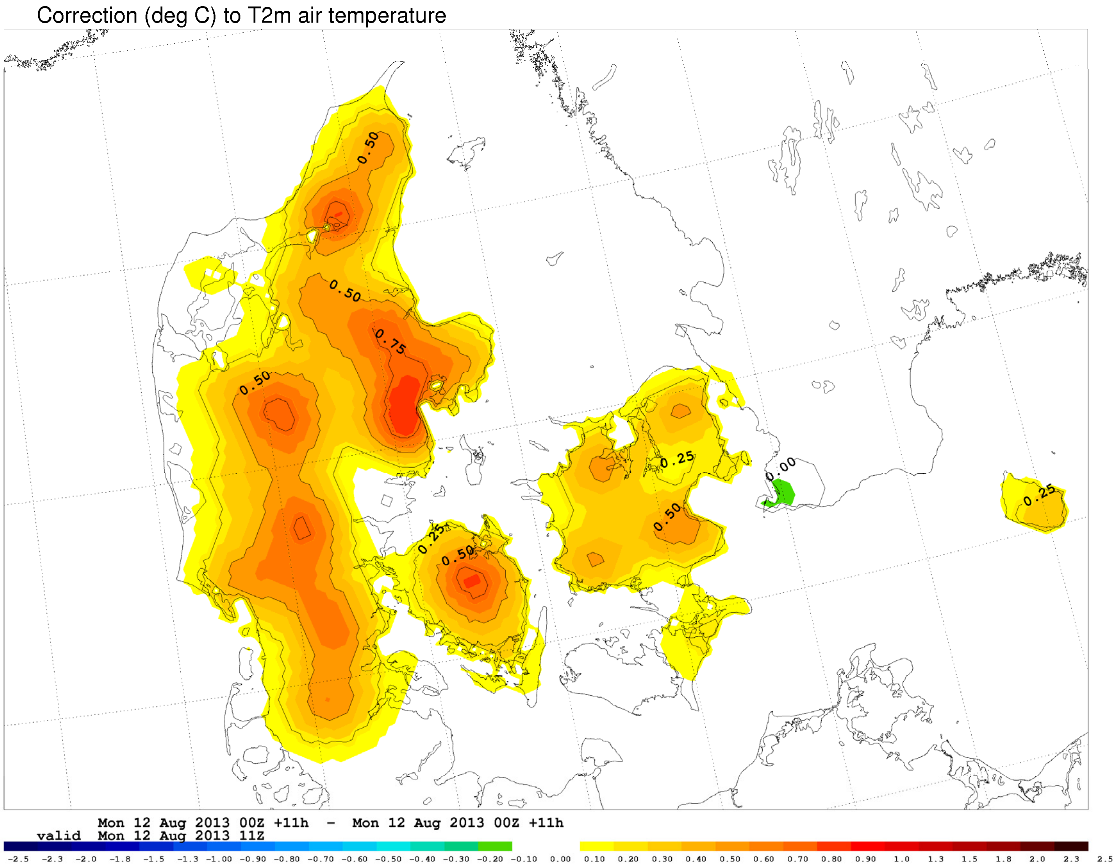
Scientific Report 13-02

Statistical Correction of Air Temperature Forecasts for City and Road Weather Applications

Alexander Mahura [1], Claus Petersen [1], Bent Hansen Sass [1], Nicolas Gilet [2]



1 Danish Meteorological Institute, DMI, Lyngbyvej 100, DK-2100, Copenhagen, Denmark
2 Institut National des Sciences Appliquées, INSA, 135 Avenue de Rangueil, 31400, Toulouse, France

Copenhagen 2013

Colophon

Serial title:
Scientific Report **13-02**

Title:
Statistical Correction of Air Temperature Forecasts for City and Road Weather Applications

Subtitle:

Author(s):
Alexander Mahura, Claus Petersen, Bent Sass, Nicolas Gilet

Other contributors:

Responsible institution:
Danish Meteorological Institute (DMI)

Language:
English

Keywords:
Numerical Weather Prediction (NWP), HIRLAM, synop stations, air temperature, statistical correction, forecast

Url:
www.dmi.dk/dmi/sr13-02.pdf

Digital ISBN:
978-87-7478-645-0 (on-line)

ISSN:
1399-1949 (on-line)

Version:

Website:
www.dmi.dk

Content:

Abstract

The method for statistical correction of the air/ road surface temperatures forecasts was developed based on analysis of long-term time-series of meteorological observations and forecasts (from HIRLAM-SKA/ RCM models outputs) and tested for May-Aug 2012/ Oct 2012 - Mar 2013 periods. For the city weather applications, new operationalized procedure for statistical correction of the air temperature forecasts has been elaborated and implemented for the HIRLAM-SKA model runs at 00, 06, 12, and 18 UTCs covering forecast lengths up to 48 hours. It includes segments for extraction of observations and forecast data, assigning these to forecast lengths, statistical correction of temperature, one- and multi-days evaluation, decision-making, interpolation, visualisation and storage /backup. Pre-operational air temperature correction runs were performed for the mainland Denmark (for 60 synoptical stations) since mid-April 2013 and have shown good results. The method performed well for domain of mainland Denmark (tested for 60 synoptical and 395 road stations), and hence, it can be also used for any geographical point within this domain, as well as we can assume that the same method can be applied in other geographical areas.

Resumé

En metode til statistisk korrektion af fejlen i luft- og overfladetemperatur er blevet udviklet. Metoden er baseret på analyser af lange tidsserier af observerede temperatur og forudsagt temperatur fra HIRLAM-SKA og HIRLAM-R03. Metoden er efterfølgende blevet testet i perioderne maj til august 2012 og oktober 2012 til marts 2013. Metoden er primært udviklet med henblik på at forbedre 2 meter temperaturen i DMI's operationelle produkter såsom byvejr. Den forudsagte modeltemperatur fra alle HIRLAM-SKA's kørsler startende henholdsvis 00, 06, 12 og 18 UTC korrigeres for alle prognoselængder dvs. 0 til 48 timers prognoser. Den statistiske korrektionsmetode består af flere elementer der bl.a. inkluderer udtrækning af observation og prognoser for 60 punkter i Danmark for den sidste måned. Baseret på den sidste måneds statistik beregnes en temperaturkorrektion for de 60 punkter, der efterfølgende interpoleres tilbage til modellens grid. Systemet har været i drift siden April 2013 og har vist gode resultater.

1. Introduction

The temperature is one of the important parameters predicted by the Numerical Weather Prediction (NWP) models and it is daily and world-wide used in different practical applications. For example, for the city weather application (CWA) the air temperature is a key importance, although the dew point temperature and temperature of the underlying surface are also important. Air temperature controls and affects nearly all other weather parameters. It is needed for customers to know how warm or cold it is outside, reproduction of plants and animals, etc. - simply, it is influencing our daily life. For the road weather application (RWA) the road surface temperature (or temperature of the asphalt surface) is a parameter of the key importance, although the air temperature and dew point temperature also characterize the surrounding of the road weather conditions. Based on the road surface temperature the maintenance of the roads and traffic conditions is taking place, deciding where and when it is necessary to make salting to reduce dangerous conditions at driving, optimising the vehicles speed and traffic, environmental impact on surroundings, etc.

It is a very rare situation when the NWP model is almost perfectly predicting any selected meteorological parameter observed in the nature. Perfectly – by meaning that there will be a relatively small error compared with observations. The occurred errors might be of different nature. In particular, errors could be linked with observations itself; errors might occur at initialisation steps of the model during assimilation of data; there are also errors connected with limitations of the model itself; errors could be generated due to numerical methods applied for equations solution; as well as they can be due to afterwards interpolation to grid/points in the modelling domain. A sum of all these errors (and probably also adding an unknown error – e.g. not yet identified) will be the overall error. This overall error as a difference between the forecasted and observed values (so-called, BIAS) is always under continuous attempt to be minimised (with a goal to approach almost 0 value). The other parameter – mean absolute error (MAE) - is also important measure and should be used jointly with BIAS to characterise quality of forecast.
Note, in reality, applying statistical post-processing of the model output we can try to correct only the systematic errors due to physics of the model or parameterisations that are used to represent physical processes. The other errors are preferably to be known in advance. For example, for temperature it can be done by introducing an uncertainty parameter (let's assume 0.25°C in our study) covering observation, assimilation, numeric, and interpolation errors (with some unknown error).

For the recent decades many approaches were suggested for improvement of the temperature forecasts provided by NWP. The operationalized procedures are mostly applied post-factum as post-processing, e.g. after the temperature forecast was originally produced it is corrected. These approaches (for example, Kalman filtering) mostly based on statistical backgrounds and are applied by different meteorological weather services worldwide, for improvement first of all, forecasted parameters within the responsibility zone (e.g., for example, for national weather services the national boundaries are the priority number one).

The **main goal** of this study is to develop, test, evaluate and operationalize method for statistical correction of temperature for the city and road weather applications (for air and road surface temperatures, respectively) with a focus, first of all, on Denmark mainland with a possibility to expand the same method to other territories as well. The **specific objectives** of this study are the following:

- Analyse available long-term time-series of observations and forecasts of various meteorological parameters to evaluate types of distributions, existence of periodicities, and possible correlations between parameters.
- Develop and test method for statistical temperature correction for different geographical loca-

tions (e.g. synoptical and road weather stations), forecast lengths, hours/days/months of the year; evaluate method performance for the city (using 1 May – 31 Aug 2012 data) and road (using 1 Oct 2012 – 15 Mar 2013 data) weather applications; and analyse possibilities for further improvements of the temperature corrections.

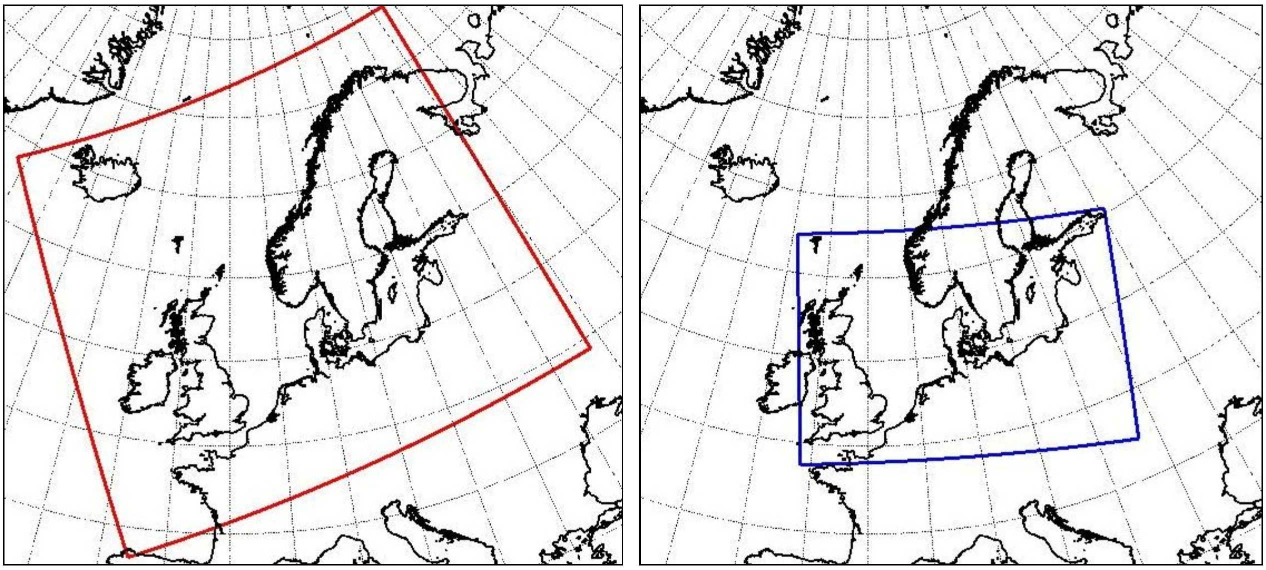
- Operationalize the developed method (coding/scripting, setup procedures for extraction of observations and forecasts, assigning to forecast lengths, correction, decision-making, interpolation, visualisation and backup) for practical use with pre-operational tests and verification (based on 20 Apr – 20 Sep 2013 data), at first, in the city weather applications (***byvejr forecasts*** – http://www.dmi.dk/vejr/til-lands/byvejr*)* for improvement of air temperature forecasts.

2. Methodology

2.1. Operational Forecasting

2.1.1. Numerical weather prediction

The HIgh Resolution Limited Area Model (**HIRLAM**; http://hirlam.org; *Uden et al., 2002*) is used as the Numerical Weather Prediction (NWP) operational modelling system at DMI. The HIRLAM output is used for different practical applications such as providing weather conditions forecasts for geographical regions, countries, cities, roads, etc. as well as for emergency response purposes, air pollution including pollen, agriculture and transportation's needs, etc. The current operational HIRLAM forecasting system, run at DMI, includes the digital filtering initialization, semi-Lagrangian advection scheme, and a set of physical parameterizations such as Savijaervi radiation, STRACO condensation, CBR turbulence scheme, and ISBA scheme for land-air interactions.



(a) *(b)*

Figure 2.1.1*: Boundaries of the modeling domains for the city (HIRLAM-SKA model) and road (HIRLAM-R03 model) weather applications.*

The NWP HIRLAM-SKA model domain covers (see Figure 2.1.1a) Denmark, north and central Europe with a resolution of approximately 3 km. This NWP model runs 4 times per day at 00, 06, 12, and 18 UTCs with a maximum forecast length up to 54 hours ahead, and a forecast output is started to be provided for about 3 hours later. There are 970 x 818 grid points in domain, 65 vertical levels, and time step is 90 seconds. The boundary fields are taken from ECMWF. It is run on CRAY-XT5 supercomputer, and produced model output files are archived on the mass storage system of DMI. Output of the HIRLAM-SKA model is used for the city weather application and other.

Although forecasts are produced for every hour up to a maximum forecast length of 54 hours. But in our study it was considered only: 1) every 6 hour interval up to 24 hours (e.g. +6, +12, +18, +24 hours forecast length; for statistical analysis of time-series to identify relationships between observed/forecasted meteorological parameters), and 2) every 1 hour interval up to 48 hours (for operationalization of the developed method for statistical correction of the air temperature at 2 m forecasts). The list of forecasted meteorological direct and derived parameters is large (to about of almost 200).

In our study, **for time-series statistical analysis,** for each of synoptical stations for each time (year, month, day, hour) and forecast length corresponding records the following 20 forecasted parameters were extracted: mean sea level pressure, wind speed and direction at 10m, air and dew temperature at 2m, total cloud cover, surface temperature, boundary layer height, absolute humidity, surface pressure, U&V-momentum flux at surface, at 1st-SKA-model-level (=12m) - air temperature and U&V-components of wind speed, global/long-wave/ diffuse radiations, latent and sensible heat fluxes (with additional data (9 parameters) on station latitude/longitude coordinates, height asl, land-sea mask, fractions of tiles – water, ice, bare soil, low vegetation and forest). For the analysed period (1 May – 31 Aug 2012) there are in total 118080 records (123 days x 4 UTC runs x 4 forecast lengths x 60 stations) corresponding to forecasts.

Note that **for operational city weather applications** forecasts only for 7 following meteorological parameters (as result of findings from statistical analysis) were used in development of a method for statistical correction of the air temperature at 2 m (*T2m*) forecast. These parameters are the air temperature (*T2m*) and dew point temperature (*Td2m*) at 2 m (in °C), wind speed (*V10m*, in m/s) and direction (*D10m*, in degrees, 0-360) at 10 m, cloud cover (*ClCov*, in %, 0-100), temperature of the surface (*Ts*, in °C), and boundary layer height (*PBLH*, in meters).

2.1.2. Road weather modelling

The Road Conditions Model (**RCM;** *Sass 1992, 1997*) is a so-called energy balance model originally developed at DMI and it is a part of the DMI's Road Weather Modelling (RWM) operational system. It is a local model that the forecast for each point does not depend on other points. Note that all advection parts of atmospheric processes are modelled by the NWP HIRLAM model (see Ch. 2.1.1) version -R03 and are available at each time step. The area for this NWP model was chosen in such a way that Denmark is in the centre of the domain (see Figure 2.1.1b). The size of R03-domain (650 x 460 points) was chosen to ensure weather conditions likely to influence Danish area during the forecasted period. The horizontal resolution is approximately 3 km, with 40 vertical levels and a time step of 120 seconds.

The energy balance model used by the RCM model differs from the corresponding energy balance model used by the NWP model. In particular, the RCM has more layers (in total 15) in the road and below, and the heat equations are solved for these layers, whereas the NWP model uses a simple 2 layer model. Additionally, the RCM model has been optimized for asphalt (dominating on roads) surfaces, which have characteristic much different from other types of natural surfaces such the bare soils, grass, sand, etc.

Operationally, the RCM model is running 24 times per day (starting at 00, 01, …, 23 UTCs) and provides forecasts up to 24 hour forecast length (with an output every 0.5 hour or every 30 minutes). At first, output covering only the first 5 hours of forecasts (e.g. +0.5, +1.0, +1.5, …, +4.5, +5.0 hours forecast lengths) is delivered. Then about 20 minutes later all forecasts up to 24 hour ahead are delivered. This model is running as a module inside the NWP model, and it is receiving input at every time step from NWP. The RCM generates forecasts for geographical positions of road stations (and since 2008, also for positions of road stretches) and input from the NWP is interpolated to these points.

At the first time steps, an analysis of the initial conditions (ICs) is done. Available observations from the road stations of the road surface temperature (*Trs*), air temperature (*T2m*) and dew point temperature (*Td2m*) at 2 m are used to run the RCM. The equations of heat conduction for the road points are solved using the forecast from the last model run as ICs, and then run the RCM 3 hour ahead with observed *Trs* as boundary condition. This provides an analysis of the temperature profile of all the road layers. Forecasts are calculated at each time step. In these calculations the following processes such as short-wave heating and long-wave cooling/heating of the road surface; turbulence fluxes of heat and moisture from/to the surface; and evaporation, melting, freezing and sublimation of water and ice from the road surface - are considered. In particular, both short- and long-wave radiation are affected by shadows (obstacles can shadow/block for incoming solar radiation /for outgoing long-wave radiation). The corresponding forecasts for wind speed and wind velocity, cloud cover, and cloud base (*ClBase*; in meters) are interpolated to the road stations positions from the HIRLAM-R03 gridded output. Output of the HIRLAM-R03+RCM model is used for the road weather application (e.g. ***vejvejr forecasts*** at http://vejvejr.dk/vejvejr; access password is required).

In our study, **for time-series statistical analysis,** forecasts of the *Trs, T2m, Td2m*, accumulated water (rain water, dew) and ice (snow, rime, frozen water) on the road surface, *V10m*, *D10m*, *ClCov,* and *ClBase* were used. For the analysed period (1 Oct 2012 – 15 Mar 2013) there are in total 7868400 records (166 days x 24 UTC runs x 5 forecast lengths x 395 stations) corresponding to forecasts (note that only forecasts for exact hours, not at every 30 minutes intervals, were used).

For **operational road** (similar to the city) **weather applications** only forecasts for 7 meteorological parameters - *T2m, Td2m, V10m*, *D10m*, *ClCov, Trs, ClBase* (note, instead of the surface temperature the road surface temperature is used as well as instead of the boundary layer height the cloud base is used) were used in development of a method for statistical correction of the road surface temperature (*Trs*) forecast.

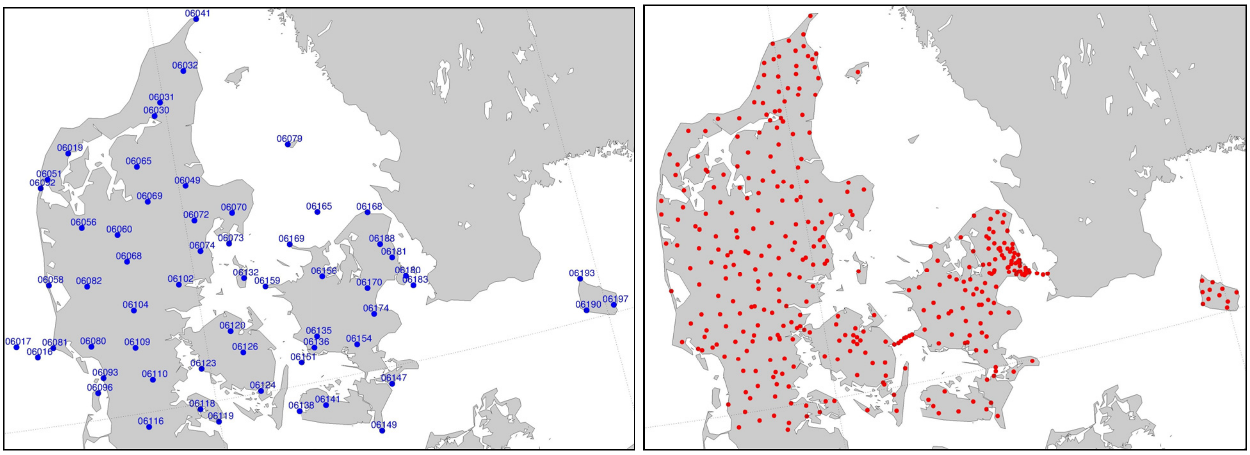
2.1.3. Measurement synoptical and road stations

Synoptical stations data

There are more than 60 synoptical stations situated in Denmark as a part of the World Meteorological Organization (WMO) weather observational network (Figure 2.1.2a). The measurements at these stations are carried out continuously during the day and provided to the national weather service at DMI through speedy network. The synoptical stations are placed in different surroundings, and can be classified as inland/coastal/ sea, urban/sub-urban/rural, etc. types.
After initial data quality control these became available for assimilation into NWP model and further verification purposes. The list of meteorological parameters includes measurements for the air and dew point temperatures and relative humidity at 2 m, wind speed and direction at 10 m, temperature of the surface and at different depths, radiation related parameters, cloud cover-type-height, etc.

In our study, **for time-series statistical analysis,** for each of synoptical stations for each time (year, month, day, hour) corresponding records the following 7 observed meteorological parameters were extracted: surface pressure, wind speed and direction, visibility, total cloud cover, air and dew temperatures. These were taken at 6 hour intervals (at 00, 06, 12 and 18 UTCs) in correspondence with forecasted meteorological parameters taken for time-series analysis. If there are no missing observations in the analysed period (1 May – 31 Aug 2012) there could be in total 29520 records (123 days x 4 UTCs x 60 stations) corresponding to observations.

For **operational city weather applications** only the 4 following parameters - air and dew point temperatures at 2 m, wind speed and direction at 10 m - were used; with a frequency of 1 hour interval (at 00, 01, …, 23 UTCs).



(a) (b)

Figure 2.1.2: *Geographical distribution of the Danish (a) synoptical stations (position and ID number within the WMO observational network) and (b) road weather stations (positions according to the Danish road stations network operated by DRD).*

Road stations data

The Danish road network is represented by a large number of roads/driving lanes in various communes. In total, there are more than 500 sensors (measuring the road surface temperature) placed at about 400 road stations (where measurements of meteorological parameters are carried out) as shown in Figure 2.1.2b. At some stations there are two or more sensors placed along the driving lanes or on the opposite side of the road not far from each other. Although road stations are not equally spatially distributed within the network, but these are placed along most of the roads to cover as much as possible (roughly estimated: for a length of approximately 10 km there is one road station). Although practically all road stations are placed near the roads or driving lanes the surroundings could be different, and hence, stations could be divided into groups (see *Mahura et al., 2009, 2010*), e.g. where the stations are placed in airports or on bridges, in urban/sub-urban/rural locations, or surrounded by forest, fields, water, and individual trees. At each road station there are continuous measurements of meteorological parameters such as road surface (with sensors installed in the asphalt of the road), air and dew point temperatures, relative humidity, wind speed and direction, etc.

In our study, **for time-series statistical analysis,** for each of road stations for each time (year, month, day, hour) corresponding records the following 6 observed meteorological parameters were extracted: air and dew point temperatures and relative humidity at 2 m, road surface temperature, wind speed and direction at 10 m. These were taken at 1 hour intervals (at 00, 01, …, 23 UTCs) in correspondence with forecasted meteorological parameters taken for time-series analysis. If there are no missing observations in the analysed period (1 Oct 2012 – 15 Mar 2013) there could be in total 1573680 records (166 days x 24 UTCs x 395 stations) corresponding to observations.

For **operational city weather applications** only the 6 following parameters - air and dew point temperatures and relative humidity at 2 m, road surface temperature, wind speed and direction at 10 m - were used; and in particular, at every 1 hour interval (at 00, 01, …, 23 UTCs).

2.2. Analysis of Meteorological Parameters

For our study we had several time-series of hourly observed and forecasted meteorological parameters (see Ch. 2.1). All these were given at successive points/stations in time spaced at uniform time intervals. It is important to analyse these data, extract and evaluate meaningful statistics and other characteristics. In order to elaborate a method for correction of temperature forecast, let' analyse structure of various meteorological parameters considering them as continuous functions, covering some distributions and having daily regularities on a diurnal cycle.
Three key criteria are evaluated and considered of importance for the method for each meteorological parameter: 1) type of distributions (the closer to the normal distribution - the better); 2) periodicity (the more regularities are observed on a diurnal cycle – the better); and 3) correlation (the stronger correlation between meteorological parameters – the better). For both the city and road weather applications, these three criteria were tested for all available forecasted and observed meteorological parameters on different temporal scales: hourly, daily, monthly, and overall studied period.

