Scientific report; WR-2002-05

KNMI

On the use of physical and statistical downscaling techniques for NWP model output

Wim de Rooy and Kees Kok

Koninklijk Nederlands Meteorologisch Instituut

De Bilt, 2002

Scientific report = Wetenschappelijk Rapport; WR 2002-05

De Bilt, 2002

PO Box 201, 3730 AE De Bilt
The Netherlands
Wilhelminalaan 10
http://www.knmi.nl
Telephone +31 30 22 06 911
Telefax +31 30 22 10 407

Authors: Wim de Rooy and Kees Kok

UDC: 551.509.1
551.509.313
551.509.314

ISSN: 0169-1651

ISBN: 90-369-2216-x

On the use of physical and statistical downscaling techniques for NWP model output

Wim de Rooy
and
Kees Kok

Contents

1 Introduction

Wind speed at 10m (u_{10m}) and temperature at 1.5m (T_{2m}) are very important output parameters of a Numerical Weather Prediction (NWP) model. However, in general one can say that near-surface verification results of NWP models, where model output is compared with synoptic observations, have shown only modest improvement during the past couple of years. A possible explanation for this is that temperature and wind near the surface are strongly influenced by local conditions, like roughness length, albedo, soil properties, etc.. In practice, a Numerical Weather Prediction model with a limited horizontal resolution will never be able to take all the important local conditions into account, simply because it works with grid box mean conditions. Increasing the resolution or using a tile land surface scheme (e.g. Koster and Suarez, 1992), in which different surface covers can be defined within one grid box, can reduce the mismatch between grid box mean and local conditions only to a limited extent. For example, the effective roughness length for momentum (Wieringa, 1976), which has a large impact on u_{10m}, may change significantly for a horizontal displacement of only a few metres. Moreover, the effective roughness length can be strongly wind direction dependent. If the error in the predicted temperature and wind speed is indeed dominated by such a representation problem (from hereon RP), then it becomes clear why little progress is made in predicting these quantities. Consistently, the prediction of a parameter much less determined by local conditions, like pressure at mean sea level, does show improvement during the last years. Another indication that the RP plays an important role is the relatively small increase of the T_{2m} and u_{10m} errors during the forecast. Note that the impact of the RP stays constant during the forecast.

A variety of subjects is described in this paper and they involve T_{2m} and u_{10m} (for specific humidity at 1.5m (q_{2m}), results are mentioned only briefly). Besides validation of NWP output, we also investigate statistical and different physical methods for downscaling. However, the most important approach that runs throughout this paper can be seen as follows. First, a distinction is made between model errors and representation problems and hypotheses about the characteristics of these errors are postulated. Subsequently, for wind speed the hypotheses are confirmed by validation of NWP model output against observations. For the small-scale RP a physical method is introduced, whereas the large-scale NWP errors are handled with a statistical technique. Finally, a combined physical and statistical approach to generate local wind speed fields is validated. The derived regression in this combined method is also effective on locations without observations.

The NWP model used in this study is the Hirlam (Källén, 1996). Hirlam is a short-range weather forecasting system developed and operationally used in several European countries. Although the Hirlam version used in this paper is replaced by a newer version a few years ago, arguments are given why most of the results and recommendations are probably comparable for state of the art models. Because the here used model has already been replaced, no attempts have been made to optimise the methods for operational forecasting purposes. All physical methods described in this paper can be applied, without changes, to other NWP models.

The model validation refers to forecasts up to +48h and started every six hours. Results as a function of the full forecast length are investigated but the hour of the day turned out to be more important than the forecast length and in the major part of the paper we restrict ourselves to the 18+6h (0 UTC) and 6+6 (12 UTC) forecasts to investigate respectively stable and unstable conditions.

To structure the analysis of model results, it is important to make clear definitions of the investigated errors (section 2.1). In this section, also some hypotheses about different error characteristics are stated. The principles of all physical methods investigated in this paper are explained in section 2.2. However, only one method is described in detail here. The downward transformation, belonging to this latter physical approach, is verified against measuring tower observations in section 4. The other physical approaches, only applicable for wind speed, are partly based on the results of section 5 and will therefore be introduced only until section 6.2. Statistical methods used in this paper are briefly discussed in section 2.3. Further, a description of the dataset, the model, the observations, and the verification set-up is given in section 3. The analysis of the model results, including verification for the physical method described in detail in section 2.2 and a simple statistical technique, can be found in section 5. A combined physical/statistical downscaling method for wind speed can be found in section 6.1. Different physically based u_{10m} estimators, meant for this combined method, are introduced in section 6.2 and verified in section 6.3. How the physical/statistical approaches perform for locations where no observations are available is described in section 6.4. Finally, the conclusions and some considerations can be found in section 7.

2 Methods

2.1 Decomposition of the total error

Here we present definitions of the model error and the representation problem used in the rest of the paper. Also hypotheses are presented about the characteristics of the different errors. First the total error at time "t" is defined as:

$$Total\_error(t) = NWP\_output(t) - observation(t) \quad (2.1)$$

Where, in our case, the NWP output is the output of the Hirlam model. The observation is the measurement at a synop station, or, more generally speaking, the actual value of a certain parameter at a certain location. Observation errors are ignored.

The *Total_error* can be split in two different components, namely:

$$Total\_error(t) \equiv NWP\_error(t) + Representation\_Problem(t) \quad (2.2)$$

where

$$NWP\_error(t) \equiv NWP\_output(t) - \text{“true”}\_grid\_box\_mean(t) \quad (2.3)$$

and

$$Representation\_Problem(t) \equiv \text{“true”}\_grid\_box\_mean(t) - observation(t) \quad (2.4)$$

The *Representation_Problem* represents the local deviation from the true grid box mean.

The above-mentioned breakdown of the *Total_error* resembles the approach of Tustison et al. (2001). They discuss the interpolation of model precipitation to observation sites and study the interpolation error (or "representativeness error") as a function of model resolution.

In verification studies and by operational forecasters the *Representation_Problem* (RP) is often ignored and the model is blamed for the total error. By tuning the roughness length of a grid box in order to get the right 10m-wind speed for a specific station, a *NWP_error* may unintentionally be introduced. Obviously, this is not correct because the model should produce a grid box mean and not a site-specific output. To emphasise this, the term Representation “problem” is used and not “error”.

An example leading to a RP is the difference between the mean roughness of a grid box and the site-specific roughness. A typical *NWP_error* can arise e.g. from a forecast error or from a deficiency in the turbulence scheme, e.g. too much mixing under stable conditions and not enough mixing during unstable situations. As hypothesis we state that normally *NWP_errors* do not fluctuate very much on small scales. Nevertheless, *NWP_errors* are sometimes linked with the RP. This is clearly the fact with land-sea transitions, leading to possibly large contrasts in stability and thus *NWP_errors* (think of the deficiency in the turbulence scheme as an example of a *NWP_error*). In contrast with the *NWP_error*, the RP will normally be variable on a very small scale, depending mostly on the heterogeneity of the landscape. There are phenomena that contribute (in a random way) to the RP, which cannot be associated with the characteristics of the observation site and its surroundings, namely the influence of cloud cover and rain showers. The grid box mean cloud cover can be perfect but

not valid for a specific location. In this paper this problem is not considered. However, in practice one could consider to use a short term high resolution cloud cover model, e.g. analysed with satellite images, in combination with a physical downscaling method like the one described in this paper.

In statistical methods, it is common practice to minimise the *Total_error* by using regressions determined with NWP_output and observations. In this way, the statistical approach deals with the RP as well as the *NWP_error*. Therefore the regression must be recalibrated when the model has undergone a significant change. This is in contrast with physical methods as we present them here, which should reduce just the RP and are consequently in principle model independent. A physical method that compensates certain model errors should be incorporated in the physics of the NWP model.

When do we need statistical techniques and when do more physically based methods become important? For large-scale parameters like pressure, the RP is unimportant and therefore statistical methods suffice. On the other hand, for small-scale parameters and relatively small model errors, a physical method incorporating all the relevant parameters to describe the surface characteristics of the site, might be competitive with statistical methods. In cases where both RP and model error are important, statistical techniques are necessary but physical methods might give surplus value. Part of the goal of this paper is to analyse the *Total_errors* of a NWP model using the above-mentioned definitions and hypotheses, in order to determine suitable methods to post-process NWP model output for a specific location. For instance, the relative importance of the RP and the *NWP_error* are investigated, and how these errors tend to fluctuate in the horizontal.

Finally, what kind of method is appropriate if we want output on locations where no observations are available? For *NWP_errors*, which are supposed to be rather homogeneous, this should not be a too large problem. However, when the RP turns out to be important, more physically based methods become inevitable because between the synop stations no observations are available to calibrate the statistical regressions. In this paper we develop and validate a method for wind speed that is a combination of physical and statistical techniques (to compensate for the RP and the model error respectively). This method can be used for locations without observations or to produce wind speed fields.

2.2 Physical methods

Introduction

In this paper, two types of physical methods to compensate for the RP are investigated. Based on the results of section 5, simple, but physically founded, methods are developed and tested in section 6. These methods do not use estimates of the locally valid atmospheric stability and are only applicable for wind speed. However, in this section we describe a method that does estimate the local atmospheric stability and can be used to estimate the local temperature, wind speed and humidity. Nevertheless, the fundamentals for both types of physical methods are the same.

As we have seen in the previous section, the RP originates for the larger part from differences in land environmental characteristics. Even the above-mentioned exception, namely the influence of cloud cover, acts on the synoptic observations via the surface. Most physical methods (including the ones described in this paper) assume that the influence of local terrain conditions, and therewith the RP, decreases with height, see e.g. McNaughton &

Jarvis (1984) or Hutjes (1996). The idea of most physical methods is to transform model output from a sufficiently high level (high enough to get rid of most of the RP) downwards but this time taking the local conditions into account of the specific location. For a small dataset, De Rooy (1999) showed that such an approach is capable of improving the standard model output on synoptic levels during daytime (unstable) hours.

The influence of nearby obstacles, represented by the local, wind direction dependent roughness length (z_{0mloc}) is felt up to the so-called blending height (Wieringa 1976, 1986). Above the blending height, the wind will be rather uniform in the horizontal. Wieringa also defines a macro wind at the top of the planetary boundary layer. Normally, geostrophic drag laws are used to transform this macro wind down to the blending height (e.g. Verkaik and Smits, 2001, Landberg and Watson, 1994). For this transformation a large-scale roughness must be applied. For example Wieringa (1986) uses a roughness (not wind direction dependent) belonging to an area of about 8km radius for this transformation. Because the horizontal resolution of NWP models is rapidly increasing during the last couple of years (at KNMI, a Hirlam version with a resolution of 11x11km^2 is operational), we believe that current NWP models are already capable of producing reasonable meso winds (i.e. winds at the blending height). Therefore, we think that for the generation of local wind speed, it is a good idea to skip the use of geostrophic drag laws in the upper part of the planetary boundary layer and to concentrate on the downward transformation from the blending height. Note that in this study the NWP model has a resolution of 50x50km^2, which can result in sub optimal meso winds. Nevertheless, most improvement is probably made in the transformation from the blending height to synoptic level.

The downward transformation concept

A well-known concept for downward (or upward) transformations is the use of flux-profile relationships (hereafter FP), which are based on Monin-Obukhov surface layer theory (Monin & Obukhov 1954). For a downward transformation from the 31st (the lowest) model level of Hirlam at height $z_{31}\approx32$m to synoptic height (1.5m for T and q, and 10m for u), the FP can be written as follows:

$$u_{10m} = u_{z_{31}} - \frac{u_*}{k}\cdot\{\ln(\frac{z_{31}}{10}) - \Psi_M(\frac{z_{31}}{L}) + \Psi_M(\frac{10}{L})\} \qquad (2.5)$$

$$T_{2m} = T_{z_{31}} - \frac{\theta_*}{k}\cdot\{\ln(\frac{z_{31}}{1.5}) - \Psi_H(\frac{z_{31}}{L}) + \Psi_H(\frac{1.5}{L})\} + \Gamma_d\cdot(z_{31} - 1.5) \qquad (2.6)$$

$$q_{2m} = q_{z_{31}} - \frac{q_*}{k}\cdot\{\ln(\frac{z_{31}}{1.5}) - \Psi_H(\frac{z_{31}}{L}) + \Psi_H(\frac{1.5}{L})\} \qquad (2.7)$$

with

$$L = \frac{u_*^2 T}{k\cdot\theta_* g} \qquad (2.8)$$

Here, u_{z31} is the wind speed (m/s) at the lowest (31st) model level at height z_{31} (m), u_* is the friction velocity (m/s), k is the Von Kármán constant (=0.4), Ψ_M is the stability correction for momentum (following Beljaars and Holtslag, 1991), L is the Obukhov length (m) defined by Obukhov (1946), θ_* is a temperature scale defined by $\theta_* = -H(\rho\cdot c_p\cdot u_*)^{-1}$, where

H is the surface sensible heat flux (Wm^{-2}), ρ is the air density (kg m^{-3}) and c_p is the specific heat (J kg^{-1} K^{-1}). Further Ψ_H is the stability correction for heat (also taken from the form in Beljaars and Holtslag, 1991), Γ_d = 0.01 (K m^{-1}) is the adiabatic lapse rate, q_* is a humidity scale defined by $q_* = -\lambda E(\rho \cdot \lambda \cdot u_*)^{-1}$, where λE is the surface latent heat flux (Wm^{-2}), λ is the latent heat of vaporisation (J kg^{-1}), and g is the acceleration of gravity (m s^{-2}).

To make use of the above-mentioned flux-profile relationships, estimates must be made of the surface fluxes of sensible heat (H), latent heat (λE), and momentum (u_*). For this, we apply a parameterisation scheme, which will be referred to as the Flux-Routines (FRs).

The use of the Flux-Routines

A detailed description of the FRs can be found in De Rooy & Holtslag (1999) (from hereon RH99). As input the FRs need standard synoptic observations or model output at a single level plus cloud cover. The FRs are composed of parameterisations concerning the radiation components and the surface fluxes. For the latent heat flux, a Penman-Monteith approach (Monteith 1981) with default a simple surface resistance formula is used. The aerodynamic resistance is determined with Monin-Obukhov theory. Without modifications, the FRs are suitable for short grass (WMO standard for synops stations) and moderate climates. However, all the site-specific surface parameters in the FRs are commonly applied and described in the literature for different surface covers and soil types. Therefore adjustments should be relatively easy for use above other surfaces.

In contrast with RH99, here the correction for the long-wave downward radiation due to low clouds is omitted because the skill of the Hirlam forecasts for low clouds is often very poor. For simplicity we take the Hirlam temperature at the lowest atmospheric model level as soil temperature (in combination with a modified heat transfer coefficient, A_G, of 9 $Wm^{-2}K^{-1}$). Another difference between the use of the FRs here and in RH99, is that here an effective (meso) roughness length (Beljaars, 1987), z_{0meff}, is applied during all stability conditions. This is in contrast with De Rooy & Holtslag, who suggest that during stable conditions temperature and wind profiles are adapted to the local surface and therefore a local z_{0m} (i.e. here the roughness length of short grass) should be used. However, they use input at synoptic levels and here input at the lowest model level (about 32m) is used. Besides this, the applicability of a local or effective roughness length also depends on the distance between the observations and the obstacles that influence z_{0meff} (see also section 4). Note that the effective roughness length takes into account the influence of upstream obstacles on the wind profile. Close enough to the surface the wind profile will be adapted to the local surface (in our case, short grass). For most stations involved in this study, the obstacles are closer to the observation sites than in Cabauw (used in RH99), which argues for the use of an effective z_{0m}. Note that during unstable conditions, we always use z_{0meff}. For all transformation procedures a constant roughness length for heat (z_{0H}) of 0.001m is used (RH99).

As mentioned before, in RH99 the FRs have been tested with observations at standard synoptic height as input, but here they are used with Hirlam output at about 32m as input. It is not a priori certain if the FRs also perform well when the input is taken from a higher level in the atmosphere than standard synoptic height. Therefore, we used tower observations at 40m as input for the FRs and subsequently verified the results of the downward transformation to synoptic height, with observations (see section 4). This verification also enabled us to examine limitations and possible necessary adaptations for downward transformations. The results are used to determine the exact transformation procedures during different conditions, which are described later on in this section.

Differences between the physical method and standard model post-processing

There are many differences concerning the determination of e.g. T_{2m} in the Hirlam diagnostic scheme and the downward transformation using the FRs. One of the differences is that Hirlam uses grid box means, while the FRs use the local terrain parameters. For example, Hirlam applies a constant roughness length for momentum (z_{0mHL}), whereas the FRs use a, most of the time smaller, season and wind direction dependent z_{0mloc}, according to Beljaars (1987). In a similar way, Hirlam uses a constant albedo, while the FRs use an albedo for short grass dependent on solar elevation (approx. 0.23, see Beljaars & Bosveld 1997). Experiments in De Rooy (1999) indicate that the improvement with the FRs in comparison with the Hirlam diagnostic output cannot be attributed to one single parameter, e.g. albedo or z_{0m}. Besides the different input parameters, also the method of the FRs is different from the diagnostic scheme in Hirlam. The FRs solve the surface energy balance of a vegetated surface and with the resulting fluxes and flux-profile relationships, the lowest model level output is transformed downwards to synoptic level. On the other hand, the diagnostic scheme of Hirlam uses an exchange coefficient to determine the fluxes and subsequently these fluxes are used, together with flux-profile relationships, to transform the surface value upwards.

Pitfalls

In the ideal case the downward transformation should start at the blending height (Wieringa, 1976), where the surface inhomogeneity is supposed to be mixed out. Among other things, the blending height depends on the stability but here we start the transformation at a fixed height, namely the height of the lowest model level (≈32m). However, under most circumstances, meteorological parameters at 32m are still influenced by local conditions, though less than near surface parameters. Therefore, transforming the lowest model level parameters down to synoptic level can only give a reduction of the RP, not a cancellation. In this paper, no attempt is made to optimise this physical method, although recommendations for further improvement are given. Beforehand it was clear that the impact of the proposed downward transformation method with the FRs would be too small to compensate for the full RP. However, in this way we can be confident that the physical method will not lead to overcompensation, which can confuse the investigation for the surplus value of the physical method.

As will be shown in section 4, under neutral and unstable conditions the errors in the resulting T_{2m}, q_{2m}, and u_{10m} after the downward transformation from the lowest model level are negligible in comparison with the typical model error. Unfortunately, this is not the case for very stable conditions because Monin-Obukhov theory (hereafter MO) is not valid anymore (the FRs and the flux-profile relationships rely heavily on MO-theory). With alternative downward transformation procedures (described beneath) we try to reduce the errors during very stable conditions. Apart from the above-mentioned problem with the downward transformation during stable conditions, there is also a more fundamental complication. Mc Naughton & Jarvis (1984) conceptually and Hutjes (1996) experimentally give two reasons why transformation methods can fail during stable conditions. Firstly, in stable conditions the surface layer itself is too shallow to act as a buffer against variations in input from the surface. Secondly, the surface layer is decoupled from the residual layer above (which might have enough buffering capacity). Nevertheless, for research purposes we will also investigate the performance of transformations during nighttime, stable, conditions.

Another problem occurs when due to a certain model deficiency, the error at model level height is larger than at synoptic height. For those situations the transformation procedure can make things worse (de Rooy, 1999). In contrast with De Rooy (1999), here we cannot determine those periods with relatively bad model level output, because the Cabauw tower (van Ulden & Wieringa, 1996), situated in the centre of the Netherlands and with measurements up to 200m high, was not operational during our validation period. Undoubtedly, these periods with large model deficiencies will be apparent in our data and most likely deteriorate the results with the transformation method.

Limitations and adaptations on the use of the flux-profile relationships

An important issue is the validity of the flux-profile relationships. Under what circumstances are they applicable and when are adaptations necessary or are downward transformations even useless? It is important to realise that temperature and humidity on the one hand, and wind speed on the other, have totally different error characteristics. For example, during very stable conditions absolute wind speed errors are small (because wind speeds are low) but temperature errors can be very large. Partly based on the validation of the transformation procedure against tower observations (section 4), it turned out that limitations and adaptations of the use of the FP relationships are only necessary for temperature and humidity.

First of all, temperature transformations during unstable sunrise hours have to be excluded. Sunrise hours are defined as $L<0$, hour of the day<14, and $\sin\varphi<0.4$, where φ is the solar elevation. During sunrise hours a very shallow, unstable profile beneath a stable temperature profile might exist. Therefore, the current transformation procedure is not suitable for T during these conditions. Incidentally, sunrise hours can also lead to relatively large wind speed errors (section 4). However, excluding these hours has less impact on wind speed results. For this and for simplicity reasons, no restrictions are applied on wind speed transformations. Yet, for specific humidity downward transformation is omitted during sunset hours because these hours can give relatively large humidity errors (section 4). Sunset hours are defined as: $L>0$, kin>1, and hour>11, where kin is the incoming short wave radiation.

Also excluded are temperature and humidity transformations during very stable ($0<L<10$) conditions. During these conditions MO-theory is no longer valid and the downward transformation might result in much too low temperatures. This can be explained as follows. During very to moderately stable conditions the stratification normally decreases with height (e.g. André & Mahrt, 1982) and (intermittent) turbulence is restricted to a shallow layer near the surface. Above this layer, the temperature profile is determined primarily by radiation flux divergence leading to much smaller vertical gradients. The Obukhov length, which is formally only valid in the lower 10% of the turbulent layer, corresponds with the strong stratification near the surface and will therefore exaggerate the temperature gradient for higher levels. Transforming downwards, this will result in underestimations at synoptic heights. Also for more moderate stability regimes, with $10<L<z_{31}$, the above mentioned problem will result in underestimations at synoptic levels. To limit this error, a neutral profile is assumed between the lowest model level and $z=L$ for hours with $10<L<z_{31}$. Subsequently, the normal flux-profile relationships are used down to the output level (the complete transformation is illustrated in Figure 2.1). A neutral profile means a constant value for the specific humidity and a dry adiabatic profile for T. This downward transformation procedure, with the transformation from Figure 2.1 for $10<L<z_{31}$ and the exclusion when $0<L<10$, will be referred to as Method A.

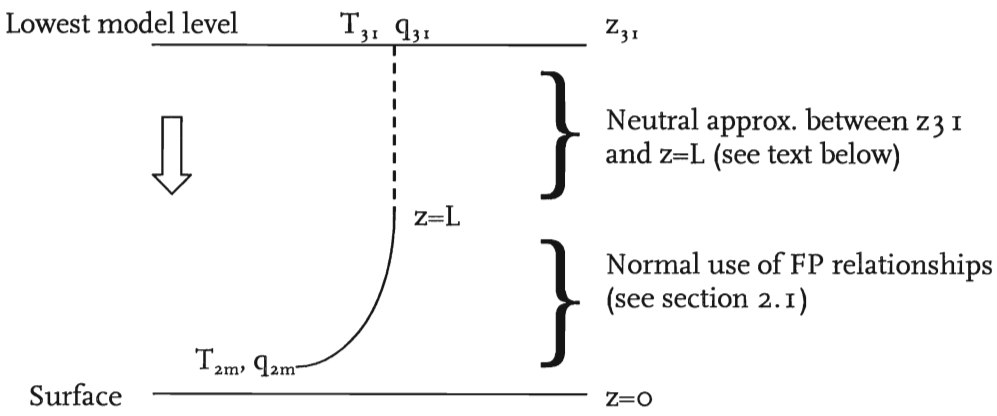

Figure 2.1 Downward transformation procedure for temperature and humidity during stable conditions with 10<L<z_{31} (together with the exclusion of 0<L<10, called Method A)

Alternatives

Increasingly popular is the use of tile schemes in NWP models. The tile approach is not only a way to describe the influence of heterogeneity of the surface on the atmosphere, but it also gives opportunities to generate more locally valid output, e.g. by taking the T_{2m} of the "grass tile" for verification against a synop station. Recent results of such an approach (Best et. al., 2000) show generally only modest improvement in verification results. In section 5 a possible explanation for the modest improvement for wind speed is given. On the other hand, an advantage of the use of tile schemes in comparison with the transformation method is that a tile scheme works prognostically, while the transformation method is purely diagnostic. To illustrate this one might think of a case where, after a long period with rain, suddenly dry air is advected. With a tile approach the soil moisture, which can have a large impact on the energy balance, is determined every time step for every tile. The FRs, however, approximate the soil moisture on the basis of the humidity deficit of the air and will consequently underestimate the soil moisture content. Yet, the tile approach as it is applied in NWP models is not dedicated to a certain location. For example the roughness length is not wind direction dependent and also the soil type might not be the most appropriate choice for a specific site. This problem can be solved with the transformation method proposed in this paper or by defining tiles with zero areas (so they do not influence the atmosphere) but with site-specific optimised parameters. We will return to this subject in the discussion section.

2.3 Statistical methods

A very common way to obtain local predictions for weather parameters is by using model output statistics (MOS) techniques (Glahn and Lowry, 1972), in which predictors obtained or constructed from NWP model output are correlated with local observations. Usually a large data set is collected on which statistical routines (e.g. multiple linear regression) are applied in order to obtain the (optimal) combination of predictors that can explain the largest part of the variance of the observations. With MOS this is done separately for every forecast time. It can only be applied for places where observations are available. Usually the very local effects are incorporated into the predictors. For instance, the effect of any directional inhomogeneities, e.g. a nearby building, city or forest, can be taken into account by predictors that are dependent on wind direction. The resulting regression equations are then applied to future cases at the same location assuming that recent and future model as well as local changes will have little effect on the forecast results. In this

paper, however, it is not the purpose to build the best possible forecasting system that will be used as post-processing tool to improve the model output. One of the reasons for this is that we are dealing here with an older version of the Hirlam model which is not operational anymore. Moreover, in our approach we try to establish the feasibility of the separation between NWP_error and RP without pursuing to make an optimal operational forecasting system.

In the statistical analysis in this paper use is made of multiple linear regression only. It is used in two different kinds of experiments. The first is the evaluation and intercomparison of a number of individual meteorological quantities. For instance, the model 10 meter wind speed and the wind speeds from the physical method using the flux routines (as described in the previous section) are verified against observations. However, this is also done after they are both regressed using the observations. Of course this verification after applying the regression is done on independent data. In this way not only the quantities themselves are compared but also their, what one might call, *predictive potential*. By doing this we can reduce the more or less systematic deficiencies of the quantity with respect to the observations. This is, we believe, a much fairer way of comparing quantities than only comparing their values without using regression. In this way we can compare the information that is available in the subsequent quantities and which can be obtained by very simple means. This comparison in terms of predictive potential is done also for winds coming from a number of additional physical methods (described later) and also for temperature and relative humidity. The evaluation throughout this paper is in terms of bias, standard deviation (SD), root mean square error (RMSE) and mean absolute error (MAE).

The statistical techniques are also used in a completely different experiment which only involves wind speed. In this so-called pool experiment we first construct a new estimator for the local wind speed, which is a direct translation from the (grid box average) model winds using site-specific roughness lengths. This physically based wind estimator can be regarded to deal with the part of the RP which is related to z_{0m}. This estimator is calculated for a number of stations and then regressed on the observations in order to obtain a correction which is common for this pool of stations. This correction, which is tested on independent data, can be regarded as being related to the larger scale NWP_error. Again a number of local wind estimators will be proposed and tested. This will be described in section 6.

As mentioned before, the regressions determined on a training set are tested on an independent set. The dependent set was taken, rather arbitrarily, as approximately the first half of the total data set, ranging from December 1 1996 until April 1 1998; the independent set comprises the remaining data ranging from the 2nd of April 1998 until the the 1st of Oct. 1999. In this way in both sets all seasons are included and the sets are large enough to determine reasonably stable regression equations and to infer statistically significant conclusions. A disadvantage of the above choice is the possible disruption if there are important model or local changes in the total period. In that case the regression could have been developed on a set that has different statistical properties compared to (part of) the set on which it is tested. However, the results of all experiments we performed did not suggest that this was a serious problem.

There was one important misfortune, however, and that was the almost absence of +30, +36 and +42 forecasts in the first half of the investigated period. Therefore there are not enough data available to derive regressions and the statistical results that will be shown in this paper do not include these forecast periods. For the +48 forecasts enough data were available.

3 Data and verification set-up

Introduction

In this paper, we use a three-year dataset for the experiments and validation. The dataset consists of NWP output and observations. In this section we start with the description of the NWP model itself (called Hirlam) and its output. Subsequently, a brief discussion on the verification set-up is given. Finally, the synop stations and their surroundings are described in detail, together with the model parameters of the nearest grid point because this information determines to a large extent the RP.

