

# Assimilating remotely sensed cloud optical thickness into a mesoscale model

D. Lauwaet<sup>1,2</sup>, K. De Ridder<sup>2</sup>, and P. Pandey<sup>1,2</sup>

<sup>1</sup>Physical and Regional Geography Research Group, Department of Earth and Environmental Sciences, K. U. Leuven, Celestijnenlaan 200 E, 3001 Heverlee, Belgium

<sup>2</sup>Vlaamse Instelling voor Technologisch Onderzoek (VITO), Boeretang 200, 2400 Mol, Belgium

Received: 13 April 2011 – Published in Atmos. Chem. Phys. Discuss.: 2 May 2011

Revised: 19 September 2011 – Accepted: 3 October 2011 – Published: 14 October 2011

**Abstract.** The Advanced Regional Prediction System, a mesoscale atmospheric model, is applied to simulate the month of June 2006 with a focus on the near surface air temperatures around Paris. To improve the simulated temperatures which show errors up to 10 K during a day on which a cold front passed Paris, a data assimilation procedure to calculate 3-D analysis fields of specific cloud liquid and ice water content is presented. The method is based on the assimilation of observed cloud optical thickness fields into the Advanced Regional Prediction System model and operates on 1-D vertical columns, assuming that the horizontal background error covariance is infinite, i.e. an independent pixel approximation. The rationale behind it is to find vertical profiles of cloud liquid and ice water content that yield the observed cloud optical thickness values and are consistent with the simulated profile. Afterwards, a latent heat adjustment is applied to the temperature in the vertical column. Data from several meteorological stations in the study area are used to verify the model simulations. The results show that the presented assimilation procedure is able to improve the simulated 2 m air temperatures and incoming shortwave radiation significantly during cloudy days. The scheme is able to alter the position of the cloud fields significantly and brings the simulated cloud pattern closer to the observations. As the scheme is rather simple and computationally inexpensive, it is a promising new technique to improve the surface fields of retrospective model simulations for variables that are affected by the position of the clouds.

## 1 Introduction

Mesoscale atmospheric models are used extensively to reconstruct high-resolution regional atmospheric conditions as an input for e.g. hydrological, land surface or air pollution models. Although sophisticated techniques are used to parameterize clouds and precipitation, a large source of uncertainty in the model results remains in predicting the location of cloud systems at high spatial resolutions. As clouds have a strong impact on the surface energy budget and hence the local temperatures, an inaccurate simulation of the overlying cloud cover is problematic for certain applications that need correct surface input data. The assimilation of satellite data into the atmospheric model can play an important role in providing improved model results on a local scale.

Cloud assimilation studies have focused mainly on cloud retrievals from radar data, either with one-dimensional variational schemes (1DVAR) (Benedetti et al., 2003) or with more complex models in 3DVAR (Hu et al., 2006a, b) and 4DVAR (Sun and Crook, 1998; Vukićević et al., 2004). Recently, Benedetti and Janisková (2008) used a 4DVAR system to assimilate Moderate Resolution Imaging Spectroradiometer (MODIS) cloud optical depth observations into the European Centre for Medium range Weather Forecast (ECMWF) model. Their results show a positive impact on certain variables like the distribution of cloud ice water content but the assimilation did not always improve the analysis fit to the observations. However, the large computational infrastructure needed to run and maintain these systems are limiting their use for smaller research centres and universities.

Other simpler and faster methods exist that attempt to retrieve model cloud microphysics from satellite observations or other sources. Soutu et al. (2003) constructed cloud fields for their forecasts over the Galician Region in Spain based



Correspondence to: D. Lauwaet  
(dirk.lauwaet@vito.be)

on relative humidity values from the NCEP Aviation Model. Their procedure followed the Local Analysis and Prediction System (LAPS, Albers et al., 1996) and clearly improved the model's skill to predict precipitation amounts. Another method is used by Yucel et al. (2003), who applied a nudging assimilation technique to ingest remotely sensed cloud cover and cloud top height data into their mesoscale atmospheric model. The cloud ingestion was found to improve the ability of the model to capture the variation in surface fields associated with cloud cover. However, they suggested that it would be necessary to modify the model dynamics and thermodynamics to be consistent with the ingested cloud fields.

In this context, the goal of the research reported here is to present a rather simple and computational fast cloud assimilation scheme. The scheme applies optimal interpolation with latent heat adjustment for the assimilation of cloud optical thickness (COT) observations into a mesoscale atmospheric model to study the effect on the simulated surface fields associated with cloud cover. The model used for this study is the Advanced Regional Prediction System (ARPS), a non-hydrostatic mesoscale atmospheric model developed at the University of Oklahoma (Xue et al., 2000, 2001). Although the ARPS model has its own cloud analysis package (ADAS, Brewster, 1996), it is not used in our study. ADAS needs information on the vertical extent of the clouds to estimate cloud types and related in-cloud vertical velocities which can not be derived from our two-dimensional cloud optical thickness data.

The remainder of the paper is organized as follows. A description of the atmospheric model and a set-up for the model simulations are given in Sect. 2. In Sect. 3, the details of the cloud assimilation scheme are presented. The results are evaluated and discussed in Sect. 4 while conclusion are given in Sect. 5.

## 2 Numerical model and data description

The Advanced Regional Prediction System (ARPS) includes conservation equations for momentum, heat, mass, water (vapour, liquid and ice), sub-grid scale turbulent kinetic energy and the state-equation of moist air. The finite-difference equations of the model are discretized on an Arakawa C-grid, employing a terrain following coordinate in the vertical direction. Advection is solved with a fourth-order central differencing scheme and leapfrog time stepping. Turbulence is represented by the 1.5-order turbulent kinetic energy (TKE) model, and the Sun and Chang (1986) parameterization for the convective boundary layer. The 6-category water/ice scheme of Lin et al. (1983) accounts for the model microphysics while the Kain-Fritsch cumulus parameterization scheme solves the cumulus convection (Kain and Fritsch, 1990). In order to suppress numerical noise, a fourth-order monotonic computational mixing was applied, following Xue (2000).

Land surface processes are parameterized by the Soil-Vegetation-Atmosphere Transfer model of De Ridder and Schayes (1997). The two primary parameters of the land surface model are the type of vegetation, which is derived from the Coordination Information Environment (CORINE) land cover data and the soil texture, which is assumed to be that of a loamy soil, homogeneous across the domain. Among the secondary parameters, vegetation fraction is based on the normalized difference vegetation index (NDVI) from the SPOT-VEGETATION satellite imagery.

Data with a  $0.25^\circ$  horizontal resolution from the global operational analysis by the ECMWF are used as initial conditions and as 6-hourly lateral boundary conditions for the model runs. The ARPS model domain has a grid spacing of 16 km and a domain size of  $1600 \text{ km} \times 1600 \text{ km}$ , centred over Paris (Fig. 1). In all simulations, 35 vertical levels are employed with a grid spacing of 25 m near the surface increasing to 1 km near the upper model boundary, located at 20 km altitude. The simulations are initialized on 1 June 2006 at 00:00 LT and run until 30 June 2006 at 24:00 LT. This month is chosen to test our assimilation scheme as during some periods of this month, the model has problems in simulating the right amount and position of clouds, which leads to large errors in some surface variables as will be shown in Sect. 4.

The cloud optical thickness images for June 2006 are retrieved from visible and shortwave infrared imagery from the Spinning Enhanced Visible and InfraRed Imager (SE-VIRI) onboard the Meteosat Second Generation satellite with a semi-analytical cloud retrieval algorithm (Pandey et al., 2011). This algorithm is based on the retrieval algorithm of Kokhanovsky et al. (2003) for the estimation of cloud optical thickness. The details of the scheme can be found in Pandey et al. (2011). As Meteosat is a geostationary satellite, the algorithm provides COT images every quarter of an hour during daytime (06:00–20:00 LT). These images are assimilated every 15 min into the ARPS model following the procedure that is explained in Sect. 3.

To test the effect of our cloud assimilation procedure, 2 m air temperature and specific humidity data for 2 observational stations close to Paris (Melun and Trappes) and a station in Bordeaux are gathered from the National Climatic Data Center (NCDC) dataset (Fig. 1). Furthermore, 2 m air temperature, specific humidity and incoming shortwave radiation data for 3 more stations (Fontainebleau, Grignon and Oensingen, Fig. 1) are taken from the CarboEurope Integrated Project.