Distributions
In math sciences and statistics, a normal distribution for a variable (with mean and variance) is a statistic distribution with a probability density function having a curved flaring shape. In real world applications, the occurrence of the normal distribution can be divided into a few categories. For example, it can be - exactly normal distribution (which is very rare situation in the real world) and approximately normal (when it is justified by the central limit theorem). Moreover, it can be called as distribution modelled as normal (e.g. distribution with maximum entropy for a given mean and variance). In physics sometimes it is called as a Gaussian distribution. Physical related quantities that are expected to be a sum of many independent processes often might have a distribution which looks very close to the normal. Some meteorological parameters follow roughly (close to) normal distributions with a few values (called outliers) located at the high and low ends and many of them in the middle. The normal distribution is also symmetric about its mean value.

Periodicities
In general, in math sciences, a periodic function is a function which always repeats its values at regular intervals or periods. For atmospheric sciences, periodic functions can be used to describe phenomena or meteorological parameters that exhibit periodicities on different scales. Evaluation of existing periodic variations/fluctuations for meteorological parameters is more natural because many of them have more clear defined and well pronounced diurnal cycles, and it is especially valid for temperature related parameters. Such variations can be regularly or semi-regularly repeated every day, or simply there is always an occurrence of the minimum and maximum values during the day, repetitive behaviour, and hence, there is always a possibility to try to predict/forecast value for the next similar moment or period of time.

Correlations
If it is assumed that the linear relationships are dominating among some of meteorological parameters, the Pearson correlation coefficients r (as dependencies between datasets of pairs of variables) can be calculated. The range of r is between ±1 inclusive. The sign of the correlation coefficient determines whether the correlation is positive or negative: $r = -1$ (perfect negative correlation), $r = 0$ (no correlation), and $r = +1$ (perfect positive correlation). As a correlation coefficient value gets closer to ±1, it is getting stronger (note, for example, that a sign just show the direction of correlation). The magnitude of the correlation coefficient determines the strength of the correlation. Let's assume that: $0 < |r| < 0.3$ (weak), $0.3 < |r| < 0.7$ (moderate/fair), and $|r| > 0.7$ (strong) correlations. The correlations between parameters (relationships for: forecasted-observed, forecasted-forecasted, and observed-observed) at different spatial and temporal scales and its variations would be of key

importance for developing a method of statistical correction.

2.3. Methodology for City and Road Weather Applications

2.3.1. General approach: system of equations and solution

Generally speaking, after detailed analysis of time-series of various parameters (let us call these as *X*'s and *Y*'s), let us select only parameters which are the most suited (there are well known natural, physical, chemical, etc. backgrounds for interrelations) for the particular stated problem and satisfying a set of selected criteria. These criteria, in particular, are the following:

1. Selected parameters MUST have **distribution close to the normal,**
2. Selected parameters MUST have **periodicity on the diurnal cycle**, and
3. Selected parameters MUST be **well correlated.**

If the selected parameters are called X_1, X_2, X_3, and Y, and these parameters had satisfied the three mentioned above criteria, than a system of linear equations can be built and solved afterwards. Let us assume a general form of such linear equation (with *a,b,c* – coefficients) as the following:

$$Y = a{\cdot}X_1 + b{\cdot}X_2 + c{\cdot}X_3 \tag{2.3.1}$$

Let us consider a horizontal axis as *t*-time and pre-history (consider in days: -1d, -2d, -3d, …, -*n*d) of the periodic function variability with time (Figure 2.3.1) then similar equations can be written for any consecutive day in the pre-history and system of equations (based on *Ys* and *Xs*) can be constructed in the following form:

$$\begin{array}{rl} 0\ day: & Y_0 = a{\cdot}X_{1,0} + b{\cdot}X_{2,0} + c{\cdot}X_{3,0} \\ -1\ day: & Y_1 = a{\cdot}X_{1,1} + b{\cdot}X_{2,1} + c{\cdot}X_{3,1} \\ -2\ day: & Y_2 = a{\cdot}X_{1,2} + b{\cdot}X_{2,2} + c{\cdot}X_{3,2} \\ -3\ day: & Y_3 = a{\cdot}X_{1,3} + b{\cdot}X_{2,3} + c{\cdot}X_{3,3} \end{array}$$

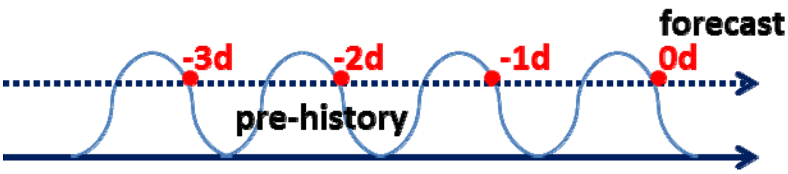


Figure 2.3.1: *Periodic function variability with respect to pre-historical (previous days) and forecast (current day) data.*

Note, that if the system is inconsistent than there is no solution; if the system is dependent than there is an infinite number of solutions; and if the system is consistent than there is one solution (that is our task to find this unique solution).

At second, a square matrix (if only 3 consecutive days in the pre-list history are considered with only 3 parameters) based on *X*'s is constructed. This matrix should also satisfy several specific conditions. Note, that the matrix should be well posed (i.e. solution exists; this solution is unique; and solution behavior can hardly change with changes in *X*'s). The matrix should not be singular (i.e. determinant $det = 0$); as well as *X*'s in matrix should be well scaled (i.e. *X*'s need to be of the same order of magnitude).

At third, solve the matrix to find coefficients (*a, b, c*) using one of the methods. In particular, in our study, the singular value decomposition (SVD) method was evaluated and then applied. This SVD method is used for solving most linear least-squares problems (see detailed description in *Press et*

al. (1992); Golub & Van Loan (1996)).
Note, if the number of days considered in the pre-history is larger, than, for example, a set of triplets of equations can be considered to find coefficients. There is an overlap between equations; at least, two previous days equations are included in each set. Here, coefficients are calculated based on averaging (as seen in Figure 2.3.2). The longer the pre-history sequence of days is considered, the closer to an averaged (smoothed) situation the coefficients will approach.

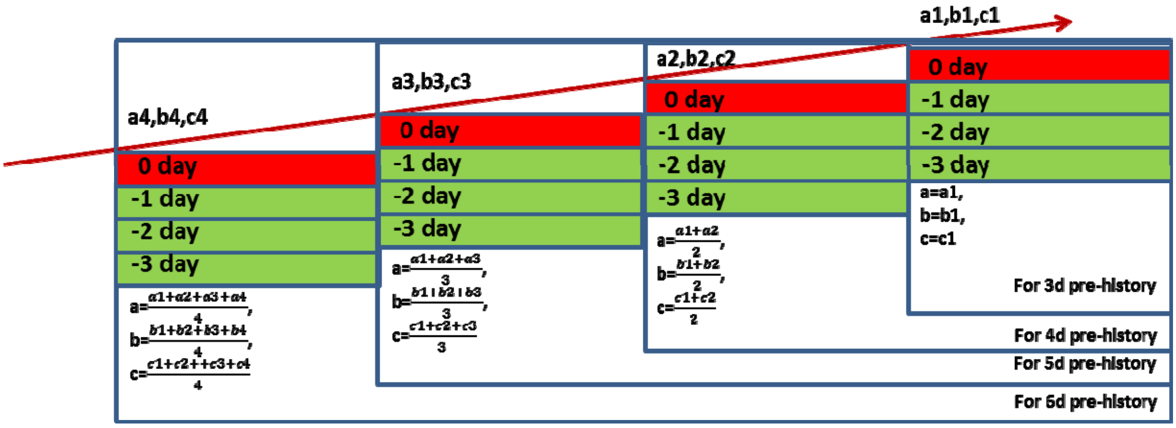


Figure 2.3.2: *Calculation of coefficients a, b, c based on the number of the pre-history days.*

2.3.2. Approach to statistical correction of temperatures

From statistical analysis of time-series of meteorological parameters, only several of these satisfying stated criteria (see Ch. 2.3.1) were selected. For the city weather (NWP – byvej) application, these included air temperature at 2m ($Y=T2m$), dew point temperature at 2m ($X_1=Td2m$), wind speed at 10m ($X_2=V10m$), and surface temperature ($X_3=Ts$). For the road weather application (RWM – vervej), these included road surface temperature ($Y=Trs$), dew point temperature at 2m ($X_1=Td2m$), wind speed at 10m ($X_2=V10m$), and air temperature at 2m ($X_3=T2m$).

Further outline will be given on example of the air temperature correction for NWP application, but similar steps are identical for the RWM application as well, where the *Y*-parameter from the air temperature (*T2m*) is replaced by the road surface temperature (*Trs*).
Let us assume that we have original air temperature forecast ($\boldsymbol{T2m_{for/orig}}$) produced by NWP model. Let us assume that this forecast still can be additionally improved (in particular, at first, the bias can be reduced to a smaller value). It can be done by recalculation of original forecast.
For that, following outlined methodology in Ch. 2.3.1, at the first step, we build system of linear equations, construct and solve matrix using SVD method, and finally - calculate coefficients:

$$
\begin{array}{l}
(j{=}{-1})\ {-1}\ day:\ \boldsymbol{T2m}\big|_{obs}^{-1d} = a\cdot \boldsymbol{Td2m}\big|_{for}^{-1d} + b\cdot \boldsymbol{V10m}\big|_{for}^{-1d} + c\cdot \boldsymbol{Ts}\big|_{for}^{-1d} \\
(j{=}{-2})\ {-2}\ day:\ \boldsymbol{T2m}\big|_{obs}^{-2d} = a\cdot \boldsymbol{Td2m}\big|_{for}^{-2d} + b\cdot \boldsymbol{V10m}\big|_{for}^{-2d} + c\cdot \boldsymbol{Ts}\big|_{for}^{-2d} \\
(j{=}{-3})\ {-3}\ day:\ \boldsymbol{T2m}\big|_{obs}^{-3d} = a\cdot \boldsymbol{Td2m}\big|_{for}^{-3d} + b\cdot \boldsymbol{V10m}\big|_{for}^{-3d} + c\cdot \boldsymbol{Ts}\big|_{for}^{-3d}
\end{array}
\qquad (2.2.1)
$$

where: *Y*'s (as *T2m*) on the left-hand side of each equation are replaced with observations (***obs***) for air temperature (taken from previous consecutive days *-1d, -2d, -3d* in the pre-history); X's (*Td2m, V10m, Ts*) on the right-hand side of each equation are forecasts (***for***) for dew point temperature, wind speed and surface temperature; and ***a,b,c*** – coefficients of the matrix to be calculated using SVD method; $j = [-1, -M]$ – time steps (in days) in pre-history (M=3).

Then, at the second step, the 1st iteration correction (***$T2m_{cor}$***) to original forecast (i.e. on 0 day) can be calculated following equation:

$$T2m_{cor} = \frac{a \cdot Td2m\,\big|^{0d}_{for} + b \cdot V10m\,\big|^{0d}_{for} + c \cdot Ts\,\big|^{0d}_{for}}{D}$$

where: $D = N \cdot SP - N$; N – number of selected meteorological parameters (N=3); and SP – scaling parameter (=10; for meteorological parameters such as *T2m, Td2m, V10m*, and *Ts*)

Then, at the third step, adjust the calculated $T2m_{cor}$ based on critical values (see Ch. 2.3.3) for selected forecasted meteorological parameters. In particular, for the NWP application, the selected forecasted meteorological parameters include: cloud cover, mixing height, wind speed, *T2m–Td2m*, *T2m* bias and mae variability from previous days. For the RWM application these parameters are the following: cloud cover, cloud base, wind speed, *T2m–Td2m*, *Trs* bias and mae variability from previous days.

And finally, based on adjusted new value of $T2m_{cor}$ to recalculate the new revised forecasted value as: $T2m_{for} = T2m_{for/orig} \pm T2m_{cor}$.
The schematics of the temperature correction procedure for the city (*Ta=T2m*) and road (*Ts=Trs*) weather applications is shown in Figure 2.3.3.

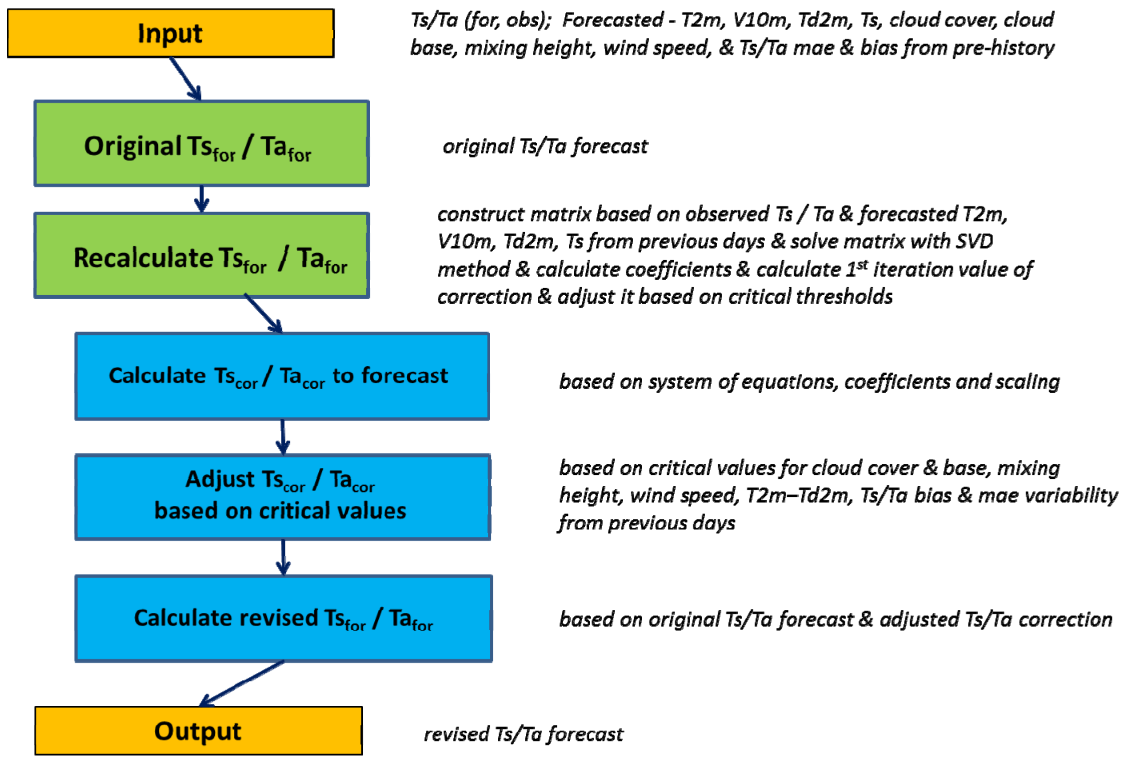


Figure 2.3.3: *Schematics of the temperature correction procedure for the city (Ta=T2m) and road (Ts=Trs) weather applications.*

2.3.3. Temperature adjustment based on critical values

Detailed statistical analysis of long-term time series of observations and forecasts showed that the calculated corrections to temperature can be additionally adjusted depending on specific meteorological conditions on a day-to-day basis.
In particular, for NWP application, the adjustment is done based on changes in forecasted - wind

speed at 10m (*U=V10m*, in m/s); boundary layer height (*H=PBLH*, in meters); cloud cover (*CLC=ClCov*; in % from 0 to 100); and difference between air and dew point temperatures at 2m (*DT=T2m-Td2m*, in °C).
For RWM application, it included the same parameters, except the boundary layer height is represented by the forecasted value of the cloud base height (*H=ClBase*, in meters). A summary is given in Table 2.3.1, where the correction to the temperature is a function of changes in *U, H, CLC*, and *DT* parameters (or *if statements*).
For example, as seen in Table, for meteorological situation when cloud cover is more than 75%, wind speed is less than average wind speed, boundary layer height is less than average value, and difference between air and dew point temperatures is less than 2°C, then the adjustment to the final value of correction to temperature will be calculated as a minimum value between correction calculated after the first iteration and value of 0.25°C.

Table 2.3.1: *Adjustment of initial value of the correction to temperature (T_{cor} = $T2m_{cor}$ = Trs_{cor}) depending on different meteorological conditions (as a function of: CLC - cloud cover, U - wind speed, H - boundary layer height/cloud base height; DT - difference between air and dew point temperatures; avg – average value of forecasted meteorological parameter).*

T	***Adjustment***	***If statements***			
		CLC, %	***U, m/s***	***H, m***	***DT, °C***
T_{cor} =	min (T_{cor} , 0.25)	≥ 75	$\leq U_{avg}$	$\leq H_{avg}$	≤ 2
	max (T_{cor} , -0.25)	≥ 75	$\leq U_{avg}$	$\leq H_{avg}$	> 2
	min (T_{cor}, -0.25)	≥ 75	$> U_{avg}$	$\leq H_{avg}$	> 2
	¼ · max(T_{cor} , 0.5)	≤ 25	$\leq U_{avg}$	$\leq H_{avg}$	≥ 2
	¼ · max(T_{cor} , 0.5)	≤ 25	$> U_{avg}$	$\leq H_{avg}$	≥ 2
	0.1	≤ 25	$\leq U_{avg}$	$\leq H_{avg}$	≥ 5
	0	≤ 25	$\leq U_{avg}$	$\leq H_{avg}$	< 1
	-1 · T_{cor}		$\geq 1.75 \cdot U_{avg}$		
	2 · T_{cor}		$\leq 0.50 \cdot U_{avg}$		
	-0.2 · T_{cor}			$\geq H_{avg}$	
	0.75 · T_{cor}	= 100			
	1.25 · T_{cor}	= 0			

2.4. Operationalization of Statistical Correction Procedure

The Figure 2.4.1 outlines a sequence of major segments of the statistical correction to air temperature procedure for NWP application. The complete setup of the temperature correction procedure was done on operational machine and it runs through crontab 4 times each day, e.g. after the NWP HIRLAM-SKA (3 km resolution) model run is completed and observed and forecasted data became available. In particular, for the 00 UTC model run, the correction procedure will start at 04 UTC hours (for 06 – at 10; for 12 – at 16, and for 18 – at 22 UTC times).
In addition, visualisation of the temperature correction and final air temperature 2-dimensional fields, and backup/achieving is also done 4 times per day with crontab, but it takes place one hour later for each run (e.g. at 5, 11, 18, and 23 UTCs).
The code is written in python (for segment: extraction, assigning, correction, and decision-making) and perl (for interpolation and visualisation (with metgraf tool)), and a sequence of scripts is used to initialize corresponding runs with through crontab.
The CPU time required for each run to finalize segments - extraction, assigning, correction, decision-making, and interpolation - is less than 15 minutes (tested for Jun-Jul 2013 runs on a basis of Danish synoptical stations). Note, that the largest portion of the CPU time used is required for

interpolation segment when simultaneously 49 tasks (1 per each of 49 forecast lengths) are run on a supercomputer. The more stations (points) included into calculation and the larger is the area of the domain where the correction will take place, than the larger CPU time will require for the interpolation procedure.
The visualisation and achieving/backup segments are additional segments which were added for a convenience of analysis and safety of results/data (these will require additional extra time).

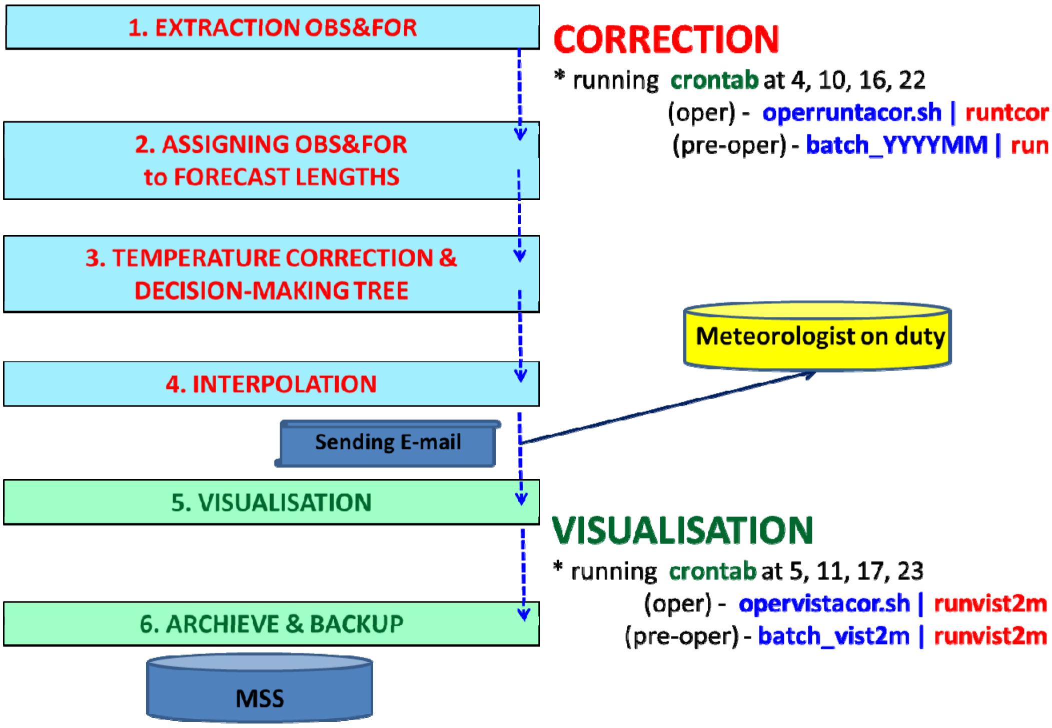


Figure 2.4.1: *Schematics of the operationalization procedure for the statistical correction of the air temperature forecast for the city weather applications.*

2.4.1. Extract observation and forecast data

At first, the original observation and forecast data are extracted and re-written for stations in separate files. Each record in re-written file with observations has: UTC time, year, month, day, hour, forecast length, station ID, and air temperature - *oT2m* (in °C), dew point air temperature - *oTd2m* (in °C), wind speed - *oUwind* (in m/s), wind direction - *oDwind* (in degrees, 0-360). Each record in re-written file with forecasts has: UTC time, year, month, day, hour, forecast length, station ID, and wind speed, wind direction, air temperature, dew point air temperature, cloud cover – *ClCov* (in %, 0-100), surface temperature – *Ts* (in °C), and boundary layer height – *PBLH* (in meters).

2.4.2. Assign to forecast lengths

At second, re-written observation and forecast data are assigned to each forecast lengths ranging from 00 till 48 hours. Each record in file has: UTC time, year-month-day-hour of the start time, forecast length, station ID, year-month-day-hour of the end time, forecast length end (as a sum of hour of the start time and forecast length), forecasted (wind speed, wind direction, air temperature, dew point air temperature, cloud cover, surface temperature, boundary layer height) and observed (air temperature, dew point air temperature, wind speed, wind direction) data. Note for the date of the forecast (i.e. 0 day) there are no yet observations, and therefore, these are replaced with missing values (-99.99). After the next NWP model run, once observations will become available, the missing values are replaced with really existing.

2.4.3. Correct temperature, evaluation statistics, decision-making

For each SKA-model run, the correction procedure is also run. It will start from statistical evaluation of outcomes from previous one-day and multi-days (up to 30 days) runs. For each UTC run, it generates summary statistics in separate files by stations and by forecast lengths.

For all day in pre-history, for each station, the **signs of temperature biases** are evaluated from previous days (assigning "1" as for positive bias, "0" – negative bias, and "9" – missing data or not yet available because observations will become available later) and saved in a separate file (1 file).
For improved/not-improved cases, for each station for each forecast length (00-48 hours of FL), the statistics on "good" (e.g. after correction applied the bias became smaller) vs. "bad" (e.g. after correction applied the bias became larger) cases is saved in separate files (number of stations x number of forecast lengths). Each record includes: forecast length, start time, end time, year-month-day-hour of the start time, forecast length, station ID, year-month-day-hour of the end time, forecast length end (as a sum of hour of the start time and forecast length), status (good/bad), identifier for range of forecast lengths (1 – FL: 00-23; 2 – FL: 24-47; 3 – FL:48-71 hours); air temperatures before/after correction, observed temperature, correction to temperature, temperature bias before/after correction, bias trend, forecasted parameters (wind speed, clod cover, boundary layer height, air and dew point temperatures difference, code of one of the adjustments (following Table 2.3.1), status on sign of correction (positive/negative).

For the previous one-day run, the generated statistics includes distribution of number (and percentage) of cases (summarized over all stations) by forecast lengths for cases - available/ missing/ saved/ with mae less or more than 1°C/ corrected / not-corrected / corrected within the limits of uncertainty (0.25°C) / all corrected and total / before and after correction for hit-rate more or less than critical value (2 or 1°C), bias and mae as well as status-flag (0 – not improved, 1 – improved) for hit-rate, bias and mae.
It also includes behaviour of a new bias (was it improved or not after correction compared with old bias value) for each station by forecast lengths (00-48 hours). The status-flags used are: 0 – bias was not improved, 1 – bias improved, and 5 – bias is within the limits of uncertainty, LOU (0.25°C). For each station there is a count on a number/percentage of forecast lengths where the bias was improved, as well as for each forecast length there is a count on a number/percentage of stations where the bias was also improved.