3.1 The NWP model and the dataset

Runs are made with the Hirlam 2.7.3 reference system, which was the operational system at the Royal Netherlands Meteorological Institute (KNMI) until September 1999. In the model are included: Sundqvist condensation (Sundqvist, 1978), Kuo convection scheme (Kuo, 1974), the Hirlam-2 surface parameterisation (Källén, 1996), and a Louis-type vertical diffusion scheme (Louis, 1979). The model has a horizontal resolution of 55km and has 31 layers in the vertical. The height of the lowest model level (level 31) varies, with a few metres, around 32m depending on temperature and surface pressure. The height of the second lowest model level (level 30) is about 140m. Every 6 hours a forecast up to +48h is started. Output of the +6, +12, +18, +24, +30, +36, +42, and +48h forecast is available. However, for a significant part of the validation set the +30, +36, and the +42h forecasts are missing. In sections 5.1 and 5.2, results as a function of forecast length are investigated. The hour of the day turned out to have more influence on the verification results than the forecast length and in the rest of the paper we restrict ourselves to the 18+6h and 6+6 forecasts to investigate the results for respectively stable and unstable conditions.

The verification period starts at December 1996 and ends at September 1999. Only for the validation of the downward transformation procedure itself, a dataset of the 200m Cabauw tower is used. A general description of the Cabauw observation site and the data is given by van Ulden & Wieringa (1996). The Cabauw data used here refers to the year 1996.

3.2 Experimental set-up

The downward transformation itself, using tower observations in Cabauw as input, is verified in section 4. For the downward transformation, the FRs are used and therefore estimates of the Obukhov length, a measure of the atmospheric stability, are made (2.4). This estimated Obukhov length is not available in section 5. Hence, for section 5 we have to rely on a rather crude indicator of stability, the hour of the day. When results for the summer season are discussed, they refer to the months from April until October (and for winter the remaining months).

In section 5 wind speed, temperature, and humidity measured at 11 different synop stations in the Netherlands are compared with standard diagnostic Hirlam output and output which is the result of the downward transformation from the lowest model level or a simple statistical regression. For this verification, the grid point nearest to the observation is taken (no horizontal interpolation). This is also the case for section 6, where a combined

statistical/physical method to optimise NWP output for a specific location, also at locations without observations, is presented and tested.

3.3 Description of the synop stations

Figure 3.1 Positions of the land/sea stations (represented by squares and italic text) and land stations (represented by circles and normal text). Definitions of land- and land/sea locations can be found in the text. In addition the position of Cabauw is plotted (data of Cabauw is only used in section 4).

Characteristics of the observation sites are very important because they determine to a large extent the representation problem for the NWP model. Besides the characteristics of the stations, also the model parameters of the nearest grid point (used for the verification and as input for the downward transformation) are important and are therefore described in this section. The different locations can be divided in the ones where the nearest grid point as well as the observation site has purely land characteristics (from hereon: land locations), and on the other hand locations where in some way land/sea representation problems play a role or might be suspected (referred to as land/sea locations). Without special remarks, all measurements are made above short grass, which has a mean albedo of 0.23 (see section 2.2). In this paper, we use local roughness lengths according to Beljaars (1987). Except for Beek, Twente and Volkel, the applied local roughness lengths are season-dependent (summer/winter). Further the wind direction dependence of the z_{omloc} has a resolution of 20° for all stations except Cabauw (30° resolution). Generally speaking, for land locations the local roughness lengths are smaller than the model z_{0mHL}. Unless mentioned otherwise, all

wind measurements are made at 10m. Temperature and humidity observations at all stations are made at 1.5m above the surface.

3.3.1 Land locations

Table 3.1. Land locations. Some characteristics of the observation sites and the nearest model grid points. See also text below.

Station	z_{0m} in model [m]	height location [m]	height in the model [m]	albedo in the model [-]
Beek	0.6022	114	99.5	0.195
Deelen	0.4956	50	12	0.188
Herwijnen	0.4956	1	12	0.188
Eelde	0.3329	4	12	0.187
Volkel	0.4010	20	14	0.195
Twente	0.3754	36	46	0.203

Beek

For this location, the differences between the model and the actual (Beljaars, 1987) roughness lengths are the largest. For the most common wind directions z_{0mloc} is about 0.15m.

Deelen and Herwijnen

Both locations have the same nearest grid point and therefore the same *NWP_error*. For the most common wind directions, Deelen has somewhat larger roughness lengths (about 25cm) than Herwijnen (about 10 cm). The two stations also differ in soil type. While Herwijnen has clay with normally quite wet soil conditions, station Deelen is situated on sand. Note the relatively large height difference (a RP) between the actual and the model height for Deelen.

Eelde and Volkel

The roughness length in Volkel fluctuates strongly, with relatively high z_{0mloc} in the north-west to north-east directions (about 0.45m) but rather small values (around 0.10m) for the remaining directions. Eelde on the other hand, shows less roughness variance (z_{0mloc} ranges from 0.1 to 0.3).

Twente

In comparison with the other land locations, Twente shows only weak variance of the roughness length with wind directions. Moreover, the difference between the local roughness length and the model z_{0m} is relatively small. For the most common wind directions there seems to be a small overestimation of the z_{0mloc} by the model. However, comparison of the here used (Beljaars, 1987) z_{0mloc} for Twente with roughness lengths determined according to Wieringa (1976), using wind data of the period 1995-1996, reveals that the Wieringa roughness lengths are significantly higher. This is in contrast with the other stations where these roughness lengths are comparable, with sometimes even a somewhat lower z_{0mloc} according to the Wieringa method.

3.3.2 Land/sea locations

Table 3.2 Land/sea locations. Some characteristics of the observation sites and the nearest model grid points. See also text below.

Station	Sea point in model [m]	z_{0m} in the model	observation height ff [m]
Valkenburg	yes	Charnock	10
De Kooy	no	0.123	10
Vlissingen	no	0.173	28 or 20
Stavoren	yes	Charnock	12
Hoek van Holland	yes	Charnock	15

For all the land/sea points the differences between model height and observation height are negligible. For sea model points, z_{0mHL} is determined with the Charnock relation (1955), which makes the roughness length wind speed dependent (for sea points, z_{0mHL} varies approximately between $3 \cdot 10^{-5}$ and $2 \cdot 10^{-3}$m).

Valkenburg

This location behaves primarily as a land point since the observations are made a few kilometers from the coast inland.

De Kooy

The local situation in De Kooy is rather complex with nearby (a couple of hundred meters away) water in the northeast to southeast direction. Occasionally the water dries up due to low tides. From northwest to northeast the buildings of the nearby city Den Helder cause relatively high roughness lengths. A priori, it is unclear if De Kooy can be considered as a pure land point for temperature and wind.

Vlissingen

The situation for Vlissingen is very complex. The 20m high measuring mast is located on a 8m high dike. Depending on the wind direction, the effective measuring height can change therefore from 20 to 28 m. The measurements are made at a small area (about 15x15m) with short grass. In the northern directions there is a harbor. Occasionally, ships can contribute to the roughness length in these directions. For winds from the southerly directions, Vlissingen can be considered as a sea point because of the long fetch over seawater.

Stavoren

In Stavoren the temperature is observed at another location than the wind speed. Although the temperature is measured only 50m from the coast, the sensor is screened from the wind. Hence, the air may have the opportunity to adapt to land conditions.

The wind mast is placed at a 2m high dike. For easterly directions the roughness length is relatively high (maximum z_{0mloc} is about 0.17m). While the observation site is almost surrounded by water, only in the westerly directions the fetch over the IJsselmeer is long enough to give the wind profile the stability belonging to a sea point. Note that the IJsselmeer is relatively shallow. Therefore its water temperature will be higher in summer and lower in winter than (deep) seawater. Nevertheless, for the common westerly wind directions Stavoren can probably be considered to be a sea point for wind.

Hoek van Holland

Also in Hoek van Holland the temperature is observed at another location than the wind speed. This time the temperature is measured above dune sand, about 700m from the coast. Therefore this station can probably be regarded as primarily a land station for temperature.

The 10m-wind mast of Hoek van Holland is situated on a 5m high dam in sea. Just like Stavoren, the vicinity for the wind measurements of Hoek van Holland is water for almost all directions. Here z_{0mloc} does not exceed a few centimeters. However, from 60 up to 270° the fetch over water is short (about hundred meters) and it is unlikely that the wind profile can adapt to sea stability.

4 Validation of the transformation procedure with tower observations

In this section we investigate the error due to the transformation method exclusively. For that, we use a one-year (1996) tower dataset of Cabauw (see Figure 3.1). As mentioned in section 2.2 the FRs are only tested with input at synoptic heights. However, in section 5, model level output at about 32m will be used as input for the FRs. In Cabauw no observations are available at 32m, therefore we use observations at 40m in this section. Using these Cabauw tower data as input, wind speed, temperature and specific humidity at 40m are transformed downwards to synoptic height and at this height compared with observations. In this way, we can get an idea of the error due to the transformation procedure if the input for the FRs is taken at a height comparable with the lowest model level.

Unfortunately cloud cover is not observed in Cabauw. Therefore total cloud cover in Cabauw is estimated by averaging the hourly observations reported at three surrounding synop stations. Because of this necessary approximation of the cloud cover, the actual error of the transformation method will probably be smaller than reported in this section. Another disadvantage of Cabauw data is that the wind direction dependence of the roughness length is divided in 12 sectors of 30°, whereas all the other synop stations work with 18 sectors. Moreover, most of the involved synop stations make a distinction between summer and winter for z_{0mloc}, whereas the Cabauw roughness length has no seasononal dependence.

To obtain conditions comparable to those in section 5, the short wave radiation (Kin) is not prescribed but parameterised using total cloud cover and solar elevation. For the same reason, the amount of medium and high-level clouds are not used as input for the FRs.

Different alternative transformation methods are investigated. As mentioned before, in general, transformation during unstable conditions is relatively simple and can be done in a straightforward way. On the other hand, during (very) stable conditions the errors can be large. Therefore the testing of transformation procedures is split up here in a stable and unstable part.

4.1 Downward transformation during unstable conditions

During unstable conditions ($L<0$), we examine two slightly different procedures. Namely, a straightforward approach with an effective z_{0meff} and no exclusions (1544 cases), and a version where sunrise hours are excluded (1267 cases). Sunrise hours are defined as $L<0$, hour of the day<14, and $\sin\varphi<0.4$, where φ is the solar elevation. The results do not fluctuate very much from month to month, so the results are only presented for the complete year. As a poor man's approach, we also present the results if the temperature at 40m is adopted as 2m temperature after correction according to the dry adiabatic lapse rate ($T_{2m}=T_{40m}+0.38°C$).

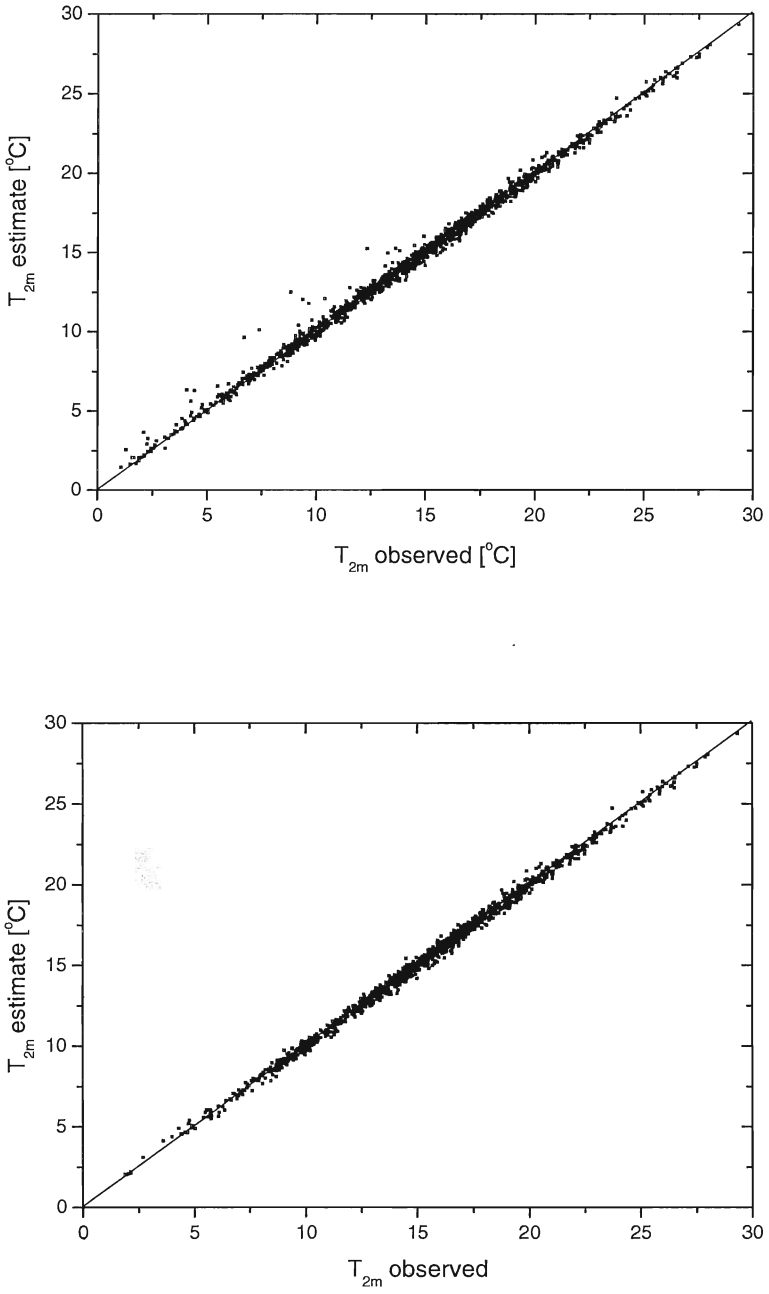

Figure 4.1 Comparison of the observed T_{2m} with estimates using the downward transformation from 40m height. Both graphs contain only unstable conditions (L<0) but in the bottom picture sun rise hours (definition in the text) are excluded.

Table 4.1 Downward transformation results during unstable conditions with no exclusions. Number of cases is 1544. SD stands for Standard Deviation

1996 Unstable	RMSE	Bias	SD
u_{10m} [m/s]	0.367	-0.137	0.340
T_{2m} [°C]	0.362	0.103	0.347
q_{2m} [gr/kg]	0.333	0.289	0.167
$T_{2m}=T_{40m}+0.38$ [°C]	0.586	-0.371	0.453

Table 4.2 Downward transformation results during unstable conditions with the exclusion of sunrise hours (definition in text). Number of cases is 1267

1996 Unstable hours	RMSE	Bias	SD
u_{10m} [m/s]	0.326	-0.208	0.251
T_{2m} [°C]	0.266	0.043	0.263
q_{2m} [gr/kg]	0.356	0.322	0.150
$T_{2m}=T_{40m}+0.38$ [°C]	0.594	-0.491	0.334

As can be seen in Tables 4.1 and 4.2 and scatter plots of Figure 4.1, the errors due to the downward transformation are small during unstable conditions. This can partly be explained by the simple fact that vertical gradients are small during unstable conditions (except for a shallow layer near the surface). This is also reflected in the reasonable performance of $T_{2m}=T_{40m}+0.38$. Apart from this effect, the transformation during unstable conditions will normally take place in the surface layer, where MO-theory should be applicable. The scatter plots for wind speed (not shown) are qualitatively the same as the temperature plots (Fig. 4.1a and b). Excluding sunrise hours filters out the largest errors, also for wind speed. However, for wind speed a negative bias is apparent.

If we look at the wind speed error as a function of the wind direction, we see a clear dependence, not only for the *Total_error* but also for the error after the downward transformation using the wind direction dependent z_{0meff}. Experiments, done by Verkaik (personal communication) with Cabauw winds for another period, showed exactly the same behaviour for the wind profile between 40 and 10m. Verkaik found that for most directions the wind profile between 40 and 200m could be described rather well with the z_{0meff}, whereas the profile between 40 and 10m "feels" only the much smoother conditions closer to the tower. His hypothesis is that the roughness lengths determined with gustiness factors (e.g. the effective z_{0m} used in this paper) have a large footprint. It is not clear if these kinds of considerations are also valid for other synop stations. Therefore, it is also uncertain if this error due to the less suitable z_{0m} should be counted as part of the transformation error or not. At any rate, the other synop stations are in general situated much closer to obstacles. Therefore, we expect that for those stations the effective roughness length is more suitable for wind profiles near the surface than at Cabauw.

All in all, for every parameter the error made by the transformation is close to the typical observation error. Moreover, the error in Kin, which is undoubtedly increased due to the indirect determination of cloud cover in Cabauw, has a clear correlation with the temperature error (see Figure 4.2) and, to a lesser degree, with the wind speed error (not shown).

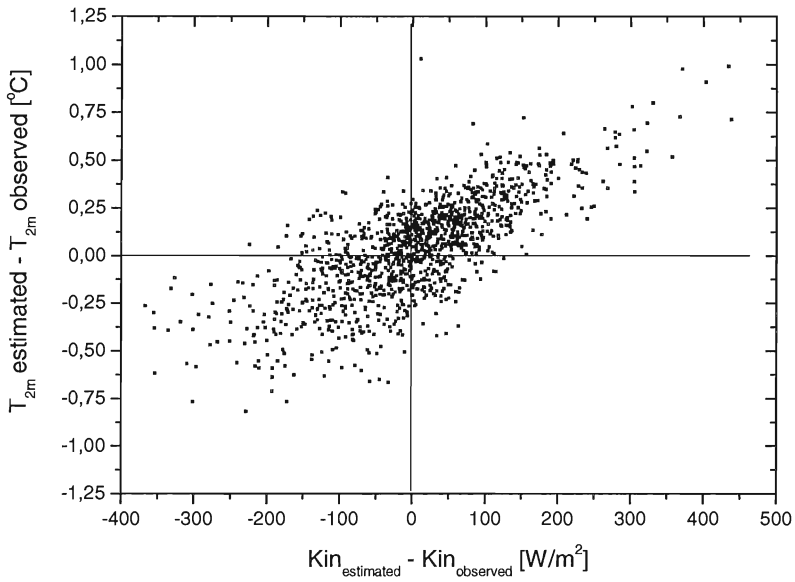

Figure 4.2 Comparison of the error in Kin with the error in T_{2m} using the downward transformation from 40m during unstable conditions with the exclusion of sun rise hours.

If the observed incoming short-wave radiation (Kin) is used as input for the FRs and sunrise hours are excluded, the RMSE for T decreases to 0.178 [°C] and the Standard Deviation (from hereon SD) in u (remember that the bias in the wind speed is primarily determined by z_{0m}) decreases to 0.228 [m/s]. Incidentally, even with Kin as input there still remains a contribution to the errors of the indirect determination of the cloud cover, because in the FRs the cloud cover is also used for the parameterisation of the incoming long wave radiation.

4.2 Downward transformation during stable conditions

During stable conditions the situation is much more complex (see section 2.2). Among other things, we have to decide on the use of a local, short grass ($z_{0mgrass}$=0.01m) or an effective roughness length (see section 2.2 and RH99). Another important point is the problem of transformation during very stable conditions with expected underestimations of T, q and u at synoptic height (see section 2.2). Different alternatives to reduce the transformation errors during very stable conditions are verified here.

Table 4.1 Downward transformation results during stable conditions (L>0). Method A is explained in the text below and section 2.2

1996 Stable	Nr. Cases	RMSE	Bias	SD
T_{2m} z_{0meff}, no exclusions	2989	0.936	-0.263	0.898
T_{2m} $z_{0mgrass}$, no exclusions	2930	1.040	-0.419	0.955
$T_{2m}=T_{40m}+0.38$	2989	1.769	1.261	1.241
T_{2m} z_{0meff}, L>10	2200	0.788	-0.233	0.753
T_{2m} z_{0meff}, Method A	2200	0.764	0.029	0.763
T_{2m} z_{0meff}, u_{40m}>10 m/s	89	0.225	-0.094	0.204
U_{10m} z_{0meff}, no exclusions	2989	0.532	-0.199	0.494
U_{10m} $z_{0mgrass}$, no exclusions	2930	0.515	-0.087	0.507
U_{10m} z_{0meff}, L>10	2200	0.549	-0.285	0.470
U_{10m} z_{0meff}, Method A	2200	0.777	0.094	0.771
U_{10m} z_{0meff}, u_{40m}>10 m/s	89	0.847	-0.721	0.444
q_{2m} z_{0meff}, no exclusions	2989	0.439	0.056	0.436
q_{2m} z_{0meff}, Method A	2200	0.297	0.100	0.279
q_{2m} z_{0meff}, Method A, no sunset	1792	0.207	0.036	0.203

Method A means that for hours with $10<L<z_{31}$, a neutral profile is assumed between the lowest model level and z=L. Subsequently the normal flux-profile relationships are applied from z=L down to synoptic height (see also section 2.2). Method A also excludes hours with 0<L<10. Further sunset hours are defined as L>0, Kin>1, and hour >11.

As expected, the errors for stable conditions are much larger. We present the results for temperature, wind speed and humidity separately.

Temperature

Table 4.1 shows that the SD using the FRs is much smaller than with the $T_{2m}=T_{40m}+0.38$ approach. The difference in performance between both approaches is much larger than for unstable conditions. The exclusion of very stable hours (0<L<10) significantly reduces the temperature errors. It is less clear which roughness length should be used. The results for temperature in Table 4.1 show some improvement when z_{0meff} is used instead of $z_{0mgrass}$, but this advantage disappears if hours with 0<L<10 are excluded (not shown). Further, the neutral approximation when $10<L<z_{31}$ (Method A, see section 2.2, note that here $z_{31}=40$) eliminates the bias in the temperature. For high wind speeds (near neutral conditions), the temperature errors are very small. This could be expected because of small vertical temperature gradients during neutral conditions. Taking $T_{2m}=T_{40m}+0.38$ for these situations results in almost the same T_{2m} error as via downward transformation.

If the observed Kin is used as input for the FRs, the RMSE for T decreases with 0.1 [°C] (for wind speed the improvement is very small). But even with the observed Kin as input, there still remains a contribution of the indirect determination of the cloud cover to the errors because cloud cover is used for the parameterisation of the incoming long wave radiation in the FRs.

Wind speed

For wind speed the improvement by applying $z_{0mgrass}$ instead of z_{0meff} is not significant if we realise that the reduction of the bias is trivial and might be a compensating error, as underestimation of the wind speed during stable conditions could be expected (section 2.2) and $z_{0mgrass}$=0.01m is smaller than the effective roughness length. Further, the insignificant reduction in the SD by applying a wind direction dependent z_{0meff} instead of $z_{0mgrass}$, is

disappointing. Despite the use of z_{0meff}, still a wind direction dependent wind speed error is found, similar to the one for unstable conditions. This phenomenon makes it impossible to choose on the basis of the results for Cabauw between a local or effective roughness length for the other 11 stations.

Excluding hours with $0<L<10$ does not result in smaller wind speed errors. This is not surprising because these very stable conditions only occur with low wind speeds, so the absolute error in the wind speed will also be small. On the other hand, during high wind speeds ($u_{40m} > 10$ m/s) near neutral conditions, the absolute error for u_{10m} will automatically also be high. However, the SD is even smaller than for all stable conditions, possibly due to the absence of stability effects. The large negative bias during these high wind speeds again suggests that the applied z_{0meff} is sub optimal for describing the wind speed profile between 40 and 10m in Cabauw. As Table 4.1 clearly shows, the neutral approximation for $10<L<z_{31}$ is not suitable for wind speed.

Humidity

The errors for specific humidity are quite reasonable, certainly if method A in combination with the exclusion of sunset hours is applied.

Conclusions

The question if $z_{0mgrass}$ or an effective roughness length should be used, is not solved with the experiments in this section. Not only the height above the surface plays a role whether the profiles are adapted to the local surface or not, but also the length of the fetch. If the observation instruments are surrounded by only 15m of grass land (or no grass land at all, as with some synop stations), the profiles can hardly be adapted to a short grass surface (corresponding with $z_{0mgrass}$). Partly based on these arguments, we present results with z_{0meff} for all parameters and conditions. On the basis of the results in this section, we further apply a straightforward, no exclusions transformation for wind speed. For temperature we use Method A (section 2.2) with the exclusion of unstable sunrise hours. Method A is also used for transformation of specific humidity but with the exclusion of sunset hours (section 2.2).

If surface fluxes are determined with the profile method (e.g. Holtslag and van Ulden, 1983), which could be considered as pseudo observations, and subsequently used in flux-profile relationships, the downward transformation errors are almost the same as with the use of the FRs. This is the case for stable and unstable conditions. So the use of the parameterisations in the FRs gives no significant contribution to the downward transformation error. Downward transformation errors arise from limitations of MO-theory (FP relationships), or wrong input parameters like roughness length.

5 Results

In this section verification results are presented of Hirlam, the physical post-processing using the FRs (described in section 2.2), and simple statistical methods (section 2.3). Although results are available of runs starting at 0, 6, 12 and 18 UTC, we here mainly present results as a function of the forecast period of the 0 UTC run to limit the amount of data. The results for different analysis times are quite similar (differences will be discussed). Nevertheless, in Appendix A results of runs started at 18 UTC can be found.

Because the results of the statistical regression must be validated on an independent dataset, all the presented results refer to the second half of the three-year dataset (unless mentioned otherwise), ranging from the 2nd of April 1998 until the 1st of Oct. 1999. The regressions are calibrated on the first half of the dataset, ranging from December 1 1996 until April 1 1998. Due to a too short independent set, the regression results on +30, +36, and +42h forecasts are not presented. Results for humidity are mentioned only briefly in the text, again to limit the amount of data and because of the minor importance of this parameter. For all the results, the atmospheric stability and therewith the hour of the day, plays a very important role. As mentioned before (section 2.2), applying the physical post-processing is questionable for stable conditions. Therefore results of the physical downscaling valid for 12 UTC (generally unstable conditions), are discussed more extensively in section 5.2.

5.1 Results as a function of the forecast period

Temperature

Land locations

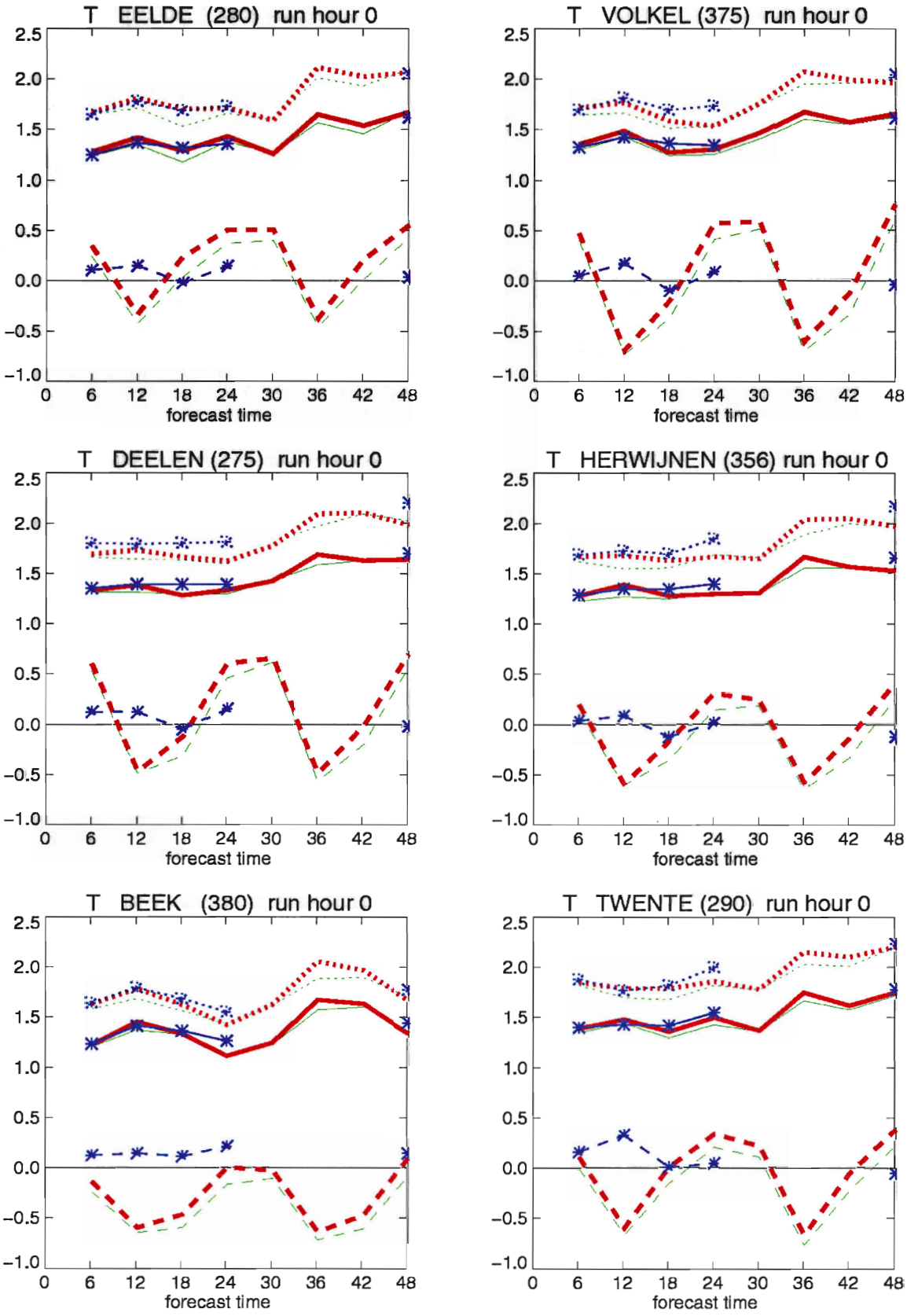

Figure 5.1a. Results for temperature [°C] for the land stations as a function of the forecast period starting at 0 UTC. Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam T_{2m} only. Regression data are missing for +30 until +42h forecasts.

