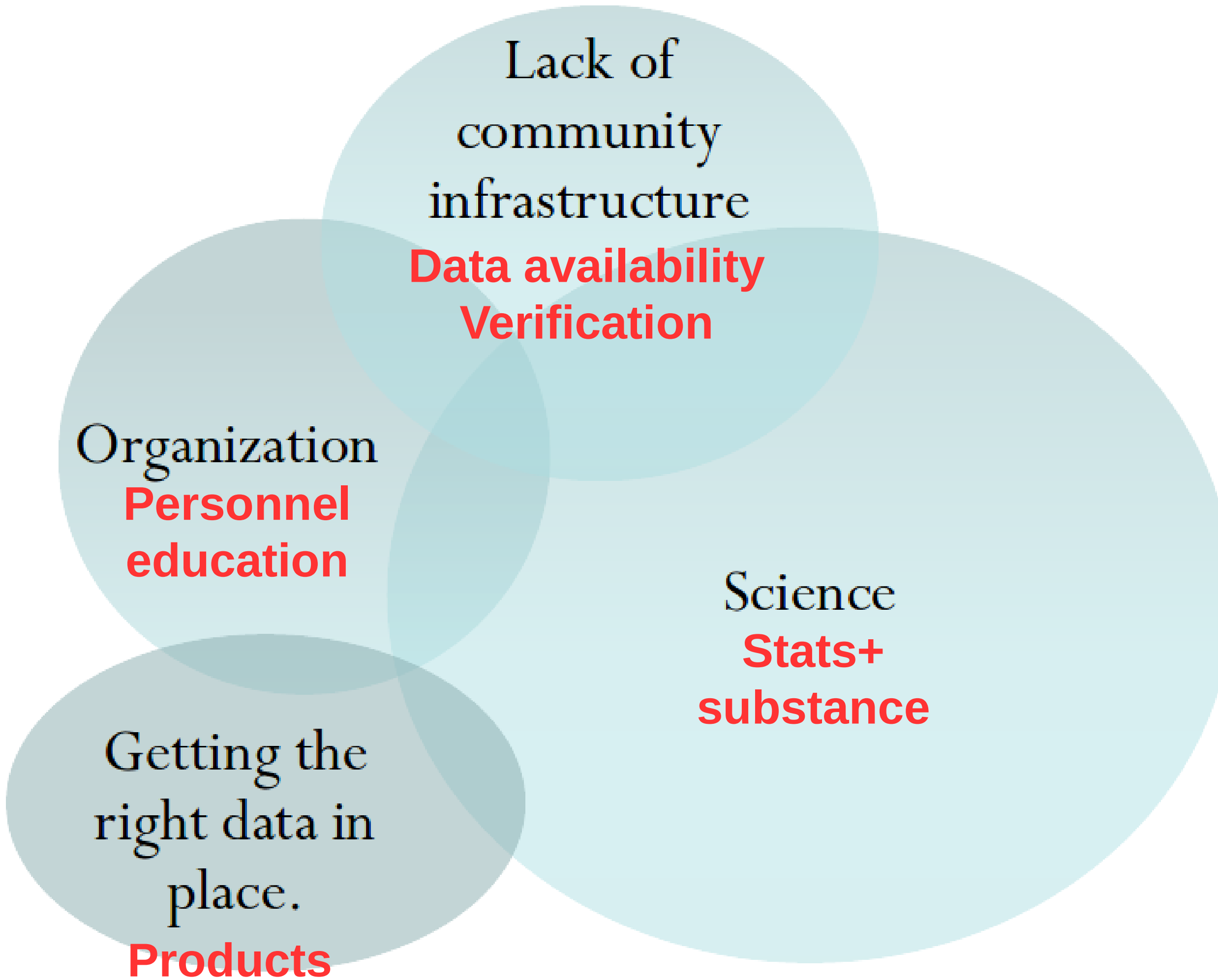
FINNISH METEOROLOGICAL INSTITUTE

Model Output Statistics (MOS) at FMI

Jussi S. Ylhäisi, with the help of several others
jussi.ylhaisi@fmi.fi
EMS2017, 6th Sep 2017

Post-processing Challenges

Note the substantial overlap.



https://www.weather.gov/media/sti/nggps/02%20NGGPS%20Program%20StatPP%20201608%20v7.pdf

Alternative point-of-view

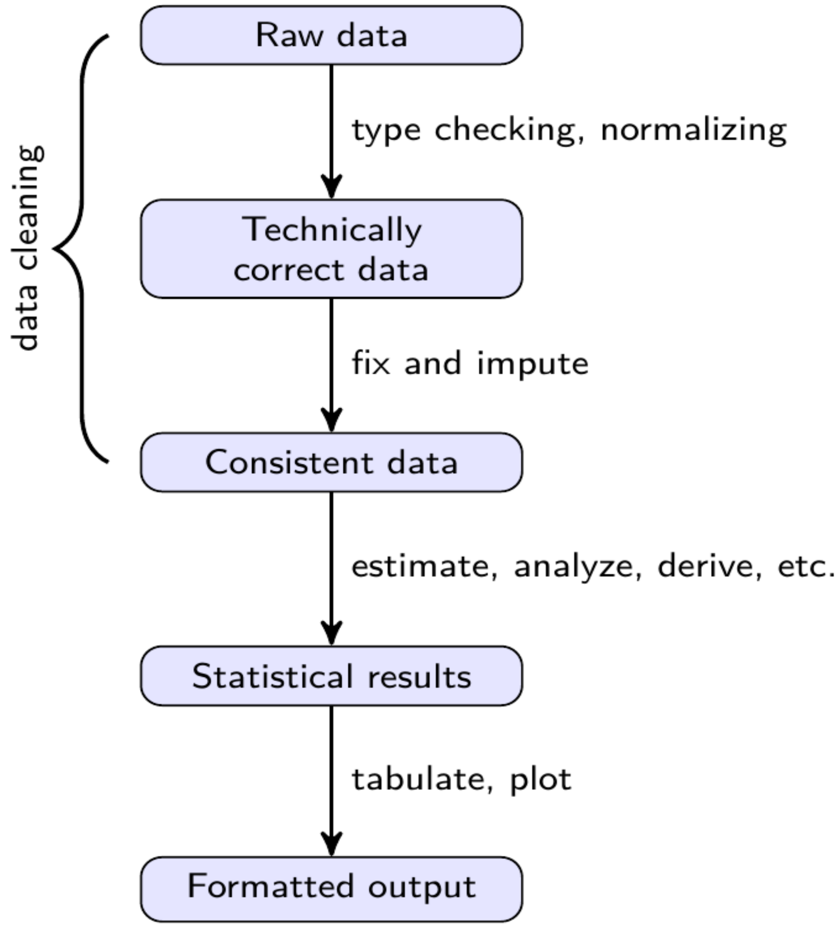
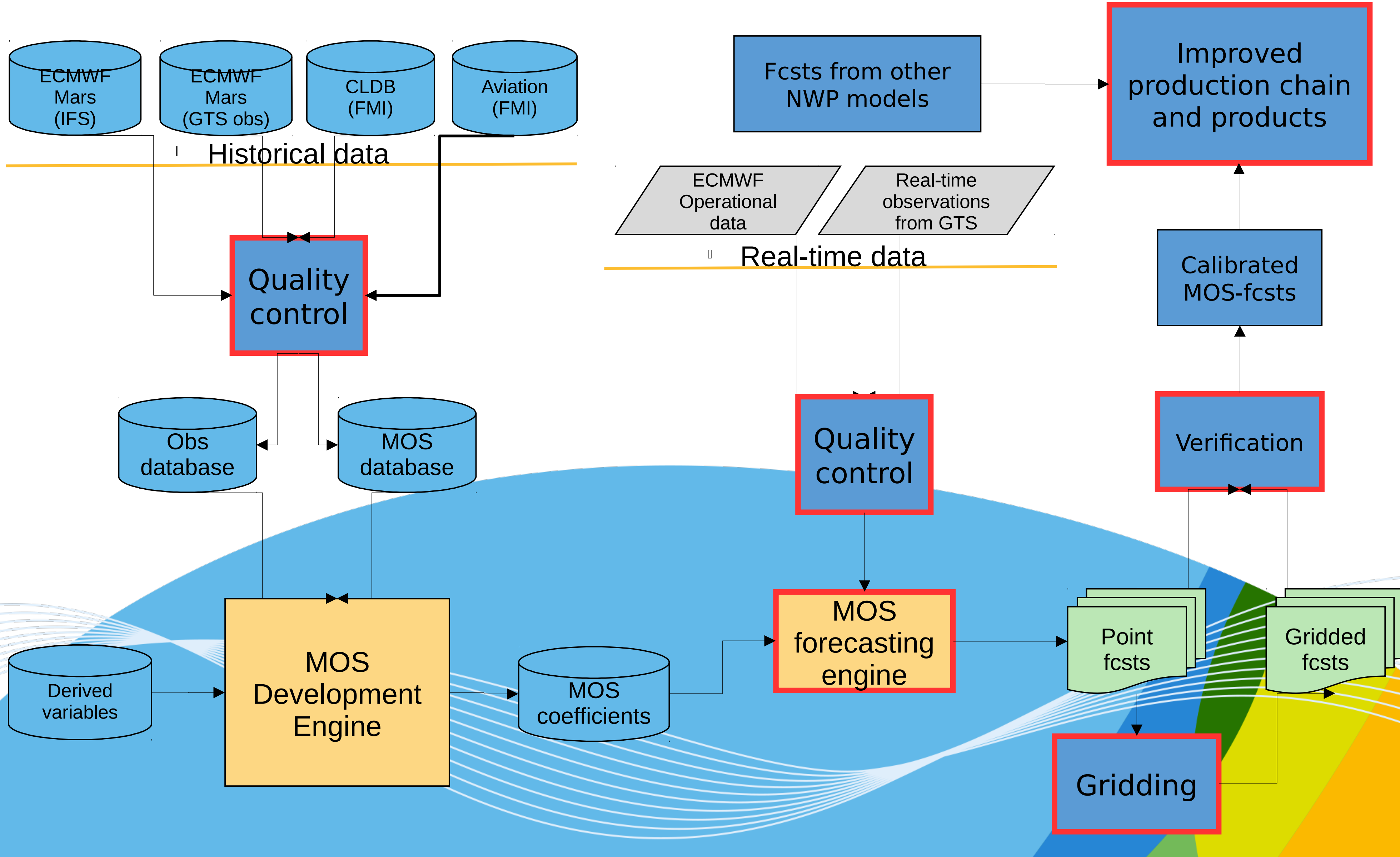


