Quality Assurance in HIRLAM-C : Status and developments

1) Meteorological model performance/deficiencies:
   a) Recent results from objective - and subjective verification (MSLP, V10m, T2m, Clouds and fog, precipitation)
   b) Ongoing work to improve model

2) Score cards : a new trend in HIRLAM-C
3) Common verification tools (HARP)
4) Special challenges related with orography and related progress using sub-km resolution
5) Conclusions
6) Workshops and reports

1) **Model performance:**
a) **Results from objective and subjective verification**

Operational quality of Harmonie CY40h1.1

- **Results of standard scores obtained with MONITOR for operational models in RCR centres of MetCoOp and in AEMET**

- **Additional input from forecasters , based on reports from**

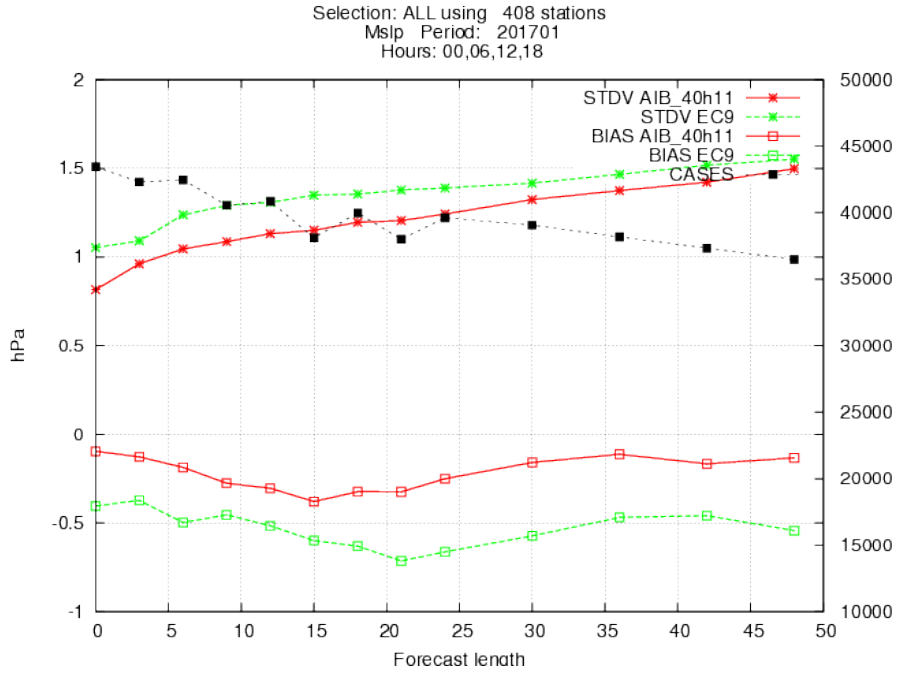
https://hirlam.org/trac/wiki/CommunicationWithUsers

1) Model performance: MSLP

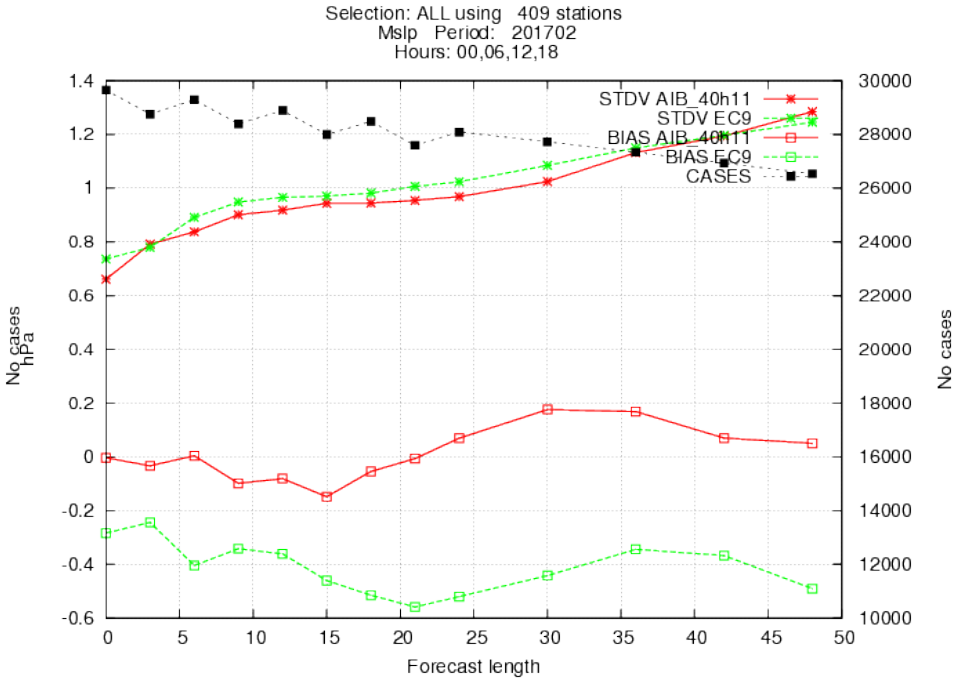
Results of MONITOR, January-February 2017

(AEMET RCR results)

Comparison (BIAS and STDV against ECMWF)



January 2017



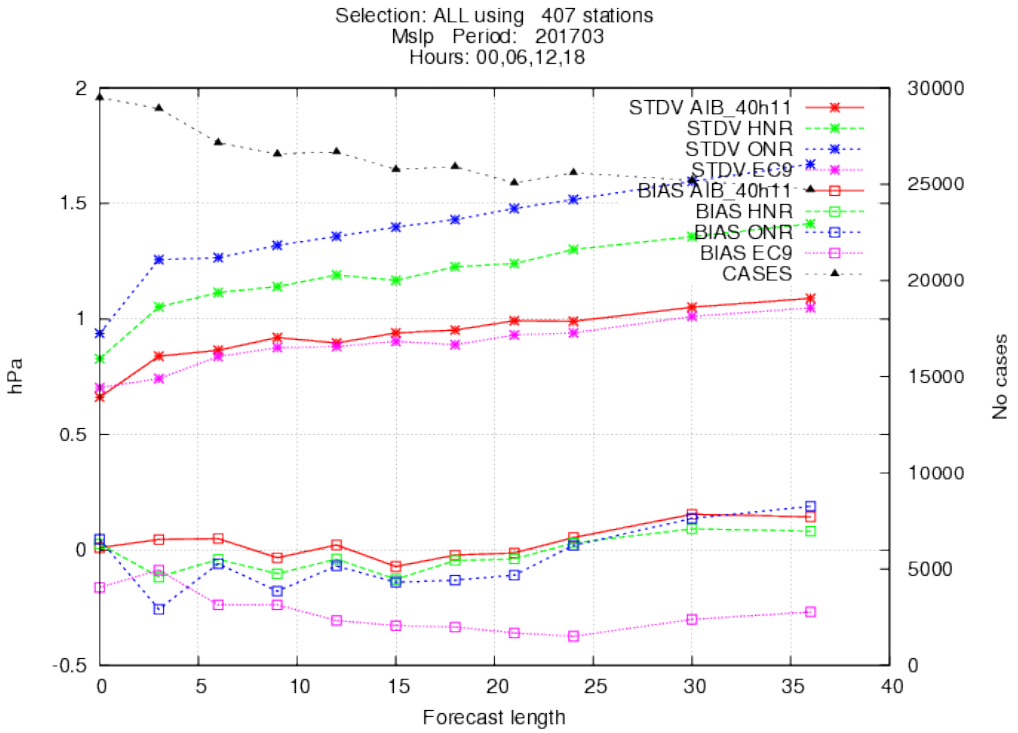
February 2017

1) Model performance: MSLP

Results of MONITOR, March 2017

(AEMET RCR results)

Comparison (BIAS and STDV) against ECMWF and other models



Conclusion MSLP at AEMET 1st Quarter 2017:

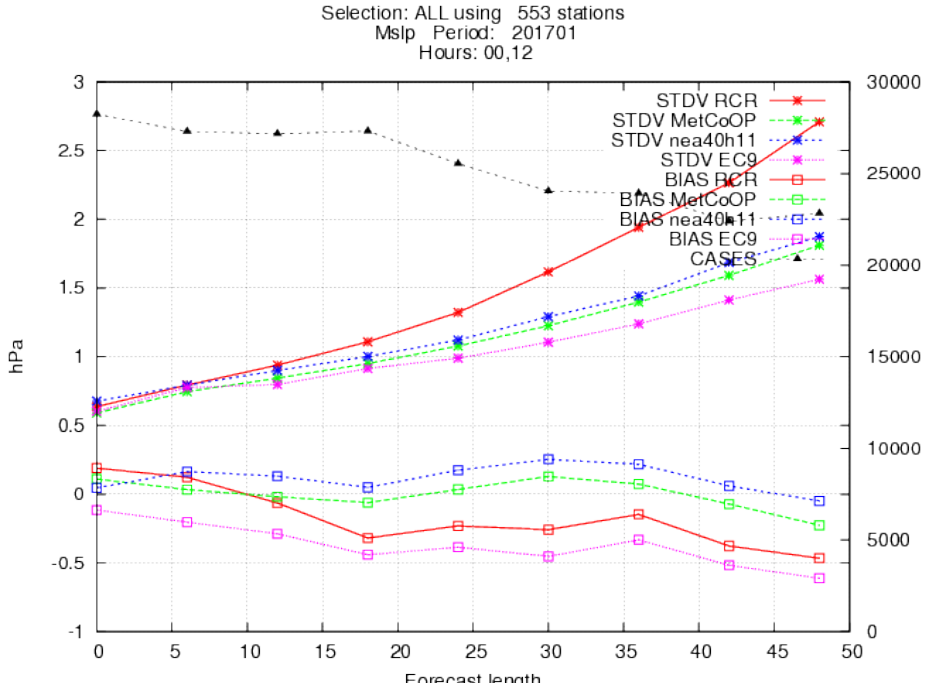
MSLP bias of CY40h1.1 scores better than ECMWF, - also standard deviation of CY40h1.1 is superior to that of ECMWF in Jan-Feb and also when compared with HNR and ONR

1) Model performance: MSLP

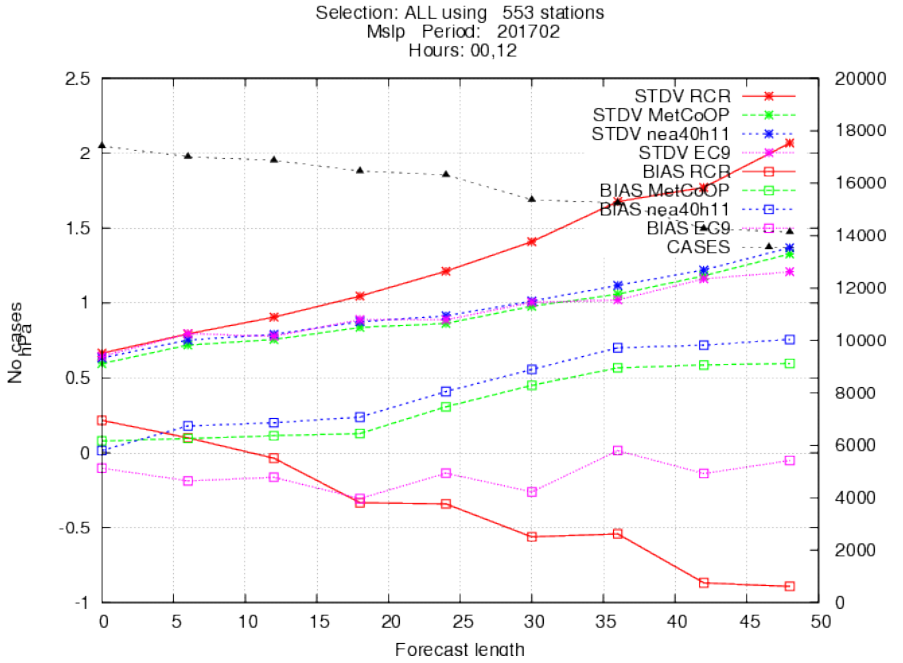
Results of MONITOR, Jan-February 2017

(MetCoOp RCR results)

