some highlights from

A NWP model inter-comparison of surface weather parameters during the Year of Polar Prediction Special Observing Period Northern Hemisphere 1

APPLICATE General Assembly 2019 - related to Task 5.1 (stream 1 exp), Task 5.2 (state-of-the-art) + clustering

Morten Køltzow (Norwegian Meteorological Institute),
Barbara Casati (Environment and Climate Change Canada), Eric Bazile (Meteo France),
Thomas Haiden (ECMWF) and Teresa Valkonen (Norwegian Meteorological Institute)

YOPP YEAR OF POLAR PREDICTION

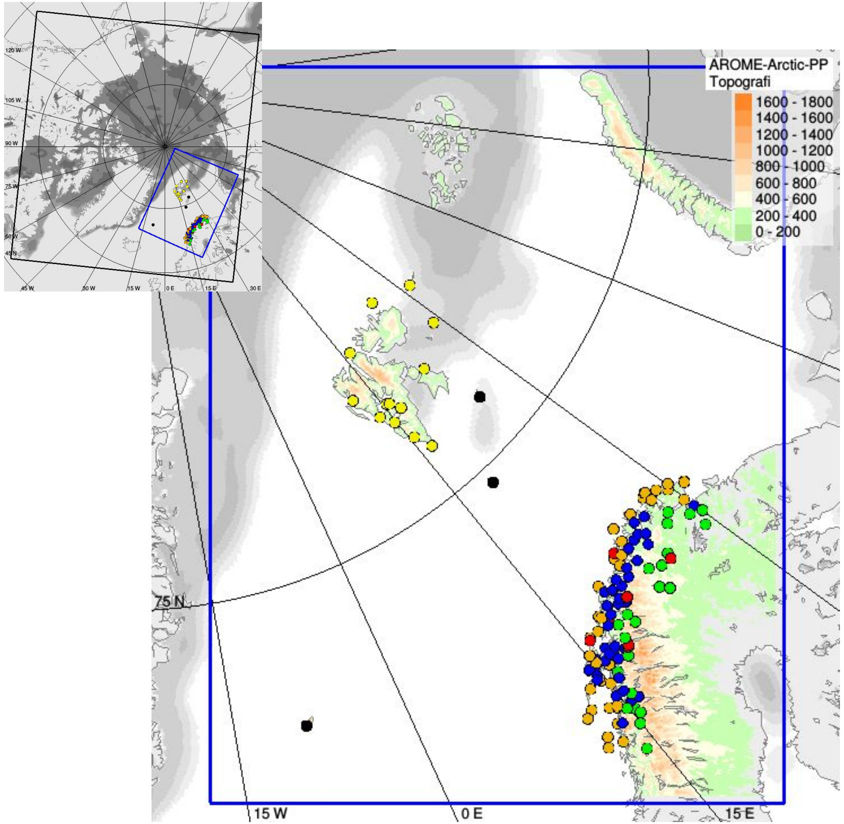
APPLICATE.eu Advanced prediction in polar regions and beyond

grant agreement number 727862

Alertness

Photo:
Norsk Luftambulanse

Numerical Weather Prediction (NWP) systems & observations



Period:

YOPP-SOP-NH1 (1.February - 31. March 2018)

4 NWP systems, short range forecasts (1-2 days ahead)

- **IFS HRES (ECMWF), Global**
  - ~9km, global system, data assim, operational
- **AROME-Arctic (MET Norway), Limited area model**
  - 2.5km, data assim, operational, LBC (IFS HRES)
- **CAPS (ECCC), Limited area model**
  - 3 km, downscaling (GDPS), YOPP-dedicated ("real-time")
- **MF AROME (Meteo France), Limited area model**
  - 2.5km, downscaling (ARPEGE), YOPP-dedicated

Norwegian quality controlled synop observations

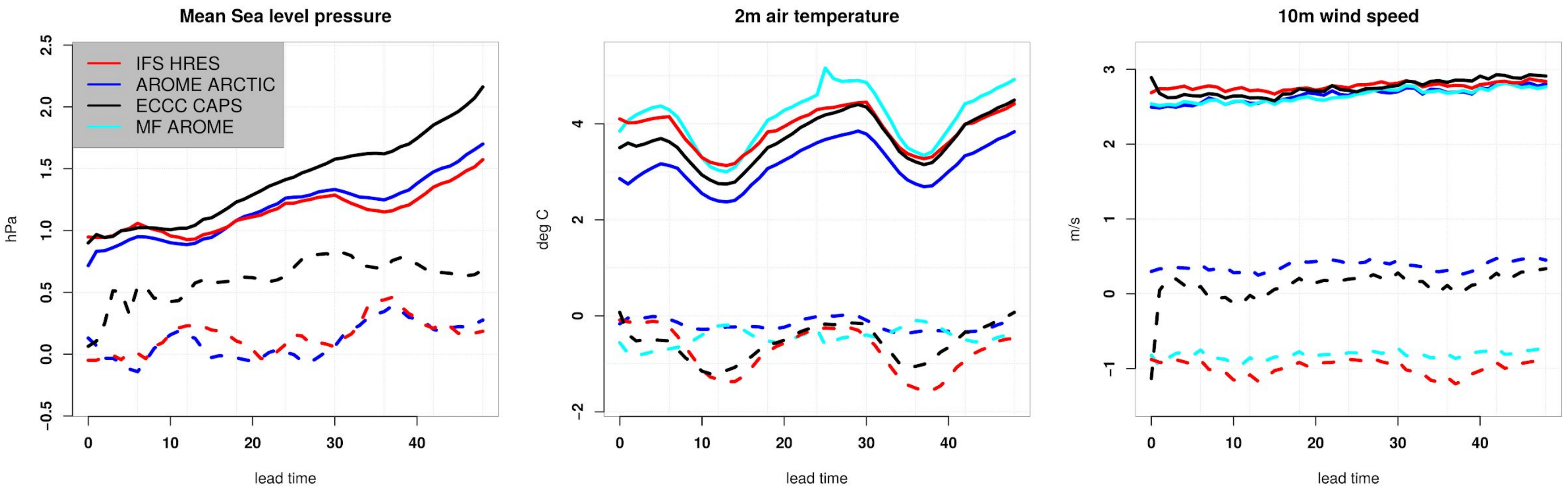
- eklima.met.no: MSLP, T2, WS10, precip24, precip1, TCC
- Split in regions; islands (3), coast (40), fjords (39), inland (25), mountains (9), Svalbard (14, yellow)

Advanced Scatterometer (ASCAT) coastal wind product

- 12.5 km grid (NWP systems and ASCAT regridded to common grid)
- EUMETSAT, Verhoef et al 2012

Forecast errors,all stations, function of lead time

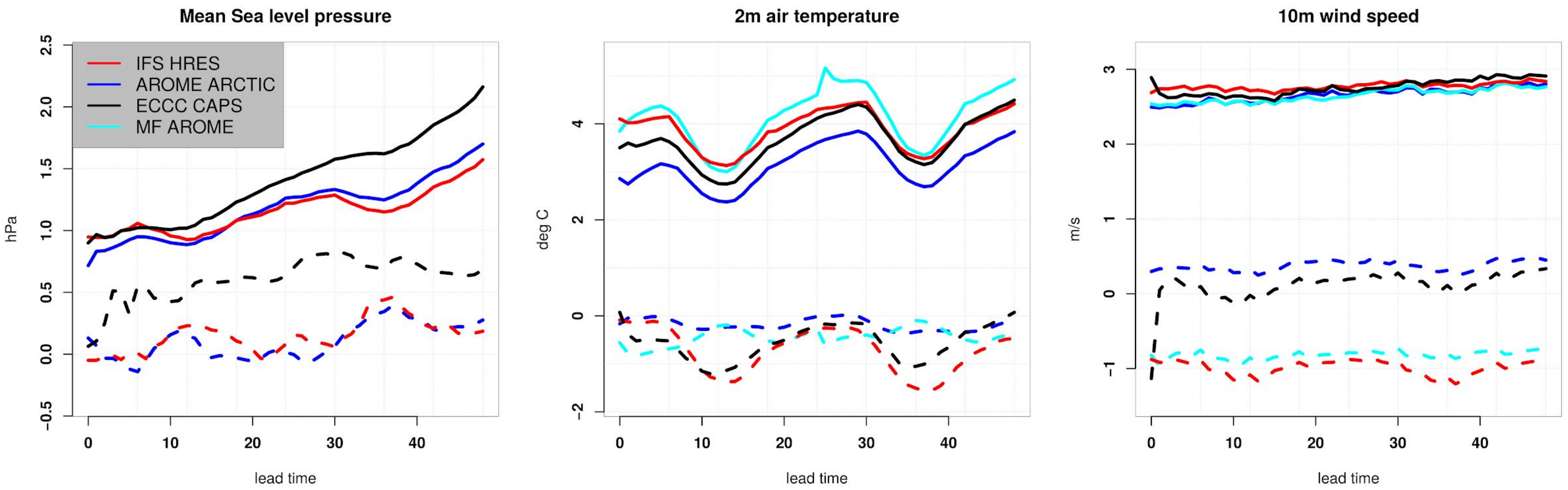
standard deviation of error (solid lines), bias (dashed lines)