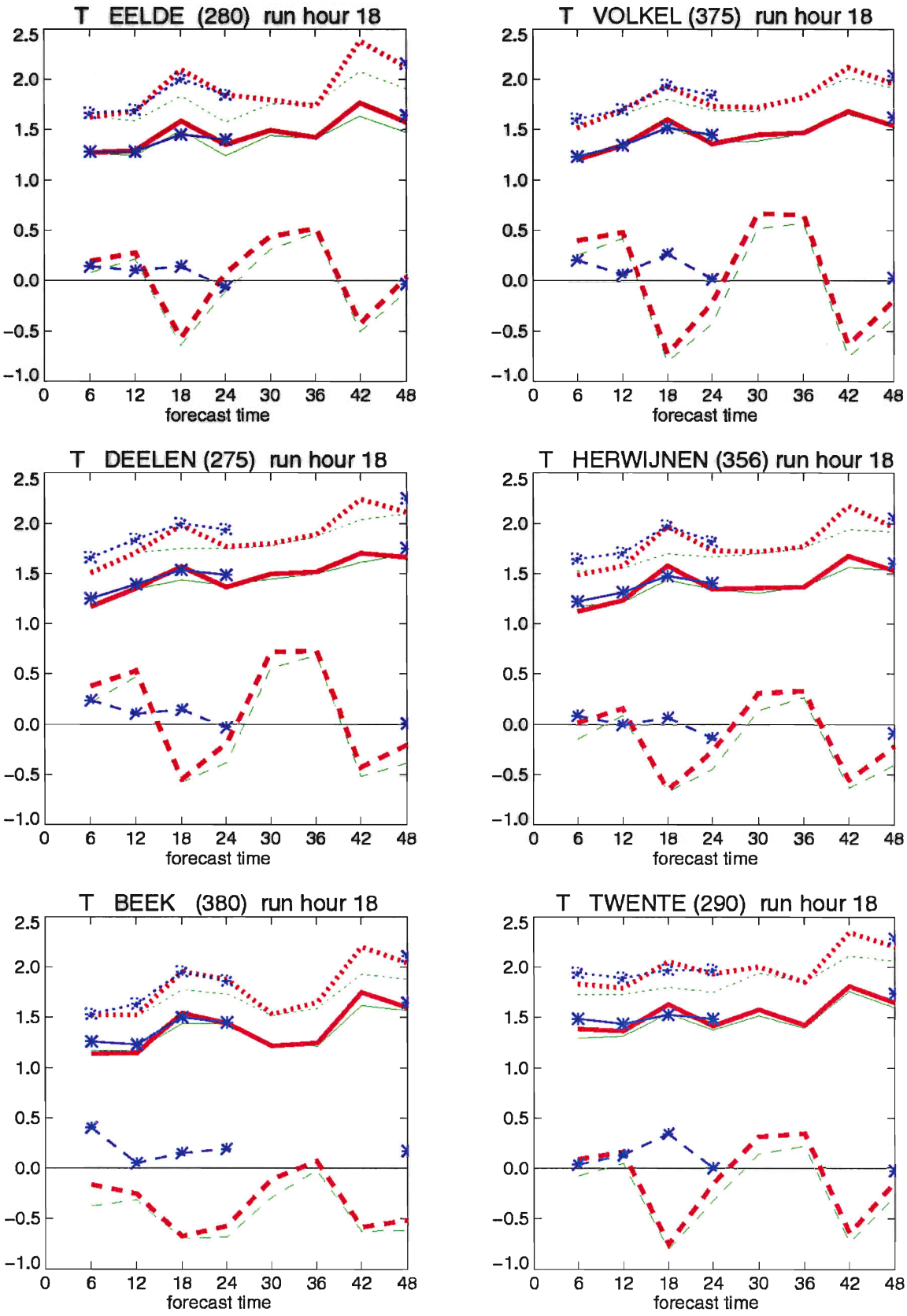

Figure 5.1b. Results for temperature [°C] for the land stations as a function of the forecast period starting at 18 UTC. Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam T_{2m} only. Regression data are missing for +30 until +42h forecasts.

As discussed in section 3.2, the stations can be categorised in land stations and stations where land/sea representation might play a role. Figure 5.1a and b for land stations shows that all locations have similar diurnal cycles for the temperature bias of the model, with relatively low temperatures at noon and high temperatures during the night. Unfortunately temperature biases are extremely difficult to interpret. The observed biases can be the result of (on average) too much clouds, too much soil moisture, errors in the turbulence scheme, etc., etc.. Consequently, improving the model in one respect does not necessarily lead to an improvement in the temperature bias due to the possible elimination of one of the many possible compensating errors. Anyhow, a known very important factor for the 2m-temperature bias is the soil moisture content (SMC). In Fig. 5.2 the bias in the temperature at 12 UTC as a function of the SMC is plotted for Twente. This plot is typical for other land stations and verification times. Overestimations of T_{2m} occur for all SMC values but large underestimations are restricted to high SMC values. For low values, the physical post processing reduces the 2m temperature. Note that in the FRs, the wetness of the soil is implicitly handled in the surface resistance by the correlation between the humidity deficit of the air and the SMC. If the FRs have a better parameterization of the SMC than Hirlam, the improvement with the FRs might be related to a reduction of the *NWP_error* and not the RP (note that a physical method should compensate for the RP only (section 2.1)). However, scatter plots (not shown) with temperature improvements applying the physical method against SMC in Hirlam do not show a clear correlation.

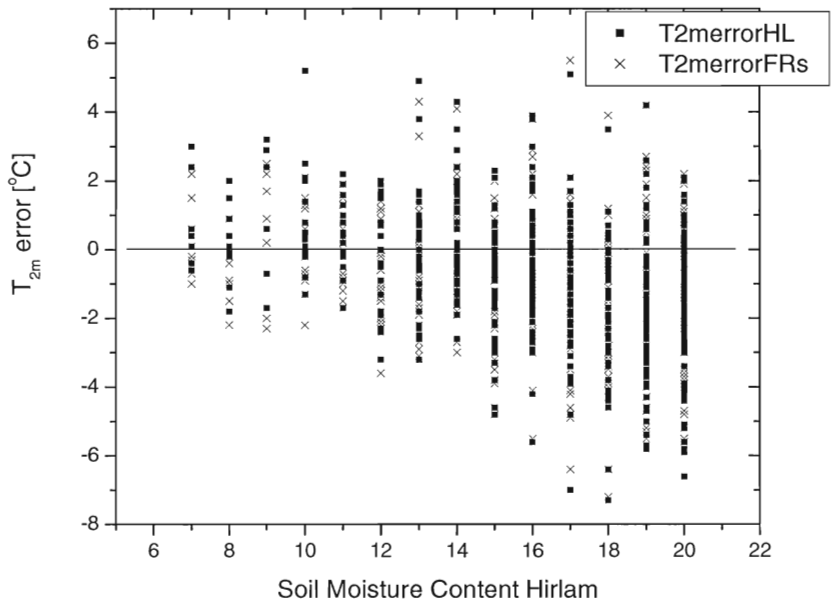

Figure 5.2 Error in T_{2m} [°C] for model output (T2merrorHL, squares), or the physical method using the FRs (T2merrorFRs, crosses) as a function of the Soil Moisture Content in Hirlam for Twente 12 UTC.

While the physical method does not seem to lead to a significant beneficial effect on the bias, the regression technique using just the T_{2m} of the model, efficiently eliminates almost the complete temperature bias. The major part of the temperature RMSE (not shown) is determined by the Standard Deviation (SD). All stations reveal a slowly increasing SD with forecast time, with local maxima at noon. Most pronounced is the second maximum, which for the 0 UTC run is the +36h forecast. Especially the runs started at 18 UTC (Fig. 5.1b) show much more pronounced first local maxima (+18h forecast) than the one of the 0 UTC run shown in Fig. 5.1a. As can be seen from Figures 5.1a and b, the FRs more or less eliminate the local maxima in the SD and as a result the improvements at noon with the

physical method are much larger for runs started at 18 UTC (about twice the improvement in SD). While it is not really clear why the first maxima in the SD is smaller for the 0UTC runs - possibly the relatively short forecast period plays a role - it could be expected that the FRs give the most improvement at noon. This issue is discussed in detail in the next section.

Figures 5.1a and b further demonstrate that the simple statistical method has no impact or even deteriorates the SD somewhat. The same is true for the MAE. However, we have to stress that only the Hirlam 2m temperatures are used in the regression. Improvement could have been obtained quite easily if more predictors were included. In fact, if the FRs output is added as a second predictor this gives by far the best results. This indicates that the physical post processing contains valuable information compared to NWP output. This will be discussed further in the next section.

As mentioned in section 3.3, Deelen and Herwijnen have the same nearest grid point and therefore the same *NWP_error*. Except for the bias, the temperature results for both stations are highly comparable. This is reflected in e.g. the high correlation (0.82 for +6h forecasts) between the *Total_errors* for T in Deelen and Herwijnen. This does not imply that there is no RP for Deelen en Herwijnen, but if there is a RP then these results indicate that there are no important differences in RP between both locations. Except for the roughness length (which has not much impact on T) and the lat/lon coordinates of the location, also the physical method uses the same surface characteristics for both stations. Apparently, the difference in soil type between Deelen and Herwijnen (see section 3.3) is not very relevant for the RP here. On the other hand, a factor which will certainly contribute to (random) differences in the RP between Herwijnen and Deelen, is the influence of cloud cover (section 2.1). The difference in bias between Deelen and Herwijnen can be explained by the difference in height of the location (section 3.3, Table 3.1)

Land/sea locations

As mentioned in section 3.3 Table 3.2, Valkenburg, Stavoren and Hoek van Holland are sea points in the Hirlam model, whereas Vlissingen and De Kooy are land points. Figure 5.3 shows striking large amplitudes in the diurnal cycle in the temperature bias for Valkenburg, Stavoren and Hoek van Holland. This is caused by the fact that for these locations the temperature measurements mostly reflect the land character. During daytime land surface warms up much more than the water surface temperature, leading to a large underestimation of T_{2m}. The opposite effect explains the overestimation at midnight. The distance to the coast is largest for Valkenburg and smallest for Stavoren, which probably causes the largest amplitude in the diurnal temperature bias cycle for Valkenburg and the smallest in Stavoren. In Stavoren also the relatively shallow water of the IJsselmeer might play a role in the relatively small amplitude in the diurnal cycle of the temperature bias. Mark that shallower water is less inert for temperature changes. It is important to note that even for such a small distance to the water like in Stavoren (although the measurement site is somewhat screened) the temperature measurements at 1.5m mostly reflect the land character (notwithstanding the existence of a tempering effect of nearby water on the temperature diurnal cycle).

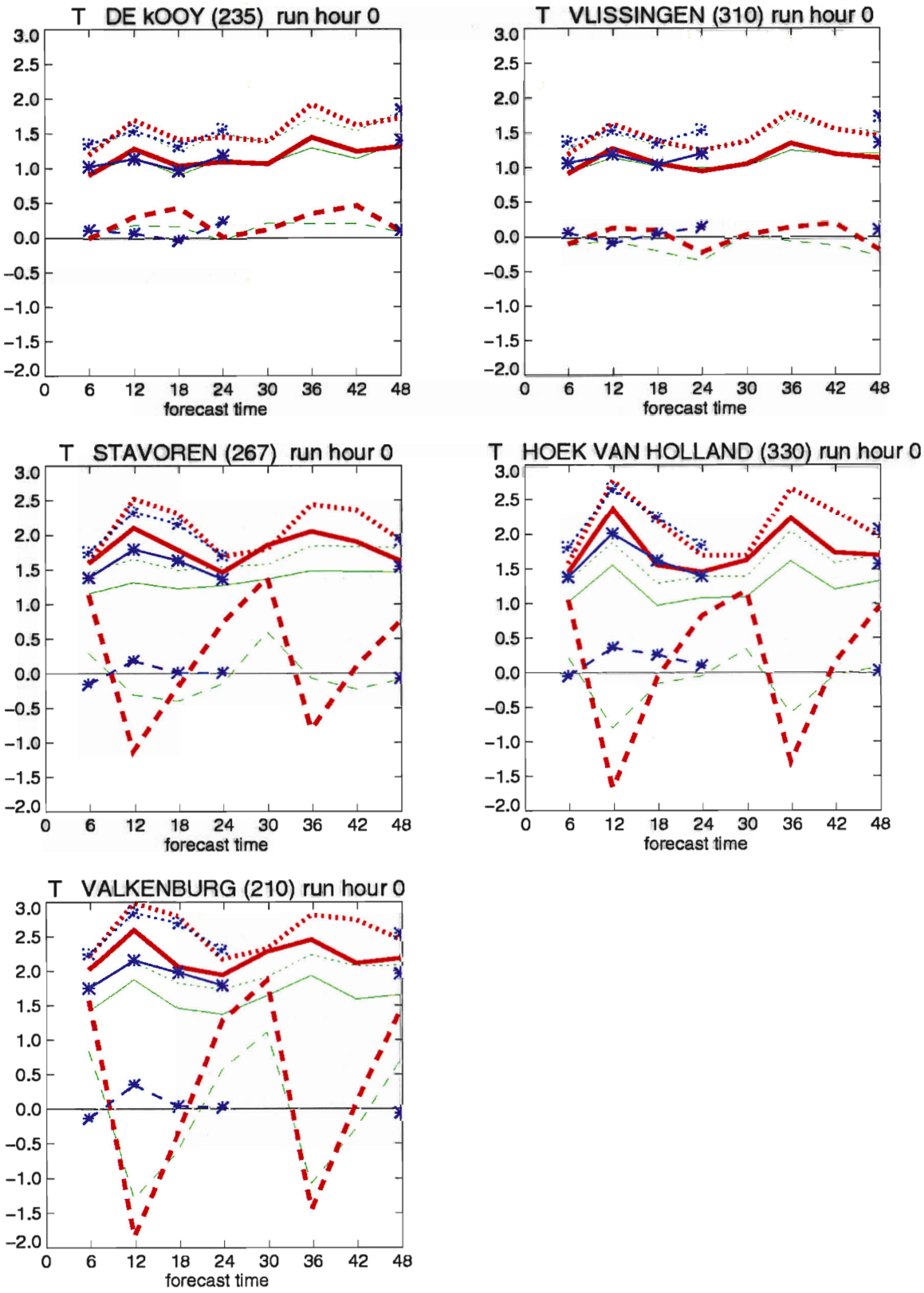

Figure 5.3 Results for temperature [°C] for the land/sea stations as a function of the forecast period starting at 0 UTC. Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam T_{2m} only. Regression data are missing for +30 until +42h forecasts.

At first glance, the improvement with the downward transformation in the bias for Valkenburg, Stavoren and Hoek van Holland looks modest. This becomes clear if the results are split up in a winter (underestimation by the model) and a summer season (overestimation by the model). For both seasons the improvement is substantial (not shown). If the results are presented for all seasons together (like in Fig. 5.3), this effect contributes to a better SD. Figure 5.3 further demonstrates that the simple linear regression effectively eliminates the bias, and reduces the SD for Stavoren and Valkenburg at 12 and 18 UTC somewhat but not for other times of the day. Again the statistical results improve significantly when output of the NWP model and the FRs are combined.

The large improvement for the three aforementioned stations with the physical downscaling method could be anticipated. Obviously, taking the nearest land grid point for Valkenburg will already reduce the model error significantly. Nevertheless, it is remarkable that the physical transformation from 32m to 2m is capable of compensating for almost the complete land/sea RP, as the results after transformation are comparable with other land stations. Therefore the physical post processing can be useful for countries like e.g. Finland with large areas with lakes and land where the nearest land grid point in the model might be hundreds of kilometers away.

For Vlissingen and De Kooy (land stations in Hirlam) the situation is completely different. While the diurnal cycle was amplified in the three aforementioned Hirlam sea point stations, for Vlissingen and De Kooy the cycle is inverted and has much smaller amplitude. Whereas the temperature rise during daytime will be in reality tempered due to the nearby water, the Hirlam will calculate a too high land surface temperature. During nighttime the opposite effect occurs. Due to the influence of the nearby water the 'normal' diurnal cycle of the temperature bias, as observed for the land stations (Fig. 5.1), is now overcompensated leading to the relatively small, inverse cycle for Vlissingen and De Kooy.

Wind speed

Land locations

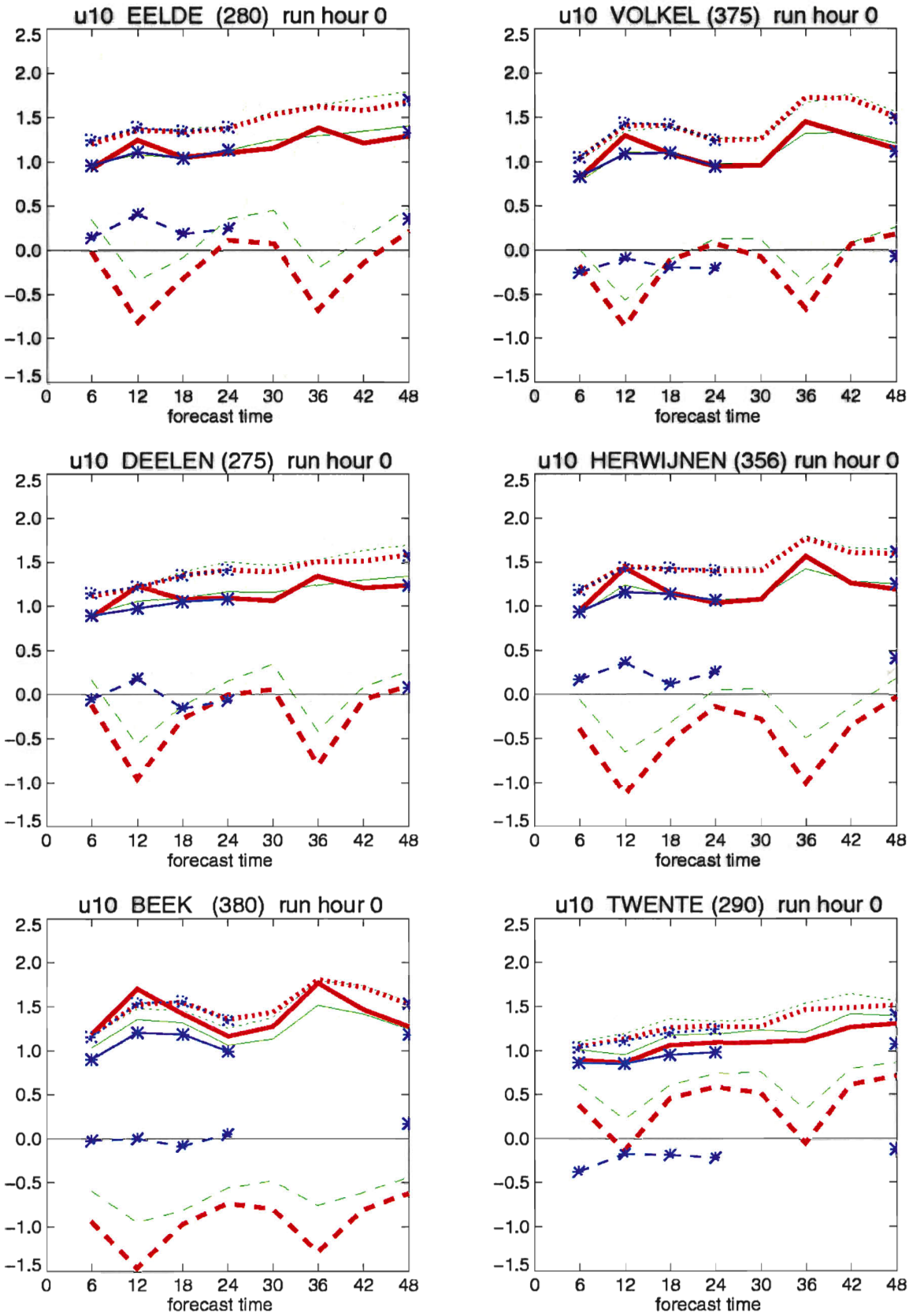

Figure 5.4. Results for wind speed [m/s] for the land stations as a function of the forecast period starting at 0 UTC. Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam u_{10m} only. Regression data are missing for +30 until +42h forecasts.

Analogous to Fig. 5.1 for temperature, the uniform diurnal cycle in the Hirlam wind speed bias is remarkable. The amplitude is about 1 m/s for all land stations with minimum at noon and maximum around midnight. On the other hand, the average bias over all forecast periods differs from station to station, with the largest underestimation for Beek and overestimation for Twente. In contrast with temperature, the bias in the wind speed is probably determined primarily by just two things:

- Difference between local roughness and model roughness, a RP

The difference between the mean actual roughness of the location and the roughness of the corresponding grid point (which is a clear example of a RP, see section 2.1) has a large impact on the mean bias of the wind speed. For example in Beek the difference between the actual and the model roughness is the largest (section 3.3), resulting in the largest underestimation by the model. For Twente the opposite is true, here differences in roughness are relatively small. Also in Herwijnen and Deelen (with the same *NWP_error*) the results are in agreement with the relatively smaller z_{0mloc} for Herwijnen (section 3.3). Note that also the influence of the roughness length difference will cause a small diurnal cycle because wind speeds are generally higher during daytime leading to a larger impact of the roughness difference during daytime. However, this effect is small because the difference in the amplitude of the diurnal cycle of the bias between Twente ($z_{0mHL} \approx z_{0mloc}$) and Beek ($z_{0mHL} >> z_{0mloc}$) is small.

- Deficiencies in the turbulence scheme of the model, a *NWP_error*.

Another important factor for the wind speed bias is the turbulence scheme in the NWP model. This scheme describes the mixing of the momentum and therefore influences the wind speed profile near the surface. In practice, turbulence schemes are not perfectly tuned for all stability regimes. Instead they have the tendency to underestimate the stability under (very) stable conditions and sometimes also the unstability during unstable conditions (personal communication Lenderink). This is e.g. also the case in the UKMO model (Best et al., 2000) where an overestimation at 0 UTC and an underestimation at 12 UTC in the modeled 10m wind speed is found. Such a (in principle too neutral) turbulence scheme helps the NWP model to prevent e.g. the total breakdown of turbulence during very stable conditions. However, as an unwanted side effect, vertical wind profiles are sub optimal. A too neutral turbulence scheme can explain the diurnal cycle in the wind bias in Fig. 5.4. During stable, nighttime hours it mixes too much momentum, leading to too high u_{10mHL}. On the other hand, during unstable, daytime hours, the scheme mixes not enough momentum, with too low u_{10mHL} as a result. As mentioned in section 2.1, errors in the turbulence scheme correspond to a *NWP_error*, which is supposed to be large-scale. Indeed the diurnal cycle in the wind speed seems to be quite uniform for all land locations.

More indications that roughness differences and the turbulence scheme are indeed for the major part responsible for the observed wind speed biases are mentioned further in this and the next sections.

Figure 5.4 shows that the physical post-processing improves the mean bias (z_{0m} effect) for most stations except Twente, possibly due to the too low local z_{0m} (see section 3.3). On the other hand, the diurnal cycles in the wind speed bias stay almost the same with the FRs approach. Because this diurnal cycle is most likely primarily caused by a *NWP_error*, the physical approach (which should only compensate for the RP) only reduces the average daily bias. However, the simple statistical method efficiently eliminates the bias including the diurnal cycle. Remember that linear regression techniques also compensate the *NWP_error*.

Another very interesting phenomenon can be observed looking at the wind direction dependent wind speed error in the model. As an example the results for Volkel 0 UTC are

shown. As Fig. 5.5 shows, there is a clear correlation between the wind direction dependent roughness length and the wind speed error made by the model. Similar results are found for all verification times and all stations with clear wind direction dependent z_{0m}. This result seems trivial: too high z_{0m} leads to too low wind speeds. However, at the same time this result is very important because here the rather strong influence of a RP on the *Total_error* is shown incontestably. Apparently, the effect of this RP is still visible in the noise of *NWP_errors*. Note that the correlation between wind direction dependent roughness length and wind speed error is also obscured by the fact that certain wind directions have higher wind speeds (and consequently larger errors) or occur more often. A tile scheme, which works with constant roughness lengths, will never be able to eliminate a significant part of the scatter, namely the scatter associated with the wind direction dependence of z_{0mloc}.

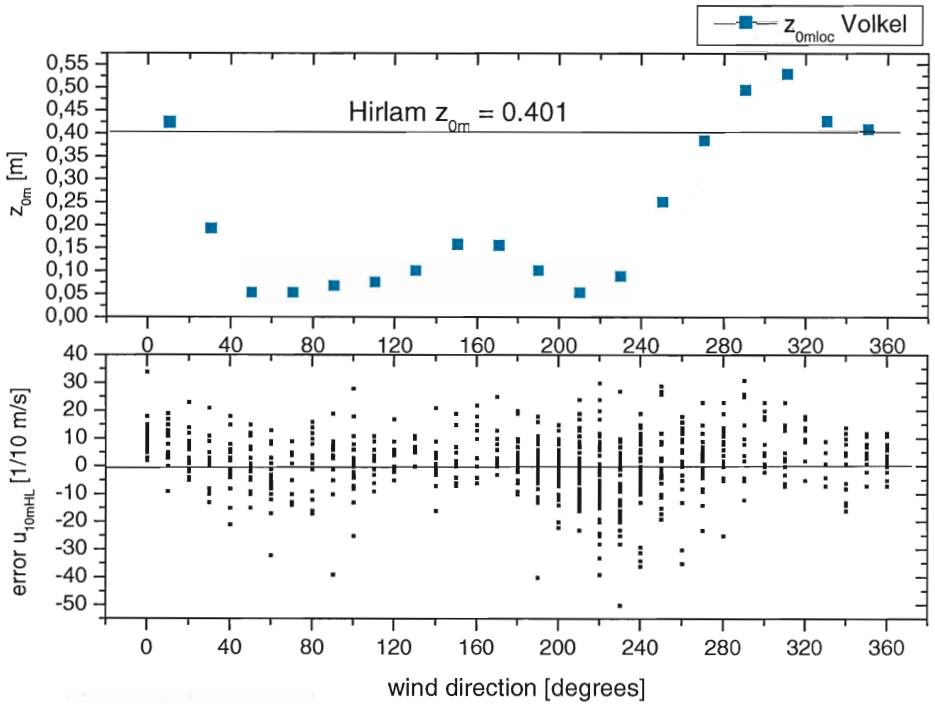

Figure 5.5 Local and model roughness length in Volkel as a function of the wind direction, and the error (u_{10mHL}-u_{10mobs}) [0.1 m/s] as a function of the wind direction for 0 UTC (18+6 forecasts).

The local maxima in the SD of Figure 5.4 are not as pronounced as for temperature. The physical post-processing nor the simple statistical method result in significant improvements in the SD. Unfortunately the wind direction dependent z_{0mloc} applied in the physical downscaling is not capable to significantly reduce the scatter. Nevertheless, using a constant z_{0mloc}=0.01m (not shown) increases the SD somewhat, which can be interpreted as an indication of valuable information in the wind direction dependent roughness length. In contrast with temperature there is difference in SD between Herwijnen and Deelen, probably due to the different locally valid roughness length. The correlation between u_{10m} *Total_errors* of Deelen and Herwijnen is only 0.33 (for +6h forecasts). However, part of this low correlation of u_{10m} errors in comparison with temperature errors (mentioned above), can be explained by the very low correlation at low wind speeds. Nevertheless, the low correlation for wind speed errors can be interpreted as an indication for relatively large differences in RP between Deelen and Herwijnen.

As expected, the transformation method only reduces the wind error to a small extent. To increase the impact of the local roughness length, the transformation should start at a higher level. Another possible improvement for the transformation method is to use the NWP wind direction at 10m instead of the lowest model level because during stable conditions the difference between the wind direction at 32 and 10m can be large.

Because the results with the physical method for unstable hours are discussed in more detail in the next section, we here discuss some marginal notes about the downscaling approach during stable conditions.

Besides the more fundamental problems of the physical concept during stable conditions (section 2.2), there is also the problem of large transformation errors during stable conditions (section 4). Both these problems might be circumvented by first transforming near surface values of the NWP model upwards, using the NWP surface characteristics, and subsequently downwards but this time with the locally valid surface parameters. Mark that errors during the upward transformation (which can be extremely large during very stable conditions) will be compensated during the downward transformation (de Rooy, 1995). By transforming upwards and downwards, only the effect of e.g. the roughness length can be incorporated in the wind speed (simply by taking different z_{0m} for the upward and downward transformation). Sensitivity experiments show that the correction on the u_{10m} due to a different z_{0m} can be made larger with the "up and down" technique (e.g. from 10m to 200m and back) than with transformation from 32m to 10m. Moreover, the height to which the upward transformation is done is not critical, as vertical gradients become small at greater heights. Only for temperature transformations during very stable conditions ($L<5$), there is a substantial impact if the transformation is done via e.g. 200m instead of 100m. This subject requires further study. In section 6 alternative methods for wind speed are presented. Also those methods have the advantage that they increase the influence of roughness length differences on the 10m wind speed. The methods of section 6 are also applicable during stable conditions.