Comparison (BIAS and STDV against ECMWF)



January 2017



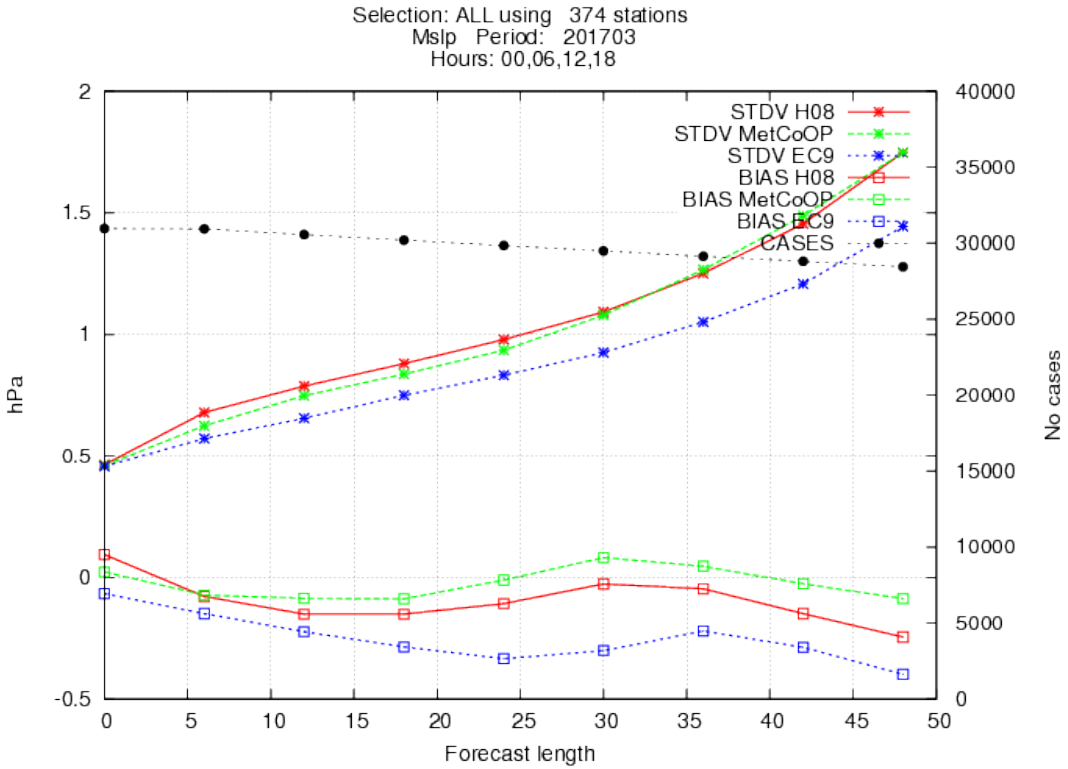
February 2017

1) Model performance: MSLP

Results of MONITOR, March 2017

(RCR MetCoOp results)

Comparison (BIAS and STDV) against ECMWF and other models



Conclusion MSLP at MetCoOp 1st Quarter 2017:

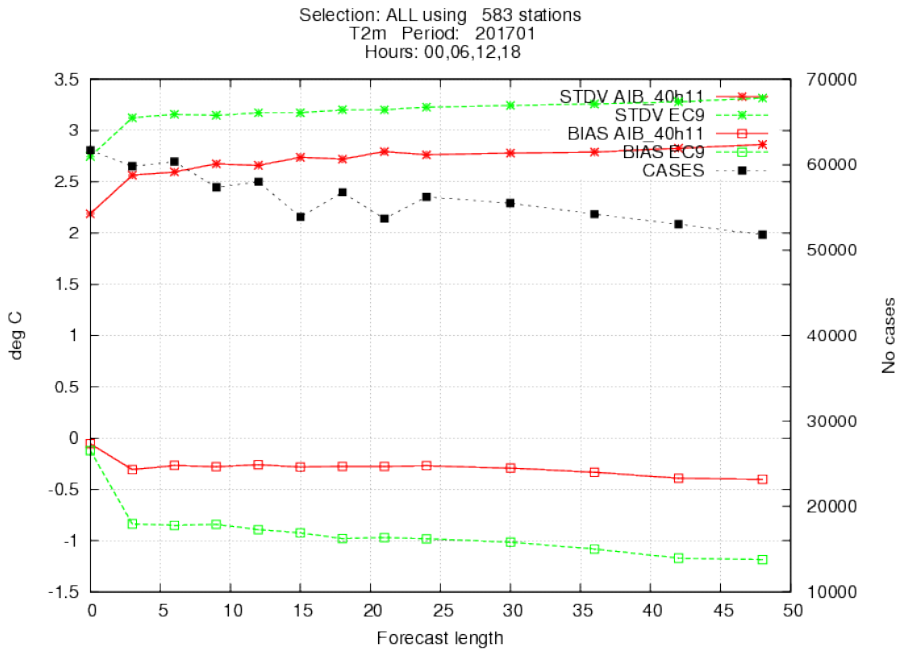
Results relative to ECMWF: Improved bias in January and February, STDV inferior in January and March, close to ECMWF in February: General conclusion: MSLP scores acceptable/good in 1st Quarter of 2017

1) Model performance: T2m

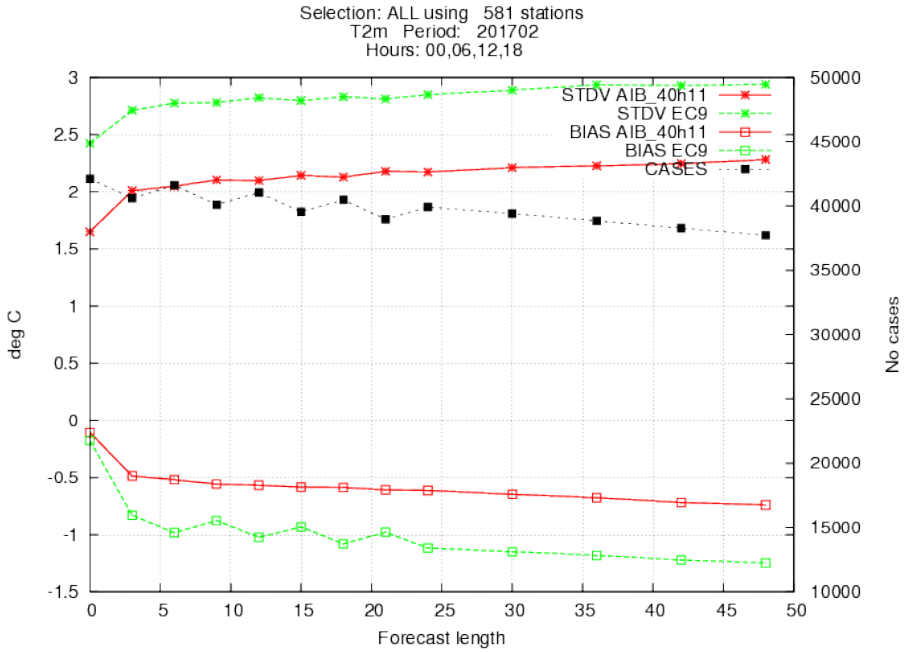
Results of MONITOR, Jan-Feb 2017

(RCR AEMET results)

Comparison (BIAS and STDV) against ECMWF and other models



January 2017



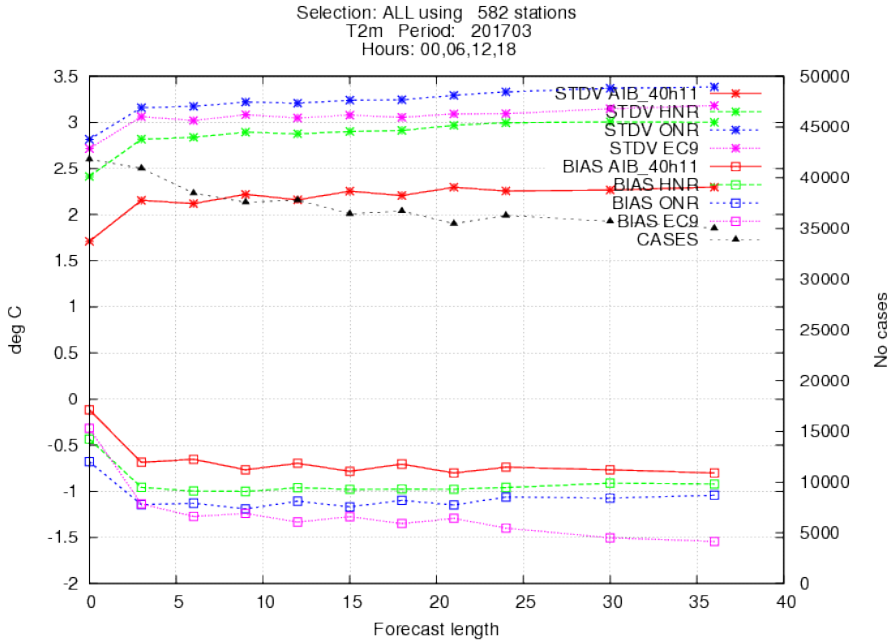
February 2017

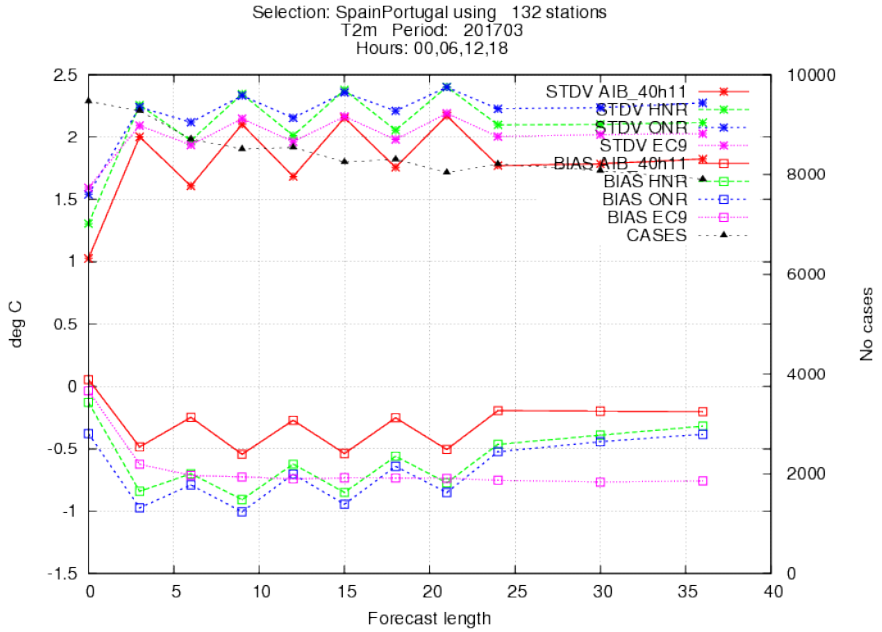
1) Model performance: T2m

Results of MONITOR, Jan-Feb 2017

(RCR AEMET results)

Comparison (BIAS and STDV) against ECMWF and other models





Conclusion T2M at AEMET 1st Quarter 2017:

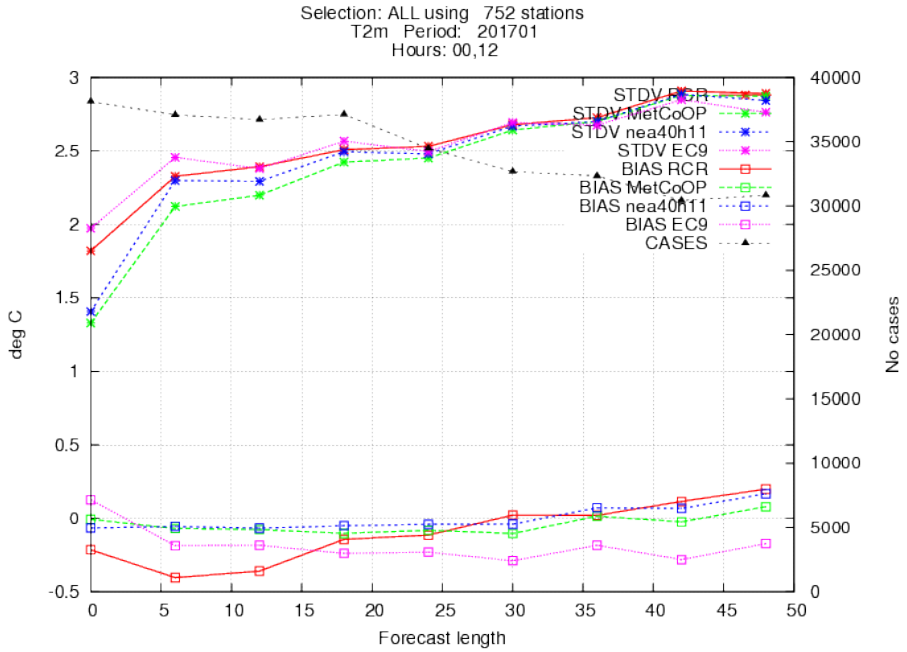
T2M much improved compared with ECMWF (BIAS and STDV) durinng 1st Quarter of 2017. Bias in March 2017 over stations in Spain + Portugal relatively small

1) Model performance: T2m

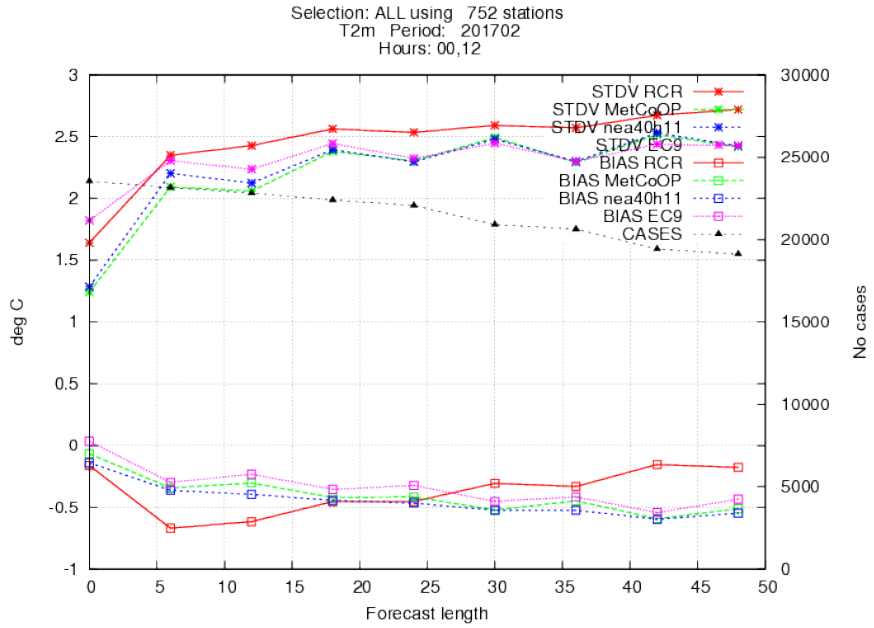
Results of MONITOR, Jan-Feb 2017

(RCR MetCoOp results)

Comparison (BIAS and STDV) against ECMWF and other models



January 2017



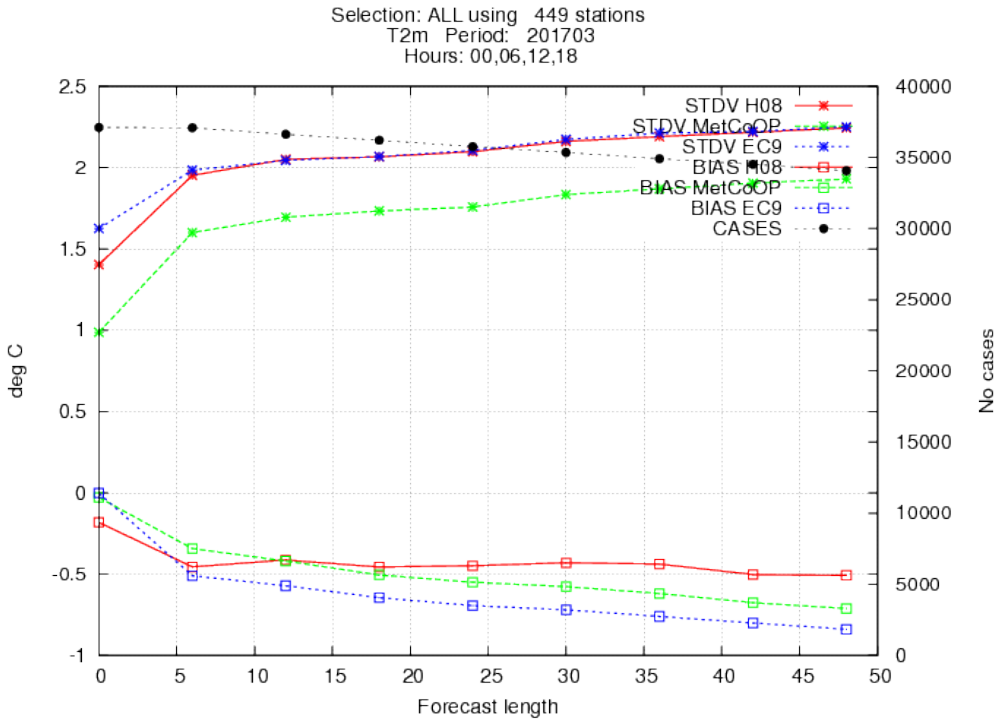
February 2017

1) Model performance: T2m

Results of MONITOR, March 2017

(RCR MetCoOp results)

Comparison (BIAS and STDV) against ECMWF and other models



Conclusion T2M at MetCoOp 1st Quarter 2017:

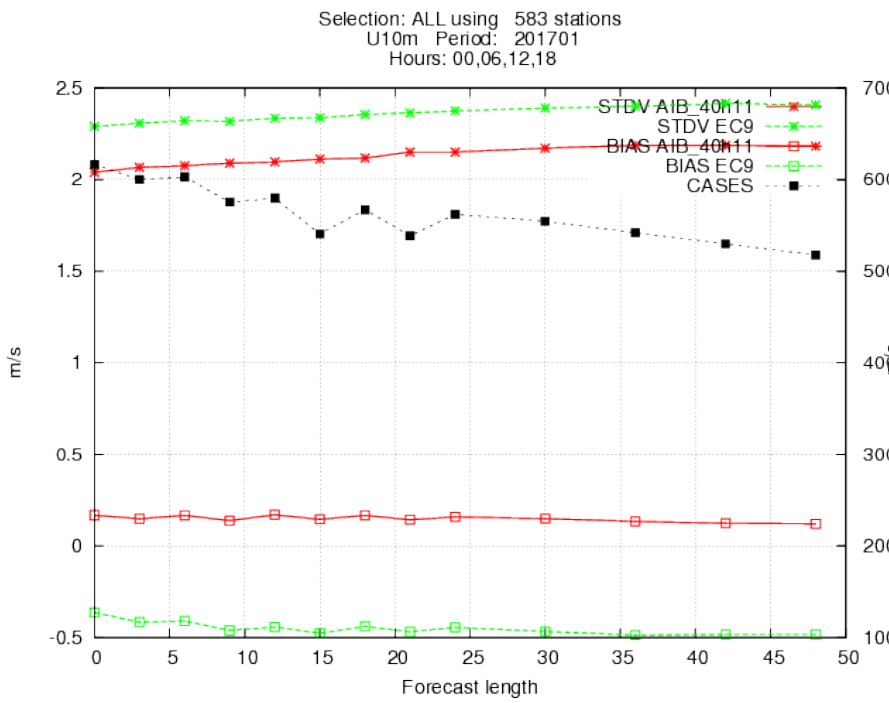
MetCoOp T2M in January-February only slightly improved with respect to STDV for the first 24 hours. However in March 2017 STDV is significantly lower in MetCoOp compared with ECMWF

1) Model performance: V10M

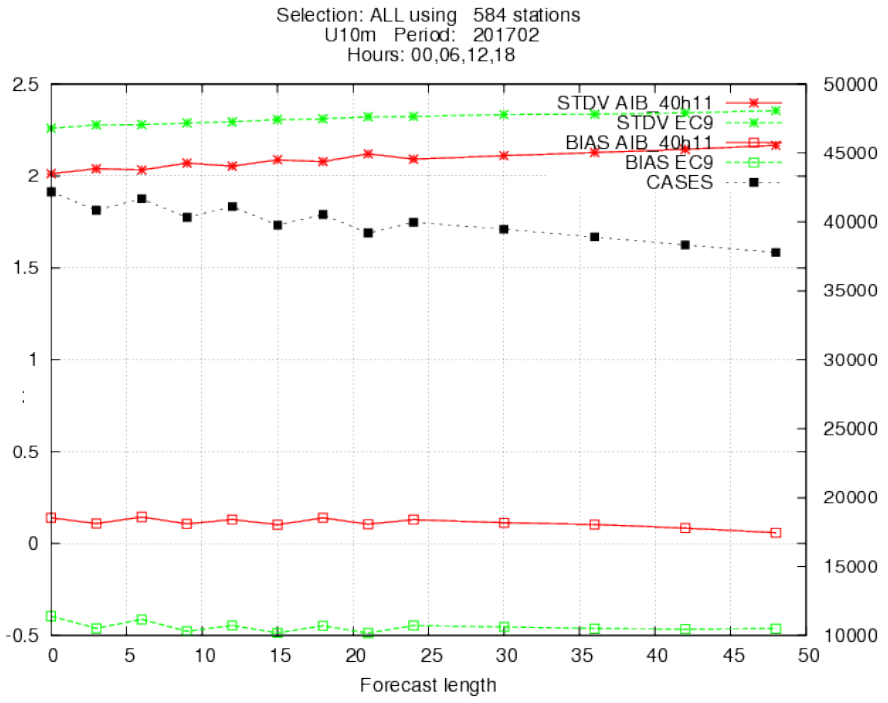
Results of MONITOR, Jan-Feb 2017

(AEMET results)

Comparison (BIAS and STDV) against ECMWF



January 2017



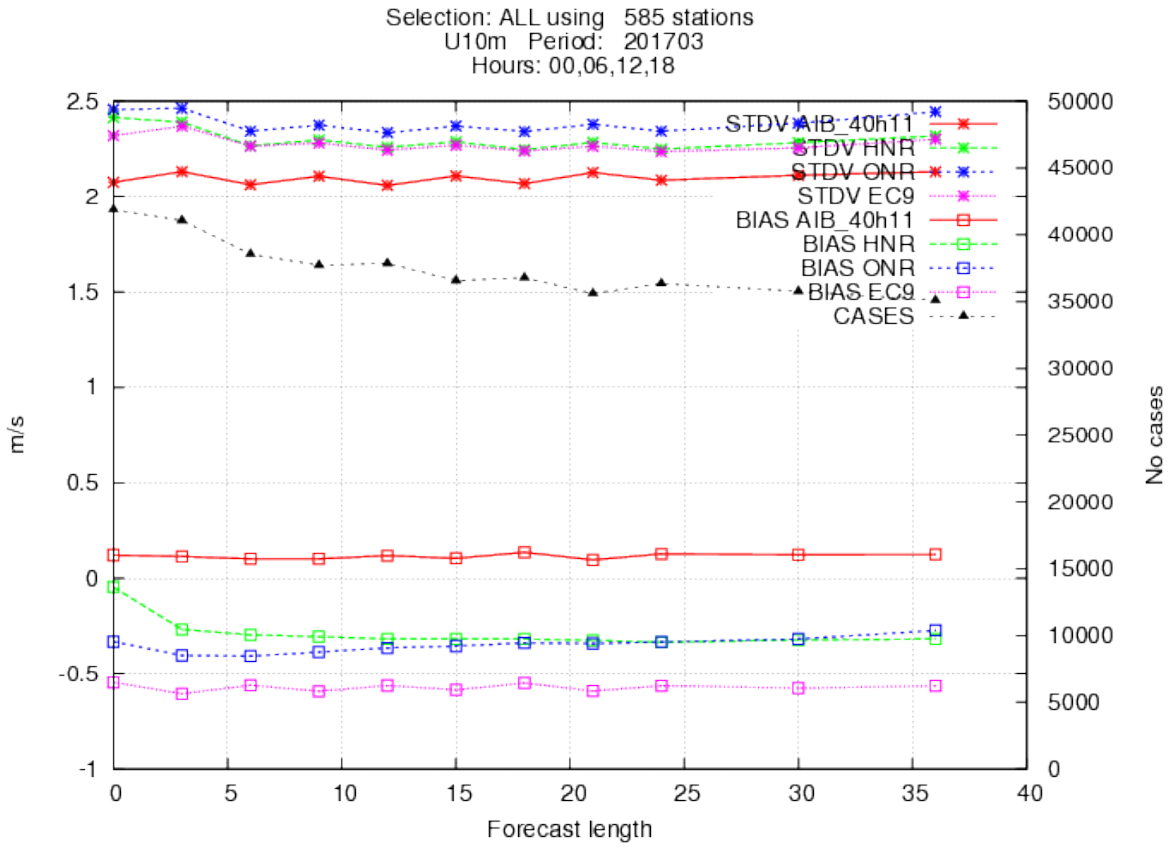
January 2017

1) Model performance: V10m

Results of MONITOR, March 2017

(RCR AEMET results)