Standard deviation of error (solid lines) and bias (dashed lines) as function of lead time. *Models are IFS HRES (red), AROME Arctic (blue), CAPS (black) and MF AROME (cyan, MSLP not available from MF AROME) and parameters Mean Sea Level Pressure (MSLP), 2m air temperature (T2) and 10m wind speed (WS10). Verification period is YOPP SOP-NH1 and all forecasts are initialized at 00 UTC.*

Forecast errors,all stations, function of lead time

standard deviation of error (solid lines), bias (dashed lines)

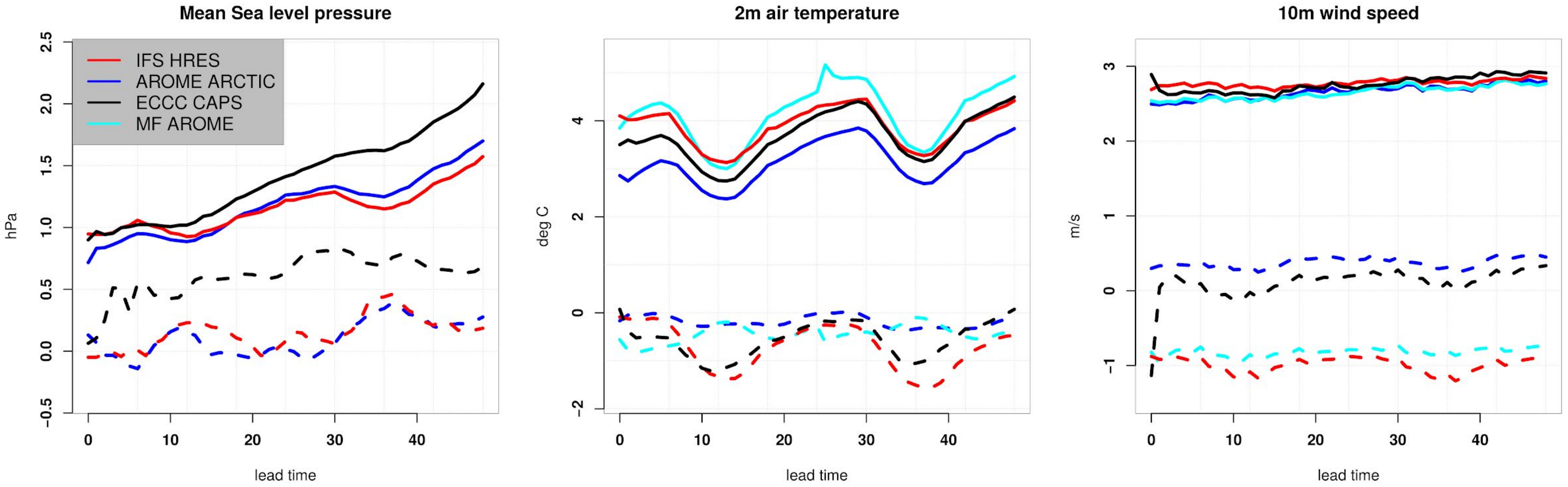


Standard deviation of error (solid lines) and bias (dashed lines) as function of lead time. *Models are IFS HRES (red), AROME Arctic (blue), CAPS (black) and MF AROME (cyan, MSLP not available from MF AROME) and parameters Mean Sea Level Pressure (MSLP), 2m air temperature (T2) and 10m wind speed (WS10). Verification period is YOPP SOP-NH1 and all forecasts are initialized at 00 UTC.*

Small initial errors MSLP, but rapid growth. Large initial errors T2 and wind speed, but slow growth

Forecast errors,all stations, function of lead time

standard deviation of error (solid lines), bias (dashed lines)



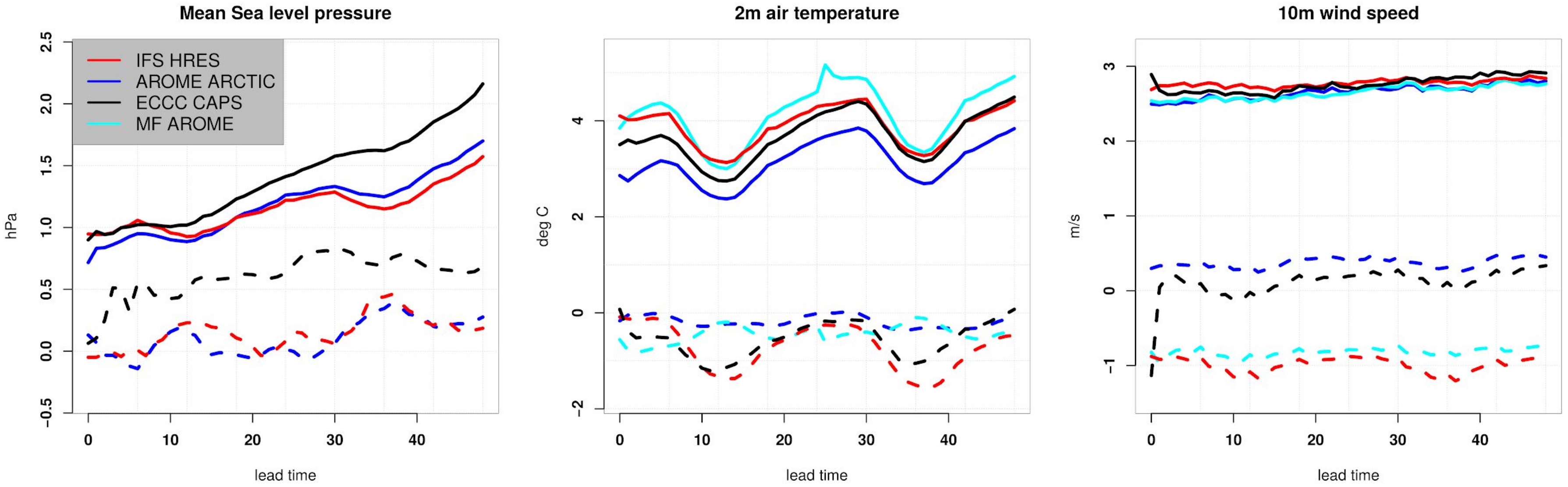
Standard deviation of error (solid lines) and bias (dashed lines) as function of lead time. *Models are IFS HRES (red), AROME Arctic (blue), CAPS (black) and MF AROME (cyan, MSLP not available from MF AROME) and parameters Mean Sea Level Pressure (MSLP), 2m air temperature (T2) and 10m wind speed (WS10). Verification period is YOPP SOP-NH1 and all forecasts are initialized at 00 UTC.*

Small initial errors MSLP, but rapid growth. Large initial errors T2 and wind speed, but slow growth

Added value of high resolution models indicated for T2 and WS10, not for MSLP.