Fortunately there are also conditions at midnight hours in which the additional error due to the current transformation approach is small, namely during neutral, strong wind conditions. It is important to realise that the results of high wind speed verifications are influenced by the choice for a model or observation threshold in the wind speed. For example a threshold in the observed wind speed of e.g. 8 m/s will inevitably lead to an underestimation of the modelled wind speed. On the other hand a threshold in the modelled wind speed will lead to a positive bias in the model wind speed. To ensure a neutral profile in the model and for the transformation, only model 10m wind speeds higher than 8 m/s are considered. For Volkel the transformation method now results in a modest reduction of the SD from 2.21 to 2.16 m/s, but the bias decreases from –0.88 to –0.08 m/s (19 cases). As described above, the model threshold should normally result in a positive bias but because the 32m wind of the model (the starting value for the downward transformation) is too low (due to $z_{0mHL}>z_{0mloc}$), there still is an underestimation. Note that compensating for the full (roughness) RP would result in a positive bias. Anyhow, these results suggest that for high wind speed situations during night time, large improvements can be obtained when information about the locally valid roughness length is used. This subject will be discussed further in section 6.2.

Land/sea locations

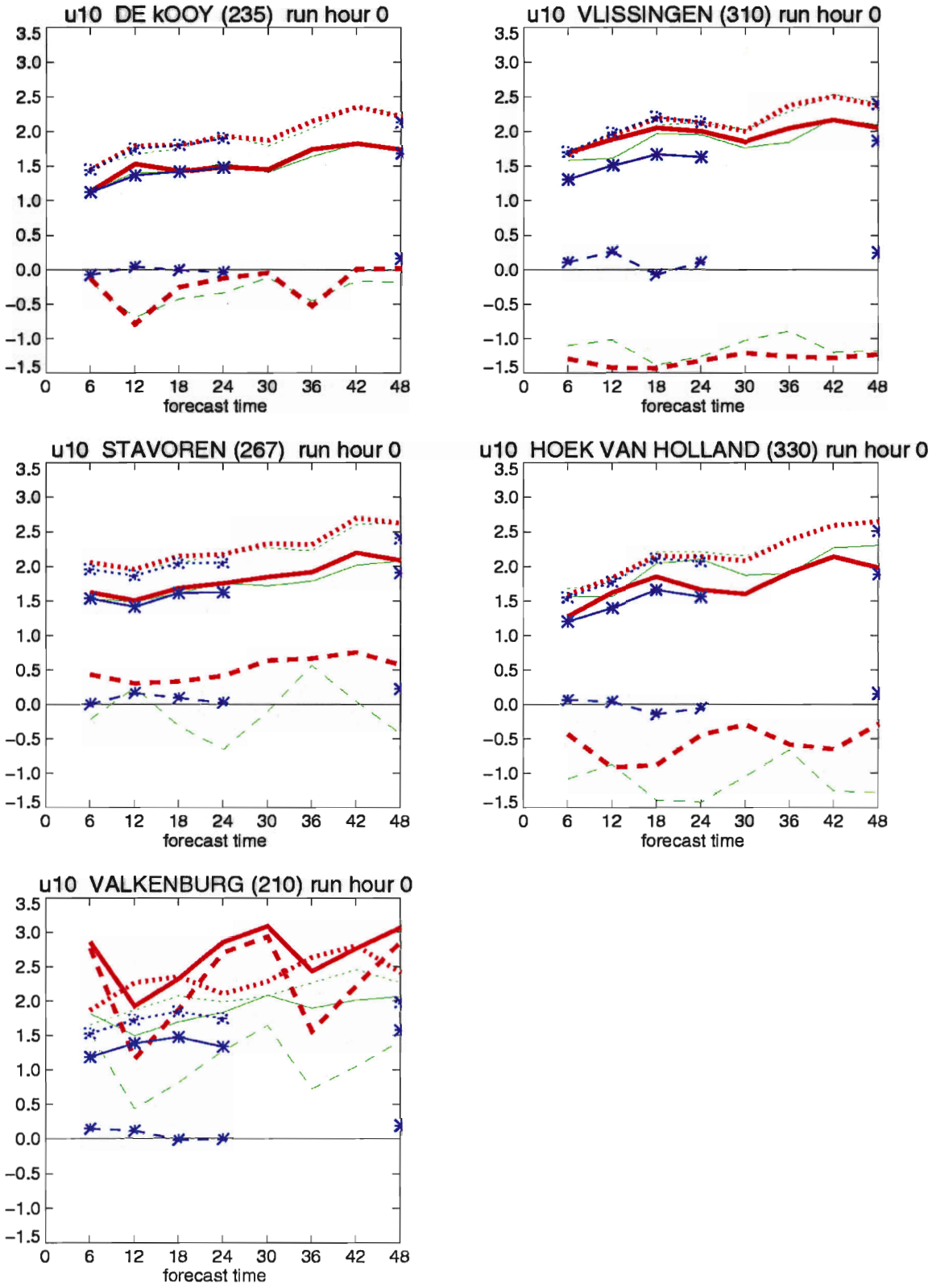

Figure 5.6. Results for wind speed [m/s] for the land/sea stations as a function of the forecast period starting at 0 UTC. Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam u_{10m} only. Regression data are missing for +30 until +42h forecasts.

Because Valkenburg winds reflect the land character but the nearest grid point is a sea point, the diurnal cycle in the bias is amplified and the wind speed is largely overestimated. Applying the physical post processing and especially the simple statistical techniques dramatically improves the results for Valkenburg.

Before the discussion of the stations with influence of the sea on the observed wind speed, it is important to mention a known problem with the water temperature, solved in the current operational model but still present in our validation set. During the summer months, the water temperature for the IJsselmeer and, in somewhat lesser degree, the North Sea coast is underestimated in the model (personal communication Sander Tijm). In the winter the biases are reversed. This phenomenon significantly influences the wind profile and thus the wind speed bias for coastal stations. Unfortunately, it also complicates the interpretation of the results.

An important factor for the wind speed bias in Hoek van Holland is the observation height of 15m instead of the standard 10m, which will result in approximately -0.5 m/s bias. The diurnal cycle in the wind speed bias reflects the influence of land stability on the wind speed profile. Although the locally valid z_{0m} is small in Hoek van Holland, the fetch over water is for many directions too short for the full adaptation of the wind profile to the stability above water (section 3.3). At noon the wind profile is in reality (mostly land) more unstable than in Hirlam (sea) leading to relatively large underestimations. The influence of the seawater temperature becomes evident with very large negative biases (up to -2.4 m/s) for wind coming from the sea during summer months (not shown). Overestimations only occur for directions with relatively large z_{0mloc} (between 70 and 100°) and stable conditions. For these directions there is almost no fetch over water and the in reality more stable conditions together with the high local z_{0m} overcompensate the influence of the 15m measuring height. Applying the physical downscaling as if Hoek van Holland is a land station does not work. The increased negative bias with the downward transformation arises from the higher (but true) local z_{0m} and the reversed diurnal cycle is the result of land stability put (too often) in the wind profile. The simple statistical regression technique nicely eliminates the bias for Hoek van Holland.

In Stavoren the situation is less complex (section 3.3). For wind from easterly directions, the model overestimates the wind speed. This can be attributed to a stronger roughness effect ($z_{0mHL}<z_{0mloc}$) than the 12m measuring height effect. On the other hand for westerly directions the wind speed is underestimated (mostly due to the 12m measuring height effect). At 12 UTC the errors are relatively small due to compensating stability errors. If the wind is coming from easterly directions, the actual (land) stability is more unstable than the model (sea) stability, which compensates for the too low z_{0m} in Hirlam. For westerly directions the z_{0mHL} is correct. However, in summer the water temperature is too low leading to too low wind speeds while in winter the opposite effect takes place. The diurnal cycle is hardly recognizable in Stavoren as for the most common directions the fetch over sea is very long. Analogous to Hoek van Holland, using the physical method, which assumes land characteristics, is a bad idea and introduces the (land) diurnal cycle in the wind speed bias. The simple regression technique finally, again eliminates the bias and improves the results.

In Vlissingen the wind speed is largely underestimated, without doubt for an important part caused by the measuring height of 28m. Another striking feature in Fig. 5.6 is the almost complete absence of a diurnal cycle in the wind speed bias. The most obvious reason for this is the influence of the sea on the stability for southerly wind directions (remember that the nearest grid point is land for Vlissingen). The most common influence of the sea will be to decrease the stability during nighttime and to decrease the unstability during daytime. Hence the model error in the turbulence scheme is counteracted leading to a

smaller diurnal cycle. Another contribution to the elimination of the 'natural' diurnal cycle in the bias, is the measuring height at 28m. At this height the climatological difference in wind speed between daytime (normally higher wind speeds) and nighttime is much smaller (from about 80m the diurnal cycle is even reversed). Consequently the underestimation during daytime and the overestimation during nighttime are compensated.
With the physical post processing some extra non-valid land stability is brought into the wind speed bias, although the mean bias is reduced (smaller z_{0m}). Also the scatter shows some improvement, possibly because Vlissingen shows a pronounced wind direction dependence of the wind speed error, associated with the locally valid z_{0m}. The statistical approach does not show the SD reduction but the bias almost disappears.

Results for De Kooy show, although with a somewhat smaller amplitude, the diurnal cycle in the wind speed bias, also evident at land stations. Apparently, in De Kooy observations of the wind mainly reflect the land character. For the temperature in De Kooy, on the other hand, we have seen a reversed diurnal cycle in comparison with other land stations (Figs. 5.2 and 5.3). In contrast with land stations, the diurnal cycle in the bias for T and u now seem to be non consistent. Consistent T and u diurnal cycles means too neutral profiles leading to too low wind speeds and temperatures at noon and the inverse at midnight. De Kooy shows us that the stability problem for T and u are not necessarily linked. Therefore it is less likely that the diurnal cycles in the biases (stability errors) arise from systematic errors in the surface temperature. More plausible is the hypothesis of systematic errors in the turbulence scheme.

Résumé

As mentioned in section 2.2, temperature and wind speed have totally different, even inverse, error characteristics. Mean wind speed errors are large when wind speeds are high. However, during high wind speeds, the atmospheric stability is neutral and vertical temperature gradients are small, leading to generally small temperature errors. On the other hand, temperature errors can be large especially during very stable conditions when vertical temperature gradients are large. During these (very) stable conditions, wind speeds are low and consequently wind speed errors are small.

So wind speed errors are dominated by those cases where stability corrections are unimportant, which simplifies the interpretation of the results for u. Moreover, it reduces the necessity to estimate the locally valid energy balance and the corresponding locally valid stability correction for wind speed. The difference between the local and the model z_{0m} plays an important factor in the RP, while the large-scale *NWP_error* in the wind speed is most likely dominated by systematic errors in the turbulence scheme.

The RP for T is, except for land/sea representation, not as transparent as for u (this subject will be discussed further in the next section). Moreover, when the RP for T is probably the most important, namely during (very) stable conditions, the here proposed physical downward transformation is inapplicable. As suggested, an upward/downward transformation approach might improve the results for these conditions.

5.2 Verification results for 12 UTC

As mentioned before there are several reasons to assume that the physical post processing can only be profitable for atmospheric conditions ranging from slightly stable (near neutral) to unstable conditions (section 2.2). In this section, we have to rely on a rather crude indicator of the atmospheric stability, namely the hour of the day. The 12 UTC and, to a lesser degree, the 18 UTC results will consist mainly of unstable conditions. Further summer months will contain more unstable cases than winter months. This is also reflected in

the less good results with the physical approach for winter than for summer months (especially for 18 UTC). To limit the effect of stable situations, we focus in this section on the performance of the physical method as a function of the forecast period in comparison with the Hirlam and simple statistical regression, for 12 UTC only.

Meteorological situations and the RP can be linked, e.g. 8/8 cloud cover will result in small horizontal variability in T_{2m}. A nice feature of all plots in this section is that all the points in a plot for one location refer to the same validation times and therefore the same observations. As a consequence the RP is constant during the forecast time, for all locations (this is also valid for other plots in this section). Hence, the increase of errors during the forecast can be attributed fully to the increase of the *NWP*_error. Remember that the RP is the difference between the observation and the "observed" mean of the grid box.

Temperature

Land locations

For all land locations the bias in the temperature is negative and deteriorates when the physical approach is applied, whereas the bias is almost eliminated with the statistical approach. As discussed in the previous section, the negative bias is probably the result of (many) model errors. An actual improvement (or reduction of the RP) does not necessarily results in an improvement of the T_{2m} bias. T_{2m} biases of the +12 and +24 forecast seem slightly smaller, possibly due to the relative good analyses at 0 and 12 UTC when more observations are available for the model analysis than at 6 and 18 UTC.

Here the physical method shows a reduction of the SD for all stations whereas the reduction with the statistical approach is smaller or absent. Remember that the statistical method is very simple. Applying e.g. wind direction as extra predictor for the regression might reduce the SD. Considering the results in Fig. 5.7 it is not surprising that the results for the statistical approach improve when the output of the physical method is included. Consistent with the fact that the RP stays constant during the forecast, the influence of the downward transformation on the SD and bias is rather constant. Yet, it is not absolutely necessary that the improvement with the downward transformation is constant during the forecast period because *NWP-errors* also influence the quality of the downward transformation. Unfortunately Figure 5.7 also shows us that the improvement on T_{2m} with the physical method is only small. This is caused by the small vertical temperature gradients during neutral to unstable conditions (this is not the case for q and u). Therefore it is doubtful if starting the downward transformation from higher levels will significantly increase the impact for these conditions. Although the impact on the SD and MAE with the FRs looks small, it is useful to present the results as RMS skill (see Appendix B). Throughout this paper, RMS skill is defined as the skill against direct model output.

$$RMS\_Skill\_A = \frac{RMS_{Hirlam} - RMS_A}{RMS_{Hirlam}} * 100\%$$

The figures in Appendix B show the rather constant positive but small skill for land locations. Clearly the temperature output of the physical downscaling gives surplus value if it is used in the regression, together with the standard T_{2m} of Hirlam.

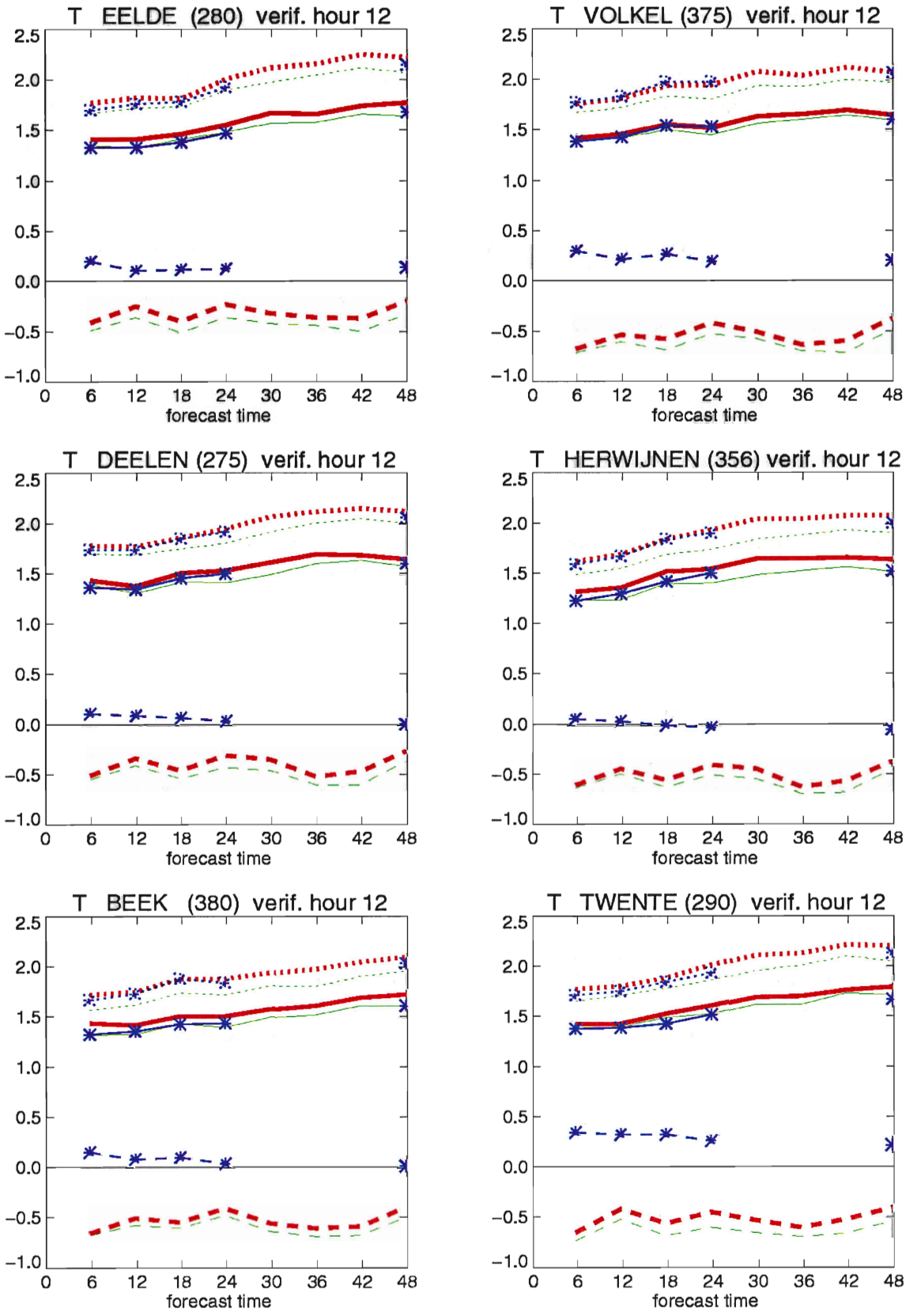

Figure 5.7 Results for temperature [°C] for the land stations valid for 12 UTC (different analysis times). Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam T_{2m} only. Regression data are missing for +30 until +42h forecasts.

Before we can answer the question if the downward transformation reduces the RP for temperature, it is important to understand why the downscaling reduces the scatter in T_{2m} at 12 UTC. Therefore the temperature results for Beek are investigated in more detail. First of all, improvements and deteriorations with the downscaling are not clustered in time, e.g. three-days periods with only excellent results. Mark that the temporal resolution of the results (6 hours) is coarse and it is likely that some clustering could be seen with hourly resolution. Looking at good results, we see that just the albedo or z_{0m} difference between the model and the downward transformation can only explain a very small part of the improvement. This is investigated by changing the albedo or z_{0m} for the downward transformation to model values.

The largest improvements occur during very stable conditions and can be explained by the neglection of the term $\Psi(z_{0mHL}/L)$ (see section 2.2) in the flux-profile relationships used in the model. Because z_{0mHL} can be large and L is small during very stable hours, the term $\Psi(z_{0mHL}/L)$ can become important. However, this only occurred for a few cases and has no significant impact on the overall results. Moreover, this phenomenon cannot explain the good results at 12 UTC because very stable conditions do not occur at noon. The most likely remaining explanation for the decreased scatter is that the surface energy balance is better resolved in the FRs leading to more suitable θ_* and L than the ones in the model and thus a better transformation to 2m. Note that the FRs, in contrast with the model, solve the energy balance for a vegetated surface with short grass. From the downward transformation test in section 4 and from RH99, we knew beforehand that the errors in the energy balance and the downward transformation would be small at 12 UTC. Consequently, if the errors at the lowest model level are smaller than near the surface, as could be expected for 12 UTC (De Rooy, 1999), the downward transformation must lead to improved T_{2m}.

Land/sea location

The largest improvements with the physical method are found at locations where the nearest grid point is sea, namely Valkenburg, Stavoren, and Hoek van Holland (Fig. 5.8). As shown in Appendix B the skill for the physical method is around 25% and, except for Valkenburg, the physical method alone is approximately as good as the regression on both model and physical post processed output. For De Kooy and Vlissingen the cooler output with the FRs this time turns out beneficial (accidentally) while the improvement in SD is about the same as for land stations. The skill with the downscaling for these stations is almost 10% (Appendix B).

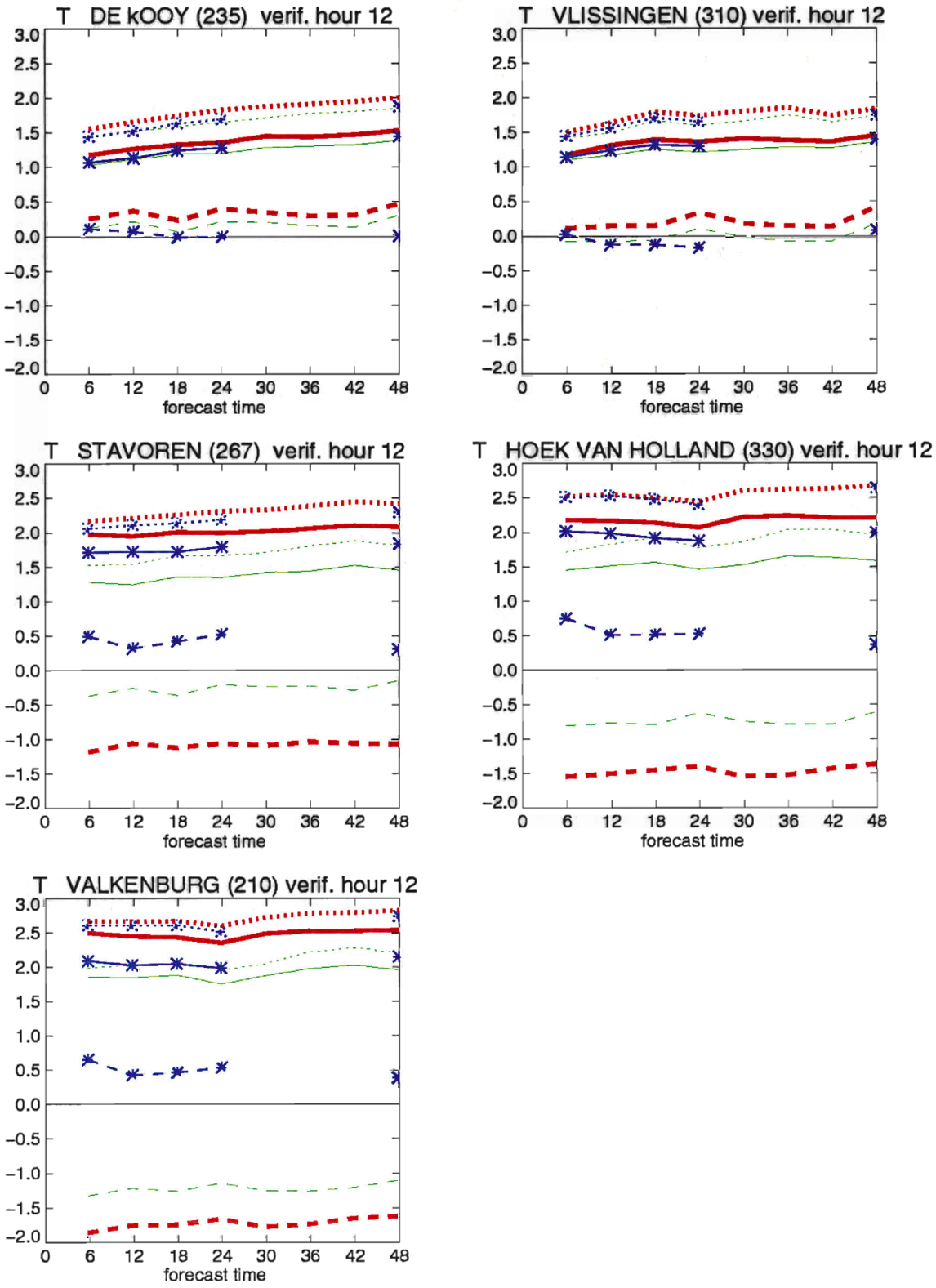

Figure 5.8 Results for temperature [°C] for the land/sea stations valid for 12 UTC (different analysis times). Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam T_{2m} only. Regression data are missing for +30 until +42h forecasts.

Wind speed Land locations

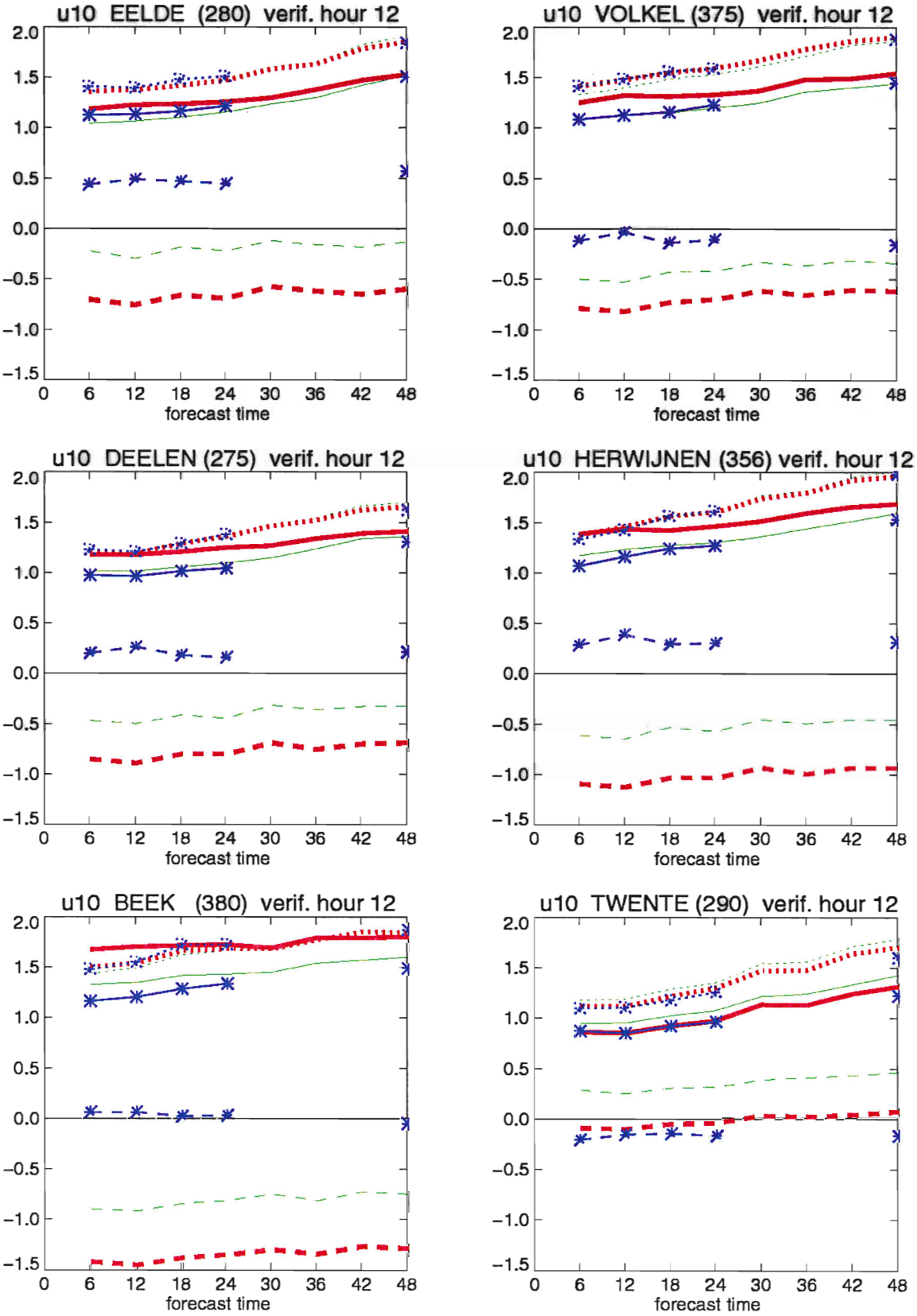

Figure 5.9 Results for wind speed [m/s] for the land stations valid for 12 UTC (different analysis times). Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam u_{10m} only. Regression data are missing for +30 until +42h forecasts.

Except for Twente, the NWP model output underestimates the wind speed. The downscaling or the linear regression on just the Hirlam wind speed reduces this bias. As mentioned in the previous section, the wind speed bias is probably primarily caused by a deficiency in the turbulence scheme (which leads to underestimations at noon) and a z_{0m} in the model that is not representative for the local situation. The physical method does not eliminate the complete bias partly because it does not compensate for model errors. Moreover, to compensate for the full z_{0m} RP, the downward transformation should start at a higher level (section 2.2 Pitfalls).

Only in Volkel the SD is somewhat reduced with the physical approach (possibly due to the highly variable z_{0mloc} for different wind directions), while for other stations neither of the post processing techniques have impact. Therefore, only for Volkel the regression benefits from the output of the physical method. Eelde is the only location where the physical method performs better than other approaches.

