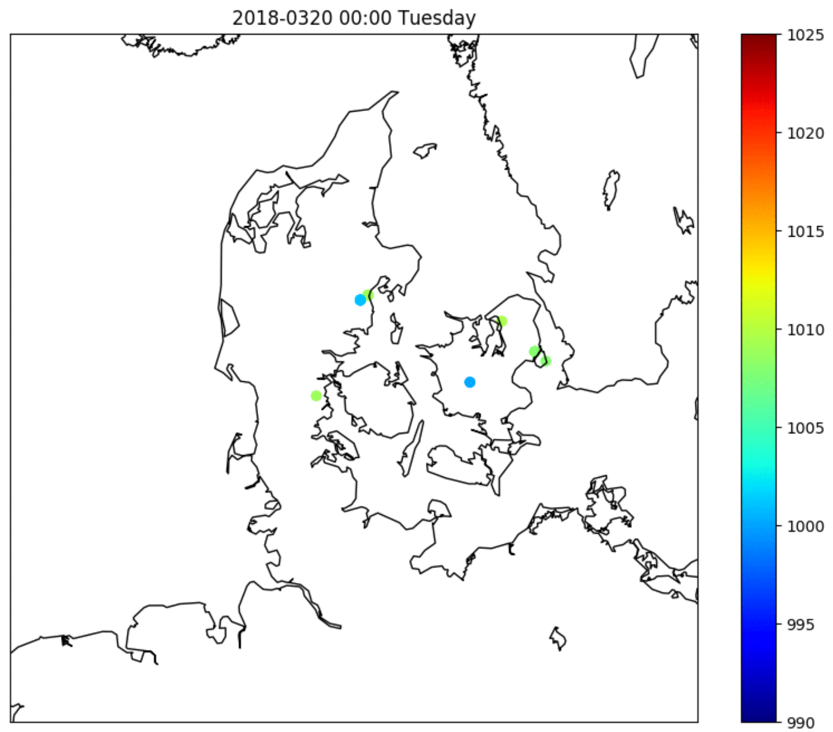
Help from new emerging observations

BIG DATA

Example:

Pressure data from mobile phones

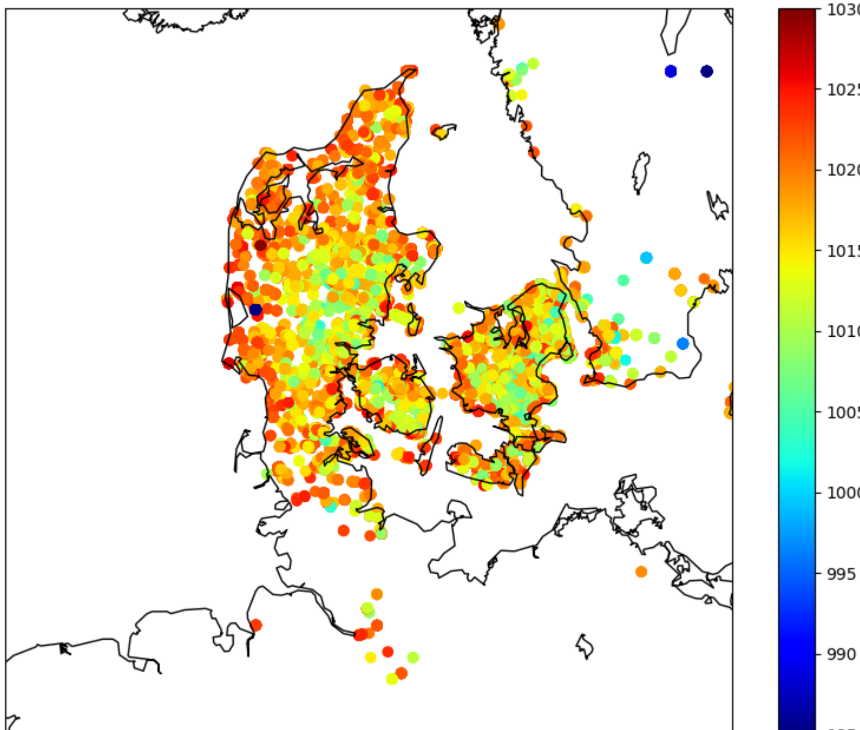
Verification and Quality Assurance in HIRLAM-C 2018: Status and Outlook