Forecast errors,all stations, function of lead time

standard deviation of error (solid lines), bias (dashed lines)



Standard deviation of error (solid lines) and bias (dashed lines) as function of lead time. *Models are IFS HRES (red), AROME Arctic (blue), CAPS (black) and MF AROME (cyan, MSLP not available from MF AROME) and parameters Mean Sea Level Pressure (MSLP), 2m air temperature (T2) and 10m wind speed (WS10). Verification period is YOPP SOP-NH1 and all forecasts are initialized at 00 UTC.*

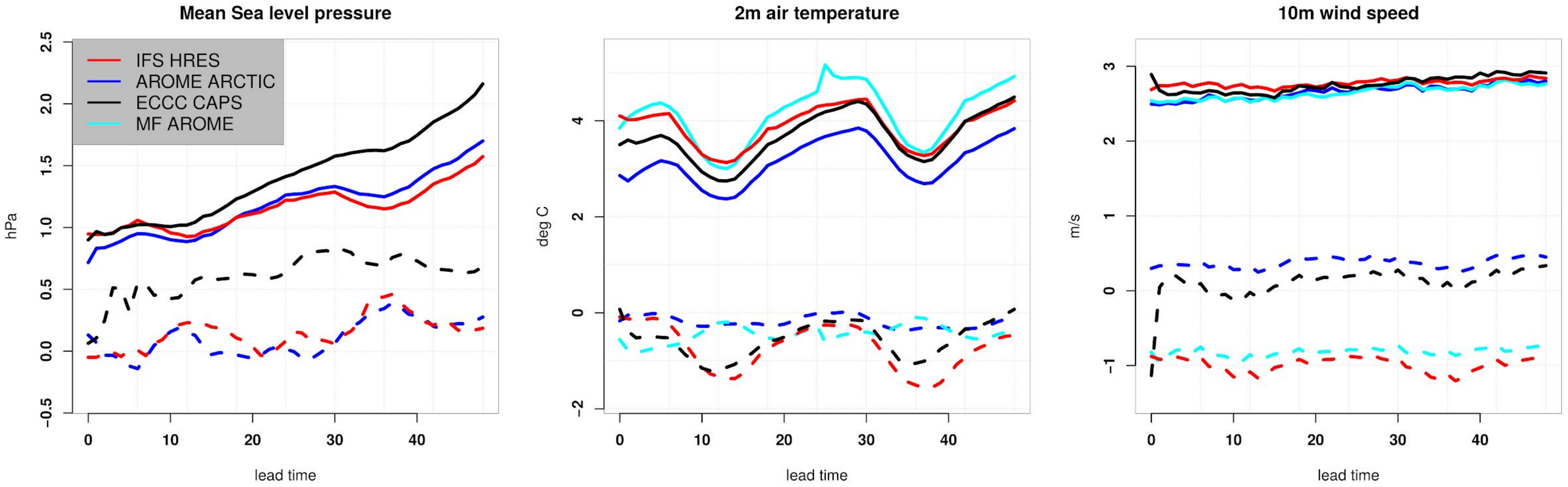
Small initial errors MSLP, but rapid growth. Large initial errors T2 and wind speed, but slow growth

Added value of high resolution models indicated for T2 and WS10, not for MSLP.

Faster error growth in high res. models than in global IFS-HRES.

Forecast errors,all stations, function of lead time

standard deviation of error (solid lines), bias (dashed lines)



Standard deviation of error (solid lines) and bias (dashed lines) as function of lead time. *Models are IFS HRES (red), AROME Arctic (blue), CAPS (black) and MF AROME (cyan, MSLP not available from MF AROME) and parameters Mean Sea Level Pressure (MSLP), 2m air temperature (T2) and 10m wind speed (WS10). Verification period is YOPP SOP-NH1 and all forecasts are initialized at 00 UTC.*

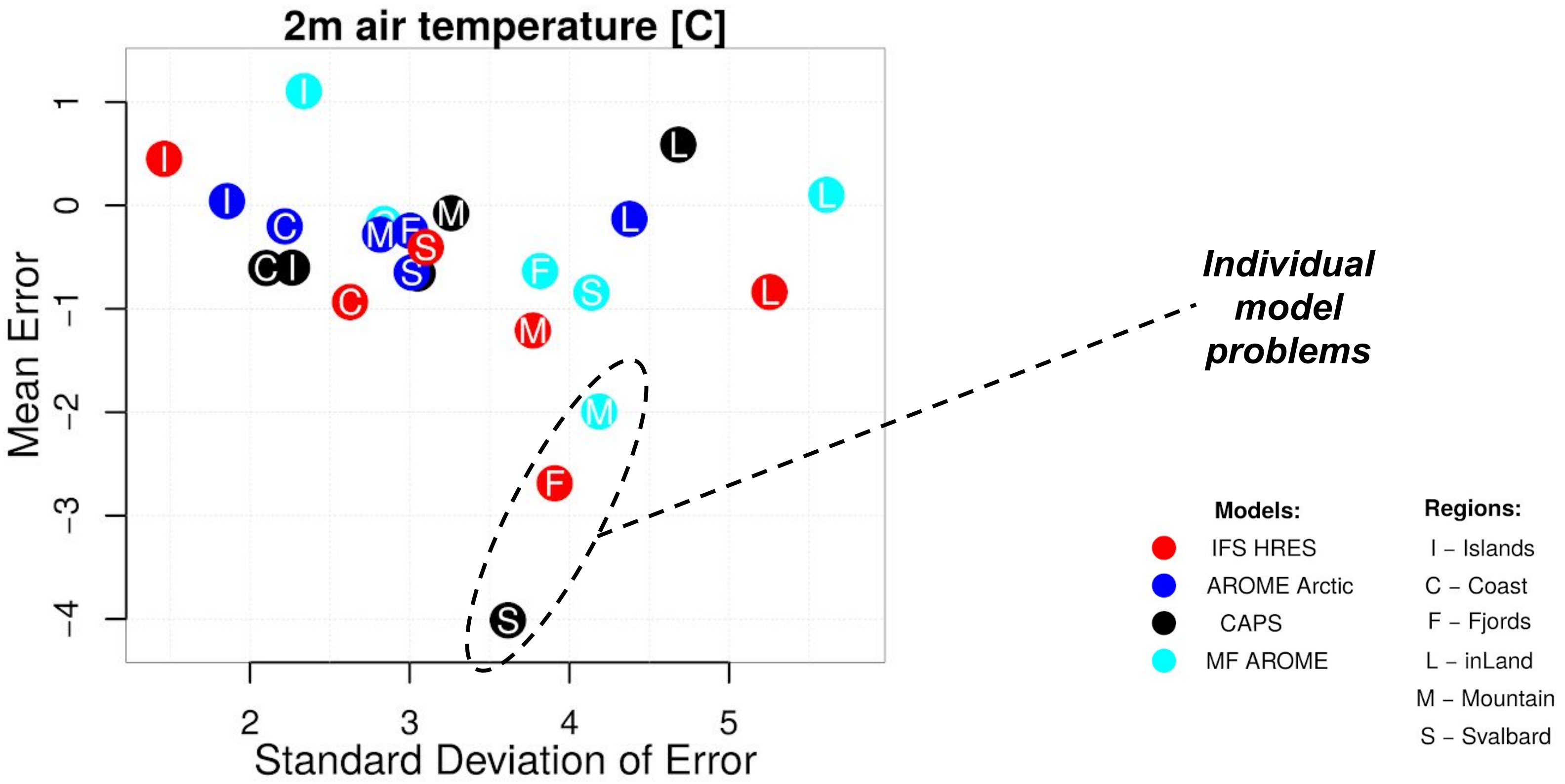
Small initial errors MSLP, but rapid growth. Large initial errors T2 and wind speed, but slow growth

Added value of high resolution models indicated for T2 and WS10, not for MSLP.