Comparison (BIAS and STDV) against ECMWF



Conclusion V10M at AEMET 1st Quarter 2017:

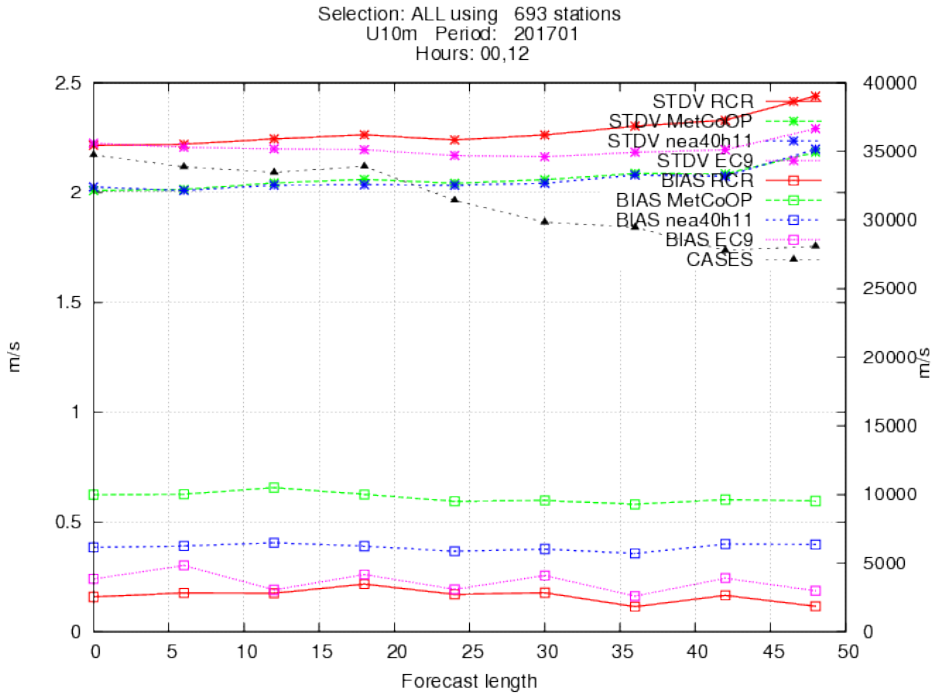
AEMET CY40h1.1 scores substantially better than the other models available

1) Model performance: V10M

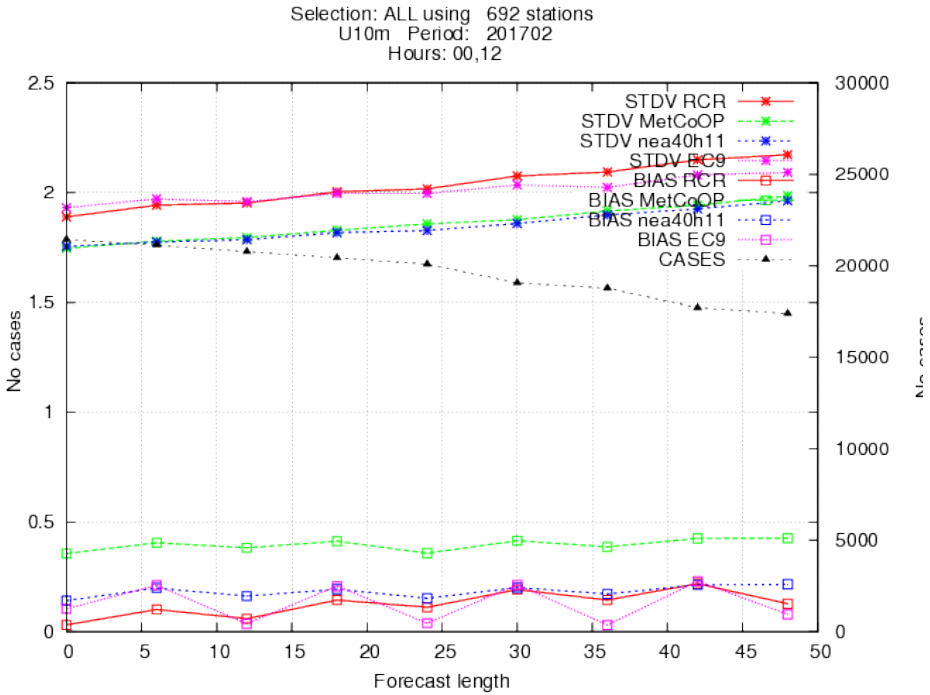
Results of MONITOR, Jan-Feb 2017

(MetCoOp results)

Comparison (BIAS and STDV) against ECMWF



January 2017



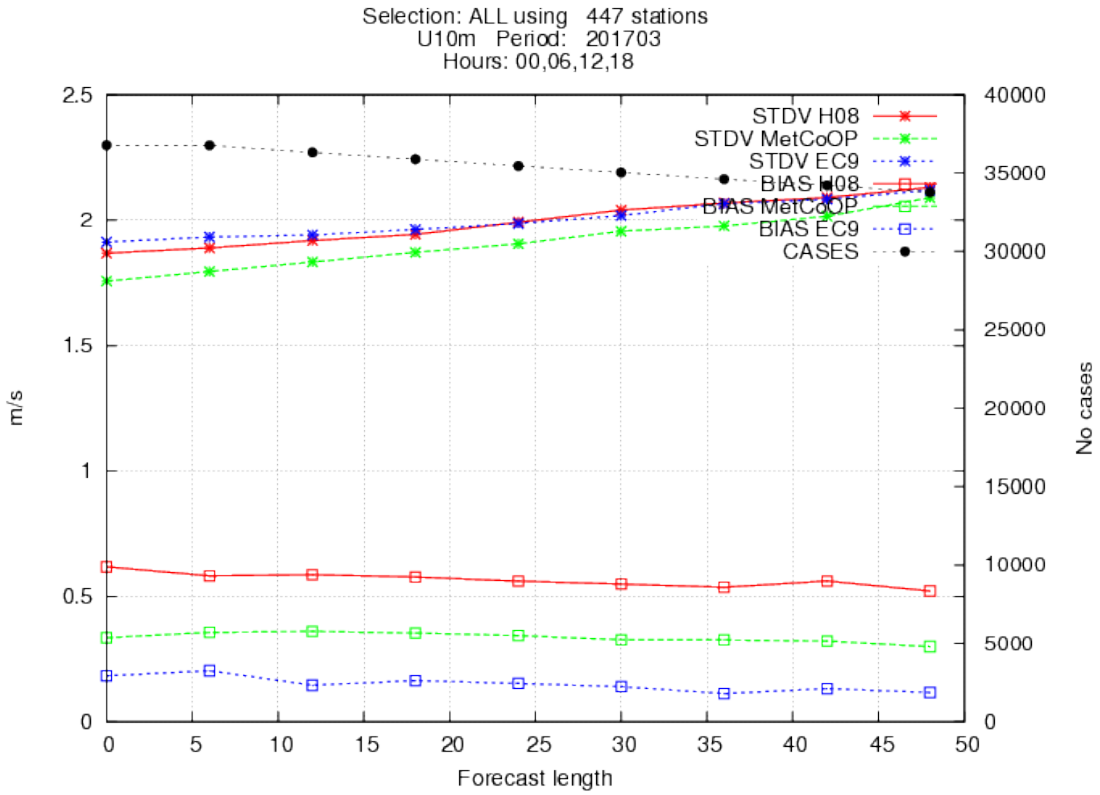
February 2017

1) Model performance: V10m

Results of MONITOR, March 2017

(RCR MetCoOp results)

Comparison (BIAS and STDV) against ECMWF

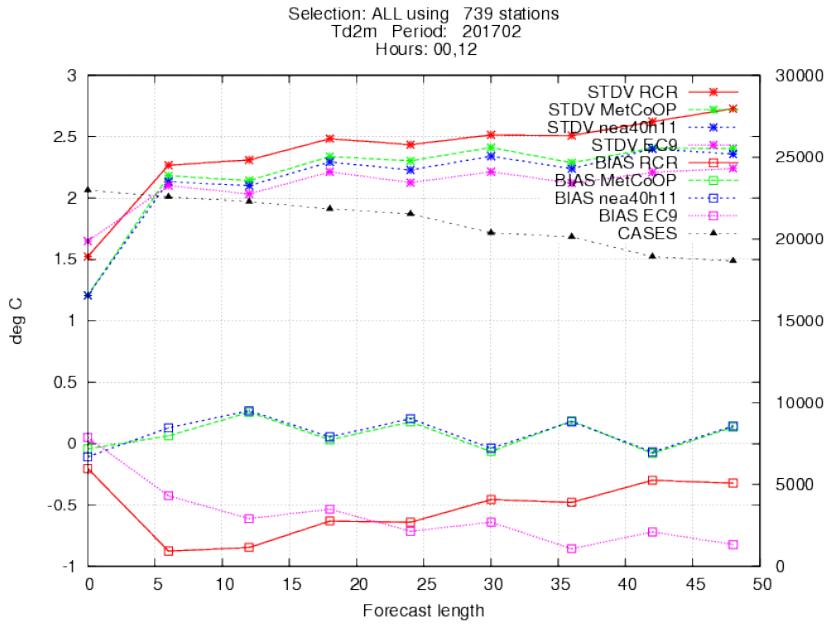


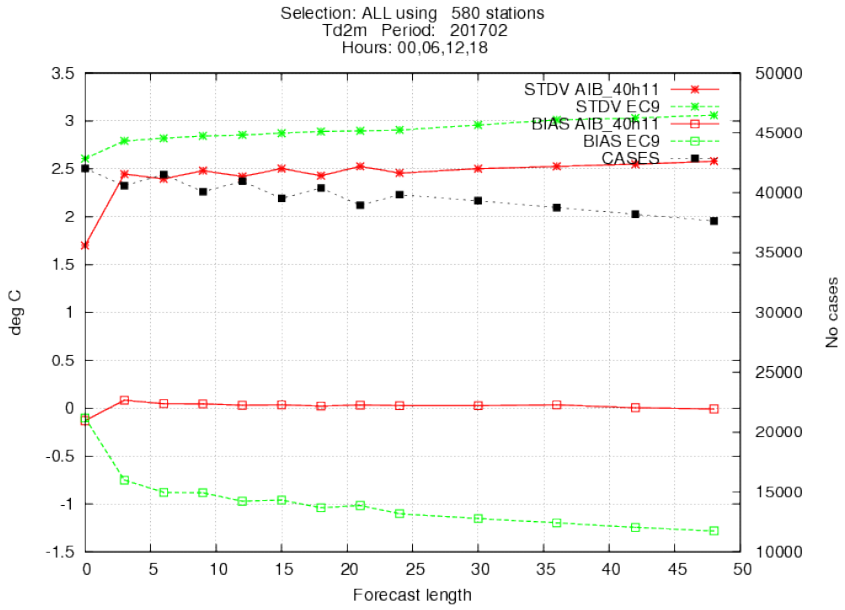
Conclusion V10M at MetCoOp 1st Quarter 2017:

STDV of V10M improved compared with ECMWF, bias more positive relative to ECMWF, perhaps due to excessive winds over snow covered areas (Swedish ”Norrland”)

1) Model performance: TD2M

Results of MONITOR, February 2017 (AEMET verification right, MetCoOp and other models left)

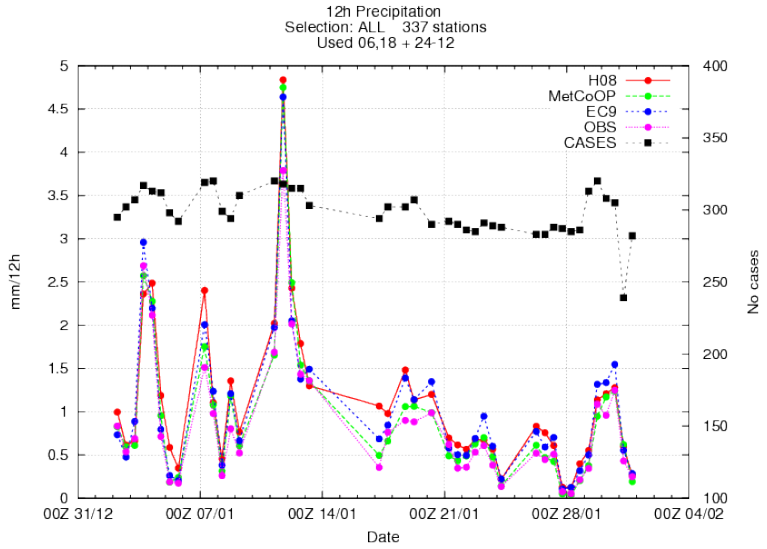


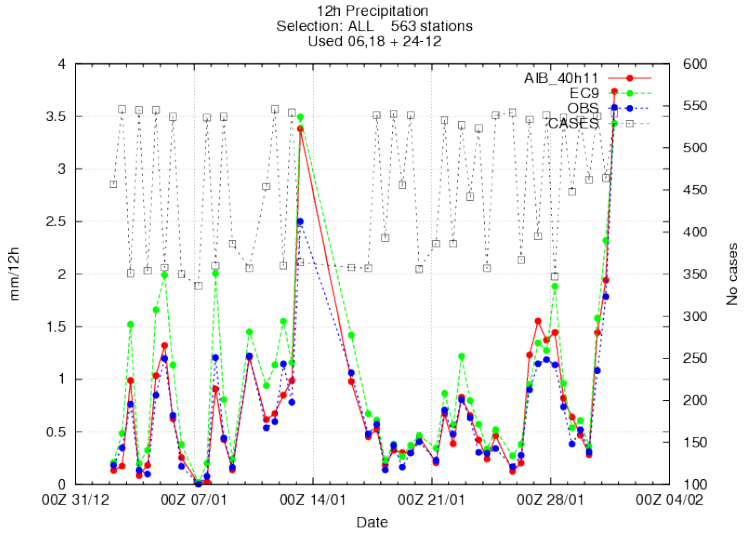


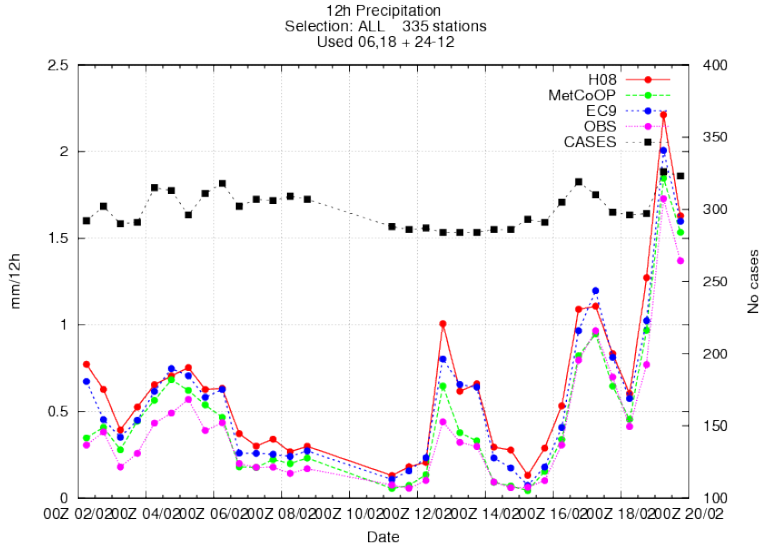
TD2M: Small bias in both AEMET and MetCoOp, STDV improved in AEMET relative to ECMWF, slightly larger in MetCoOp for February .

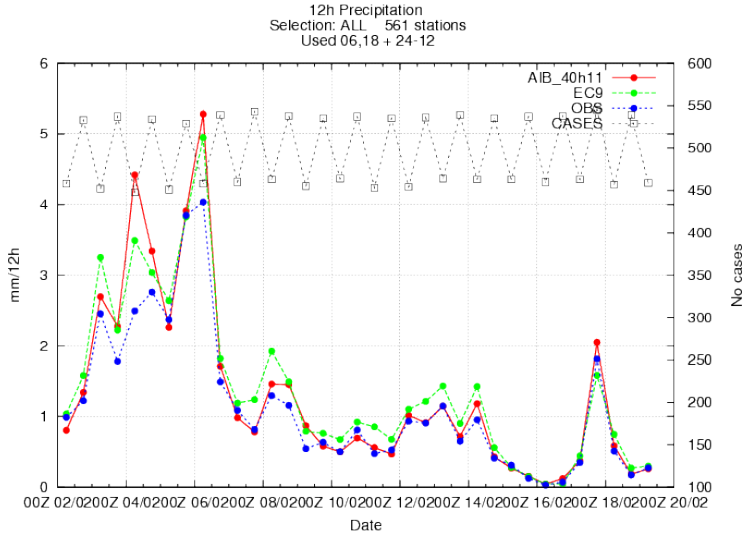
1) Model performance: Precipitation

Results of MONITOR, Jan-February 2017 (AEMET verification right, MetCoOp and other models left)



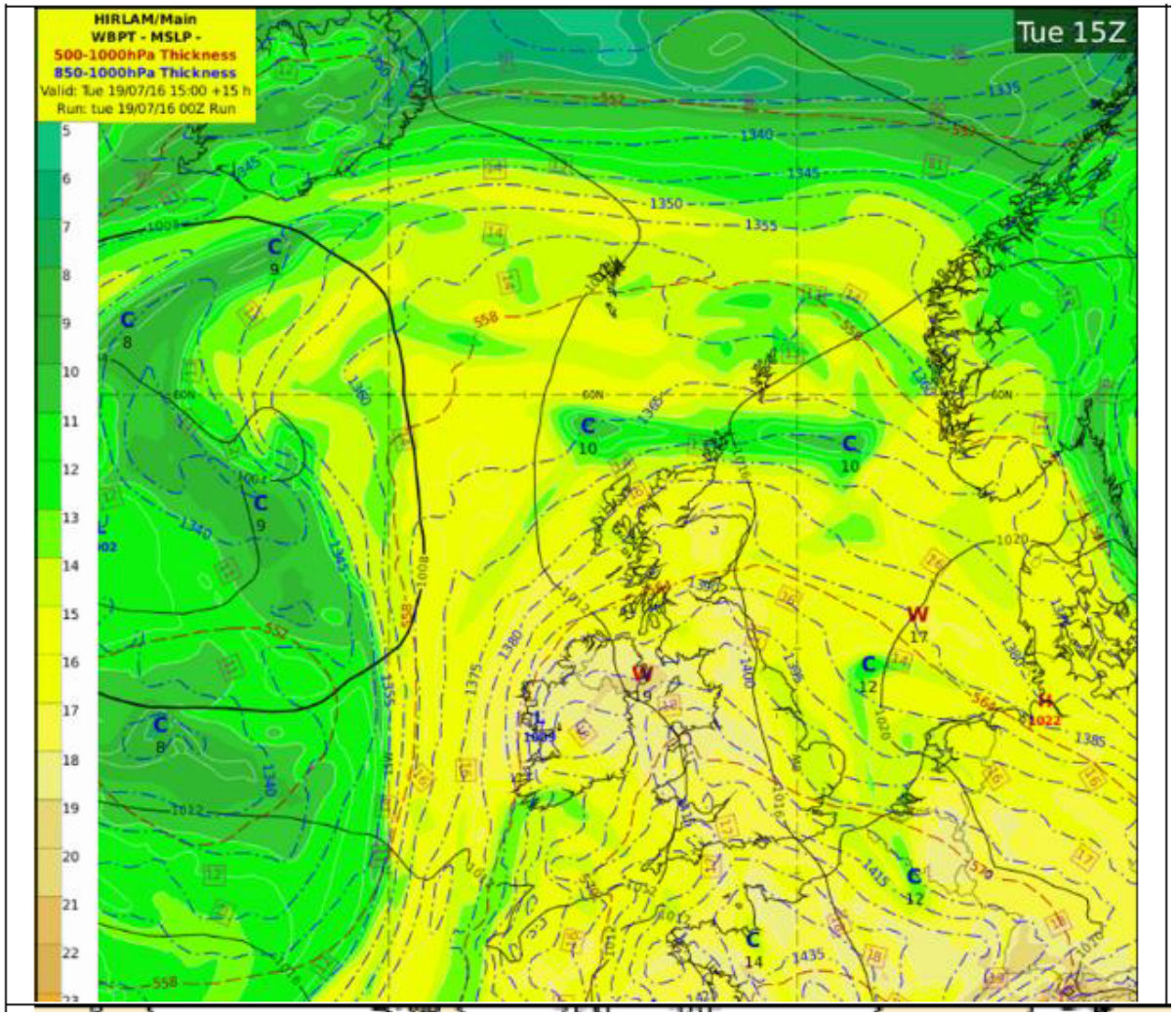






Model performance: FOG /visibility

Overpredicting fog over sea in a warm air mass in Summer 19.July 2016

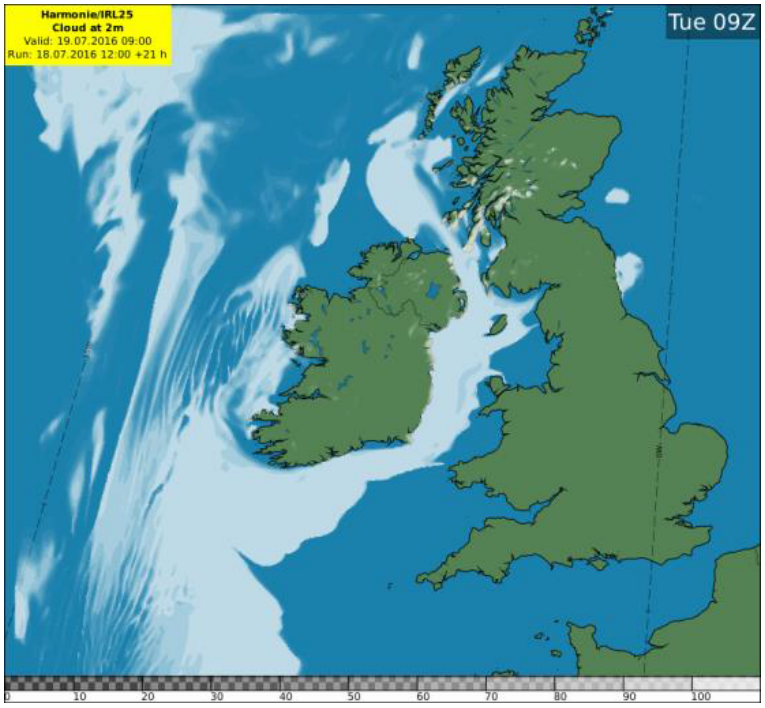


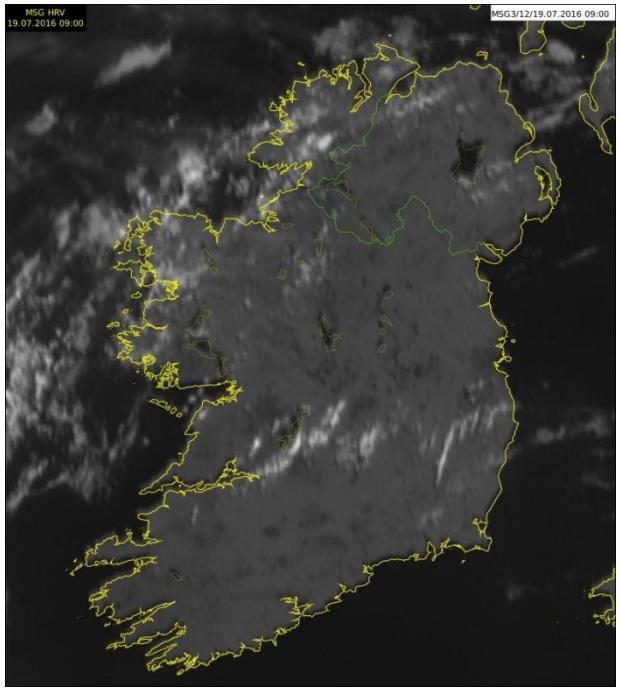
15h forecast valid 15 UTC 19 July 2016 850-1000 hPa thickness