## 3 Cloud optical thickness assimilation procedure

The data assimilation procedure applied in this study calculates 3-D analysis fields of specific cloud liquid and ice water content ( $q_c$  and  $q_i$ ) and operates on 1-D vertical columns. The rationale behind the method is to find vertical profiles

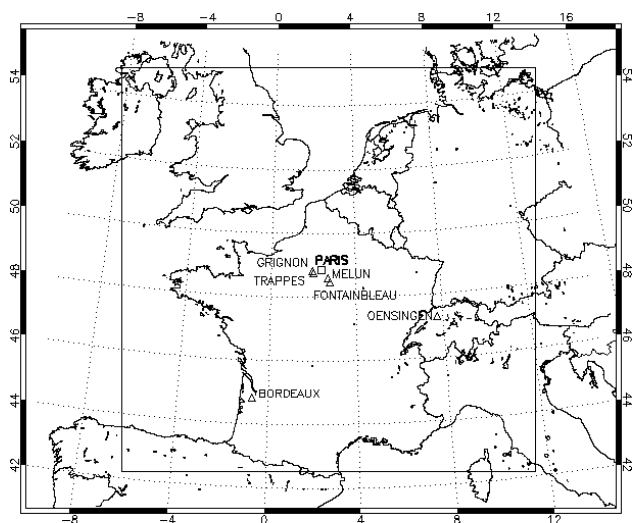


Fig. 1. Location of the model domain and the observational stations.

of  $q_c$  and  $q_i$  that yield the observed cloud optical thickness fields  $\tau_o$ , and that are consistent with the background (simulated) profile, in the sense that clouds are put in layers with a large humidity. This a priori assumption is required as  $\tau_o$  does not contain height information.

### 3.1 Background COT

Consider a vertical model column containing  $n$  grid cells irregularly spaced at positions  $z_i$  ( $i = 1, \dots, n$ ), starting from the surface. Each layer (thickness  $\Delta z_i$ ) is characterized by a simulated specific cloud water content  $q_{cbi}$ , which can be either liquid or solid (ice) water. Noting that the quantity  $\rho_i q_{cbi} \Delta z_i$  is the incremental liquid/ice water path (in  $\text{kg m}^{-2}$ ) of layer  $i$  ( $\rho_i$  being the air density of layer  $i$ ), the incremental optical depth contributed by layer  $i$  is given by (Salby, 1996):

$$\Delta \tau_{bi} = \frac{3}{2\rho_w} \frac{\rho_i q_{cbi}}{r_{ei}} \Delta z_i \quad (1)$$

with  $\rho_w = 1000 \text{ kg m}^{-3}$  the density of liquid water, and  $r_{ei}$  the effective radius of cloud droplets in layer  $i$ , which is parameterized in ARPS as a function of temperature, to account for the different values of this quantity for liquid versus solid water. The subscript “b” denotes the background fields. A list of all the symbols is provided in Appendix A.

The full model-based columnar optical depth is then given by:

$$\begin{aligned} \tau_b &= \sum_{i=1}^n \Delta \tau_{bi} \\ &= \frac{3}{2\rho_w} \sum_{i=1}^n \frac{\rho_i q_{cbi}}{r_{ei}} \Delta z_i \\ &= \frac{3}{2\rho_w} \begin{pmatrix} \rho_1 \Delta z_1 & \dots & \rho_n \Delta z_n \\ r_{e1} & \dots & r_{en} \end{pmatrix} \begin{pmatrix} q_{cb1} \\ \vdots \\ q_{cbn} \end{pmatrix} \\ &\equiv \mathbf{H} \mathbf{q}_{cb} \end{aligned} \quad (2)$$

$\mathbf{H}$  is a so-called observation operator, which linearly maps  $\mathbf{q}_{cb}$  onto a background (i.e. simulated) optical thickness ( $\tau_b = \mathbf{H} \mathbf{q}_{cb}$ ).

### 3.2 Optimal interpolation

Given observations of cloud optical thickness for a given position on the globe, the best linear unbiased estimate of cloud water content is given by (Kalnay, 2003):

$$\mathbf{q}_{ca} = \mathbf{q}_{cb} + \mathbf{K}(\tau_o - \mathbf{H} \mathbf{q}_{cb}) \quad (3)$$

with  $\mathbf{q}_{ca}$  the vector containing the analyzed cloud water content values at level  $i$ , and  $\mathbf{q}_{cb}$  likewise containing the values generated by the model (“background” or first guess value). The gain matrix is given by

$$\mathbf{K} = \mathbf{B} \mathbf{H}^T (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} \quad (4)$$

with  $\mathbf{B}$  the background error covariance matrix and  $\mathbf{R}$  ( $\equiv \sigma_\tau^2$ ) the observation error covariance, which in this case is a scalar since  $\tau_o$  itself is a scalar quantity.

We will assume that  $\mathbf{B}$  is a diagonal matrix. This is not entirely realistic, since errors of cloud water content at different vertical layers may be correlated, especially if these layers are close to each other in comparison to the typical thickness of a cloud layer. Nevertheless, it is difficult to estimate these inter-layer correlations and, moreover, the thicknesses of the layers that are prone to cloud development (sufficiently far away from the surface) are rather thick, thus making this diagonality assumption less of a problem. A diagonal background error covariance matrix has the advantage of leading to a fairly simple final expression for the analyzed specific cloud water content, as shown below. In Appendix B, the results of a test with a non-diagonal  $\mathbf{B}$  matrix are presented in order to have an estimate of the impact of this assumption.

Thus, in case  $\mathbf{B}$  is a diagonal matrix with elements  $\sigma_{ci}^2 \delta_{ij}$  (with  $\delta_{ij}$  the Kronecker delta), one has:

$$\begin{aligned} \mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R} &= \sum_{i=1}^n \sigma_{ci}^2 h_i^2 + \sigma^2 \\ \mathbf{B} \mathbf{H}^T &= (\sigma_{c1}^2 h_1 \dots \sigma_{cn}^2 h_n)^T \end{aligned} \quad (5)$$

with  $h_i$  the operation operator for layer  $i$ . This leads, finally, to:

$$q_{cai} = q_{cbi} + \frac{\sigma_{ci}^2 h_i}{\sum_{i=1}^n \sigma_{ci}^2 h_i^2 + \sigma^2} \left( \tau_0 - \sum_{i=1}^n h_i q_{cbi} \right) \quad (6)$$

The main challenge is to specify the  $\sigma_{ci}$  in a suitable manner, in particular in such a way that model layers with a high simulated humidity are more affected than the drier layers.

### 3.3 Cloud water background error variance

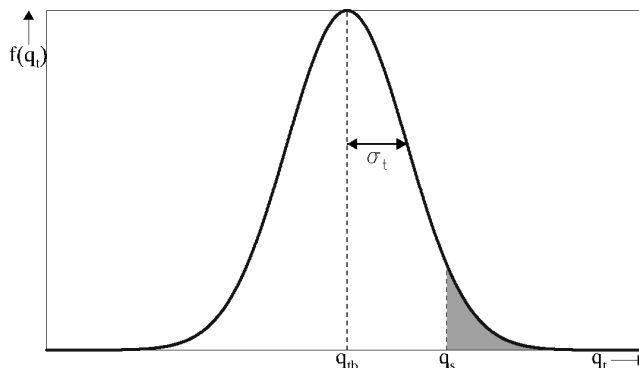
The specification of the cloud water background error variance  $\sigma_{ci}$  of each model layer is not straightforward, in particular when a layer contains no simulated liquid or ice water ( $q_{cbi} = 0$ ). One might be tempted to set  $\sigma_{ci} = 0$  in such a situation, but from the analysis equation above it is clear that  $q_{cai}$  will also be zero then, even if a cloud is observed ( $\tau_0 > 0$ ). Clearly, a non-zero cloud water background error variance must be assigned, even for non-saturated layers. Simply taking  $\sigma_{ci}$  as a fraction of  $q_{cbi}$  will not work for the reasons just explained. The background error variance matrix will therefore be established starting from a probability density function for total specific water content  $q_t$ , which is defined as the sum of vapour and cloud (liquid/ice) contributions, i.e.  $q_t = q_v + q_c$ . It should be noted that by assigning a cloud variance to a non-cloudy background layer, the analysis implicitly allows to adjust the water vapour profile in absence of background clouds. The goal is now to find the cloud water content error, given the error on total water content. The error on the latter needs to be specified a priori, in our case this will be done as a fixed fraction of total water content (see Sect. 3.4).