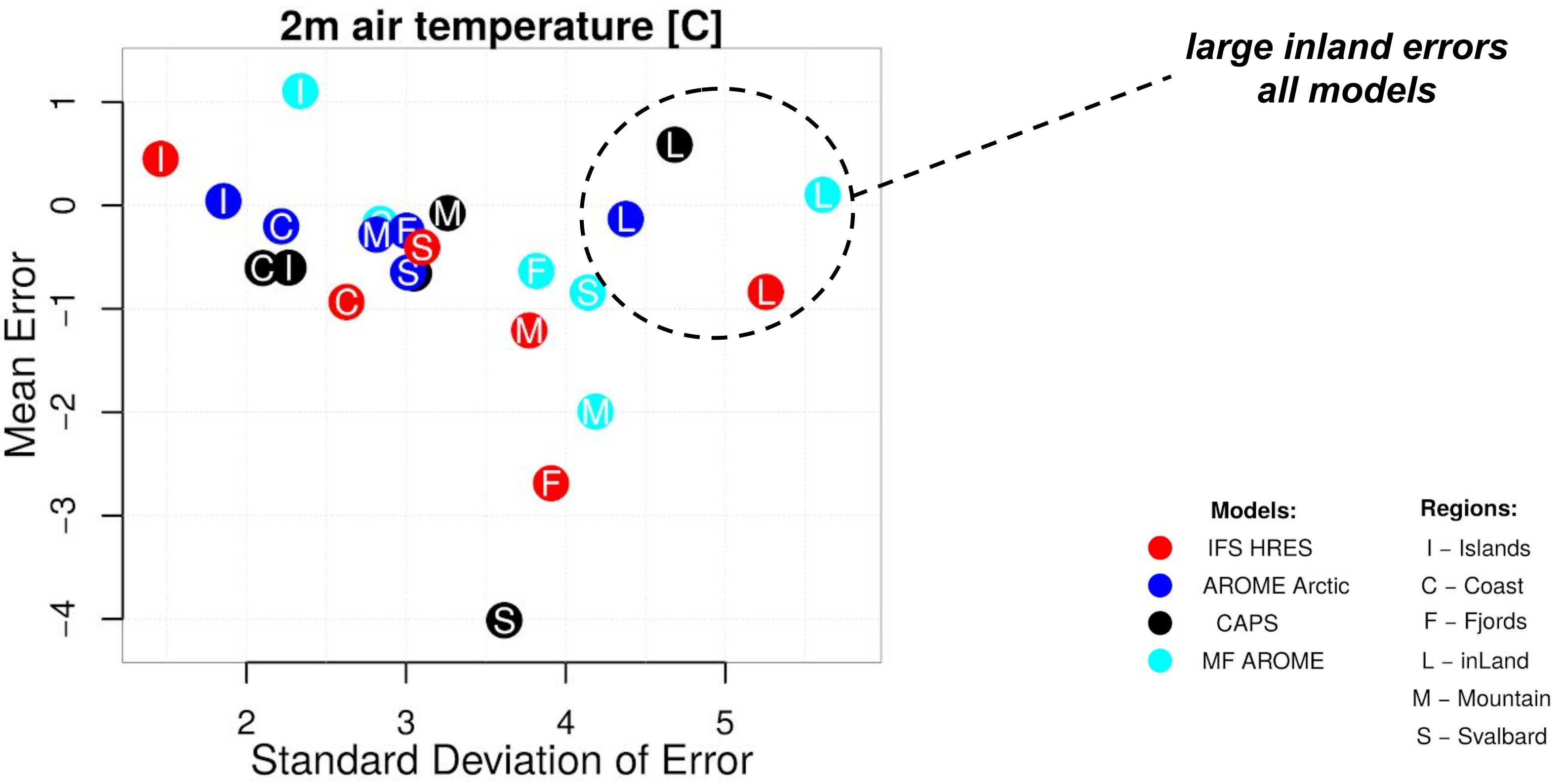
Faster error growth in high res. models than in global IFS-HRES.

A diurnal cycle and systematic errors revealed for some parameters and models.

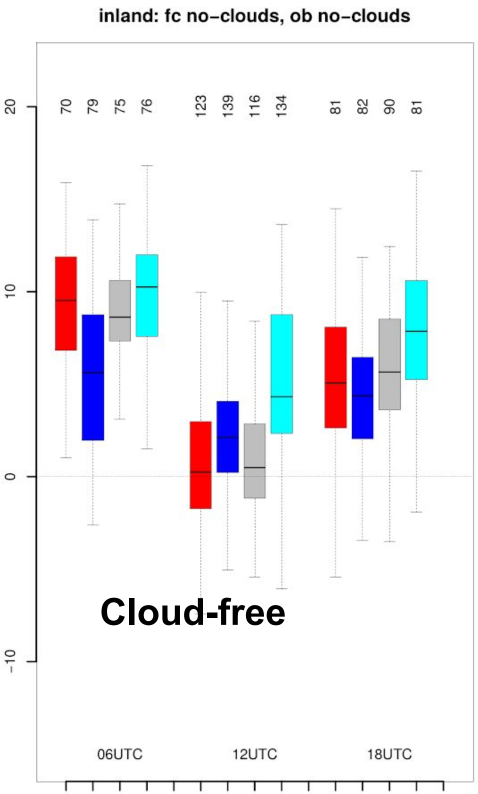
T2m forecast errors, function of region, "day 2 forecasts"

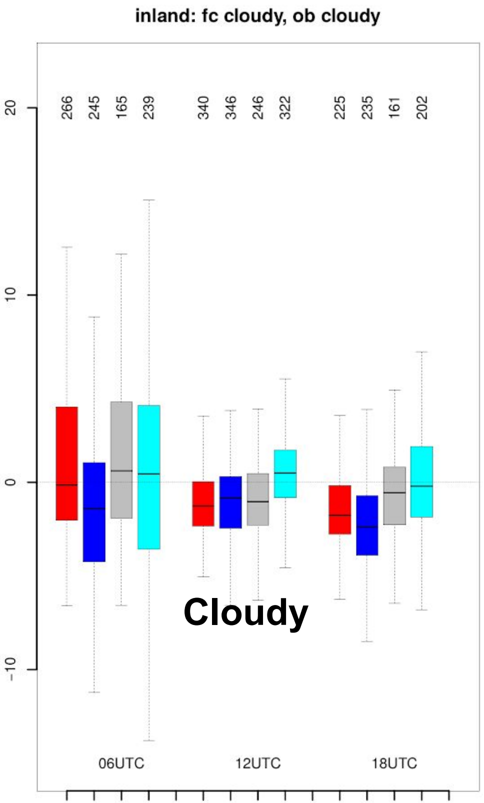


T2m forecast errors, function of region, “day 2 forecasts”



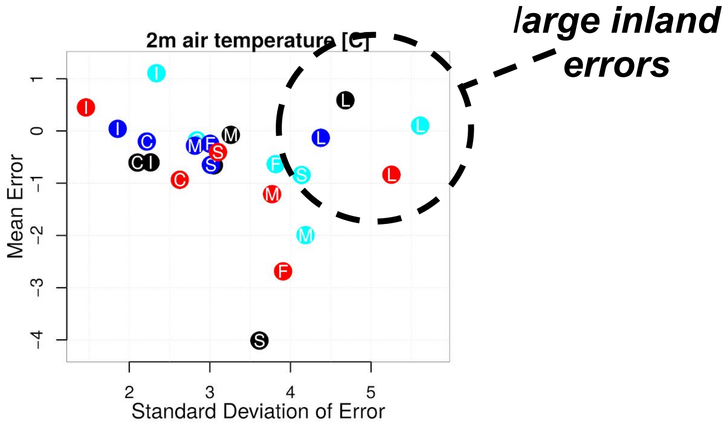
Temperature (T2), inland, function of clouds, “day 2 forecasts”



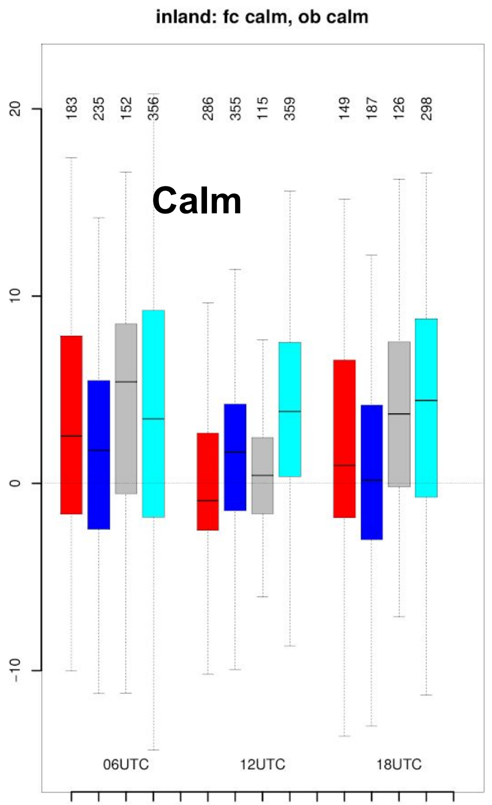


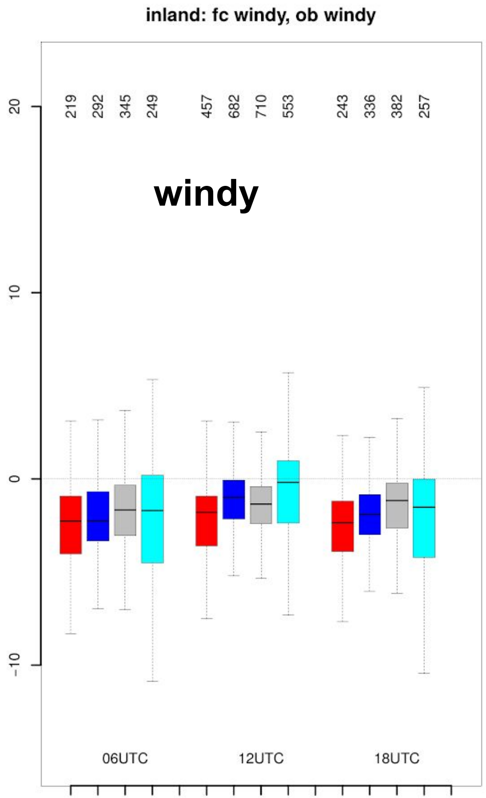
Conditional verification of T2 for inland stations. Box-and-whiskers plot of T2 errors (forecasted minus observed) conditioned by TCC (4 boxes to the left) and conditioned by wind (4 boxes to the right). Each box is divided into models and time of day. Number of cases is plotted at top and outliers is omitted to increase readability in plots.