Figure 1: Statistical analysis value chain

Jonge and van der Loo, 2013

FINNISH METEOROLOGICAL INSTITUTE

MOS production system



FINNISH METEOROLOGICAL INSTITUTE

ECMWF training data

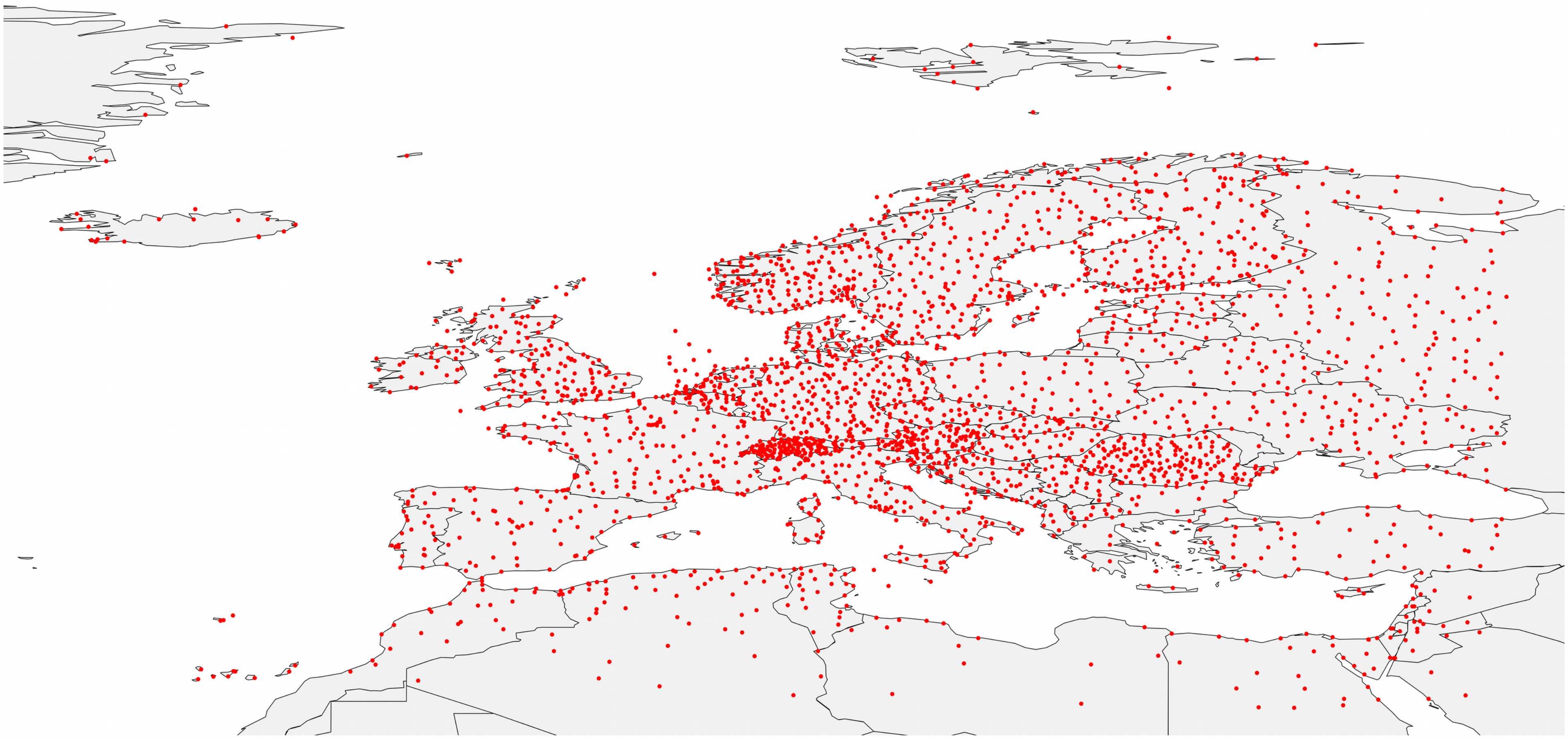
The training data for MOS comprises from ECMWF IFS data, currently spanning over a time period 01-12-2011 … real-time for <3000 European stations. The model variables used for training the linear model are listed below. More complicated predictor variable conversions have only been tested on desktop so far.

abbrev.	variable	unit
MSL	Mean sea level pressure	P-PA
T2	2m temperature	T-K
D2	2m dewpoint temperature	TD-K
MX2T3	Maximum temperature at 2m in the last 3 hours	TMAX3H-K
MN2T3	Minimum temperature at 2m in the last 3 hours	TMIN3H-K
SKT	Skin temperature	SKT-K
SSTK	Sea surface temperature	TS-K
TP	Total precipitation	RR-KGM2
LSP	Large-scale precipitation	RRL-KGM2
CP	Convective precipitation	RRC-KGM2
U10	10m U-wind speed	U-MS
U_100M	100m U-wind speed	U-MS
V10	10m V-wind speed	V-MS
V_100M	100m V-wind speed	V-MS
FG10_3	10 metre wind gust in the last 3 hours	FFG-MS
SD	Snow depth	SD-M
RSN	Snow density	SND-KGM3
BLH	Boundary layer height	MIXHGT-M
DEG0	Zero degree line	H0C-M
CBH	Cloud-based height	CLDBASE-M
LCC	Low cloud cover	NL-PRCNT
MCC	Medium cloud cover	NM-PRCNT
HCC	High cloud cover	NH-PRCNT
CAPE	Convective available potential energy	CAPE-JKG
CIN	Convective inhibition	CIN-N
TCW	Total column water	TOTCWV-KGM2
STRD	Surface thermal radiation downwards	RADLW-WM2
SLHF	Surface latent heat flux	FLLAT-JM2
SSHF	Surface sensible heat flux	FLSEN-JM2
SSR	Surface net solar radiation	RNETSW-WM2
STR	Surface net thermal radiation	RNETLW-WM2

T_500	Temperature at 500hPa	T-K
T_700	Temperature at 700hPa	T-K
T_850	Temperature at 850hPa	T-K
T_925	Temperature at 925hPa	T-K
T_950	Temperature at 950hPa	T-K
Z_500	Geopotential height at 500hPa	Z-M2S2
Z_700	Geopotential height at 700hPa	Z-M2S2
Z_850	Geopotential height at 850hPa	Z-M2S2
Z_925	Geopotential height at 925hPa	Z-M2S2
Z_950	Geopotential height at 950hPa	Z-M2S2
RH_500	Relative humidity at 500hPa	RH-PRCNT
RH_700	Relative humidity at 700hPa	RH-PRCNT
RH_850	Relative humidity at 850hPa	RH-PRCNT
RH_925	Relative humidity at 925hPa	RH-PRCNT
RH_950	Relative humidity at 950hPa	RH-PRCNT
W_500	Vertical velocity at 500hPa	VV-PAS
W_700	Vertical velocity at 700hPa	VV-PAS
W_850	Vertical velocity at 850hPa	VV-PAS
W_925	Vertical velocity at 925hPa	VV-PAS
W_950	Vertical velocity at 950hPa	VV-PAS

Spatial coverage

asemia alueella 2879 / 8680 kpl



Observation-model pairs are formed for each individual...

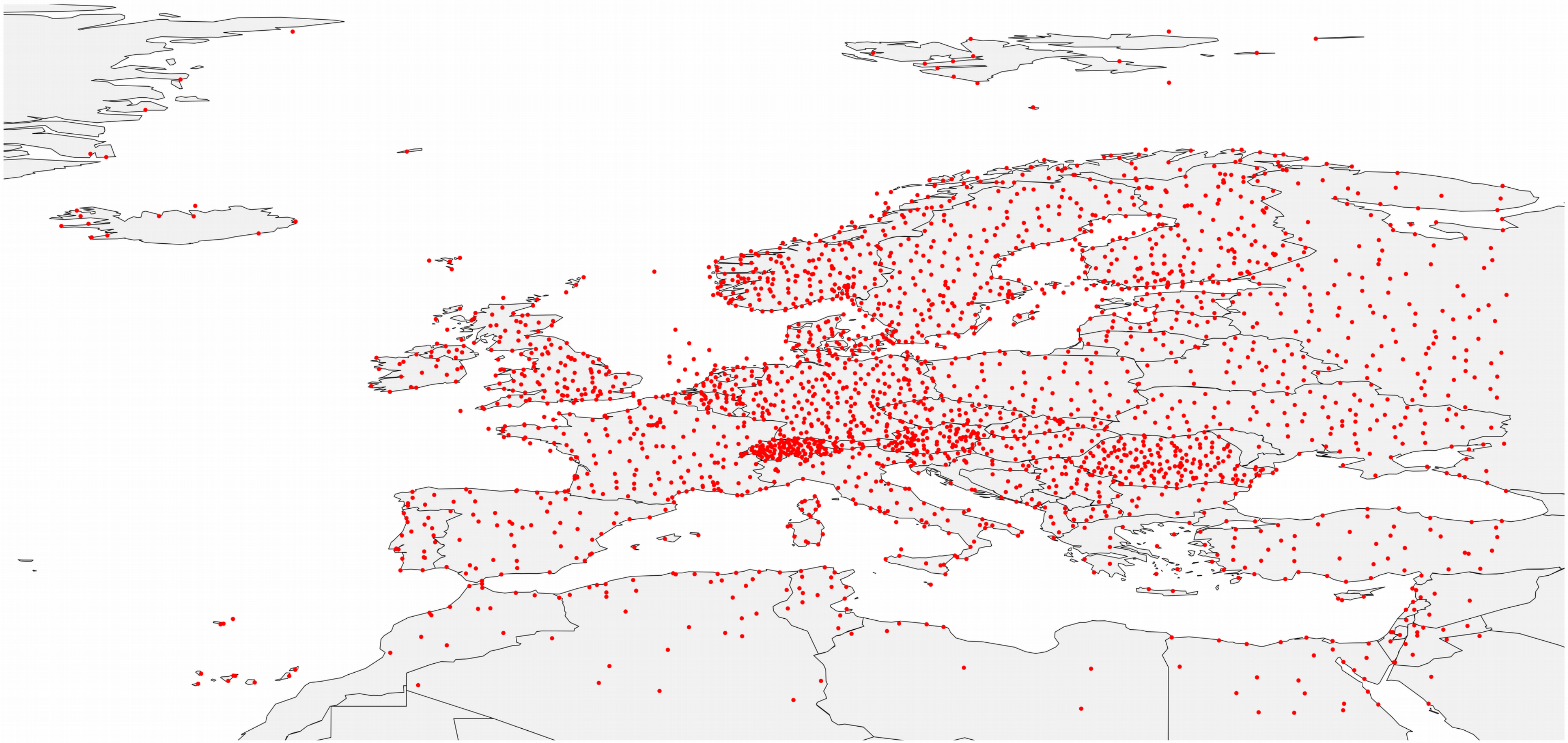
- ...station (2879*)
- ...forecast length (65*)
- ...analysis hour (2*)
- ...season (4*)

=> All-year-round working 10-day MOS forecasting system for European domain is compromised from 1 497 080 statistical models. In practice the number of models is less, mostly because observations are missing / are done too sparsely on some stations that would otherwise fill the criteria for sufficiently long training period: It is well possible that MOS forecasts for some stations are produced only for the other analysis hour or for a few seasons.