For the previous multi-days run, a similar statistical output is generated, with including also a number of days (for those the statistic is calculated) for each station. Note, that the multi-day statistics for station will only start to be calculated when the minimum number of days is, at least, 7. There are only two status-flags for bias: 0 – bias was not improved, and 1 – improved. Note, if the number of days will reach a limit of 30 days, than only the most recent 30 day period in pre-history will be used to generate statistics. It means that the performance of the statistical procedure can be evaluated for a period of a month from the current run and it is done every time when the model is run (00, 06, 12, and 18 UTCs runs). In addition, for a multi-day runs the statistics is produced for each forecast length (in separate files) by stations. For each station output includes information on: model run time, forecast length, station ID, number of days used in evaluation, percentage of improved/ not–improved / improved within LOU / all improved and total, percentage of negative/positive biases and hit-rates before/after correction applied; avg/max/min values of correction; avg bias and mae and difference between maes before/after correction; max/min/range for bias before/after correction applied.

Moreover, even if the "abrupt" changes in the NWP model code will take place (for example, changes in physics, parameterisations, dynamics, assimilation, etc.), than within a relatively short period of time the correction procedure will adapt to new reality (for example, if previously the model was mostly over-predicting and from now it started to under-predict) and will start to correct based on new statistics from forecasted meteorological parameters.

Although temperature corrections are calculated for each station, but for the interpolation procedure only for selected "valid" stations corrections will be used, and for the rest of stations: 1) the values can be replaced with zeros, or 2) records for these stations will not be written at all.

2.4.4. Interpolate corrections in domain

The interpolation will take place once the temperature correction value is finally chosen after the decision-making tree. Interpolation is done twice taking into account: 1) calculated corrections for all stations and 2) only corrections approved through the decision-making tree. The interpolation method used is optimal interpolation with variable weighting coefficients and taking into account distance between stations, altitude of stations above the sea level, and land-mask domination (e.g. sea vs. land surroundings in a grid cell, where the station is placed). There is no interpolation if a grid cell is completely occupied by sea. Once the interpolation is finished a notification e-mail is send that output data became available for meteorologist-on-duty. For each of two interpolation runs, the output (grib-format files) is presented as the final corrected air temperature field and saved in 49 separate files (each file for each forecast length 00-49 hours).

2.4.5. Visualise, achieve, backup

The visualisation of corrections to air temperature and final (or corrected) air temperature fields can be also done. It is also done once the interpolated data became available. For each of two interpolation runs, in total 96 (49+49) files of png-format are produced (correction and final air temperature fields) for each of 49 forecast lengths. The scale-legend for the correction to air temperature is ranging from −2.5°C to +2.5°C, and for values within a range of ±0.1°C a white colour was selected. In particular, inspection of multiple visualised outputs showed that smaller values (±0.1°C) are more frequent in night times, reflecting the fact that for Denmark the correction was very small.
Finally, the generated output was also achieved and backup-ed.

3. Results and Discussions

3.1. Relationships between Meteorological Parameters

3.1.1. Distributions

In this study, in order to develop a method for statistical correction of the temperature forecasts analysis of forecasted and observed meteorological parameters was performed for long-term time series of data. In particular, the type of distribution, periodicity on a diurnal cycle and correlations between meteorological parameters were evaluated. Note, that statistical analysis of data for road weather applications was done by *Mahura et al. (2012),* and here, results of such evaluation are done for the city weather applications. In total, 20 forecasted (based on HIRLAM-SKA model runs at 00, 06, 12, and 18 UTCs with forecast lengths of 6, 12, 18, and 24 hours) and 7 observed meteorological direct and derived parameters for 60 synoptical stations were evaluated for a period from 1 May till 31 Aug 2012.

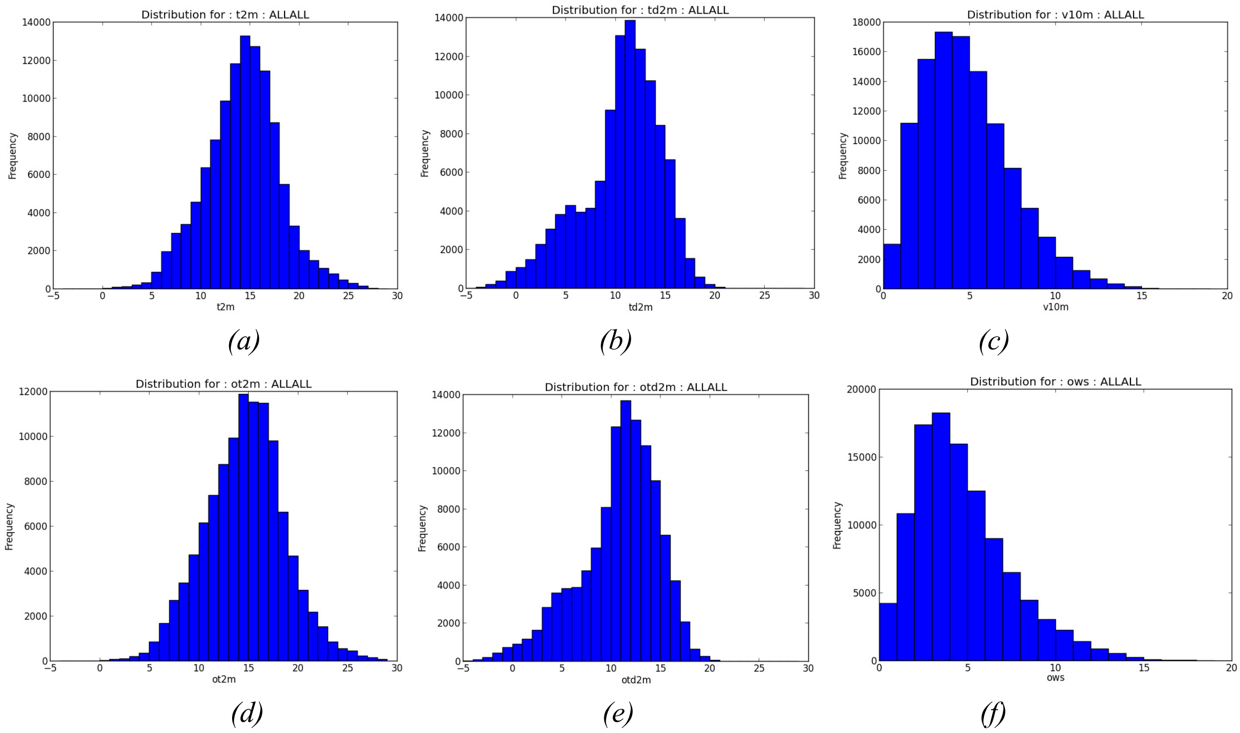


***Figure 3.1.1**: Distribution of (abc) forecasted and (def) observed meteorological parameters: (ad) air temperature at 2 m, (be) dew point temperature and (cf) wind speed for 60 synoptical stations covering period from 1 May till 31 Aug 2012 (included 00, 06, 12, 18 UTCs and all – 6, 12, 18, 24 h - forecast lengths).*

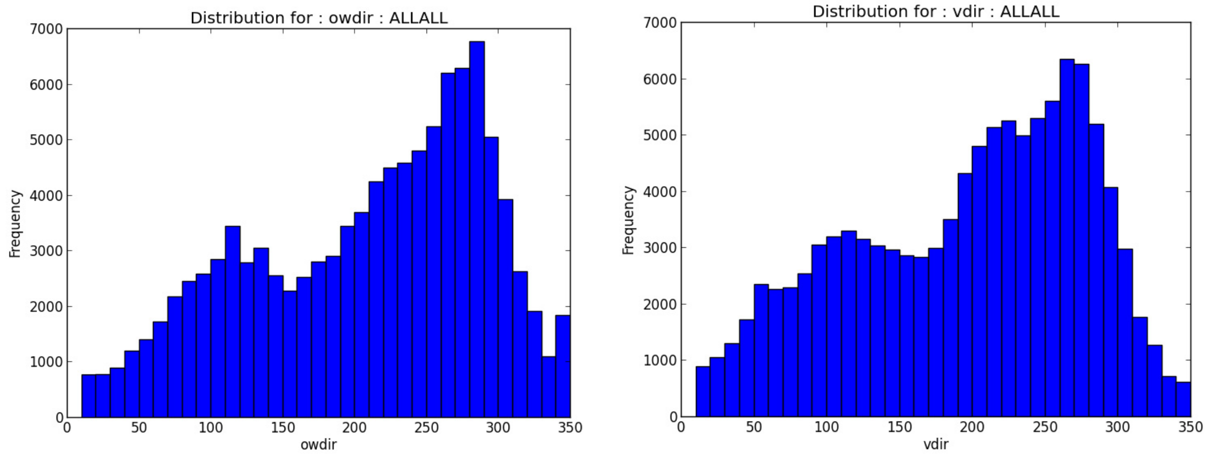
***Table 3.1.1**: Variability of distributions of meteorological parameters (observed and forecasted) as a function of the HIRLAM-SKA model runs (at 00, 06, 12, 18 UTCs) with forecast lengths (of 6, 12, 18, 24 h) for 60 synoptical stations covering period from 1 May till 31 Aug 2012 /Remarks on distributions: N – close to normal; U – U-shape; biM – bi-modal; SR – skewed to the right; SL – skewed to the left; O+N – close to normal with large number of zero values; O+SR – skewed to the right with large number of zero-values*

	0006	0012	0018	0024	00ALL	0606	0612	0618	0624	06ALL	1206	1212	1218	1224	12ALL	1806	1812	1818	1824	18ALL	ALL06	ALL12	ALL18	ALL24	ALLALL
FORECASTS																									
mslp	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
v10m	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
vdir	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM
t2m	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
td2m	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
clcov	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U
ts	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
pblh	O+SR	O+N	O+SR	O+SR	O+SR	O+N	O+SR	O+SR	O+SR	O+SR	O+SR	O+SR	O+SR	O+N	O+SR	O+SR	O+SR	O+N	O+SR	O+SR	O+SR	O+SR	O+SR	O+SR	O+SR
q2m	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
psurf	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
umom	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
vmom	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
tnlev	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
unlev	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
vnlev	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
difrad	SR	SR	SR	N	O+SR	SR	SR	N	SR	O+SR	SR	N	SR	SR	O+SR	N	SR	SR	SR	O+SR	O+SR	O+SR	O+SR	O+SR	O+SR
latf	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL	SL
senf	SL	SL	N	SR	SL	SL	N	SR	SL	SL	N	SR	SL	SL	SL	SR	SL	SL	N	SL	SL	SL	SL	SL	SL
OBSERVATIONS																									
opsurf	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
ows	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
owdir	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM	biM
ovis	N	N	N	?	N	N	N	?	N	N	N	?	N	N	N	?	N	N	N	N	N	N	N	N	N
oclcov	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U
ot2m	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
ot2m	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N

As expected, majority of the meteorological parameters have shown a distribution close to the normal (see Table 3.1.1), except wind direction, cloud cover, latent and sensible heat fluxes, boundary layer height. For the air temperature, dew point temperature and wind speed these are shown in

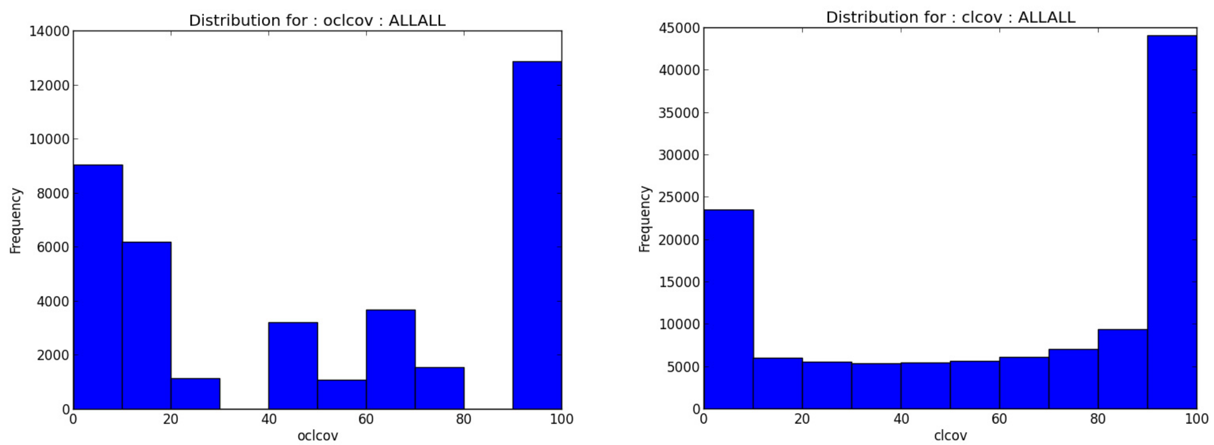
Figure 3.1.1.

The **wind direction** both observed (Figure 3.1.2a) and forecasted (Figure 3.1.2b) has a bi-modal distribution. It corresponds to two dominating sectors: winds from the western directions 270±30° and from the eastern direction 95±30°. As expected, the forecasted **cloud cover** has a form of U-shape distribution, where lowest (cloud free as 0-10%) and highest values (overcast as 90-100%) are the most frequent (Figure 3.1.3a). But, the highest values are almost 1.5-2 times more frequent compared with the lowest ones. For observed cloud cover, the distribution looks like W-shape distribution (Figure 3.1.3b), where both the end with lowest/highest values exists as well there is a secondary peak in the middle range (40-60%). Although note that the observations for analysis were available only at 4 terms – 00, 06, 12, 18 UTCs.



(a) *(b)*

Figure 3.1.2*: Distribution of (a) observed and (b) forecasted wind direction for 60 synoptical stations covering period from 1 May till 31 Aug 2012 (included 00, 06, 12, 18 UTCs and all – 6, 12, 18, 24 h - forecast lengths).*



(a) *(b)*

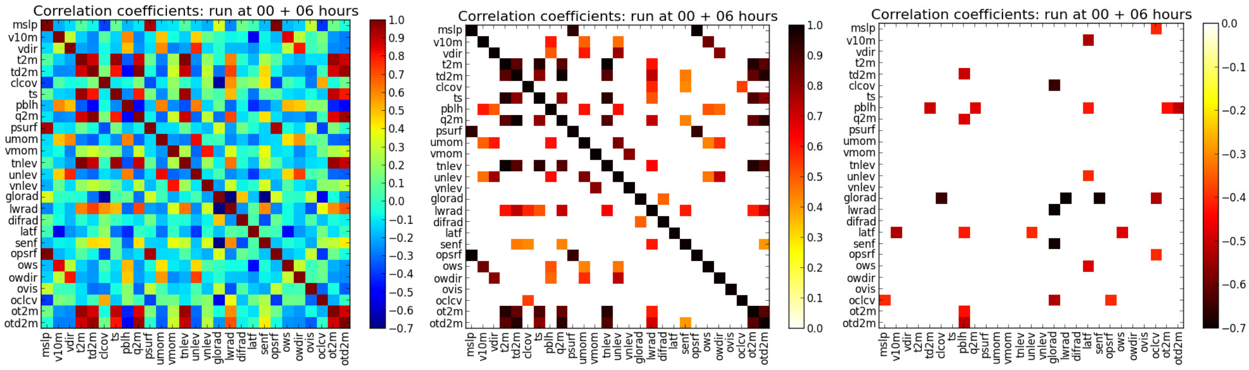
Figure 3.1.3*: Distribution of (a) observed and (b) forecasted cloud cover for 60 synoptical stations covering period from 1 May till 31 Aug 2012 (included 00, 06, 12, 18 UTCs and all – 6, 12, 18, 24 h - forecast lengths).*

The averaged over the day the sensible heat flux distribution (not shown) is mostly skewed to the left, but on a diurnal cycle it is closer to the normal in the evening hours, at mid-night it is skewed to the right, and in morning and daytime hours - skewed to the left. The latent heat flux has a distribution skewed to the left. The boundary layer height mostly follows a distribution skewed to the right with more frequent occurrence of low heights, except at mid-day it more closely to the normal. Note, although visibility mostly follows a distribution close to the normal but at mid-night

it is not clearly defined with a more complex structure and shape.
Note that majority of meteorological parameters have shown periodicity on a diurnal cycle, especially it is clearly defined for temperature related parameters

3.1.2. Correlations

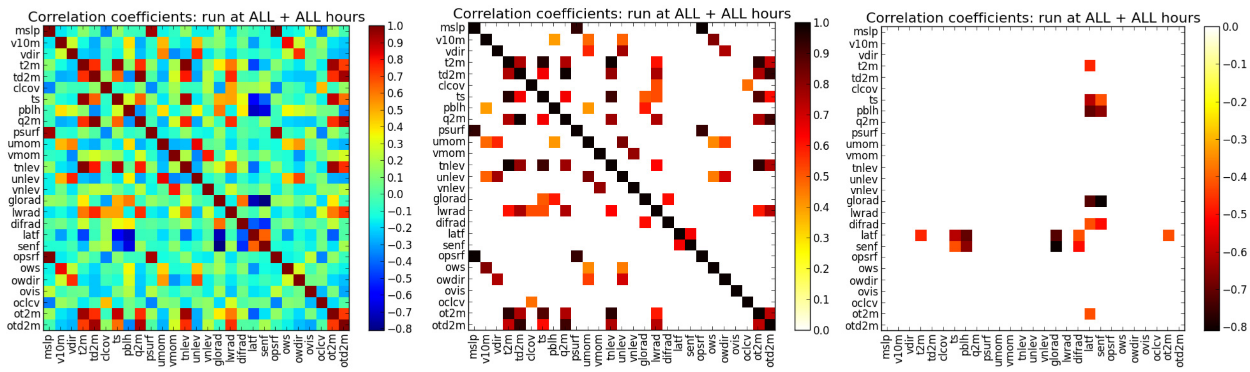
Analysis of correlations between different meteorological parameters (both forecasted and observed) allows us to examine possible strength (magnitude/value of correlation: negligible, weak, fair, strong, very strong) and direction (positively vs. negatively) of relationship between parameters, and hence, try to look into physical backgrounds and processes leading to such correlations. In weather prediction, it also allows us to compare on how well the NWP model is predicting a particular meteorological parameter, and if not, to get more insight on possible limitations in parameterizations or description of processes.



(a) *(b)* *(c)*

Figure 3.1.4*: Matrix of the correlation coefficients between meteorological (both forecasted and observed) parameters based on HIRLAM-SKA model run at 00UTC + 6 h forecast length: (a) all range coefficients; (b) only positive coefficients (cut-off: +0.4); and (c) only negative coefficients (cut-off: -0.4)*

An example of outcome of the correlation analysis is shown in Figure 3.1.4. Here, the matrix of the correlation coefficients between meteorological (both forecasted and observed) parameters based on the HIRLAM-SKA model run at 00UTC (+ 6 h forecast length) is plotted. Figures 3.1.4b and 3.1.4c show only positive vs. negative correlation coefficients stating from ±0.4 (the lower values were cut-off). Figure 3.1.4a shows variability of the coefficient within the full range of calculated values for this example (although maximally it is ±1).

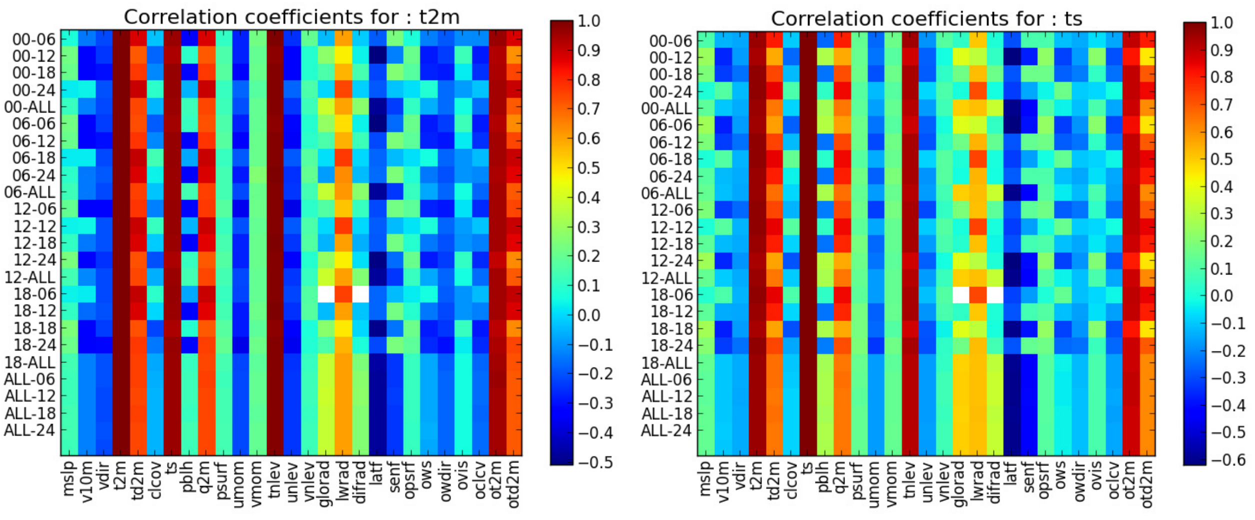


(a) *(b)* *(c)*

Figure 3.1.5*: Matrix of the correlation coefficients between meteorological (both forecasted and observed) parameters based on averaging of all HIRLAM-SKA model runs (00, 06, 12, 18 UTCs) and all forecast lengths (6, 12, 18, 24 hours): (a) all range coefficients; (b) only positive coefficients (cut-off: +0.4); and (c) only negative coefficients (cut-off: -0.4).*

Figure 3.1.5 shows similar matrixes of correlation coefficients, but these were averaged over the entire dataset of available forecasts (all model runs and all forecast lengths) and observations for all meteorological parameters. Figure 3.1.6 shows temporal variability (considering different times of model run and different forecast lengths: depicted on vertical axis) of the correlation coefficients between forecasted *T2m* air temperature and *Ts* surface temperature vs. other meteorological (both forecasted and observed) parameters (depicted on horizontal axis).

Detailed analysis of correlations showed that in general both forecasted and observed temperatures (air, dew, and surface) are strongly correlated with each other. On average, the correlation coefficient between observed both *T2m* and *Td2m* is about +0.73, ranging from +0.66 to +0.91 for noon vs. night-time, respectively. The correlation is similar for forecasted values of these two parameters: +0.75, ranging from +0.73 to +0.91. The forecasted *T2m* and *Ts* are strongly (+0.96) correlated as well during the day, ranging from +0.94 to +0.97. The forecasted *Td2m* and *Ts* are also correlated well: +0.65, but correlation varies significantly during the day. It is larger during night - morning hours (0.85÷0.81), and it is smaller during noon - evening hours (0.56÷0.74).
The forecasted wind speed (*v10m*) is fairly positively correlated (+0.4) with the planetary boundary layer height (*pblh*). This correlation increases substantially during night-time up to +0.78, but it became negligible (less than +0.1) during noon. The observed wind speed (*ows*) is also fairly correlated (+0.35) with *pblh*, and correlation shows same variability on a diurnal cycle: from larger values (+0.68) at night to negligible ones at noon. The forecasted *Td2m, T2m, Ts* are fairly negatively correlated with *v10m*: about -0.3, and this correlation is weaker for observed temperature parameters. Both *v10m* and *ows* are fairly positively correlated (+0.3) with forecasted (*vdir*) and observed (*owdir*) wind directions. Correlation of both observed and forecasted wind speed is negligible for forecasted cloud cover (*clcov*), and although it is slightly higher (up to +0.2) for observed cloud cover (*oclcov*). Correlation between forecasted vs. observed wind speeds is about +0.82, being slightly larger (up to +0.86) during night hours.



(a) *(b)*

***Figure 3.1.6**: Variability of the correlation coefficients between forecasted (a) T2m air temperature and (b) Ts surface temperature vs. other meteorological (both forecasted and observed) parameters (depicted on horizontal axis) as a function of HIRLAM-SKA model runs and forecast lengths (depicted on vertical axis).*

For other selected meteorological parameters – wind speed and direction, dew point air temperature, cloud cover, boundary layer height, latent/sensible heat fluxes, long-wave radiation – the variability of the correlation coefficients with other parameters (note, *results of correlations analysis for all remaining from 26 in total parameters are available upon request*) is given in Appendix A as a function of the model runs and forecast lengths. After detailed analysis of correlations between meteorological parameters and following the requirements previously stated (see Ch. 2), only certain several parameters were selected to be included into the equation for air temperature – air temperature, dew point, surface temperature and wind speed; and a series of others – wind speed, boundary layer height, cloud cover (*T2m* – *Td2m*, in addition) – was selected to be used in the adjustment of the calculated correction value depending on most typical meteorological conditions.

