

Assimilation of SSMIS humidity-sounding channels in all-sky conditions over land using a dynamic emissivity retrieval

F. Baordo,^{a,b*} and A. J. Geer^a

^aResearch Department, European Centre for Medium-Range Weather Forecasts, Reading, UK

^bForecast Systems, Weatherzone, Sydney, New South Wales, Australia

*Correspondence to: F. Baordo, Weatherzone, 8 West Street, Sydney, 2060, NSW, Australia. E-mail: fabriziobaordo@hotmail.com

The extension of all-sky assimilation of Special Sensor Microwave Imager/Sounder (SSMIS) humidity-sounding channels to land surfaces is investigated in this article. The scattering index, which is able to discriminate between cloudy and precipitating regions over land, is used as a predictor to develop a ‘symmetric’ model for observation error. This formulation is able to increase the observation error in those scenes that are more difficult to model because of radiative transfer and ‘mislocation’ errors. **The implementation of an instantaneous emissivity retrieval from SSMIS observations is also presented.** In clear-sky scenes, emissivity retrievals appear better at capturing daily differences in surface conditions compared with emissivity atlas values. In the presence of clouds, retrievals show different behaviour. In the lower microwave frequencies (less than 50 GHz), emissivity estimates appear nearly as reliable as those in clear skies, **but at higher frequencies, as the magnitude of scattering increases, so does the error in the retrieval and the resultant emissivity estimate can be unphysically low or high.** However, the retrieval still appears feasible at high frequencies in light cloud situations; the number of retrievals discarded due to this kind of problem is around 10%. **In these cases, an estimate from an emissivity atlas can be substituted instead.** Together, the **new observation-error model** and the **instantaneous emissivity retrievals** were adopted for the assimilation of SSMIS 183 GHz channels over land in all-sky conditions. Assimilation experiments showed that the assimilation system is not degraded and the improvements in analysis and forecast scores are about the same as those obtained by the equivalent clear-sky approach. The developments described in this study were an essential first step to create framework to allow the all-sky assimilation over land of other microwave humidity sounders: this started operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF) in 2015, covering both SSMIS and four Microwave Humidity Sounder (MHS) instruments.

Key Words: surface emissivity; cloud and precipitation; microwave; data assimilation

Received 2 August 2015; Revised 15 June 2016; Accepted 23 June 2016; Published online in Wiley Online Library 15 August 2016

1. Introduction

The assimilation of microwave observations in numerical weather prediction (NWP) systems is still more intensive over ocean than over land surfaces. Over ocean, fast and accurate emissivity models have been developed for NWP (e.g. Liu *et al.*, 2011; Kazumori and English, 2014) that allow the assimilation of channels with strong surface sensitivity. **In contrast, the microwave signal emerging from land surfaces not only depends on frequency, incidence angle and polarization, but is also affected by the large variability in surface types (e.g. deserts, vegetation, high orography) and conditions (e.g. roughness, moisture, snow, ice).** For this reason, in data assimilation, the high complexity of modelling the interaction between all these surface parameters and the microwave radiation has generally restricted the use of

observations to temperature and humidity-sounding channels, which receive a weak contribution from the surface. However, an estimate of surface emissivity is still required to assimilate such data. Increasingly, in NWP, the approach is to make retrievals of land emissivities directly from satellite measurements (e.g. Karbou *et al.*, 2005; Prigent *et al.*, 2005; Ruston *et al.*, 2008, among many others). The general assumption is that, for most surface types, **land emissivity is sufficiently invariant with frequency that emissivity retrieved from surface-sensitive (‘window’) channels can be used as a reasonable approximation for sounding channels.** This approach has also been implemented for some years within the operational European Centre for Medium-Range Weather Forecasts (ECMWF) system to assimilate clear-sky observations in temperature and humidity-sounding channels (Krzeminski *et al.*, 2009) from the Advanced Microwave Sounding

Unit-A (AMSU-A) and Microwave Humidity Sounder (MHS): emissivities retrieved from AMSU-A channel 3 (50.3 GHz) are given to the temperature channels (50–60 GHz), whilst retrievals from MHS channel 1 (89 GHz) are assigned to the humidity channels (183 GHz). Karbou *et al.* (2010a, 2010b) investigated the use of land emissivity retrievals within the Météo-France assimilation system, showing that it helps to reduce bias and standard deviation of first-guess (FG) departures (e.g. observation minus forecast differences) and increases the number of assimilated observations from AMSU-A and AMSU-B surface-sensitive and sounding channels. Improved correlations between observations and simulations were found over snow-covered areas and analyses and forecasts over tropical regions appeared to be improved.

A further increase in the use of sounding observations over land can be achieved by extending the assimilation to clear, cloudy or precipitating scenes (the so called ‘all-sky’ approach). However, at microwave frequencies in situations where the atmospheric scattering is most important (such as over land and in temperature and humidity sounding channels), the inaccuracy of scattering radiative transfer models has been a problem (e.g. Geer *et al.*, 2012; Baordo *et al.*, 2012). Geer and Baordo (2014a) were able to improve the accuracy of the microwave scattering signal globally across all frequencies from 10–183 GHz in all weather conditions by modelling snow as a non-spherical hydrometeor and using optical properties from the Liu (2008) discrete dipole approximation for a sector snowflake. These developments allowed ECMWF to start operational all-sky assimilation of Special Sensor Microwave Imager/Sounder (SSMIS) humidity-sounding channels over ocean (Geer, 2013). The benefits of all-sky assimilation on forecast quality are roughly twice that of assimilating in clear-sky alone. **The benefit comes through the 4D-Var assimilation, which can infer dynamical initial conditions from humidity, cloud and precipitation features in the observations.** This has also been demonstrated in the assimilation of clear-sky infrared humidity observations by Peubey and McNally (2009). The benefit of SSMIS humidity channels is greatest in the southern midlatitudes, where the storm tracks provide ideal conditions for model-tracing. The good results over the ocean strongly encouraged feasibility studies to extend the assimilation of SSMIS humidity-sounding channels to land surfaces.