FINNISH METEOROLOGICAL INSTITUTE

Spatial coverage

2615 / 2879 kpl



Multiple linear regression problem

A simple statistical model with

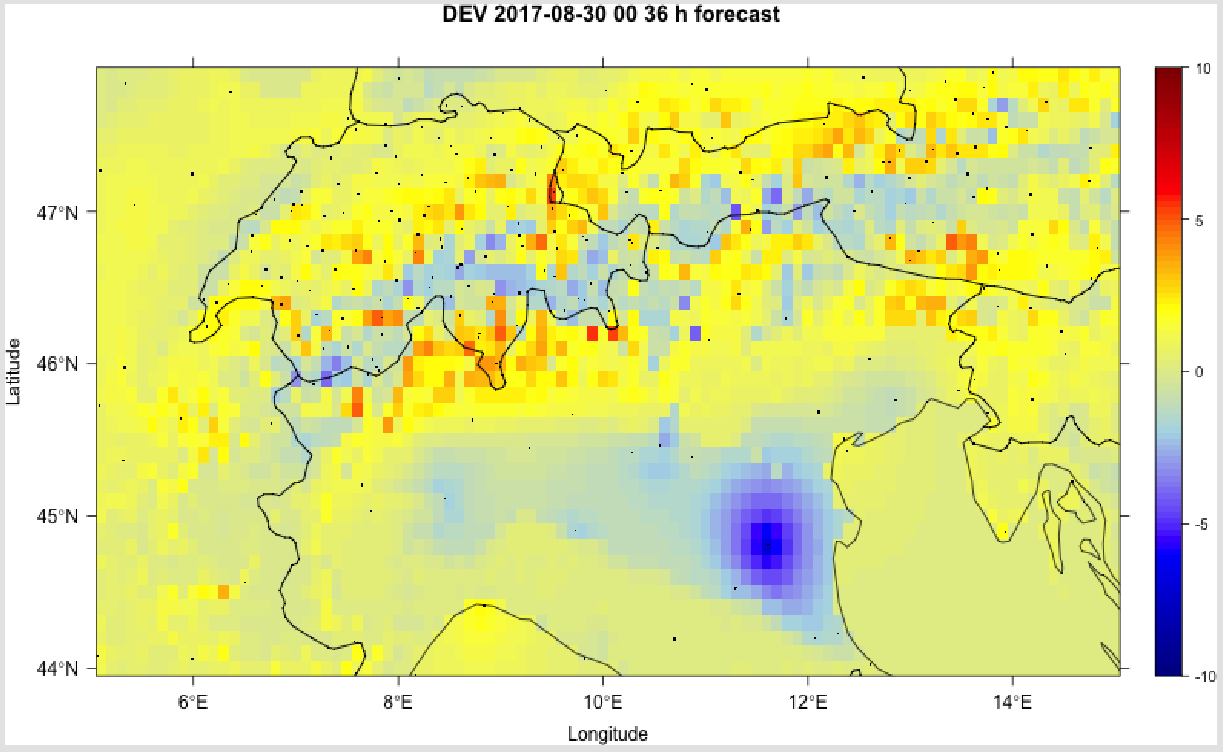
$$\hat{y} = b_o + b_1x_1 + b_2x_2 + \dots + b_kx_k$$

$\hat{y}$ being the predictand
x = 1, 2, …, k being predictors
b = 1, 2, …, k coefficients for the predictors

Operative version uses a constant set of 9 predictors for all forecast lengths.
Development version using Elastic net Lasso predictor screening is running in parallel.

Gridding the forecasts

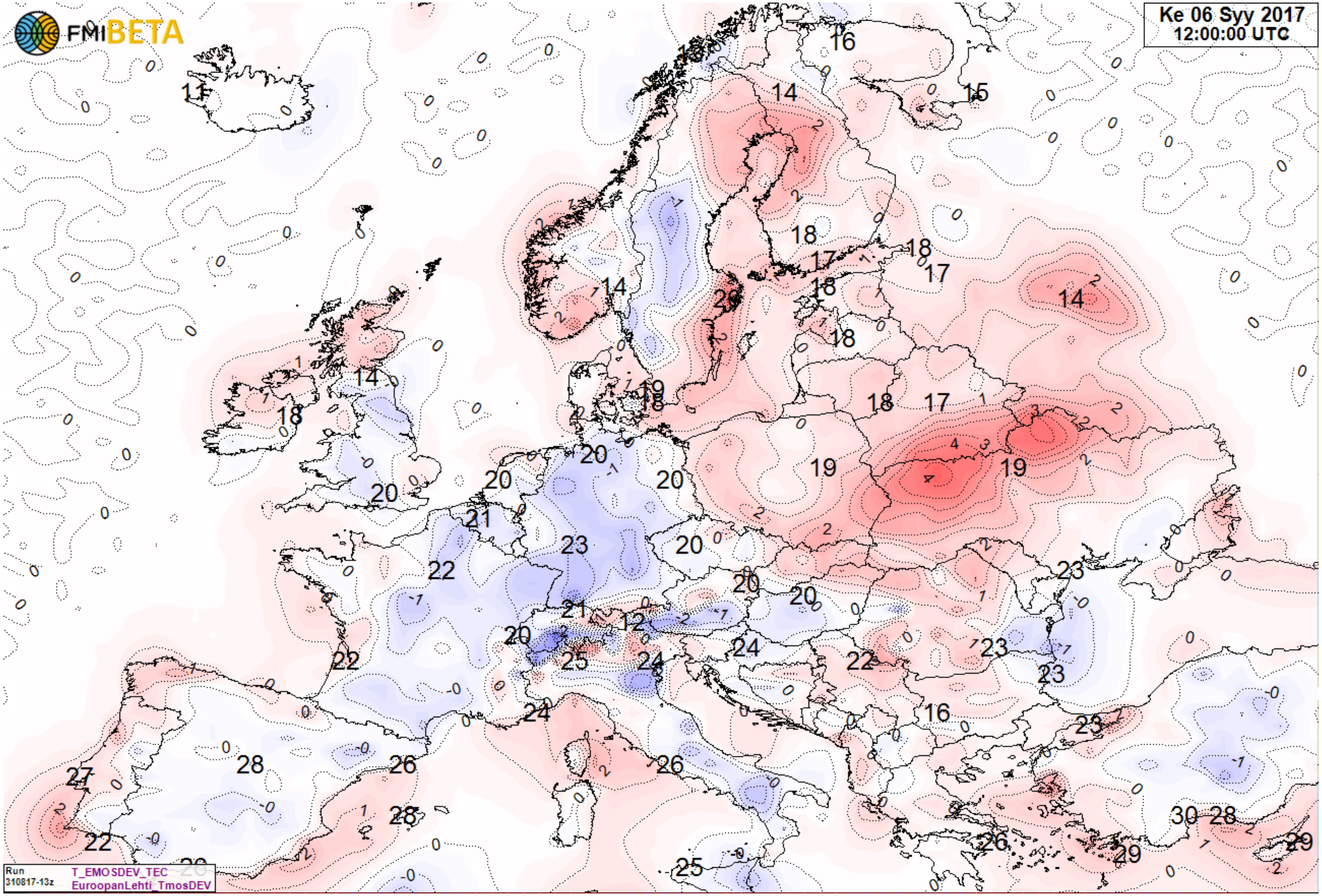
The gridding of the MOS station temperature forecasts is done by Kriging method depending on both geographical and altitude distance and using ECMWF forecast as a background field. Land/sea station points are uncorrelated between each other.



Lat: 27.5 - 73.5
Lon: -40.0 - 72.5
Gridpoints x: 1126
Gridpoints y: 461
Resolution (Deg): 0.1~12km

