Chapter 5

Assimilation of water vapour measurements

Atmospheric water vapour plays an important role in the atmospheric energy budget especially for the creation and distribution of clouds and precipitation which raised the question whether a numerical weather prediction model could benefit from the atmospheric water vapour data retrieved from satellite measurements. A test was performed in collaboration with the Swedish Meteorological and Hydrological Institute SMHI, during which MODIS water vapour measurements were used alongside the operational data like radio soundings and aircraft reports within the data assimilation for the numerical weather prediction model HIRLAM. The experiment was performed for the two first weeks of July 2002.

The first section gives a small introduction into the principles of data assimilation. The next section describes how the MODIS data was prepared for the use within the data assimilation system and shows how the data fulfils its requirements. The last section contains the results from the assimilation experiment.

5.1 3D Variational data assimilation

The general goal of data assimilation is well expressed by a citation from A. Lorenc (The MetOffice): “Assimilation is the process of finding the model representation

which is most consistent with the observations" [30]. For this purpose, a short-range model forecast is used as an initial first-guess, which is modified during the assimilation such that the best agreement between the model and all available observations is achieved. "All available observations" typically includes a vast variety of different quantities, observed by very different instruments, ranging from meteorological synop-stations and ships over radiosondes and aircrafts to satellite-based instruments. "Best agreement" does not mean that after the assimilation the model is meant to exactly reproduce the observations, which would not only be impossible but also wrong. Both, the model and the observations have errors, which necessarily leads to a deviation between both. Also, deviating measurements of the same quantity coming from two different sources make an exact reproduction of the observations by the model impossible, or observations of different quantities might be dynamically imbalanced. Therefore, the "most consistent" model state is looked for, which is seen as the best possible representation of the true atmospheric state given the observations and the model first guess. In the following, a short introduction to data assimilation theory using Bayes Theorem is presented based on [30].

Bayes Theorem says that "the probability that x and y both occur equals the probability that y occurs times the probability that x occurs *given that y has occured:*

$$p(x \cap y) = p(y)\, p(x \mid y) = p(x)\, p(y \mid x)\,. \tag{5.1}$$

For the sake of simplicity, let x for the moment be one single grid point temperature and y a temperature observation at the location of the grid point. For the assimilation, an actual observation y_o with an observation error variance v_o shall be used. The model forecast valid at the time of the assimilation is the temperature x_f which has a forecast error variance v_f. Assuming that the forecast error is gaussian and bias-free, the so-called *prior distribution* $p(x)$ for the model forecast can be given as

$$p(x) = N(x \mid x_f, v_f) \tag{5.2}$$

(Here, $N(x \mid a,b)$ denotes a Gaussian for the variable x with expectancy a and variance b.)

This distribution includes knowledge from previous observations and can be interpreted as a measure of our certainty that *any* model state x represents the true state of the atmosphere. Although only one representation of this probability density function is actually given by the model forecast, namely its expectancy x_f, the forecast error variance gives an estimation of the probability that any other model state would reflect the truth. If we now assume that the instrument providing the temperature measurement y_o also has a bias-free gaussian error *for all possible observations* y, then the probability of any observation y given that any model state x is the true state is also a Gaussian that has the model state as expectancy:

$$p(y \mid x) = N(y \mid x, v_o). \tag{5.3}$$

The *absolute probability* of any observation can be calculated as follows:

$$p(y) = \int p(x)p(y \mid x)\,dx. \tag{5.4}$$

The resulting probability density function is again a Gaussian with expectancy x_f and variance $v_f + v_o$. After Bayes Theorem, the product $p(x)p(y \mid x)$ equals the probability that both, x and y occur: $p(x \cap y)$and can also be expressed as $p(y)\,p(x \mid y)$. Here, $p(x \mid y)$denotes the probability that a model state x is true given that any observation y has occurred. It is called the *posterior distribution*, after adding the information from the observation to the model. For the given observation y_o it is a function of x. The maximum of this distribution is the "most consistent" model state x_a to be found during the assimilation, the most probable model state given the observation and the a priori knowledge from previous observations and the model evolution.

For the actual observation y_o, the posterior distribution is:

$$p(x \mid y_o) = p(y_o \mid x)\,p(x). \tag{5.5}$$

(The observation has already occurred, therefore its absolute probability $p(y_o)$ equals 1). Note, that $p(y_o \mid x)$ is now regarded as a function of x; in that case it is not a probability density, because its integral not necessarily equals one. It is called

the *likelihood function for x*.

Taking the negative logarithm of equation 5.5 leads to

$$-\ln\left(p\left(x \mid y_o\right)\right) = -\ln\left(p\left(y_o \mid x\right)\right) - \ln\left(p\left(x\right)\right). \tag{5.6}$$

Using equations 5.3 and 5.2, both terms on the right side of 5.5 can be expressed as Gaussians, and under the simplification that both x_f and y_o denote simple scalar values of the same quantity, the above equation can be written as:

$$-\ln\left(p\left(x \mid y_o\right)\right) = \frac{1}{2}\frac{\left(x_f - x\right)^2}{v_f} + \frac{1}{2}\frac{\left(y_o - x\right)^2}{v_o} \tag{5.7}$$

Finding the maximum of $p\left(x \mid y_o\right)$corresponds to finding the minimum of its negative logarithm. Minimising 5.7 gives:

$$x_a = \frac{x_f v_o + y_o v_f}{v_f + v_o}, \tag{5.8}$$

the most probable model state is the mean of forecast and observation, weighted with both errors. In this simple example, it can be calculated analytically. However, for real assimilation problems, both, x and y are very large vectors, and equation 5.7 becomes

$$J\left(\overline{x}\right) = \frac{1}{2}\left(\overline{x} - \overline{x}_f\right)^T B^{-1}\left(\overline{x} - \overline{x_f}\right) + \frac{1}{2}\left(\overline{y}_o - H\left[\overline{x}\right]\right)^T R^{-1}\left(\overline{y}_o - H\left[\overline{x}\right]\right). \tag{5.9}$$

$J(\overline{x})$ is the so-called *cost function* that has to be minimised during the assimilation. The model state vector $\overline{x}$ contains all model variables on all grid points, the observation vector $\overline{y}$ contains all available observations. The matrix B is the forecast error covariance matrix which contains the covariances between all elements of $\overline{x}$. The appropriate matrix for the observations is the error covariance matrix R. As the observations are typically neither made directly at the grid points nor at the exact time of the assimilation, the forward operator H transforms the model state into the observation space. For observations of model variables, e.g. temperature, H performs a spatial and temporal interpolation. For observations different from model

variables, H also has to calculate the quantity equivalent to the observations. In the case of the measurements of integrated water vapour used in this work, the forward operator performs the vertical integration of model humidity. Other examples of forward operators are the calculation of satellite radiances for their direct assimilation or the calculation of zenith path delay for the assimilation of GPS measurements[1].

Due to the large dimension of $\overline{x}$ and $\overline{y}$, B and R can generally not be inverted analytically, therefore the cost function has to be minimised numerically. A fast convergence of the minimization can be achieved by changing the model variable such that the forecast error covariance matrix is an identity matrix. Details on how this is performed within the HIRLAM 3D variational data assimilation used in this work are given in [24]. The observation handling and some assimilation experiments are described in the companion paper [29].