Our purpose in this article is to **document the technical and scientific changes that were needed to implement the all-sky framework over land, summarizing and extending the initial work carried out by Baordo *et al.* (2012, 2013).** The over-land framework has been developed and evaluated through the use of SSMIS observations, but it can be generalized for the assimilation of other humidity-sounder sensors. A general overview of the all-sky assimilation at ECMWF is provided in section 2. The methodology to implement the over-land framework is described in section 3. Results of assimilation experiments are provided in section 4.

2. General overview of the all-sky assimilation

2.1. Radiative transfer model

ECMWF has been assimilating microwave imager observations operationally in all-sky conditions for over 5 years (Bauer *et al.*, 2010; Geer *et al.*, 2010; Geer and Bauer, 2011). The observation operator designed for assimilating microwave radiances in clear, cloudy and precipitating scenes is RTTOV-SCATT (Bauer *et al.*, 2006), which uses the delta-Eddington approximation (Joseph *et al.*, 1976) to solve the radiative transfer equation including scattering. The **bulk optical properties** for cloud water, cloud ice and rain are pre-tabulated for each hydrometeor type as a function of **temperature, frequency and water content**. Cloud water, cloud ice and rain are modelled as **spherical particles** using Mie theory and a constant density: the first two hydrometeors

use a **gamma size distribution** (e.g. Petty and Huang, 2011), while a **Marshall and Palmer** (1948) size distribution is used for rain. Since the ECMWF operational cycle 40r1, snow has been modelled as a non-spherical hydrometeor, which uses optical properties calculated from the discrete dipole approximation for a sector snowflake (Liu, 2008). To simulate bulk optical properties for the Liu sector shape, the tropical version of the Field (2007) size distribution is used.

Ocean surface emissivity is computed by version 5 of FASTEM (Liu *et al.*, 2011; Bormann *et al.*, 2012). Over land, as will be documented in section 3, we adopt the same strategy as implemented in previous studies (Krzeminski *et al.*, 2009; Karbou *et al.*, 2010a, 2010b): first we retrieve land emissivities from satellite observations in surface-sensitive channels and second, assuming that the spectral variability of emissivity is minimal, we reassign these estimates to the closest channel higher in frequency in order to perform radiative transfer calculations (for instance, emissivities retrieved in the SSMIS 91 GHz channel are applied to the humidity-sounding channels). If the retrieval fails, we use emissivities provided by the Tool to Estimate Land Surface Emissivities at Microwave (TELSEM) atlas, which is based on a pre-calculated monthly-mean emissivity climatology derived from 10 years of SSMI observations (Aires *et al.*, 2011). TELSEM atlas values are available within RTTOV and they provide emissivity estimates for all land surfaces between 19 and 100 GHz and for all angles and linear polarizations.

The observation operator in the all-sky framework selects the **nearest model profile** to the observation (in time and space) and then runs RTTOV-SCATT. The hydrometeor inputs to RTTOV-SCATT are the vertical profiles of cloud water, cloud ice, total rain and total snow, plus the effective cloud fraction (C), which controls the computation of the all-sky simulated observation ($T_{\text{all-sky}}$), which is a weighted average of brightness temperature from two independent subcolumns, one clear (T_{clr}) and one cloudy (T_{cld}). The effective cloud fraction provides the weight given to the two subcolumns:

$$T_{\text{all-sky}} = (1 - C)T_{\text{clr}} + CT_{\text{cld}}. \quad (1)$$

Over ocean surfaces, C is computed as a hydrometeor-weighted average of cloud, convective and large-scale precipitation fractions across all vertical levels, providing an approximate but computationally efficient solution to account for the effects of subgrid variability in cloud and precipitation (Geer *et al.*, 2009a, 2009b, the ‘ C_{av} ’ approach).

Over land surfaces, C is computed as the largest cloud fraction in the model profile (the ‘ C_{max} ’ approach.) This is essentially a tuning measure to compensate for a relative lack of deep convection over land in the model, compared with over the ocean. The choice of the Liu sector snowflake to model scattering from snow hydrometeors was made based on achieving the best fit between modelled and observed brightness temperatures over the ocean (Geer and Baordo, 2014a). However it was difficult to find one strategy that gave good results over land and ocean simultaneously. However, there was good consistency between observations and first-guess departures when the use of the ‘ C_{max} ’ approach was adopted over land. Instead of having different hydrometeor models for land and ocean, technically it was easier to vary the cloud fraction. This is the strategy we implemented in the ECMWF system.

2.2. Observations

In order to investigate the all-sky assimilation over land of 183 GHz humidity-sounding channels, this study uses observations from SSMIS. In line with the ECMWF operational usage, only Defence Meteorological Satellite Program satellite (DMSP) F-17 has been used. The SSMIS sensor (Kunkee *et al.*, 2008) represents an important advancement over its predecessor, SSMI, as it combines the SSMI imaging capabilities with the profiling capabilities of microwave sounders. SSMIS

Table 1. Properties of SSMIS channels.