IFS-HRES
AROME-Arctic
CAPS
MF-AROME



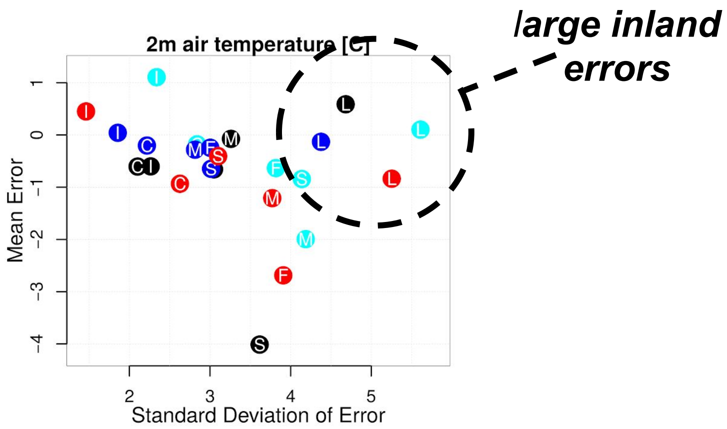
Temperature (T2), inland, function of wind, “day 2 forecasts”



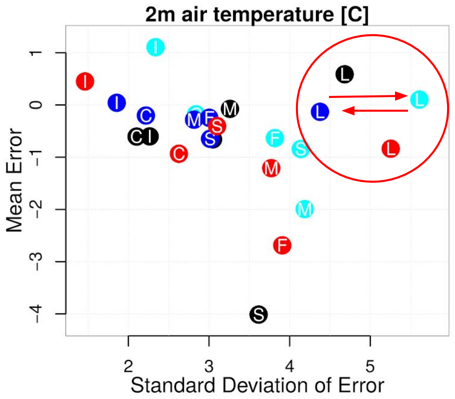


Conditional verification of T2 for inland stations. Box-and-whiskers plot of T2 errors (forecasted minus observed) conditioned by TCC (4 boxes to the left) and conditioned by wind (4 boxes to the right). Each box is divided into models and time of day. Number of cases is plotted at top and outliers is omitted to increase readability in plots.

IFS-HRES
AROME-Arctic
CAPS
MF-AROME



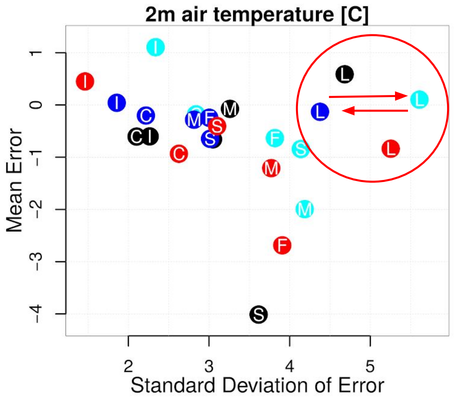
Importance of surface initial conditions



Differ in process descriptions, e.g. representation of surface boundary layers

Why large differences between AROME Arctic and MF AROME?

Importance of surface initial conditions



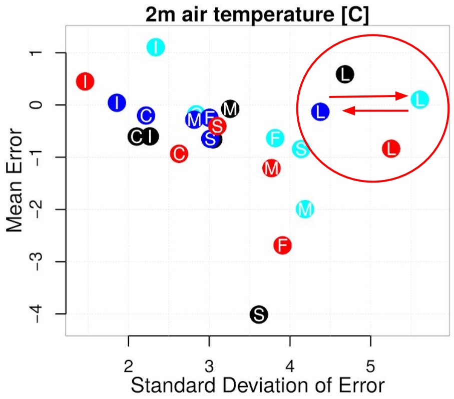
Differ in process descriptions, e.g. representation of surface boundary layers

Why large differences between AROME Arctic and MF AROME?

surface data assimilation

Interpolated from global model ARPEGE (short cut off analysis coarser resolution, interpolation issues)

Importance of surface initial conditions



Differ in process descriptions, e.g. representation of surface boundary layers

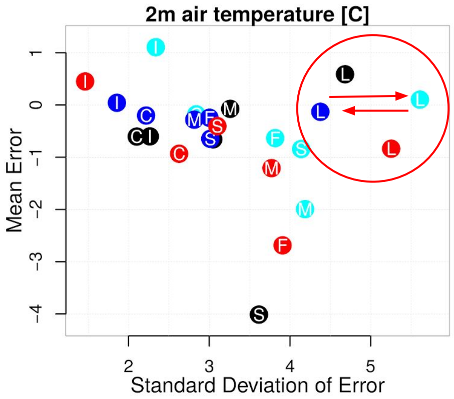
Why large differences between AROME Arctic and MF AROME?

surface data assimilation

Interpolated from global model ARPEGE (short cut off analysis coarser resolution, interpolation issues)

test 4 re-runs where surface initial conditions are taken from AROME-Arctic

Importance of surface initial conditions



Differ in process descriptions, e.g. representation of surface boundary layers

Why large differences between AROME Arctic and MF AROME?

surface data assimilation

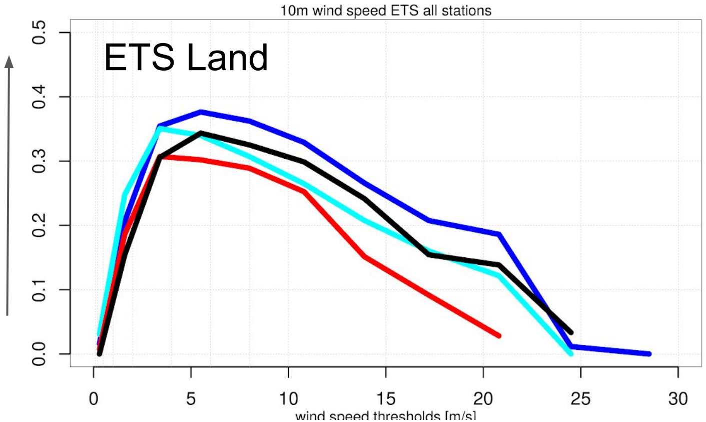
Interpolated from global model ARPEGE (short cut off analysis coarser resolution, interpolation issues)

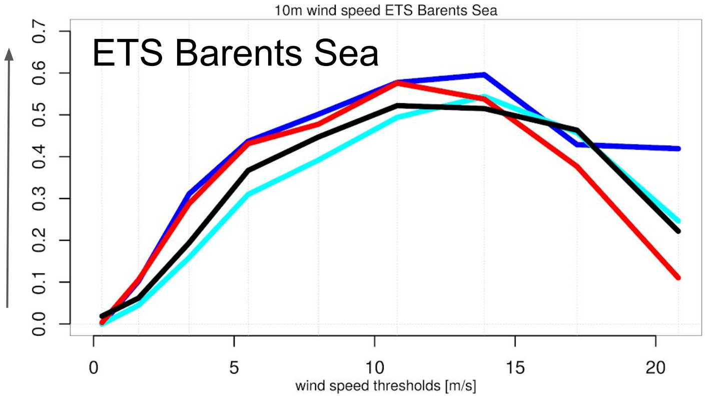
test 4 re-runs where surface initial conditions are taken from AROME-Arctic

Verification re-run cases only	***AROME-Arctic***	***MF AROME init by surface data assimilation***	***MF AROME init by interp. glob model***
Mean abs error	***3.1C***	***3.3C***	***4.0C***
Stand dev error	***3.8C***	***4.0C***	***4.8C***
Mean error	***1.3C***	***1.6C***	***1.7C***

The difference in initial conditions explains most of the difference between AROME-Arctic and MF AROME

Wind speed, categorical scores, “day 2 forecasts”