Land/sea locations

Valkenburg is the only Hirlam sea location with in reality clear land characteristics for wind speed (section 3.3). Accordingly, the improvement with the physical method in Valkenburg is substantial, although the simple statistical method performs even better. The other two stations with a corresponding Hirlam sea point, Stavoren and Hoek van Holland show much smaller impact with the post-processing methods. The underestimation in Hoek van Holland, caused primarily by the 15m measuring height, is eliminated by the statistical method but, because of the higher (and true) z_{0m} in the downward transformation, increased with the downscaling. Obviously, transforming downwards to 15m would improve the bias in Hoek van Holland substantially.

For Vlissingen and De Kooy the SD is somewhat improved with the physical method (remember that especially in Vlissingen z_{0mloc} is highly variable with wind directions), which is also the reason why using the downscaling output in the regression gives surplus value for De Kooy and Vlissingen (Fig. B5 AppendixB)

Overall, The linear regression seems to be the most appropriate approach to improve the Hirlam wind speed. This is also illustrated in Figs. B3 and B4 in Appendix B.

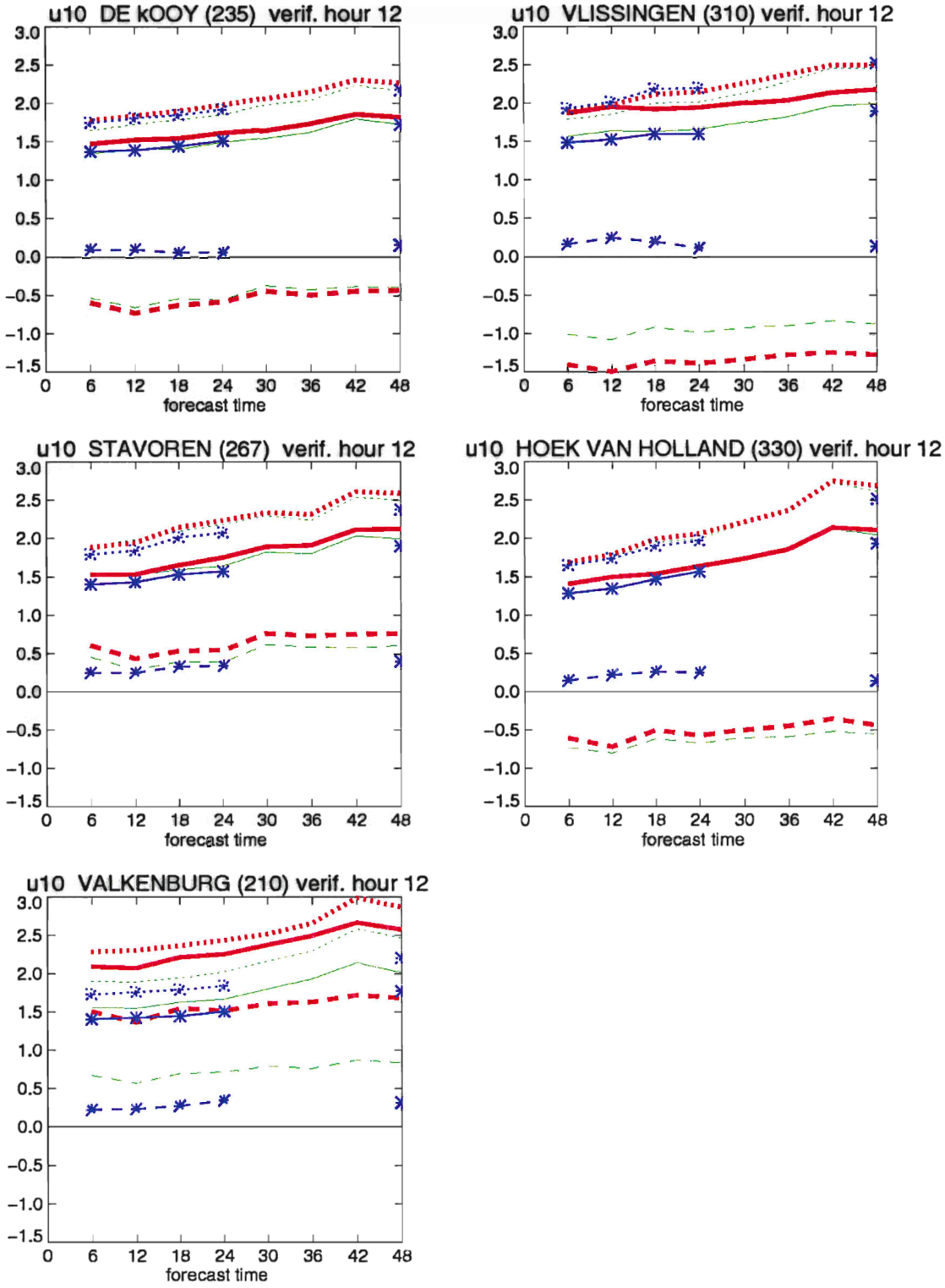

Figure 5.10 Results for wind speed [m/s] for the land/sea stations valid for 12 UTC (different analysis times). Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam u_{10m} only. Regression data are missing for +30 until +42h forecasts.

6 Generation of local wind speed fields

6.1 The pool method

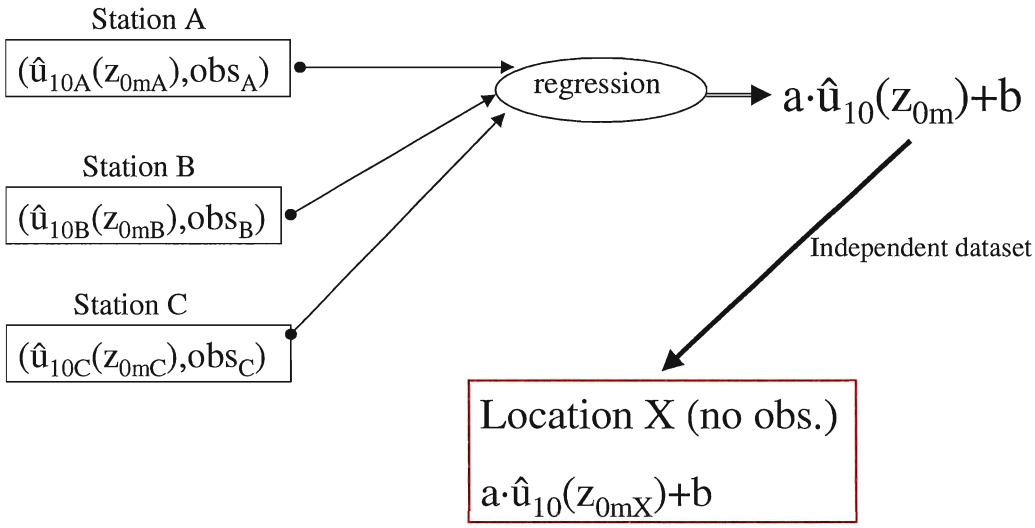

Figure 6.1 Outline of the combined physical/statistical method for the estimation of wind speed at locations without observations (here denoted as location X). Further explanation can be found in the text.

From the results in the previous sections, it becomes clear that analyzing the wind speed error is less complicated than analyzing the temperature error. The results suggest that the wind speed error is primarily determined by a RP, associated with the difference between the locally valid z_{0m} (z_{0mloc}) and the model z_{0m} (z_{0mHL}), plus a *NWP_error* bias probably originating for the larger part from a deficiency in the turbulence scheme. Both errors can be reduced with a simple statistical method and the physical method seems redundant. As expected (section 2.1) the *NWP_errors* are horizontally rather homogeneous, whereas the RP changes from location to location. This implicates that, as far as the *NWP_errors* are concerned, a regression determined for a certain location with observations, can also be useful for another location (e.g. without observations). However, this is not the case for the horizontally variable RP part. To circumvent these problems we investigate a method that compensates for the *NWP_error* and the RP but is also applicable for locations without observations, e.g. when fields of wind speed have to be produced. The outline of this method is depicted in Fig. 6.1. In the first step of the method, NWP model output for a certain station is mostly stripped from the RP, using z_{0mloc} information of that particular station. Different ways to combine NWP model output with z_{0mloc} information to estimate a locally valid u_{10m}, using so-called estimators, are described beneath. The resulting u_{10m} without (or with less) RP is denoted as e.g. $\hat{u}_{10A}$ for station A (Fig. 6.1). In this way we get pairs (e.g. $(\hat{u}_{10A}(z_{0mA}),obs_A)$) of estimated and observed 10m wind speeds. Differences between both components of these pairs are supposed to be the result of *NWP_errors*. Because *NWP_errors* are horizontally homogeneous, we can derive one single regression from a pool of pairs of different stations (stations A, B and C in Fig. 6.1). In the regression the estimator, $\hat{u}_{10}$, is used as single

predictor for the observations. Preferably, the pool must cover the complete range of possible z_{0mHL}-z_{0mloc} values. However, as long as land/sea representation or deviating measuring height is not included in the estimator(s), land/sea locations (for definition see section 3.2) or stations with deviating measuring height should be handled in a separate pool and a separate regression. After the regression on the pool, the regression coefficients can be used on a new location (location X in Fig. 6.1) to compensate for the *NWP_errors*. Finally, the RP on the new location is handled by substitution of the locally valid z_{0mloc} in the estimator.

Naturally, the method using estimators based on local roughness length can only work under the condition that reasonable estimates for z_{0mloc} can be made for the locations of interest. In practice the z_{0mloc} information can be derived from high resolution land use or height maps. A recently derived method described by Verkaik and Smits (2001) has a resolution of about 100x100m. Verkaik (personal communication) showed that his z_{0m} derived from land use maps correlates well with z_{0m} derived from gustiness analyses, as the ones used in this paper. Therefore we expect, that the here-described method will perform comparably when wind speed fields are generated using roughness lengths derived from land use maps.

6.2 Introduction of the estimators

As described above, we need an estimator for the 10m-wind speed in which z_{0mloc} information is included. Here different alternative estimators are introduced. We start with the description of two possible estimators that assume neutral stability.

Estimator A

Analogous to (2.5) in section 2.2, but without stability correction, the flux-profile relationships for Hirlam (HL) and the local station (loc) can be written as:

$$u_{10mHL} = \frac{u_{*HL}}{\kappa} \cdot \ln\left(\frac{10}{z_{0mHL}}\right) \qquad (6.1a)$$

$$u_{10mloc} = \frac{u_{*loc}}{\kappa} \cdot \ln\left(\frac{10}{z_{0mloc}}\right) \qquad (6.1b)$$

As a crude approximation one can state:

$$u_{*HL} = u_{*loc} \qquad (6.2)$$

However, in reality u_* is strongly influenced by the roughness length itself. Anyhow, (6.2) together with (6.1a) and (6.1b) leads to:

$$\hat{u}_{10mloc} = u_{10mHL} \cdot \frac{\ln\left(\frac{10}{z_{0mloc}}\right)}{\ln\left(\frac{10}{z_{0mHL}}\right)} \equiv estimator\_A \qquad (6.3)$$

So, estimator_A assumes neutral stability and is based on the debatable assumption $u_{*HL}=u_{*loc}$. Because z_{0mHL} is larger than z_{0mloc} most of the time, also u_{*HL} will be larger than

u_{*loc} and consequently the correction on u_{10mHL} will be too strong, resulting in generally too high u_{10mloc}.

Estimator_B

For neutral conditions, the 140m-wind speed (second lowest model level) can be written as:

$$u_{140mloc} = \frac{u_{*loc}}{\kappa} \cdot \ln(\frac{140}{z_{0mloc}}) \qquad (6.4)$$

If we apply surface layer theory, which states that the friction velocity is constant in the vertical, (6.4) and (6.1b) can be combined, leading to:

$$u_{10mloc} = u_{140mloc} \frac{\ln(\frac{10}{z_{0mloc}})}{\ln(\frac{140}{z_{0mloc}})} \qquad (6.5)$$

As mentioned in section 2.2, it is plausible that wind speed is less influenced by local terrain conditions (like z_{0m}) at larger heights. At 140m it can be expected that, although depending somewhat on the (stability) conditions, the wind speed for local and Hirlam conditions is almost the same. Or, to put it in other words, the blending height is at or beneath 140m.

$$u_{140mloc} = u_{140mHL} \qquad (6.6)$$

Now (6.5) can be rewritten as:

$$\hat{u}_{10mloc} = u_{140mHL} \frac{\ln(\frac{10}{z_{0mloc}})}{\ln(\frac{140}{z_{0mloc}})} \equiv estimator\_B \qquad (6.7)$$

Note that for other (measuring) heights, (6.7) can easily be adapted by replacing 10 with the requested height.

Estimator_B has a theoretically sounder basis than estimator_A. Nevertheless some remarks should be made here. First, the surface layer theory (SL theory) is a commonly accepted framework but formally it can only be applied on the lowest 10% of the planetary boundary layer. On the other hand, different studies (e.g. Holtslag, 1984) showed that in practice the SL theory could be extrapolated to larger heights. However, for stable conditions it is arguable whether the assumption that u_* is constant in the vertical, can be applied up to 140m. For very stable conditions, it can be expected that u_{140m} and u_{10m} are decoupled, which makes estimator_B even useless. For these decoupled conditions, estimator_A has the advantage that it works with only the 10m-wind speed. A simple way to reduce the problems with estimator_B during stable conditions is to replace 140m by 32m (the lowest model level) wind speeds. Finally, also estimator_B does not include any stability effect and will consequently (apart from the z_{0m}-effect) overestimate u_{10mloc} for stable and underestimate u_{10mloc} for unstable situations.

Estimator_C

For this estimator, estimator_B is used as a basis, but now the stability correction of Hirlam between 140 and 10m is added.

$$u_{10mHL} = u_{10mHL\_neutral} + u_{*HL} \cdot \Psi_{HL} \tag{6.8a}$$

$$u_{10mloc} = u_{10mloc\_neutral} + u_{*loc} \cdot \Psi_{HL} \tag{6.8b}$$

where Ψ_{HL} is the stability correction of Hirlam [m/s]. For the neutral estimation of u_{10m}, estimator_B is used with z_{0mloc} and z_{0mHL} for resp. $u_{10mloc\_neutral}$ and $u_{10mHL\_neutral}$.

In the further derivation of estimator_C, we use approximation (6.2) in (6.8a) and (6.8b), although this approximation is labeled as crude in the derivation of estimator_A. However, there are several reasons why this approximation can be made in the second term on the RHS of (6.8b) but is questionable for estimator_A. Firstly, the major part of the wind speed is determined by the neutral part and the stability correction is only a small correction on the neutral estimation. Therefore the approximation on a small correction has only a small influence on the ultimate u_{10mloc}. Secondly, also Ψ_{HL} includes a u_{*HL} term, namely in the Obukhov length (see section 2.2). As formulae (6.8b) shows, the stability correction from Hirlam is also applied for the local situation. Now if e.g. $u_{*loc}<u_{*HL}$, this implies that $L_{loc}<L_{HL}$ and consequently $\Psi_{loc}>\Psi_{HL}$ (or vice versa). So the error due to the approximation $\Psi_{HL}=\Psi_{loc}$ is always counteracted by the approx. $u_{*loc}=u_{*HL}$. Anyhow, there is no real need to use approximation (6.2) because u_{*loc} can be easily determined from e.g. (6.5). As could be expected, results for estimator_C when u_{*loc} is used in (6.8b) are almost exactly the same as when u_{*HL} is used for the second term on the RHS of (6.8b). Differences are limited to a few 0.01 m/s but most of the time in favor of the version with u_{*HL} in front of the Ψ_{HL} term. All in all, we decided to use approximation (6.2) in (6.8b).

Combining estimator_B (with z_{0mHL}) for $u_{10mHL\_neutral}$ and (6.8a) results in:

$$u_{*HL} \cdot \Psi_{HL} = u_{10mHL} - u_{140mHL} \cdot \frac{\ln(\frac{10}{z_{0mHL}})}{\ln(\frac{140}{z_{0mHL}})} \tag{6.9}$$

in which the plausible assumption $u_{140mHL\_neutral}=u_{140mHL}$ is made. So finally the Estimator_C can be constructed, consisting of a neutral part with a correction for the locally valid roughness length plus the stability correction from the NWP model.

$$\hat{u}_{10mloc} = u_{10mHL} + u_{140mHL} \cdot \{\frac{\ln(\frac{10}{z_{0mloc}})}{\ln(\frac{140}{z_{0mloc}})} - \frac{\ln(\frac{10}{z_{0mHL}})}{\ln(\frac{140}{z_{0mHL}})}\} \equiv estimator\_C \tag{6.10}$$

Note that ψ_{HL} is the stability with the known errors due to deficiencies in the turbulence scheme. However, we want to compensate for the RP in the estimator, not for the *NWP_error* (wrong stability). The *NWP_error* must be dealt with by the regression.

As explained in the discussion of estimator_B, it is questionable if the use of u_{140m} for the estimation of u_{10m} is useful during (very) stable conditions. A simple reduction of this problem is to use wind speed at 32m. This estimator will be called estimator_C32.

$$\hat{u}_{10mloc} = u_{10mHL} + u_{32mHL} \cdot \{ \frac{\ln(\frac{10}{z_{0mloc}})}{\ln(\frac{32}{z_{0mloc}})} - \frac{\ln(\frac{10}{z_{0mHL}})}{\ln(\frac{32}{z_{0mHL}})} \} \equiv estimator\_C32 \qquad (6.11)$$

Many of the above-mentioned approximations in estimators A, B and C might be of minor importance once they are used in a regression. In the statistical optimization process, systematic deficiencies of the estimators are compensated by adaptation of the regression coefficients.

6.3 Results per station

Before the estimators, introduced in the previous section, are tested in a pool with several land stations (Fig. 6.1), we first investigate their performance if they are applied per station separately. In this way, an idea can be given of the optimal performance using the estimators and therewith of the deterioration due to the (sub optimal) pooling. Moreover, the difference between the performance when only the estimator is used per station (so without regression) can be compared with the performance for the same station but in the pool experiment (so with regression). This shows us the improvement due to the use of the universal regression on a location without observations.

Additionally, results for land/sea locations are presented here, although they cannot directly be used in the pool experiment as explained before. Besides with the estimators from the previous section, we also did experiments where the estimator_B and the stability correction (6.9) were offered separately (as two predictors) for the regression. The idea behind this experiment was that the NWP error depends on the stability in the model and therefore a better optimization was expected. However, the improvement in comparison with estimator_C turned out to be too small to present these results.

The performance of the standard u_{10mHL} is compared with the results from the estimators, with and without linear regression. To make a comparison possible, the regressions are determined on the first half of the three-year dataset whereas all the presented results, including the results of the standard NWP output, refer to the second, independent, half of the data set. To investigate the performance under stable and unstable conditions, results for 12 and 0 UTC are discussed separately.

Verification time 12 UTC

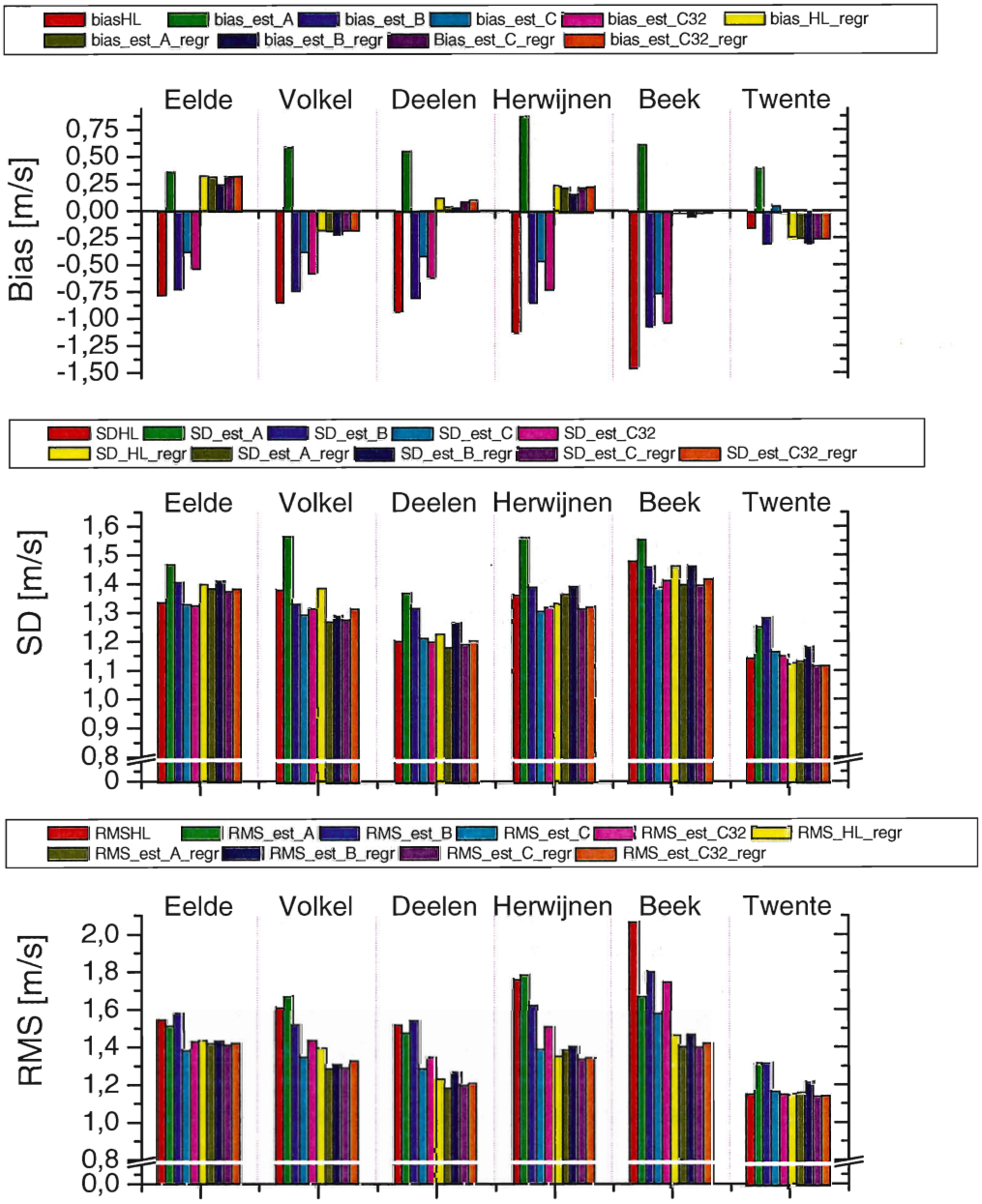

Figure 6.2 Wind speed results [m/s] per land station for Hirlam output (HL) and different estimators (_est_) with (_regr) and without regression. The results are based on 6+6h forecasts valid for 12 UTC.

Land locations

Results of sections 5.1 and 5.2 show that for most land locations the Hirlam predictions have a negative bias at 12 UTC (Fig. 5.4). Partly due to the turbulence scheme, which on the average gives not enough mixing for this hour of the day, partly due to the generally higher z_{0mHL} than z_{0mloc}. If we exclude Twente (which again shows an atypical behavior, probably due to the use of non-valid z_{0mloc} values), Fig. 6.2 shows that the negative bias for u_{10mHL} differs considerably, namely from -0.78 m/s in Eelde to -1.12 m/s in Herwijnen and even -1.44 m/s in Beek. With Estimator_C without regression we hope to compensate the z_{0m} RP, leaving just the *NWP_error*, which should be rather homogeneous for the different land stations. Apart from Twente and Beek, the bias is indeed nicely constant, around -0.4 m/s ± 0.05 m/s. So if indeed the RP is eliminated with Estimator_C, -0.4 m/s

must be the bias due to the NWP error at 12 UTC. The most likely reason for the anomalous bias with estimator_C in Twente is already mentioned before. Also in Beek, the bias using estimator B, C or C32 differs from the rest of the land stations (-0.75 m/s for estimator_C). Maybe the transformation should start at a higher level in Beek because the most common upstream grid points in Hirlam also have very high z_{0mHL} values (due to an orographic contribution). As a consequence, u_{140mHL} in Beek might be influenced by these upstream high z_{0mHL}-z_{0mloc} values, leading to too low u_{140mHL} and therefore also too low $\hat{u}_{10m}$ with predictor B, C, or C32. Nevertheless for the other land stations the negative bias becomes remarkably constant applying e.g. estimator_C. As could be expected on theoretical grounds, estimator_C performs better than other estimators or direct Hirlam output, also for other error measures like SD, RMS and MAE. Also expected are the large underestimations with the "neutral" estimator B. Because normally the conditions will be unstable at 12 UTC, the assumption of neutral conditions results in not enough mixing and consequently too low wind speeds at 10m. Due to the "transformation" from a lower height, estimator_C32m does not compensate the negative bias as much as the 140m version does. Moreover, also for other error measures estimator_C with u_{140mHL} is preferable. Estimator_A performs the worst, it overcompensates the negative bias and besides that, results in the largest scatter (as explained in section 6.2, due to the assumption $u_{*HL}=u_{*loc}$ in combination with $z_{0mHL}>z_{0mloc}$).

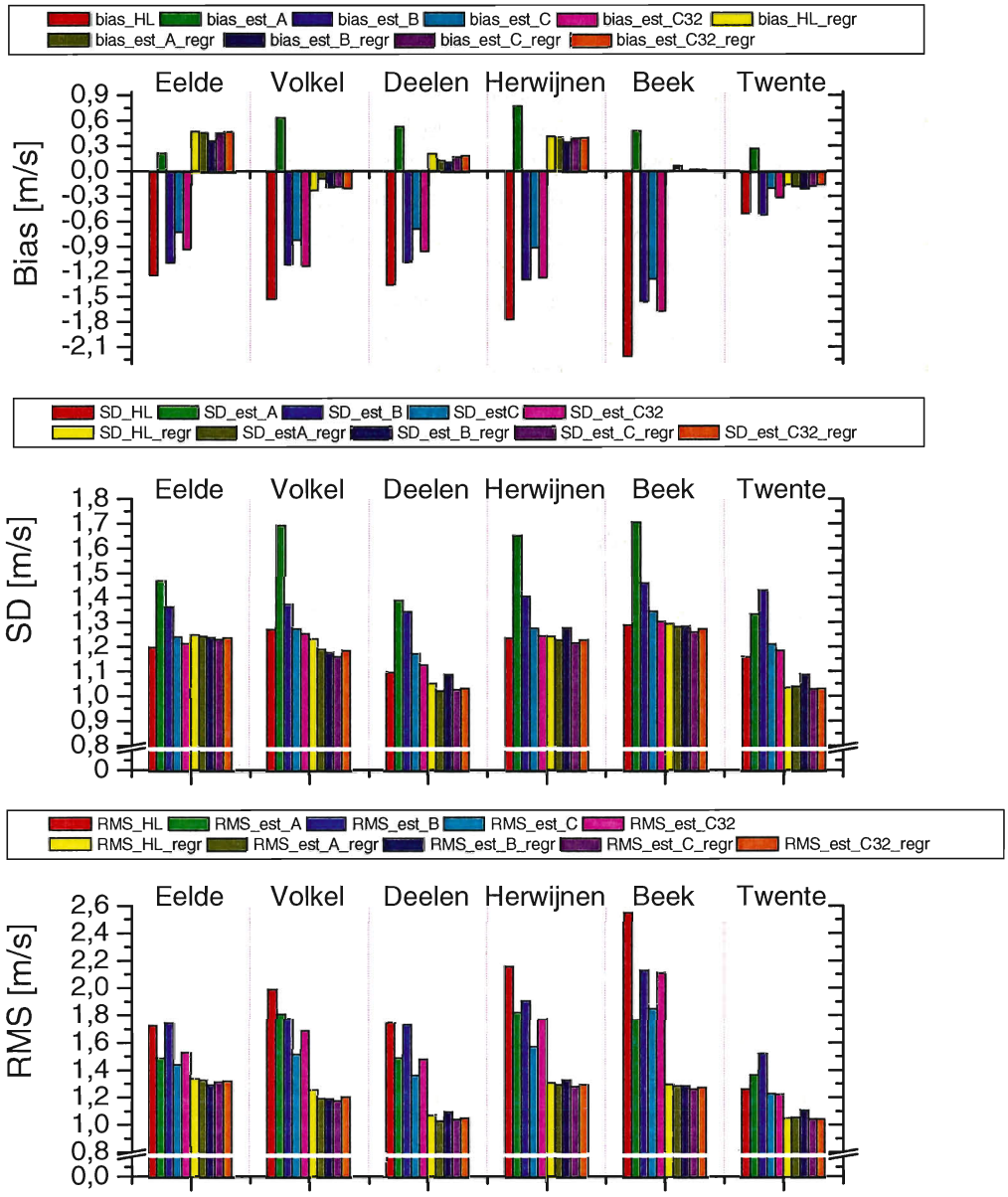

Figure 6.3 Wind speed results [m/s] per land station for Hirlam output (HL) and different estimators (_est_) with (_regr) and without regression for hours with observed 10m wind speeds > 4 m/s. The results are based on 6+6h forecasts valid for 12 UTC.