Channel/name	Frequency (GHz)	Polarization	Sensitivity
1 / 50H	50.3	H	Surface
2 / –	52.8	H	Temperature
3 / –	53.596	H	Temperature
4 / –	54.4	H	Temperature
5 / –	55.5	H	Temperature
6 / –	57.29	RC	Temperature
7 / –	59.4	RC	Temperature
8 / 150H	150.0	H	Humidity
9 / 183 ± 7	183.31 ± 6.6	H	Humidity
10 / 183 ± 3	183.31 ± 3.0	H	Humidity
11 / 183 ± 1	183.31 ± 1.0	H	Humidity
12 / 19H	19.35	H	Surface
13 / 19V	19.35	V	Surface
14 / –	22.235	V	Surface
15 / 37H	37.0	H	Surface
16 / 37V	37.0	V	Surface
17 / 91V	91.65	V	Surface
18 / 91H	91.65	H	Surface
19 / –	63.28 ± 0.285	RC	Temperature
20 / –	$\nu = 60.79 \pm 0.358$	RC	Temperature
21 / –	$\nu \pm 0.002$	RC	Temperature
22 / –	$\nu \pm 0.005$	RC	Temperature
23 / –	$\nu \pm 0.016$	RC	Temperature
24 / –	$\nu \pm 0.050$	RC	Temperature

allows microwave measurements at frequencies ranging from 19–183 GHz over a swath width of 1707 km. SSMIS channels can be grouped as follows: 13 channels (channels 1–7 and 19–24) located near the oxygen absorption band (50–60 GHz), which allow atmospheric temperature sensing from about 80 km down to the Earth's surface; three humidity-sounding channels (channels 9–11) close to the strong 183 GHz water-vapour line; and seven SSMI-like imaging channels (channels 12–18), with 91.65 GHz replacing the SSMI 85.5 GHz channels and the addition of 1 channel at 150 GHz (channel 8). Table 1 summarizes the general characteristics of SSMIS channels.

The SSMIS F-17 data are pre-processed along the lines of Bell *et al.* (2008), who described the necessary corrections for F-16 observations. However, instrument improvements implemented for F-17 make the pre-processing slightly different: corrections are made for scan non-uniformity and reflector emission (based directly on an accurate thermistor-measured reflector temperature) but not for warm load intrusions, which are infrequent. In the all-sky framework, SSMIS observations are also averaged (or 'superobbed') in boxes of approximately 80 km by 80 km in order to make the horizontal scales of observed cloud and precipitation more similar to their effective resolution in the model (Geer and Bauer, 2010).

In this study, bias corrections are applied to remove systematic differences between the satellite observations and the model first guess. Bias coefficients are derived adaptively within the analysis system using variational bias correction (Auligné *et al.*, 2007). For the SSMIS 150 and 183 GHz channels, the bias predictors are a constant offset, the layer thickness across four vertical bands in the atmosphere, a fourth-order polynomial in instrument scan position and a predictor based on the solar hour to remove biases between ascending and descending parts of the orbit. Surface-sensitive channels (such as 91 GHz) have additional predictors including surface wind speed, skin temperature and total column water vapour. The same bias model is applied for clear-sky and all-sky assimilation; no cloud or precipitation-related biases are modelled. The bias correction addresses mainly instrument issues and biases at the broadest global scales.

3. All-sky framework over land

3.1. Observation error modelling

In the all-sky assimilation, the accuracy of the radiative transfer calculations in cloud and precipitating areas is not the sole source

of error. Forecast models have difficulties in predicting cloud and precipitation in exactly the right place with the right intensity (e.g. Fabry and Sun, 2010). These model errors are sometimes broadly described as 'mislocation' and (along with radiative transfer error and representivity error) lead to apparently highly non-Gaussian behaviour of first-guess departures, independent of the surface type (land or ocean). Geer and Bauer (2011) showed that, over ocean surfaces, these errors could be treated successfully as observation error when performing all-sky data assimilation and they could be represented by a symmetric error model that inflates observation error as a function of the mean of the cloud amount in the observations and the forecast model. This approach also provides a threshold quality-control check that works in clear-sky and cloud- or precipitation-affected situations. In the operational ECMWF system, the observation error for all-sky microwave imager observations over ocean is a linear function of the symmetric cloud amount given by the average of the observed and simulated polarization difference at 37 GHz.

Over land, the 37 GHz polarization difference cannot be used as a cloud predictor for the symmetric error model, as it relies on the highly polarized nature of ocean surface emission. Instead, the greatest atmospheric signal from hydrometeors comes at higher frequencies from snow, so the scattering index (SI; e.g. Bennartz *et al.*, 2002, 2003) is a good parameter to identify precipitation and convective areas. A SI given by the difference between SSMIS channel 18 (91H) and channel 8 (150H) was found to give good performance as a symmetric predictor for the observation-error model (Baordo *et al.*, 2012). Bias correction (the usual VarBC correction) is applied to the simulated TBs to ensure consistency between observed and simulated SI.

To check the capability of the SI to identify clear-sky areas as well as regions affected by clouds and precipitation, we can compare the modelled SI with the total hydrometeor content in the model, given by the sum of the model FG cloud water, cloud ice, rain and snow. We can also distinguish between the liquid hydrometeor content (LHC), considering only cloud water and rain, and the ice hydrometeor content (IHC), taking into account only cloud ice and snow. Figure 1(b) shows that the model's mean hydrometeor content is approximately a linear function of the simulated SI. Cloud-free regions (zero hydrometeor content) are characterized by SI less than 0 K, while for increasing values of SI the hydrometeor content increases as well, representing the increasing influence of scattering as the modelled cloud and precipitation amount increase. The IHC and the total hydrometeor content are relatively similar, suggesting that the LHC adds little information to the main signal at these frequencies. There is not such an obvious link between the observed SI and the IHC (Figure 1(a)). This is the consequence of 'mislocation' errors. There are some cases where the observation is cloudy (higher scattering index), but there is no corresponding cloud in the model, counterbalanced by other cases where the observation is cloud-free (lower scattering index), but the model has cloud.