Difference between ECMWF backgound field and Kriging-interpolated forecast

Gridding the forecasts



Difference between ECMWF backgound field and Kriging-interpolated 6-day forecast

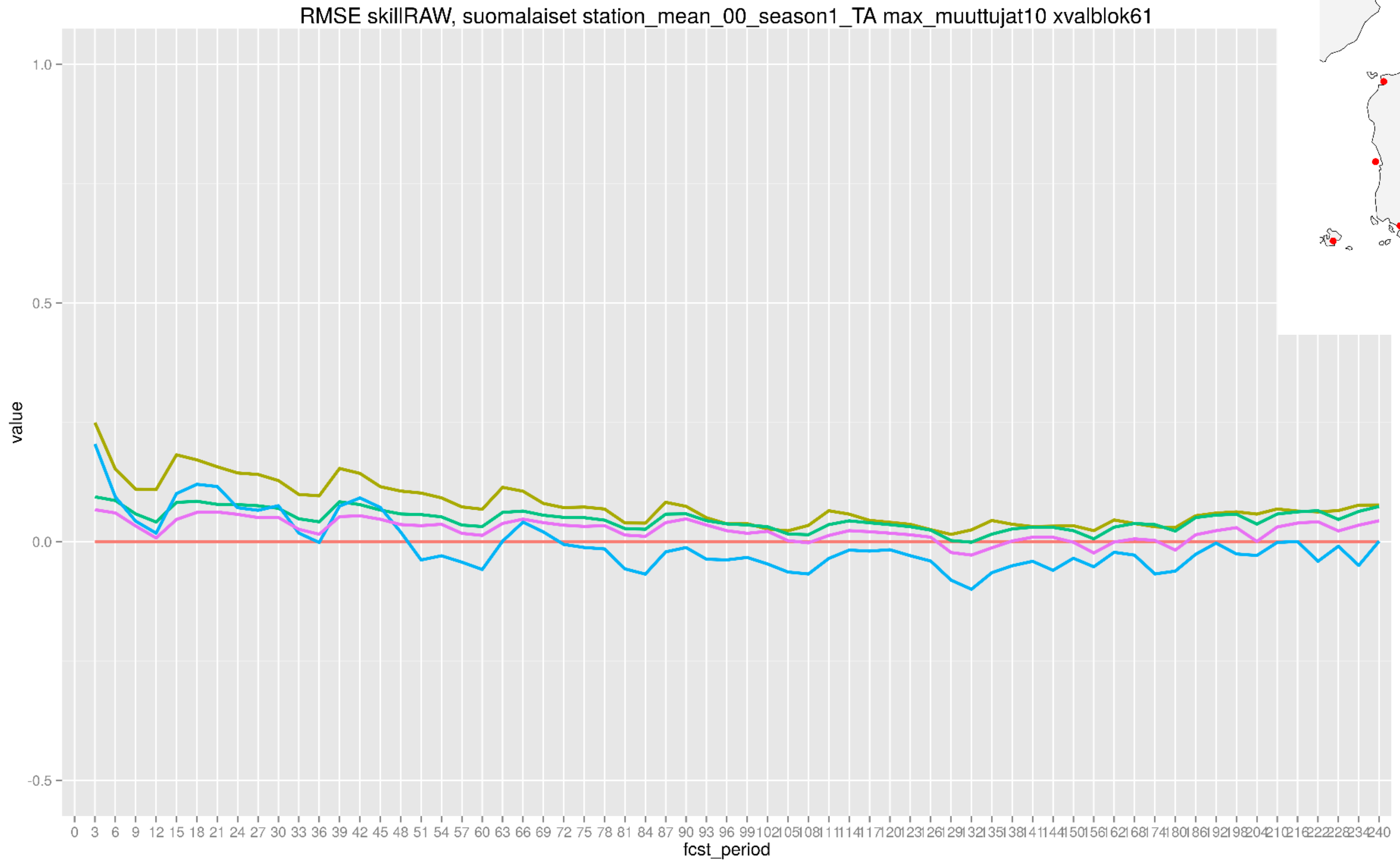
Deterministic MOS in a nutshell

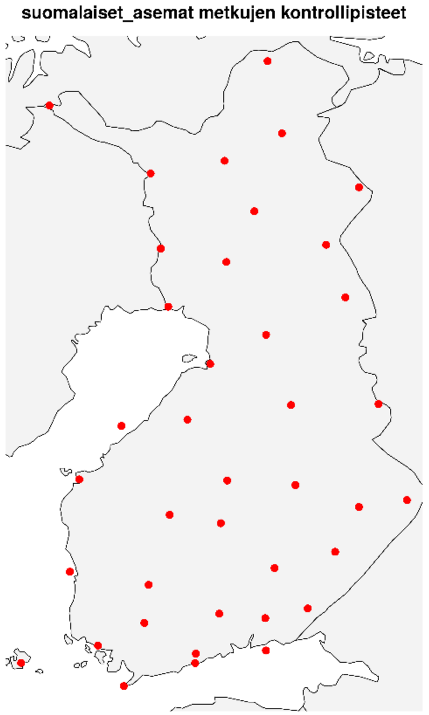
+ Makes model forecasts better
+ Identifies the predictor variables most important for forecasting
+ Can be trained for any observed variable
+ Objective method for removing systematic biases (in FMI, forecasters still manipulate gridded forecasts)
+ Downscaling from grid resolution to station point is genuinely done (both horizontal and vertical
+ Cheap to calculate operationally

- Depends on the quality+version of the NWP model and does not replace model development. Though the difference between DMO and calibrated forecasts is likely to persist in the future.
- Needs QC-controlled observation data from a sufficiently long training period. Preferably from the same geographical location.
- Consecutive model versions can have different statistical relations with observations. In principle, MOS needs new hindcast data from the new model version each time the model version is updated.
- Works “on average”, so the tails of the variable distribution might not be captured as well. Does not replace (and is not even intended to) in-depth process understanding in hazardous weather situations.
- Existing risk for catastrophically bad forecasts.
- Has to be frequently updated with new data.

FINNISH METEOROLOGICAL INSTITUTE

More data is better





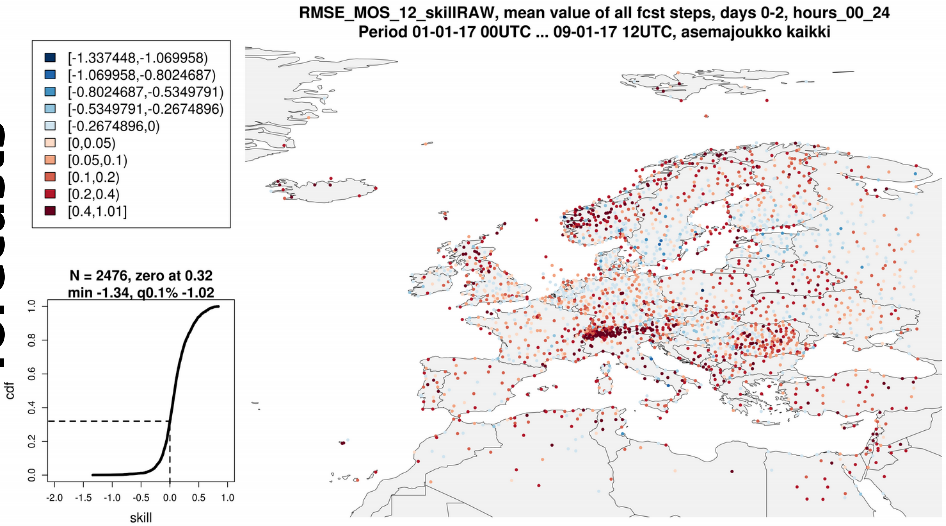
ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

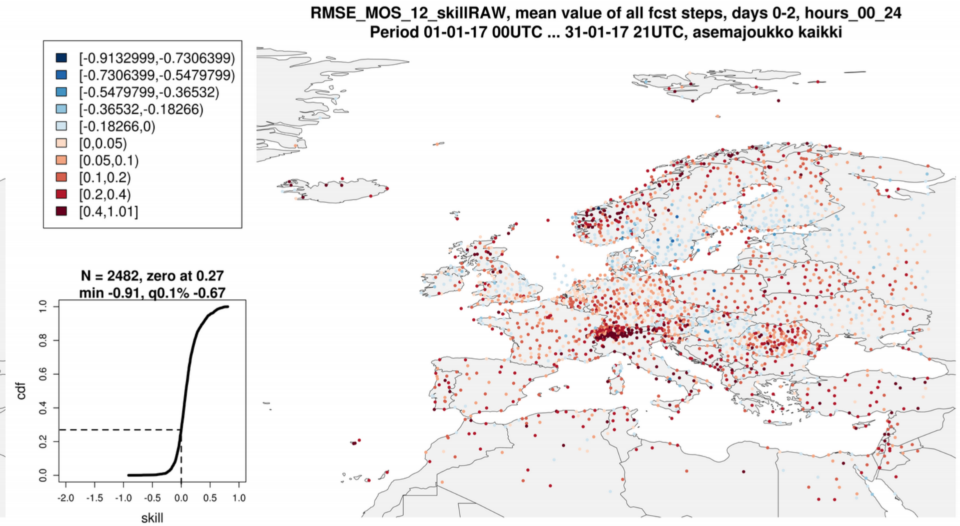
Jan2017 RMSE

1.1. - 9.1.

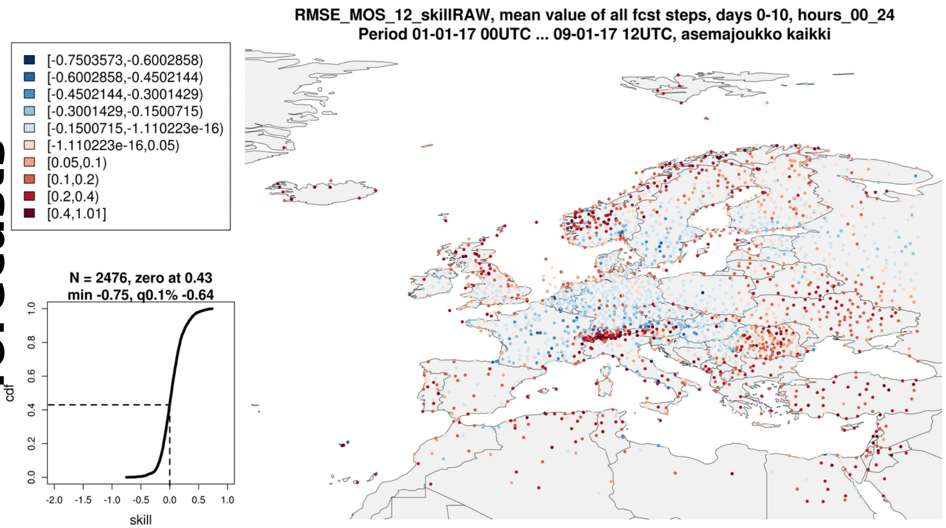
1.1. - 31.1.