We employ a normal distribution, given by:

$$f(q_t) = \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(q_t - q_{tb})^2}{2\sigma_t^2}} \quad (7)$$

with  $q_{tb}$  the background (simulated) value of  $q_t$ , and  $\sigma_t$  the standard deviation of the distribution, which is a measure for the error on simulated total water  $q_t$ . Figure 2 presents the concept of the normal distribution of  $q_t$ .

Cloud water is that portion of  $q_t$  which is in excess of the saturated value, denoted  $q_s$  ( $\equiv q_{\text{sat}}(T)$ ), with  $T$  the temperature of the considered layer, so that  $q_c = (q_t - q_s)H(q_t - q_s)$ , with  $H(\cdot)$  the Heaviside step function, which is unity for a positive argument and zero otherwise. Using this, taking the simulated cloud water content  $q_{cb}$  as expected value for this quantity, and omitting the layer index  $i$  for the moment, the error variance of simulated cloud water can be calculated as follows:



**Fig. 2.** Probability distribution function for total water content (blue line), with expectation value  $q_{tb}$ , and standard deviation  $\sigma_t$ . The light shading corresponds to the area above the saturated specific humidity (denoted  $q_s$ ), which contains cloud water.

$$\begin{aligned} \sigma_c^2 &= \int_{-\infty}^{+\infty} (q_c - q_{cb})^2 f(q_t) dq_t \\ &= \int_{q_s}^{+\infty} [(q_c + q_s) - (q_{cb} + q_s)]^2 f(q_t) dq_t \\ &= \int_{q_s}^{+\infty} (q_t - q_{tb})^2 f(q_t) dq_t \end{aligned} \quad (8)$$

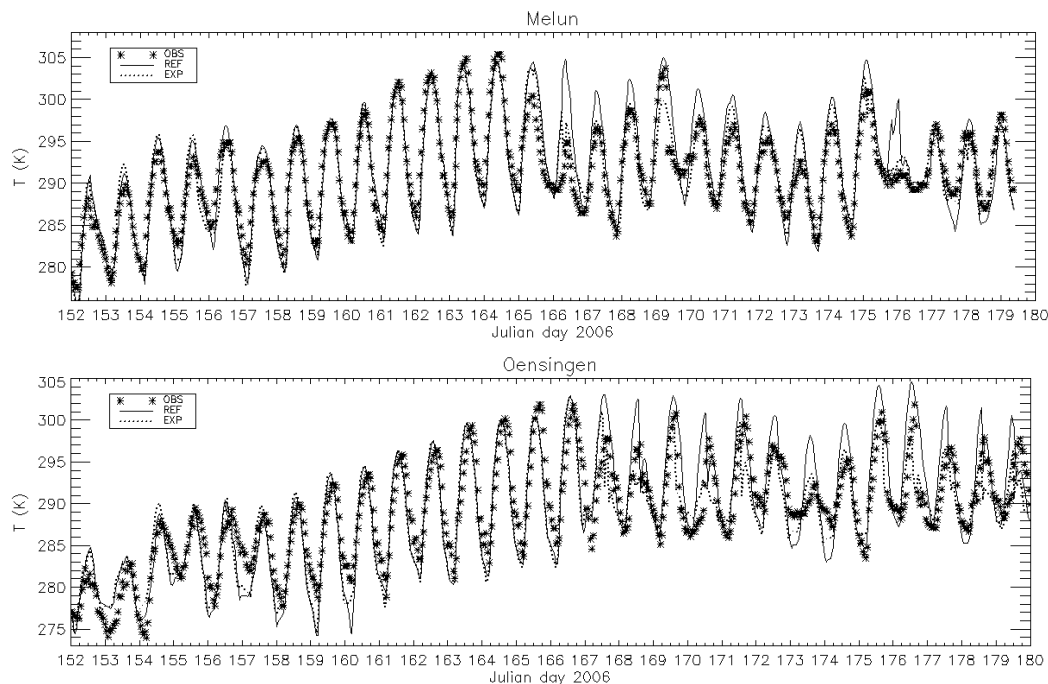
such that:

$$\begin{aligned} \sigma_c^2 &= \frac{1}{\sqrt{2\pi}\sigma_t} \int_{q_s}^{+\infty} (q_t - q_{tb})^2 e^{-\frac{(q_t - q_{tb})^2}{2\sigma_t^2}} dq_t \\ &= \frac{2\sigma_t^2}{\sqrt{\pi}} \int_{x_s}^{+\infty} x^2 e^{-x^2} dx \sigma_c^2 \\ &= \sigma_t^2 \left[ \frac{1}{2} \text{erfc}(x_s) + \frac{x_s}{\sqrt{\pi}} e^{-x_s^2} \right] \end{aligned} \quad (9)$$

with  $x_s = \frac{(q_s - q_{tb})}{\sqrt{2}\sigma_t}$  and  $\text{erfc}(\cdot)$  the complementary error function. It should be noted that in these formulas the error on the background saturated specific humidity, hence the background temperature field, is ignored to simplify the assimilation scheme. In Appendix C, the effect of this approximation is assessed and it is shown that it does not have a significant effect on the outcome of the assimilation procedure.

### 3.4 Implementation in the ARPS model

In the above, the standard deviation on the simulated total water content and the observed cloud optical thickness are still unknown. These standard deviations are expressed as a fraction of  $q_t$  and  $\tau_0$ , respectively, i.e.  $\sigma_t = a q_t$  and  $\sigma_\tau = b \tau_0$ . In our study, a value of 0.3 is adopted for coefficient  $a$  (i.e.  $\pm 30\%$  error of  $q_t$ ). This value is obtained from a comparison between modelled and observed specific humidity profiles at Trappes, as more than 80% of the observed data points are within this error margin of the simulated profiles. For coefficient  $b$ , a rather conservative value of 0.25 is adopted (i.e.  $\pm 25\%$  error of  $\tau_0$ ), with a lower limit of 5 for  $\sigma_\tau$ , based on a



**Fig. 3.** 2 m air temperature at Melun (upper panel) and Oensingen (lower panel) for June 2006.

validation study of our cloud optical thickness product with Cloudsat COT data (Pandey et al., 2011). To test the sensitivity of the assimilation procedure to these two coefficients, a sensitivity study is performed by varying the values between 0.1 and 0.5. The results are presented at the end of Sect. 4.2.

The resulting  $q_{ca}$  from the assimilation procedure is defined as cloud liquid water when the temperature is warmer than  $-10^{\circ}\text{C}$ , and as cloud ice when the temperature is colder than  $-30^{\circ}\text{C}$ . A linear ramp is applied for the temperature in between. Whenever a non-saturated (and cloudless) layer becomes cloudy after the analysis, the specific humidity is set to its saturated value. Last, a latent heat adjustment to temperature based on the added or subtracted amounts of  $q_c$  and  $q_i$  is applied, according to the formula:

$$\begin{aligned}\Delta T_{q_c} &= \Delta q_c \frac{L_v}{C_p} \\ \Delta T_{q_i} &= \Delta q_i \frac{L_v + L_f}{C_p}\end{aligned}\quad (10)$$

where  $L_v$  and  $L_f$  are the latent heat of vaporization and fusion at  $0^{\circ}\text{C}$  respectively, and  $C_p$  is the specific heat of dry air at constant pressure.

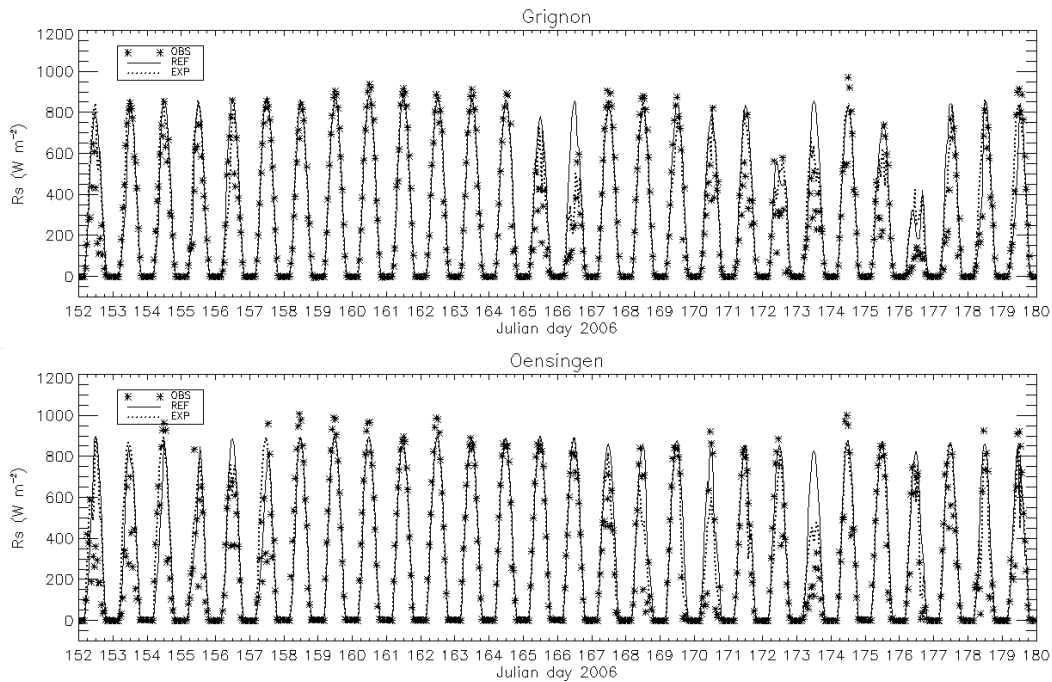
#### 4 Results of the assimilation procedure

This section describes the results for a COT assimilation experiment (EXP) for the month of June 2006, compared to a reference simulation (REF) with a setup identical to the cloud assimilation experiment to provide a benchmark for the effect of the introduction of cloud optical thickness data.