5.2 MODIS data for data assimilation

5.2.1 Bringing MODIS data into the HIRLAM grid

Maybe the biggest difference between satellite derived products and "classical" atmospheric measurements is their extremely high horizontal spatial resolution (in contrast to a mostly very poor vertical resolution). The MODIS data used within this work has a horizontal resolution in the sub-satellite pixel of approximately one kilometre. Vertical resolution does not exist as only integrated columnar water vapour content was used. The HIRLAM was operated with a spatial resolution of approximately 30 kilometres. As the pure amount of possibly available data would certainly be too much for the data assimilation system, a thinning, averaging or filtering of

[1] The question whether the direct assimilation of satelite measurements is advantageous compared to the assimilation of derived products is heavily debated. The main argument for the direct assimilation is that the error characteristics of the measured radiances are generally much better known than the ones of the derived products. Another argument is that e.g. temperature soundings from infared measurements or humidity soundings from GPS limb measurements rely on a model first guess or a model temperature profile, respectively, which is disadvantageouos for the assimilation, as the measurements should be independent from the model first guess. The satellite measurements used in this work are in the solar spectrum. Here, no fast forward operator exists yet, therefore retrieved products were assimilated. In cases where the operation operator (H) shows strong non-linearities the use of derived products rather than radiances may also be advantageous.

the data has to be performed. Another reason for this is that it is very likely that the observation errors are spatially correlated in the pixel scale which would violate the assumptions described in the above section. Therefore, the resolution of the MODIS data was reduced to the HIRLAM grid box size. This was done by averaging the MODIS measurements over the grid box area for each individual box. Firstly, a cloud mask and a land-sea mask were applied to the MODIS data in order to ensure that only measurements above cloud free land surfaces were taken into account. Secondly, the geographic coordinates of all appropriate MODIS pixel were transformed into the rotated HIRLAM coordinate system and the HIRLAM grid box number was calculated for each pixel in order to identify the fitting MODIS pixels for each grid box. Thirdly, the mean integrated water vapour value was calculated from all MODIS measurements within one grid box. The last step had to be performed carefully, as the surface height within one grid box can vary significantly. Averaging measurements taken above different heights should be avoided because the integrated water vapour typically does not decrease linearly with the surface height. For that reason, a two-dimensional histogram of integrated water vapour and surface height was calculated from all MODIS pixels within one grid box. Then, the water vapour measurements and the height were averaged only for pixels around the histogram's maximum[2]. Consequently, for each grid box one MODIS integrated water vapour measurement representative for one surface height was calculated. Since the used HIRLAM assimilation system is invoked in a 6 hour interval, in a last step, the different MODIS overpasses were put together accordingly: MODIS on TERRA typically takes measurements over Europe between 08 and 13 UTC. All overpasses earlier than 09 UTC were collected for a 06 UTC dataset, while all measurements taken later were put into a 12 UTC dataset.

The two resulting datasets for the first day of the experiment period are illustrated in Figure 5.1. One can see that these two datasets cover a very large area ranging from Spain to the Black Sea and from northern Africa to northern Scandinavia.

Together with the observed integrated water vapour, the observation error σ_{obs}

[2]The histogram was calculated with a bin size of 1 mm for integrated water vapour and 50 m for surface height. The avaraged values were calculated for one bin.

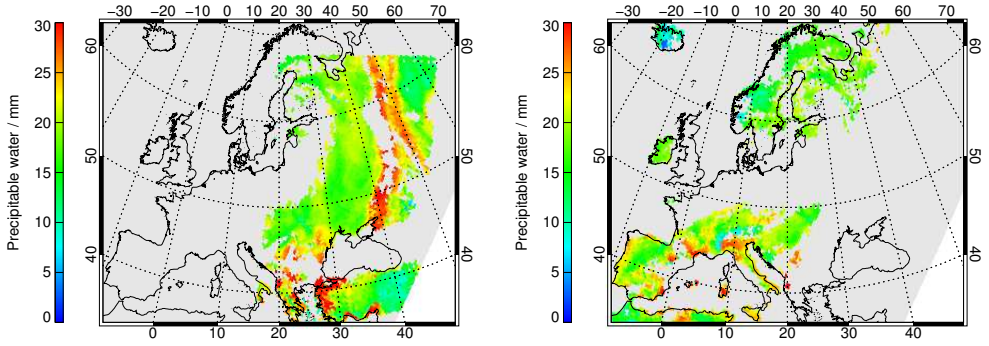

Figure 5.1: Integrated water vapour derived from MODIS measurements and averaged over HIRLAM grid boxes for 01. 07. 2002, 06 UTC (left panel) and 12 UTC (right panel).

was also calculated from an estimated inaccuracy of 2.7 mm [3] and the standard deviation of the measurements averaged over σ_{obs}^{box}:

$$\sigma_{obs} = \sqrt{2.7^2 + (\sigma_{obs}^{box})^2}. \tag{5.10}$$

5.2.2 MODIS observation error statistics

There are three major requirements with regard to observation error statistics within the assimilation theory: the measurements should be bias-free, the error distribution should be close to gaussian and the spatial correlation of the observation error should be small. These three assumptions are implicitly made during the assimilation process, and a violation can lead to bad results. Using the validation data described in chapter 4, all three requirements were tested:

The comparison of MODIS water vapour measurements with GPS measurements showed a small overall bias of 0.2 mm, however, the bias showed a strong variability with varying water vapour values which has to be further investigated in the future.

The error distribution, i.e. the frequency of occurrences of deviations of the MODIS derived water vapour from the MWR and GPS measurements is shown

[3] This value was chosen based on the theoretical inversion accuracy of the retrieval algorithm.

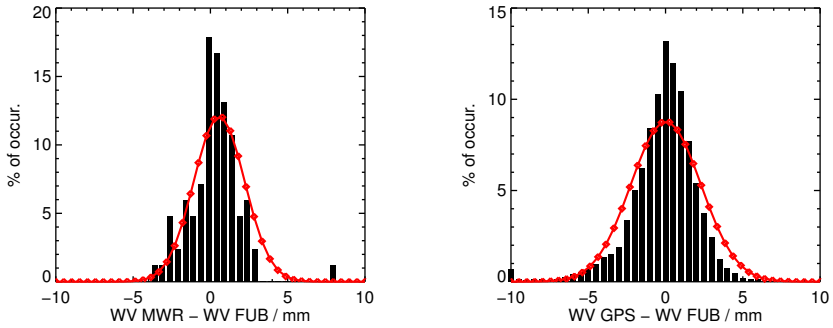

Figure 5.2: Frequency of occurrence of deviations between MODIS and MWR (left panel) and GPS (right panel) water vapour measurements. The red lines shows Gaussians around the mean differences.

in Figure5.2.Although very small deviations occur more often than predicted by the appropriate Gaussian, the second requirement of the data assimilation can be expected to be fulfilled by the MODIS measurements.

The spatial correlation of observation errors was calculated from the deviation of MODIS from GPS water vapour measurements assuming that the latter show no spatially correlated observation error[4]. The GPS stations used for this purpose are illustrated in Figure5.3. For each station pair, all timely coinciding differences between MODIS and GPS measurements were collected from the period October 2002 to September 2003. The correlation was calculated and hereafter the correlations were binned into 100 km intervals. The mean correlation was then calculated for each interval. Since only those pairs of measurements could be used where for both stations the appropriate MODIS pixels were cloud free, the actual number of useful data points for each station pair was smaller than the number of data available for each individual station. This reduced the statistical significance of the individual calculated correlations. Therefore, the distance-dependent correlation was also calculated alternatively by first binning all available difference-datasets and then calculating the correlation.

In Figure 5.4, the individual correlations for the different station pairs are shown

[4]It has to be noted, though, that this assumption may be poor as GPS results may be affected by e.g. satellite clock errors, thus leading to horizontally correlated observation errors.

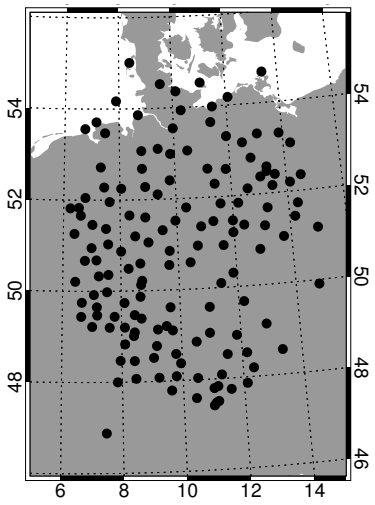

Figure 5.3: Ground-based GPS stations used for the calculation of the spatial correlation of MODIS water vapour observation errors.

as a function of distance as a colour-coded 2-dimensional histogram, with the number of occurrences increasing from blue to red. The mean correlation as a function of distance is shown in black and the alternatively calculated correlation in red. The latter is always slightly smaller, as more independent measurements were used for their calculation. However, it can clearly be seen that a non-negligible correlation is visible in the data. One possible reason for such a spatial correlations is e.g. atmospheric aerosol loading which can occur in comparable spatial scales. Another reason could be extended broken cloud fields, which can also lead to correlated deviations over larger areas.