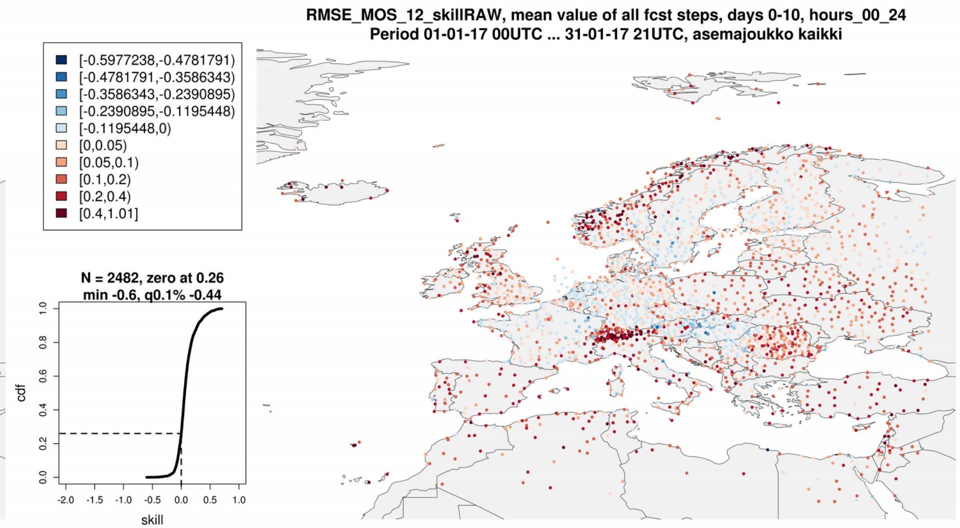
1-2 day forecasts





1-10 day forecasts





ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

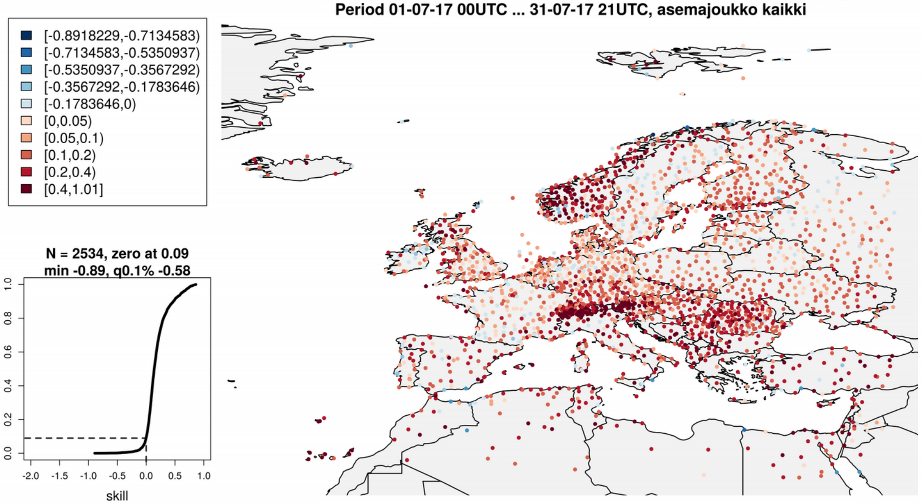
Jul2017 errors

RMSE

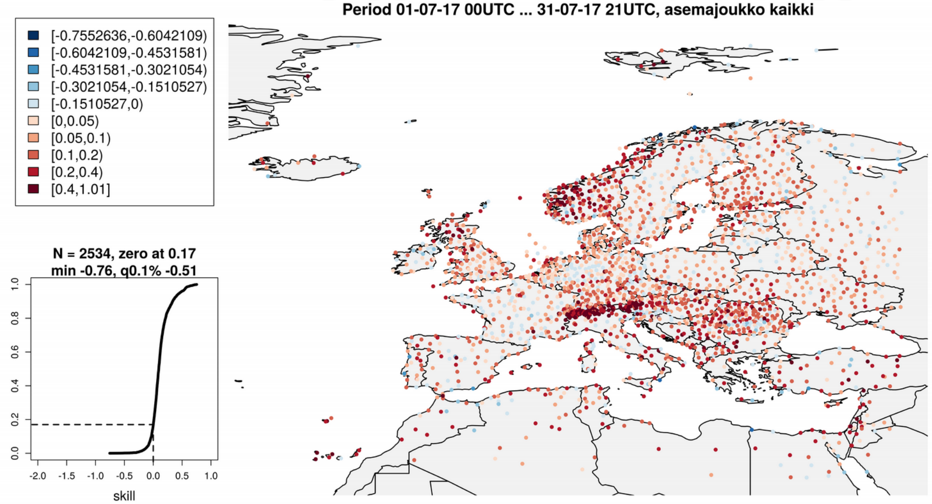
MAXE

1-2 day forecasts

RMSE_MOS_12_skillRAW, mean value of all fcst steps, days 0-2, hours_00_24
Period 01-07-17 00UTC ... 31-07-17 21UTC, asemajoukko kaikki

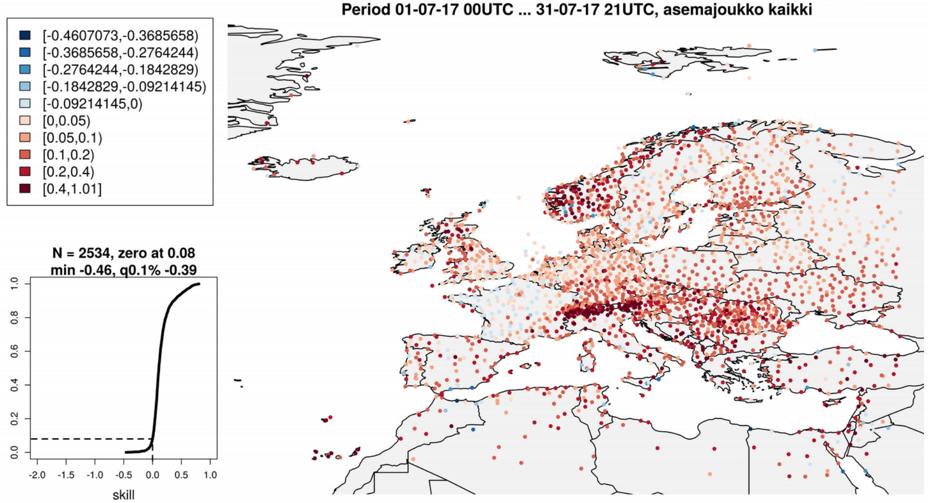


MAX_MOS_12_skillRAW, mean value of all fcst steps, days 0-2, hours_00_24
Period 01-07-17 00UTC ... 31-07-17 21UTC, asemajoukko kaikki

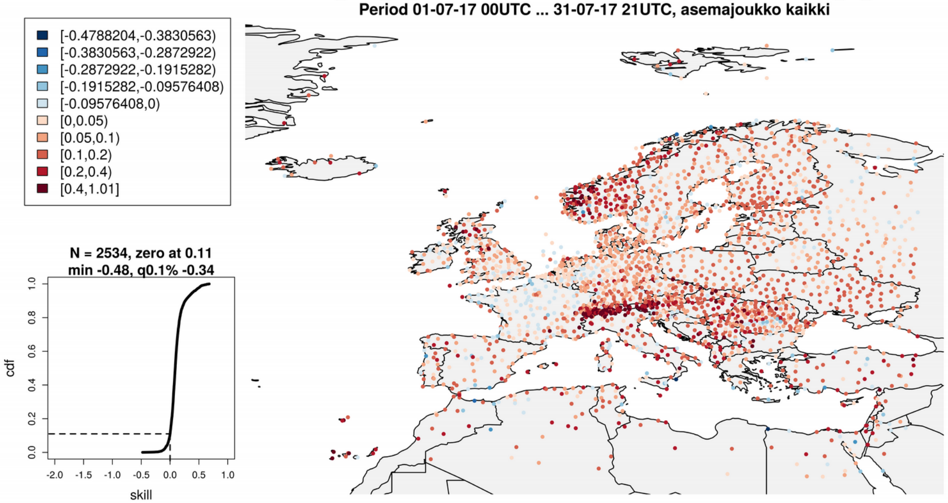


1-10 day forecasts

RMSE_MOS_12_skillRAW, mean value of all fcst steps, days 0-10, hours_00_24
Period 01-07-17 00UTC ... 31-07-17 21UTC, asemajoukko kaikki



MAX_MOS_12_skillRAW, mean value of all fcst steps, days 0-10, hours_00_24
Period 01-07-17 00UTC ... 31-07-17 21UTC, asemajoukko kaikki