#### 4.1 Comparison to observations

Figure 3 shows the impact of the COT assimilation on 2 m air temperatures, measured at Melun and Oensingen (locations in Fig. 1). It is apparent that the reference simulation does not correctly capture a sharp temperature decrease halfway through the month (Julian day 166 in Melun and 168 in Oensingen) and keeps on overestimating the temperatures around noon by a few degrees during the rest of the month. This is substantially improved by the COT assimilation which picks up these temperature decreases on the respective days and the following period. The problems for the reference simulation are caused by a wrong location of a cold front and associated overlying cloud cover during these days, as will be shown later on. The assimilation procedure is thus capable of improving the cloud fields and yielding more accurate temperature values. Note that sometimes also the temperature during night time improves in the EXP simulation (e.g. on Julian day 177 in Melun) although the assimilation scheme is only active during day time (as satellite data from a visible channel are needed to derive COT). This is due to the transportation of the assimilated moisture throughout the model domain.

These findings are further demonstrated in Table 1, which shows the results for 4 of the observation stations throughout the model domain. The COT assimilation reduces the root-mean-square error (RMSE) between modelled and observed values for all stations. Also the correlation coefficients between the modelled and observed time series are higher for



**Fig. 4.** Incoming shortwave radiation at Grignon (upper panel) and Oensingen (lower panel) for June 2006.

the assimilation experiment. The positive bias for the stations around Paris that is present in the reference simulation is decreased. However, there is a slight negative effect of the assimilation on the biases in Oensingen en Bordeaux. Overall, the statistics of these time series are clearly improved by the assimilation.

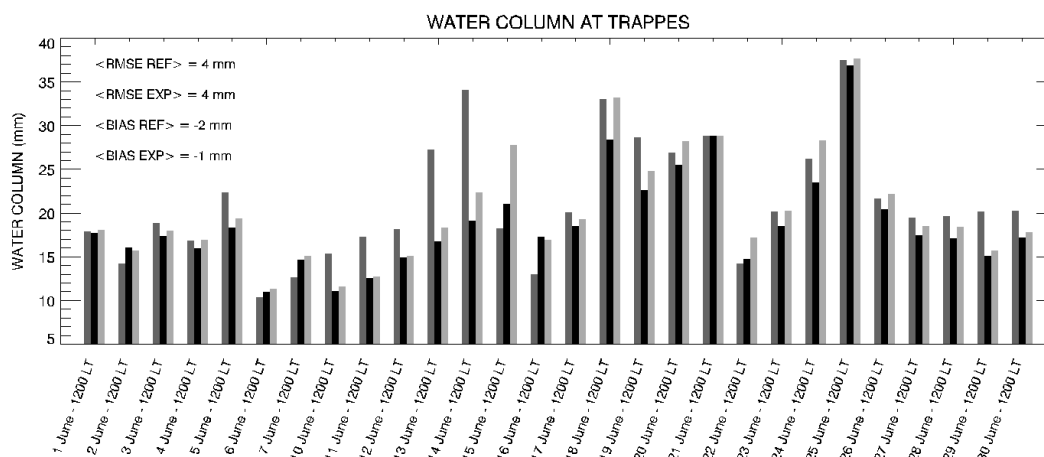
Another variable that is closely linked to the cloud fields, is the surface shortwave radiation. The results for the observation stations of Grignon, Fontainbleau and Oensingen are presented in Fig. 4 and Table 2. As for the temperature, the COT assimilation experiment clearly improves the time series, especially on the problematic Julian days 166 (Grignon) and 173 (Grignon and Oensingen). The statistics show a substantial reduction of the bias and RMSE and higher correlation coefficients. These numbers confirm that the improvement in the assimilation experiment is linked to a better position of the cloud cover in the model.

However, the impact of the COT assimilation is not positive for all variables, as shown in Table 3. The specific humidity at the surface is in good agreement with the observations for the reference simulation, whereas it is overestimated by about  $1 \text{ g kg}^{-1}$  for most observation stations when the assimilation scheme is applied. Only in Oensingen, the assimilation improves the underestimated humidity in the reference simulation. The extra moisture is caused by the fact that the assimilation scheme sets the humidity of a layer to its saturated value whenever a non-saturated layer becomes cloudy. As the reference simulation underestimates the amount of clouds compared to the observations, an increase of the hu-

**Table 1.** Statistics of the 2 m air temperature for the entire month of June 2006.

$T$ ( $^{\circ}\text{C}$ )	Mean	Bias	RMSE	$R^2$
<b>Melun</b>				
Obs	18.06	–	–	–
REF	18.59	0.53	2.64	0.85
EXP	18.01	–0.05	1.94	0.89
<b>Grignon</b>				
Obs	17.18	–	–	–
REF	18.52	1.34	2.95	0.84
EXP	17.84	0.66	2.14	0.87
<b>Oensingen</b>				
Obs	15.83	–	–	–
REF	16.13	0.30	3.42	0.79
EXP	15.28	–0.55	2.84	0.80
<b>Bordeaux</b>				
Obs	21.27	–	–	–
REF	21.11	–0.16	3.09	0.84
EXP	20.53	–0.74	2.52	0.88

midity is the logical consequence in this case. This can not be avoided if we want to retain the assimilated clouds, otherwise they would evaporate instantly. In their cloud assimilation experiments, Benedetti and Janiskova (2008) also



**Fig. 5.** The water column at Trappes from radio-sounding measurements (dark grey bars), the Reference simulation (black bars) and the COT assimilation experiment (light grey bars).

**Table 2.** Statistics of the incoming shortwave radiation for the entire month of June 2006.

Rs ( $\text{W m}^{-2}$ )	Mean	Bias	RMSE	$R^2$
<b>Grignon</b>				
Obs	360.30	–	–	–
REF	416.12	55.82	156.35	0.77
EXP	400.72	40.42	124.02	0.84
<b>Fontainbleau</b>				
Obs	381.90	–	–	–
REF	430.69	48.79	157.24	0.77
EXP	410.46	28.56	119.35	0.85
<b>Oensingen</b>				
Obs	311.84	–	–	–
REF	360.75	48.91	168.43	0.77
EXP	344.00	22.16	140.24	0.83

noticed that the humidity was affected in a slightly negative way by the assimilation.