We can now derive the observation-error formulation from the symmetric scattering index SI_{sym} , which is the mean of the observed and the simulated SI. Figure 2 gives an example of how the standard deviations of SSMIS channel 9 FG departures have been binned as a function of SI_{sym} ; results for SSMIS channels 10 and 11 are similar in shape if not magnitude (not shown). For low scattering indexes, i.e. the clear-sky regime, the standard deviation is relatively constant at around 3 K. As the symmetric mean SI increases, so does the standard deviation. This increase is reasonably well-modelled by a straight line. However, beyond a certain threshold the errors seem to saturate. This behaviour is modelled by a piecewise linear fit that is estimated graphically and shown in the figure. Channels 9–11 have a minimum clear-sky standard deviation E_{min} of 3 K and a maximum E_{max} that identifies strong scattering situations and decreases for higher-peaking channels. For channels 9–11, E_{max} varies as follows: 57, 35 and 16 K. The start and end points of the up slope are labelled SI_{clr} (estimated as 1, 2 and 3 K respectively for channels 9–11) and

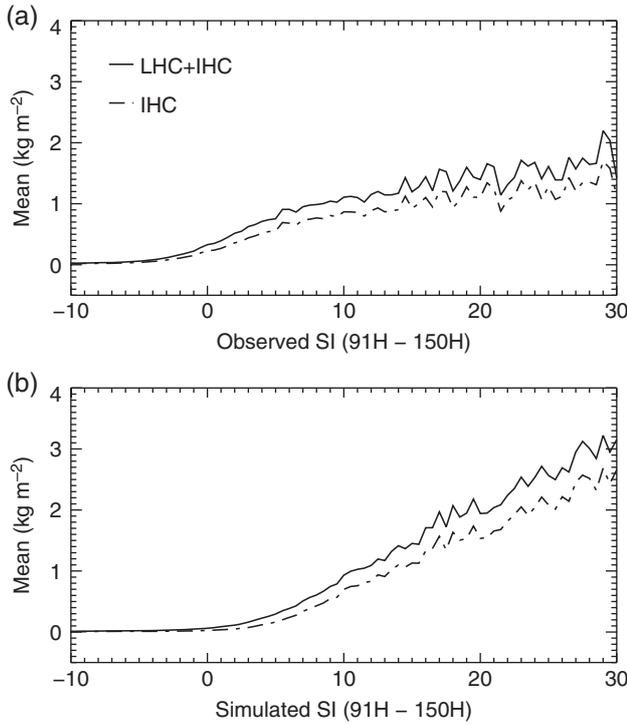


Figure 1. Mean of total (solid line) and ice (dash-dotted line) hydrometeor content (kg m^{-2}) binned as a function of (a) observed and (b) simulated scattering index. The liquid hydrometeor content (LHC) is given by the sum of model FG cloud water and rain, while the ice hydrometeor content (IHC) takes into account cloud ice and snow. The sample of SSMIS data is for the month of June 2013, considering observations over land restricted to latitudes equatorward of 60° . The bin size is 0.5 K.

SI_{cld} (estimated as 25 K for all three channels), with the standard deviation of FG departures T_{err} being predicted as follows:

$$T_{\text{err}} = E_{\text{min}} \in SI_{\text{sym}} \leq SI_{\text{clr}}; \quad (2)$$

$$T_{\text{err}} = E_{\text{min}} + (E_{\text{max}} - E_{\text{min}}) \frac{SI_{\text{sym}} - SI_{\text{clr}}}{SI_{\text{cld}} - SI_{\text{clr}}} \in SI_{\text{clr}} < SI_{\text{sym}} < SI_{\text{cld}}; \quad (3)$$

$$T_{\text{err}} = E_{\text{max}} \in SI_{\text{sym}} \geq SI_{\text{cld}}. \quad (4)$$

The error model for T_{err} predicts the total error, i.e. the combination of background error and observation error. To estimate the observation error from T_{err} , we subtract (in quadrature) a rough estimate of the background error in radiance space, which is 1 K for all channels. For instance, in clear-sky scenes (e.g. $T_{\text{err}} = T_{\text{min}} = 3$ K), the observation error for the humidity-sounding channels is 2.8 K. This is a little larger than the constant observation error of 2 K that has been used to assimilate MHS humidity-sounding observations over both land and ocean in clear-sky conditions. The slightly larger observation error reflects the standard deviation in situations where SI is close to 0 K; it is a cautious approach to assimilating all-sky SSMIS data over land, even when nominally in clear skies. The effectiveness of the observation error is discussed later in section 4.2.

3.2. Surface emissivity retrieval

For a non-scattering plane-parallel atmosphere, assuming a flat and specular surface, for a given zenith (θ) angle and frequency (ν), the brightness temperature (T_b) observed by a satellite sensor can be expressed as follows:

$$T_b = \epsilon_{(\theta,\nu)} T_s \Gamma_{(\theta,\nu)} + (1 - \epsilon_{(\theta,\nu)}) T_{(\theta,\nu)}^\downarrow \Gamma_{(\theta,\nu)} + T_{(\theta,\nu)}^\uparrow, \quad (5)$$

where $\epsilon_{(\theta,\nu)}$ represents the surface emissivity at observation zenith angle θ and frequency ν ; T_s , $T_{(\theta,\nu)}^\downarrow$ and $T_{(\theta,\nu)}^\uparrow$ are, respectively, the