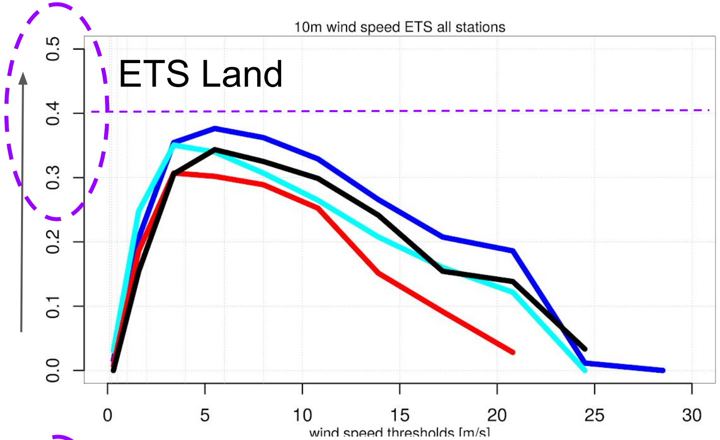


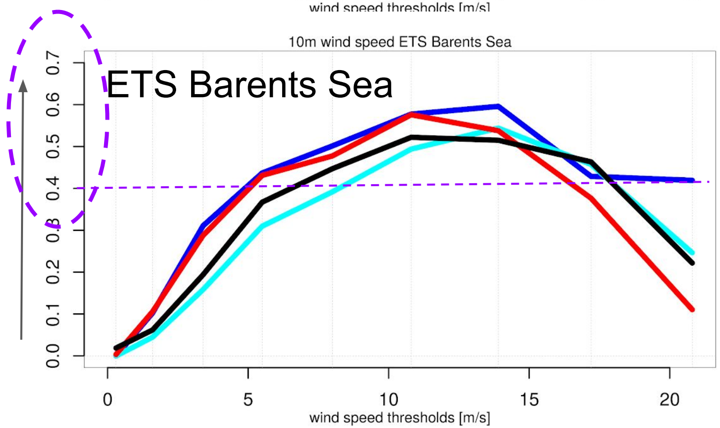
Equitable Threat Score (ETS) and Frequency Bias (FB) for wind speed over all synop stations used in the model-intercomparison. Models are IFS HRES (red), AROME Arctic (blue), CAPS (black) and AROME MF (cyan). Lead times from +25 to +48hr.

IFS-HRES
AROME-Arctic
CAPS
MF-AROME

Same as above, but WS10 forecasts are now compared with scatterometer based observed wind for an area in the Barents Sea (24-38E and 72-76N). Notice that the highest threshold (20.8m/s) include 311 observations and 80, 477, 288 and 895 for the four models, respectively.

Wind speed, categorical scores, “day 2 forecasts”





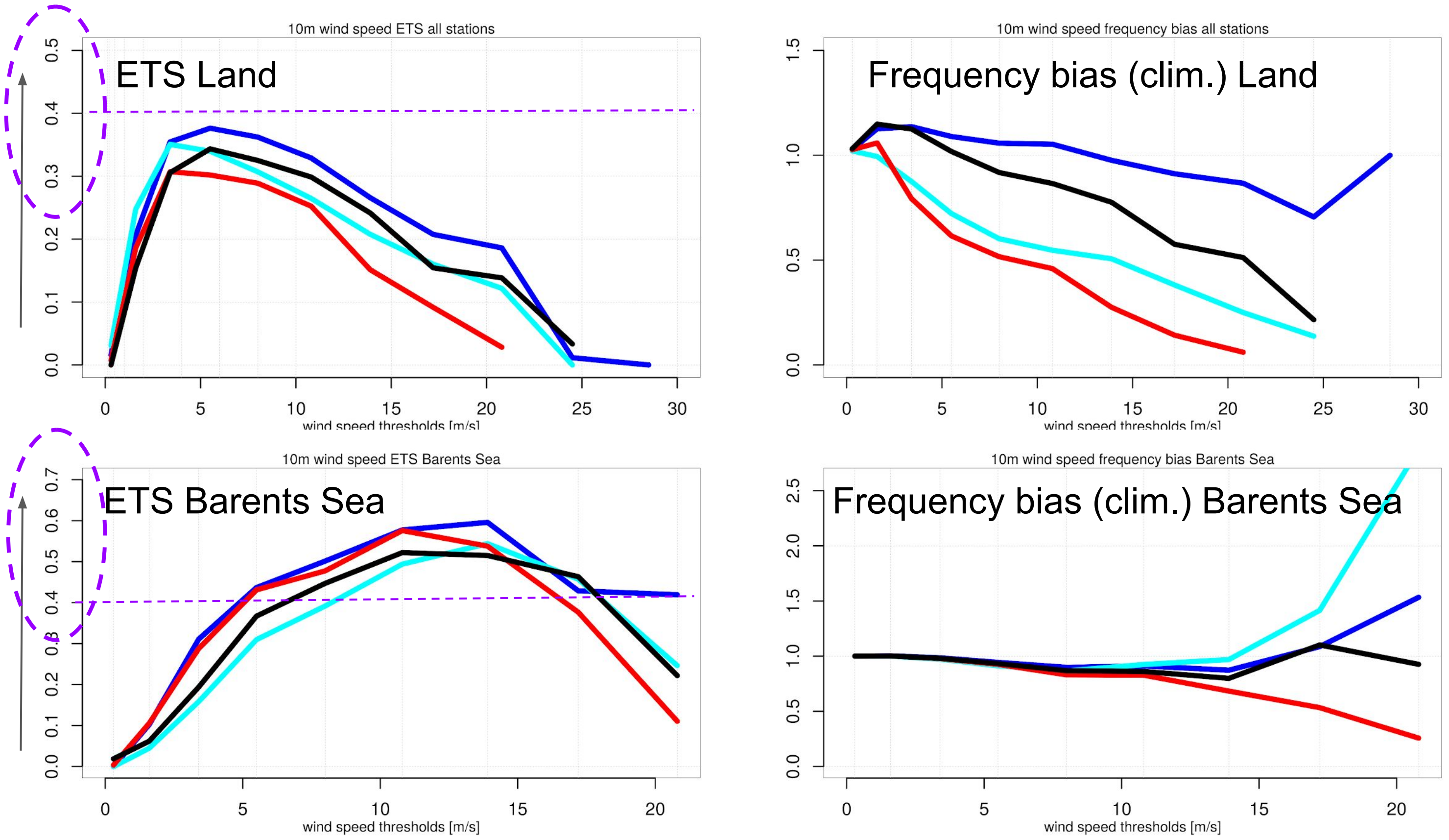
Higher skill (ETS)
over ocean (vs ASCAT)
than over land (vs SYNOP)

Equitable Threat Score (ETS) and Frequency Bias (FB) for wind speed over all synop stations used in the model-intercomparison. Models are IFS HRES (red), AROME Arctic (blue), CAPS (black) and AROME MF (cyan). Lead times from +25 to +48hr.

IFS-HRES
AROME-Arctic
CAPS
MF-AROME

Same as above, but WS10 forecasts are now compared with scatterometer based observed wind for an area in the Barents Sea (24-38E and 72-76N). Notice that the highest threshold (20.8m/s) include 311 observations and 80, 477, 288 and 895 for the four models, respectively.

Wind speed, categorical scores, “day 2 forecasts”



Equitable Threat Score (ETS) and Frequency Bias (FB) for wind speed over all synop stations used in the model-intercomparison. Models are IFS HRES (red), AROME Arctic (blue), CAPS (black) and AROME MF (cyan). Lead times from +25 to +48hr.

IFS-HRES
AROME-Arctic
CAPS
MF-AROME

Same as above, but WS10 forecasts are now compared with scatterometer based observed wind for an area in the Barents Sea (24-38E and 72-76N). Notice that the highest threshold (20.8m/s) include 311 observations and 80, 477, 288 and 895 for the four models, respectively.