The first possible solution would be a further thinning of the data, however, as the large spatial resolution of the satellite data is one of its big advantages, this attempt is not really satisfactory. Another solution would be to model the horizontal correlation in future assimilation experiments.

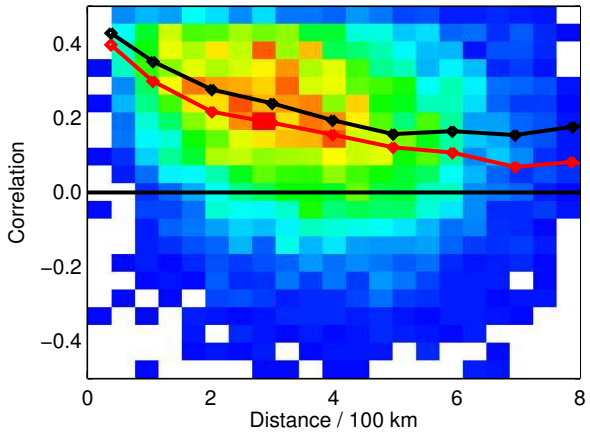

Figure 5.4: Spatial correlation of MODIS water vapour observation errors as a function of distance. The results are based on comparisons with ground-based GPS measurements. For details see text.

5.3 The assimilation experiment

The data assimilation experiment was performed at the SMHI for a 2-week-period from 1st until 15th July 2002. Two different model runs were performed: one reference run, during which all observations types operationally available to the HIRLAM 3D- Var system were used, and one "MODIS" run during which the integrated atmospheric water vapour above land derived from MODIS was used as an additional data source. In the following three subsections we will investigate

- the quality of the MODIS data with respect to the reference run,
- how the model reacted to the new data source and
- which impact the additional moisture information had on the forecasts.

5.3.1 Innovation vectors

The innovation vector is defined as the difference between observations and the appropriate model field for observations which have not been assimilated, yet. Accordingly, it shows how well the model can predict the new data. This difference

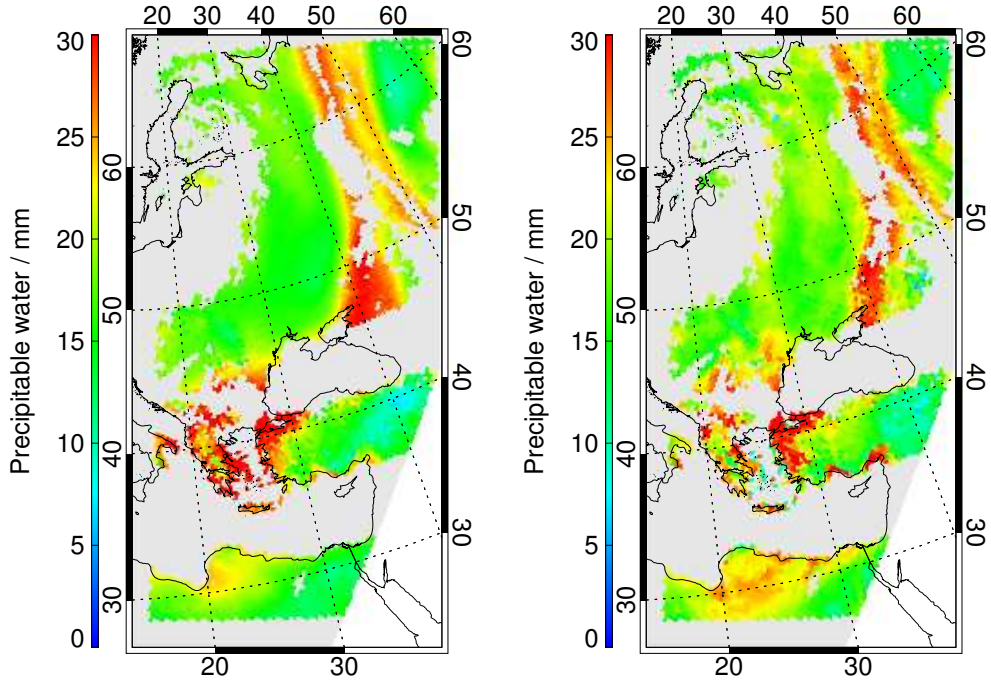

Figure 5.5: MODIS scene and HIRLAM forecast at 01. 07.2002, 06 UTC
left: 6 hours HIRLAM forecast of precipitable water vapour
right: MODIS obeservations

was calculated for the MODIS data for all assimilation cycles from the reference run where the model was never affected by the new data. For each assimilation cycle, the difference was calculated between the MODIS observations and the appropriate 6 hour forecast valid at the same time. An example is shown in Figures 5.5 and 5.6 for the first assimilation cycle. The first two images show the model 6 hours forecast valid at 01.07.2002 6 UTC and the MODIS observations of precipitable water vapour from the overpass at 01.07.2002, 8:50 UTC , the next two figures show a true colour image created from the MODIS measurements together with the difference between MODIS observations and the model forecast.

A first look at both water vapour fields shows a good agreement between HIRLAM and MODIS, however, the difference field reveals deviations especially in areas affected by clouds. Areas showing broken cloud fields seem to lead to a positive, areas with larger cloud systems to a negative bias. It should be noted here that the

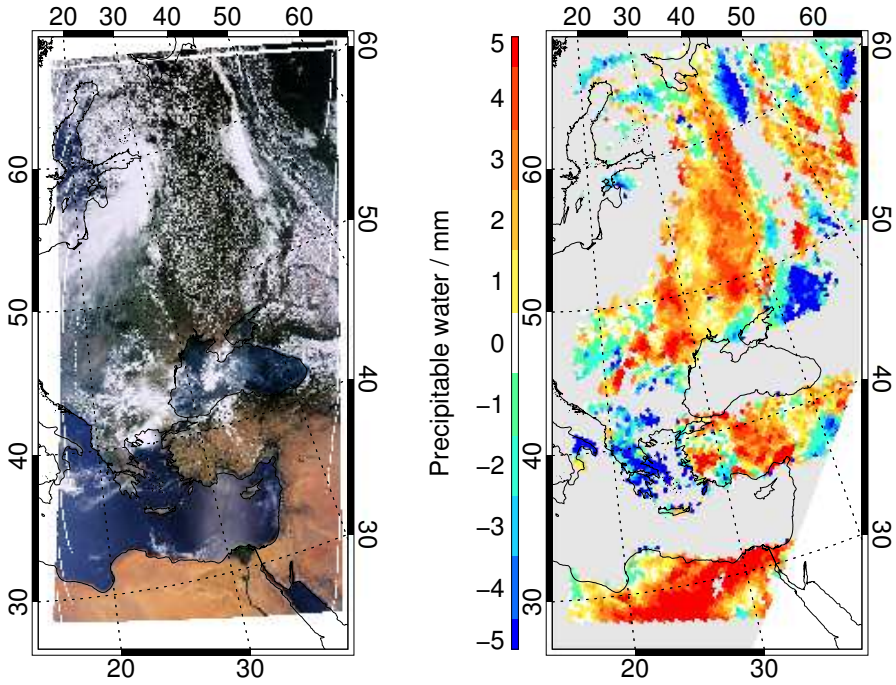

Figure 5.6: MODIS scene and HIRLAM forecast at 01. 07.2002, 06 UTC
left: true colour image of the MODIS scene
right: MODIS observations - HIRLAM forecast of precipitable water vapour

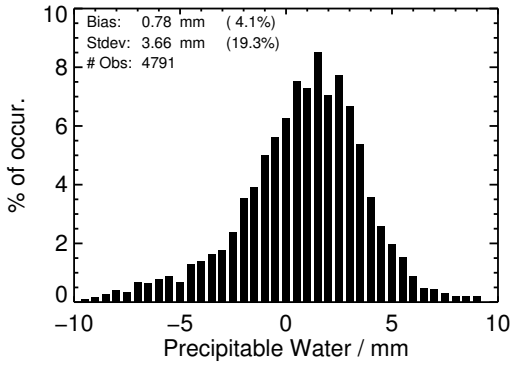

Figure 5.7: Difference histogram between MODIS observed integrated water vapour and the HIRLAM 6 hour forecast from the example showed in figures 5.5 and 5.6.