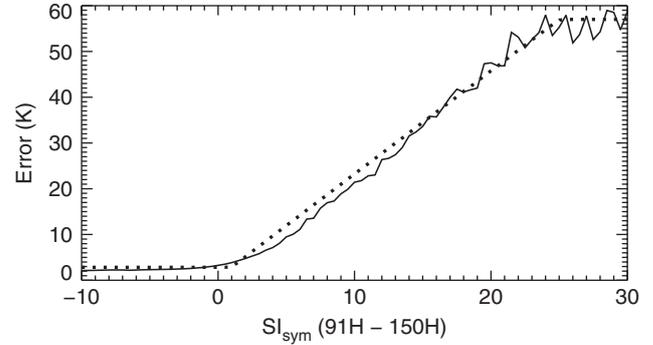


Figure 2. Error model for SSMIS channel 9 (183 ± 7 GHz), showing how the standard deviations of FG departures binned as a function of the symmetric scattering index (solid line) are modelled by a linear fit (dotted line). FG departures are for the month of June 2013, considering SSMIS observations over land restricted to latitudes equatorward of 60° . The bin size is 0.5 K.

surface skin temperature (i.e. the effective radiating temperature of the surface at the relevant frequency), the atmospheric downwelling radiation at the surface and upwelling radiation at the top of the atmosphere; $\Gamma_{(\theta,\nu)}$ is the net atmospheric transmissivity. The emissivity can be retrieved from Eq. (5) (for simplicity we omit the dependence on zenith angle and frequency):

$$\epsilon = \frac{T_b - (T^\uparrow + T^\downarrow)\Gamma}{(T_s - T^\downarrow)\Gamma}. \quad (6)$$

Equation (6) represents the scheme that has been commonly used in the literature to retrieve emissivity directly from satellite measurements for those channels that receive a strong contribution from the surface (e.g. Karbou *et al.*, 2005; Prigent *et al.*, 2005). In the case of NWP systems, the atmospheric contribution to the observed brightness temperature (T^\downarrow , T^\uparrow and Γ) is computed within the radiative transfer model using the atmospheric profiles from short-range forecasts as input. The surface skin temperature, like the emissivity, is also affected by the land surface variability (e.g. soil moisture, roughness, wetness, snow), so that it is a difficult parameter to estimate accurately (English, 2008). Skin temperature can be also retrieved from satellite measurements (e.g. Karbou *et al.*, 2010a) and used within NWP systems. In our study, we did not investigate the impact of using different sources of skin temperature (e.g. retrieved or monthly-mean based on climatology), but following the existing clear-sky approach at ECMWF: we relied on the estimate provided by the short-range forecasts.

In order to be consistent with the way the observations are used in the all-sky assimilation, we take account of two independent columns, one clear and one cloudy, which, weighted by the effective cloud fraction C , give the simulated brightness temperature (Eq. (1)). Assuming that skin temperature and emissivity are the same in each subcolumn, from Eqs (1) and (5), we have

$$\epsilon = \frac{T_b - (1 - C)(T_{\text{clr}}^\uparrow + T_{\text{clr}}^\downarrow \Gamma_{\text{clr}}) - C(T_{\text{cld}}^\uparrow + T_{\text{cld}}^\downarrow \Gamma_{\text{cld}})}{(1 - C)(T_s - T_{\text{clr}}^\downarrow) \Gamma_{\text{clr}} + C(T_s - T_{\text{cld}}^\downarrow) \Gamma_{\text{cld}}}. \quad (7)$$

The atmospheric terms are simulated for the clear and cloudy subcolumns (T_{clr}^\uparrow , $T_{\text{clr}}^\downarrow$, Γ_{clr} and T_{cld}^\uparrow , $T_{\text{cld}}^\downarrow$, Γ_{cld}). When the contribution of the cloudy column is zero ($C = 0$), the emissivity computation is reduced to the emissivity scheme of Eq. (6).

Independently of the way the emissivity retrieval is performed (Eq. (7) or (6)), it is not free from error. The larger the error in the model skin temperature or atmospheric profile, the larger the uncertainty in the emissivity estimate. The assumption of a specular surface can also generate errors in those cases where such an approximation might be no longer valid (e.g. in presence of very rough terrain and surface scattering). Finally, where there is