Spatial representativity
an example

1. Forecasts represent a grid box average and differs from point observations

Spatial representativity
an example

1. Forecasts represent a grid box average and differs from point observations.
2. Assume Tromsø MET & Tromsø airport represent a model grid box (2.7 km apart) and that the “perfect forecast” for that grid box is the average of the observations in the grid box (Göber et al., 2008)

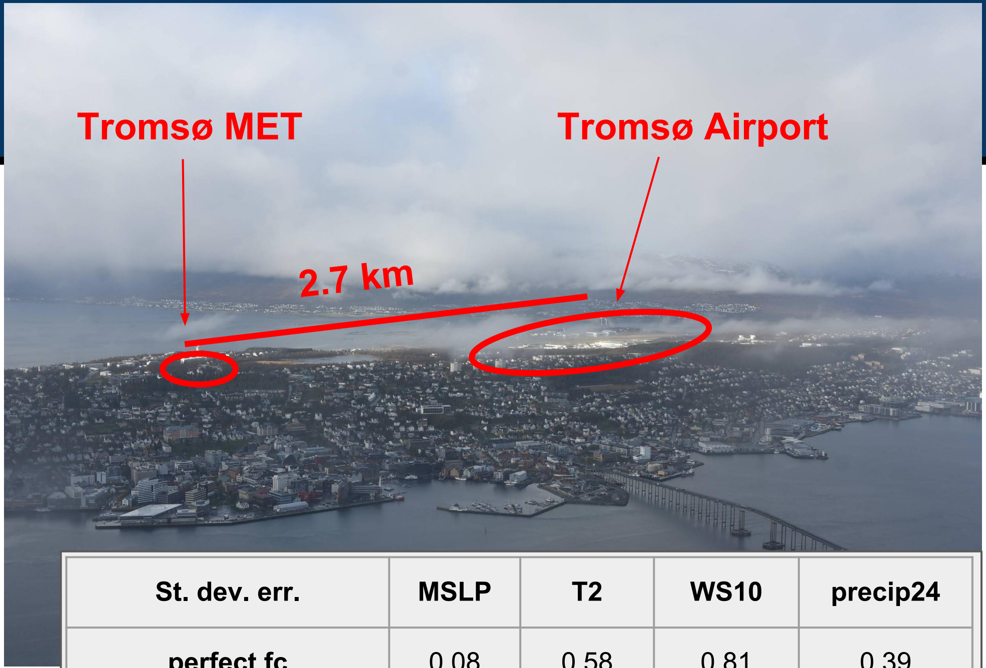
Spatial representativity
an example

1. Forecasts represent a grid box average and differs from point observations.
2. Assume Tromsø MET & Tromsø airport represent a model grid box (2.7 km apart) and that the "perfect forecast" for that grid box is the average of the observations in the grid box (Göber et al., 2008)
3. Verify the "perfect forecast" against Tromsø MET and Tromsø airport (error > 0). Errors are due to representativity issues.

St. dev. err.	MSLP	T2	WS10	precip24
perfect fc	0.08	0.58	0.81	0.39

Spatial representativity an example

1. Forecasts represent a grid box average and differs from point observations.
2. Assume Tromsø MET & Tromsø airport represent a model grid box (2.7 km apart) and that the "perfect forecast" for that grid box is the average of the observations in the grid box (Göber et al., 2008)
3. Verify the "perfect forecast" against Tromsø MET and Tromsø airport (error > 0). Errors are due to representativity issues.
4. Verify NWP models and compare with "perfect forecast".

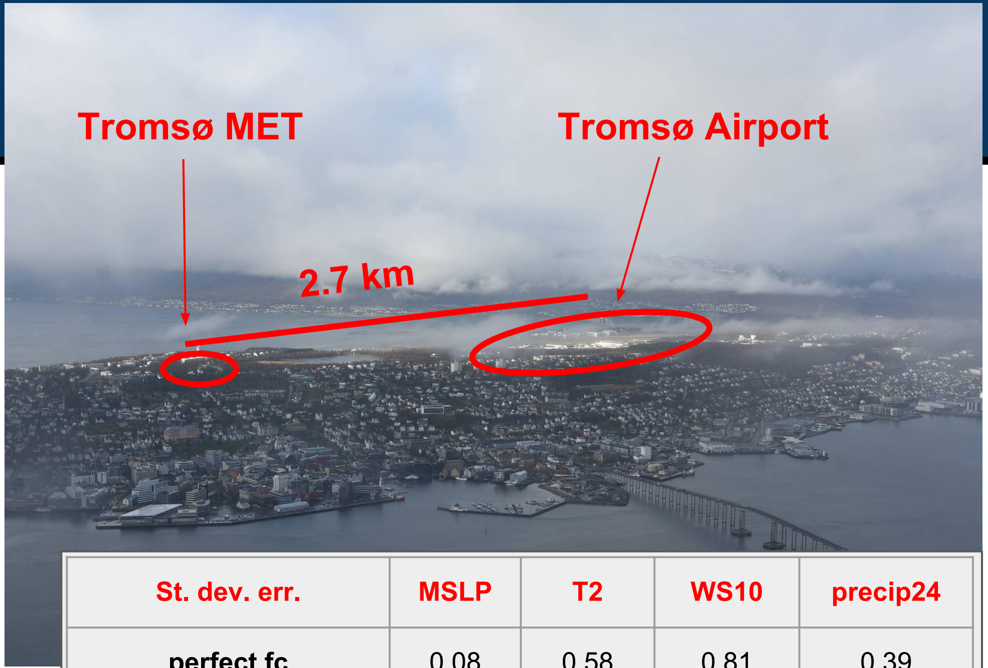


St. dev. err.	MSLP	T2	WS10	precip24
perfect fc	0.08	0.58	0.81	0.39
IFS HRES	0.72	3.04	2.25	2.57
AROME-Arctic	0.97	2.09	1.91	2.55
CAPS	1.27	1.67	2.06	2.36
MF AROME	NA	2.75	1.95	1.98
% of model error	6-11%	19-35%	36-42%	15-20%

Spatial representativity an example

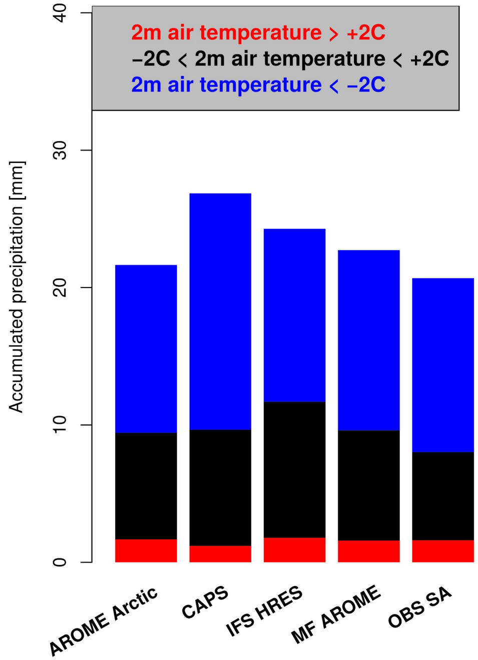
1. Forecasts represent a grid box average and differs from point observations.
2. Assume Tromsø MET & Tromsø airport represent a model grid box (2.7 km apart) and that the “perfect forecast” for that grid box is the average of the observations in the grid box (Göber et al., 2008)
3. Verify the “perfect forecast“ against Tromsø MET and Tromsø airport (error > 0). Errors are due to representativity issues.
4. Verify NWP models and compare with “perfect forecast”.

A substantial part of the difference between short range forecasts and synop observations can be explained by observation representativity issues (as also indicate by other results)



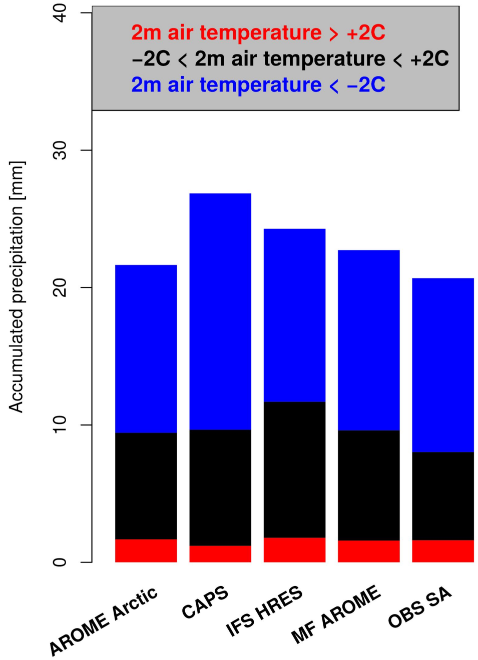
St. dev. err.	MSLP	T2	WS10	precip24
perfect fc	0.08	0.58	0.81	0.39
IFS HRES	0.72	3.04	2.25	2.57
AROME-Arctic	0.97	2.09	1.91	2.55
CAPS	1.27	1.67	2.06	2.36
MF AROME	NA	2.75	1.95	1.98
% of model error	6-11%	19-35%	36-42%	15-20%