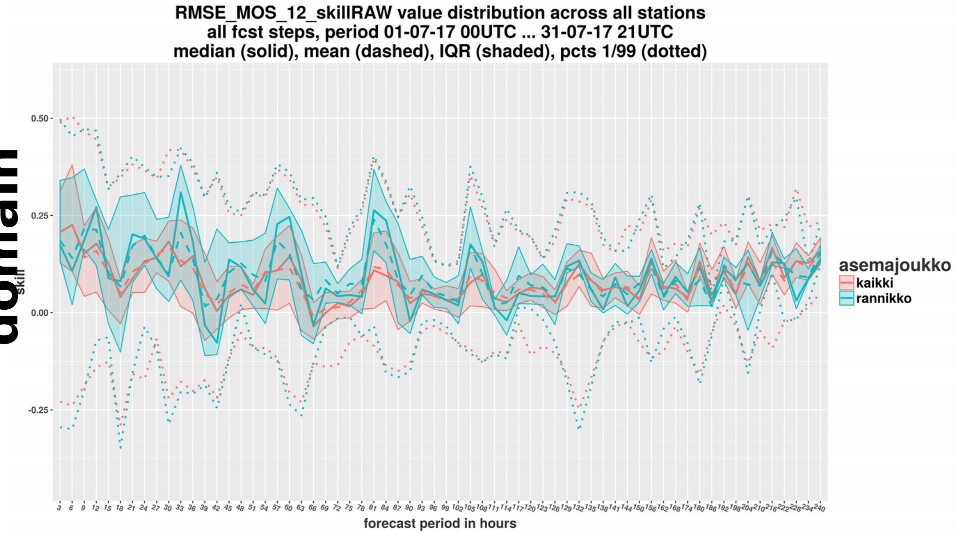


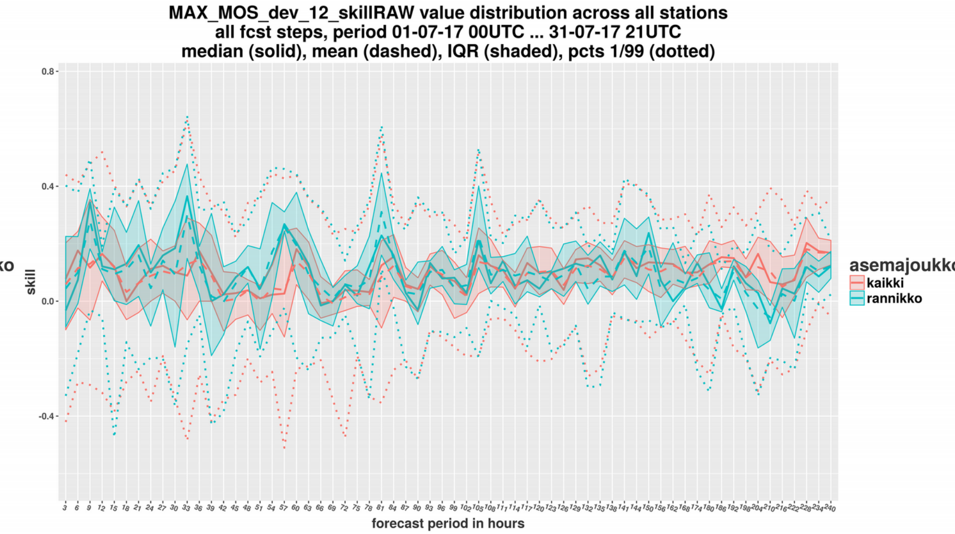
Jul2017 errors

RMSE

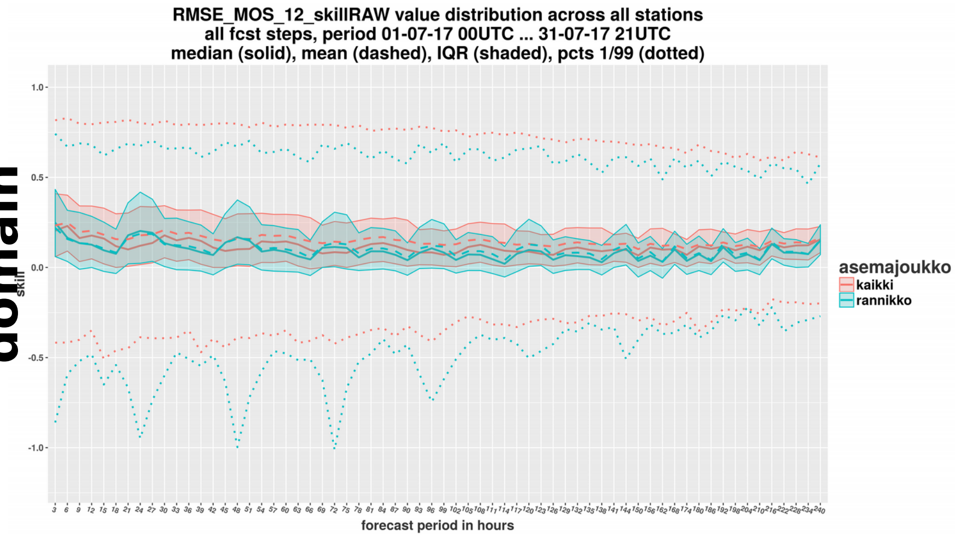
MAXE

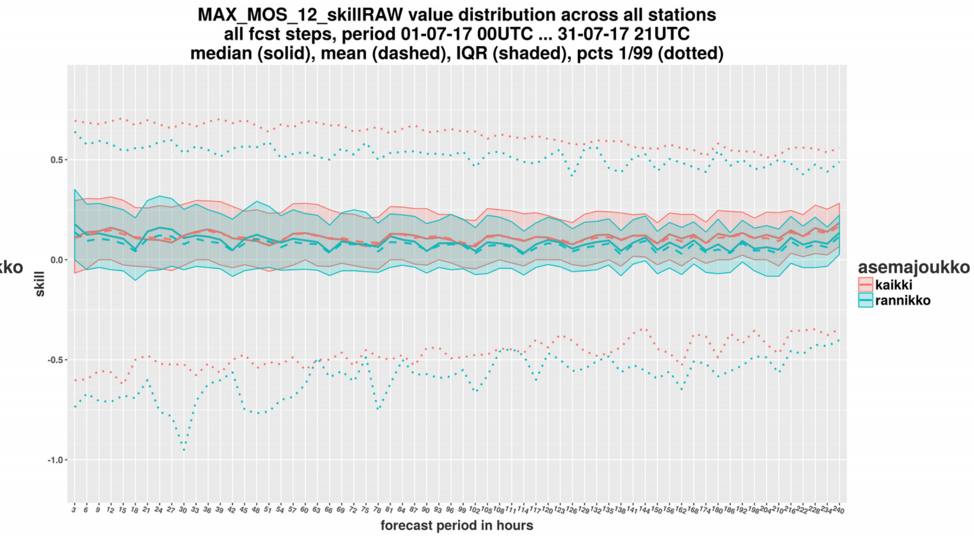
Finnish domain



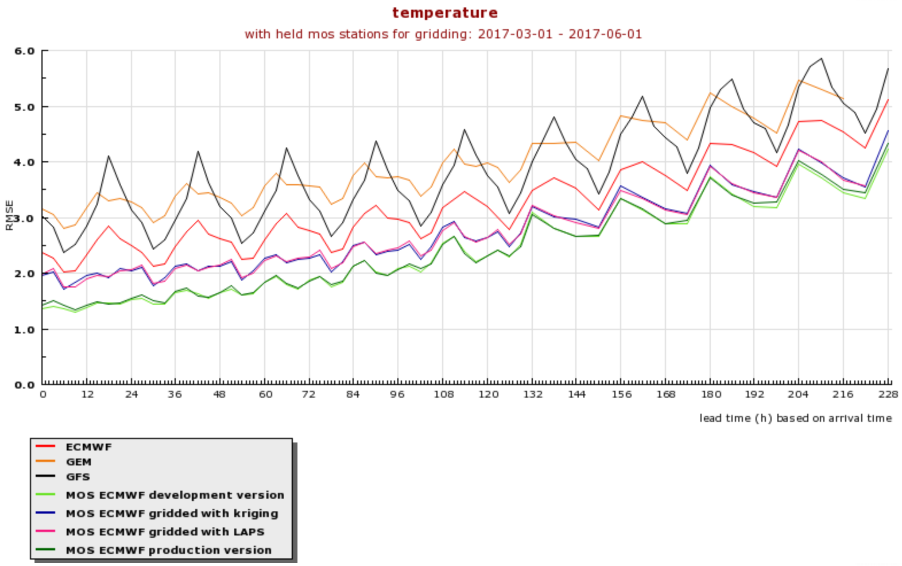


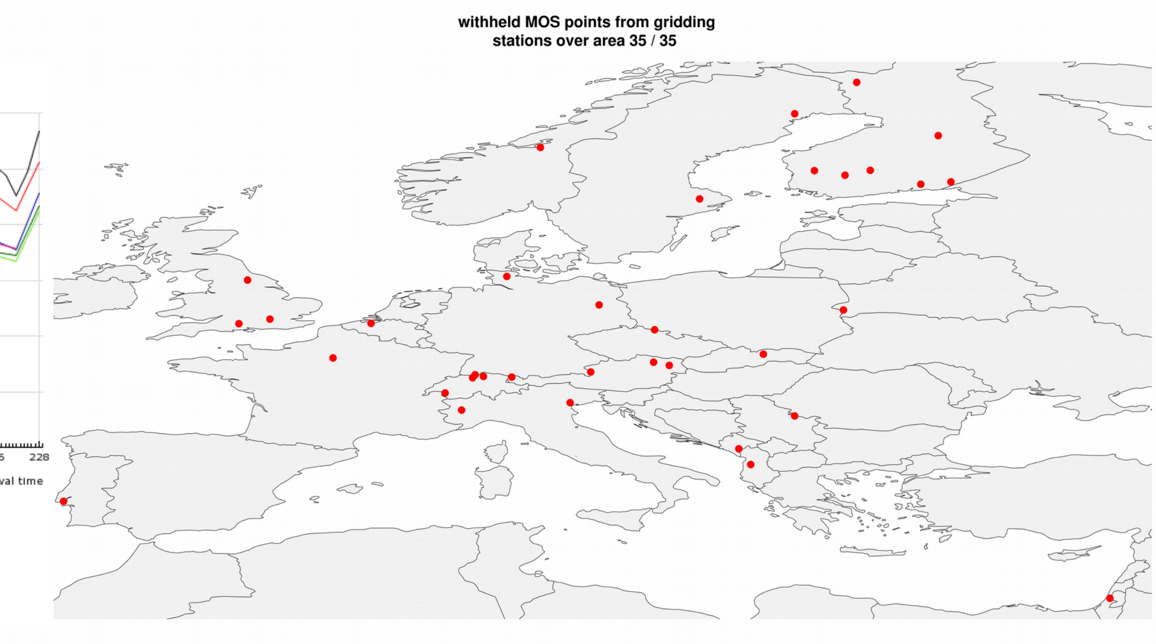
European domain





The effect of gridding





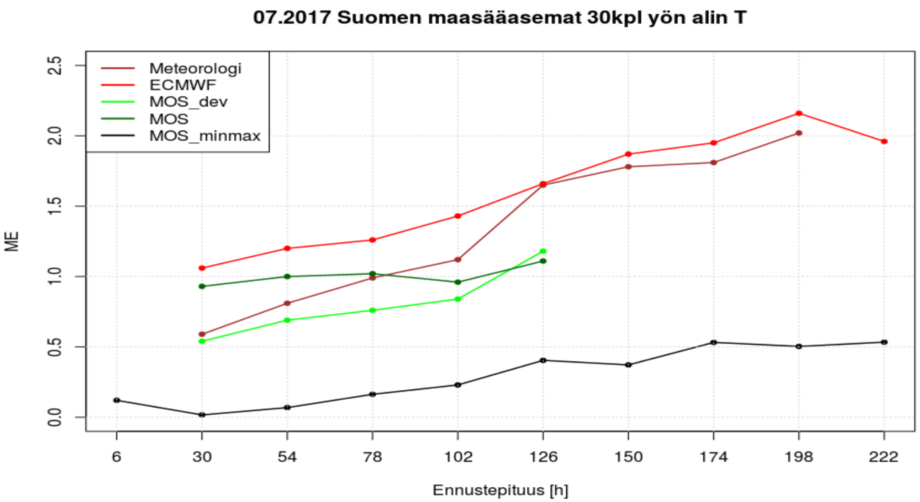
On the left figure there are MOS point forecasts trained specifically for the observation station and station-interpolated MOS forecasts from regridded MOS forecast fields where these specific stations are withheld from. Re-gridding the calibrated point forecasts induces a further source of error, but MOS point forecasts still clearly outperform the DMO.

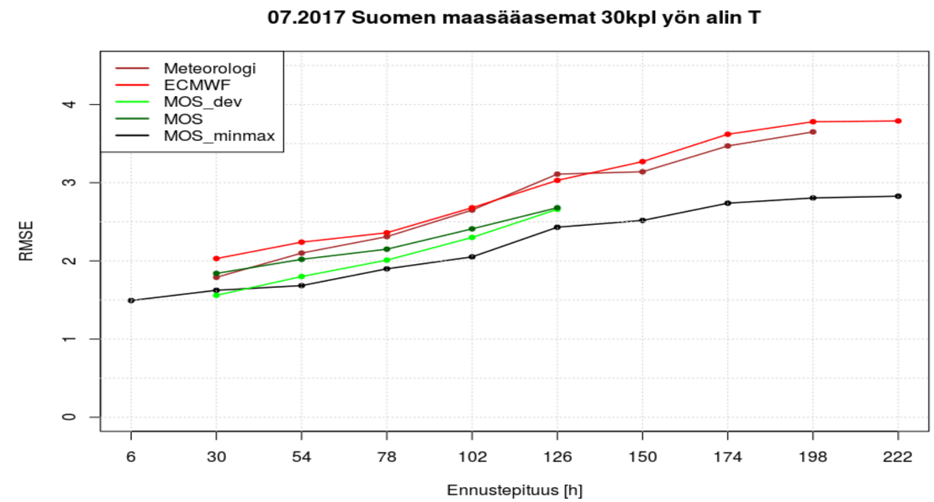
Jul2017 errors in Finland for daily extreme T values