For further investigation purposes and because for some applications high wind speeds are important, the experiments are repeated with a threshold of 4 m/s in the observed wind speed (Fig. 6.3). High wind speeds correspond to near neutral situations. Therefore, we expect that the *NWP_error* (which, for Hirlam, mainly acts during stable very stable and unstable conditions) decreases leading to a reduction of the negative wind speed biases. However, there are mainly two effects of the wind speed threshold that act towards more negative biases. Firstly, the generally higher z_{0m} in the model will have a larger absolute effect on the wind speed (thus increasing the impact of the RP) leading to larger underestimations. Secondly, as explained in section 5.1, choosing a threshold in the observed wind speed will inevitably lead to an underestimation in the predicted wind speed. Besides all this, larger wind speeds will normally lead to larger errors (biases and scatter).

The results show that the latter effects, leading to larger underestimations, are stronger than the reduction of the NWP error. Consistent with the relatively large RP, the improvement with the estimators in comparison with standard u_{10mHL} is also large. This improvement can be ascribed totally to the bias reduction. On the other hand, the scatter is slightly larger with estimators but this is probably inherent to the substantial higher wind speeds using the proposed estimators. Anyhow, the RMSE for all stations (except Twente) decreases strongly. Remarkable are the relatively large positive biases in Eelde and Herwijnen after the regression, possibly caused by structural changes in the surroundings of these stations (in the most common wind directions for high wind speeds).

Land/sea locations

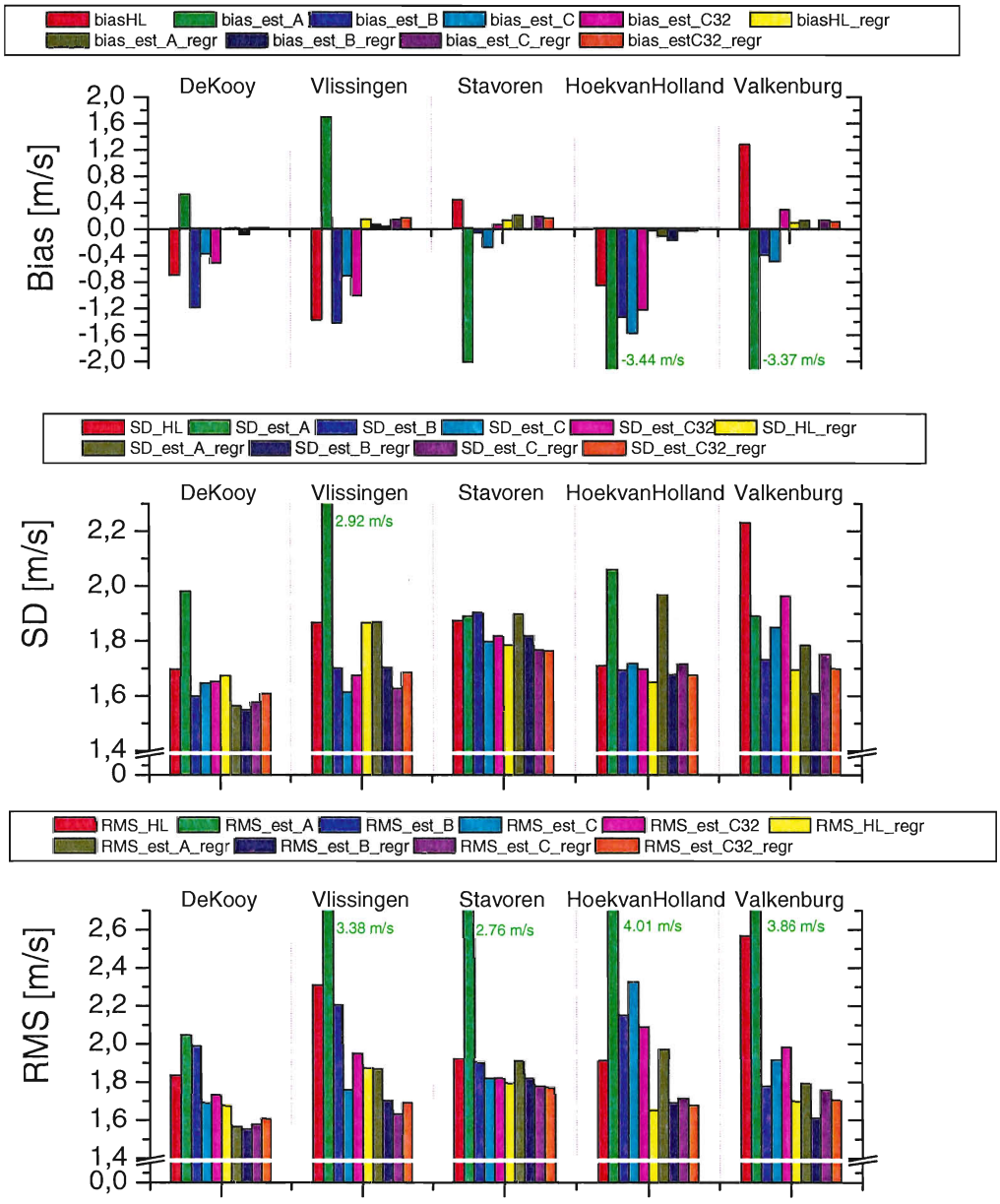

Figure 6.4 Wind speed results [m/s] per land/sea station for Hirlam output (HL) and different estimators (_est_) with (_regr) and without regression.. The results are based on 6+6h forecasts valid for 12 UTC. Some errors with estimator_A are too large to be plotted within the current vertical range. These errors are shown as green numbers next to the corresponding bar.

Starting with De Kooy and Vlissingen, which are in land Hirlam, we see that without regression again estimator_C gives the best results. Apparently, the land stability from Hirlam is on average still applicable for these locations, although for De Kooy, estimator_B (so without stability information) gives the smallest scatter. Obviously, estimator_C does not compensate for the large measuring height in Vlissingen (28 or 20m depending on the wind direction). For that, the estimators should be adapted as suggested in the previous section. After regression, there is not much difference in performance between the estimators for De Kooy and Vlissingen, but they are generally better than applying the regression on the Hirlam wind itself.

Stavoren, Hoek van Holland and Valkenburg are all sea locations in Hirlam. In Stavoren the most common wind directions have a long fetch over water, which results in the better performance of estimator_C (which for Stavoren uses the water stability of Hirlam). On the other hand, for Hoek van Holland and Valkenburg, the sea stability of Hirlam is not suitable for the actual situation and consequently estimator_B scores the best for the latter two stations.

After regression, the differences between different approaches are small. However, it cannot be concluded that this is also the case for the pool experiments as in the pool experiment the regression is applied on new locations (with different RPs).

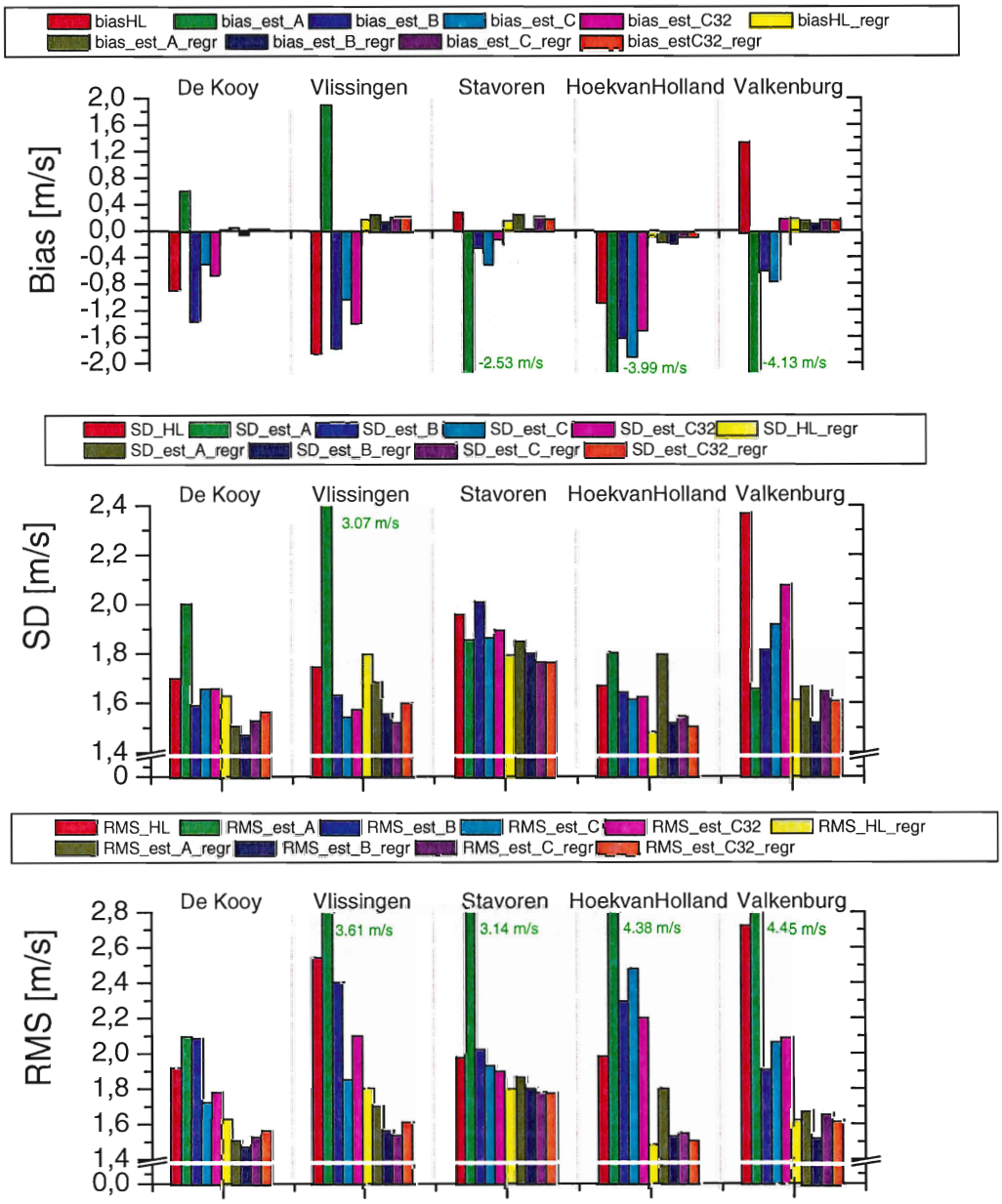

Figure 6.5 Wind speed results [m/s] per land/sea station for Hirlam output (HL) and different estimators (_est_) with (_regr) and without regression for hours with observed 10m wind speeds > 4 m/s. The results are based on 6+6h forecasts valid for 12 UTC. Some errors with estimator_A are too large to be plotted within the current vertical range. These errors are shown as green numbers next to the corresponding bar.

With a threshold in the observed wind speed of 4 m/s, locations Vlissingen and De Kooy benefit strongly from estimator_C. In contrast with land stations, now also the scatter reduces considerably, especially in Vlissingen, despite the much higher wind speeds with estimator_C. Also after regression, the results with estimator_C are very good. In Hoek van Holland, Stavoren and Valkenburg $z_{0mHL} < z_{0mloc}$ and consequently the wind speed reduces with the estimators. The overestimation in Stavoren decreases with increasing wind speed, which can partly be explained by the more common westerly wind directions for higher wind speeds. In this direction there is a long fetch over water (low z_{0mloc}) in Stavoren. Here, only the scatter and RMSE reduce with estimator_C. In Hoek van Holland the underestimation due to the measuring height obviously worsens with the estimators ($z_{0mHL}<z_{0mloc}$) , while in Valkenburg ($z_{0mHL}<<z_{0mloc}$) the overestimation is strongly reduced. The difference in bias

between estimator_B and C in Valkenburg is caused by the Hirlam sea-stability, which is normally too neutral for 12 UTC land conditions.

Verification time 0 UTC

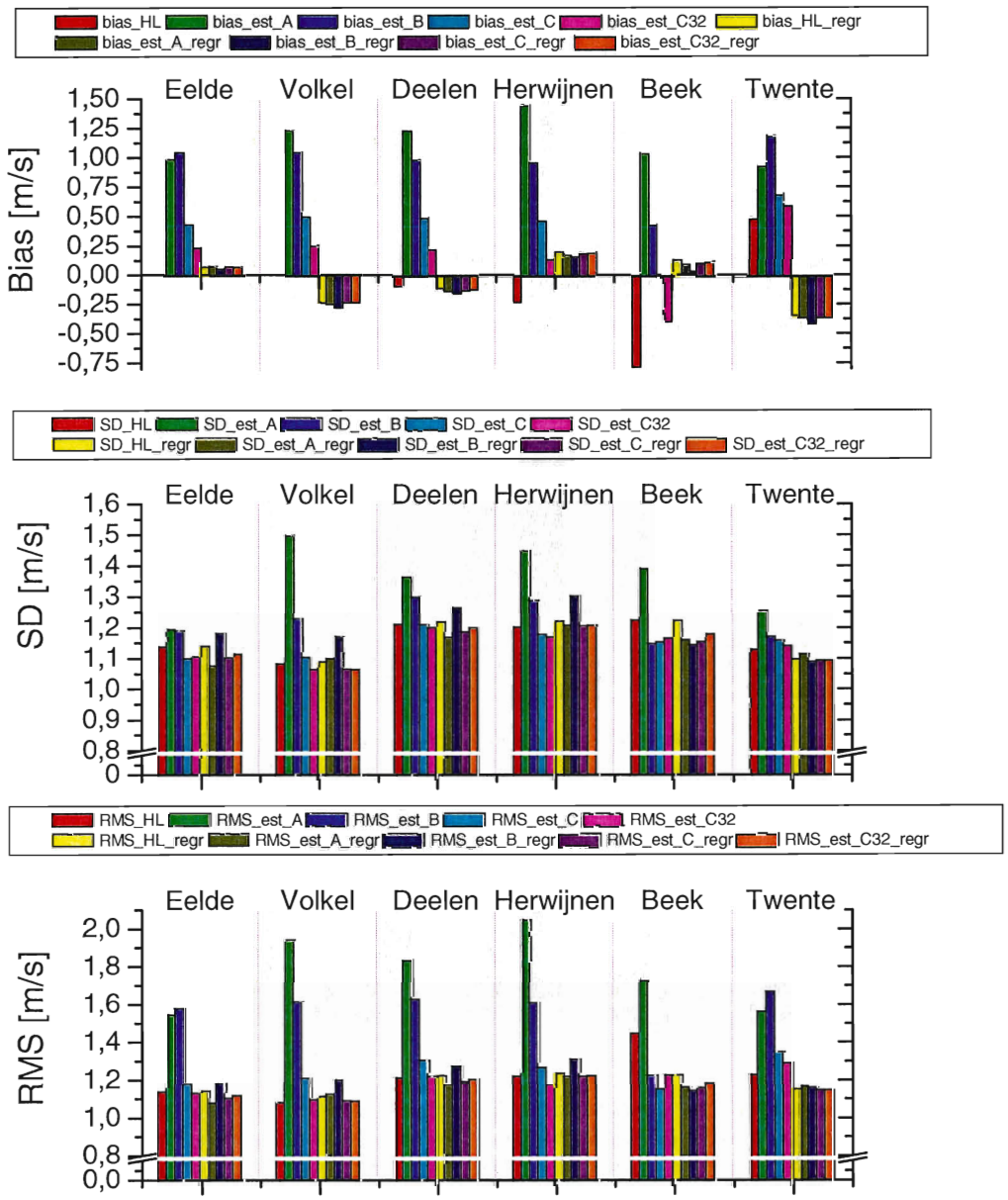

Figure 6.6 Wind speed results [m/s] per land station for Hirlam output (HL) and different estimators (_est_) with (_regr) and without regression. The results are based on 18+6h forecasts valid for 0 UTC

Land locations

For midnight hours the effect of the generally too high roughness length is counteracted by the excessive mixing of the turbulence scheme during stable conditions (sections 5.1 and 5.2). Therefore, applying the estimators without regression will eliminate one of the compensating errors in the Hirlam and consequently the bias increases. As illustrated in Fig. 6.6, only in Beek the z_{0m}-effect is so dominant that the bias improves with the estimators (except estimator_A). On the other hand, there is a small improvement in the scatter using estimator_C, possibly due to the wind direction dependent z_{0m} information.

As 0 UTC contains very stable conditions when wind speed at 140 and 10m can be decoupled (section 2.2), it is plausible that for these situations applying estimator_C32m gives better results than estimator_C (which uses u_{140mHL}). However, the difference with estimator_C is small enough to use just one estimator for all situations.

The variability in the u_{10mHL} bias for Eelde, Deelen, Herwijnen and Volkel (ranging from -0.22 to +0.01 m/s) is larger than the variability with estimator_C (+0.47 ± 0.03 m/s), just as we saw for the 12 UTC results. Again this can be interpreted as a confirmation of our hypothesis of horizontally homogeneous *NWP_errors,* a variable RP, and the reduction of the RP using the estimators. So +0.47 m/s might be associated with the average impact of the excessive mixing of the turbulence scheme at 0 UTC.

Again estimator_A and B clearly perform worse than the rest and the large overestimation with estimator_B can be attributed to the neutral approximation. After regression, differences between all the approaches are small (section 6.2).

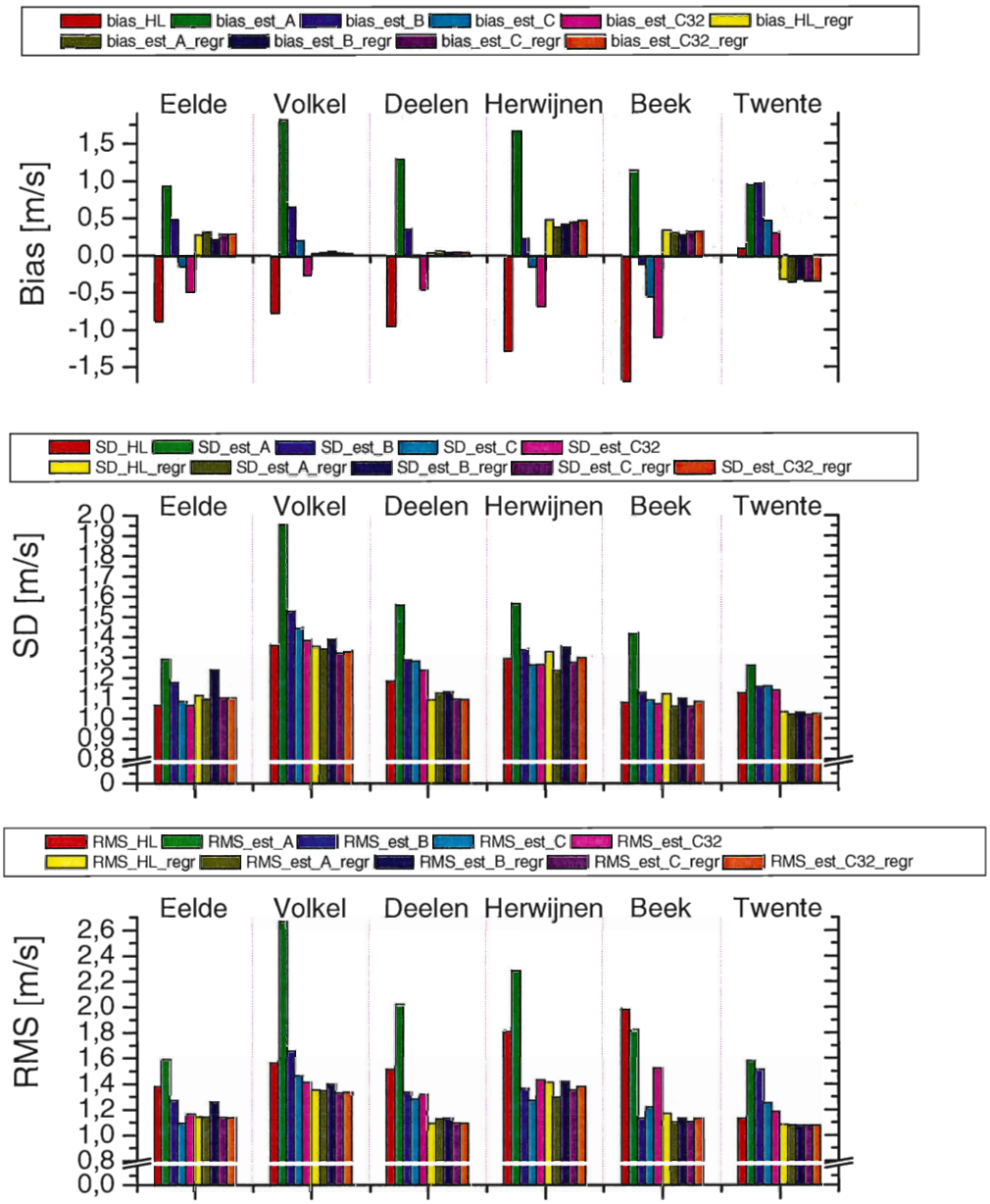

Figure 6.7 Wind speed results [m/s] per land station for Hirlam output (HL) and different estimators (_est_) with (_regr) and without regression for hours with observed 10m wind speeds > 4 m/s. The results are based on 18+6h forecasts valid for 0 UTC

After the results of the experiments with high(er) wind speeds at 12 UTC, the outcome for 0 UTC could be predicted. For high wind speeds, the difference in wind speed verification results between 12 and 0 UTC diminishes because of the approximately neutral stability. With a less strong NWP error (leading to lower wind speeds for 0 UTC) and a stronger RP, the Hirlam now lacks the compensating errors and consequently wind speeds and biases are reduced with estimators. The improvements increase with increasing wind speed. These results imply that the overall results for 0 UTC (with only minor advantage for the estimators) are deteriorated by the results for low wind speeds. Possibly, the use of estimator_C32 for low wind speed situations can improve the results for 0 UTC.

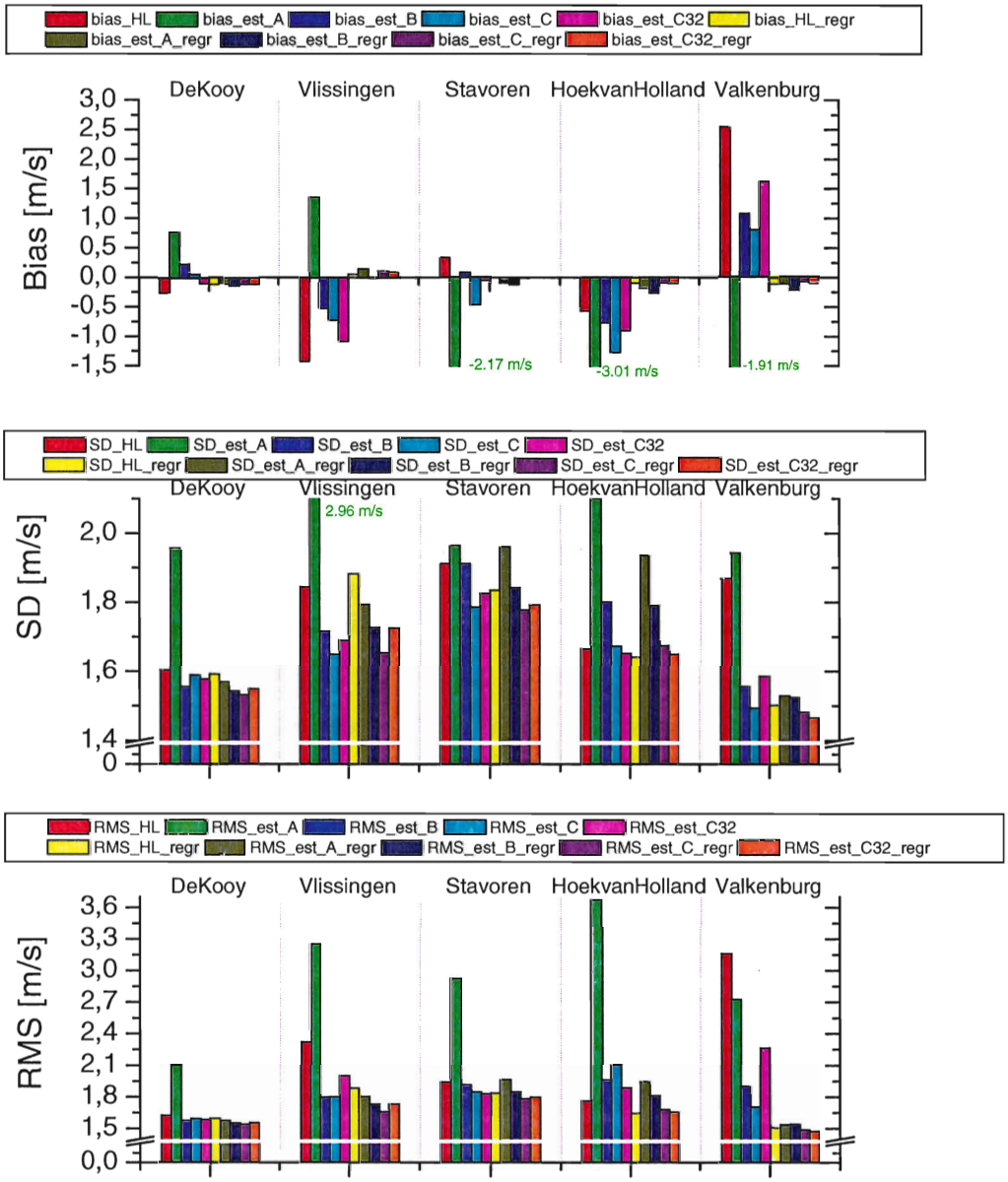

Figure 6.8 Wind speed results [m/s] per land/sea station for Hirlam output (HL) and different estimators (_est_) with (_regr) and without regression. The results are based on 18+6h forecasts valid for 0 UTC. Some errors with estimator_A are too large to be plotted within the current vertical range. These errors are shown as green numbers next to the corresponding bar.

Land/sea locations 0 UTC

Just as for 12 UTC, estimator_B shows the lowest scatter for De Kooy. However, a more significant decrease in scatter is achieved with estimator_C in Vlissingen (note that Vlissingen has a clear difference in z_{0m} for northerly and southerly directions, section 3.3). The latter estimator nicely improves the Hirlam scores for Vlissingen, although the bias due to the measuring height remains. When the neutral part of estimator_C (i.e. estimator_B) is adapted to the actual measuring height, the bias is almost eliminated (not shown). For both Vlissingen and De Kooy, the bias is eliminated after the regression and using estimator_C as predictor instead of u_{10mHL} improves the regression results.

For Stavoren estimator_C again gives the best result, with and without regression. This is not the case for Hoek van Holland where biases (mainly due to the 15m measuring height) increase using estimators (z_{0m} in Hirlam is lower than in the estimators). It follows naturally that estimator B without the Hirlam stability correction has relatively the best bias from all the estimators. The Hirlam sea stability correction for Hoek van Holland is not stable enough (for most directions a land correction is more appropriate) leading to higher wind speeds (and thus compensating partly for the 15m measuring height effect. Estimator_C32m does not compensate for the higher z_{0mloc} as much as the standard estimator_C and as a result the bias increase is less for Hoek van Holland. Clearly, application of the proposed estimators for the Hoek van Holland situation is useless and the best regression is the one on u_{10mHL}. Possibly better estimations of the 10m wind speed in Hoek van Holland can be achieved when estimator_B is adapted for the 15m measuring height. Further, the Hirlam stability correction for sea should be added for wind directions with long fetch over water and land stability corrections (of the nearest land grid point) for other directions.

Fortunately, in Valkenburg the measuring height is the standard 10m. Despite the fact that the stability of the sea should not be applicable here, estimator_C gives the best results. Here, the major part of the error is caused by the bias and estimator_C reduces the bias the most. Nevertheless, estimator_C also has the smallest scatter. Improvements are extremely large for Valkenburg but after regression again differences are small between the estimators and u_{10mHL}.

Results of the experiments with high wind speeds for land/sea locations are qualitatively the same as described for 12 UTC (not shown).

6.4 Results of the pool experiments

In this section the results of the pool experiment are described for 12 UTC (6+6 forecasts) and 0 UTC (18+6 forecasts). The regressions on the estimators (see section 6.2) are derived from a pool of 5 land stations and subsequently applied on the sixth land station (see also section 6.2). So the regressions are applied on an independent station and therefore the regression and the application could be made on the total data set. Nevertheless, we have chosen to apply the regressions on a totally independent data set to eliminate the possibility that climatological/meteorological influences disturb the results. More precisely, the regression equations are derived on a pool of 5 land stations on data ranging until 1-4-1998 and applied on the 6^{th} land station on data after 1-4-1998. The disadvantage of this approach is that possible changes in the model or observations deteriorate the results of the regression method unjustly. However, probably this has not played an important role since the biases appeared to be generally small.

Not presented in this section are the results where the neutral part (=estimator_B) and the stability correction part of estimator_C are used as two separate predictors for the regression. The results with the separate predictors do not differ significantly with the results for estimator_C. We also examined the influence of excluding Twente (because of invalid z_{0mloc} values) or both Twente and Beek (because Beek is an outlier in the z_{0m} spectrum) in the pool experiment. The process of regression on pools of observations turned out to be very robust because the results were almost exactly the same with the afore-mentioned exclusions.