MODIS measurements are only based on cloud free pixels within one HIRLAM grid box creating areas with no data when a cloud field extends over the size of one grid box. The observed deviations can be explained with cloud radiation interactions: Broken cloud fields can be expected to increase the mean photon path length on the sun-sensor path even for cloud free pixels by scattering photons into the field of view of the satellite. This increases the absorption measured by the satellite and so leads to larger water vapour estimates. The largest negative deviations occur in areas with relatively thick and extended clouds. Taking this into account, one could conclude that the MODIS cloud mask might fail in detecting all cloudy pixels. The water vapour derived for such a pixel will always be significantly lower than for a cloud free pixel, as through the backscattering of solar radiation at the upper part of the cloud a great part of the water vapour column is not “visible” to the satellite. If the cloud mask classifies some of these pixels as cloud free, these false measurements will also be used in the averaging and the resulting mean columnar water vapour will show a negative bias.

The frequency of occurrence of the differences between MODIS and HIRLAM for this scene is plotted in Figure 5.7. The mean difference is 0.8 mm, the standard deviation 3.7 mm.

The comparison results for the average values of precipitable water vapour for

the whole experiment period are shown in Figure 5.8.

The upper two panels show the averaged HIRLAM forecasts and MODIS observations, respectively, the middle panels show the absolute and relative differences. For each HIRLAM grid box, the number of available observations was of course different, as it depends on the location of the different MODIS overpasses and on the cloud cover during each overpass. In the lower left panel, the number of observations (or forecasts, respectively) used for averaging the water vapour measurements is therefore plotted. In the lower right panel, the histogram of the absolute deviations is shown. The bias of the averaged values is -0.34 $\pm$0.57 mm, the standard deviation is 3 mm. The range given for the bias is the standard deviation of all biases calculated individually for each assimilation cycle. They show a noticeable variation, making a very simple bias correction by just subtracting a mean bias from all observations questionable.

With regard to the absolute values, no significant, artificial differences between the MODIS observations and the HIRLAM forecasts can be observed. Both, the model and the satellite measurements show a strong bias in east-west direction. For comparison purposes, the mean columnar water vapour was also calculated and triangulated to the whole area from the radiosondes used in chapter4. The result is shown in Figure5.9. Given the small number of radiosonde stations used for the triangulation, the result must be interpreted with care, but at least the east-west gradient seems to be confirmed by the radio soundings. The values close to zero over the black sea are purely a triangulation artefact as there are no radiosonde measurements to support this. The same applies for the low values at the western border and the very high values west of Gibraltar. A certain disagreement can be observed over Great Britain.

The differences between MODIS and HIRLAM show a strong latitudinal dependancy with MODIS being dryer than HIRLAM above 60°N and between approximately 40°N and 45°N. In between and especially south of 30°N, MODIS is mostly more humid than HIRLAM. Further studies will have to show whether this latitudinal dependency is statistically significant or not.

Recapitulating the results shown here and from chapter 4, the different validation and comparison results are presented in Table 5.1.

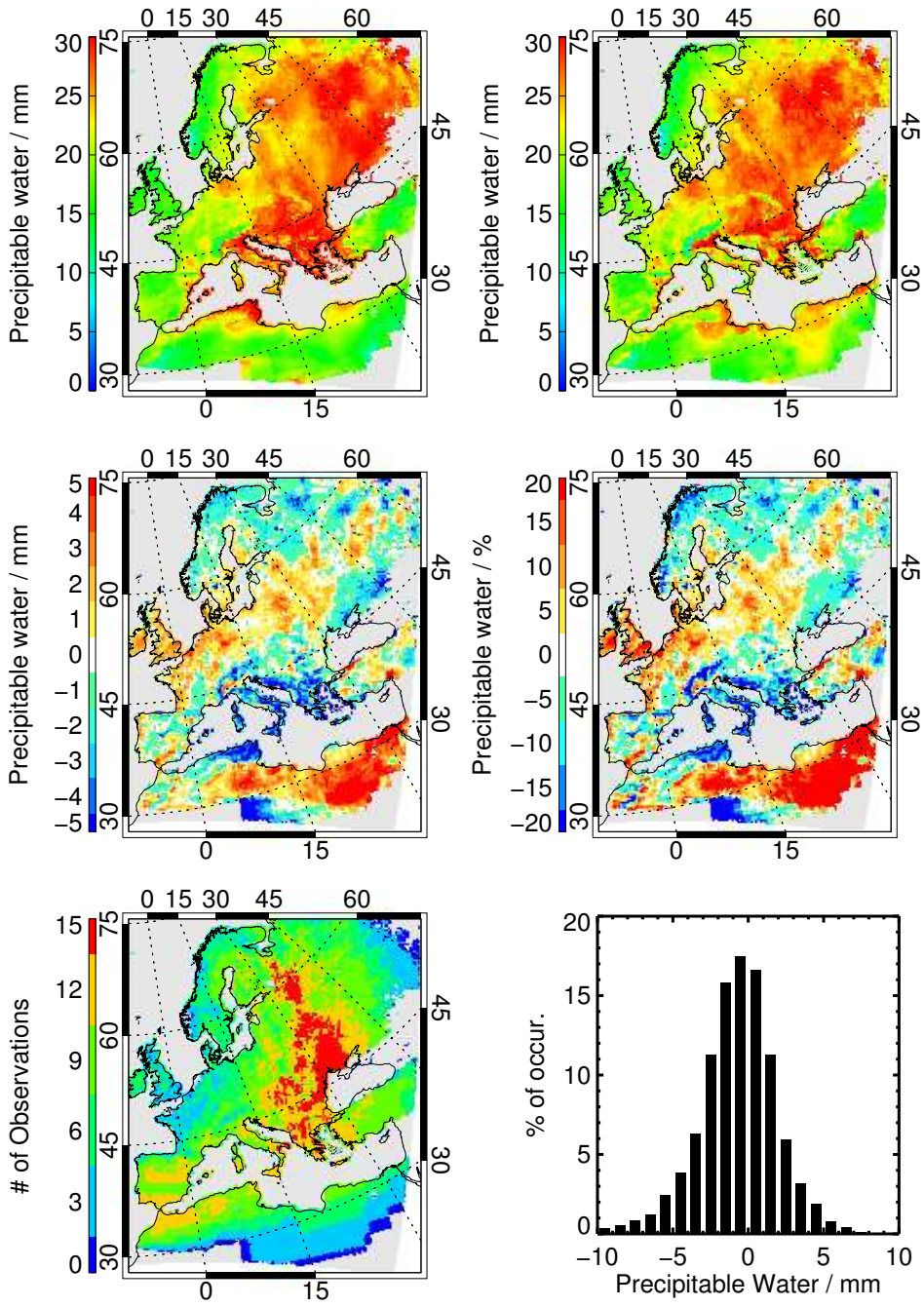

Figure 5.8: Mean precipitable water for the first two weeks in July:

6 hours HIRLAM forecast	MODIS obeservations
Observations - forecast	Relative differences
Number of observations used for averaging	Difference histogram

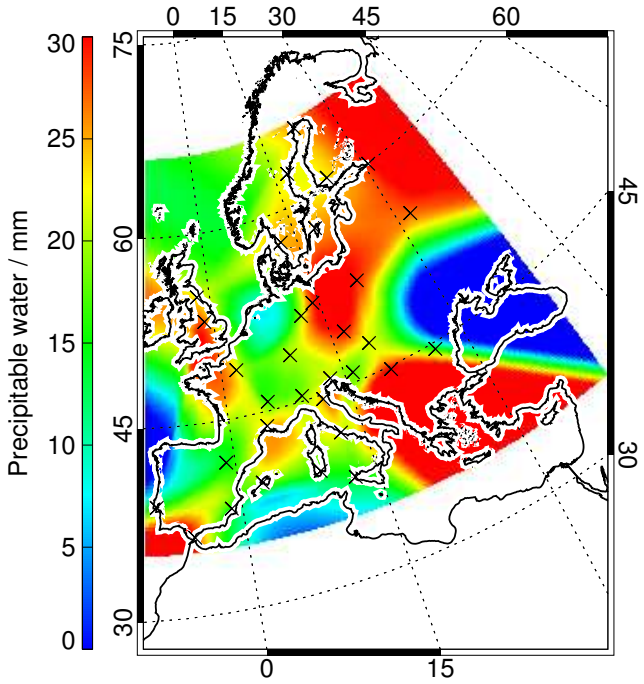

Figure 5.9: Mean precipitable water calculated from radio soundings for the assimilation experiment period. The crosses indicate the location of the radiosonde stations.