Model performance: FOG /visibility

Overpredicting fog over sea in a warm air mass in Summer 19.July 2016

NB: The excessive fog over sea appears to be eliminated in HARMONIE CY40





Spurious fog forecasted (left) following coast line -

not verified in satellite picture (right)

The fog over sea following the coast line was not observed (compare forecast with satellite picture)

Model deficiencies FOG/visibility

Excellent prediction of visibility over Ireland

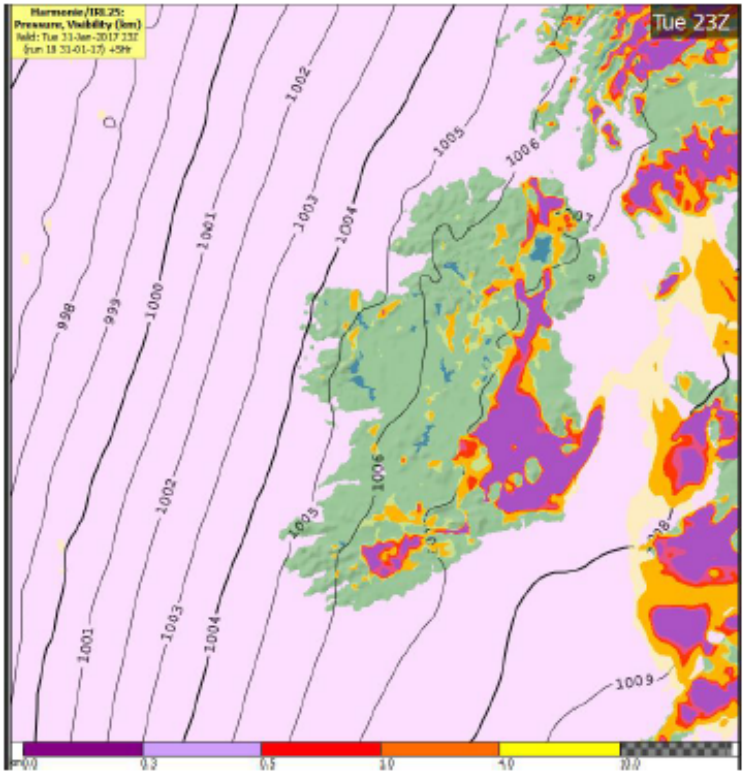
31-01-2017

Night shift 31/01/2017

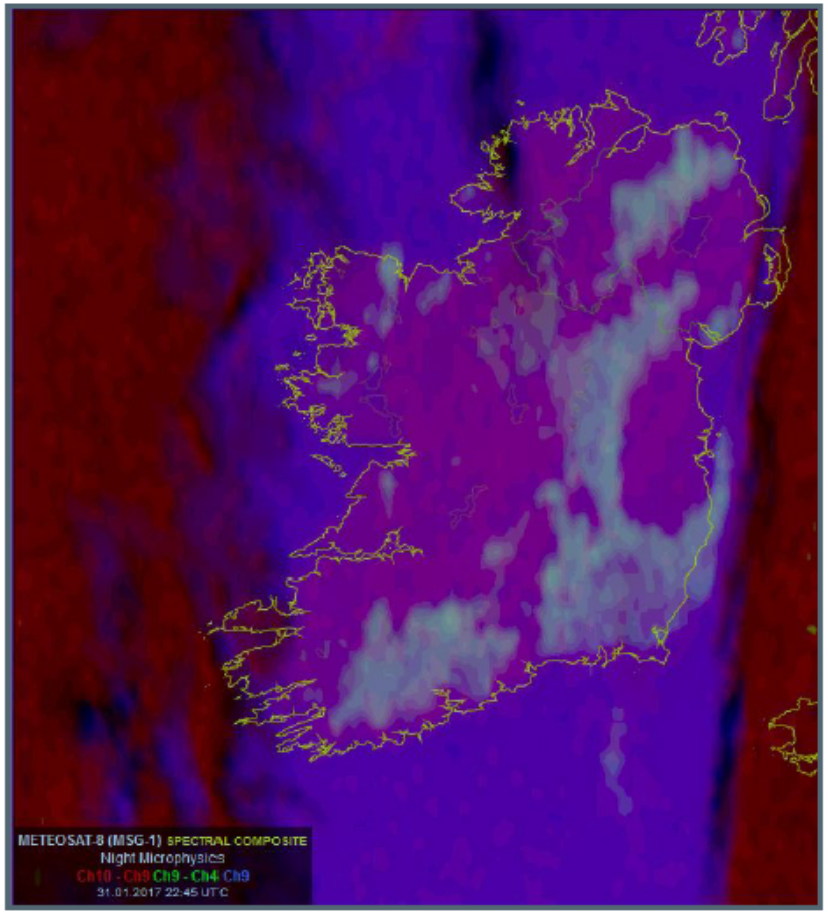
Visibility product for Harmonie performed excellent on this particular occasion, I have not seen it do so well previously, or since.

The night microphysical derived satellite gave me high confidence in the Harmonie product through the night.

At the time I should have taken more screen shots, but below are what I did take.



Hamonie predicated a low visibility area across much of Leinster, with a break around Dublin, due to the southerly flow. For the following couple of hours, it predicted that mist and fog would push back in from the Irish Sea on to the land, which it did.



Observations from Dublin airport show low visibility moving in through the early hours.

Model performance:
b) Interaction with users

- **Communication at web-page https://hirlam.org/trac/wiki/CommunicationWithUsers**
- **Contains quarterly reports with forecasting issues sent by the HIRLAM institutes**
- **Objective verification reports documenting the quality of new model cycles**
- **Example of good or poor forecasts sent by forecasters**
- **Challenges:**

  a) **Different institutes may run different versions of Harmonie/Arome**
  b) **Forecaster reactions is often based on ”single cases” implying the risk that they are not general enough**

CY40h1.1 and CY40h1.2
Why was an update considered necessary ?

The short story is:

First version of CY40 showed many good results especially for wind profiles in the low troposphere.

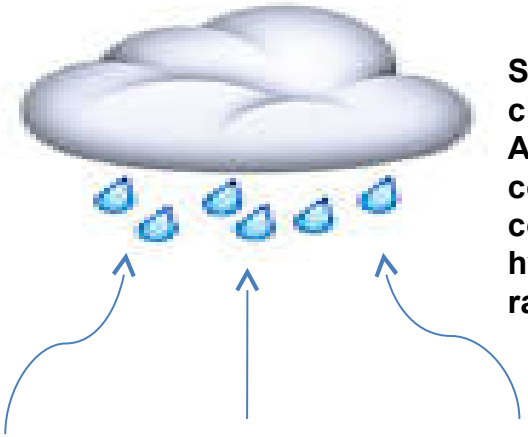
However, the results had a regional dependency. Over Iberian domain the results verified poorer for temperature near the ground, and for clouds, especially for winter conditions.

As a consequence it was decided to make continued experimentation during Autumn 2016 in order to alleviate these deficiencies, aiming at CY40h1.2

Meanwhile more problems were identified by some users, e.g. related with convective precipitation claimed to come to late and die out to quickly. Also supercooled rain is claimed to be not forecasted with sufficiently high frequencey (important in winter conditions with the risk of slippery roads)

HARMONIE grid column

Introduce cloud condensate and precipitation at the lateral boundaries

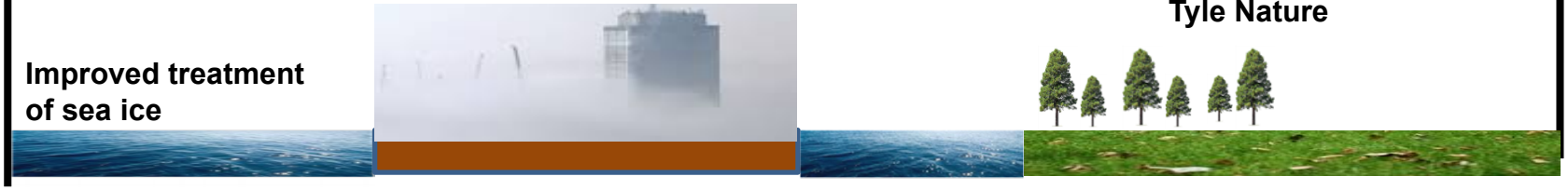


Shallow precipitating clouds
Aim at
correct timing
correct type of hydrometeors e.g. freezing rain

Improved settings for turbulence scheme

Suppress unrealistic precipitation ”bombs”: COMAD dynamics option

Improved treatment of sea ice



Two patches in Tyle Nature

patch2 patch1

Sea Urban Lake Nature

1) Model deficiencies:

c) Ongoing work

Tested components of Harmonie CY 40h1.2

Target 1:

Freezing rain update
Bug fixes stratospheric warming
Accretion changes
COMAD
SICE
PGD interpolation bug
noSBL
Smoothing of orography

Target 2:

Two patches
Flake

TABLE 1a TOPIC: SURFACE initial state	INPUT from verification /research (YES/NO)	INPUT from forecasters (YES/NO)	Expected action(s)
CANARI scheme is old and needs improvements/ replacement ?	YES	—	Implement a more advanced surface analysis scheme (SODA) ?
SSTs and ice in lakes:	YES	YES, unrealistic SSTs /ice in lakes mentioned	Implement FLAKE, check availability of measurements
Quality check of physiographic fields, e.g. variation of Leaf Area index, ECOCLIMAP-2 data (”Disappeared Cities”)	YES	—	Improved use of current satellite data and data bases such as ECOCLIMAP2

TO BE EXPLORED

Figure shows urban fraction (Ecoclimap2 versus Ecoclimap1)