When the pool experiment also includes land/sea stations it is important to use separate pools for land circumstances (i.e. land stations or land/sea stations for wind directions from land) and sea circumstances (measurements above sea or land/sea stations with long wind fetches over water). If a sea stability correction should be applied but the

nearest grid point is a land point, the stability correction in the estimator can be simply replaced by the stability correction from the nearest sea point.

Results of the pool experiment for 0 UTC

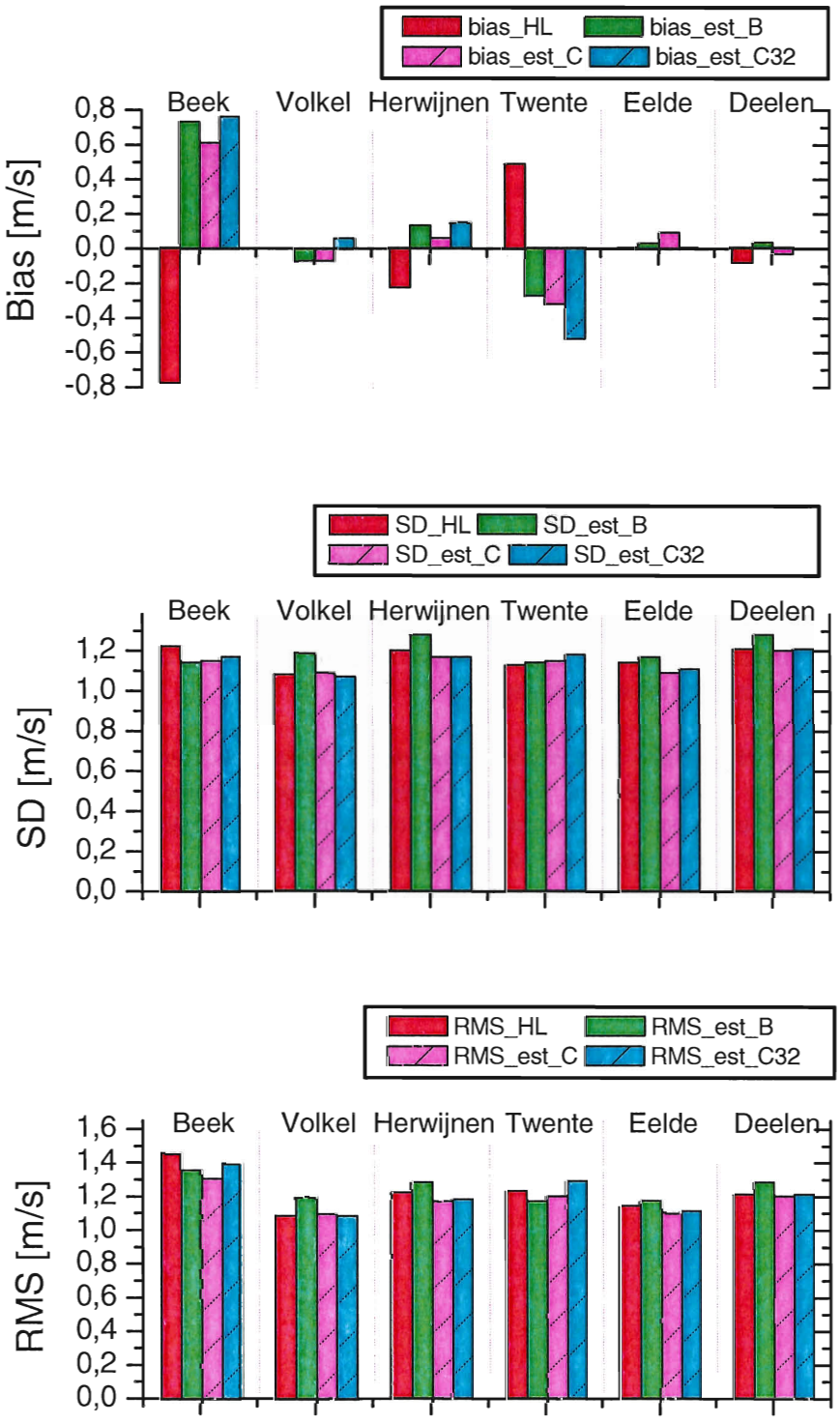

Figure 6.9 Results of the pool experiment for 0 UTC (18+6). Red bars refer to direct model output (_HL), green bars to regression on 5 land stations applied on the sixth using estimator_B (_est_B). Analogous magenta and blue, hatched, bars refer to estimator_C (est_C) and estimator_C32 (est_C32).

Figure 6.9 shows that estimator_B and, to a lesser extent, estimator_C32m, perform worse than estimator_C. However, also with estimator_C improvements in comparison with standard Hirlam wind is small, except for Beek where the RMSE skill is +10%. Because in Beek z_{0mHL}-z_{0mloc} is relatively large whereas in Twente it is small, these stations show relatively large biases using the regression. For both stations, the regression has to be extrapolated to z_{0mHL}-z_{0mloc} values that are not supported in the pool with observations. This confirms the importance that the pool must cover the complete z_{0mHL}-z_{0mloc} spectrum.

From the results in the previous section, we know that the results with the estimators are worsened by low wind speed situations. Especially at midnight hours compensating errors in the model (see section 5.1 and 6.3) limit the possibilities to improve the model. Better results with the pool regression technique for night time hours can be expected when negative wind speeds are set to 0 (note that after the regression estimated wind speeds can be negative), and when regressions for low wind speeds are made separately with another estimator than estimator_C (e.g. estimator_C32m, section 6.2 and 6.3).

Results of the pool experiment for 12 UTC

In contrast with the results for 0 UTC, the improvements by applying the regression for 12 UTC are substantial. Again the best results are achieved with estimator_C. Except for Twente (-15% RMS skill) all locations show a RMS skill of about 20% with on top of the list again Beek with +30% skill. The improvements with the regression arise mainly from bias reduction and for a smaller part from a reduction in the SD (despite the generally higher wind speeds!). For most stations the bias is practically eliminated with the regression from the pool experiment. For the same reasons as mentioned before (see the results for 0 UTC), Beek and Twente show relatively large remaining biases.

If we compare the results of the pool experiment (Figures 6.9 and 6.10) with the results of the estimators per station but without regression (Figs. 6.2 and 6.6), it becomes clear that applying the "universal" regression on locations with in principle no observations, effectively reduces (or even eliminates) the remaining bias (*NWP_error*)

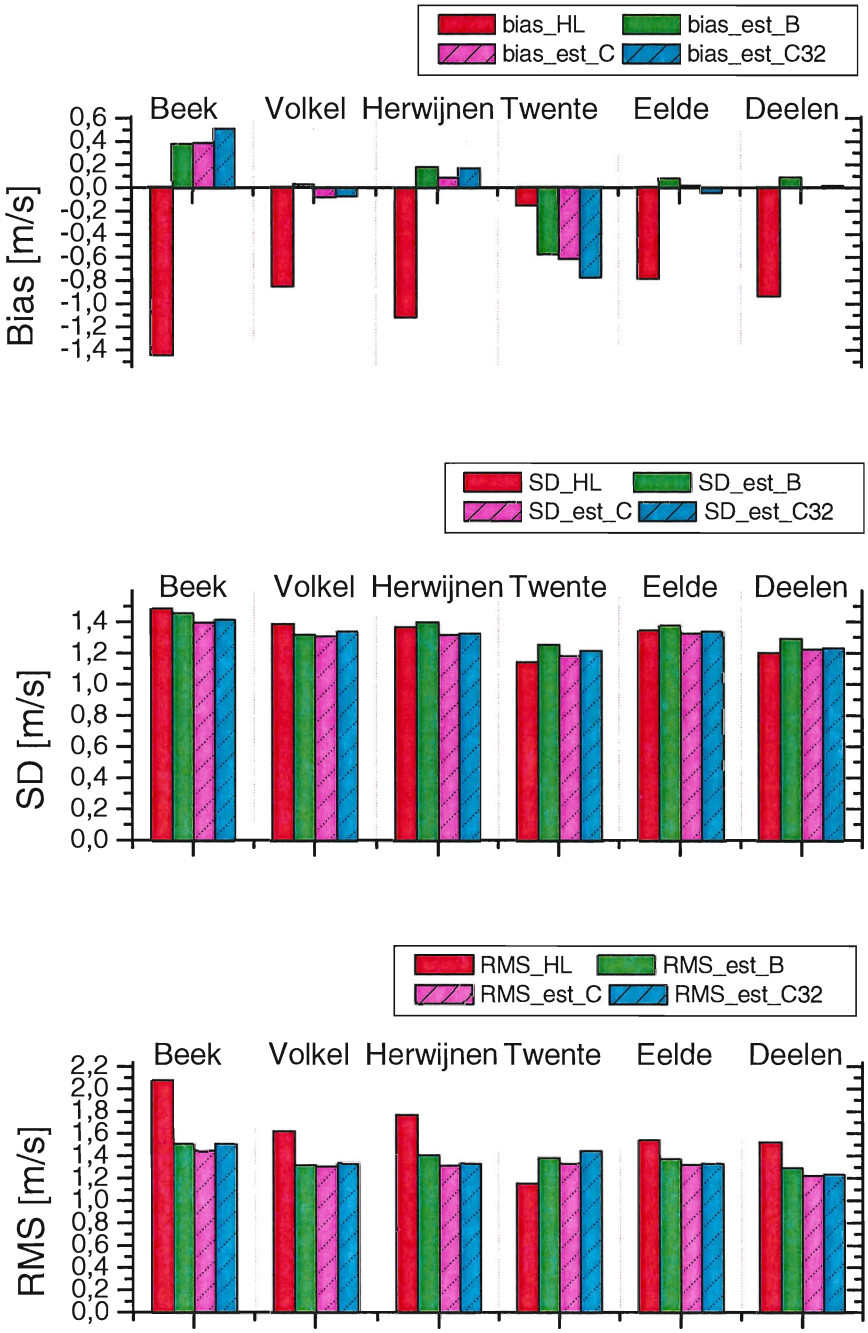

Figure 6.10 Results of the pool experiment for 12 UTC (6+6). Red bars refer to direct model output (_HL), green bars to regression on 5 land stations applied on the sixth using estimator_B (_est_B). Analogous magenta and blue, hatched, bars refer to estimator_C (est_C) and estimator_C32 (est_C32).

7 Discussion and conclusions

The main goal of this paper is to investigate ways to optimise NWP model output for a specific location. For that, an error analysis is made on a three-year dataset of the NWP model Hirlam. To structure and interpret the analysis of the model results, some error definitions are made. Important is the division of the total error in a *Representation_Problem*, which can fluctuate considerably on small horizontal scales, and a *NWP_error*, which is horizontally rather homogeneous. In verification studies and by operational forecasters the Representation problem (RP) is often ignored and the model is blamed for the total error. By tuning the roughness length of a grid box in order to get the right 10m-wind speed for a specific station, a *NWP_error* may unintentionally be introduced. The model should produce a grid box mean and not a site-specific output.

Especially in comparison with wind speed, temperature results are difficult to interpret because T_{2m} is influenced by so many parameters, like cloud cover, soil moisture, mixing of heat by the turbulence scheme, etc.. Remarkable are the diurnal cycles in the T_{2m} bias for all examined land stations. Relatively low temperatures occur at noon whereas relatively high temperatures are found for midnight and 6 UTC. For stations at or near the coast, the impact of the sea can be easily explained in terms of other stability regimes when the fetch over sea is long enough. The characteristics of the temperature results seem to adapt rapidly towards land characteristics, moving from the coast inland. Further, erroneous water temperatures in Hirlam influence the temperature results also. Applying a physical method which transforms the Hirlam 32m temperature down to 2m, has a small but beneficial effect on the T_{2m} scatter for daytime (unstable) hours. When the actual temperature observation behaves primarily as a land point and the model grid point is a sea point, the improvement with the transformation method is substantial and occurs for all hours of the day. During daytime hours, the output of the physical downward transformation method contains independent information in comparison with standard Hirlam T_{2m} output and therefore the statistical regression for temperature benefits when the transformation output is added as an extra predictor. The improvement in the scatter with the physical method probably arises from a better calculation of the surface energy balance. A simple linear regression on T_{2m} from Hirlam only, eliminates the temperature bias but has no impact on the scatter.

Despite attempts to limit the transformation error during stable conditions, no improvements are found for these hours. Physical arguments are given why the transformation method with the flux routines does not work under stable conditions. However, these problems might be circumvented when the near surface value is first transformed upwards with the model surface parameters and subsequently downward, this time with the locally valid surface parameters. The (large) error made during the upward transformation will be compensated during the downward transformation.

Wind speed has totally different error characteristics than temperature. High wind speed errors occur when wind speeds are high and stability corrections are small. However, under these conditions vertical temperature gradients and temperature errors are usually small.

For wind speed, the division of the total error in a representation problem and a *NWP_error* is valuable. Analogous to the temperature results, also for wind speed all the land stations show a diurnal cycle in the bias. However, for wind speed the diurnal cycles are quite uniform for all land stations. It is most likely that this diurnal cycle can be associated with deficiencies in the turbulence scheme. Clearly a deficiency in the turbulence scheme is a *NWP_error* and, consistent with our hypothesis, the diurnal cycle of the wind speed bias is horizontally homogeneous. During stable conditions the turbulence scheme mixes momentum

too much, leading to relatively high wind speeds, whereas during unstable conditions the mixing is too weak with relatively low wind speeds at 10m as a result. Especially too much mixing during stable conditions is quite common for turbulence schemes. It prevents the model from total disappearance of turbulence, resulting in decoupled layers. However, as a consequence, the wind profiles are sub optimal. In contrast with the diurnal cycle, the wind speed bias strongly fluctuates from station to station and can be related to the average z_{0m} difference between the locally valid z_{0mloc} and the Hirlam z_{0mHL}. The difference between the local z_{0m} and the Hirlam z_{0m} is clearly a representation problem. As expected, this representation problem is far from constant in the horizontal.

The physical downward transformation using the flux routines improves the 10m-wind speed bias (=RP) for land stations but does not, and should not, completely compensate its diurnal cycle (=*NWP_error*). However, the output of this downward transformation does not seem to contain independent information, as a simple linear regression per station on just u_{10mHL} performs about as good as a regression on both u_{10m} from the physical method and u_{10mHL}.

Another approach for the downscaling of wind speed is a combined physical statistical approach. To make estimates of 10m-wind speed for locations where no observations are available, it is necessary to deal with the horizontally fluctuating representation problem (here due to z_{0mloc}- z_{0mHL}) before the statistical regression can handle the *NWP_error*. For that, different estimators for u_{10mloc} are developed and used as a predictor in the regression process. The estimator with the best performance is based on surface layer theory and assumes a homogeneous wind field at 140m (blending height). The blending height can be easily adapted to other heights on which NWP output is available. For low wind speeds, stable conditions, it is probably beneficial to assume a smaller blending height. Further, the estimator assumes that the stability correction from 140 to 10m is the same for the local situation and in the model. The estimator can also be adapted to deviating measuring heights. After the compensation for the RP with the estimator, this estimator can be used in one pool for all the land stations together to perform the linear regression for the compensation of the model error. The resulting "universal" regression formulae can be used at any land location under the condition that z_{0mloc} is known. In practice, a wind direction dependent z_{0mloc} can be derived from high-resolution land use maps.

Before the estimators were tested in a pool experiment and applied on a new location, we first investigated their performance per station separately. Also without regression, the best estimator already improves the standard Hirlam results for 12 UTC significantly. For 0 UTC, the estimator eliminated a compensating error (namely the RP) and therefore the bias increased. For both verification hours the difference in wind speed bias between most land stations becomes very small with the estimator. Because the *NWP_errors* are horizontally rather homogeneous (and the RP is not), this can be interpreted as a confirmation that the RP is reduced with the estimators.

In this paper we validated the above-mentioned pool experiment by deriving the regression on sets of five land stations and applying the regression to the sixth land station. The results for six hours forecasts valid for 0 and 12 UTC have been presented. The results are very promising, especially for 12 UTC (RMS skills in comparison with the standard model output of about 20%) when the *NWP_error* and the RP both lead to underestimations of the wind speed. For 0 UTC the results are modest due to compensating errors in the standard model output (except for Beek due to the dominating RP) but the method can easily be improved by applying a different estimator for low wind speed, stable conditions. If the results of the pool experiment are compared with the results of the estimators per station but without regression, it becomes clear that applying the "universal" regression on locations

with in principle no observations, effectively reduces (or even eliminates) the remaining bias (*NWP_error*). Further, the results confirm the importance that the pool must cover the complete z_{0mHL}-z_{0mloc} spectrum. Recommendations have been given when coastal stations are included in the pool. Further the combined method does not flatten high wind speeds for longer forecast periods (due to a Perfect Prog statistical technique) or due to horizontal interpolation (not necessary).

The results in this paper confirm that statistical methods suffice when wind speed has to be forecast at locations where observations are available. Linear regression can tackle all kinds of model deficiencies (e.g. flow dependent) in combination with local peculiarities. However, for places outside observation sites physical downscaling seems necessary. We believe that downscaling wind speed should be a combination of physical and statistical techniques. Not only because in this way both the representation problem and the model errors can be handled, but also because this approach offers a lot of opportunities to optimise the system for a particular use or customer. For example, one can choose to use MOS techniques for the smallest possible errors for longer forecast times, or to use Perfect Prog techniques and regressions derived for high wind speeds only, to produce a storm warning system. As we did not pursue to make an operational forecasting system, we did not examine regression on more than a single predictor. For wind speed as well as temperature we expect that large improvements can be achieved when more predictors are included, e.g. the sine or cosine of the wind direction.

The increasingly popular use of a tile scheme in the surface parameterization of a NWP model is another way to make model output more representative for a certain location. This study gives no indications that this concept does not work for temperature. For example Deelen and Herwijnen, with the same nearest grid point and thus *NWP_error*, have about the same temperature error except from a small difference in bias because of the different surface elevation. It is plausible that the RP is mostly determined by the surface cover (here short grass) and hence the tile concept might be a suitable way to compensate for the temperature representation problem. An exception is the possibly significant influence of the upstream soil moisture on the local temperature (Brandsma et. al., 2002), which is a horizontally strongly fluctuating RP for temperature. Analogous to the pool experiment for u_{10m}, a pool experiment can be interesting for T_{2m}, this time with the local upstream soil moisture instead of a local roughness length. For wind speed the situation is completely different. Now the roughness is the dominant factor determining the RP and with the small-scale heterogeneity of the landscape (as in the Netherlands), this cannot be described satisfactorily in the tile concept unless a site-specific tile (with area zero) is specially constructed for a certain location (e.g. to simulate a synop station).

Acknowledgements

Henk Benschop is thanked for sharing his extensive knowledge of synop stations and their surroundings. We are grateful to Fred Bosveld for many helpful discussions on especially the 10m-wind speed estimators. Sander Tijm is thanked for his help concerning characteristics of the Hirlam and Geert Lenderink for sharing his knowledge of turbulence schemes. We are also indebted to Job Verkaik for providing the roughness lengths for the synop stations and discussions about the validity of roughness lengths. Maurice Schmeits is thanked for his help with the production of Fig. 3.1 in GIS. Finally, comments on a draft of this report by Ben Wichers Schreur and Dick Blaauboer are gratefully acknowledged.

References

André, J.C., and L. Mahrt, 1982: The nocturnal surface inversion and influence of clear-air radiative cooling. *J. Atmos. Sci.*, **39**, 864-878

Beljaars, A.C.M., 1987: The measurements of gustiness at routine wind stations-A review. WMO/IOM Rep. 31, Geneva, Switzerland, 50pp.

Beljaars, A.C.M., and F.C. Bosveld, 1997: Cabauw data for the validation of land surface parameterization schemes. *J. Climate*, **10**, 1172-1193

Beljaars, A.C.M., and A.A.M. Holtslag, 1991: Flux parameterization over land surfaces for atmospheric models. *J. Appl. Meteor.*, **30**, 327-341

Best, M.J., F.J. Bornemann, B.V. Chalcraft, and C.A. Wilson, 2000: Mesoscale Model Upgrade - Introduction of the land surface tile scheme (MOSES 2). Forecasting Research Technical Report No. 341, Met Office, London Road, Bracknell, Berkshire, RG12 2SZ, United Kingdom

Brandsma, T., G.P. Können, and H.R.A. Wessels, 2002: Empirical estimation of the effect of urban heat advection on the temperature series of De Bilt (The Netherlands), Submitted to *International J. of Climatology*

Charnock, H., 1955: Wind stress on a water surface. *Quart. J. Roy. Meteor. Soc.*, **81**, 639-640

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203-1211

Holtslag, A.A.M., 1984: Estimates of diabatic wind speed profiles from near-surface weather observations. *Boundary-Layer Meteor.*, **29**, 225-250

Holtslag, A.A.M., and A.P. van Ulden, 1983: A simple scheme for daytime estimates of the surface fluxes from routine weather data. *J. Climate Appl. Meteor.*, **22**, 517-529

Hutjes, R.W.A., 1996: Transformation of near-surface meteorology in a small-scale landscape with forests and arable land. Thesis University of Groningen, The Netherlands.

Källén, E. (editor), 1996: Hirlam documentation manual System 2.5

Koster, R.D., and M.J. Suarez, 1992: Modelling the land surface boundary in climate models as a composite of independent vegetation stands. *J. Geophys. Res.,* (D) 97, 2697-2715

Kuo, H.L., 1974: Further studies of the influence of cumulus convection on large-scale flow. *J. Atmos. Sci.*, **31**, 1232-1240

Louis, J.F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Boundary-Layer Meteorol.,* vol. **17**, 187-202

McNaughton, K.G., and P.G. Jarvis, 1984: Using the Penman-Monteith equation predictively. Agr. Water Management, **8**, 263-246

Monin, A.S., and A.M. Obukhov, 1954: Basic laws of turbulent mixing in the ground layer of the atmosphere. Trudy Geophys. Inst. Ak. Nauk SSSR, **24**, 163-187

Monteith, J.L., 1981: Evaporation and surface temperature. *Quart. J. Roy. Meteor. Soc.*, **107**, 1-27

Obukhov, A.M., 1946: Turbulence in an atmosphere with a non-uniform temperature. Tr. Akad. Nauk. SSSR Inst. Teorel. Geo-fis., 1 (translation in *Boundary.-Layer Meteor.*, 1971, 2, 7-29)

De Rooy, Wim C., 1995: Regionalisation of meteorological parameters. Scientific report, WR95-06 KNMI, De Bilt, The Netherlands

De Rooy, Wim C., and A.A.M. Holtslag, 1999: Estimation of surface radiation and energy flux densities from single-level weather data. *J. of Appl. Meteor.*, **38**, 526-540

De Rooy, Wim C., 1999: HIRLAM near-surface output. Error sources and possible improvements in T2m and V10m Forecasts. Hirlam Newsletter No.33, 101-109

Sundqvist, H., 1978: A parameterization scheme for non-convective condensation including prediction of cloud water content. *Quart. J. Roy. Meteor. Soc.*, **104**, 677-690

Tustison, Ben, Daniel Harris, and Efi Foufoula-Georgiou, 2001: Scale issues in verification of precipitation forecasts. *Journal of Geophys. Res.,* Vol. **106**, No. D11, 11,775-11,784

Van Ulden, A.P., and J. Wieringa, 1996: Atmospheric boundary layer research at Cabauw. *Boundary-Layer Meteorol.*, **78**, 39-69

Verkaik, J.W., and A. Smits, 2001: Interpretation and estimation of the local wind climate. Proceedings 3rd European & African Conference on Wind Engineering, Eindhoven University of Technology, The Netherlands, 43-56

Wieringa, J., 1976: An objective exposure correction method for average wind speeds measured at a sheltered location. *Quart. J. Roy. Meteor. Soc.*, **102**, 241-253

Wieringa, J., 1986: Roughness-dependent geographical interpolation of surface wind speed averages. *Quart. J. Roy. Meteor. Soc.,* **112**, 867-889

Appendix A

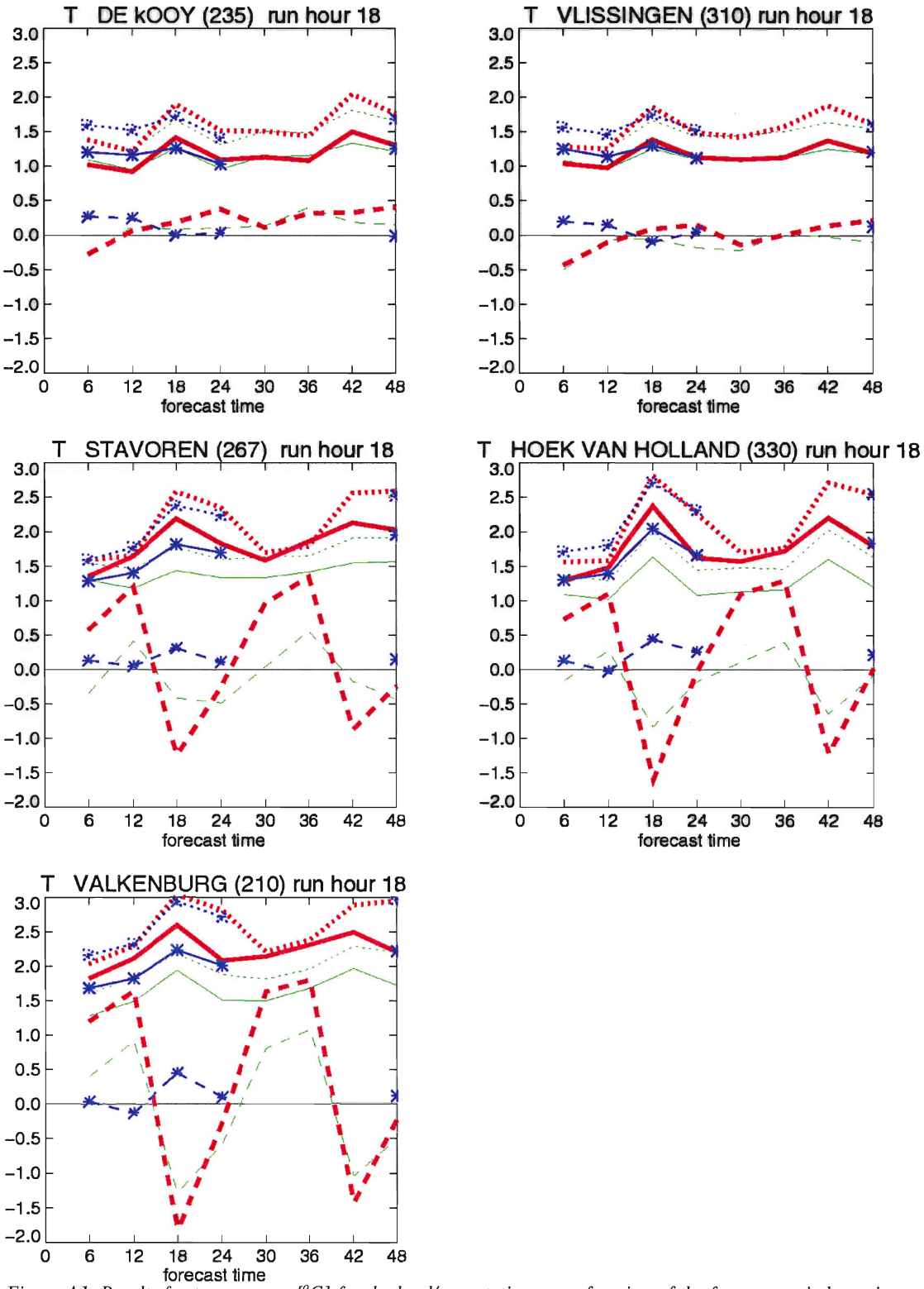

Figure A1. Results for temperature [°C] for the land/sea stations as a function of the forecast period starting at 18 UTC. Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam T_{2m} only. Regression data are missing for +30 until +42h forecasts.