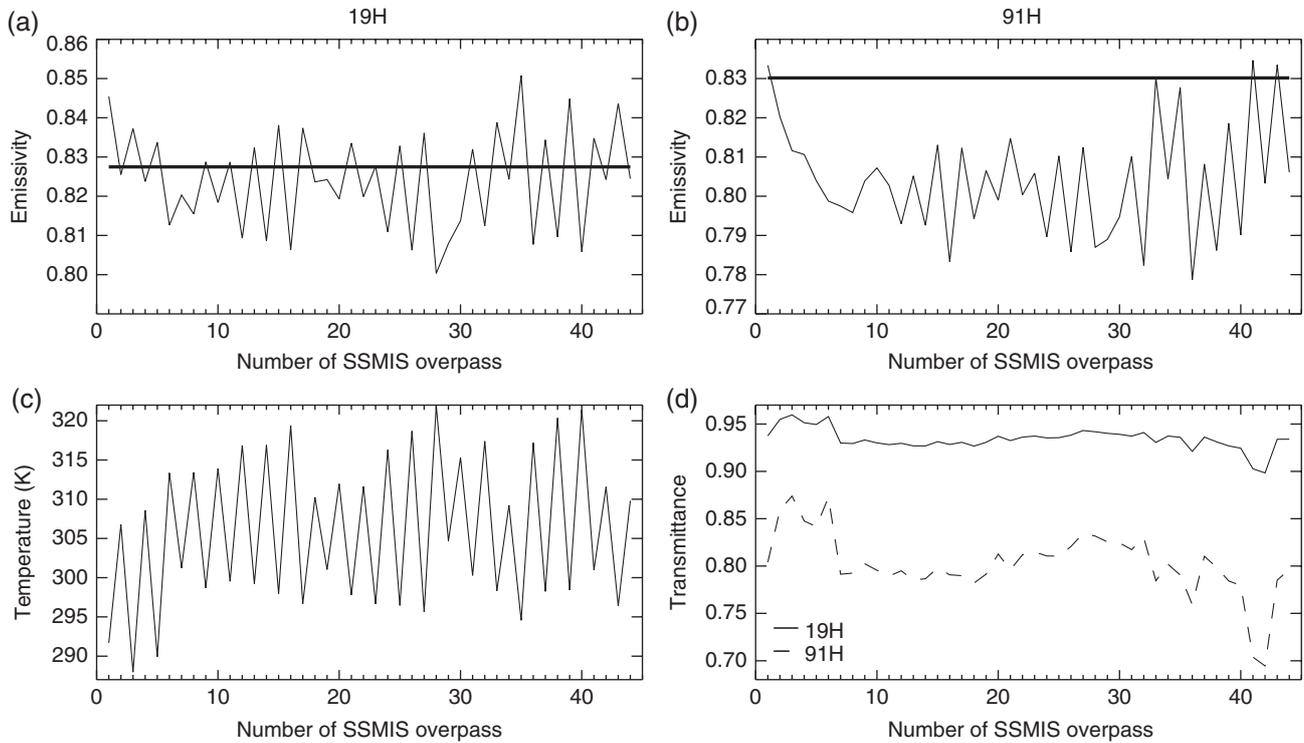


Figure 3. All-sky emissivity retrievals derived by SSMIS measurements in Northern Africa (28.42°N ; 4.50°E) from 1–30 June 2013 (for a total of 44 overpasses): for (a) 19 GHz horizontal polarization and (b) 91 GHz horizontal polarization. The black thick line represents the TELSEM atlas value. To complete the figure, for each SSMIS overpass we show (c) model skin temperature and (d) model transmittance at 19 and 91 GHz.

strong atmospheric scattering or high liquid water content, the increasing opacity of the atmosphere can amplify the magnitude of the error, to the point where the surface becomes invisible and an emissivity retrieval is clearly impossible. Because the all-sky assimilation does not implement cloud screening and the observed brightness temperatures can be from clear, cloudy or precipitating scenes, this can affect our retrievals; later sections will examine the errors caused by cloud and precipitation and ways of dealing with them. Essentially, we default to an emissivity atlas value where the emissivity retrieval appears to be erroneous.

3.3. Emissivity retrieval assessment

The technical implementation of the emissivity retrieval in the all-sky framework was first evaluated by Baordo *et al.* (2012), but here we will examine the impact of cloud in more detail and look at the benefit of using instantaneous emissivity estimates rather than values provided by climatology. This is organized as follows. In section 3.3.1, we look at emissivity retrievals obtained using cloud-free observations. Secondly, in section 3.3.2, we examine the impact of cloud on the retrievals. Finally, in section 3.3.3, we describe the completed scheme for generating emissivities to be used for the 183 GHz channels.

To compare the clear-sky and all-sky emissivity retrievals outlined in the previous section, we ran two assimilation experiments (further details are provided in section 4). Both process SSMIS observations through the all-sky path of the ECMWF system, but the first experiment implements Eq. (7), the second, forcing $C = 0$, retrieves emissivity through Eq. (6). Even though the two experiments have slightly different backgrounds (e.g. different skin temperature and atmospheric profiles coming from the natural chaotic variability that occurs when two non-identical assimilation experiments are performed), it is still reasonable to compare the two calculations of surface emissivity, as they use identical SSMIS observations.

Retrievals are performed globally, but two screening criteria are applied. Firstly, retrievals are restricted to latitudes equatorward of 60° , particularly to avoid high-latitude areas where, during winter, snow and ice cover might introduce additional errors in the emissivity retrieval. Secondly, SSMIS observations are rejected

in those locations where the model grid point contains a mixture of water and land. Retrievals are computed only if the model land–sea mask is greater than 0.95.

In the following sections, we also refer to Eqs (6) and (7), respectively, as the clear-sky (ϵ_{cs}) and all-sky (ϵ_{as}) emissivity retrieval.

3.3.1. Retrieval in cloud-free scenes

As an example of clear-sky retrievals, we chose a bare soil point in North Africa. The location is in Algeria (28.42°N ; 4.50°E) and it was observed by 44 SSMIS overpasses for the period ranging from 1 (0500 UTC) to 30 (1700 UCT) June 2013, shown in Figure 3. The observations are most likely free from clouds and precipitation during the entire month: the SSMIS observed and simulated SI are always less than 0; also, images from SEVERI show no cloud systems over the selected location. The modelled atmospheric transmittance is always high (Figure 3(d)), so the retrievals should be reliable. The retrieved emissivities are shown for channels 19H and 91H, alongside the values from the TELSEM atlas (Figure 3(a) and (b)). Because the model cloud fraction is always nearly zero at this location, there is little difference between retrievals from the all-sky and clear-sky approaches. Retrieved emissivity is similar to the atlas, but shows variability from one overpass to the next. Particularly at 19H, it appears to be anticorrelated with the diurnal variability of skin temperature (the correlation coefficient is -0.90). The skin temperature varies around 15 K between the times of the early evening and early morning orbits (Figure 3(c)). In dry desert soils, the microwave emission may come from some distance below the surface. The anticorrelation is consistent with an excessively strong diurnal cycle of skin temperature in the model, perhaps because the diurnal cycle of subsoil temperature is damped compared with the skin temperature (the forecast model skin temperature is a key part of radiative energy transfer at the surface, so it is more representative of infrared and visible wavelengths, where it is a genuine ‘skin’ temperature). Hence one benefit of performing an emissivity retrieval, rather than using an atlas, is that it reduces the impact of errors in the use of model skin temperature, which is particularly evident in desert regions. Although it would be nice

Table 2. RMSE of differences between simulated and SSMIS observed brightness temperatures in Northern Africa (28.42°N; 4.50°E) from 1–30 June 2013 (for a total of 44 overpasses).