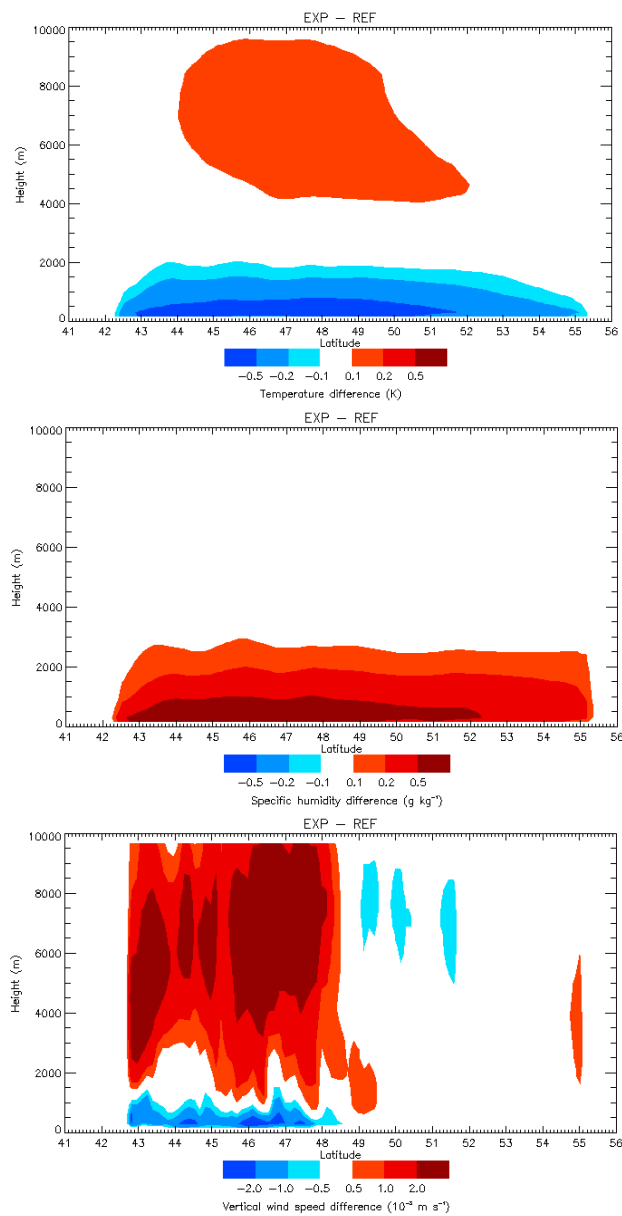
In order to evaluate the effect of the assimilation on the vertical distributed moisture, the total water columns from radio-soundings launched at Trappes are compared to the reference and assimilation experiments (Fig. 5). Overall, there is a good agreement between model results and observations, as the mean RMSEs of both the REF and EXP simulations are only 4 mm. The assimilation scheme has only a small impact on the total water column and slightly improves the small negative model bias. So the problems with the humidity that are mentioned before are limited and do not translate in drastic changes in the total vertical moisture amounts.

**Table 3.** Statistics of the 2 m specific humidity for the entire month of June 2006.

$q$ ( $\text{g kg}^{-1}$ )	Mean	Bias	RMSE	$R^2$
<b>Melun</b>				
Obs	8.67	–	–	–
REF	8.70	0.03	1.35	0.56
EXP	9.62	0.95	1.89	0.50
<b>Grignon</b>				
Obs	8.67	–	–	–
REF	8.64	–0.03	1.26	0.58
EXP	9.44	0.77	1.66	0.54
<b>Oensingen</b>				
Obs	9.68	–	–	–
REF	8.67	–1.01	2.13	0.67
EXP	9.41	–0.27	1.76	0.71
<b>Bordeaux</b>				
Obs	8.96	–	–	–
REF	8.82	–0.14	1.26	0.69
EXP	9.85	0.89	1.62	0.66

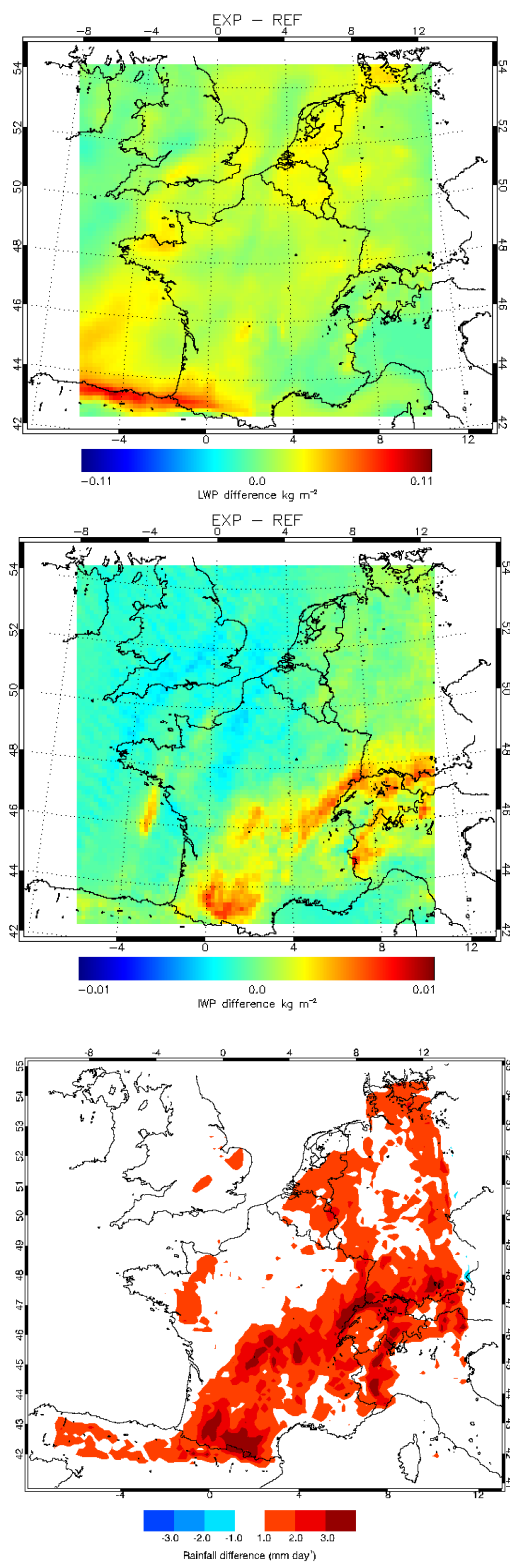
#### 4.2 Impact on temperature, humidity and cloud parameters

To assess the impact of the assimilation procedure on the model variables in the entire model domain, mean zonal differences between the experiment and the reference for temperature, specific humidity and vertical wind speed are shown in Fig. 6. The assimilation clearly has the largest effect in the lowest 2000 m of the model domain for both



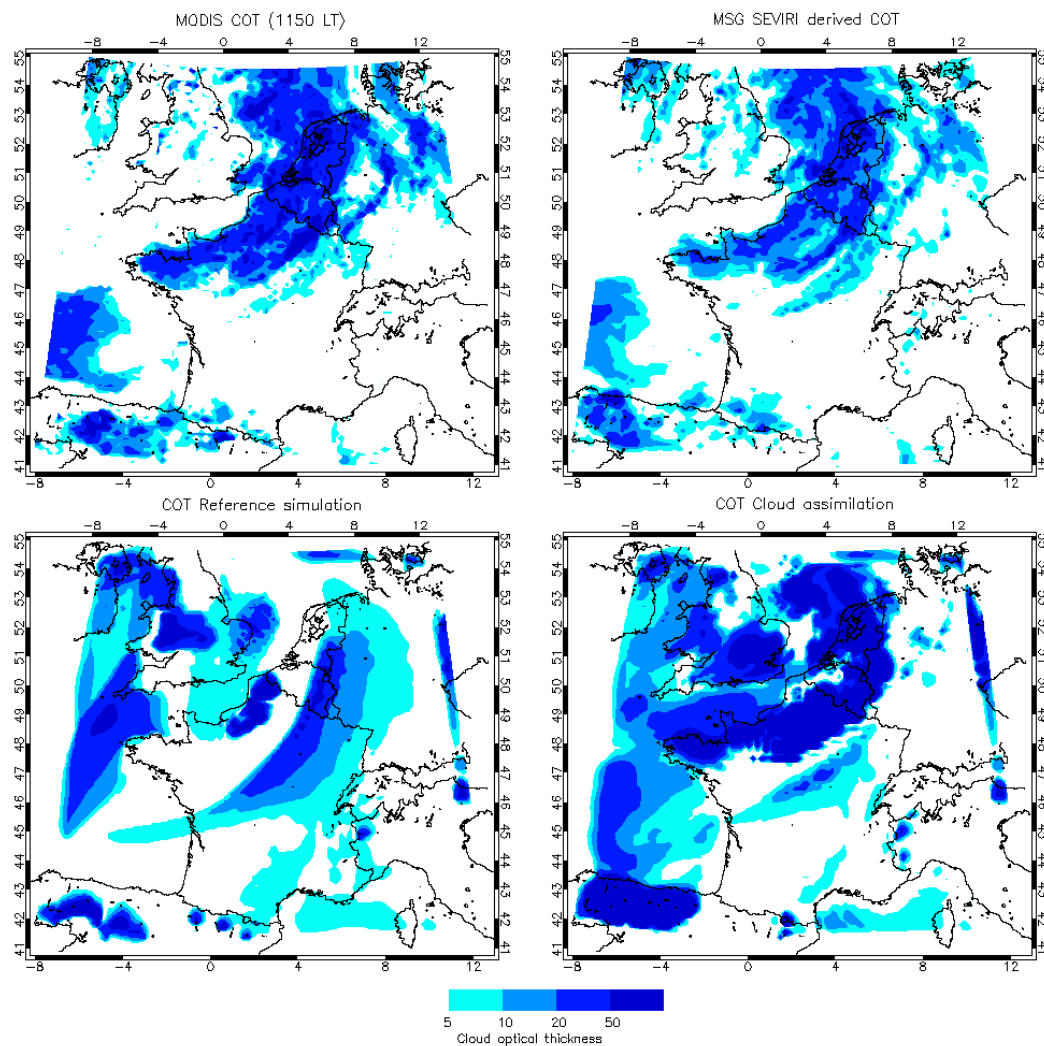
**Fig. 6.** Monthly mean differences between the COT assimilation experiment and the Reference simulation, averaged per latitude band.