Table 5.1: Biases and root mean square deviations (rmsd) from all MODIS validations and comparisons presented in this work.

	MODIS - MWR	MODIS - RS	HIRLAM - MODIS
bias / mm	0.2	0.6	-0.3
rmsd / mm	2.1	2.0	3.0

5.3.2 Assimilation increments

The next open question is how the model incorporates the additional information. Are the measurements accepted by the model? How are the humidity fields modified when deviations between the observations and the model occur? The first question was already answered in the last section, as the innovation vectors were only calculated for the MODIS observations accepted by the 3D-Var procedure. In order to answer the second question, three different informations are needed: the humidity profile before the assimilation, coming from the 6 hours forecast, the observed integrated water vapour and the humidity profile after the assimilation. Unfortunately, the 6 hours forecasts are currently not available, therefore the humidity profiles from the two assimilations with and without the MODIS data included[5] data will be compared in the following. Of course the 6 hours forecast in the MODIS run is not the same as the analysis in the reference run, however, these datasets can be used to qualitatively illustrate the general reaction of the HIRLAM to the MODIS data. As an example, the impact of two MODIS measurements from the 2^{nd} July, 12 UTC is shown in Figure 5.10.

In the left panel, the impact of a MODIS observations larger than the appropriate first guess is shown. From the 6 hours forecast the humidity profile is not available but only the integrated water vapour, however, this value (18 mm) is considerably close to the appropriate value of the reference run analysis (16 mm), especially with regard to the sign of the difference to the MODIS observation (27 mm). Using the analysis humidity profile instead of the first guess will therefore not lead to qualitatively different results. The Figure shows both humidity profiles and their difference, and one can see that the largest differences occur between 900 and 600 hPa. The right panel shows the inverse case where a MODIS observation (17 mm) is smaller than the first guess (23 mm) and the appropriate reference run analysis (25 mm). As a result, the difference in the humidity profile is negative, again the largest differences occur around 800 hPa.

For all assimilation cycles, the humidity differences were calculated accordingly for all HIRLAM grid boxes where MODIS observations were available. In Figure

[5] i.e. from the MODIS and the reference run

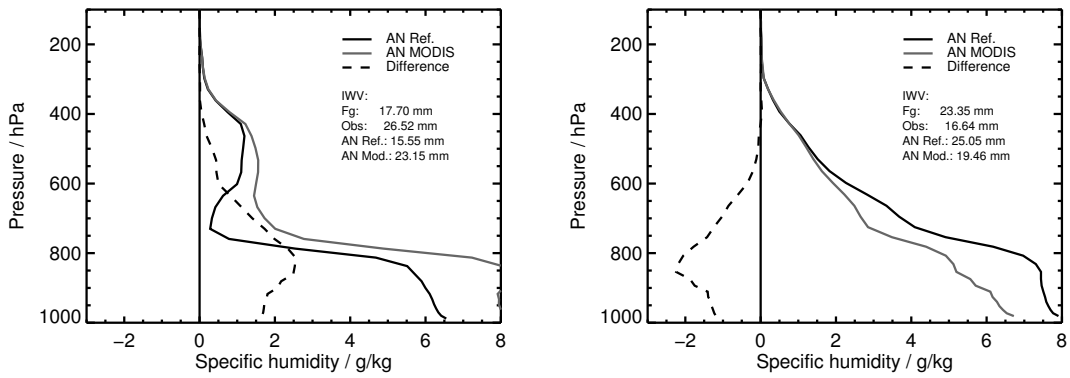

Figure 5.10: HIRLAM Humidity profiles of two grid boxes from analyses for the $2^n d$ of July, 12 UTC. The analysis from the reference run is plotted in a solid black line, the analysis from the MODIS run is plotted in grey. The dashed line shows the difference between both. The left panel shows a MODIS observation larger, the right panel an observation smaller than the 6 hours forecast

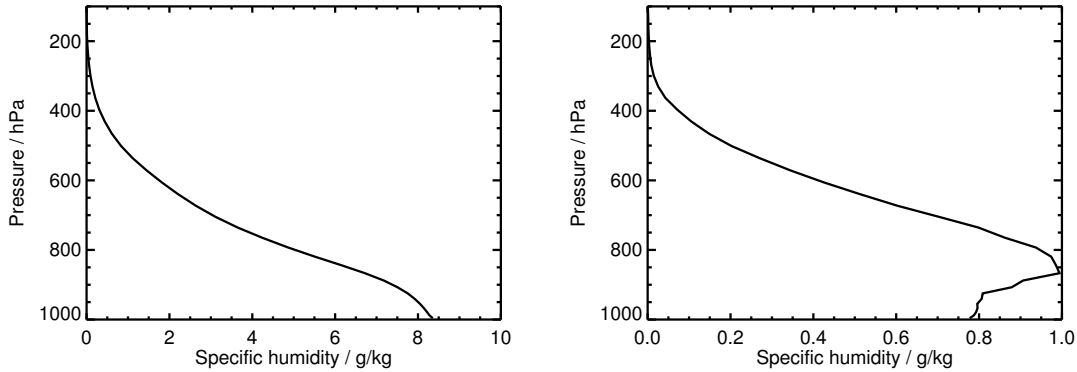

Figure 5.11: Mean humidity profile and standard deviation of analysis increments for all assimilation cycles and all MODIS observations.

5.11 the resulting mean humidity profile and the standard deviation of differences are given. The mean vertical distribution of the assimilation increments also reflects the above example: most of the differences are distributed around the 800 hPa level. The reason for this behaviour is given in [24]: The vertical distribution of the water vapour increment during the assimilation is mainly determined by the profile of assumed standard deviation of the HIRLAM background specific humidity. The model allows the largest modifications where the assumed model error is high. This profiles peaks around 800 hPa, leading to the shape of the water vapour increments shown above.

These results illustrate that the MODIS data has a reasonable impact on the humidity fields during the assimilation cycle. However, the additional information only results in short-term variations of the model humidity and the MODIS run quickly returns towards the reference run when no further MODIS data is available. The following Figures show the differences of the integrated water vapour fields from the MODIS and the reference run for 4 consecutive days. The difference is shown for the 00 UTC and the 12 UTC analysis for the complete model domain in Figure 5.12. In Figure 5.13, the difference fields are plotted for the same days, showing only the area of MODIS observations. The following two sets of Figures show the same plots for 4 days at the end of the assimilation experiment. One can see in all Figures that the additional information from the MODIS data does lead to significant deviations between the two assimilations at 12 UTC. At 00 UTC, the last MODIS information was typically 12 hours ago, and for all shown days it is obvious that the differences between the MODIS and the reference run are generally smaller compared to 12 UTC. Additionally, following the general wind direction, the differences appear to move eastwards out of the model domain. Although the additional information from the MODIS is correctly incorporated into the model, most of the forecast deviations, if any, must be expected to occur there.