3.2. Statistical Correction for Temperature Forecasts

Results of application of the developed method for statistical correction of the forecasted air and road surface temperatures are summarised in Ch. 3.2.1 (for city weather) and Ch. 3.2.2 (for road weather) for different days, months, periods, individual and averaged over all stations (synoptical and road), forecast lengths, and initial times of the HIRLAM-SKA and RCM models runs.

3.2.1. City weather application – for synoptical stations

For the city weather application, the correction of the air temperature forecasts was performed for May-Aug 2012. The T2m outputs of the HIRLAM-SKA model runs at 00, 06, 12, and 18 UTCs with forecast lengths of 6, 12, 18, and 24 hours were statistically corrected using the method developed in this study (see Ch. 2). As seen in Figure 3.2.1 (left-top-corner) on example for the synoptical station WMO-06030 for the 00 UTC model run with a 6 hour forecast length (for other forecast lengths see Appendix B), the original (i.e. old – before correction) bias and mae were -1.132 and 1.246°C, respectively, and the hit-rate was less than 44% for temperature predicted within an accuracy of ±1°C. When the statistical correction was applied, all listed above parameters showed overall improvement: the corrected (new – after correction applied) bias and mae became -0.396°C (almost 3 times better) and 0.765°C (more than 1.5 times better), with the hit-rate improved by extra 28% reaching 71%.

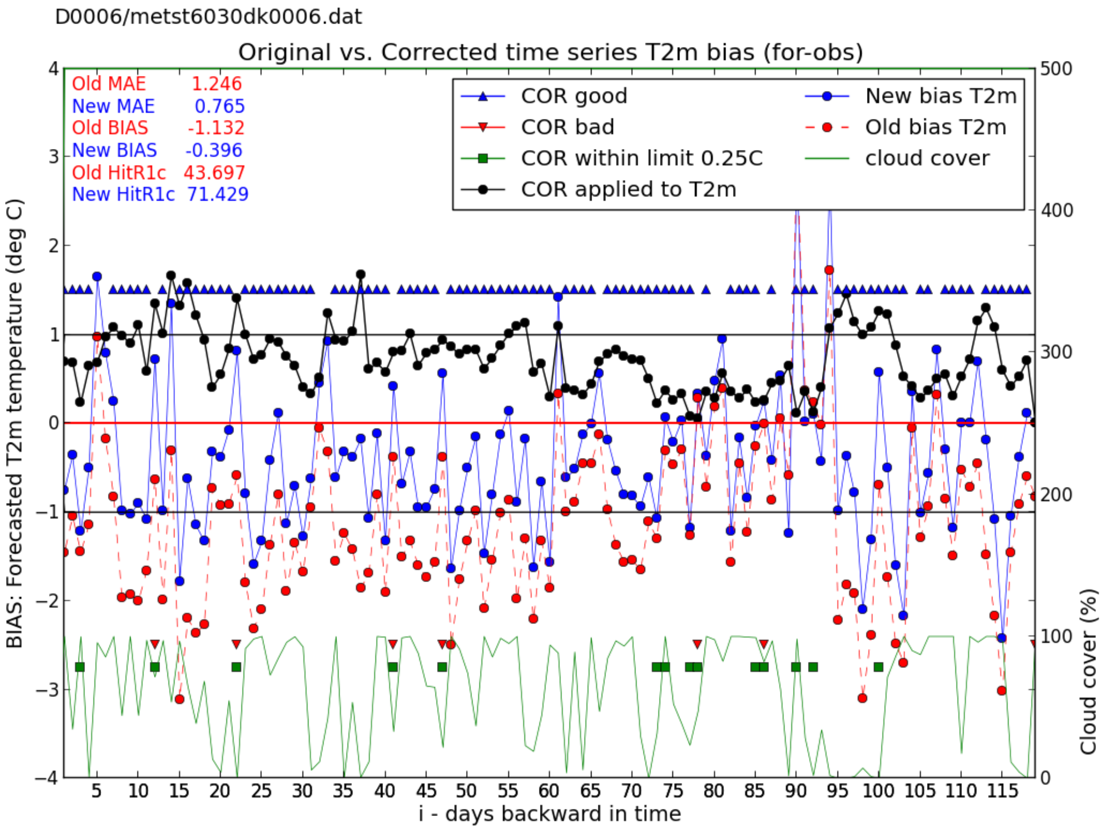


Figure 3.2.1*: Results of statistical correction to air temperature for the synoptical station WMO-06030 for the HIRLAM-SKA model run initiated 00 UTC for 6 hour forecast length (period covered 1 May – 31 Aug 2012; days are in a reverse order on a horizontal time axis).*

The temperature correction related information (covering 4 month period) is represented by several time-series (shown in legend of the Figure 3.2.1):

- COR applied to T2m – value of correction to air temperature calculated from procedure;
- New bias T2m – value of the bias after the correction procedure was applied;
- Old bias T2m – value of the bias before the correction procedure applied;
- COR good – stated the fact that after correction the bias became better/improved;
- COR bad – stated the fact that after correction the bias became worse/did not improved;
- COR within limit 0.25°C – stated the fact that after correction the bias became within ±0.25°C limit (linking with the uncertainty parameter) from the original bias value;
- Plus, there is a time-series for the forecasted cloud cover (0-100) for inter-comparison.

As seen in Figure 3.2.1, on average, the correction to the temperature value is less than 1°C, although for some cases it is more than 1.5°C. For majority of the cases the bias had improved (see COR good cases – blue triangles up), although for 7 cases (see COR bad cases – red triangles down) it became worse. But in total the number of cases with improved bias is more than 90% considering also cases where the bias after corrections applied remained within the limits of uncertainty (which is assumed to be still acceptable for the forecasted value). Note, that after correction applied, there is a possibility that the sign of the bias can be changed, but if by an absolute value the bias became smaller than there is still improvement of the originally forecasted value.

Analysis of all other SKA-model runs vs. forecast lengths by synoptical stations also showed overall improvement of the forecasted air temperature after the correction procedure had been applied (see Table 3.2.1). It allowed to assume that the similar approach can be applied also for other duration forecast lengths, and in particular, for up to 48 hours (see results in Ch. 3.4).

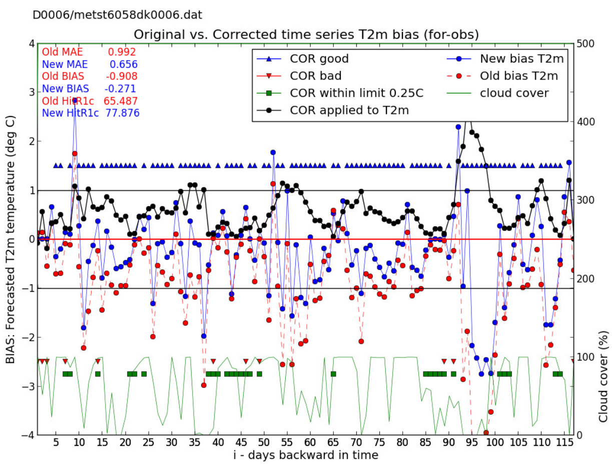
***Table 3.2.1**: Results of evaluation of the T2m correction procedure performance for 4 different HIRLAM-SKA model runs and 4 corresponding forecast lengths averaged over all synoptical stations (covering period from 1 May till 31 Aug 2012) /FL – forecast length/.*

FL	UTC run	# tot	% imp	Hit-Rate, %		T2m correction, °C			Avg T2m bias, °C		Avg T2m MAE, °C		
				old	new	avg	max	min	old	new	old	new	diff
6	**00**	110	***92,1***	71,7	***77,7***	0,25	0,90	-0,03	-0,433	-0,246	0,767	0,641	0,126
	06	109	81,0	52,3	63,3	0,56	1,70	-0,11	-0,562	-0,219	1,202	0,929	0,273
	12	111	89,8	66,1	74,1	0,33	1,18	-0,03	-0,442	-0,227	0,894	0,719	0,175
	18	108	90,3	71,7	75,5	0,24	0,83	-0,05	-0,055	-0,015	0,787	0,689	0,098
12	**00**	111	80,1	51,7	61,3	0,57	1,73	-0,05	-0,546	-0,222	1,219	0,978	0,241
	06	111	88,9	64,6	71,7	0,34	1,23	0	-0,429	-0,218	0,916	0,747	0,169
	12	109	90,5	70,9	74,3	0,23	0,84	-0,06	-0,116	-0,061	0,801	0,703	0,098
	18	109	90,9	70,1	75,9	0,25	0,92	-0,02	-0,423	-0,241	0,795	0,682	0,113
18	**00**	110	87,4	63,2	70,8	0,37	1,33	-0,02	-0,402	-0,183	0,949	0,77	0,179
	06	108	89,9	68,9	72,8	0,24	0,85	-0,05	-0,079	-0,033	0,831	0,732	0,099
	12	109	90,6	68,1	74,6	0,29	1,03	-0,05	-0,483	-0,265	0,838	0,694	0,144
	18	109	78,6	50,1	59,6	0,61	***2,00***	-0,08	-0,604	-0,226	1,269	1,011	0,257
24	**00**	108	88,6	68,2	72,0	0,25	0,90	-0,07	-0,089	-0,042	0,849	0,742	0,107
	06	108	90,4	65,5	73,2	0,30	1,10	-0,02	-0,514	***-0,278***	0,881	0,727	0,154
	12	108	77,3	***47,8***	57,3	0,65	2,09	***-0,12***	***-0,608***	-0,209	***1,339***	***1,057***	***0,282***
	18	110	85,9	61,6	70,1	0,42	1,41	-0,04	-0,463	-0,195	0,986	0,793	0,194
Average		**109**	**87,0**	**63,3**	**70,3**	**0,37**	**1,25**	**-0,05**	**-0,391**	**-0,180**	**0,958**	**0,788**	**0,169**

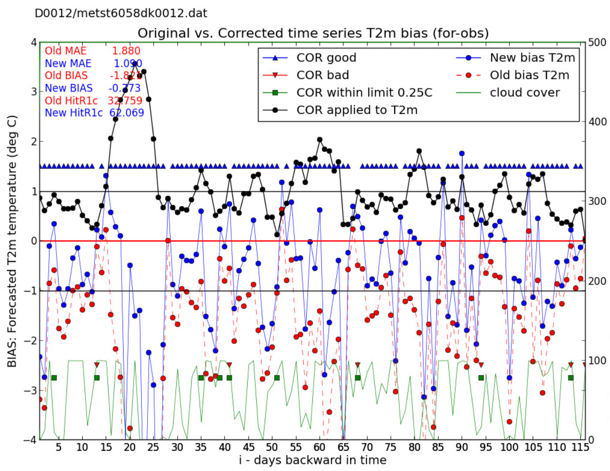
As seen in Table 3.2.1, on average over all stations, about 87% of the T2m forecasts (cases) can be improved with a maximum of about 92% for UTC run at 00 and forecast length 6 hours. The correction value ranges from +2 to -0.12°C with an average of 0.37°C for the studied period. After correction, the averaged bias is improved from -0.39 to -0.18°C as well as the MAE was also improved from 0.96 to 0.79°C. On average, the hit-rate (for biases within ±1°C) improved from 63 to 70%. Such changes are more visible – improved – for some stations than others, and especially for those where the MAE is larger than 1°C. As seen in Table 3.2.2, for WMO-06058 synoptical station, on average about 89% of the cases can be improved. The correction value ranges from +4.1 to -0.29°C with an average of 0.69°C. After correction, the averaged bias is improved from -1.16 to –0.50°C as well as the MAE was also improved from 1.28 to 0.82°C. The hit-rate improved from 54 to 71%. Figure 3.2.2 (similar to Figure 3.2.1) shows results of statistical correction to air temperature for the synoptical station WMO-06058 for the HIRLAM-SKA model run initiated 00 UTC for 6, 12, 18, and 24 hour forecast lengths and covering period from 1 May till 31 Aug 2012.

***Table 3.2.2**: Results of evaluation of the T2m correction procedure performance for 4 different HIRLAM-SKA model runs and 4 corresponding forecast lengths for the WMO-06058 synoptical station (covering period from 1 May till 31 Aug 2012) /FL – forecast length; see Figure 3.2.2 for marked in this table summary results on hit-rate, bias and mae, in red color/.*

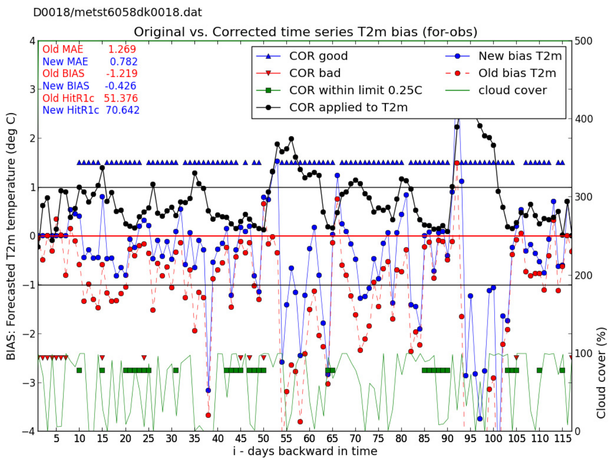
FL	UTC run	# tot	% imp	Hit-Rate, %		T2m correction, °C			Avg T2m bias, °C		Avg T2m MAE, °C		
				old	new	avg	max	min	old	new	old	new	diff
6	**00**	113	85,8	**65,5**	**77,9**	0,59	2,87	-0,19	**-0,863**	**-0,280**	**0,992**	**0,656**	0,335
	06	113	91,2	21,2	62,8	1,06	3,82	-0,18	-1,838	-0,750	1,867	1,016	0,851
	12	117	91,5	65,0	75,2	0,75	3,23	0,09	-1,163	-0,429	1,196	0,715	0,481
	18	112	94,6	72,3	74,1	0,27	1,48	-0,19	-0,722	-0,481	0,873	0,700	0,173
12	**00**	116	92,2	**32,8**	**62,1**	1,05	3,56	0,13	**-1,850**	**-0,800**	**1,880**	**1,090**	0,790
	06	109	88,1	51,4	68,8	0,80	3,03	-0,29	-1,296	-0,484	1,322	0,786	0,536
	12	109	94,5	71,6	74,3	0,29	1,72	-0,19	-0,735	-0,485	0,941	0,738	0,203
	18	114	84,2	63,2	78,1	0,56	2,85	-0,12	-0,843	-0,293	0,993	0,706	0,287
18	**00**	109	88,1	**51,4**	**70,6**	0,77	3,18	-0,23	**-1,204**	**-0,441**	**1,269**	**0,782**	0,488
	06	109	92,7	64,2	70,6	0,27	1,40	-0,19	-0,781	-0,532	0,972	0,769	0,203
	12	111	82,9	61,3	78,4	0,58	2,61	-0,19	-0,874	-0,373	1,028	0,707	0,322
	18	115	88,7	33,0	58,3	1,02	3,24	0,06	-1,741	-0,721	1,785	1,108	0,677
24	**00**	108	80,6	**63,9**	**73,1**	0,59	3,13	-0,19	**-0,743**	**-0,339**	**1,029**	**0,703**	0,325
	06	107	83,2	59,8	74,8	0,61	2,71	-0,19	-0,945	-0,350	1,091	0,736	0,356
	12	115	91,3	28,7	59,1	1,06	4,10	0,11	-1,808	-0,762	1,877	1,129	0,748
	18	107	90,7	52,3	70,1	0,77	2,96	-0,16	-1,217	-0,446	1,300	0,768	0,532
Average		**112**	**88,8**	**53,6**	**70,5**	**0,69**	**2,87**	**-0,12**	**-1,164**	**-0,498**	**1,276**	**0,819**	**0,457**

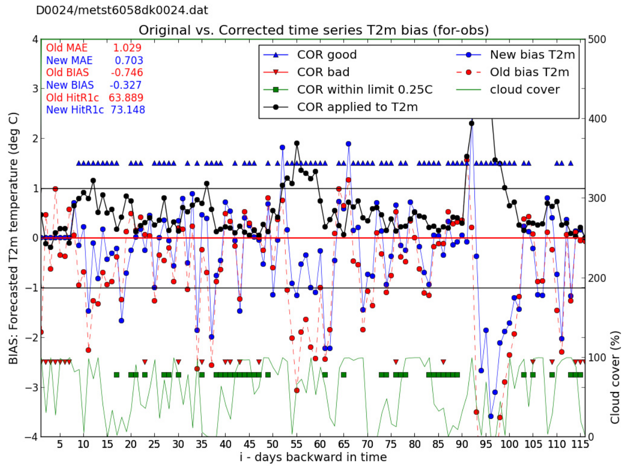


(a)



(b)





(c) *(d)*

***Figure 3.2.2**: Results of statistical correction to air temperature for the synoptical station WMO-06058 for the HIRLAM-SKA model run initiated at 00 UTC for (a) 6, (b) 12, (c) 18 and (d) 24 hour forecast lengths (period covered 1 May – 31 Aug 2012; days are in a reverse order on a horizontal time axis).*

Although the developed method showed possibility of improvement of the air temperature forecast averaging over the entire domain of Denmark, this improvement could be larger for some stations if in addition to the adjustment of the T2m correction (see Ch. 2.3.3), extra turning could be applied. Results of such tuning, with a focus on stations having higher MAE values, are presented in Ch. 3.5. Moreover, the operationalization of the developed method (see Ch. 2.4) was also done for a perspective usage by the DMI Weather Service, and summary of results from the pre-operational test runs is given in Ch. 3.4.

3.2.2. Road weather application – for road stations

For the road weather application, the correction of the road surface temperature (*Trs*) forecasts was performed for the road weather season 2013-2013 (covering period from 1 Oct 2013 till mid-March 2013). Note that the RCM model is running daily every hour producing forecasts up to 24 hours ahead and delivering output to customers/ end-users. Although accuracy of all model runs for all forecast lengths is evaluated, but mainly quality of forecasts corresponding to 3 hour ahead forecast length are of greater concern and it is officially reported through the Danish Road Directorate (DRD) to the customers.

Here, the results of application of the *Trs* statistical correction method (see Ch. 2) are shown on example of the 3 hour forecast length (e.g. model runs at every hour) for individual and averaged over all road stations. January is a month having one of the largest numbers of days with icing conditions as well as it is one of the coldest months. For illustration of analysis, as an example, January 2013 was selected and then inter-compared with the previous January 2012. The reason for such comparison is to show that even after the road RWM system setup had been changed (from 2011-2012 to 2012-2013 season) the method is still valid and applicable, e.g. the method is a model type/setup independent.

As seen in Figure 3.2.3 (left-top-corner) on example for the road station N-1006 for the 06 UTC model run with a 3 hour forecast length, during Jan 2013 the original (i.e. old – before correction) averaged *Trs* bias and mae were -1.006 and 1.047°C, respectively. When the statistical correction was applied, all listed above parameters showed overall improvement: the corrected (new – after correction applied) bias and mae became -0.357°C (almost 3 times better) and 0.665°C (more than 1.5 times better). The temperature correction related information is represented by several time-series (shown in legend of the Figure 3.2.3; *Trs shown as Ts*):

- COR to Ts – value of correction to road surface temperature calculated from procedure;
- New bias Ts – value of the bias after the correction procedure was applied;
- Old bias Ts – value of the bias before the correction procedure applied;
- COR good – stated the fact that after correction the bias became better/improved;
- COR bad – stated the fact that after correction the bias became worse/did not improved;
- COR within limit 0.25°C – stated the fact that after correction the bias became within ±0.25°C limit (linking with the uncertainty parameter) from the original bias value;
- Plus, there is a time-series for the forecasted cloud cover (0-100) for inter-comparison.

As seen in Figure 3.2.3, on average, the correction to the temperature value is less than 1°C, although for some cases it could be more than 1°C. For majority of the cases the bias had improved (see COR good cases – blue triangles up), although for a few cases (see COR bad cases – red triangles down) it became worse. Note that on occasion, the model forecast is already of almost perfect quality (i.e. the bias is very close to 0 value; *see cases corresponding to days 14-27-29*) and its improvement to even smaller bias value will became more complex to handle. That is why there is no such a method for correction where all 100% cases could be improved. Partially to solve this problem, the decision-making was implemented for the operationalization. But in total the number of cases with improved bias is about 87% considering also cases where the bias after corrections applied remained within the limits of uncertainty (green squares in Figure 3.2.3). Note, that although a sign of the bias can be change after correction applied, there is still improvement if by an absolute value the bias became smaller.

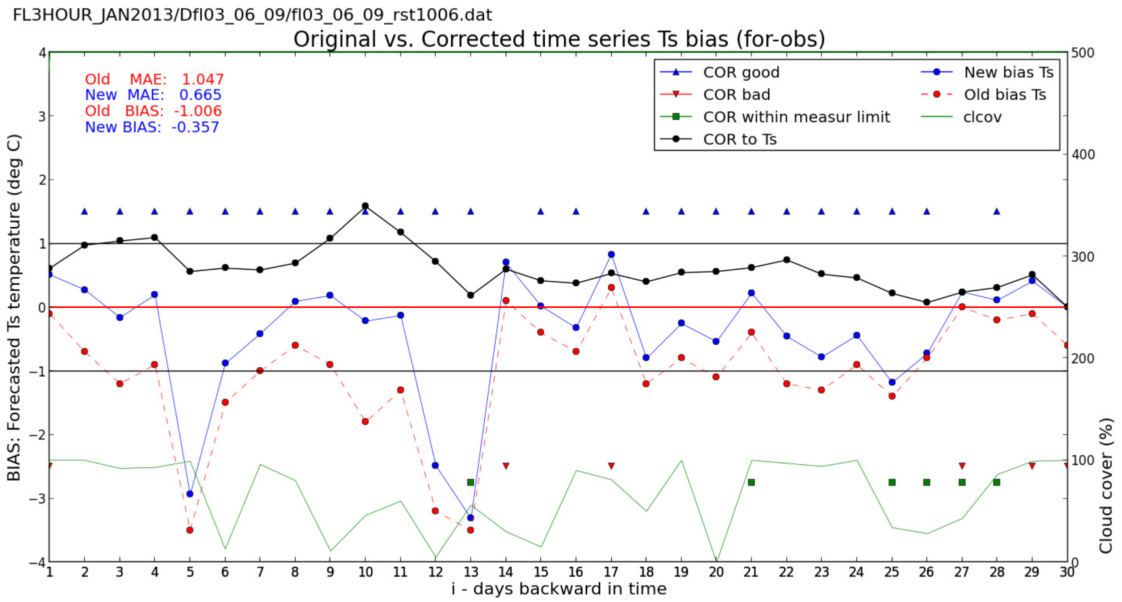
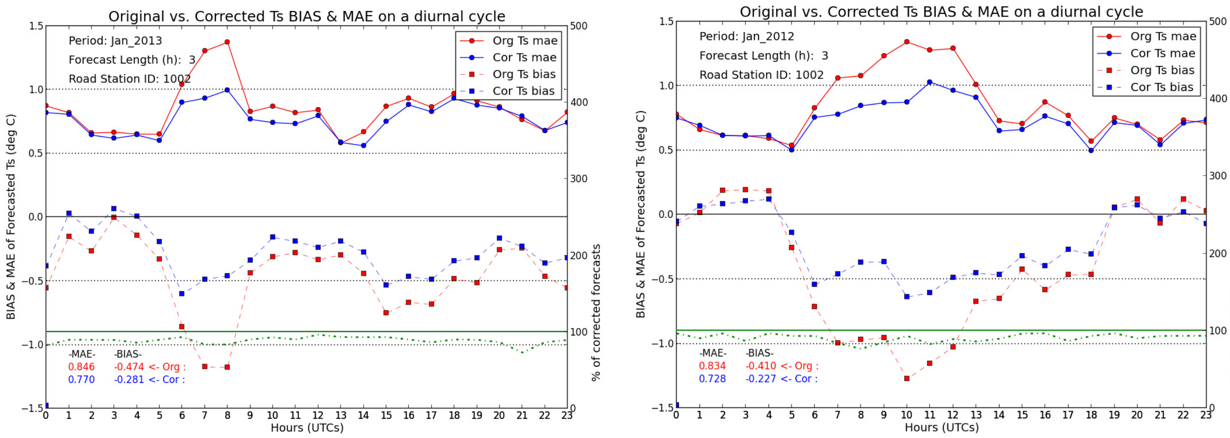


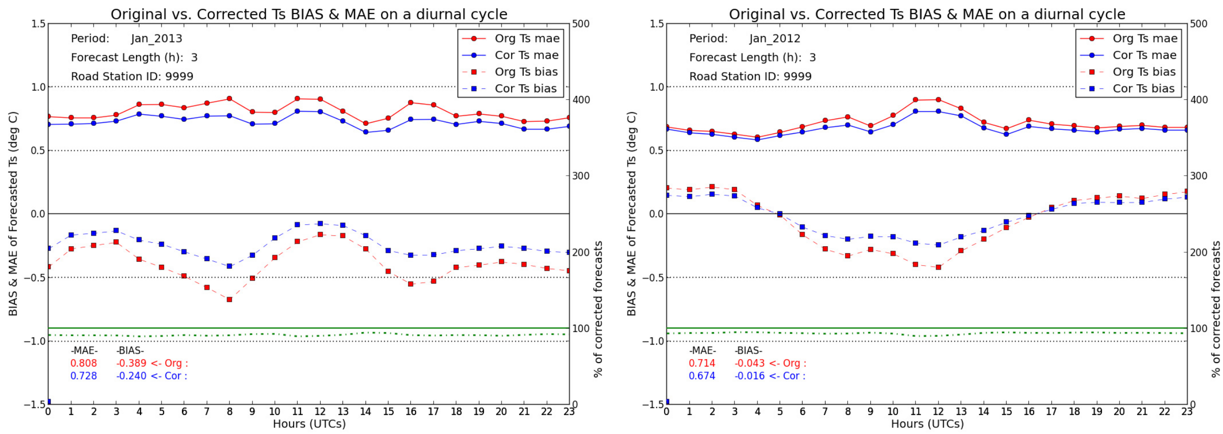
Figure 3.2.3*: Results of statistical correction to road surface temperature for the road station N-1006 for the RCM model run initiated 06 UTC for 3 hour forecast length (period covered - Jan 2013; days are in a reverse order on a horizontal time axis).*

As seen in Figure 3.2.4 (on example of the road station N-1002), for the 3 hour forecast length, the improvement in the road surface temperature is observed also on a diurnal cycle for both bias and mae. It is also seen for the same month of January in different years (e.g. when different setup was used for the RCM model) – 2013 vs. 2012. In Jan 2013 (Figure 3.2.4a), the averaged bias improved from to -0.474 to -0.281°C and mae - from 0.846 to 0.770°C, respectively. In Jan 2012 (Figure 3.2.4b) the averaged bias improved from to -0.410 to -0.227°C and mae - from 0.834 to 0.728°C, respectively. On a diurnal cycle, the bias' as well as mae's improvements are larger during morning hours. In general, it was possible to reduce mae below 1°C for practically all hours when originally it was higher than 1°C. Also, the larger biases were also reduced to smaller values. A similar pattern was observed for the vast majority of other road stations (not shown here). The results of averaging over all road stations are shown in Figure 3.2.5.