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

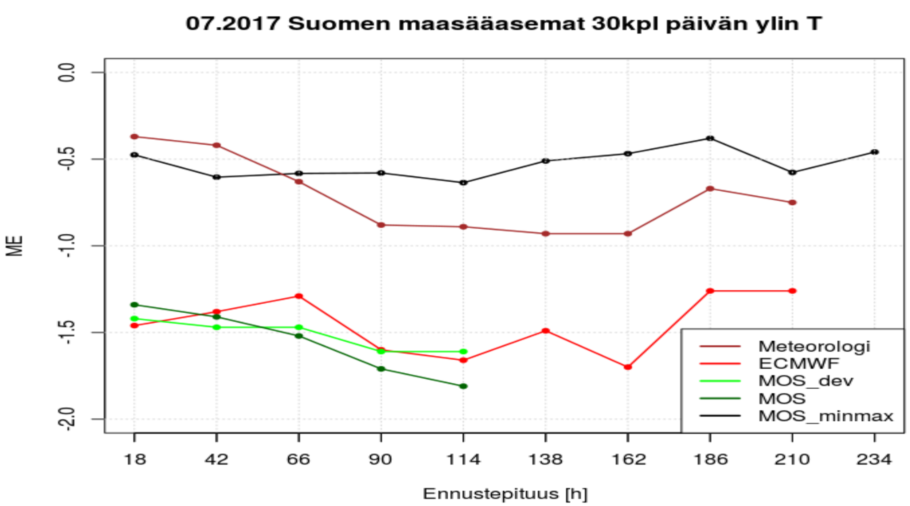
ME **RMSE**

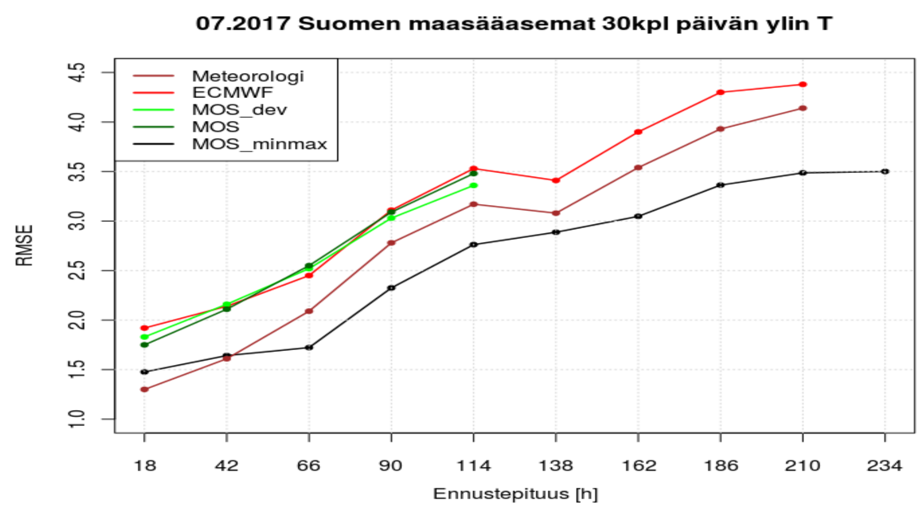
Tmin_day





Tmax_day





Positive experiences

- Foreign observations are now paid more attention to as previously they were not really used in operational forecasting (or even stored to databases)! Historical SYNOP observations have been re-collected from ECMWF archive and storing real-time data to internal databases has now been revised
- A statistical QC engine (time series method with GEV fitting) has been developed to foreign temp data
- Objective bias correction
- Education of forecasters and progress towards more objective bias correction
- Different gridding methodologies have been compared (long-used LAPS has also been used for gridding)
- Development of verification methods and operational verification system

In summary

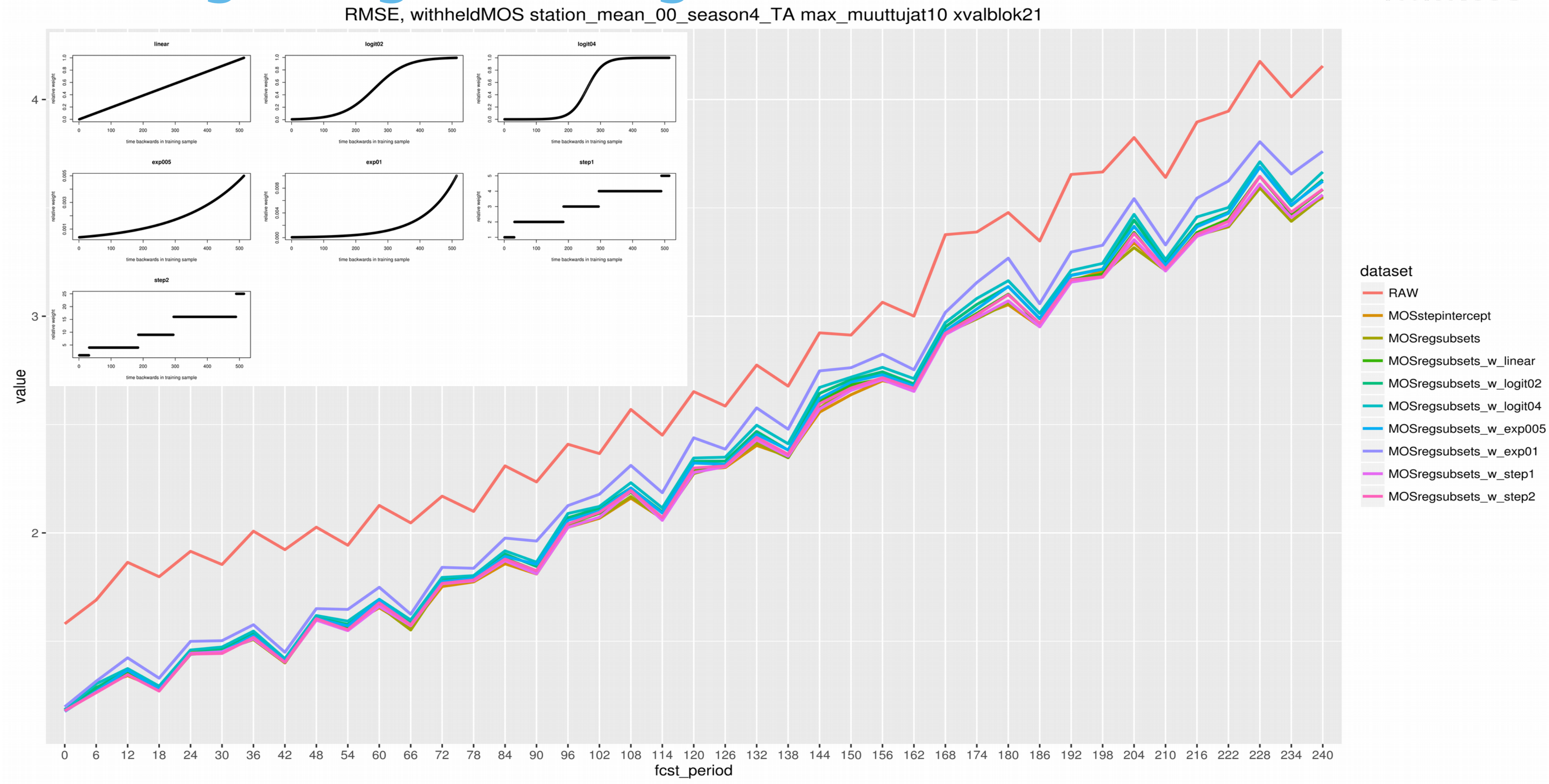
- MOS forecasts have a long production chain. Developing several components of it at the same have forced some pragmatic choices to be made (e.g. the choice of seasons). The long production chain also makes the production of really bad forecasts possible
  - Station location has changed (either in training period or after that)
  - DMO interpolated to station point is considerably off or erroneous
  - Statistical fit does not work on some part of the distribution
  - Model version changes have affected relationship (land/sea -mask changes or physical improvements in the model)
  - Overfitting to data
  - The weather phenomena is very rare and does not happen during training period
  - Grid-point-grid-point transfer has errors at some point
  - Other parts of production system have errors
- Based on experience most of these error sources do not constitute too bad problems, but those stations with problems are not straightforward to explain. Even the heterogeneous training data having several model version changes do not constitute a problem. More data with recent model version outweights the heterogeneity.
- Other variables would need to be calibrated soon. The obvious problem with using calibrated T2 forecasts is the inconsistency with other DMO variables.

FINNISH METEOROLOGICAL INSTITUTE

Thank you for your attention!