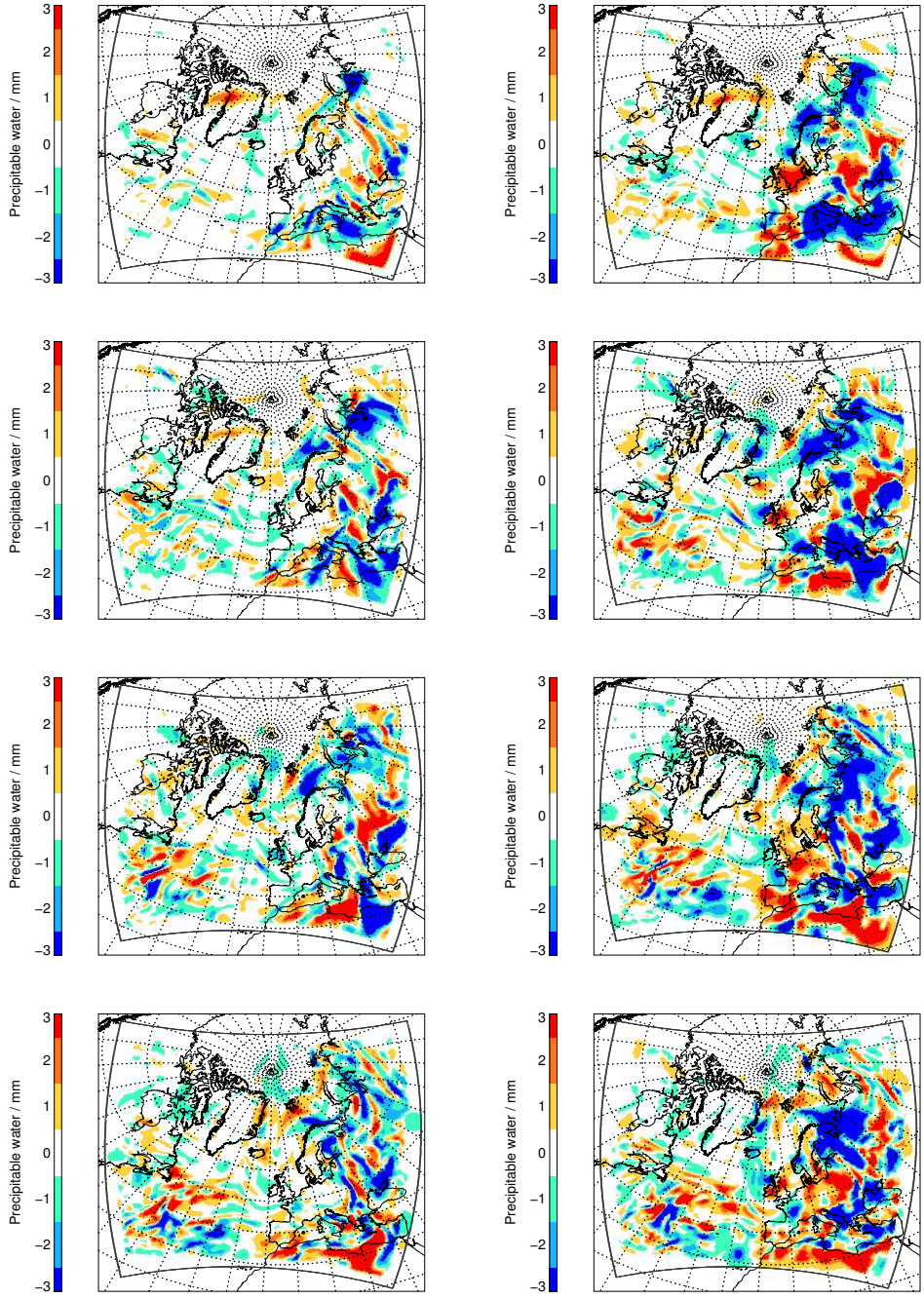

Figure 5.12: Differences in integrated water vapour between the MODIS and the reference run, the individual rows are for the 02. to 05. 07, the left panel always shows the difference at 00 UTC, the right panel at 12 UTC.

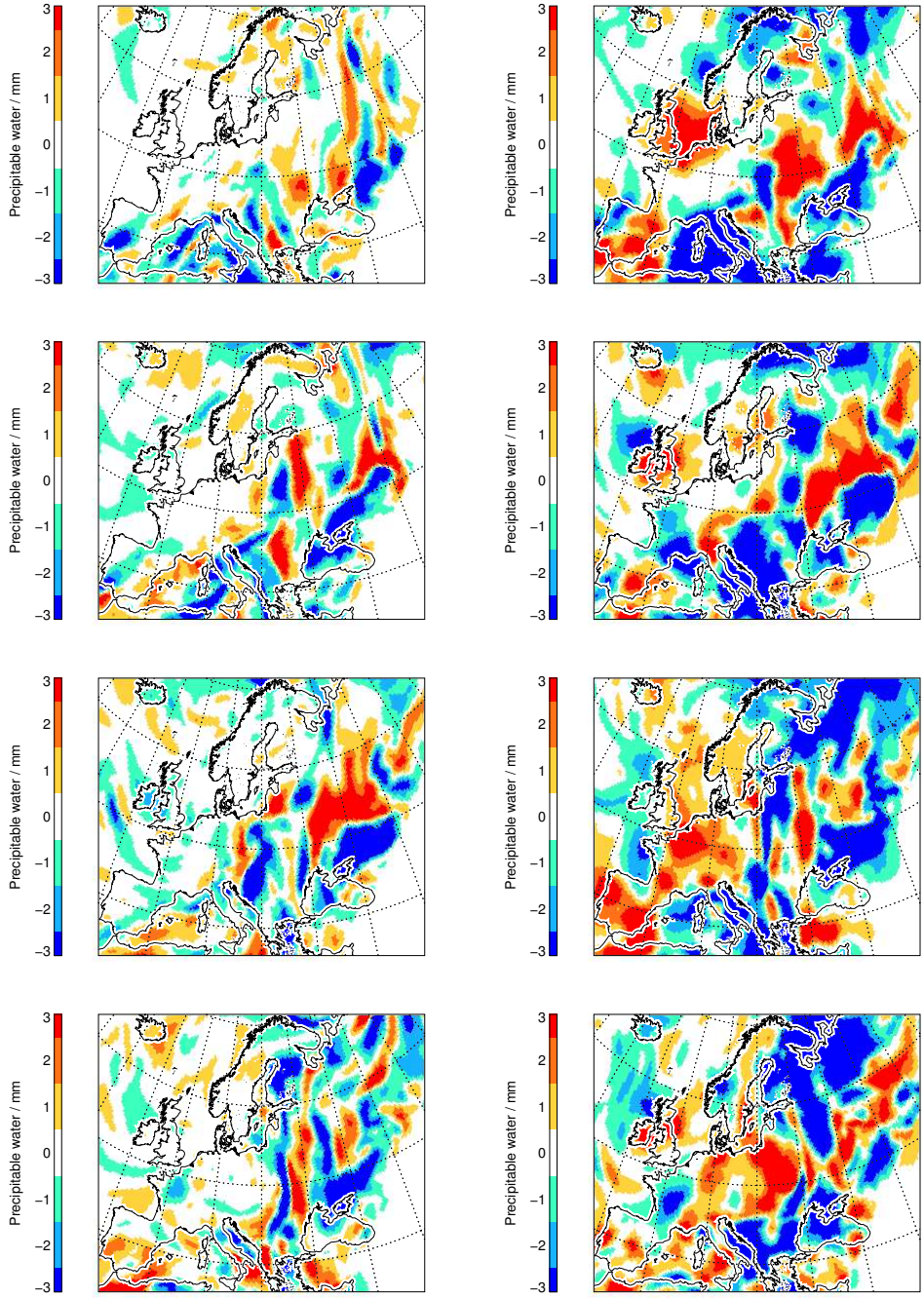

Figure 5.13: Like figure 5.12, but for the area of MODIS observations.