(a) *(b)*

Figure 3.2.4: *Results (averaged bias and mae) of statistical correction to road surface temperature on a diurnal cycle for the road station N-1002 and based on the RCM model 24 runs for 3 hour forecast length for January month of (a) 2013 and (b) 2012.*



(a) *(b)*

Figure 3.2.5: *Results (averaged bias and mae) of statistical correction to road surface temperature on a diurnal cycle averaged over 395 road stations and based on all of 24 the RCM model runs for 3 hour forecast length for January month of (a) 2013 and (b) 2012.*

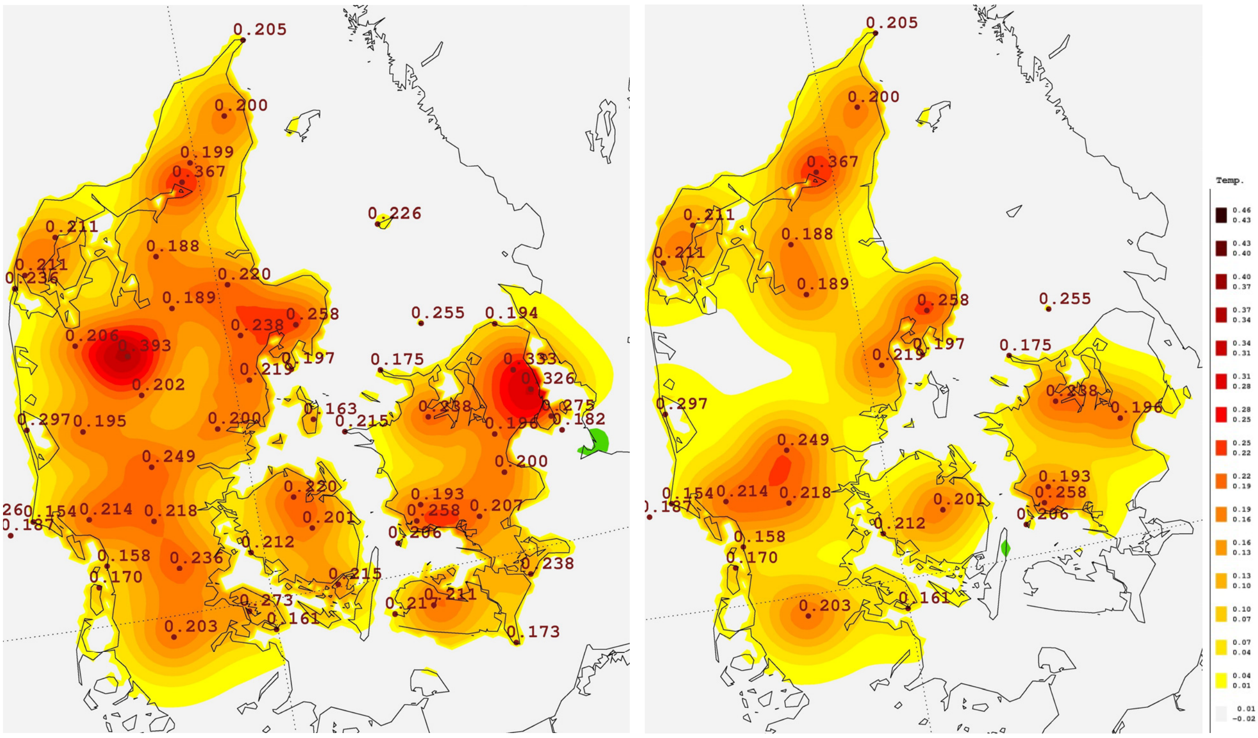
As seen in Figure 3.2.5, averaged over all 395 road stations for the 3 hour forecast length, the improvement in the road surface temperature is observed on a diurnal cycle for both bias and mae. It is also seen on a diurnal cycle (Figure 3.2.5a) for January 2013. The averaged bias improved from -0.389 to -0.240°C and mae - from 0.808 to 0.728°C, respectively, with more than 90% of the cases/ forecasts improved. Note that for the 2012-1013 road weather season the RCM model in general showed overall domination of negative bias, compared with the 2011-2012 season (when negative bias on average dominated during early morning – day-time hours). As seen in Figure 3.2.5b, during evening-nighttime hours although a small change in bias was observed, but the mae values showed improvement for all hours on a diurnal cycle. The averaged bias improved from -0.043 to -0.016°C and mae - from 0.714 to 0.674°C, respectively. On a diurnal cycle, the bias' as well as mae's improvements were largest in the morning-noon hours.

Similar results showing improvement of the original road surface temperature forecasts after applying the statistical temperature correction method were obtained for other months selected. It allowed us to suggest that the method (similarly to the city weather application) could be applied for operational improvement of the road weather forecasts at road stations at different forecast lengths. Initially, implementation (see Ch. 2.3.4) and pre-operational testing (see Ch. 3.4) of the developed method was done on example for improvement of the city weather forecasts.

3.3. Spatial Variability of the Temperature Correction Field as a Function of

3.3.1. Number of stations

Depending on a number of geographical points (associated with geographical coordinates of the synoptical stations) used in interpolation the spatial pattern of the correction to temperature can change as seen in Figure 3.3.1. Figure 3.3.1a shows results of spatial interpolation for temperature correction field, when all (i.e. 60 stations) points are taken into account. Figure 3.3.1b – is based on the half number of points (i.e. 30), which were randomly excluded from the original list of stations. As seen in Figure 3.3.1b, although a general pattern in domination of the signs of correction and its magnitudes are remained similar, but there are spatially also territories (see the Jutland Peninsula of Denmark), where value of correction approaches zeros, shapes of isolines became more smooth giving more areas for smaller values of corrections, and absolute values for maximum and minimum are selected from only 30 stations.



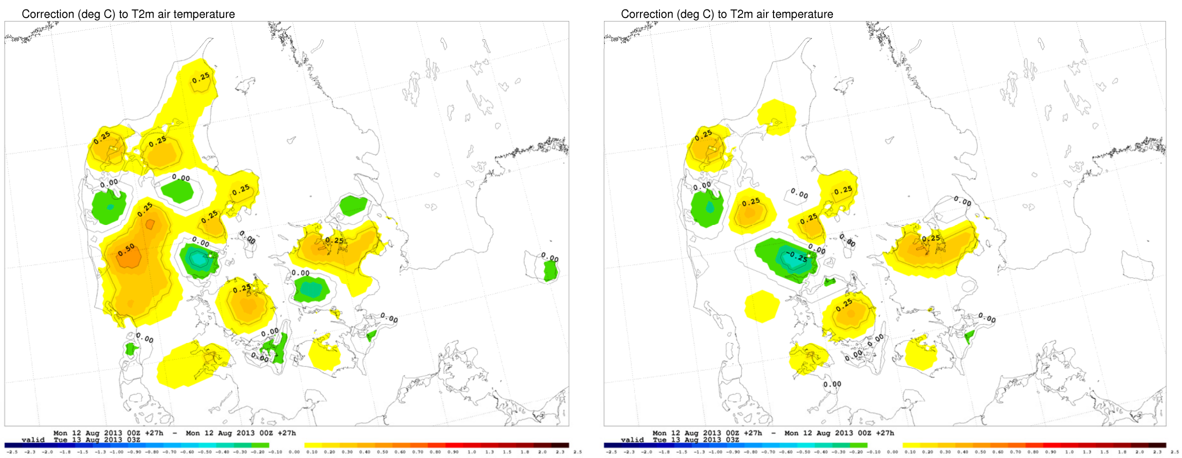
(a) (b)

Figure 3.3.1: *Example of the interpolated 2D field of correction to air temperature based on input data from (a) 60 and (b) 30 synoptical stations.*

In operational part, the number of points used for the interpolation procedure is changing at every model run (at 00, 06, 12 and 18 UTCs) and every forecast length (00-48 hours: 49 forecast lengths), as it is based on outcome from decision-making tree (DMT) procedure. This outcome depends on several specific criteria to decide if the calculated value of the temperature correction should be used further in the interpolation procedure and further to correct the initially predicted (by the NWP model) air temperature. These criteria are the following:

- A - the station will start to be considered if the number of previous days with corrected forecasts will be, at least, 7;
- B - the value of the bias estimated from the previous one-day forecast should be improved;
- C - the averaged value of the bias from the previous multi-days (in our study a 30 day period is selected, although it can be modified as needed) should be improved
- D - at least 50% (from 49 maximally possible) of forecast lengths from previous one- and multi-day runs evaluation should show improvement of temperature forecast after correction had applied;
- E - at least 80% (from 49 maximally possible) of forecast lengths from previous multi-day runs evaluation should show improvement of temperature forecast after correction had been applied.

There are several options in the decision-making: (1) – A+B+C; (2) – A+B+C+D; and (3) – A+B+C+D+E. Moreover, there are also 2 ways for writing output for temperature correction: (I) only corrections for stations satisfying one of the options listed above (i.e. depending on selection - 1, 2, or 3) will be written into output file for further interpolation and the other stations (which do not satisfy options) will be excluded; and (II) both stations with satisfying and not-satisfying listed options will be written, but for “excluded” stations the values of corrections will be replaced with zeros. An example is shown in Figure 3.3.2. Figure 3.3.2a shows spatial distribution of temperature correction field based on interpolation for all stations. Figure 3.3.2b shows the same field but also including “excluded” stations where initial correction value was replaced to 0. As seen, when more data points are used for interpolation, the larger will be the area where non-zero corrections to temperature will be applied.

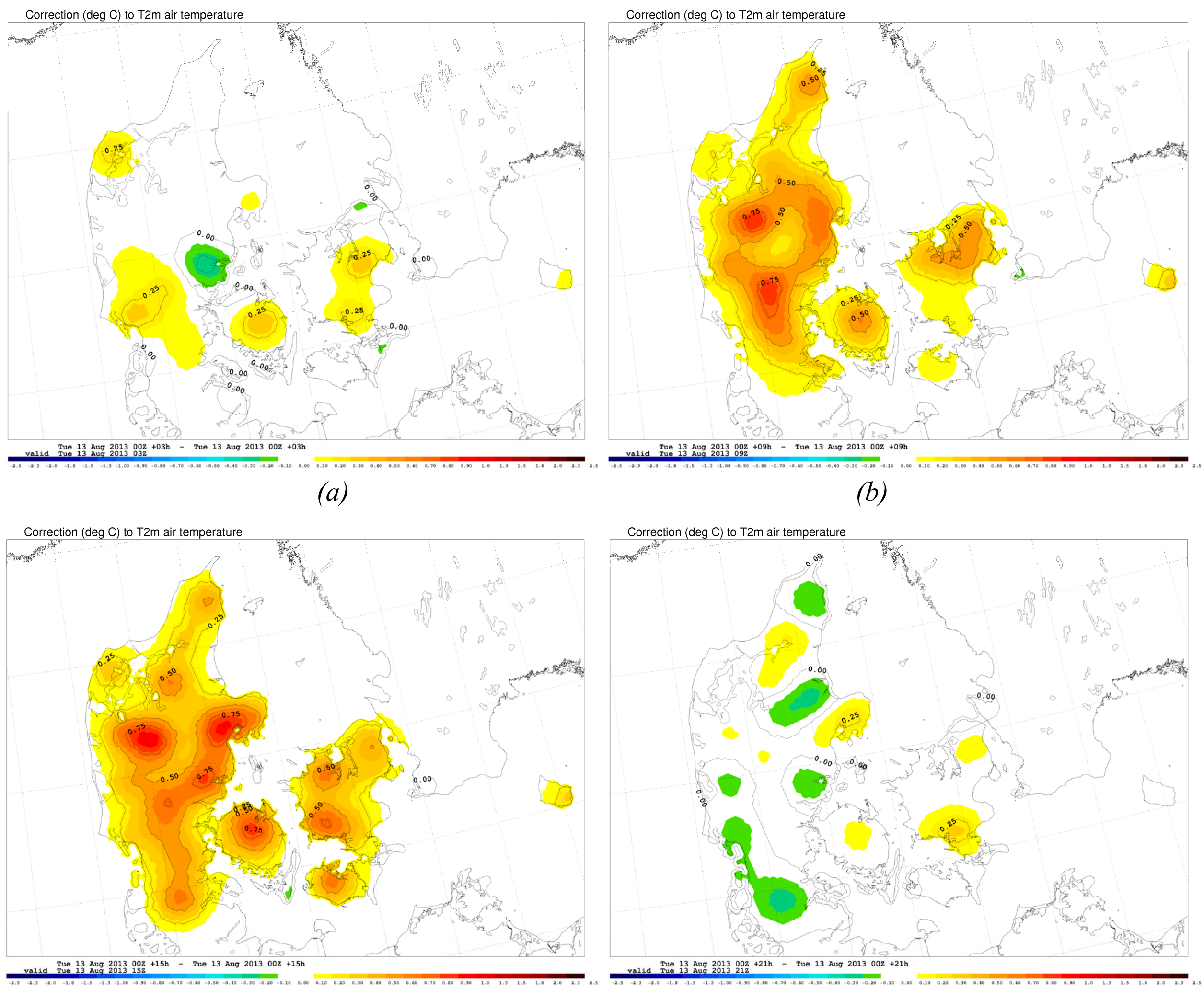


(a) (b)

Figure 3.3.2: *Example of the interpolated 2D field of correction to air temperature based on output from the decision-making tree procedure for the (a) original field: maximum number (when corrections are assigned as valid to all stations) and (b) final field: reduced number (when for non-satisfying stations the corrections are replaced with zeros) of synoptical stations.*

3.3.2. Forecast length

The spatial variability of the temperature correction field as a function of the forecast length (note, there are 49 forecast lengths starting from 00 to 48 hours) is shown as an example in Figure 3.3.3. In general, domination of the sign and magnitude of the correction value has diurnal cycle variability as well. As seen in Figure 3.3.3ad the negative corrections are more frequent in late evening and nighttime compared with morning and daytime hours; and as seen in Figure 3.3.3bc the absolute value of the correction is larger in morning and daytime compared with other times.



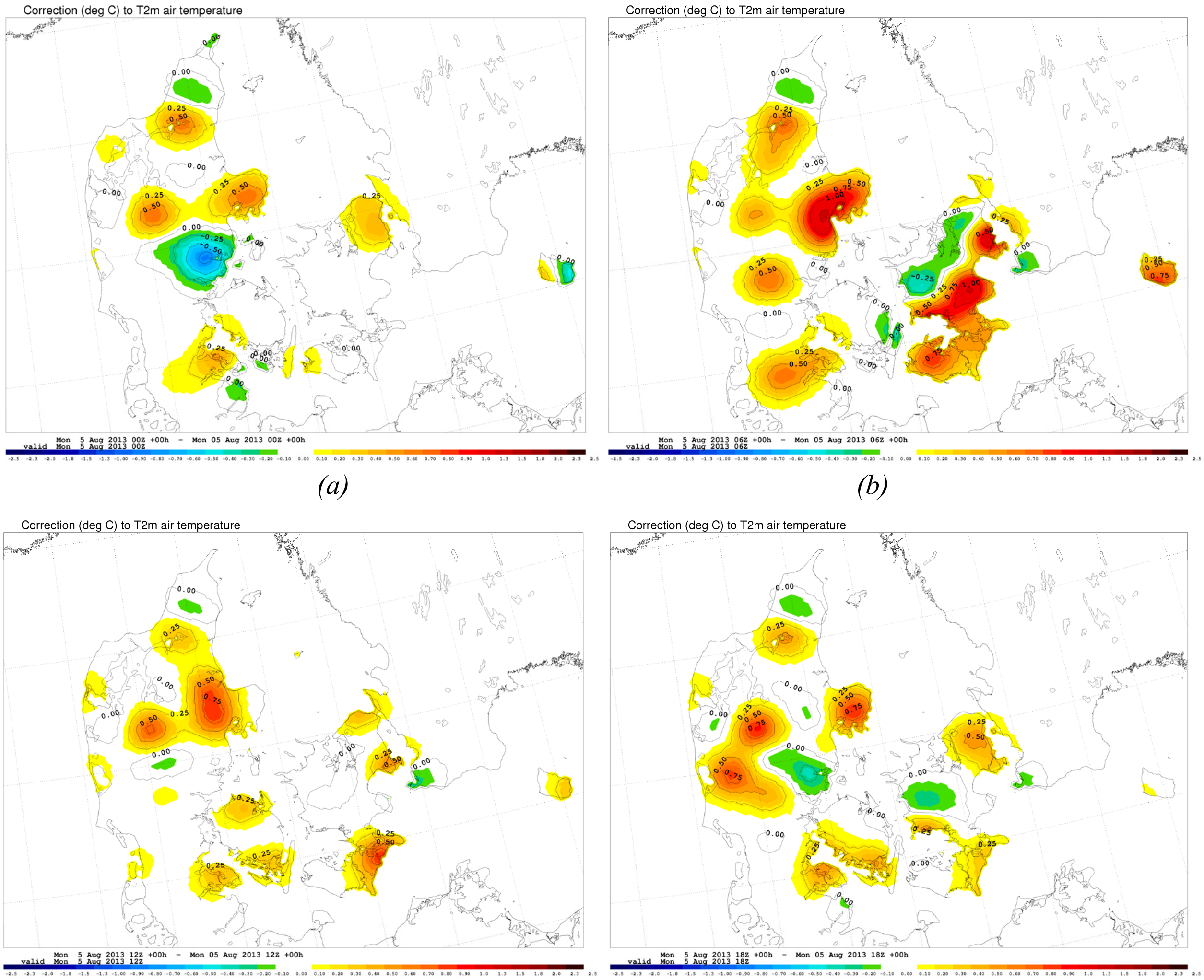
(a) (b)

(c) (d)

Figure 3.3.3: *Example of the interpolated 2D field of correction to air temperature as a function of the forecast length for 13 Aug 2013, 00 UTC plus (a) 03, (b) 09, (c) 15 and (d) 21 hours.*

3.3.3. Model runtime

The spatial variability of the temperature correction field as a function of the model runtime (note, for the HIRLAM-SKA model there are 4 runs daily – at 00, 06, 12, and 18 UTCs) is shown as an example in Figure 3.3.4 for the 00 forecast length time. In general, at 00 hour forecast length there are no large differences for the initial start time runs of the model compared with the runs for different forecast lengths, although for mid-night run at 00 UTC (Figure 3.3.4a) there is more probability of larger (by value) negative corrections to temperature compared with other runs. The positive corrections were more frequent.



(a) (b) (c) (d)

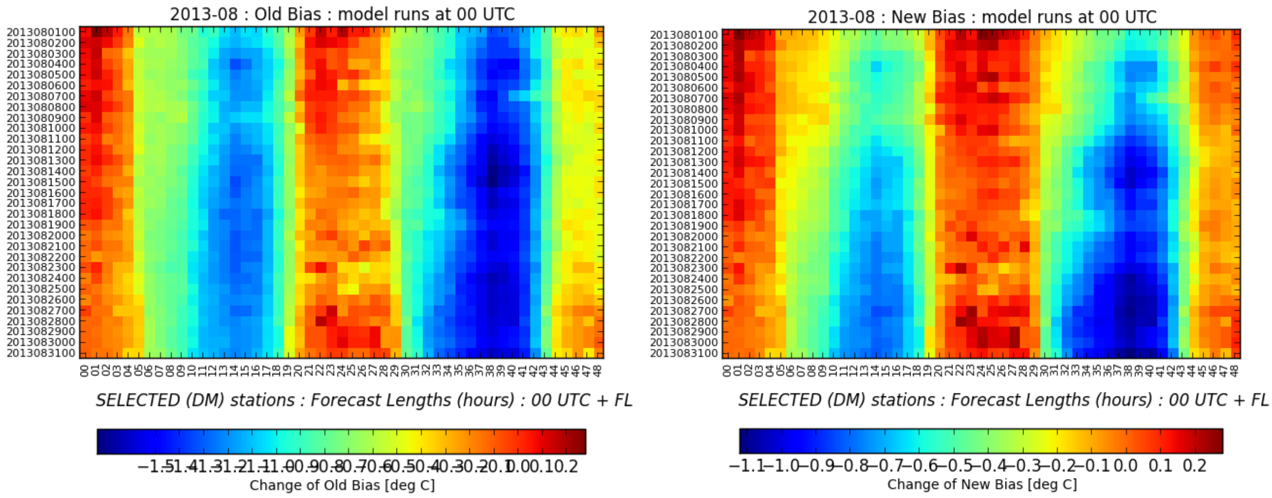
Figure 3.3.4: *Example of the interpolated 2D field of correction to air temperature as a function of the model runtime for 5 Aug 2013 (a) 00, (b) 06, (c) 12 and (d) 18 UTC hours (for forecast length 00 hours).*

3.4. Results of Pre-Operational Runs for City Weather Applications

For the city weather applications, new operationalized procedure for statistical correction of the air temperature forecasts has been elaborated and implemented for the HIRLAM-SKA model runs at 00, 06, 12, and 18 UTCs covering forecast lengths up to 48 hours. It includes segments for extraction of observations and forecast data, assigning these to forecast lengths, statistical correction of temperature, one- and multi-days evaluation, decision-making, interpolation, visualisation and storage/backup. Pre-operational air temperature correction runs were performed for the mainland Denmark (for 60 synoptical stations) since mid-April 2013 and shown good results.
Tests also showed that the CPU time required for the operational procedure is relatively short (less than 15 minutes including a large time spent for interpolation). These also showed that in order to start correction of forecasts there is no need to have a long-term pre-historical data (containing forecasts and observations) and, at least, a couple of weeks will be sufficient when a new observational station is included and added to the forecast point.

The summary of evaluation is given in Figures 3.4.1-3.4.4 on example of August 2013 (HIRLAM-SKA model run at 00 UTC), as well as summarized for other months in Appendix C-D (started to be saved from approximately mid-May 2013) The following evaluation parameters - average, maximum negative and maximum positive corrections to air temperature, values of bias before and after corrections applied, changes/differences in maes and hit-rates before/after correction applied; and percentage of the cases corrected – are shown in figures for each day of the month and each forecast length (00-48 hours). Such plotted results can be averaged over all stations or over selected stations (based on the decision-making outcome) as shown in these figures. Note similar results can be plotted and analyzed for individual stations as well.