New data source:
BIG data from public mobile phones :
tentative collection of surface pressure data from a setup at DMI (PhD student Kasper Hintz) .
About 2 million data from Denmark collected in 1 ½ week. , implying a huge amount of data for quality control and surface analysis

Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

Verification and Quality Assurance in HIRLAM-C 2018: Status and Outlook



Pressure observations over DK (not height corrected)

Bent Hansen Sass
ASM 16-20 April 2018
Météo-France

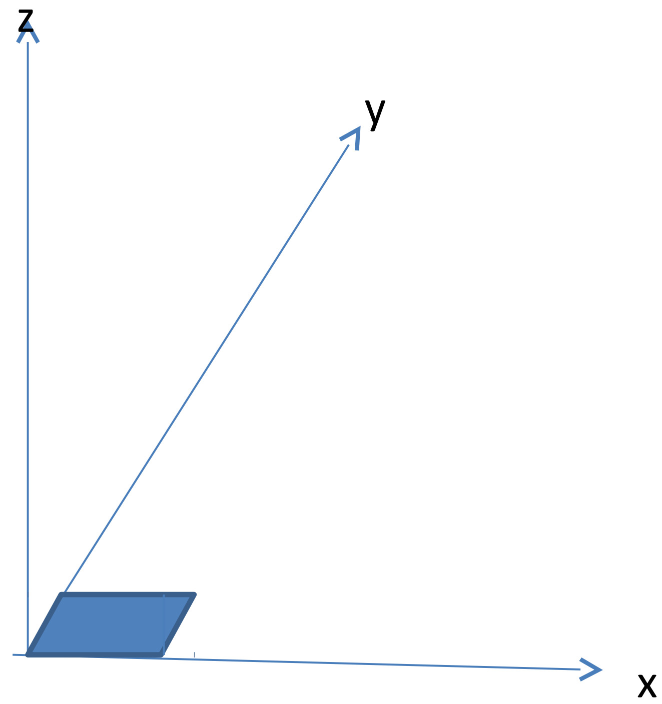
Types of verifications

The future offer opportunities for many types (dimensions) of comparisons between observations and forecasts , e.g. using

- a one-to one comparison at point
- a spatial window
- a temporal window
- a spatial plus temporal window , …

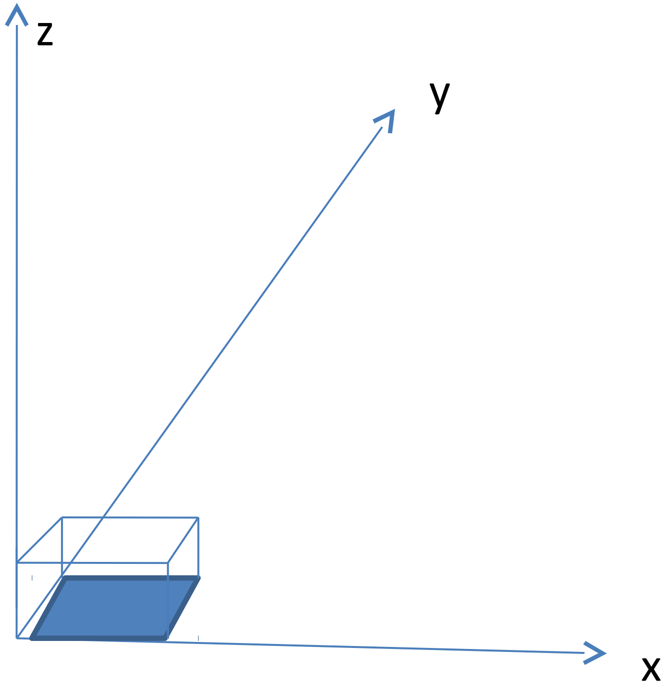
fc / obs	Set of forecasts at irregular observation points (e.g. station list)	1-dim (regular) grid of forecasts	1-dim selection of fuzzy verification, spatial or in time in each point	2-dim grid verification	2-dim grid fuzzy verification, spatial structures	3-dim grid verification	3-dim grid fuzzy verification, Spatial or in time of each point
Set of irregular observation points (e.g. station list)	Traditional point to point verification		EXAMPLE Fuzzy verification in time				NEW 3-dimensional scheme
1-dim (regular) grid of obs or analyzed values		EXAMPLE: Traditional vertical profile line of data or selected points					NEW 3-dimensional scheme
2-dim (regular) grid of obs or analyzed values				Traditional point to point verification	FSS, SAL, SWS, . . .		NEW 3-dimensional scheme
3-dim (regular) grid of obs or analyzed values						Traditional point to point verification	NEW 3-dimensional schemes