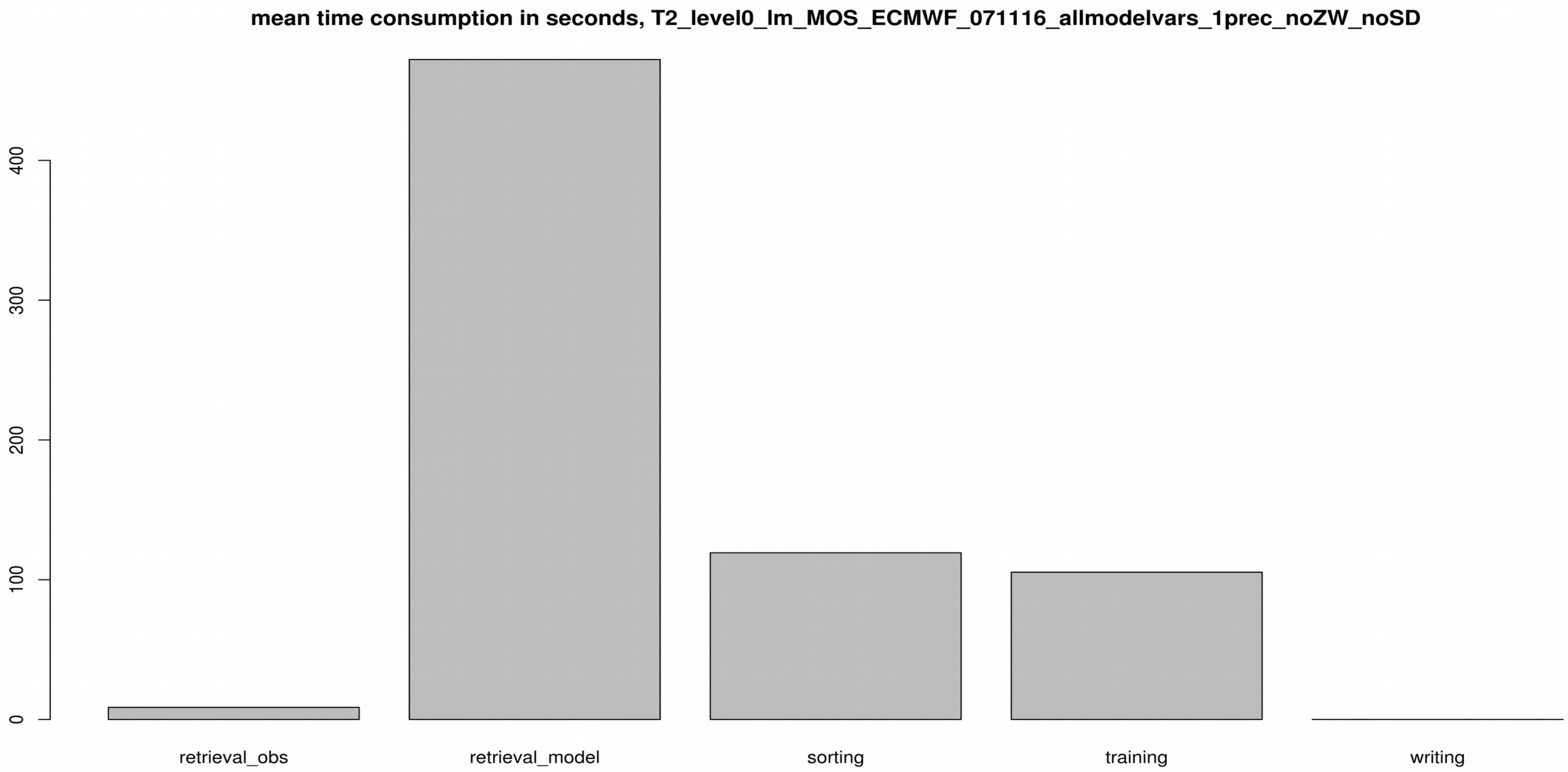
jussi.ylhaisi@fmi.fi

Weighting training data



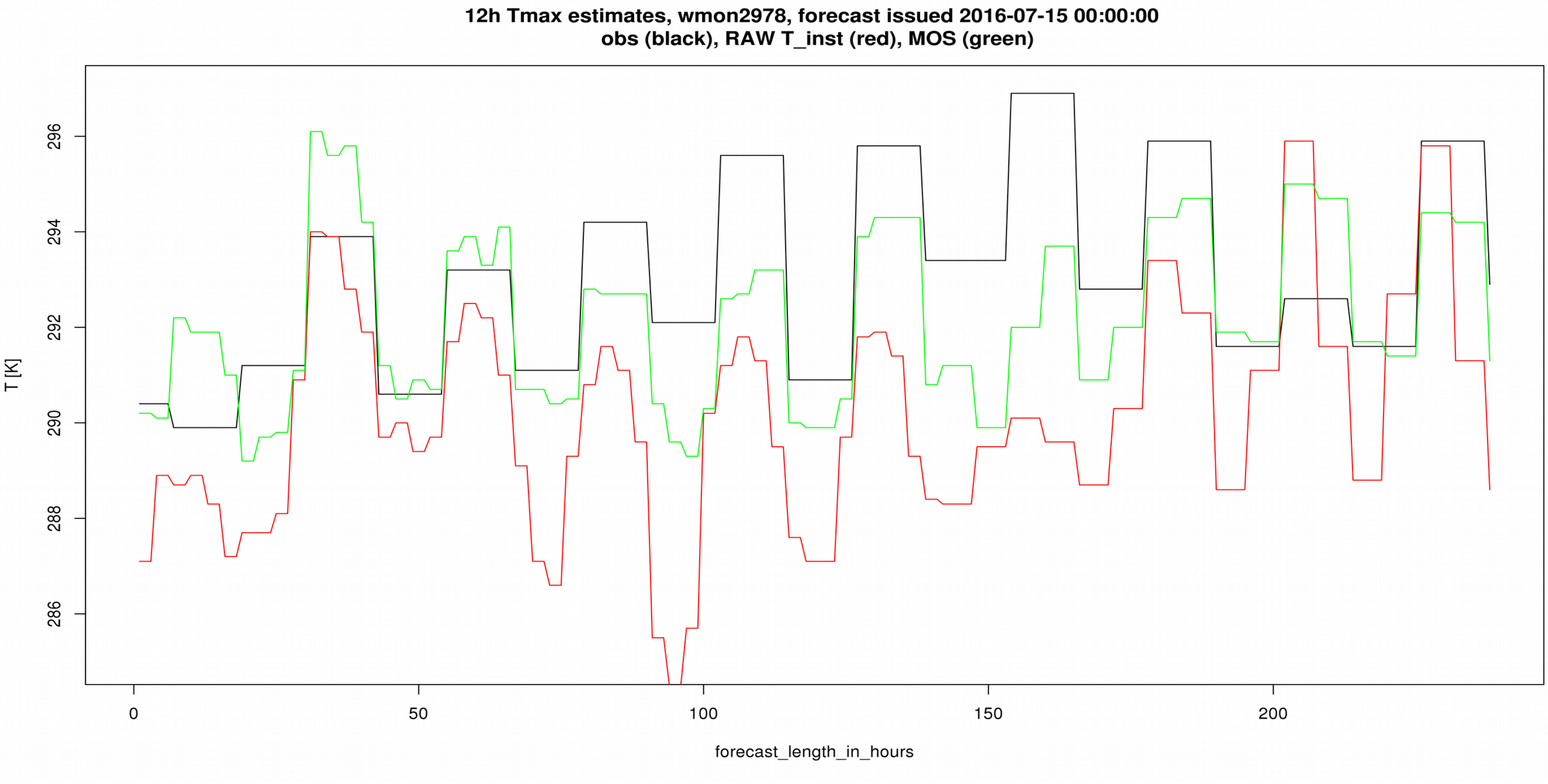
FINNISH METEOROLOGICAL INSTITUTE

Data retrievals can be painstakingly slow

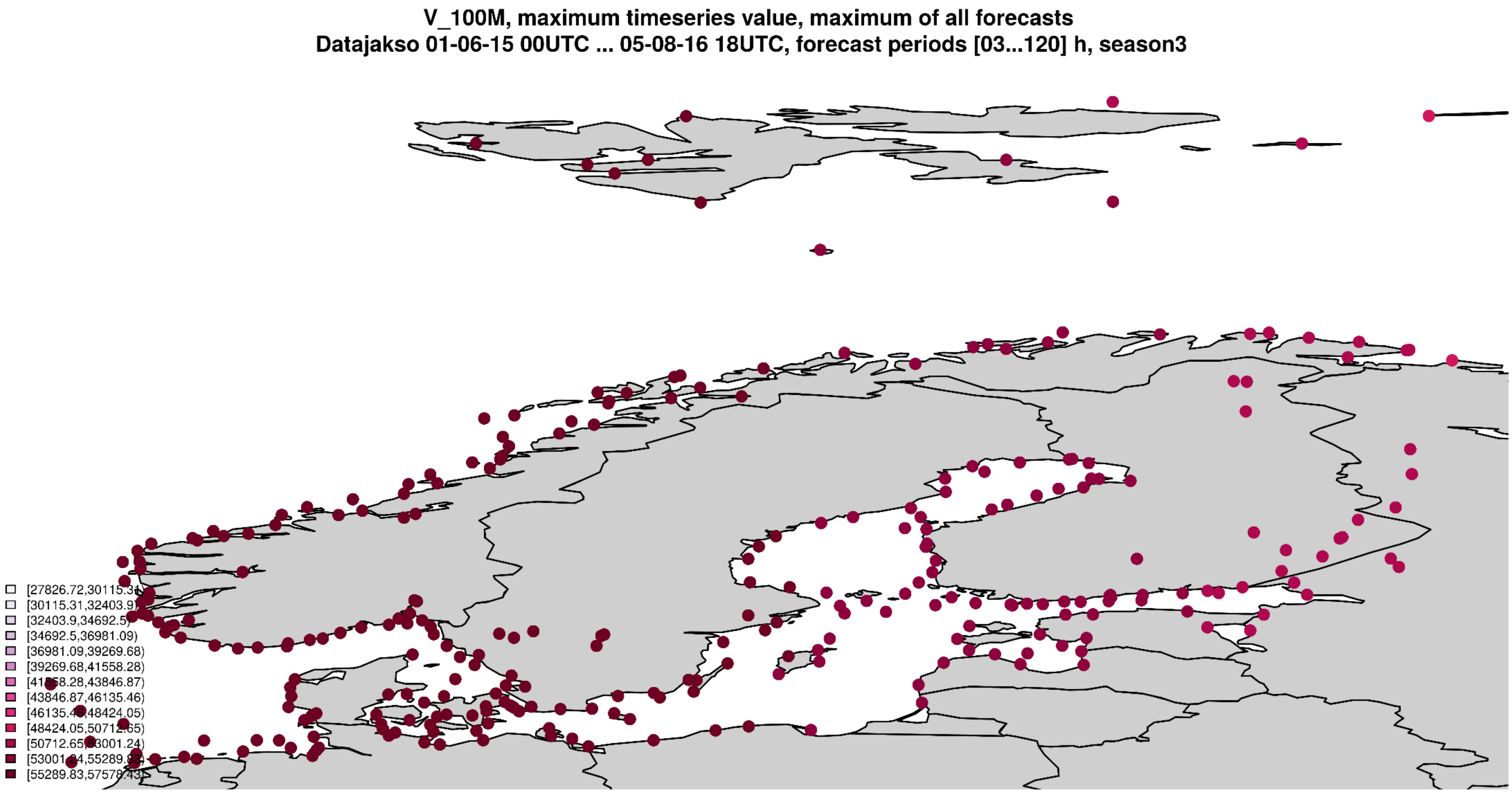


FINNISH METEOROLOGICAL INSTITUTE

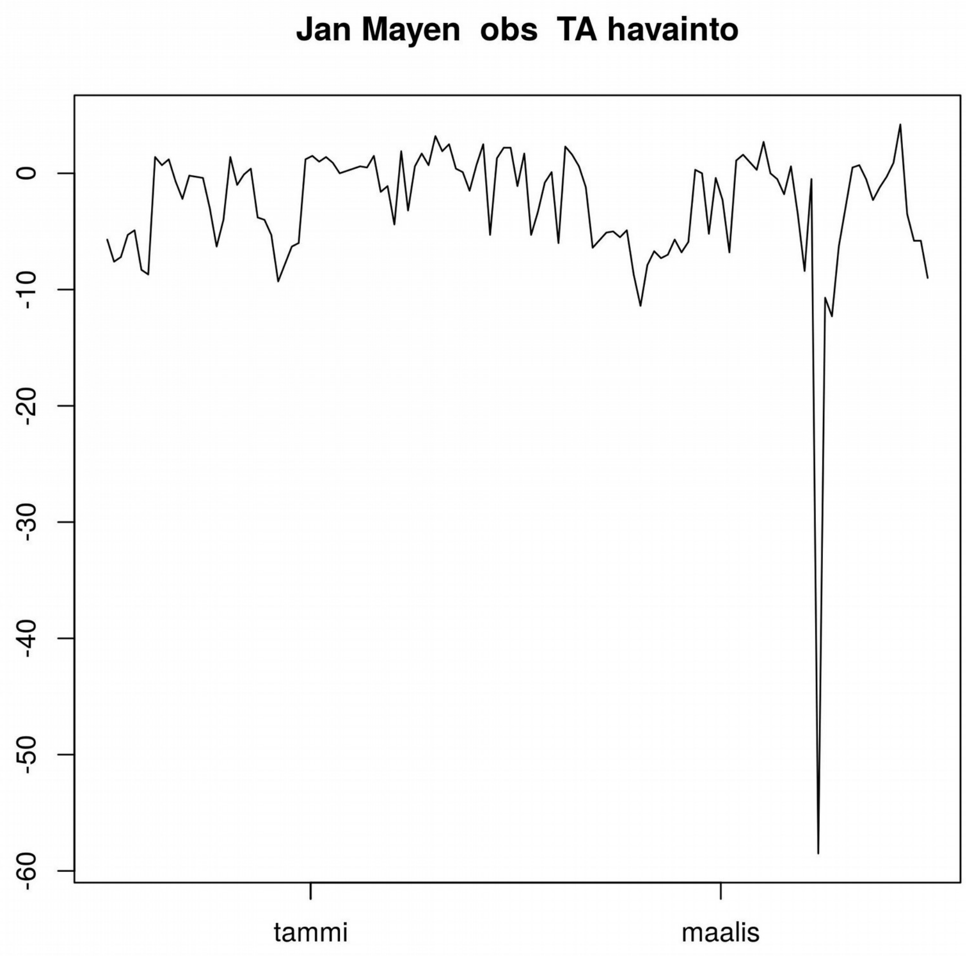
Tmax estimation



DMO in databases can also be considerably off the grid



Actual observations can also be erroneous



FINNISH METEOROLOGICAL INSTITUTE

Station location – grid-box elevation differences is key