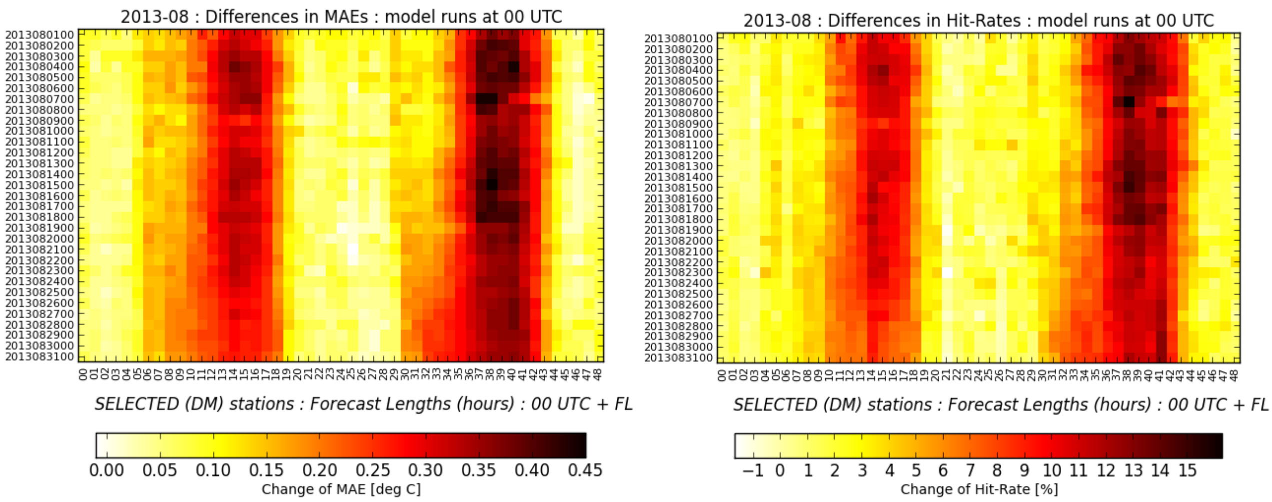
Figure 3.4.1 shows temporal (days vs. forecast lengths) variability of the *T2m (Ta)* bias before (Figure 3.4.1a) and after (Figure 3.4.1b) statistical correction was applied. As seen, the positive bias dominated mostly during late evening – nighttime hours, and the positive bias dominated during daytime. In general, after correction the absolute value of the negative bias became smaller (from -1.7 to -1.1°C) and the occurrence of the positive bias became more frequent.



(a) *(b)*

Figure 3.4.1: *Temporal variability of the air temperature at 2 m biases (a) before and (b) after correction applied (results are averaged over selected synoptical stations following the decision-making evaluation, and based on 1-31 Aug 2013 HIRLAM-SKA model runs at 00 UTCs for forecast lengths 00-48 hours).*

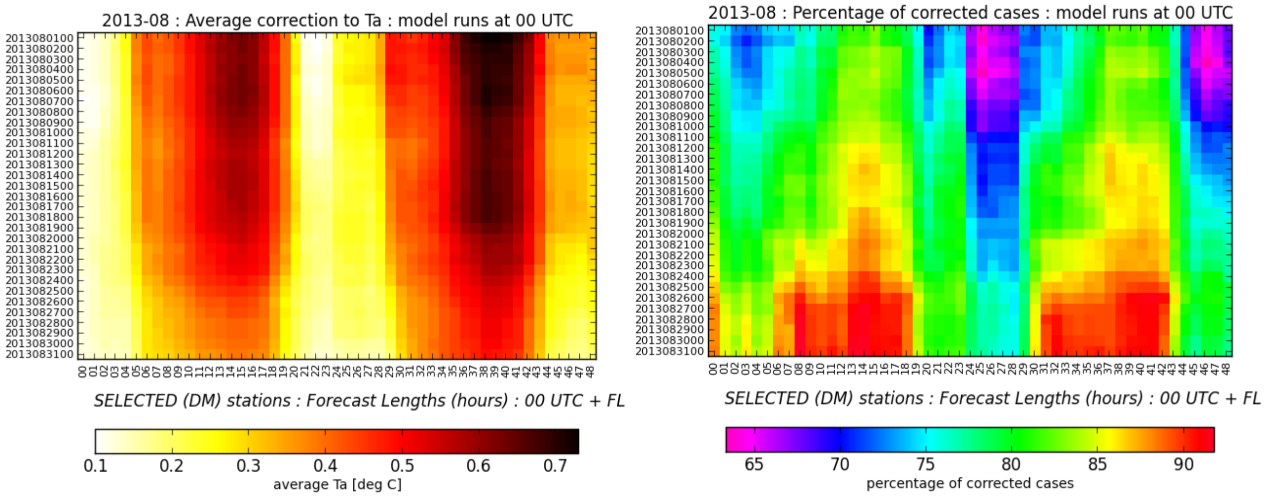
Similarly, the change (difference between before/after correction applied) in mae was larger during day-time hours compared with other hours. The largest improvement in MAE was up to 0.45°C (Figure 3.4.2a). The hit-rate change is larger at longer forecast lengths (on 2nd day of forecast corresponding to daytime hours) on occasion reaching more than 15% of improvement (Figure 3.4.2b).



(a) *(b)*

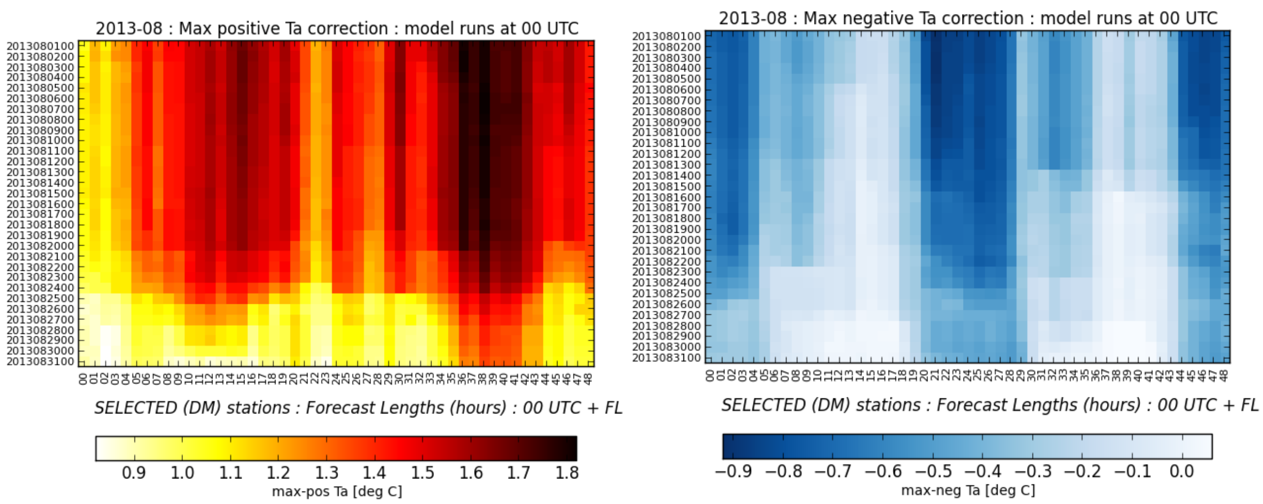
Figure 3.4.2: *Temporal variability of the air temperature at 2 m (a) differences in mae and (b) differences in hit-rate after the correction applied (results are averaged over selected synoptical stations following the decision-making evaluation, and based on 1-31 Aug 2013 HIRLAM-SKA model runs at 00 UTCs for forecast lengths 00-48 hours).*

Figure 3.4.3a shows variability for the average correction to the air temperature, which is also larger during daytime compared with nighttime hours. On average, this correction could be of more than 0.7°C, and moreover, the positive correction on average has dominated. As for the maximum (positive – Figure 3.4.4a, and negative - Figure 3.4.4b) corrections, these could be to about +1.8°C and -0.9°C, respectively. As seen, the maximum positive correction was larger in daytime hours, and the maximum negative correction was more frequent in nighttime hours.
Figure 3.4.3b shows variability in percentage of the cases corrected with improvement of the bias value (e.g. after correction the absolute value of the bias became smaller). It is more variable, but in general, more cases were corrected during daytime hours compared with nighttime, where it could be up to more than 90% vs. about 70±5%, respectively.



(a) *(b)*

Figure 3.4.3: *Temporal variability of the (a) average correction to air temperature at 2 m and (b) percentage of cases with improved bias after correction applied (results are averaged over selected synoptical stations following the decision-making evaluation, and based on 1-31 Aug 2013 HIRLAM-SKA model runs at 00 UTCs for forecast lengths 00-48 hours).*



(a) *(b)*

Figure 3.4.4: *Temporal variability of the maximum - (a) positive and (b) negative - correction to air temperature at 2 m (results are averaged over selected synoptical stations following the decision-making evaluation, and based on 1-31 Aug 2013 HIRLAM-SKA model runs at 00 UTCs for forecast lengths 00-48 hours).*

Note that similar pattern in variability on a diurnal cycle was also obtained for other HIRLAM-SKA model runs (see Appendix C – for 06, 12, 18 UTC runs on example of Aug 2013) as well as other months (half of May, Jun, Jul, Sep on example of 00 UTC run; note that Aug is shown in this section) evaluated (see Appendix D – for bias, mae, hit-rate, and percentage of corrected cases).

A summary on changes in cases with improved/not-improved bias, mae and hit-rate values (by **averaging over all forecast lengths**) is shown in Table 3.4.1 for the studied period from 15 May till 29 Sep 2013 inclusive (note that from 1st Aug 2013, evaluation was also done for other model runs, e.g. at 06 and 18 UTCs). Note that total number of analyzed changes (overall improved/not-improved bias/mae/hit-rate) for each model run in each month is equal to a number of days (when model run) within a month multiplied by number of forecast lengths (for example, in May 2013 only 17 days were counted, which gives 17 x 49 = 833). As seen from table, taking into account averaging over all forecast lengths, for each HIRLAM-SKA model run the overall improvement of three mentioned verification parameters was observed when the temperature correction was applied.

Table 3.4.1: *Month-to-month (May-Sep 2013) variability of the bias, mae and hit-rate for the air temperature at 2 m (results are based following the decision-making evaluation) for different HIRLAM-SKA model runs and averaged over all forecast lengths (00-48 hours).*

Month of 2013	UTC run	Improvement						Total # cases
		Bias		Mae		Hit-Rate		
		yes	no	yes	no	yes	no	
May	**00**	833	0	814	19	818	15	833
	12	830	3	831	2	832	1	833
Jun	**00**	1336	134	1404	66	1416	54	1470
	12	1323	147	815	655	1005	465	1470
Jul	**00**	1362	157	1511	8	1488	31	1519
	12	1367	152	1519	0	1490	29	1519
Aug	**00**	1385	134	1517	2	1491	28	1519
	06	1308	211	1514	5	1443	76	1519
	12	1402	117	1518	1	1495	24	1519
	18	1269	250	1517	2	1447	72	1519
Sep	**00**	1122	299	1409	12	1261	160	1421
	06	983	438	1404	17	1258	163	1421
	12	1200	221	1420	1	1301	120	1421
	18	1015	406	1417	4	1213	208	1421

3.5. Additional Tuning for Selected Stations with high MAEs

Although the chosen method works practically for all stations, there are still some stations where air temperatures can be additionally corrected. Especially, stations having the largest values of biases (by module) and elevated (i.e. "bad") values of maes, are of the major concern. Among the Danish synoptical stations there are about 10 stations having maes higher than 1°C. For such stations, an additional tuning or additional improvement/correction of the original temperature forecast can be done.

An example (for 00 run + 6 h forecast length) of such tuning for one of the synoptical stations (WMO-06030) having "bad" or too high maes (i.e. more than 1°C) is shown in Figure 3.5.1 and summarized in Table 3.5.1. For tuning, several tests to calculate final values of temperature correction were performed. Note, the focus is on cases having consequently in a pre-history (i.e. -1d, -2d,

-3d, …, -*n*d) considering at least 3 days period with values of bias higher than ±1°C. The final value of air temperature correction - $T2m_{cor/final}$ - for the day of forecast *[0d]* is calculated following:

$$T2m_{cor/final}\,[0d] = (\left| avgNbt2m \right| + T2m_{cor/initial}\,[0d]) \,/\, TDF$$

where:

$T2m_{cor/initial}\,[0d]$ – initial value of air temperature correction (in °C);
TDF – tuning division factors (*TDF* = 1, 2, 3, 4; dimensionless);
avgNbt2m – averaged bias over *n*-days in pre-history (in °C);

and

$$avgNbt2m = (Nbt2m[-1d] + Nbt2m[-2d] + Nbt2m[-3d] + \ldots + Nbt2m[-nd]) \,/\, n$$

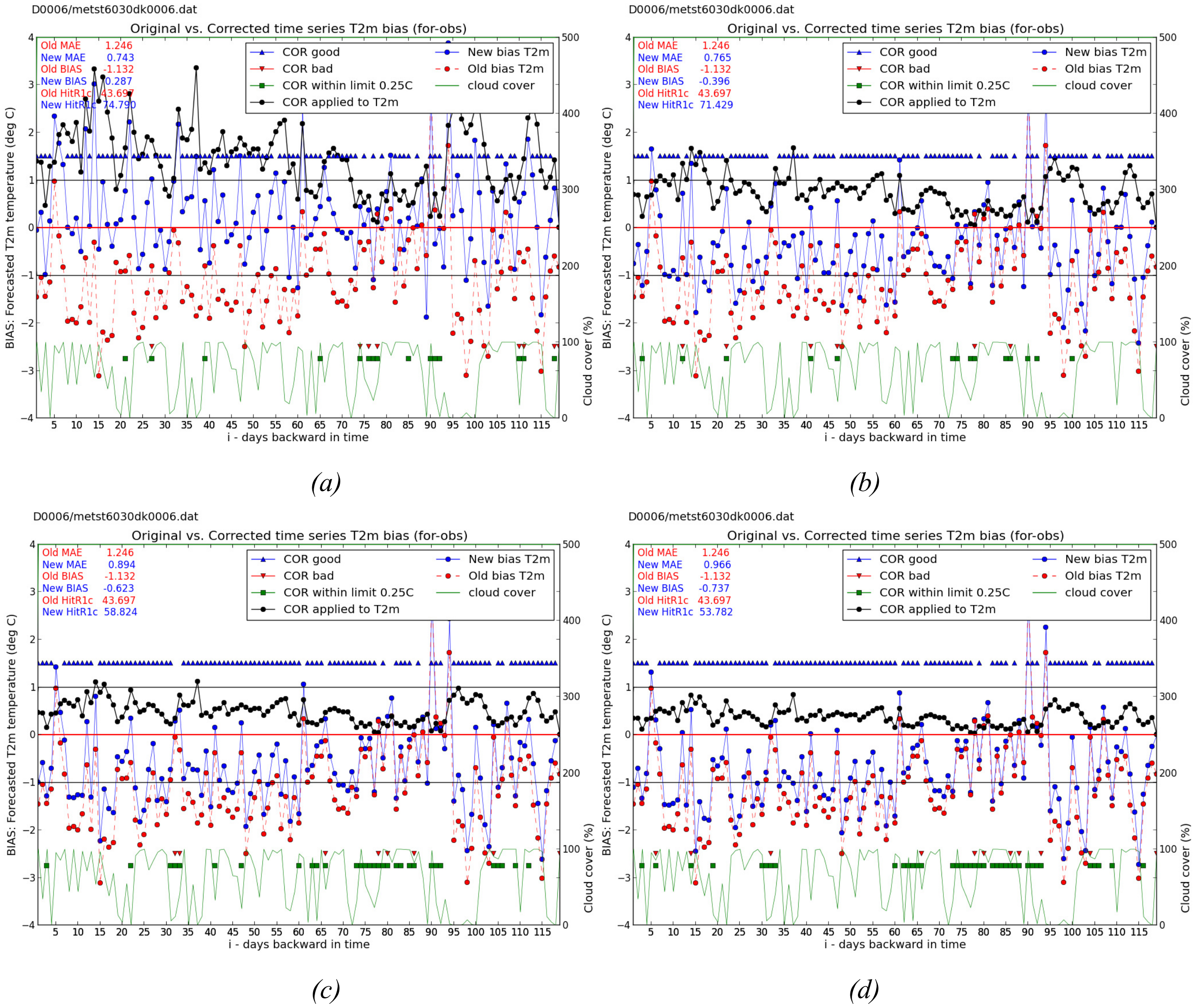


Figure 3.5.1: *Tuning (at different tuning division factors – TDFs – (a) 1, (b) 2, (c) 3 and (d) 4) of the air temperature correction for the Danish synoptical station WMO-06030 based on data for 1 May – 31 August 2012, 00 run + 6 h forecast length /days are shown in reverse order on a horizontal time axis; see description of legend in Ch. 3.2.1/.*

As seen in Figure 3.5.1, due to tuning option variability in changes leading to improvement of the temperature forecasts are well pronounced. In particular, at lower TDFs the temperature correction could be larger by an absolute value, as seen in Figure 3.5.1a, where *T2m* correction was frequently above 1°C or even more than 2°C compared with higher TDFs (Figure 3.5.1d). Moreover, at higher TDFs the correction will be more frequently within the limits of ±0.25°C as well as changes in bias will be smaller (Figure 3.5.1d) compared with lower TDFs (Figure 3.5.1a).

Table 3.5.1: *Summary / Statistics (number and percentage of cases, old and new biases, maes, and hit-rates) on tuning of air temperature correction for the Danish synoptical station WMO-06030 based on data for 1 May – 31 August 2012, 00 run - 6h forecast length.*

Tuning division factor, TDF **Evaluated parameters**	**1**	**2**	**3**	**4**
total #/% forecasts, CORtot	119 / 100	119 / 100	119 / 100	119 / 100
#/% corrected forecasts, CORyes	87 / 73.11	97 / 81.51	102 / 85.71	103 / 86.55
#/% not corr. forecasts, CORnot	32 / 26.89	22 / 18.49	17 / 14.29	16 / 13.45
#/% corr. forecasts within uncertainty limit (±0.25°C), CORzer	8 / 6.72	6 / 5.04	7 / 5.88	10 / 8.40
total #/% of corr. forecasts, COR-ok	95 / 79.83	103 / 86.55	109 / 91.60	113 / 94.96
Old (before correction) T2m bias	avg: −1.13 max: 3.00 min: −3.11	avg: −1.13 max: 3.00 min: −3.11	avg: −1.13 max: 3.00 min: −3.11	avg: −1.13 max: 3.00 min: −3.11
New (after correction) T2m bias	avg: 0.29 max: 3.87 min: −1.88	avg: −0.40 max: 2.88 min: −2.42	avg: −0.62 max: 2.92 min: −2.62	avg: −0.74 max: 2.94 min: −2.72
Value of correction to T2m	avg: 1.47 max: 3.36 min: 0.11	avg: 0.73 max: 1.68 min: 0.06	avg: 0.49 max: 1.12 min: 0.04	avg: 0.37 max: 0.84 min: 0.03
Old mean T2m mae	1.246	1.246	1.246	1.246
New mean T2m mae	0.743	0.765	0.894	0.966
#/% Old hit-rate (±1°C)	52 / 43.7	52 / 43.7	52 / 43.7	52 / 43.7
#/% New hit-rate (±1°C)	89 / 74.8	85 / 71.4	70 / 58.8	70 / 53.8

As seen in Table 3.5.1, the largest total number and percentage of corrected forecasts (cases) is the highest for TDF=4 (from 119 cases only 113 were corrected or almost 95%). It is the smallest for TDF=1 corresponding to 95 cases (or about 80%). Before correction, on average, the "old" bias was about −1.13°C and it ranged within an interval from −3.00°C to +3.11°C. After tuning the bias had improved; and this improvement was the largest for TDF=1 (bias became +0.29°C; Figure 3.5.1a) and the smallest for TDF=4 (with bias −0.74°C; Figure 3.5.1d). Similarly, an average value of correction to air temperature is the largest for TDF=1 compared with TDF=4: 1.47°C vs 0.37°C, respectively. Also, the largest improvement in mae was also shown for lower values of TDF with maximum (more than half of degree) improvement from original mae of 1.246°C to 0.743°C. The old hit-rate for air temperatures (considering biases within ±1°C) was about 44%, and it can be improved after tuning up to 75%.

Note, for synoptical stations with maes below 1°C, the higher TDFs provide larger improvements for biases, maes, and hit-rates. As seen here, this is opposite for selected (e.g. for smaller number of stations from the total list of chosen in this study) stations with maes higher than 1°C. As it is found, for such stations improvements will be larger for smaller TDFs values.

Moreover, additional analyses for improvement of air temperature forecasts due to reducing bias/mae and extra-adjustment of temperature correction value was performed by *Gilet et al., (2013)*. Exploratory data analysis of the dataset (which included observations and forecasts of selected meteorological parameters and covered a period: 1 May – 31 Aug 2012) for different forecast lengths (in particular: 6, 12, 18, and 24 hours) showed existence of similar patterns and behavior in temperature errors. It allowed, following approach suggested by *Mahura et al (2009)* for pollen modeling applications, for a group of stations having large biases and the largest ("worst") maes to develop and test a parameterization for additional correction of the error temperature. See more details in *Gilet et al., (2013)*. The developed parameterization can be further implemented for operational tasks in numerical weather prediction.

4. Conclusions

Based on analysis of long-term time-series of meteorological observations and forecasts (from HIRLAM-SKA model / RCM model outputs) the method for statistical correction of air /road surface temperatures forecasts was developed and tested (for May-Aug 2012 / Oct 2012 - Mar 2013).

The developed method is based mostly on forecasted meteorological parameters with a minimal inclusion of observations (covering only pre-history period). The adjustment of the *T2m/Trs* temperatures is purely based on forecasted meteorological parameters. The method is model independent, e.g. it can be applied for temperature correction with other types of models (for example, for HARMONIE) as well as with other horizontal resolutions (for example, 5-15 km). It is relatively fast due to application of the singular value decomposition method (SVD) for matrix solution to find coefficients. As it was shown (on example of synoptical stations) there is always a possibility for additional improvement due to extra tuning of the temperature forecasts for some locations (stations), and in particular, where for example, the MAEs are generally higher compared with others. *Gilet et al. (2013)* performed additional study, evaluated data and suggested parameterization for extra improvement of the air temperature forecasts.

For the city weather applications, new operationalized procedure for statistical correction of the air temperature forecasts has been elaborated and implemented for the HIRLAM-SKA model runs at 00, 06, 12, and 18 UTCs covering forecast lengths up to 48 hours. It includes segments for extraction of observations and forecast data, assigning these to forecast lengths, statistical correction of temperature, one-&multi-days evaluation, decision-making, interpolation, visualisation and storage/backup. Pre-operational air temperature correction runs were performed for Denmark (for 60 synoptical stations) since mid-April 2013 and shown good results.
Tests also showed that the CPU time required for the operational procedure is relatively short (less than 15 minutes including a large time spent for interpolation). These also showed that in order to start correction of forecasts there is no need to have a long-term pre-historical data (containing forecasts and observations) and, at least, a couple of weeks will be sufficient when a new observational station is included and added to the forecast point.

Note for the road weather application, the operationalization of the statistical correction of the road surface temperature forecasts (for the RWM system daily hourly runs covering forecast length up to 5 hours ahead) for the Danish road network (for about 400 road stations) is under development and testing.

The method can also be applied for correction of the dew point temperature and wind speed (as a part of observations/ forecasts at synoptical stations), where these both meteorological parameters are parts of the proposed system of equations. The evaluation of the method performance for improvement of the wind speed forecasts is planned as well, with considering possibilities for the wind direction (which is more complex due to multi-modal types of such data distribution) improvements. The method worked for Denmark (tested for 60 synoptical and 395 road stations), and hence, it can be also applied for any geographical point within this domain, as it was done through interpolation to about 100 cities' locations (for ***byvejr forecasts***). Moreover, we can assume that the same method can be used in other geographical areas. The evaluation for other domains is in plans, with a focus on Greenland and Nordic countries. In addition, a similar approach might be also tested for statistical correction of concentrations of chemical species, and here a log-transformation to data can be used, but such approach will require additional evaluation.

5. Acknowledgments

The authors are thankful to colleagues from the CMM section of DMI and Roman Nuterman (now at University of Copenhagen, UoC) for constructive discussions and comments.
The DMI CRAY-XT5 supercomputing facilities were used extensively in this study, as well as observational datasets for the Danish synoptical (provided by DMI) and road (provided by the Danish Road Directorate, DRD) stations.
Thanks to DMI's Computer Department for advice and assistance.

6. References

Gilet N., Mahura A., Petersen C., Sass B. (**2013**): Analysis and Improvement of Statistical Error in Forecasted Air Temperature. *DMI Scientific Report 13-03, ISBN: 978-87-7478-646-7, 39p,* www.dmi.dk/dmi/sr13-03.pdf

Mahura A., Petersen C., Sass B., Gilet N. (**2013**): Statistical Correction of Air Temperature Forecasts for City and Road Weather Applications. *DMI Scientific Report 13-02, ISBN: 978-87-7478-645-0, 49p,* www.dmi.dk/dmi/sr13-02.pdf

Mahura A., Baklanov A., Korsholm U. (**2009**). Parameterization of the birch pollen diurnal cycle. *Aerobiologia 25, 203-208.*

Mahura A., Petersen C., Sass B.H. (**2009**): Road Weather Verification: Revised Approach, DMI Sci. Report 09-02, ISBN: 978-87-7478-587-3, 34p, www.dmi.dk/dmi/sr09-02.pdf

Mahura A., Petersen C., Sattler K., Sass B.H., Pedersen T. (**2010**): High Resolution Physiographic Data for Fine-Scale Road Weather Forecasting, *DMI Sci. Report 10-05, ISBN: 978-87-7478-605-4, 30p,* www.dmi.dk/dmi/sr10-05.pdf

Mahura A., Petersen C., Sass B. (**2012**): Approaches to Statistical Correction of the Road Weather Model Forecasts. *DMI Scientific Report 12-05, ISBN: 978-87-7478-643-6, 33p,* www.dmi.dk/dmi/sr12-05.pdf

Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T. (**1992**): Singular Value Decomposition. *In Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed. Cambridge, England: Cambridge University Press*, pp. 51-63.