	RMSE [$T_{b_{as}} - T_b$]	RMSE [$T_{b_{cs}} - T_b$]
183 ± 7	2.836	3.071
183 ± 3	1.925	1.922
183 ± 1	2.086	2.157

$T_{b_{as}}$ and $T_{b_{cs}}$ indicate the simulations computed using ϵ_{as} and ϵ_{cs} , respectively.

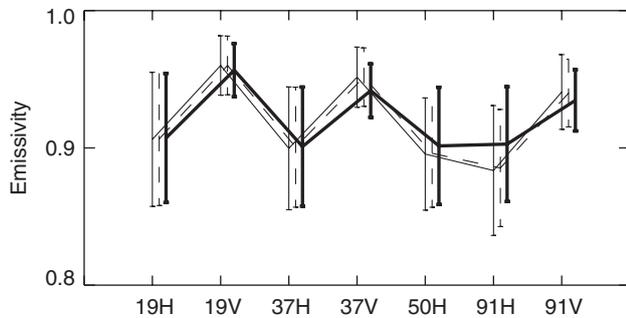


Figure 4. Mean of land-surface emissivity retrievals in SSMIS window channels for the month of June 2013. ϵ_{as} (thin line) and ϵ_{cs} (dashed line) retrieved from SSMIS cloud-free observations are compared with TELSEM atlas values (thick line). Retrievals are restricted to latitudes equatorward of 60° and model land–sea mask values greater than 0.95. Error bars represent the standard deviation.

to estimate a skin temperature simultaneously with the emissivity, that is not feasible (see Eq. (6)). As a general consideration, the variability of our retrievals with respect to TELSEM estimates might also be partially driven by the superobbed observations: the exact location of raw SSMIS observations composing the superob is going to vary from day to day and so we might see a slightly different bit of surface each time with a slightly different emissivity. Further, the 91H emissivity retrievals in Figure 3 are lower than the equivalent atlas value. This is observed more generally and might be explained by considering that TELSEM estimates were derived from SSMI data at 85.5 GHz and that there may be some change in the emissivity with frequency. So, even in a desert environment where surface emissivity is normally quite constant with time, there are some advantages of using emissivity retrievals rather than an atlas: compensating for skin-temperature errors, matching the emissivity of the exact field of view and other sensor characteristics (such as polarization) and reducing the impact of frequency extrapolation.

The root-mean-square error (RMSE) of differences between the simulated and observed brightness temperatures can provide a general measure of the error in the radiative transfer calculations. Table 2, for the Algeria case, provides the RMSE computed considering the 44 SSMIS observations in the 183 GHz humidity-sounding channels, which are assigned the emissivity retrieved at 91 GHz. We distinguish between RMSE computed using ϵ_{as} to simulate the brightness temperatures ($T_{b_{as}}$) and that coming from the use of ϵ_{cs} ($T_{b_{cs}}$), but results are very similar for both. The error is largest at 183 ± 7 GHz in SSMIS channel 9. This is a lower tropospheric humidity channel and, consequently, in the relatively clear atmosphere of this example it is likely to be more affected by errors in the surface parameters (e.g. emissivity, skin temperature). In any case, the residual uncertainty in the first guess is consistent with the observation-error model, which in this scenario would predict a total error of 3 K. All 44 SSMIS humidity-sounding observations in this example are actively assimilated and are not rejected by quality control.

To generalize the behaviour of the retrieval for SSMIS window channels in cloud-free scenes (observed SI ≤ 0 K), Figure 4 compares global means and standard deviations of emissivities from ϵ_{as} , ϵ_{cs} and TELSEM for the month of June 2013. As mentioned, statistics are computed between 60°S and 60°N,

only considering land–sea mask values greater than 0.95. The sample has been restricted slightly further by (a) only considering retrievals that have generated emissivities between 0 and 1 and (b) requiring the availability of TELSEM atlas values. The all-sky and clear-sky retrievals behave the same across all frequencies and match the climatology values fairly well. Statistics at 91H are those particularly relevant for the goal of our assimilation and they show the negative bias that was highlighted in the single observation case (Figure 3).

3.3.2. Retrieval in cloud-affected scenes

At higher microwave frequencies, the presence of clouds and precipitation can reduce the surface-to-space transmittance substantially, leading to larger errors in the emissivity estimate. However, in the case of limited scattering and reasonably high surface-to-space transmittance, an instantaneous emissivity retrieval might still be possible. To illustrate the impact of cloud contamination on the emissivity retrievals, we chose a densely vegetated area in the northeast of Bolivia (15.09°S; 63.28°W). Compared with the desert area examined earlier, this area exhibits higher emissivities and is often affected by deep convection. Looking at the SSMIS observed and simulated SI and cross-checking with images from the *Geostationary Operational Environmental Satellite (GOES)*, we chose three scenes that are cloud contaminated but are characterized by different amounts of scattering. It is important to remember that the all-sky assimilation is affected by ‘mislocation’ errors, so that what is seen in the observations might or might not be present in the model with the same intensity. The three cases explore the situation when the observations have more cloud than the model, or about the same.

- (1) Intense scattering (30 June 2013 2200 UTC). Observations have much more cloud than the model: observed SI 34 K; simulated SI 7.5 K; model total hydrometeor content 0.47 kg m⁻²; model ice hydrometeor content 0.39 kg m⁻²; model TCWV 52.4 kg m⁻²; model rain rate 0.040 mm h⁻¹.
- (2) Intermediate scattering (24 June 2013 2200 UTC). Observations have a little bit more cloud than the model: observed SI 7 K; simulated SI 3.5 K; model total hydrometeor content 0.35 kg m⁻²; model ice hydrometeor content 0.15 kg m⁻²; model TCWV 51 kg m⁻²; model rain rate 0.052 mm h⁻¹.
- (3) Moderate scattering (25 June 2013 2100 UTC). Observations and model have the same amount of cloud: observed SI 1.2 K; simulated SI 1.2 K; model total hydrometeor content 0.26 kg m⁻²; model ice hydrometeor content 0.04 kg m⁻²; model TCWV 48.5 kg m⁻²; model rain rate 0.144 mm h⁻¹.