the temperature and specific humidity. Here, the temperature values of the experiment have a tendency to be lower (up to 0.5 K) while the specific humidity is augmented (up to  $0.5 \text{ g kg}^{-1}$ ). Both temperature and moisture changes are in phase to enhance cloud formation. In the upper levels, a slight temperature increase is visible for the assimilation experiment, which is caused by latent heat release during the formation of additional clouds. As shown in the previous section, the temperature decrease near the surface improves the positive bias in the reference simulation, while the moisture increase has a negative effect when compared to the observations.



**Fig. 7.** Monthly mean differences in liquid water path (upper panel), ice water path (middle panel) and rainfall amounts (lower panel) between the COT assimilation experiment and the Reference simulation.





**Fig. 8.** Cloud optical thickness maps on 15 June 2006 at 12:00 LT.

Regarding the vertical wind speed, the largest changes clearly occur in the southern part of the model domain where most extra water is added to the model (see Fig. 7). The enhanced convection and latent heat release causes more updrafts in the higher model layers and downdrafts close to the surface. It should be noted that these changes are rather small (less than 10%) in comparison to the mean vertical wind speeds. As a response to the changes induced by the assimilation scheme, there is a noticeable redistribution in liquid water path and ice water path in the model domain (Fig. 7). The changes appear to have a rather varied structure over the largest part of the model domain. Most increases occur along the southern boundary of the domain and over the alpine region. Overall, there is a clear tendency of increased cloud amounts in the assimilation experiment. The monthly mean values of the liquid and ice water paths are raised by 25 and 8% respectively. Regarding the overestimation of incoming shortwave radiation by the reference simulation (Table 2),

these changes appear to have a positive impact on the model results. The assimilation also has an impact on the modelled rainfall amounts with a logical positive trend in the regions where more cloud ice is produced (last panel of Fig. 7). The overall effect is relatively large as the monthly mean rainfall is increased by 26%, although it should be noted that this is a dry month where eventual changes have a strong effect on the overall statistics.

The direct impact of the COT assimilation on the modelled cloud fields is presented in Fig. 8. In this figure the position of the clouds is shown on 15 July at noon, when a cold front passes Paris which is not picked up in the reference simulation (Fig. 3). The top two panels of the figure show a comparison between the cloud optical thickness product of the Moderate Resolution Imaging Spectroradiometer (MODIS) and our product from MSG Seviri. The positions of the clouds in both images are clearly very similar, although the mean COT value of MODIS is 18% higher than the MSG Seviri value.

**Table 4.** Statistics for Melun for all the sensitivity experiments.

	Obs	REF	EXP	$a = 0.1$	$a = 0.5$	$b = 0.1$	$b = 0.5$
$T$ ( $^{\circ}\text{C}$ )	18.06	18.59	18.01	18.10	16.19	16.57	18.03
RMSE	–	2.64	1.94	2.44	3.04	2.51	1.97
$R^2$	–	0.85	0.89	0.86	0.82	0.86	0.88
$q$ ( $\text{g kg}^{-1}$ )	8.67	8.70	9.62	9.18	10.30	10.13	9.61
RMSE	–	1.35	1.89	1.60	2.78	2.65	1.88
$R^2$	–	0.56	0.50	0.51	0.41	0.44	0.50

For the reference simulation, the clouds are positioned too far to the east and there is no strong front structure visible. The COT assimilation scheme is able to alter the cloud structure significantly and brings it much closer to the observations over the central region of the model domain. Although the scheme is not able to get rid entirely of the overestimation of clouds in the west side of the model domain, it is clearly able to improve the model simulation.

Finally, the sensitivity of the assimilation scheme to the coefficients  $a$  and  $b$ , related to the standard deviation of the simulated total water content and the observed cloud optical thickness, respectively, is tested. The results of this sensitivity experiment on the statistics for Melun are presented in Table 4. The values of these parameters have an effect on the model results and the scheme seems a little more sensitive to  $\sigma_{\tau}$ . Considering the results of both the temperature and humidity at the surface for this station, the initial choice of the values for  $a$  and  $b$  seems accurate as none of the sensitivity experiments can improve on these results.

## 5 Discussion and conclusions

In this paper, a technique has been presented to assimilate cloud optical thickness information into a mesoscale atmospheric model to yield improved diagnoses of surface solar radiation and temperature. The technique comprises an optimal estimation of cloud liquid and ice water in 1-D vertical columns together with a latent heat adjustment. The scheme requires some assumptions including an independent pixel approximation, but it is rather simple and computationally inexpensive, especially when compared to the 4DVAR systems that are currently developed (e.g. Benedetti and Janisková, 2008). The goal of the assimilation scheme is to improve retrospective model simulations by feeding the model with observed cloud optical thickness images every 15 min.

Results for the month of June 2006 show a positive impact of the assimilation on near-surface temperatures and incoming shortwave radiation, two variables that are closely linked to the overlying cloud cover and are crucial as input in, for instance, air pollution models. However, comparison to specific humidity observations show that the changes induced

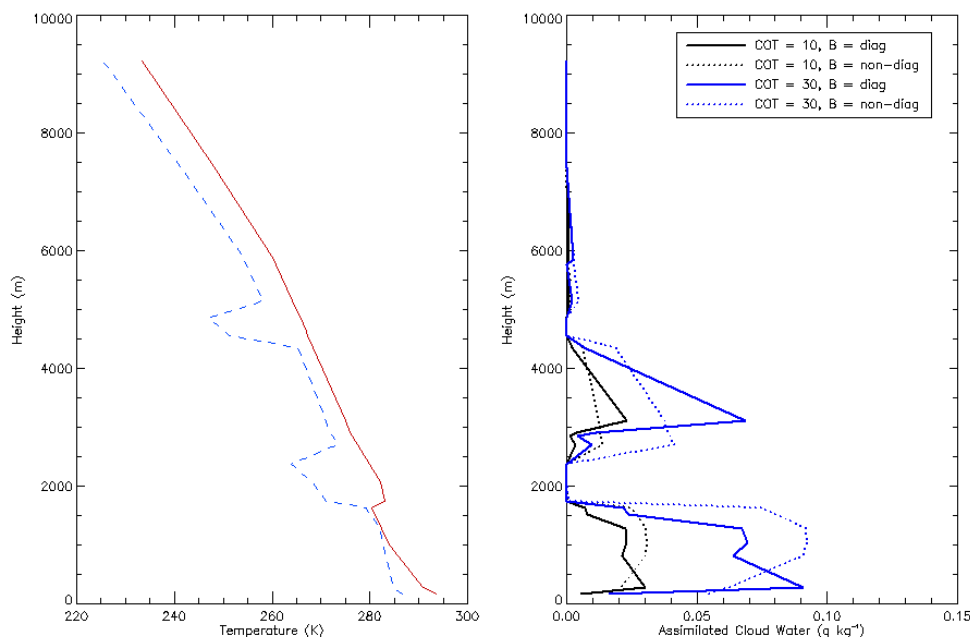
by the assimilation do not always improve the model fit to the observations. The assimilation scheme tends to induce overestimations of humidity close to the surface due to the fact that a layer is set to saturation when it becomes cloudy. This is necessary to retain the new clouds in the model and the same technique is used in the cloud analysis scheme of Soutu et al. (2003). Although the moisture field in the lowest 2000 m of the model domain is affected in a slightly negative way, the results show that the position of the cloud fields are more accurately simulated when the cloud observations are assimilated.

We can thus conclude that it is feasible to introduce the presented COT assimilation procedure in a mesoscale atmospheric model. The results show that the benefits of the assimilation to the surface temperature and radiation fields outweigh eventual inconsistencies that are caused by only adjusting the moisture and temperature fields of the model. As the procedure is simple and fast, it is a promising new technique to improve the quality of surface level model output of retrospective simulations.