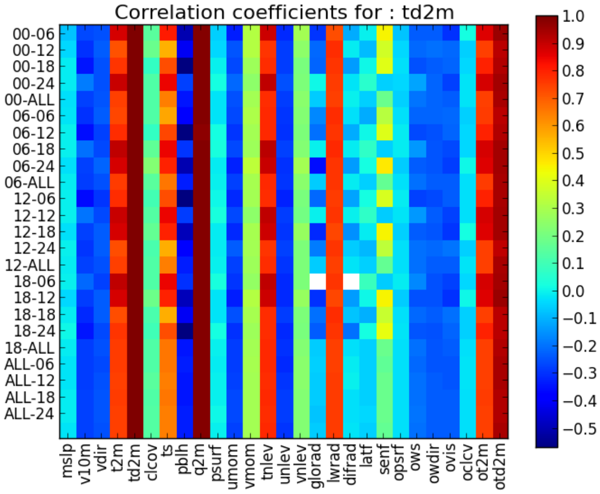
Golub G.H., Van Loan C.F. (**1996**): The Singular Value Decomposition. *In Matrix Computations, 3rd ed. Baltimore, MD: Johns Hopkins University Press,* pp. 70-73.

Sass B.H. (**1992**): A Numerical Model for Prediction of Road Temperature and Ice. *Journal of Applied Meteorology,* 31, pp. 1499-1506.

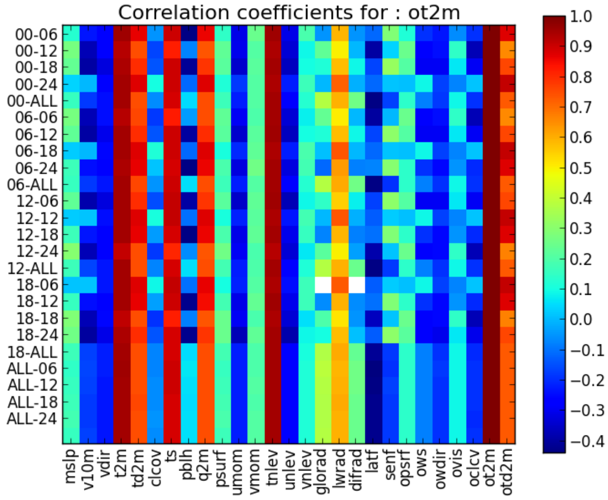
Sass B.H. (**1997**): A Numerical Forecasting System for the Prediction of Slippery Roads. *Journal of Applied Meteorology,* 36, pp. 801-817.

Undén P., L. Rontu, H. Järvinen, P. Lynch, J. Calvo, G. Cats, J. Cuhart, K. Eerola, C. Fortelius, J.A. Garcia-Moya, C. Jones, G. Lenderlink, A. McDonald, R. McGrath, B. Navascues, N. W. Nielsen, V. Degaard, E. Rodriguez, M. Rummukainen, R. Room, K. Sattler, B.H. Sass, H. Savijarvi, B.W. Schreur, R. Sigg, H. The, A. Tijm (**2002**): HIRLAM-5 Scientific Documentation. *December 2002, HIRLAM-5 Project Report, SMHI,* 146 p.

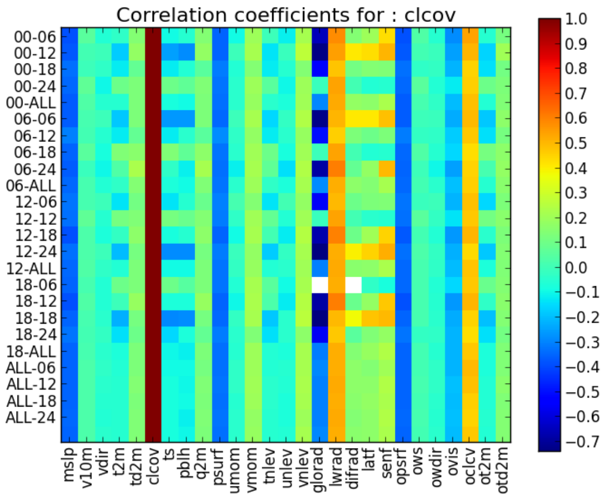
Appendix A: Correlation coefficients between meteorological parameters vs. model runs and forecast lengths



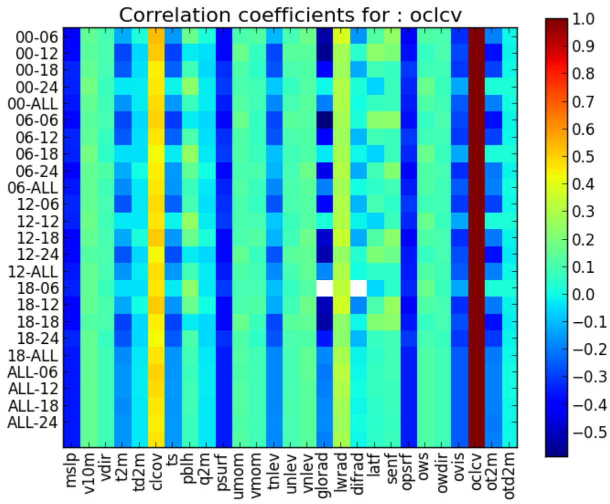
(a)



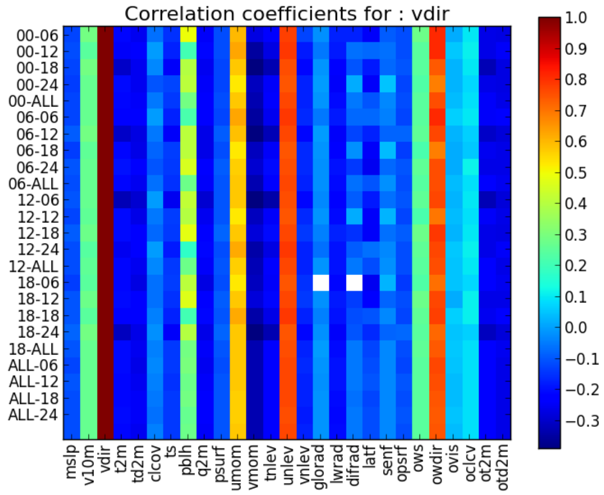
(b)



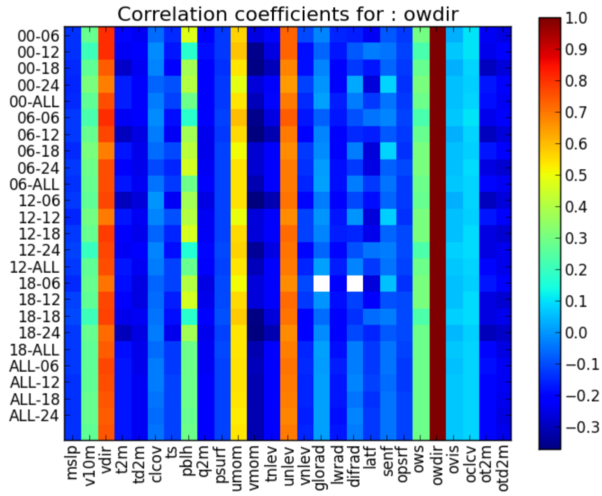
(c)



(d)



(e)



(f)

***Figure A1**: Variability of the correlation coefficients between forecasted (ace) and observed (bdf) - (a) dew point air temperature, (b) air temperature, (cd) cloud cover, (ef) wind direction - vs. other meteorological (both forecasted and observed) parameters (depicted on horizontal axis) as a function of HIRLAM-SKA model runs and forecast lengths (depicted on vertical axis).*

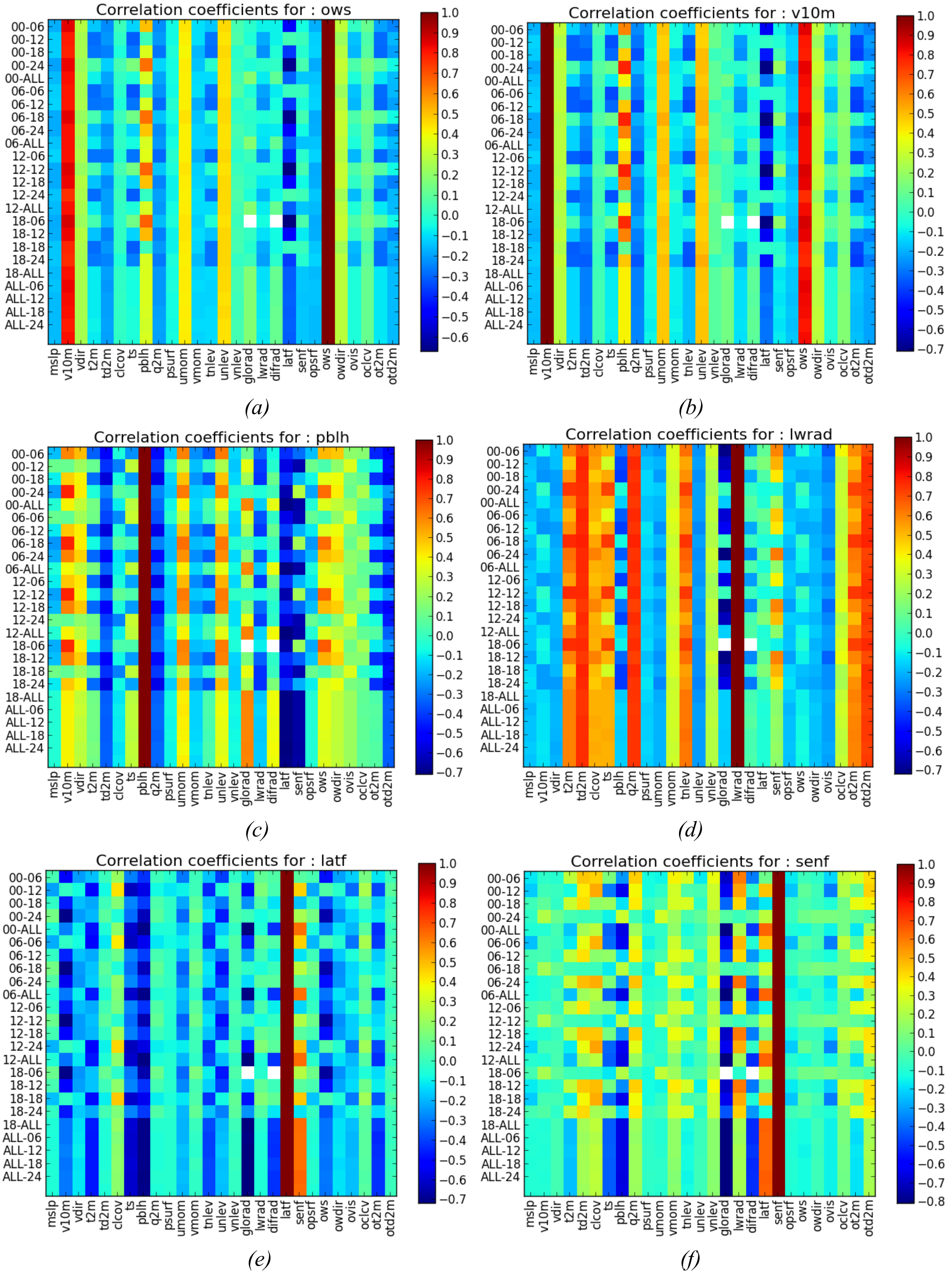


***Figure A2**: Variability of the correlation coefficients between forecasted (bcdef) and observed (a) - (ab) wind speed, (c) boundary layer height, (d) long-wave radiation, (ef) heat fluxes /latent & sensible/ - vs. other meteorological (both forecasted and observed) parameters (depicted on horizontal axis) as a function of HIRLAM-SKA model runs and forecast lengths (depicted on vertical axis).*

Appendix B: Statistical corrections for air temperature forecasts for different forecast lengths

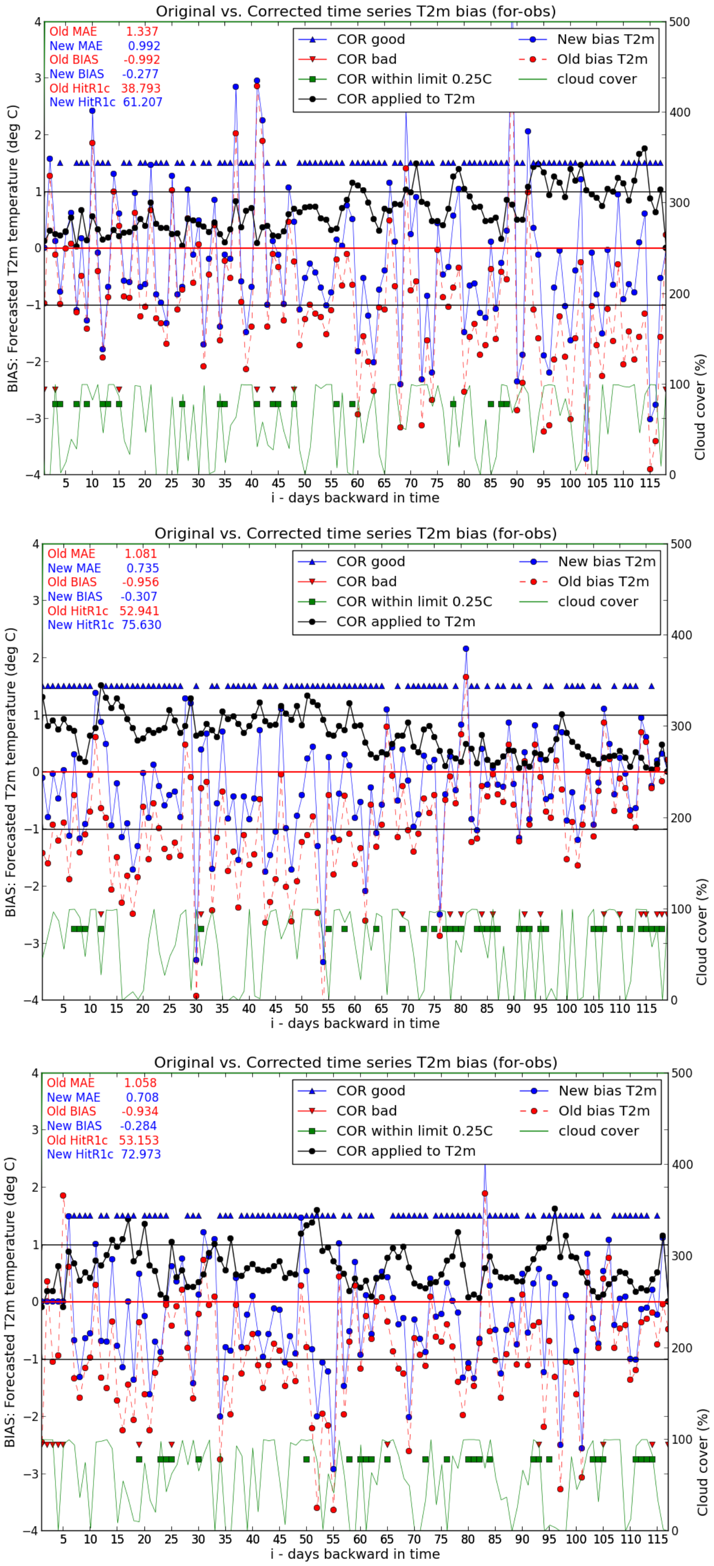
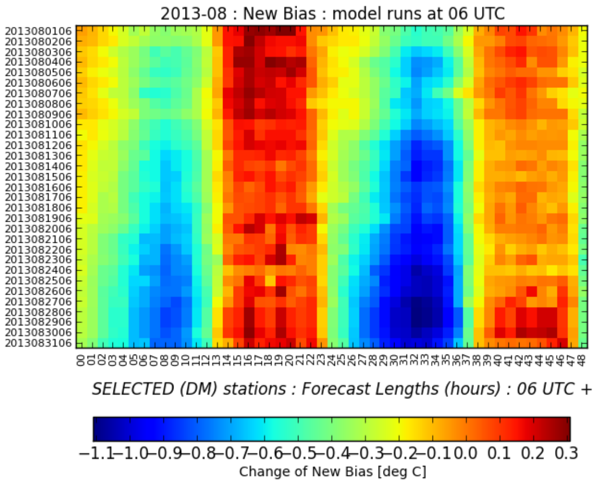
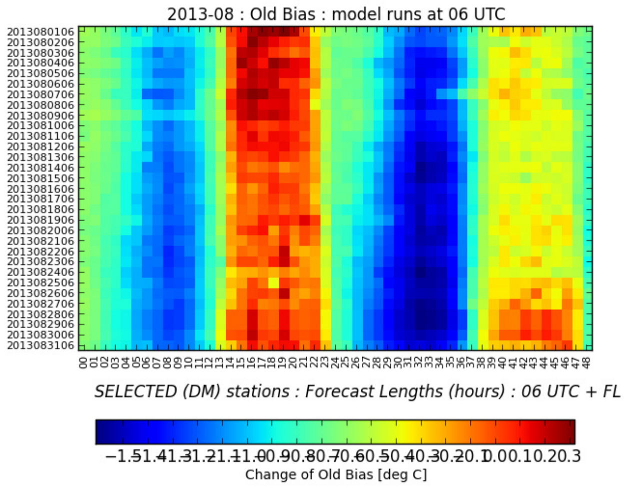


***Figure B**: Results of statistical correction to air temperature for the synoptical station WMO-06030 for the HIRLAM-SKA model run initiated 00 UTC for: (top) 12, (middle) 18, and (bottom) 24 hour forecast lengths /period covered 1 May – 31 Aug 2012; days are in a reverse order on a horizontal time axis/.*

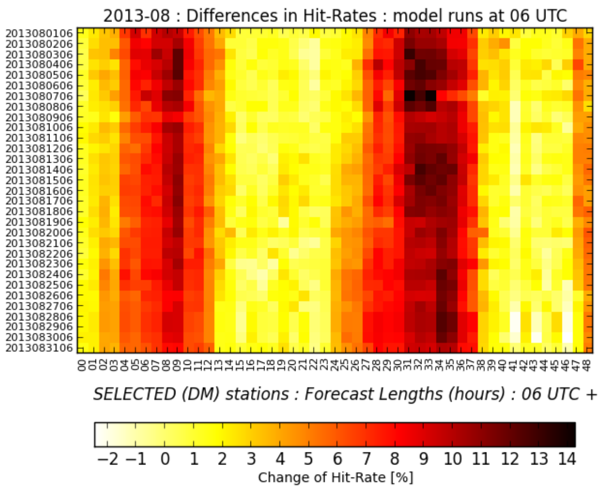
Appendix C: Changes in air temperature's bias, mae, and hit-rate for different HIRLAM-SKA model runs during Aug 2013



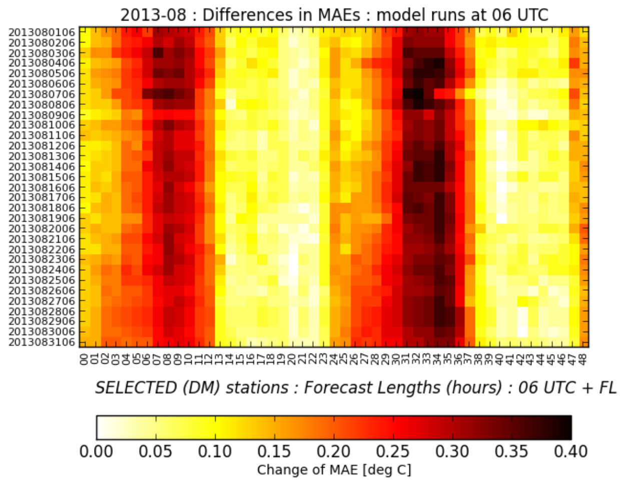
(a)



(b)



(c)



(d)

***Figure C1**: Temporal variability of the air temperature at 2 m - (a) bias after and (b) before correction applied, and (c) differences in hit-rate and (c) differences in mae after the correction applied /results are averaged over selected synoptical stations following the decision-making evaluation, and based on 1-31 Aug 2013 HIRLAM-SKA model runs at **06 UTCs** for forecast lengths 00-48 hours/.*

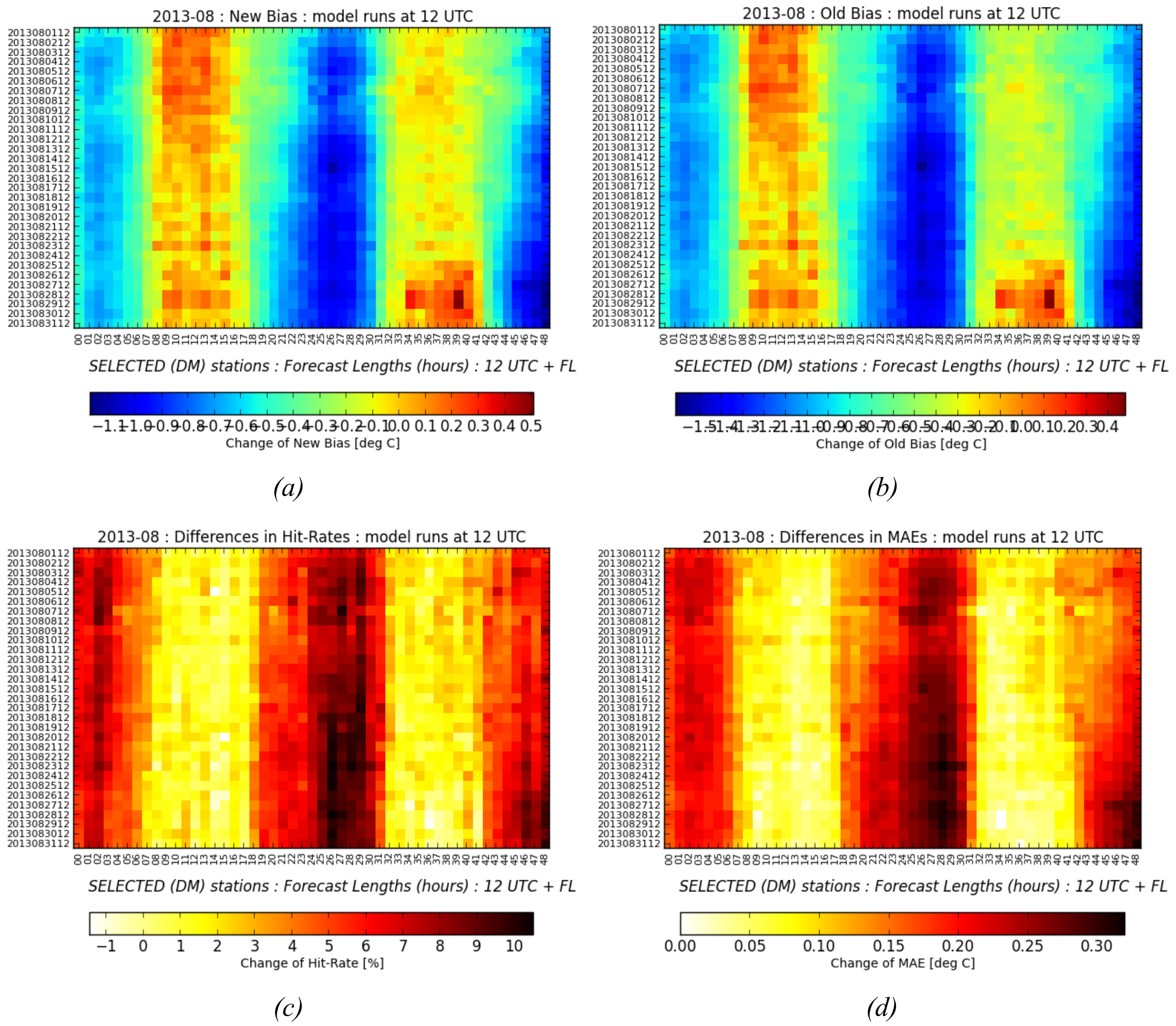
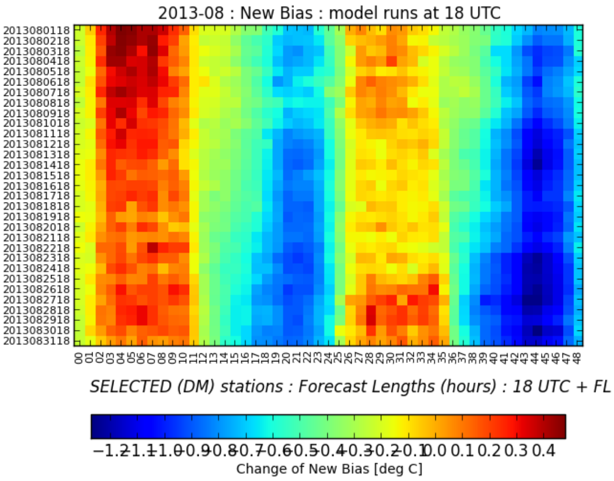
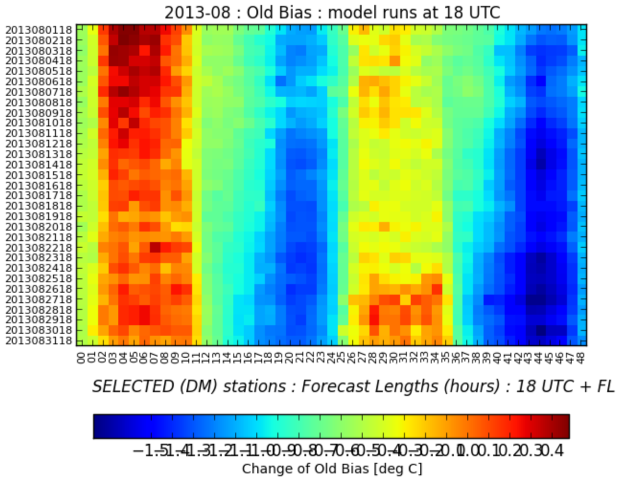