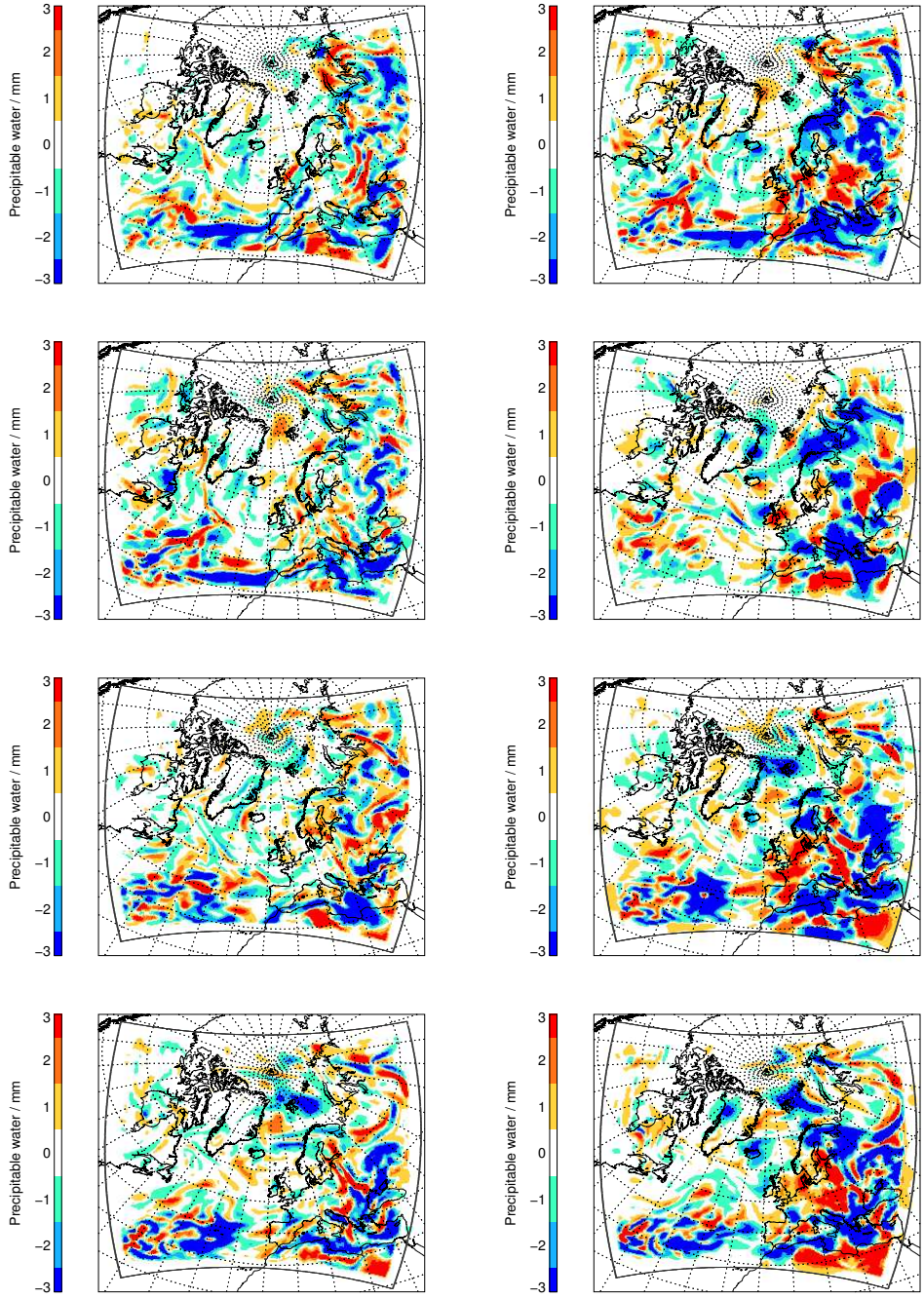

Figure 5.14: Like figure 5.12, but for 09. to 12. 07.

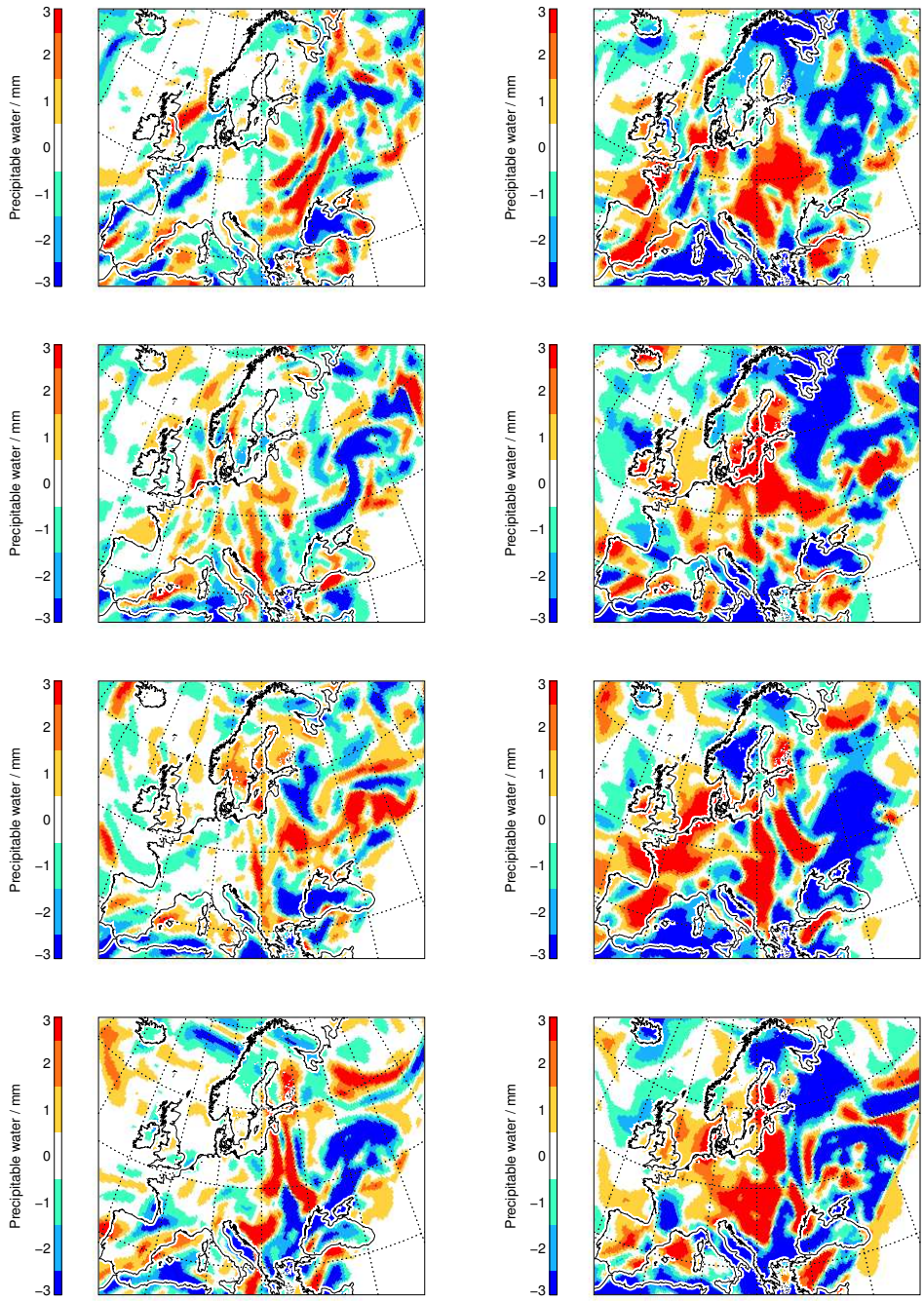

Figure 5.15: Like figure 5.13, but for the 09. to 12. 07.

5.3.3 Impact of MODIS data on forecast

A modification of the model fields of specific humidity is most likely to result in modified precipitation forecasts, therefore the impact of the additional water vapour observations on the model forecast were investigated with regard to this model variable. For a quantitative comparison with independent precipitation measurements, radar observations for the Baltic Sea Experiment (BALTEX) were made available by Daniel Michelson from SMHI. Measurements from the BALTEX Radar Network (BALTRAD) are available in 3 hours intervals [31, 26]. The radar precipitation measurements are adjusted to rain gauges using SYNOP data. In [26] a validation of the radar measurements was performed using offline gauges. Precipitation distributions tend to be skewed, therefore the 10-logarithm of precipitation was used there in order to approximate a normal sample distribution. The validation was then performed for

$$F(dB) = 10\log_{10}\left(\frac{G}{R}\right), \tag{5.11}$$

with G being the gauge and R the radar precipitation measurements. The validation revealed a mean deviation of radar from gauge precipitation measurements of less than 1 dB for most of the cases. This corresponds to less than approximately 20% underestimation of precipitation by radar. Larger deviations occurred in wintertime for distances between radar and gauge larger than 160 km, only.

After each assimilation at 00 and 12 UTC, 6, 12, and 18 hours forecasts were performed. In the first few hours after the assimilation, precipitation forecasts are generally less reliable due to model spin-up/soin-down processes. Therefore, the 12 hour precipitation was calculated from the difference between the 18 hours and the 6 hours forecasts of integrated precipitation. The appropriate radar measurements were calculated from 4 consecutive 3 hours intervals.