2D-fuzzy



Quality Assurance in HIRLAM-C 2017: Status and Outlook

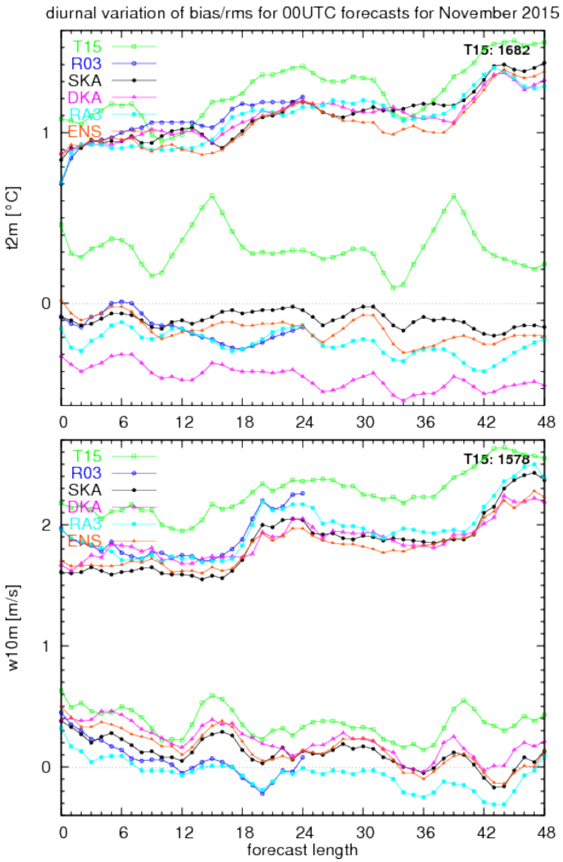
3D-fuzzy verification

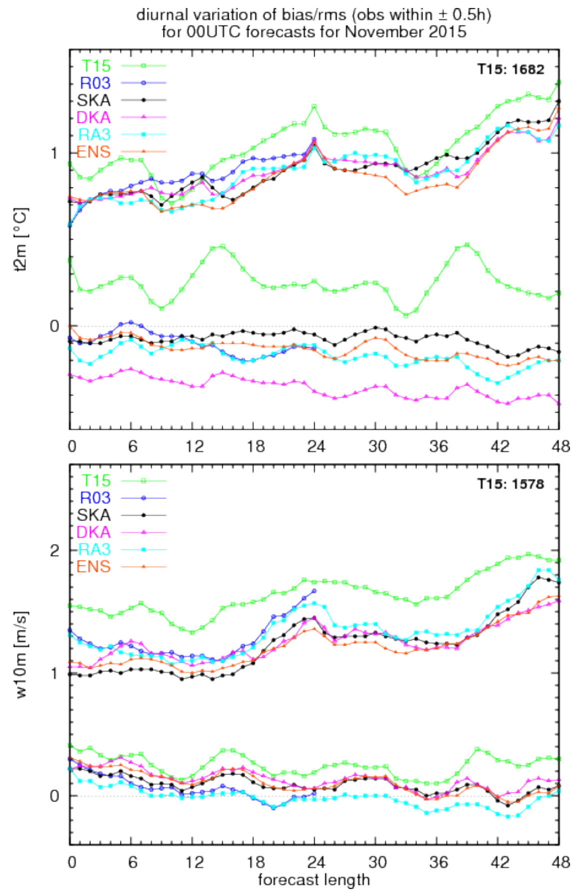


- Bent Hansen Sass
- ASM 16-20 April 2018
- Météo-France

"TIME WINDOW" verification example
Sensitivity of DMI models to verification time window
verification at precise time (left)
allowing a time window of +/- 0,5 hours (right)

Temperature and wind predictions in November 2015:
CONCLUSION : observation time window matters !
What is important to the users ?





- Bent Hansen Sass
- ASM 16-20 April 2018
- Météo-France

Harmonie Users' Meeting 2018

- **When ?   6 -7 November 2018**
- **Where ? AEMET**
- **We ask for a preliminary indication of the participation in the workshop from your organization by 1 June (number of participants and possibly names)**
- **Reports on interaction with Users are still found at**

  https://www.hirlam.org/trac/wiki/CommunicationWithUsers

Harmonie Users' Meeting 2018

Why is it important ?

- to get an overview of how Harmonie is being used in the different operational centers.
- to know the challenges that forecasters experience? Which topics are particularly relevant to take up in future model developments ?
- to let the Hirlam-C management, researchers and forecasters discuss the status of Hirlam-C developments including new information on model improvements
- to provide examples from new setups of Harmonie dedicated to special applications where e.g. high model resolution , ensemble techniques, run frequency and data-assimilation methodology plays a role to improve the quality of model output
- to provide better insight to an efficient use of Harmonie ensembles and post processed products