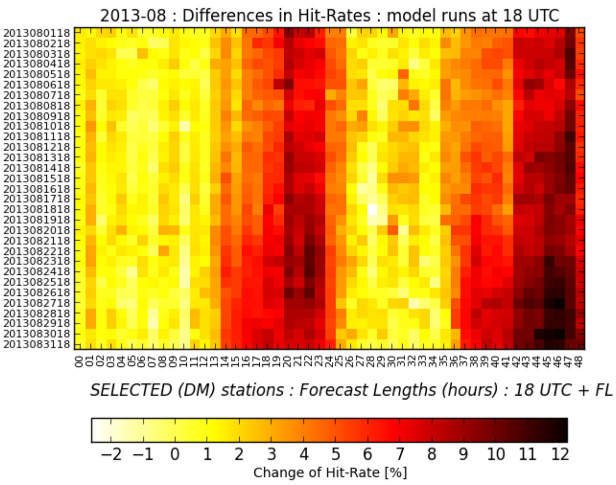
***Figure C2**: Temporal variability of the air temperature at 2 m - (a) bias after and (b) before correction applied, and (c) differences in hit-rate and (c) differences in mae after the correction applied /results are averaged over selected synoptical stations following the decision-making evaluation, and based on 1-31 Aug 2013 HIRLAM-SKA model runs at **12 UTCs** for forecast lengths 00-48 hours/.*



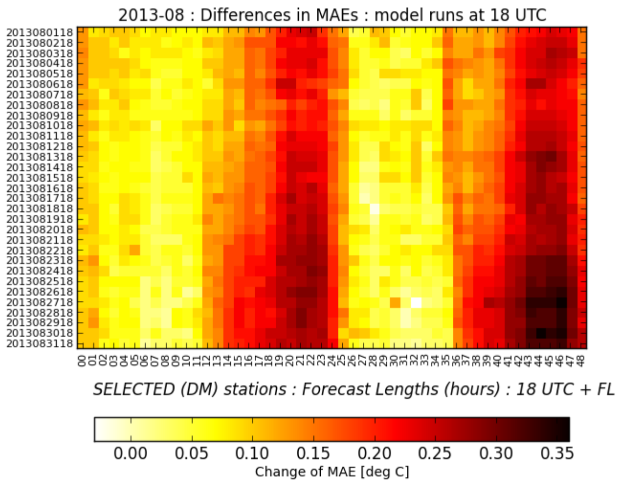
(a)



(b)



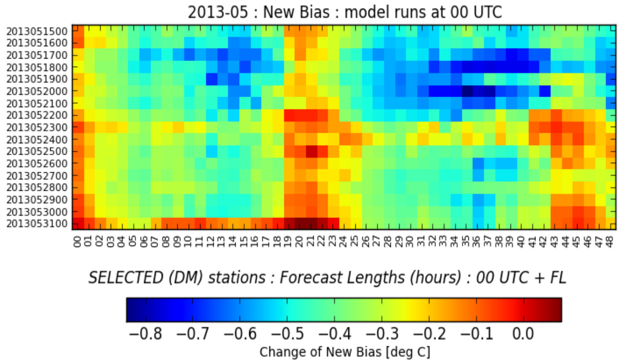
(c)

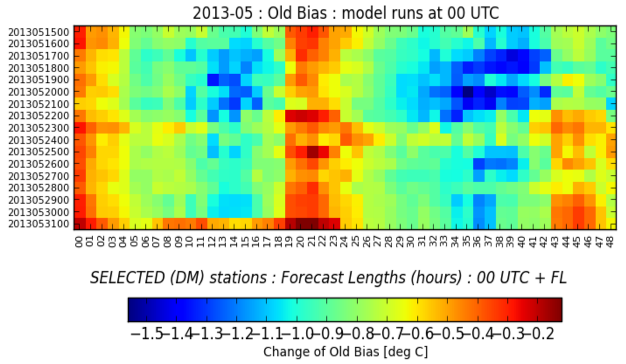


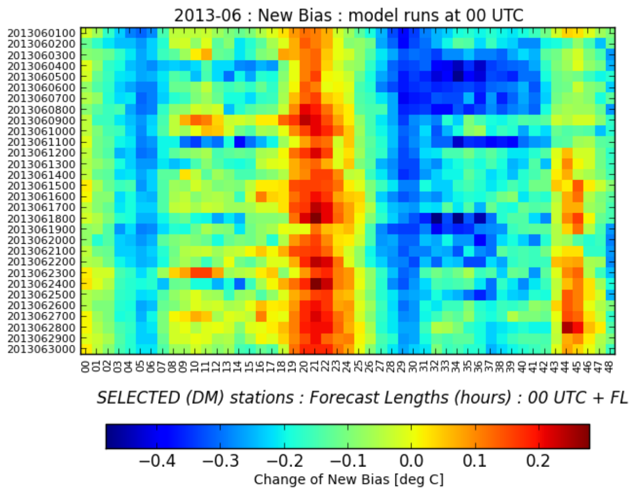
(d)

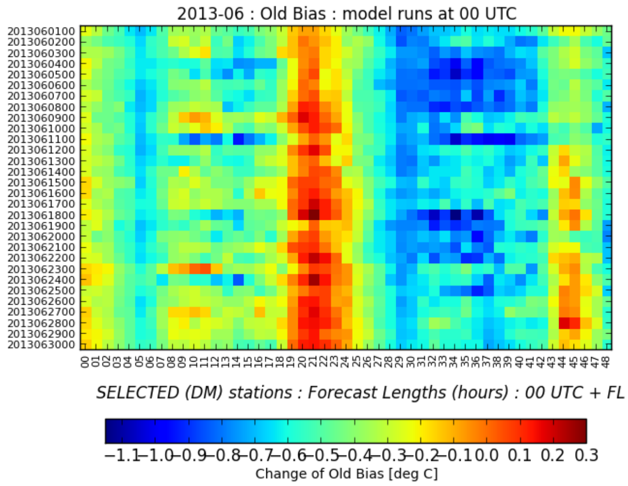
Figure C3*: Temporal variability of the air temperature at 2 m - (a) bias after and (b) before correction applied, and (c) differences in hit-rate and (c) differences in mae after the correction applied /results are averaged over selected synoptical stations following the decision-making evaluation, and based on 1-31 Aug 2013 HIRLAM-SKA model runs at* ***18 UTCs*** *for forecast lengths 00-48 hours/.*

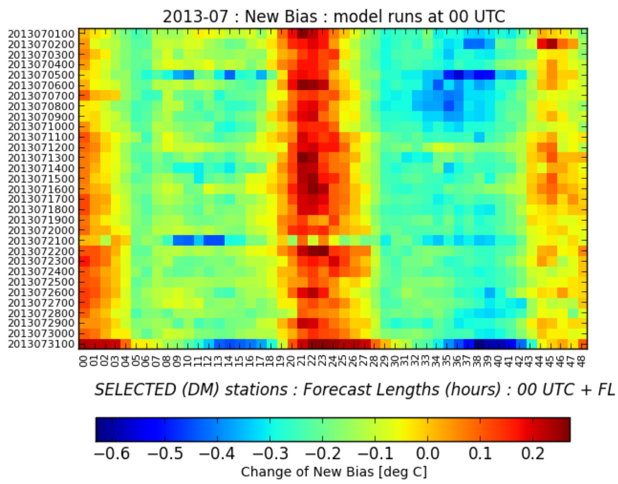
Appendix D: Monthly changes in air temperature's bias, mae, hit-rate, and percentage of corrected cases for HIRLAM-SKA model runs at 00 UTC

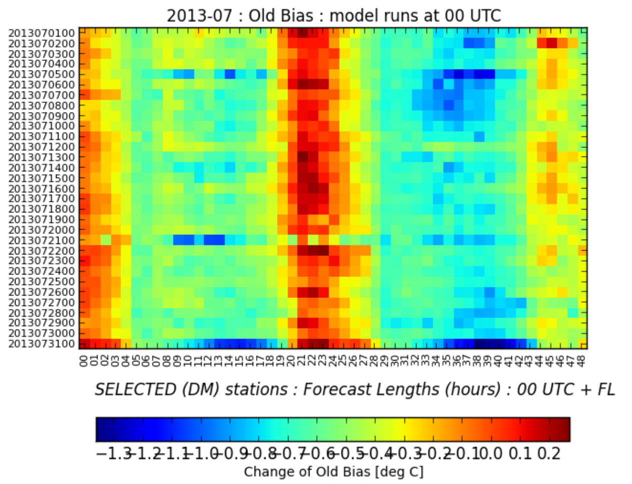












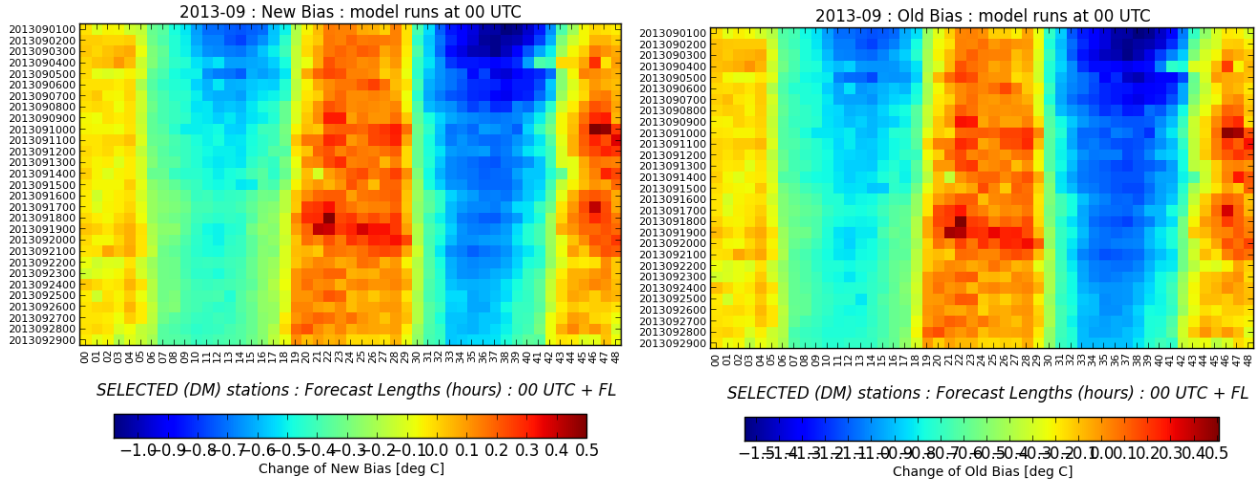
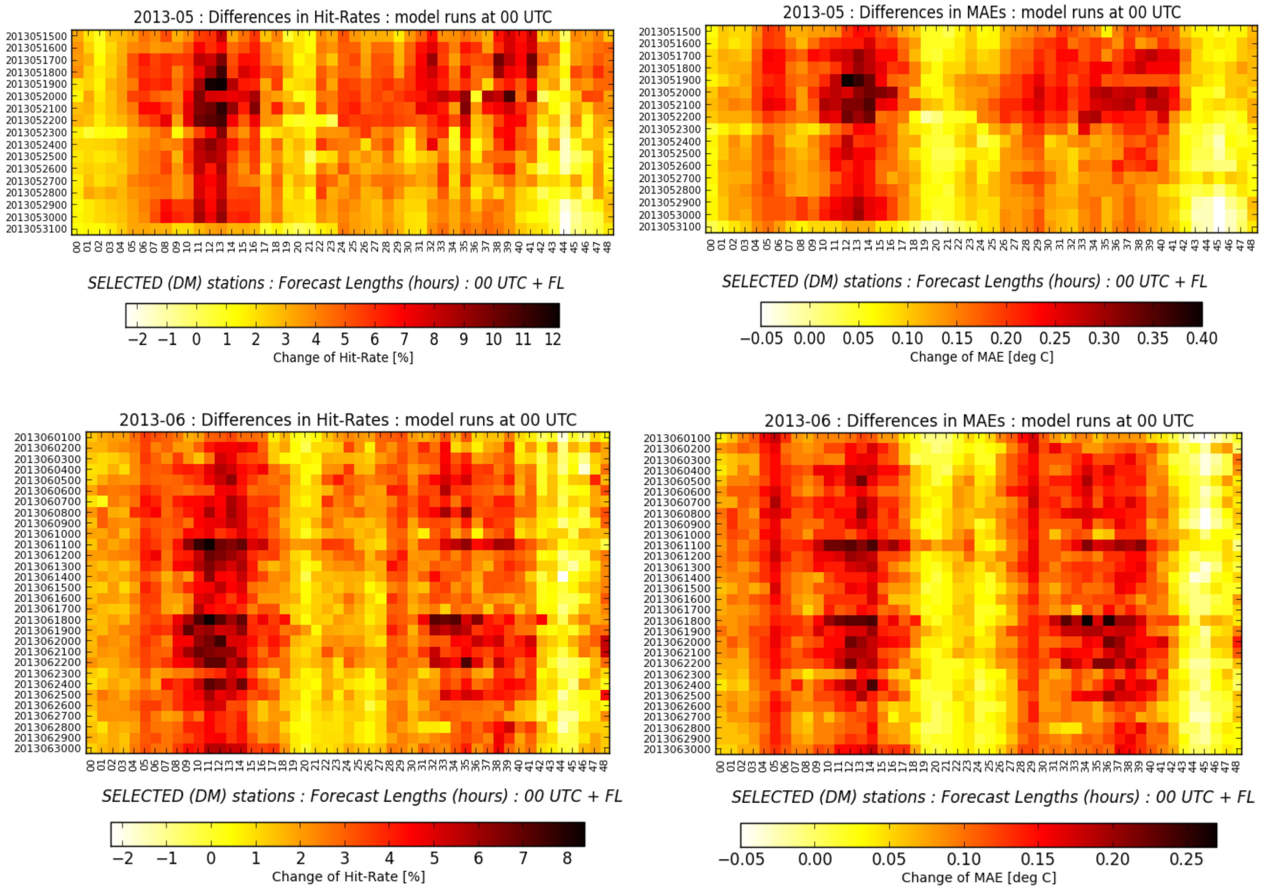


Figure D1*: Temporal variability of the bias of the air temperature at 2 m - (left) after and (right) before the correction applied (results are averaged over selected synoptical stations following the decision-making evaluation, and based on monthly (1st row - May, 2nd - Jun, 3rd - Jul, and 4th - Sep 2013) HIRLAM-SKA model runs at 00 UTCs for forecast lengths 00-48 hours).*



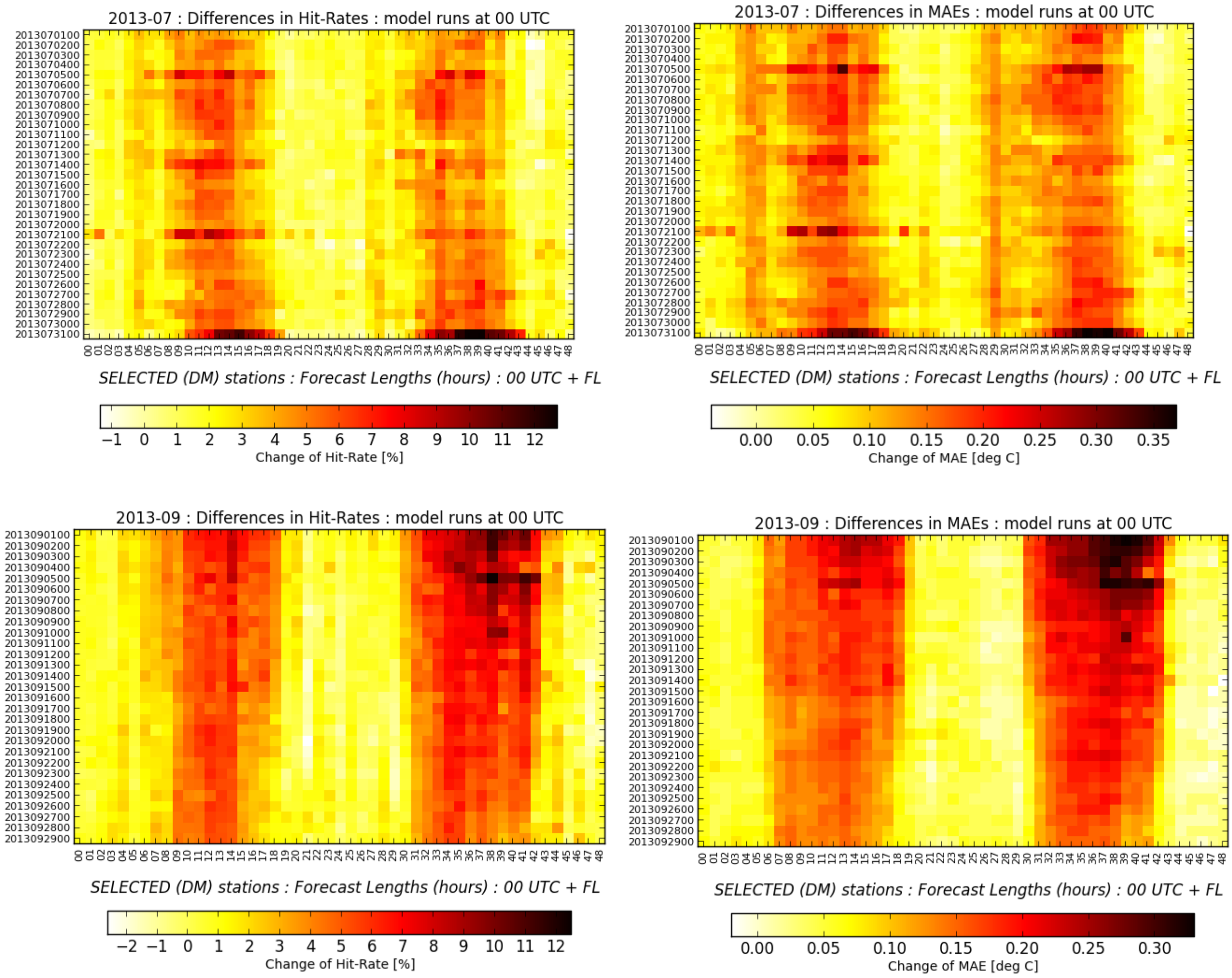
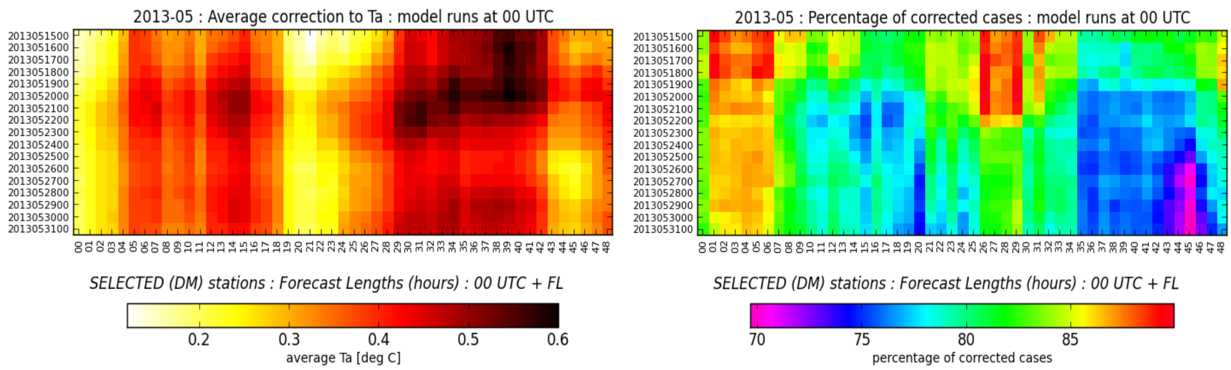


***Figure D2**: Temporal variability of the air temperature at 2 m - (left) differences in hit-rate and (right) differences in mae - after the correction applied (results are averaged over selected synoptical stations following the decision-making evaluation, and based on monthly (1st row - May, 2nd - Jun, 3rd - Jul, and 4th - Sep 2013) HIRLAM-SKA model runs at 00 UTCs for forecast lengths 00-48 hours).*



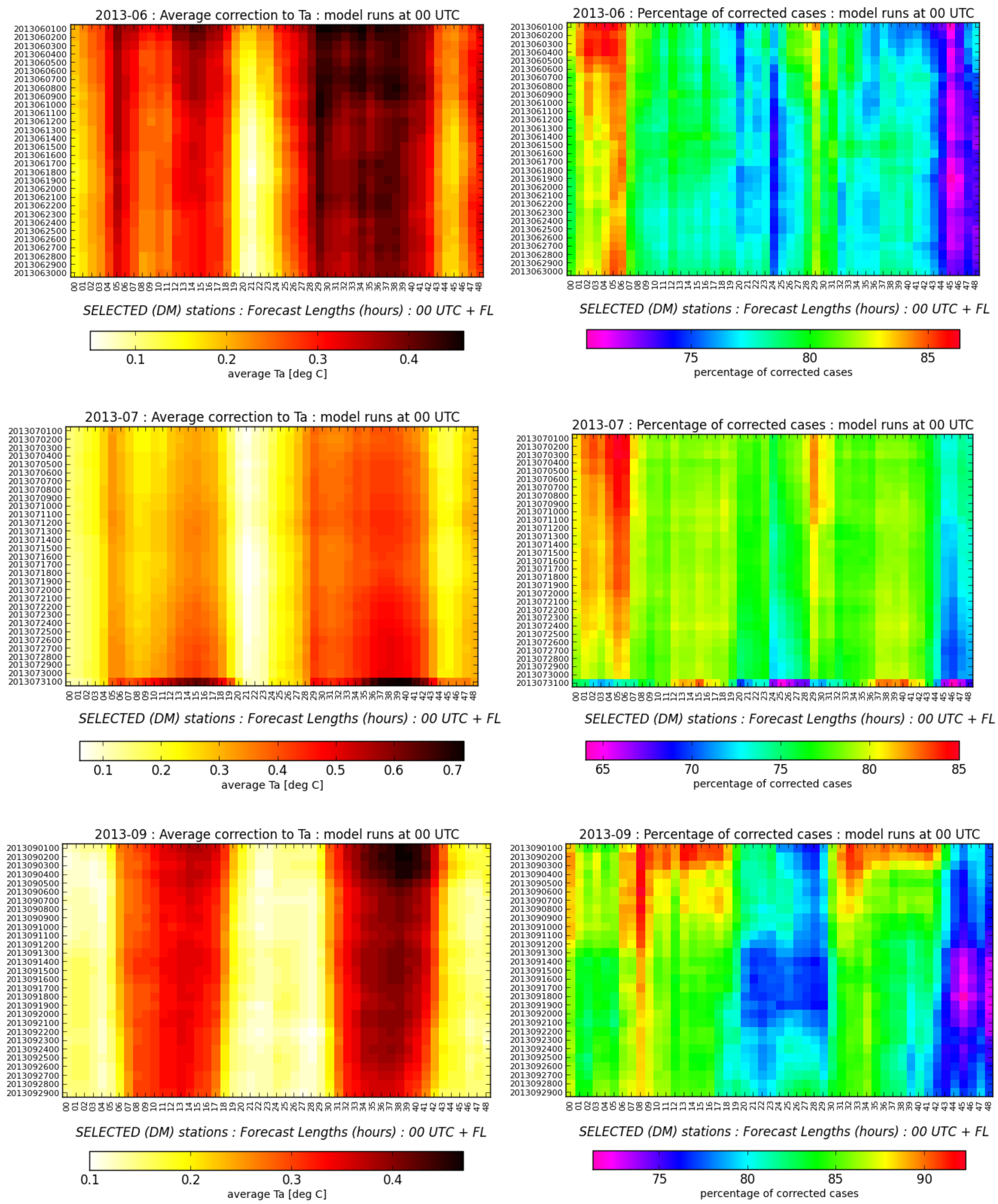


Figure D3: *Temporal variability of the air temperature at 2 m - (left) average correction and (right) percentage of corrected case after the correction applied (results are averaged over selected synoptical stations following the decision-making evaluation, and based on monthly (1st row - May, 2nd - Jun, 3rd - Jul, and 4th - Sep 2013) HIRLAM-SKA model runs at 00 UTCs for forecast lengths 00-48 hours).*

Previous reports

Previous reports from the Danish Meteorological Institute can be found on:
http://www.dmi.dk/dmi/dmi-publikationer.htm