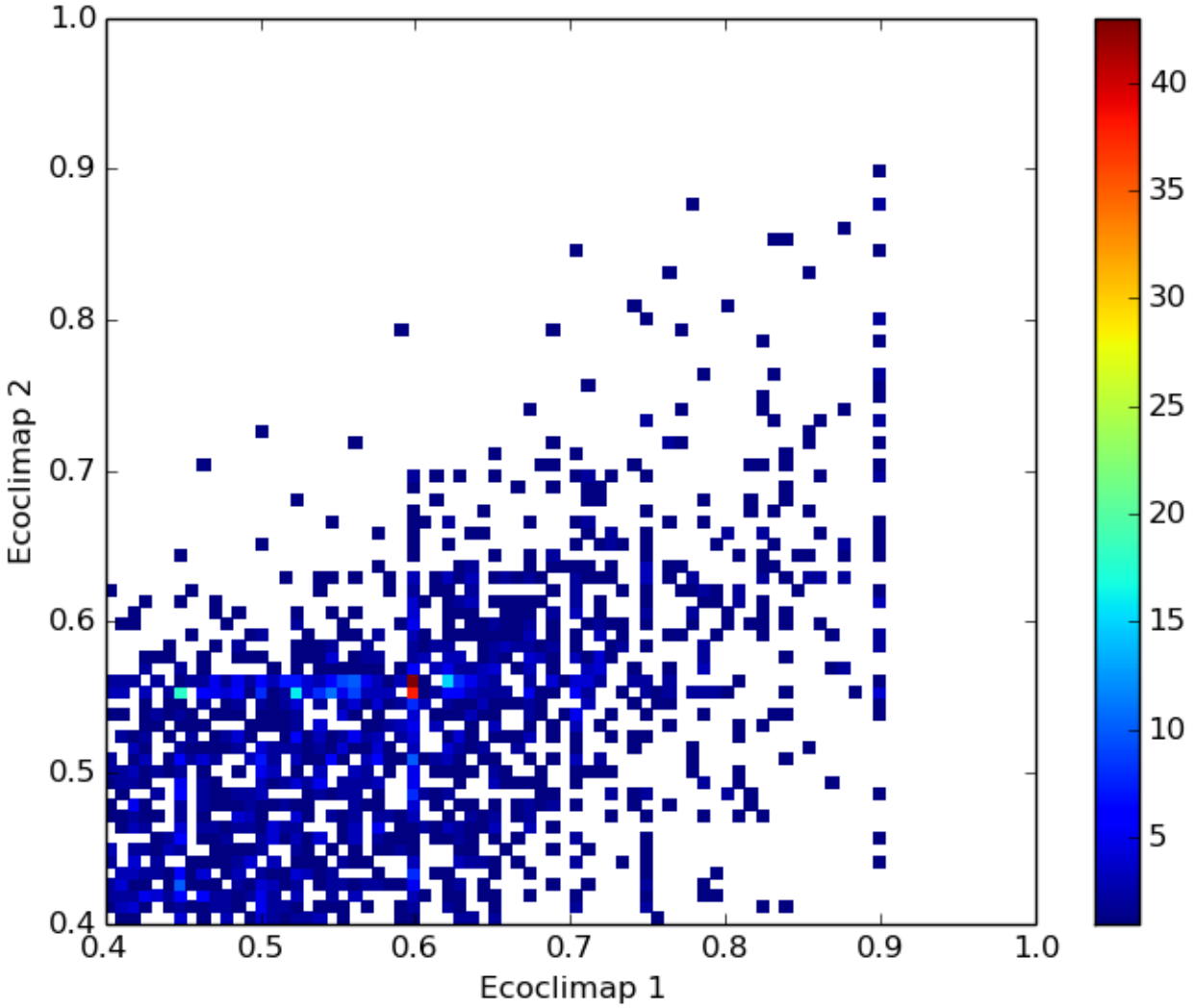


TABLE 1b TOPIC: SURFACE scheme	INPUT from verification/ research (YES/NO)	INPUT from forecasters (YES/NO)	Expected action(s)
Force-restore surface scheme has limitations	YES	NO	Implement a more advanced surface prediction scheme (multilayer scheme in SURFEX)
Improved temp on glaciers needed	YES	YES	Install update for snow treatment (`Glacier update´)
Snow melt: Does snow melt too slowly in (rapidly) changes weather	YES, at DMI	(YES)	Investigate melting process in current and future surface scheme

TO BE EXPLORED

TABLE 1c TOPIC: SURFACE diagnostics	INPUT from verification/research (YES/NO)	INPUT from forecasters (YES/NO)	Expected action(s)
Remove T2m negative temp. bias, in in winter	YES, a typical feature in winter conditions	YES, but very lowest minima may have opposite problem (too warm)	Improve physiography data, surface scheme and surface analysis + fluxes from atmosphere (e.g. radiation from clouds)
Wind over sea ice too weak while wind over snow covered land is too strong (seen in Northern Scandinavia)	YES	YES, complaints from MetCoOp forecasters	Modify roughness length of ice/snow
Visibility often too low in situations of fog/mist	YES	YES	Test new published formulas and use possibly LIMA scheme

TO BE EXPLORED

TABLE 2a TOPIC: Atmospheric noise/spinup	INPUT from verification/ research (YES/NO)	INPUT from forecasters (YES/NO)	Expected action(s)
Suppress initial noise in prognostic fields	YES	YES	Better structure functions (more fine scale) more consistent with model , use of 4DVAR or nudging
Less model spinup desirable: (Initial clouds / precipitation rate sometimes not realistic	YES	YES	Use new data sources, e.g. Satellite data in combination with new assimilation methods
Occasional noise issues seen in low troposphere	YES	YES	Document cases and work out solutions to alleviate problems

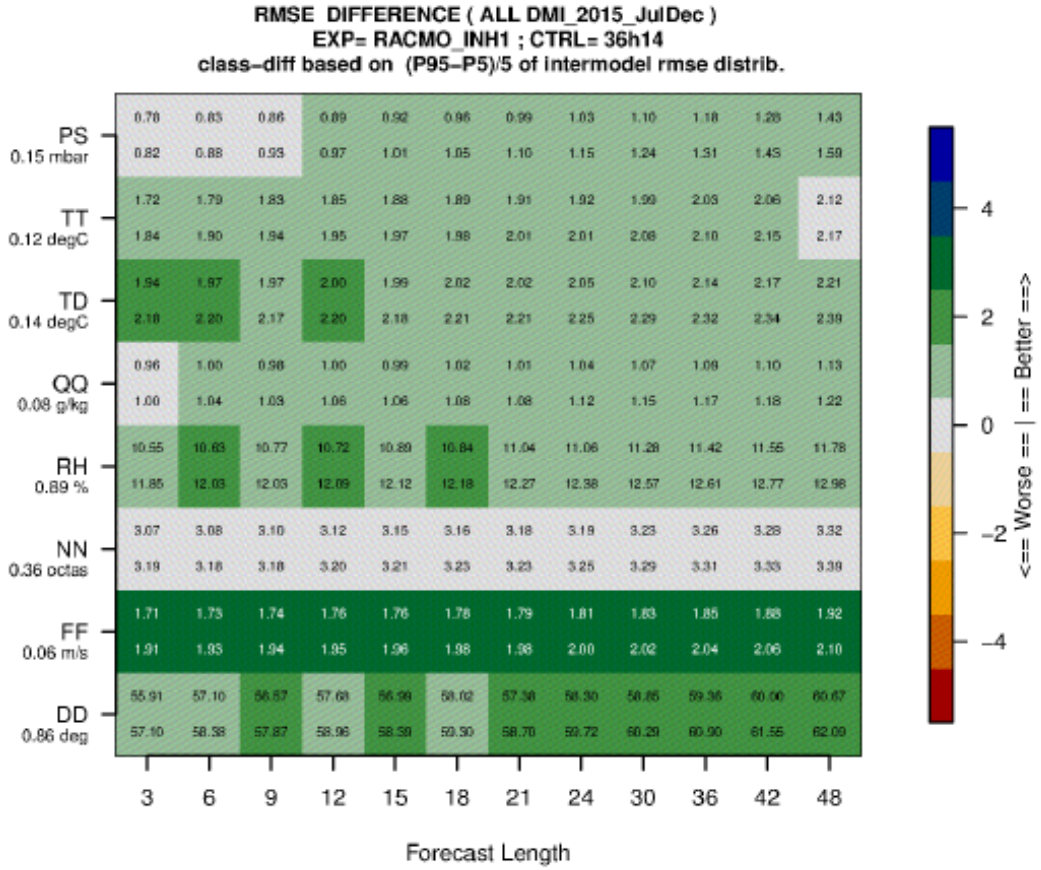
TO BE EXPLORED

TABLE 2b TOPIC: Clouds and Fog	INPUT from verification/ research (YES/NO)	INPUT from forecasters (YES/NO)	Expected action(s)
Improved fractional cloud cover and fog over land requested !	YES	YES	Improved fractional cloud cover (humidity variance term)
FOG sometimes too persistent over land, -not sea - CY38 suffered from too much fog over both sea and land.	(YES)	YES	Modelling of realistic deposition of cloud droplets (work identified in Cloud workshop)
Periods with too few clouds (CY40) Opposite in CY38, had far too many low clouds.	Variable results	Remark by forecasters in several institutes	Revise Cloud parameterization (e.g. variance term)

TO BE EXPLORED

TABLE 2c TOPIC: Convection, precipitation type	Identified by verification/research (YES/NO)	Identified by forecasters/ users (YES/NO)	Expected action(s)
Triggering of convection to be improved ?	----	(YES) some forecasters	Better modelling of variance terms for temperature and humidity
Horizontal structure of convection: Are the precipitation fields in CY40 too smooth? (no consensus on this)	----	YES/NO (significant disagreement)	Carry out spatial verification , e.g. SAL
Is the timing and duration of (shallow) convection correct ?	no timing problem, duration to be investigated	too late claimed by some users	Timeseries of convective precipitation forecasted and observed
Supercooled rain not forecasted frequently enough (CY38,40h1.1)	YES	YES	CY40h1.2 is (successfully) addressing this issue

2) Score cards : a new trend in HIRLAM-C



Score card developed in KNMI

Figure 3.1: RMSE Scorecard for "ALL", comparing Harmonie versions RACMO_INH1 to 36h14. Vertical axis denotes the variables, horizontal axis the forecast step. All forecasts are used. Numbers inside the boxes denote the RMSE of the particular variable for RACMO_INH1 (top) and 36h14 (bottom). Green colours denote improvement (reduced RMSE). See main text for more details. Variable names are given in section 3.1.

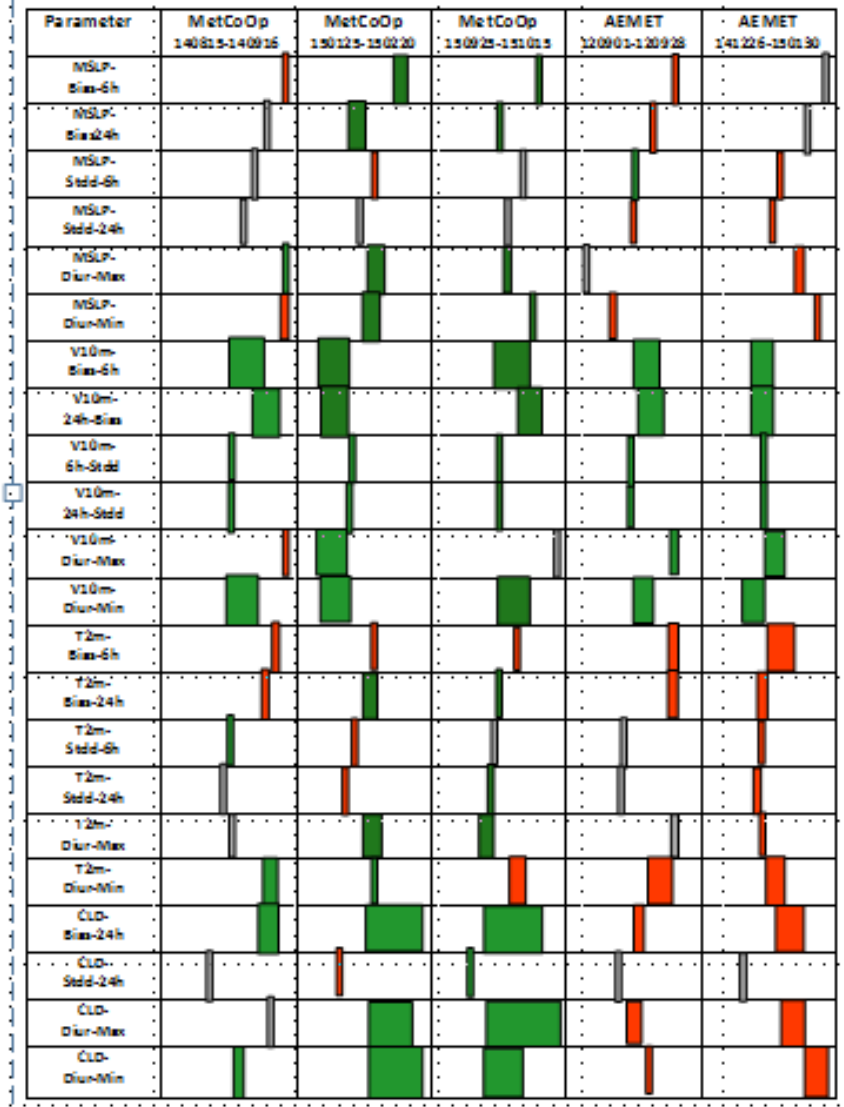
2) Score cards : a new trend in HIRLAM-C