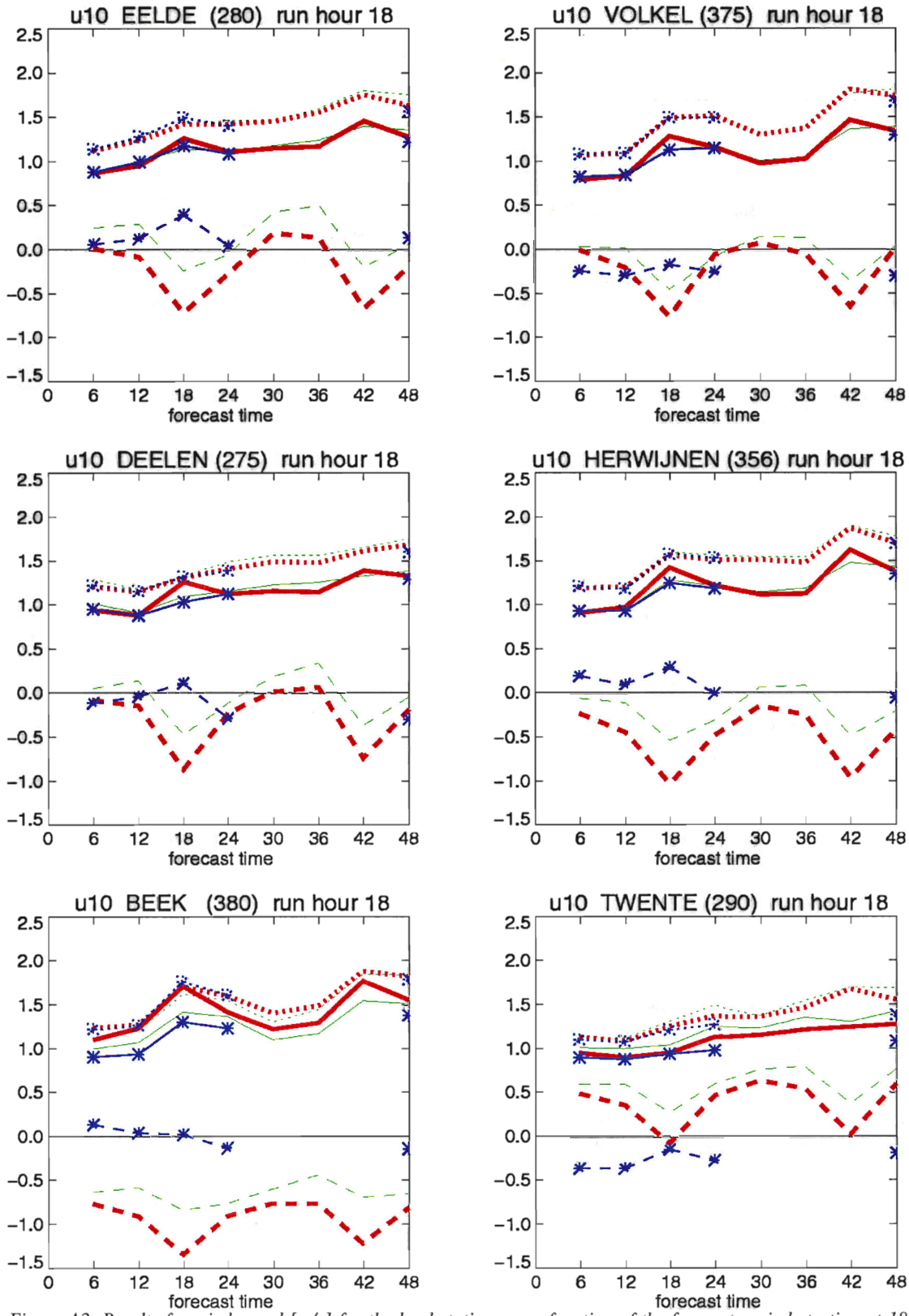

Figure A2. Results for wind speed [m/s] for the land stations as a function of the forecast period starting at 18 UTC. Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam u_{10m} only. Regression data are missing for +30 until +42h forecasts.

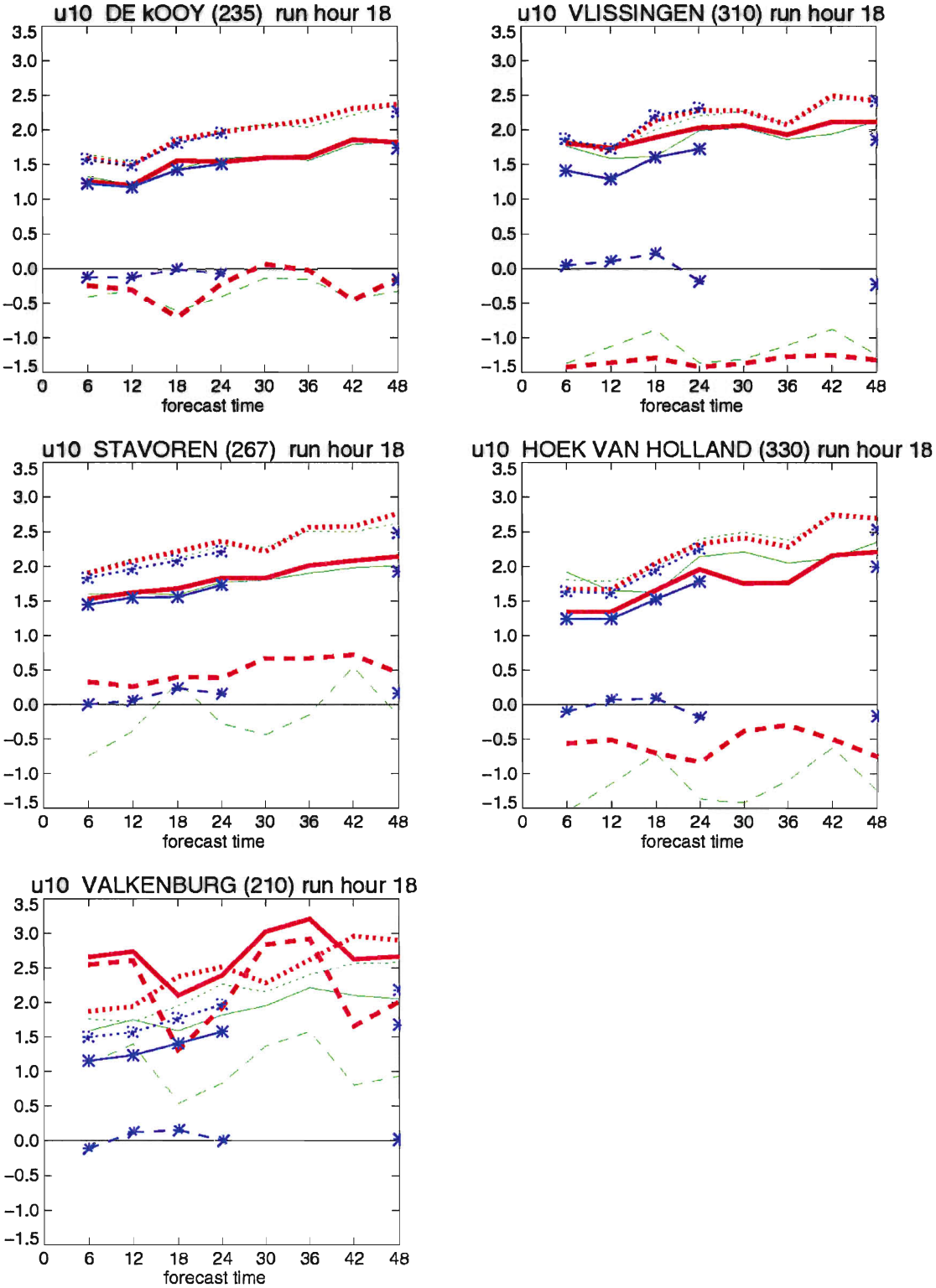

Figure A3. Results for wind speed [m/s] for the land/sea stations as a function of the forecast period starting at 18 UTC. Solid lines refer to mean absolute error (MAE), dashed to bias and dotted to standard deviation (SD). Red lines are based on NWP output, green lines on the physical method using the FRs and blue lines with asterisks on linear regression on Hirlam u_{10m} only. Regression data are missing for +30 until +42h forecasts.

Appendix B

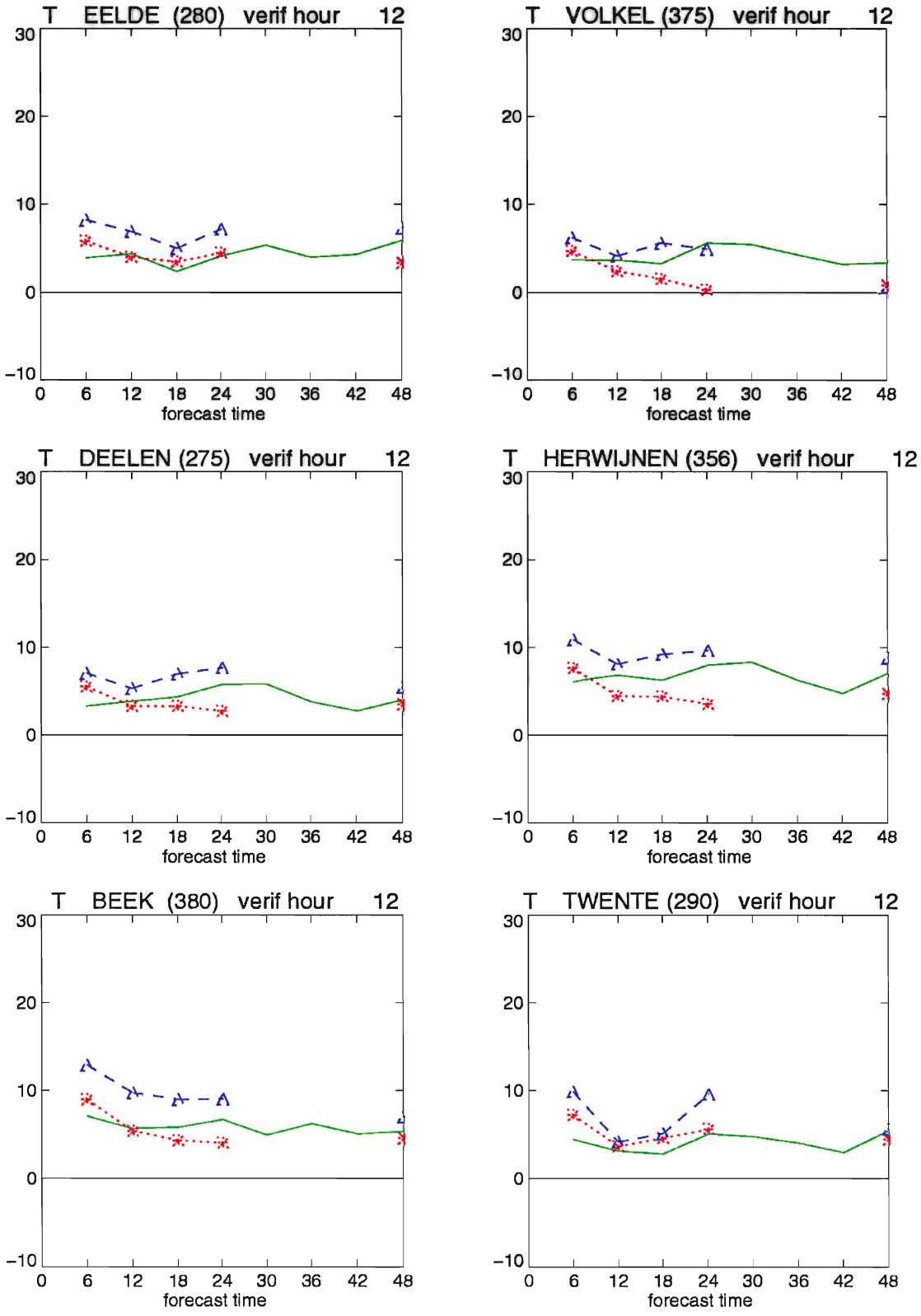

Figure B1 Results for temperature [°C] for the land stations valid for 12 UTC (different analysis times). Solid green lines refer to the skill of the physical method using the FRs, dotted red lines to the skill of linear regression on the T_{2m} of Hirlam only, and dashed blue lines are based on regression on Hirlam T_{2m} and the output of the physical method. Regression data are missing for +30 until +42h forecasts.

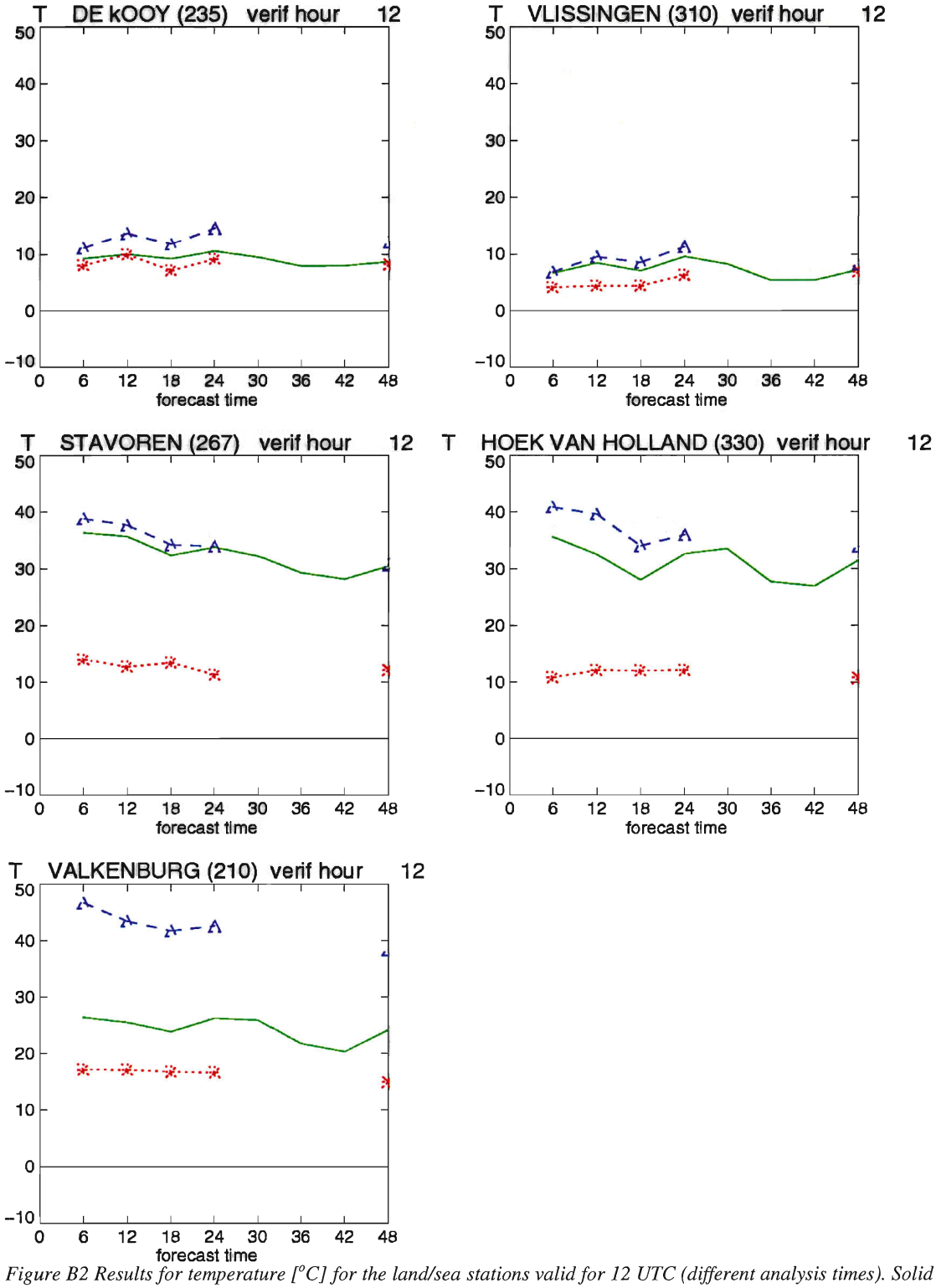

Figure B2 Results for temperature [°C] for the land/sea stations valid for 12 UTC (different analysis times). Solid green lines refer to the skill of the physical method using the FRs, dotted red lines to the skill of linear regression on the T_{2m} of Hirlam only, and dashed blue lines are based on regression on Hirlam T_{2m} and the output of the physical method. Regression data are missing for +30 until +42h forecasts.

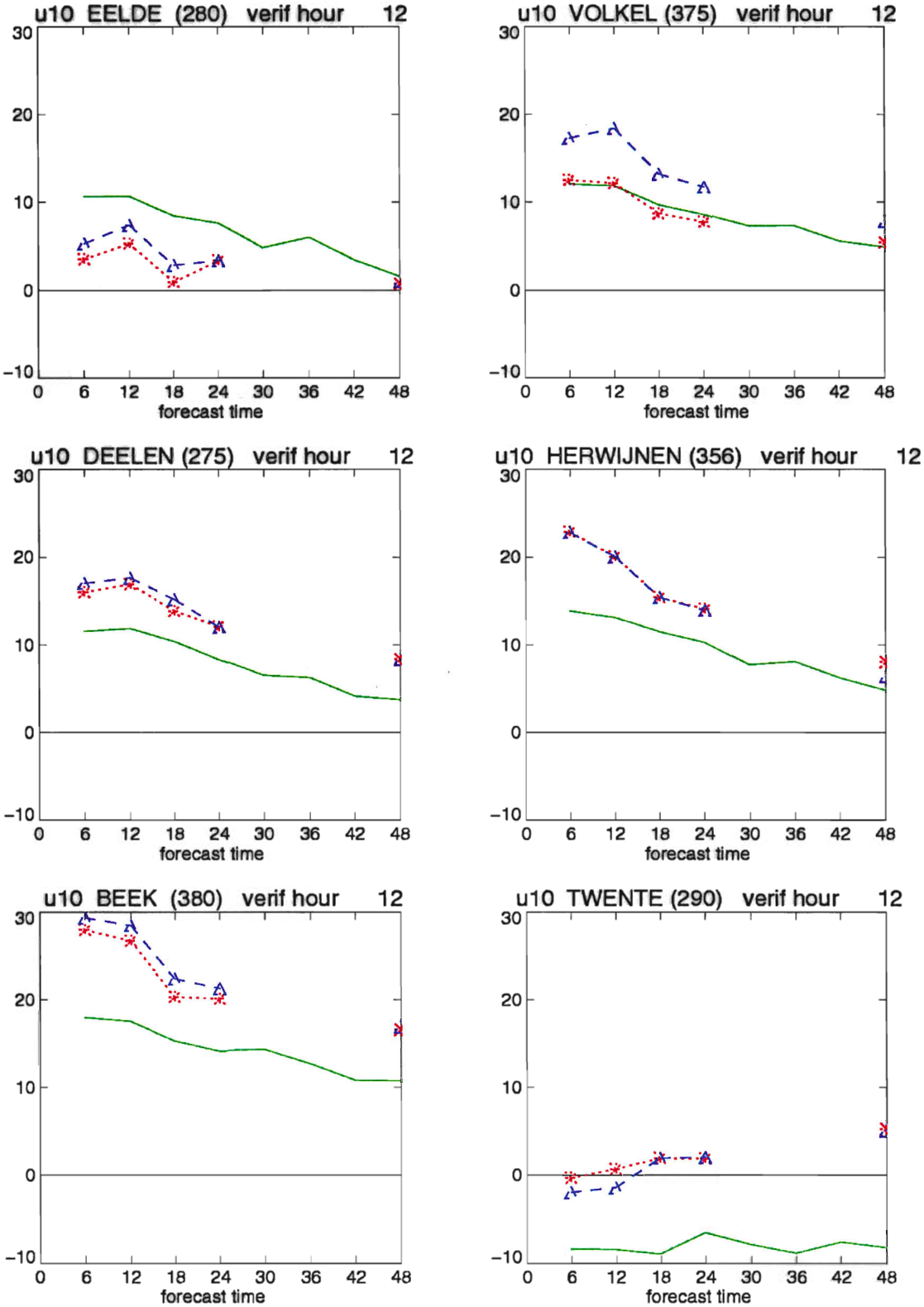

Figure B3 Results for wind speed [m/s] for the land stations valid for 12 UTC (different analysis times). Solid green lines refer to the skill of the physical method using the FRs, dotted red lines to the skill of linear regression on the u_{10m} of Hirlam only, and dashed blue lines are based on regression on Hirlam u_{10m} and the output of the physical method. Regression data are missing for +30 until +42h forecasts.

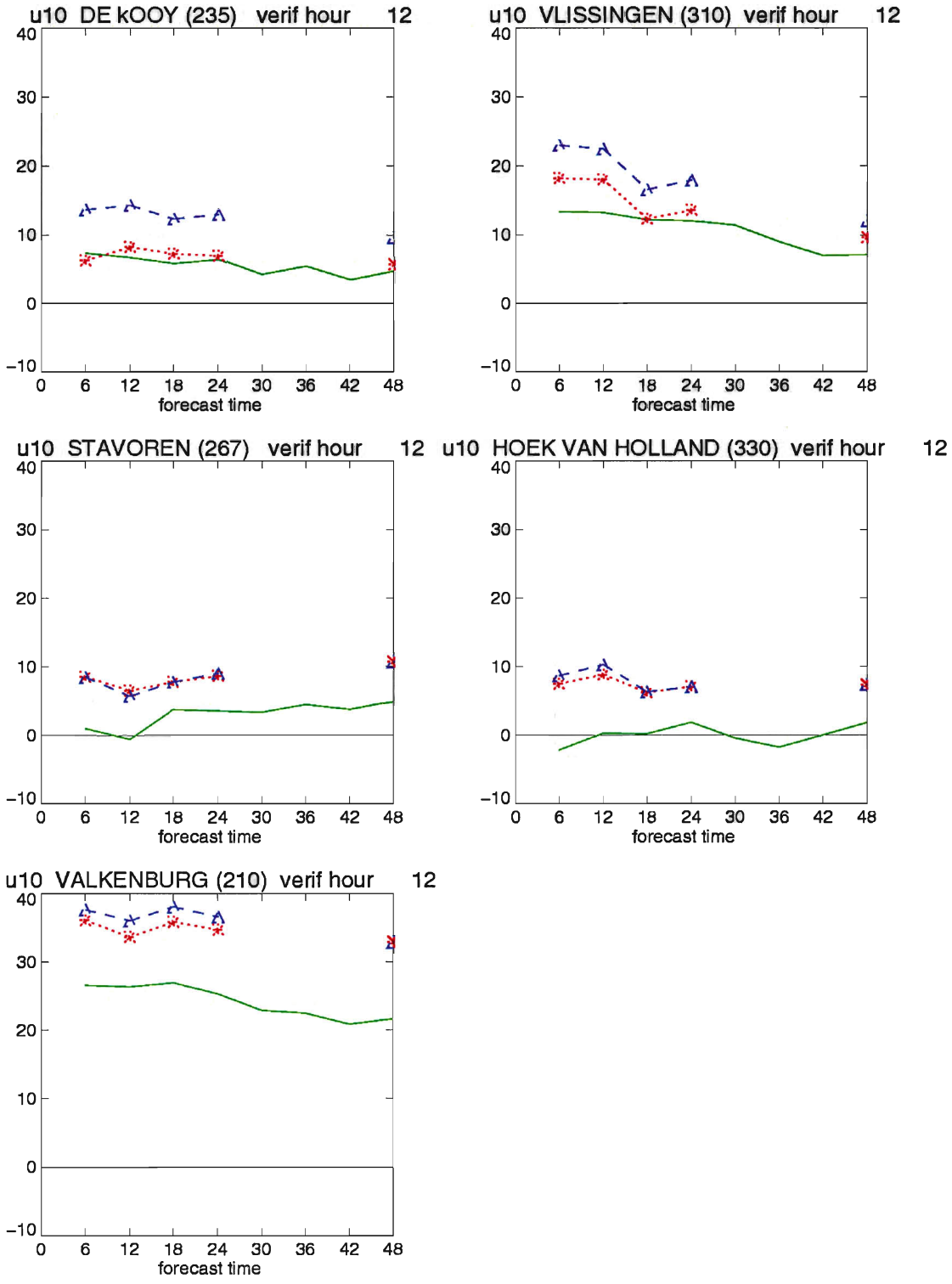

Figure B4 Results for wind speed [m/s] for the land/sea stations valid for 12 UTC (different analysis times). Solid green lines refer to the skill of the physical method using the FRs, dotted red lines to the skill of linear regression on the u_{10m} of Hirlam only, and dashed blue lines are based on regression on Hirlam u_{10m} and the output of the physical method. Regression data are missing for +30 until +42h forecasts.

OVERZICHT VAN RECENT VERSCHENEN KNMI-PUBLICATIES

KNMI-PUBLICATIE MET NUMMER

186-II Rainfall generator for the Rhine Basin: multi-site generation of weather variables by nearest-neighbour resampling / T. Brandsma a.o.

186-III Rainfall generator for the Rhine Basin: nearest-neighbour resampling of daily circulation indices and conditional generation of weather variables / Jules J. Beersma and T. Adri Buishand

186-IV Rainfall generator for the Rhine Basin: multi-site generation of weather variables for the entire drainage area / Rafal Wójcik, Jules J. Beersma and T. Adri Buishand

189 Aardbevingen in Noord-Nederland in 1998: met overzichten over de periode 1986-1998 / [Afdeling SO]

190 Seismisch netwerk Noord-Nederland / [afdeling Seismologie]

191 Het KNMI-programma HISKLIM (HIStorisch KLIMaat) / T. Brandsma, F. Koek, H. Wallbrink, G. Können

192 Gang van zaken 1940-48 rond de 20.000 zoekgeraakte scheepsjournalen / Hendrik Wallbrink en Frits Koek

193 Science requirements document for OMI-EOS / contr. by R. van der A .. [et al.] **(limited distribution)**

194-1 De zonsverduistering van 11 augustus 1999, deel 1: de waarnemingen van het gedrag van flora en fauna / J. Kuiper, m.m.v. Guus Kauffeld

195 An optimal infrasound array at Apatity (Russian Federation) / Läslo Evers and Hein Haak **(limited distribution)**

196-I Rainfall Generator for the Meuse Basin: simulation of 6-hourly rainfall and temperature for the Ourthe catchment / Rafal Wójcik and T. Adri Buishand

197 Meteorologie op zee: beknopte handleiding voor waarnemingen op zee [= manual meteorology at sea] **(limited distribution)**

TECHNISCH RAPPORT = TECHNICAL REPORT (TR)

228 Evaluation of modified soil parameterization in the ECMWF landsurface scheme / R.J.M. IJpelaar

229 Evaluation of humidity and temperature measurements of Vaisala's HMP243 plus PT100 with two reference psychrometers / E.M.J. Meijer

230 KNMI contribution to the European project WRINCLE: downscaling relationships for precipitation for several European sites / B.-R. Beckmann and T.A. Buishand

231 The Conveyor Belt in the OCCAM model: tracing water masses by a Lagrangian methodology / Trémeur Balbous and Sybren Drijfhout

232 Analysis of the Rijkoort-Weibull model / Ilja Smits

233 Vectorization of the ECBilt model / X. Wang and R.J. Haarsma

234 Evaluation of a plant physiological canopy conductance model in the ECMWF land surface scheme / J. van de Kassteele

235 Uncertainty in pyranometer and pyrheliometer measurements at KNMI in De Bilt / J.S. Henzing a.o.

236 Recalibration of GOME spectra for the purpose of ozone profile retrieval / Ronald van der A

237 Tracing water masses in the Atlantic / Yann Friocourt and Siebren Drijfhout

238 Klimaat voor Amsterdam Airport Schiphol / A. Smits

239 Seismische analyse van de aardbevingen bij Alkmaar op 9 en 10 september en Bergen aan Zee op 10 oktober 2001 / H.W. Haak, B. Dost, F.H. Goutbeek

240 EBEX-2000 : the KNMI/WAU contribution / W. Kohsiek, E.W. Meijer, P.J.B. Versteeg, O.K. Hartogensis, a.o.

241 Ontwikkeling gidsvergelijkingen voor meerdaagse neerslagkansen / D. Vogelezang en K. Kok

242 On photosynthesis parameters for the A-gs surface scheme for high vegetation / G.J. Steeneveld

243 Temperatuurvergelijkingen voor de Middellange Termijn Gids : ontwikkeling en verificatie over 2000 / J.Wijngaard

244 Verification of clear-air turbulence forecasts / A. Overeem

245 A comprehensive description of the KNMI seismological instrumentation / B. Dost and H. Haak

246 Verandering van neerslakarakteristieken in Nederland gedurende de periode 1901-2001 / A.T.H. Bruin

WETENSCHAPPELIJK RAPPORT = SCIENTIFIC REPORT (WR)

00-04 On the behaviour of a few popular verification scores in yes-no forecasting / C.J. Kok

01-01 Hail detection using single-polarization radar / Iwan Holleman

01-02 Comparison of modeled ozone distributions with ozonesonde observations in the tropics / Rob Put

01-03 Impact assessment of a doppler wind lidar in space on atmospheric analyses and numerical weather prediction / G.J. Marseille, A. Stoffelen, F. Bouttier, C. Cardinali, S. de Haan and D. Vasiljevic.

02-01 Quality control and wind retrieval for SeaWinds / M. Potabella and A. Stoffelen

02-02 Shortwave radiation and cloud parameterizations for intermediate complexity models / J.J. Beersma, R. van Dorland and J.D. Opsteeg

02-03 Sensitivity study of the residue method for the detection of aerosols from space-borne sensors / M. de Graaf

02-04 Assimilation of satellite derived surface heating rates in a Numerical Weather Prediction model / Bart van den Hurk and Han The.

02-05 On the use of physical and statistical downscaling techniques for NWP model output / Wim de Rooy and Kees Kok

KNMI Bibliotheek, postbus 201, 3730 AE De Bilt, tel. 030 - 2 206 855, fax 030 - 2 210 407; bibliotheek@knmi.nl