## Appendix A

### List of symbols

$a$	coefficient in relation to $\sigma_{\tau}$ (–)
$b$	coefficient in relation to $\sigma_{\tau}$ (–)
$\mathbf{B}$	background error covariance matrix
$C_p$	specific heat of air at constant pressure ( $\text{J kg}^{-1} \text{K}^{-1}$ )
erfc	complementary error function
$\mathbf{H}$	observation operator matrix
$H$	heaviside step function
$\mathbf{H}^T$	transposed observation operator matrix
$\mathbf{K}$	gain matrix
$L_f$	latent heat of fusion ( $\text{J kg}^{-1}$ )
$L_v$	latent heat of vaporization ( $\text{J kg}^{-1}$ )
$q$	specific humidity ( $\text{kg kg}^{-1}$ )
$q_c$	specific cloud liquid water content ( $\text{kg kg}^{-1}$ )
$q_{ca}$	analyzed cloud water content ( $\text{kg kg}^{-1}$ )
$q_{cb}$	simulated (background) cloud water content ( $\text{kg kg}^{-1}$ )
$q_i$	specific cloud ice water content ( $\text{kg kg}^{-1}$ )
$q_s$	saturated specific humidity ( $\text{kg kg}^{-1}$ )
$q_{sb}$	simulated (background) saturated specific humidity ( $\text{kg kg}^{-1}$ )



**Fig. B1.** Left: radio-sounding profile of temperature (red) and dew point temperature (blue) at Trappes on 20 June at 12:00 LT. Right: resulting cloud water profiles from experiments with a non-diagonal  $\mathbf{B}$  matrix.

$q_t$	total water content ( $\text{kg kg}^{-1}$ )
$q_{tb}$	simulated (background) total water content ( $\text{kg kg}^{-1}$ )
$r$	decorrelation distance (m)
$r_e$	effective radius of cloud droplets (m)
$\mathbf{R}$	observation error covariance matrix
$R_s$	incoming shortwave radiation ( $\text{W m}^{-2}$ )
$T$	air temperature (K)
$z$	height (m)
$\delta$	Kronecker delta
$\rho$	air density ( $\text{kg m}^{-3}$ )
$\rho_w$	density of liquid water ( $\text{kg m}^{-3}$ )
$\sigma_c$	standard deviation of the background cloud water content
$\sigma_s$	standard deviation of the simulated saturated specific humidity
$\sigma_t$	standard deviation of the simulated total water content
$\sigma_\tau$	standard deviation of the observed cloud optical thickness
$\tau_0$	observed cloud optical thickness (-)
$\tau_b$	simulated (background) cloud optical thickness (-)

## Appendix B

### Non-diagonal $\mathbf{B}$ matrix

To obtain the formulas in Sect. 3.2, we have assumed that the background error covariance matrix is diagonal, which ignores the correlation of errors of cloud water content between different vertical layers. This may have an impact on the resulting cloud water profiles of the analysis, especially if the layers are close to each other in comparison to the typical thickness of a cloud layer. In order to have an estimate of the effect of this assumption, a test is performed with a non-diagonal  $\mathbf{B}$  matrix. In our test, a simple exponential decay

of the correlation of the errors is applied to the off-diagonal elements of  $\mathbf{B}$ , which gives for row  $i$  and column  $j$ :

$$\mathbf{B}_{ij} = \sigma_{ci} \sigma_{cj} e^{-\frac{|z_i - z_j|}{r}} \quad (\text{B1})$$

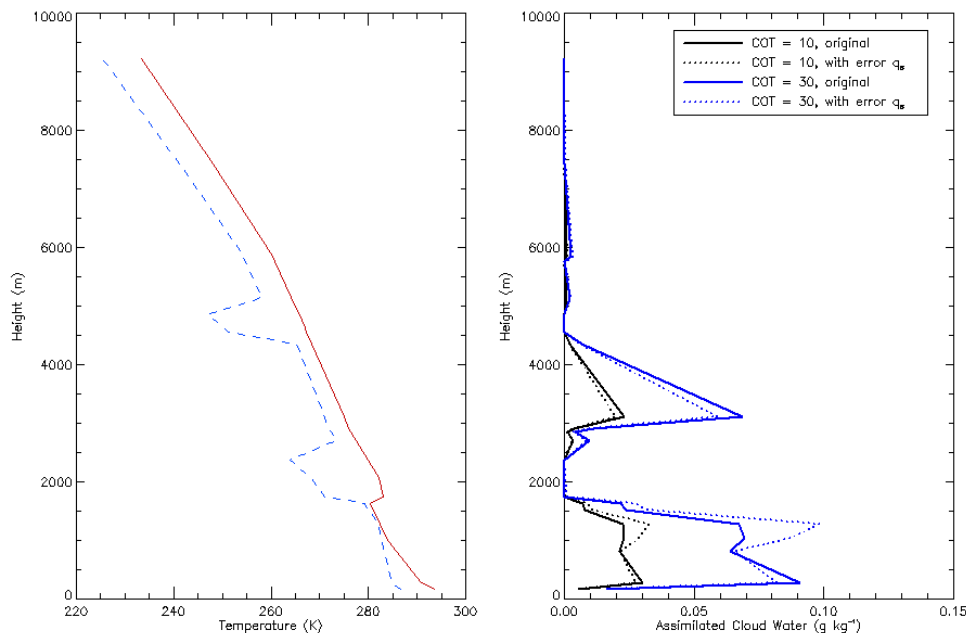
with  $r$  the vertical decorrelation length.  $r$  is given a value of 2000 m, corresponding to the value used by Županski et al. (2005) for the vertical decorrelation length of total water mixing ratio in their study on vertical correlations in error covariance modelling.

The impact of the non-diagonal  $\mathbf{B}$  matrix on the analyzed cloud water contents for a radio-sounding profile at Trappes is presented in Fig. B1. The non-diagonal matrix has some effect on the assimilated cloud water profile and mainly spreads out the cloud water peaks. However, the total amounts and locations of the assimilated cloud water content are comparable between both experiments. Therefore, we can conclude that the assumption of a diagonal background error covariance matrix will not have a strong impact on the assimilated cloud water profiles in this study.

## Appendix C

### The uncertainty on the background saturated specific humidity

In the formulation of the error variance of simulated cloud water ( $\sigma_c$ ) in Sect. 3.3, the error on the background saturated specific humidity (which depends on the error in background



**Fig. C1.** Left: radio-sounding profile of temperature (red) and dew point temperature (blue) at Trappes on 20 June at 12:00 LT. Right: resulting cloud water profiles from experiments including the uncertainty in the background saturated specific humidity.

temperature) is assumed to be zero to simplify the assimilation scheme. To assess the implications of ignoring the error on  $q_s$ , an alternative (and more complex) formulation is derived here, which takes into account this error. The error variance of simulated cloud water is now calculated as follows:

$$\sigma_c^2 = \langle q_c^2 \rangle - \langle q_c \rangle^2, \quad (\text{C1})$$

with, for  $n = 1, 2$ ,

$$\langle q_c^n \rangle = \frac{1}{2\pi\sigma_s\sigma_t} \int_{-\infty}^{+\infty} \int_{q_s}^{+\infty} (q_t - q_s)^n e^{-\frac{(q_t - q_{tb})^2}{2\sigma_t^2}} e^{-\frac{(q_s - q_{sb})^2}{2\sigma_s^2}} dq_t dq_s \quad (\text{C2})$$

In this expression,  $q_{tb}$  and  $q_{sb}$  denote the simulated (background) values of the quantities  $q_t$  (total specific water content) and  $q_s$  (saturated specific humidity content), which have assumed known standard deviations  $\sigma_t$  and  $\sigma_s$ , respectively, and which are taken here to be stochastically independent.