The impact of the additional MODIS observations of precipitable water vapour on the precipitation forecasts was first investigated for the whole HIRLAM model domain. The total precipitation integrated over the whole experiment period was calculated from the 12 hours intervals for both model runs. The results for the reference run and the difference to the MODIS run are shown in Figure 5.16 for the

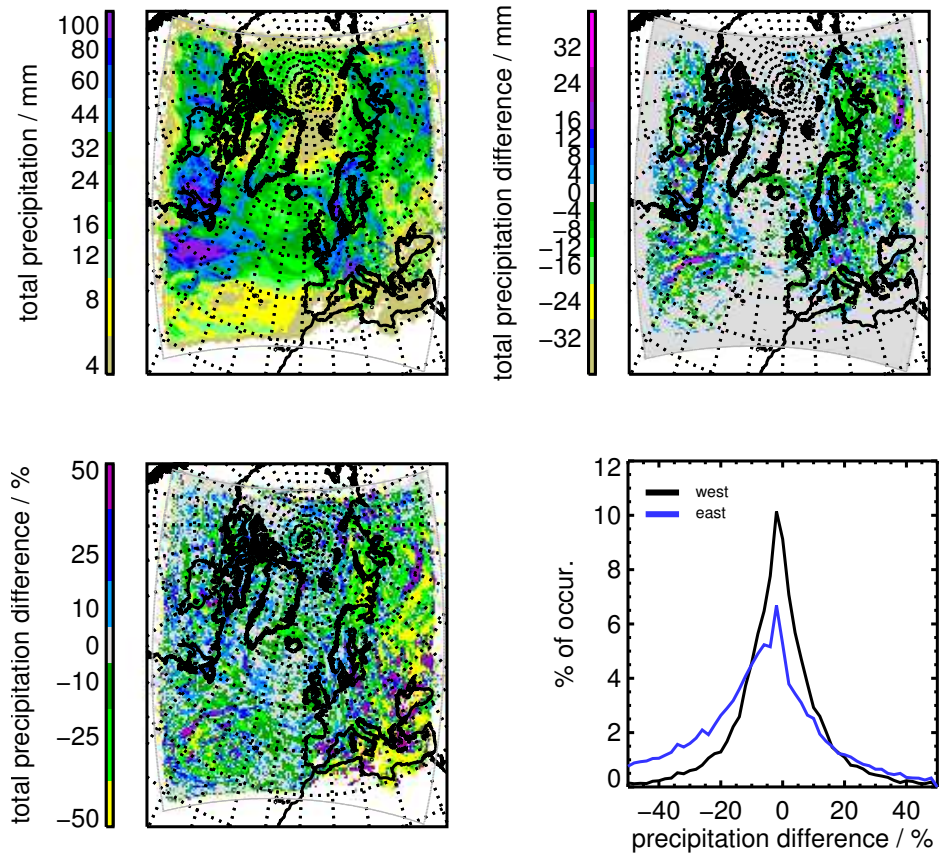

Figure 5.16: Total precipitation from 1^{st} to 15^{th} of July 2002

reference run	MODIS - reference run, absolute differences
MODIS - reference run, relative differences	Histogram of relative differences

In the lower right panel, the histogram of relative differences between the MODIS and the reference run is shown for the western (< 20°W, black) and the eastern model area (> 20°W, blue) independently. Relative differences equal to zero were excluded.

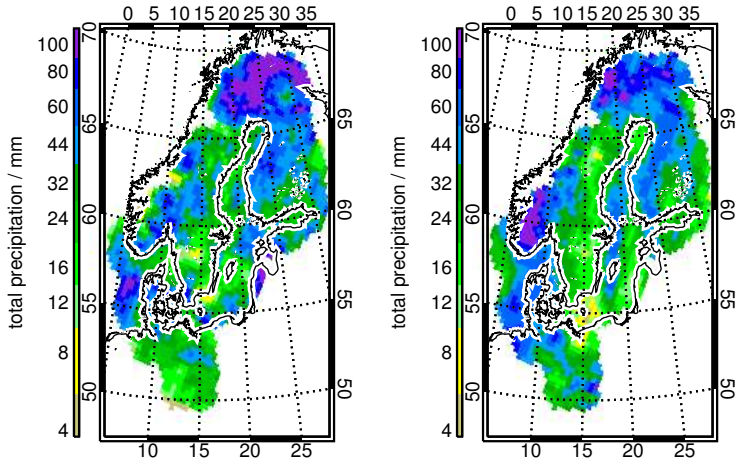

Figure 5.17: Total precipitation from 1^{st} to 15^{th} of July 2002.

Radar measurements | HIRLAM forecast, reference run

whole model area. The upper two panels show the total precipitation as predicted by the reference run and the absolute difference between the two HIRLAM runs, the lower panels show the relative differences and their histogram. It is interesting to note that significant deviations do occur in the western part of the model area although no MODIS observations were available there. Even though the absolute deviations are comparable in magnitude, the relative deviations increase significantly towards the eastern model boundaries, where relative deviations much larger in magnitude than elsewhere in the model area occur. The assimilation of MODIS generally leads to a reduction in precipitation here. To better illustrate this, the histogram of relative deviations shows two curves, the black one representing the western part of the model area (longitude < 20°W), the blue one the eastern part (longitude > 20°W). In this Figure one can well see that the MODIS run creates less precipitation than the reference run in the eastern part of the model area. The mean precipitation in the western part differs by less than 1% with ~24 mm for both model runs. In the eastern part, the difference is around 10% with a mean precipitation of

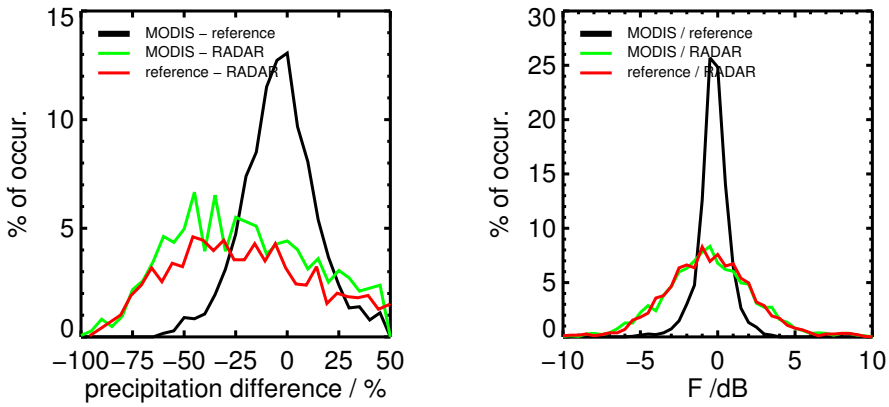

Figure 5.18: Histograms of absolute differences of total precipitation and of F(dB). For details see text.

around 18 mm for the reference run and 16 mm for the MODIS run.

Unfortunately, the above Figure also shows that the differences in predicted precipitation between the two model runs are comparably small over Scandinavia, where the radar measurements are made. Nevertheless, the integrated precipitation from radar and from the reference run are shown in Figure 5.17. One can see that the spatial shape of the precipitation fields predicted by HIRLAM generally agrees well with the radar observations. Large differences are visible e.g. in southern Norway, where the radar view is obstructed by mountains and the HIRLAM precipitation is much larger. On the other hand, radar precipitation by far exceeds HIRLAM in northern Scandinavia.

The last figure shows the histograms of relative differences between total precipitation from radar and from the reference and the MODIS run, respectively and between the two model runs. Also shown are the histograms of the variable F(dB). The differences between the two model runs are too small for an effective improvement with regard to the radar observations. It has to be noted on the other hand that the histograms of F(dB) for the ratio of HIRLAM and radar precipitation are very similar in shape to the appropriate histograms for the ratio of rain gauge and radar given in [26]. With regard to the evaluation of the impact of MODIS water

vapour observations on the precipitation forecast, no final conclusion can therefore be drawn from this comparison. It would be desirable in the future to use independent validation data in areas where stronger variations in precipitation occurred.