- to discuss possible new ways of interaction between forecasters in the future and a more efficient link to developers.

Verification and Quality Assurance in HIRLAM-C 2018: Status and Outlook

Where should we put future efforts ?

- Continue present plans to develop HARP further (v3 + v4)
- More work on new schemes to characterize properties of model 2D-3D (4D) features compared with corresponding observed feature (QA2.4)
- Combination of ensembles and spatial (2D /3D) schemes (QA2.5, HARP v4 +)
- Adequate follow up on communication with users of ALADIN-HIRLAM NWP systems
- More work on statistical significance ?

Merci beaucoup
pour votre attention !

REFERENCES

[1] http://www.cawcr.gov.au/projects/verification/
Theme: Forecast Verification methods Across Time and Space Scales , 7th International Verification Methods Workshop (*WWRP/WGNE Joint Working Group on Forecast Verification Research*)

[2] Diebold, F.X. and Mariano, R., 1995:
“Comparing Predictive Accuracy,”Journal of Business and Economic Statistics, 13, 253-265
,
[3] Gilleland, E., Ahijevych, D.A., Brown, B.G., Ebert, E., 2010: Verifying forecasts spatially, BAMS 1365-1373

[4] Murphy, A.H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8**, 281-293.

[5] Roberts , N.M., and Lean, H.W., 2008:
Scale –selective verification of rainfall accumulations from high-resolution forecasts of convective events . Mon. Wea. Rev., 136, 78-97

[6] Sass, B.H., and Yang, X, 2012: A verification score for high resolution NWP: Idealized and pre-operational tests , HIRLAM Tech. Rep. no. 69, 28pp [available from www.hirlam.org]

[7] Wernli, H. , Paulat, M., Hage, M. and Frei, C., 2008:
SAL- A Novel Quality Measure for the Verification of Quantitative Precipitation Forecasts , Mon. Wea. Rev. , 136, 4470 – 4487