Daily precipitation forecast errors, function of region, “day 2 forecasts”



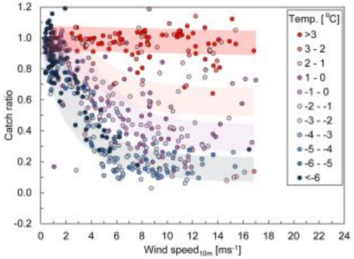
Accumulated precipitation (estimated by temperature thresholds; rain in red, sleet in black and solid precipitation in blue) for AROME Arctic, CAPS, IFS HRES, AROME MF with lead times +18 to +42hr, observed precipitation from Geonor rain gauges with single alter shields

Daily precipitation forecast errors, function of region, “day 2 forecasts”



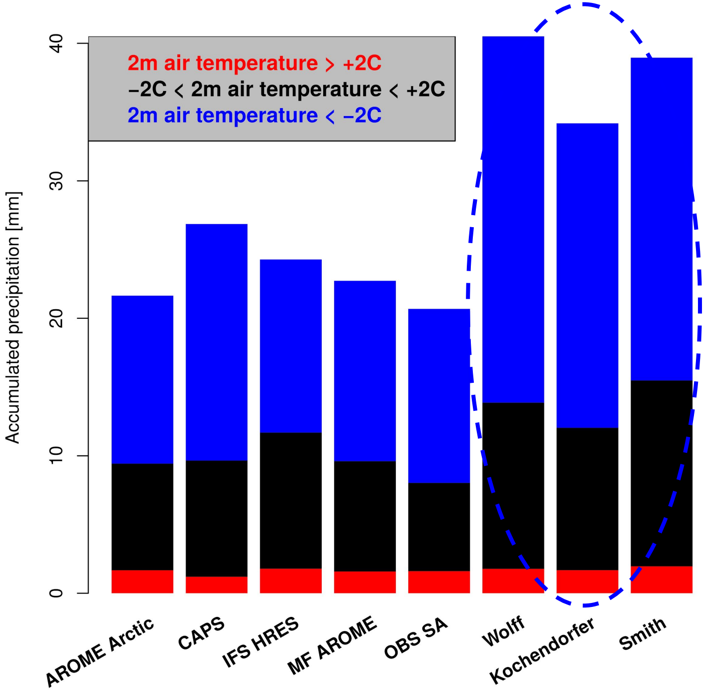
Accumulated precipitation (estimated by temperature thresholds; rain in red, sleet in black and solid precipitation in blue) for AROME Arctic, CAPS, IFS HRES, AROME MF with lead times +18 to +42hr, observed precipitation from Geonor rain gauges with single alter shields

1. Solid precipitation are heavily underestimated in windy conditions (Rasmussen et al., 2012).
2. From parallel observations with “double fence shield” and “single alter shield” adjustment algorithms for observed precipitation are established.



Wolff et al. (2015)

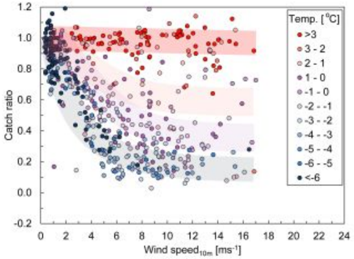
Daily precipitation forecast errors, function of region, “day 2 forecasts”



Accumulated precipitation (estimated by temperature thresholds; rain in red, sleet in black and solid precipitation in blue) for AROME Arctic, CAPS, IFS HRES, AROME MF with lead times +18 to +42hr, observed precipitation from Geonor rain gauges with single alter shields, observed precipitation corrected with Wolff et al. (2015), by Kochendorfer et al. (2017) and Smith (2007). The accumulated precipitation amounts are averaged over 21 stations.

In reality, all models underestimate the amount of solid precipitation!

1. Solid precipitation are heavily underestimated in windy conditions (Rasmussen et al., 2012).
2. From parallel observations with “double fence shield” and “single alter shield” adjustment algorithms for observed precipitation are established.



Wolff et al. (2015)

3. For 21 stations during YOPP SOP NH1; single alter shield, hourly observations of precipitation, wind speed and temperature, can estimate “real precipitation”.
4. Forecasted precipitation can be compared with adjusted observations (note that only accumulated precipitation is compared, no skill evaluation).

Summary

- Paper recently submitted (with APPLICATE acknowledgement)
- Three high resolution limited area models and one coarser resolution global model are compared during YOPP SOP NH1 in the Barents Sea, Svalbard and Northern Scandinavia.
- The forecast capabilities varies between parameter, region and models. No model system is superior for all parameters, regions and lead times. High resolution models add value to the coarser resolution global model, but not for all parameters, regions and lead times.
- The NWP systems have common weaknesses (e.g. inland temperatures, underestimation of precipitation, representation of spatial variability in wind speed, ….).
- Model specific weakness (or more pronounced in specific systems) are found (e.g. CAPS: temperature Svalbard, IFS-HRES: fjord temperatures AROME-Arctic/ MF-AROME: (coastal) precipitation, IFS-HRES/MF-AROME: underestimation of wind speed, ….).
- Important to take observation errors into account (e.g. reveal underestimation of solid precipitation).
- A substantial part of the difference between forecasts and observations arise from representativity issues which need considerations in the verification process

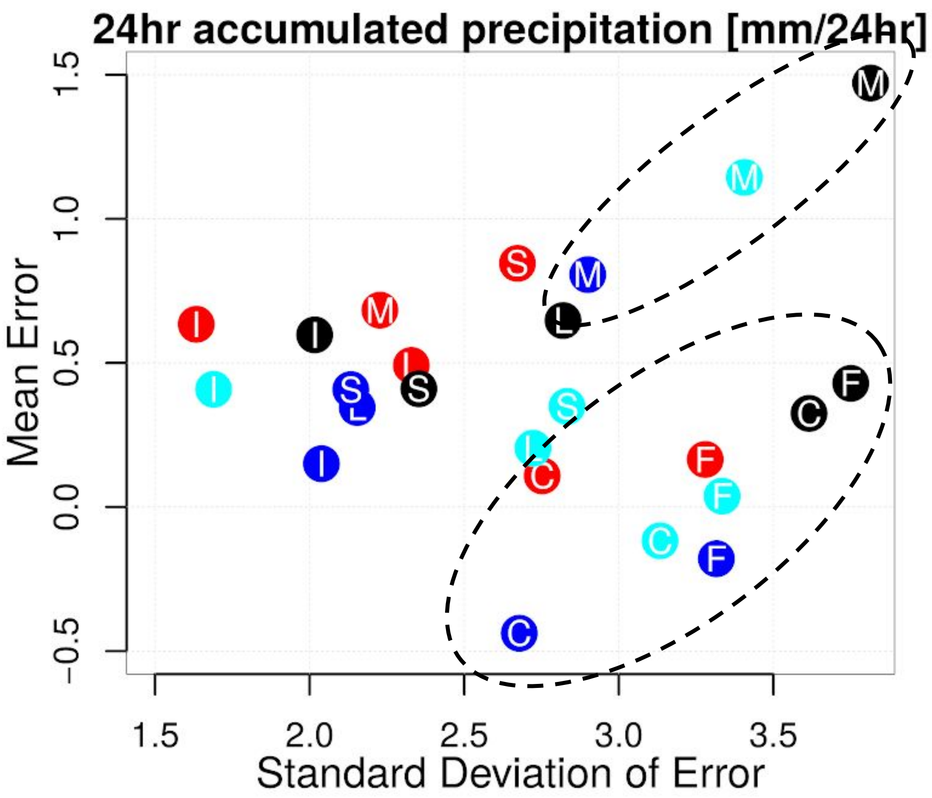
YOPP YEAR OF POLAR PREDICTION

APPLICATE.eu Advanced prediction in polar regions and beyond

Alertness

Photo: Ketil Isaksen, MET Norway

Daily precipitation forecast errors, function of region, “day 2 forecasts”



More precipitation in IFS HRES and CAPS compared to AROME-Arctic and MF-AROME

Positive bias, larger errors in mountain areas

Large errors and small positive (IFS HRES and CAPS) and negative (AROME-Arctic and MF-AROME) biases in coast and fjords

Models:	Regions:
IFS HRES	I – Islands
AROME Arctic	C – Coast
CAPS	F – Fjords
MF AROME	L – inLand
	M – Mountain
	S – Svalbard