We first proceed to a transformation of variables by  $x = \frac{(q_t - q_s)}{\sqrt{2}\sigma_t}$ , and also introduce  $x_s = \frac{(q_{tb} - q_s)}{\sqrt{2}\sigma_t}$  and  $x_{sb} = \frac{(q_{sb} - q_s)}{\sqrt{2}\sigma_t}$ . Moreover, introducing  $\alpha = \sigma_t/\sigma_s$ , and changing the order of integration, the integral can be written as:

$$\begin{aligned} \langle q_c^n \rangle &= \frac{(\sqrt{2}\sigma_t)^{n+2}}{2\pi\sigma_s\sigma_t} \int_0^{+\infty} x^n \left[ \int_{-\infty}^{+\infty} e^{-(x-x_s)^2} e^{-\alpha^2(x_s+x_{sb})^2} dx_s \right] dx \\ &= \frac{(\sqrt{2}\sigma_t)^{n+2}}{2\sqrt{\pi}\sigma_s\sigma_t} \frac{1}{\sqrt{1+\alpha^2}} \int_0^{+\infty} x^n e^{-\frac{\alpha(x+x_{sb})^2}{\sqrt{1+\alpha^2}}} dx \end{aligned} \quad (\text{C3})$$

where we made use of one of the definite integrals listed in Gradshteyn and Ryzhik (2007). After some further algebra,

again using integrals listed in Gradshteyn and Ryzhik (2007), and introducing  $\sigma_{ts}^2 \equiv \sigma_t^2 + \sigma_s^2$ , it is then fairly straightforward to work out the above integrals, which yields:

$$\begin{aligned} \langle q_c \rangle &= \sigma_{ts} \left[ \frac{1}{\sqrt{2\pi}} e^{-x_{sb}^2} - \frac{x_{sb}}{\sqrt{2}} \operatorname{erfc}(x_{sb}) \right] \\ \langle q_c^2 \rangle &= \sigma_{ts}^2 \left[ \frac{1}{\sqrt{\pi}} x_{sb} e^{-x_{sb}^2} - \frac{1}{2} (1 + 2x_{sb}^2) \operatorname{erfc}(x_{sb}) \right] \end{aligned} \quad (\text{C4})$$

The impact of this new formulation, which takes into account the error of  $q_s$ , on the analyzed cloud water contents for a radio-sounding profile at Trappes is presented in Fig. C1. For this experiment,  $\sigma_s$  is given a value of 0.2, which corresponds to an uncertainty on the background air temperature of about 3 K. The effect on the assimilated cloud water profile is rather small as only some minor shifts can be noticed. The total amounts and locations of the assimilated cloud water content are very similar between both experiments. Therefore, we can conclude that the assumptions that are made in Sect. 3.3 are not likely to have a significant impact on the results of this study.

*Acknowledgements.* This research has been performed in the framework of the Climate and Air Quality Modelling for Policy Support (Climaqs) project, sponsored by the Flemish agency for Innovation by Science and Technology, and the FP-7 Myair PASODOBLE project, sponsored by the European Commission's Seventh Framework Programme. We would like to thank the following CARBOEUROPE PIs for providing access to their data: Jean-Yves Pontailier (FR-FON), Nicolas Mascher (FR-GRI) and Cristof Ammann et al. (SWI-OEN). We also acknowledge the National Climatic Data Center (NCDC) for providing valuable

validation data. The ARPS meteorological model was made available by the Center for Analysis and Prediction of Storms at the University of Oklahoma. We would also like to acknowledge the European Environment Agency for making available the CORINE land-cover data, the US Geological Survey for the GTOPO30 dataset, the Jet Propulsion Laboratory for the SST imagery, NASA's Earth Science Division for the GLDAS data, and our VITO colleagues at TAP for the SPOT-VEGETATION NDVI imagery. Meteorological forcing data were obtained from the European Centre for Medium-Range Weather Forecasting (ECMWF).

Edited by: A. Geer

## References

- Albers, S. C., McGinley, J. A., Birkenheuer, D. A., and Smart, J. R.: The Local Analysis and Prediction System (LAPS): Analysis of clouds, precipitation and temperature, *Weather Forecast.*, 11, 273–287, 1996.
- Benedetti, A. and Janisková, M.: Assimilation of MODIS Cloud Optical Depths in the ECMWF Model, *Atmos. Rev.*, 136, 1727–1747, 2008.
- Benedetti, A., Stephens, G. L., and Vukov, T.: Variational assimilation of radar reflectivities in a cirrus model. I: Model description and adjoint sensitivity studies, *Q. J. Roy. Meteor. Soc.*, 129, 27–300, 2003.
- Brewster, K.: Application of a Bratseth analysis scheme including Doppler radar data. Preprints, 15th Conf. Wea. Anal. Forecast. Amer. Meteor. Soc., Norfolk, VA, 92–95, 1996.
- DeRidder, K. and Schayes, G.: The IAGL land surface model, *J. Appl. Meteorol.*, 36, 167–182, 1997.
- Gradshteyn, I. S. and Ryzhik, I. M.: *Table of Integrals, Series and Products*, Seventh Edition, Academic Press, 1161 pp., 2007.
- Hu, M., Xue, M., and Brewster, K.: 3D-VAR and cloud analysis with WSR-88D level-II data for the prediction of the Fort Worth, Texas, tornadic thunderstorms. Part I: Cloud analysis and its impact, *Mon. Weather Rev.*, 134, 67–698, 2006a.
- Hu, M., Xue, M., and Brewster, K.: 3D-VAR and cloud analysis with WSR-88D level-II data for the prediction of the, Fort Worth, Texas, tornadic thunderstorms. Part II: Impact of radial velocity analysis via 3D-VAR, *Mon. Weather Rev.*, 134, 69–721, 2006b.
- Kain, J. S. and Fritsch, J. M.: A One-Dimensional Entraining Detraining Plume Model and Its Application in Convective Parameterization, *J. Atmos. Sci.*, 47, 2784–2802, 1990.
- Kalnay, E.: *Atmospheric modeling, data assimilation and predictability*, Cambridge Univ. Press, 341 pp., 2003.
- Kokhanovsky, A. A., Rozanov, V. V., Zege, E. P., Bovensmann, H., and Burrows, J. P.: A semi-analytical cloud retrieval algorithm using backscattered radiation in 0.4–2.  $\mu\text{m}$  spectral region, *J. Geophys. Res.*, 108, 4008, doi:10.1029/2001JD001543, 2003.
- Lin, Y.-L., Farley, R. D., and Orville, H. D.: Bulk parameterization of the snow field in a cloud model, *J. Climate Appl. Meteor.*, 22, 1065–1092, 1983.
- Pandey, P., De Ridder, K., and van Lipzig, N. P. M.: Estimating scaled cloud optical thickness from SEVIRI by implementing a semi-analytical cloud retrieval algorithm, *Atmos. Res.*, in review, 2011.
- Salby, M. L.: *Fundamentals of Atmospheric Physics*. Academic Press, San Diego, CA, 627 pp., 1996.
- Soutu, M. J., Balseiro, C. F., Pérez-Munuzuri, V., Xue, M., and Brewster, K.: Impact of Cloud Analysis on Numerical Weather Prediction in the Galician Region of Spain, *J. Appl. Meteorol.*, 42, 129–140, 2003.
- Sun, W. Y. and Chang, C. Z.: Diffusion-Model for a Convective Layer .1. Numerical-Simulation of Convective Boundary-Layer, *J. Clim. Appl. Meteorol.*, 25, 1445–1453, 1986.
- Sun, J. and Crook, N. J.: Dynamical and microphysical retrieval from Doppler radar observations using a cloud model and its adjoint. Part II: Retrieval experiments of an observed Florida convective storm, *J. Atmos. Sci.*, 55, 835–852, 1998.
- Vukićević, T., Greenwald, T., Županski, M., Županski, D., Vonder Haar, T., and Jones, A. S.: Mesoscale cloud state estimation from visible and infrared satellite radiances, *Mon. Weather Rev.*, 132, 3066–3077, 2004.
- Xue, M.: High-order monotonic numerical diffusion and smoothing, *Mon. Weather Rev.*, 128, 2853–2864, 2000.
- Xue, M., Droegemeier, K. K., and Wong, V.: The Advanced Regional Prediction System (ARPS) – A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification, *Meteorol. Atmos. Phys.*, 75, 161–193, 2000.
- Xue, M., Droegemeier, K. K., Wong, V., Shapiro, A., Brewster, K., Carr, F., Weber, D., Liu, Y., and Wang, D.: The Advanced Regional Prediction System (ARPS) – A multi-scale nonhydrostatic atmospheric simulation and prediction tool. part II: Model physics and applications, *Meteorol. Atmos. Phys.*, 76, 143–165, 2001.
- Yucel, I., Shuttleworth, W. J., Gao, X., and Sorooshian, S.: Short-term Performance of MM5 with Cloud-Cover Assimilation from Satellite Observations, *Mon. Weather Rev.*, 131, 1797–1810, 2003.
- Županski, M., Županski, D., Vukićević, T., Eis, K., and Vonder Haar, T.: CIRA/CSU Four-Dimensional Variational Data Assimilation System, *Mon. Weather Rev.*, 133, 829–843, 2005.