Extract of Score card from an evaluation report of HARMONIE CY40h1.1

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

All scores are transformed (*) to values between 0 (useless forecast) and 1 (perfect forecast)

High end of interval shows best model and low end the poorest (CY40 compared with CY38)



The absolute values of score between two model versions are visible from the graphical display plus colors showing if the new model CY40 is an improvement (green) , is neutral (grey) or is a degradation (orange)

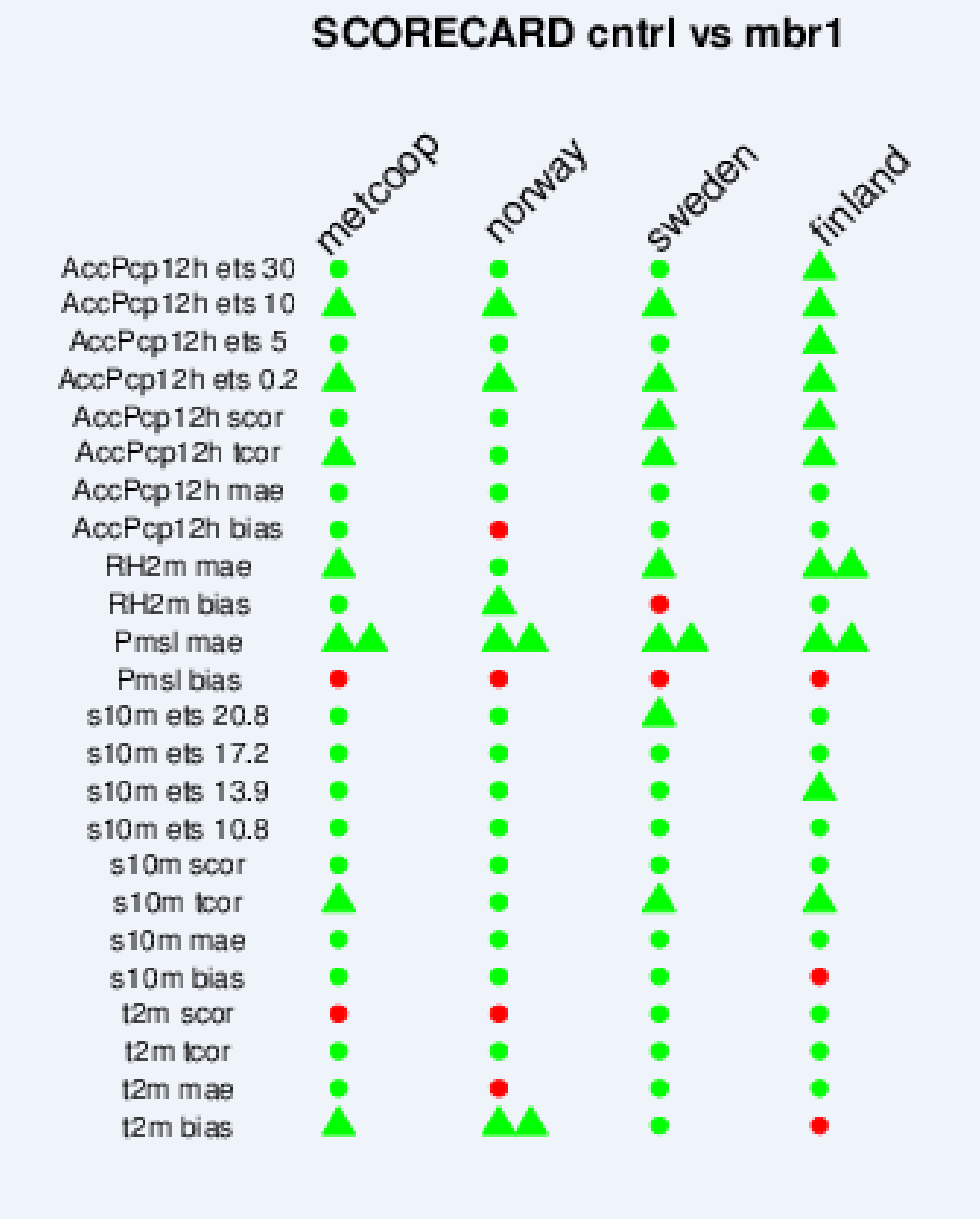
*Quality assessment of Harmonie Cycle 40h1.1 for use in operational Hirlam systems

HIRLAM–C Management June 2016

2) Score cards : a new trend in HIRLAM-C

Scorecard developed in Met.Norway

Statistical significance tests (Diebold - Mariano test)



Work oriented towards HARP, e.g. reading of SQLITE tables

Priority to compare Harmonie-Arome against ECMWF

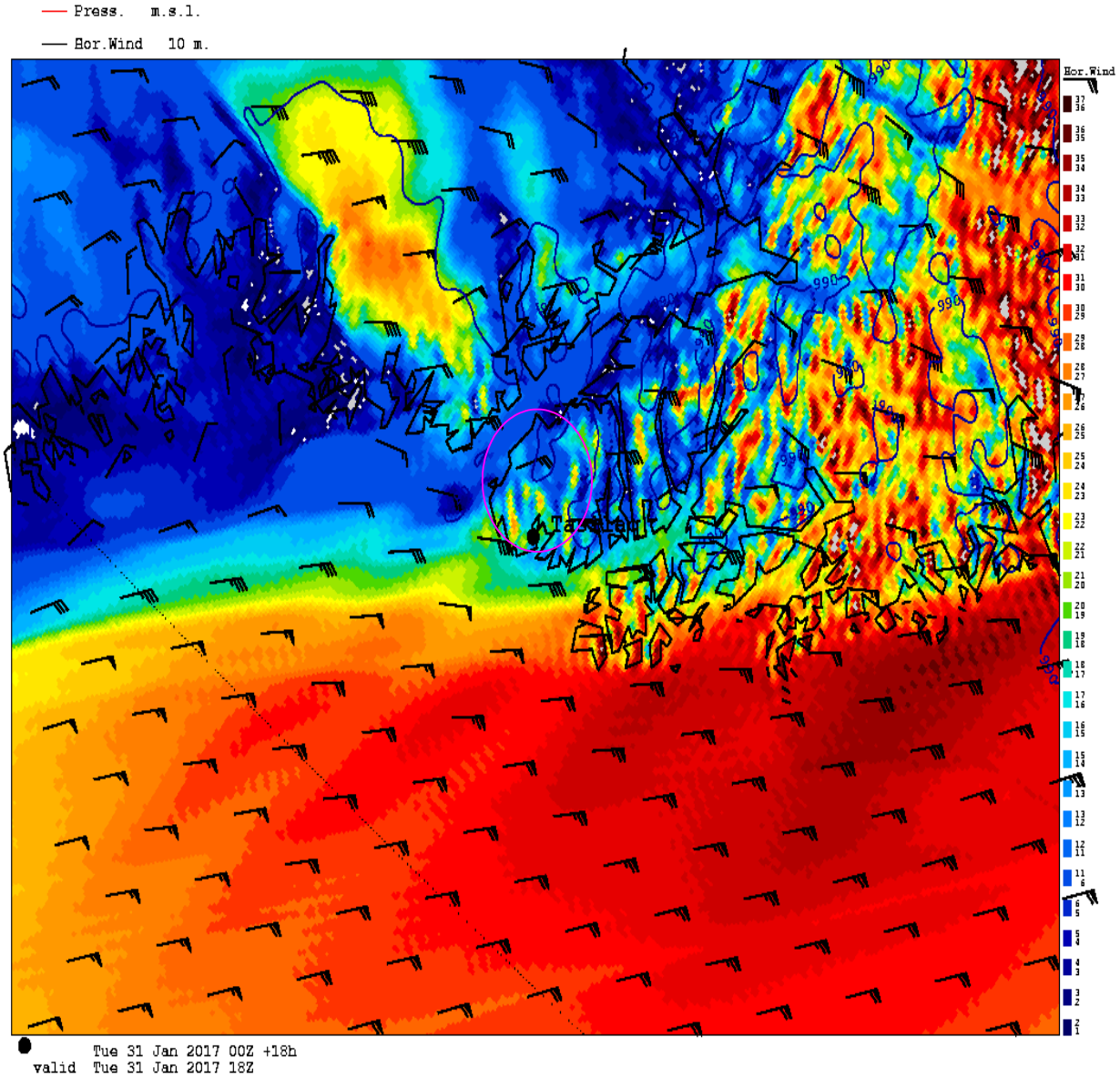
3) Common verification tools (HARP)

HARP version 2.0 is needed soon. A deadline of ~15 June 2017 (before summer holidays) was agreed on for a beta-release.

HARP Version 2.0 should be available to all from a new HARP directory in ec-gate, readable to all in the HIRLAM-ALADIN consortium. In addition, an updated documentation should be made by 15.June and obtainable from the same directory. In addition, the new HARP release should also be available from *hirlam.org.*

Both **Spatial part** (workable SAL + FSS) and **EPS part** of HARP will be updated (Christoph Zingerle's presentation)

4) Special challenges related with orography: Predicting weather parameters in Greenland using sub-km grid



Challenge:

How to communicate to public with forecast of weather that has large local variation.

Wind forecast from a 750 m HARMONIE model

Radius of ~5 km

5) CONCLUSIONS

- **Mostly quite satisfactory verification results have been obtained with CY40h1.1**
- **Further improvement in the pipeline (CY40h1.2)**
- **Announced plans and issues mentioned in last year ASM (e.g. communication with users, score cards) have been followed up.**
- **A list of model challenges is being maintained and followed up**
- **Very high spatial resolution appears to be needed in mountain areas (e.g. Greenland) to fulfil user needs. - Initial results of sub-km Harmonie-Arome appear to be promising**

6) Relevant workshops and reports

a) HIRLAM-C Training week on Harmonie verification and validation tools; DMI, 12-15 December 2016

b) Cloud workshop at Meteo-France . 16-18th January 2017: within the ALADIN-HIRLAMconsortium.Main topic of the workshop : The problem of the low cloud forecast and fog. http://www.meteo.fr/cic/meetings/2017/CWW/agenda.pdf

c) HARP Working Meeting at RMI 8-10 March 2017
Summary available from
Alex Deckmyn, Christoph Zingerle, Anrew Singleton, Bent Hansen Sass

1) Quality assessment of Harmonie Cycle 40 for use in operational Hirlam systems, 24 p , HIRLAM–C Management, First version June 2016, in final form 31 October 2016 (available from www.hirlam.org/trac/wiki/CommunicationWithUsers)

2) Wim de Rooy and Hylke de Vries et al: **Harmonie verication and evaluation** , March 2017: Royal Netherlands Meteorological Institute KNMI, De Bilt, Netherlands (Contact: wim.de.rooy@knmi.nl , hylke.de.vries@knmi.nl), with contributions from Christiaan van Dalum, Siebren de Haan, Geert Lenderink, Gert-Jan van Marseille, Jan Fokke Meirink, Rinus Scheele

3) Diebold, F.X. and R.S. Mariano. (1995). **Comparing Predictive Accuracy**. Journal of Business and Economic Statistics, 13: 253-63.