Note that, even in the ‘moderate scattering’ case, this is far from a clear-sky location: there is cloud and precipitation present. Table 3 shows, for each observation case, ϵ_{as} and ϵ_{cs} retrievals, model surface-to-space transmittance and observed brightness temperatures for the SSMIS window channels. The magnitude of the scattering signal can be clearly distinguished within the three cases: moving from 19 to 91 GHz, the higher the frequency, the larger the depression in the observation. In the intense scattering case, observed brightness temperatures at 37, 50 and 91 GHz are about 10, 20 and 30 K colder than those at 19 GHz. In the second case, the size of such differences is reduced by half and in the moderate scattering situation the impact of cloud is hard to identify. In general, the lower microwave frequencies are not affected by cloud contamination, but rather by large liquid water drops such as precipitation, and are less affected by ice or snow scattering. The emissivity retrievals appear to be consistent with those provided by the TELSEM atlas. However, at 50 GHz and above, the depressed brightness temperatures in the intermediate and intense scattering cases lead to unphysically low emissivity retrievals. For example,

Table 3. Impact of cloud contamination on ϵ_{as} and ϵ_{cs} for three observation cases over the northeast of Bolivia (15.09°S; 63.28°W).

ϵ_f	19H	19V	37H	37V	50H	91H	91V
	0.934	0.941	0.914	0.922	0.912	0.908	0.916
1. Intense scattering (observed SI 34 K)							
T_b (K)	287.11	286.11	277.73	278.25	268.44	253.11	254.43
Γ_{clr}	0.806	0.807	0.788	0.788	0.460	0.400	0.400
ϵ_{as}	0.938	0.953	0.893	0.919	0.751	0.193	0.279
ϵ_{cs}	0.937	0.952	0.894	0.919	0.764	0.356	0.423
2. Intermediate scattering (observed SI 7 K)							
T_b (K)	286.94	285.60	281.72	281.09	274.10	273.91	273.41
Γ_{clr}	0.809	0.811	0.790	0.790	0.462	0.406	0.406
ϵ_{as}	0.941	0.953	0.924	0.935	0.832	0.655	0.702
ϵ_{cs}	0.937	0.949	0.924	0.934	0.843	0.742	0.766
3. Moderate scattering (observed SI 1.2 K)							
T_b (K)	286.30	284.75	281.19	281.05	277.43	282.47	281.55
Γ_{clr}	0.816	0.818	0.797	0.797	0.468	0.423	0.423
ϵ_{as}	0.936	0.947	0.915	0.932	0.892	0.920	0.935
ϵ_{cs}	0.938	0.949	0.924	0.938	0.900	0.920	0.938

The observed brightness temperatures (T_b) and the model surface-to-space transmittance (Γ_{clr}) are also shown for the SSMIS window channels. Emissivities from the TELSEM atlas (ϵ_f) are provided as an additional reference.

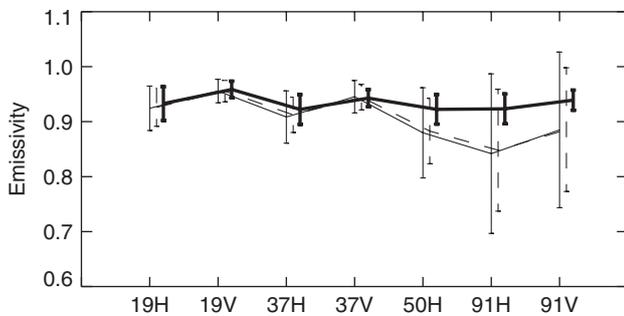


Figure 5. As Figure 4, but considering only SSMIS cloud-affected observations.

$\epsilon_{as} = 0.193$ in the 91H channel. This problem affects the clear-sky and all-sky emissivity retrievals equally. Even in the all-sky case, these are situations where the model has substantially less scattering than the observations and will struggle to generate a good retrieval. The method of emissivity retrieval assumes that the modelled atmospheric parameters that go into the retrieval are correct, but that assumption is violated in the intense and intermediate scattering cases. However, the moderate scattering case demonstrates that instantaneous emissivity retrievals are still feasible. In this particular case, the emissivity retrieval at 91 GHz is similar to the atlas values. The retrieved emissivity at 91 GHz can be given to the 183 GHz channels, leading first-guess departures for SSMIS channel 9, 10 and 11 that are 0.947, -0.801 and 1.812 K, respectively, indicating a good fit between model and observations. This shows that the retrieval is sufficient for our purpose of assimilating 183 GHz channels. Note that, in all three cases, at higher microwave frequencies the modelled clear-sky sensitivity to the surface is reduced (for instance, the value of the surface-to-space transmittance at 91H is roughly half that at 19H), but that does not seem itself to be the cause of problems in retrieving emissivity; rather it is the strong depression of observed brightness temperatures in the intense and moderate scattering cases.

Figure 5 generalizes the impact of cloud contamination on SSMIS emissivity retrievals by showing the global mean and standard deviation of emissivities from ϵ_{as} , ϵ_{cs} and TELSEM for the month of June 2013, selecting only those cases where observed SI > 0 K, indicating scattering. Retrievals at frequencies less than 50 GHz are consistent with the climatology estimates, but at 50 and 91 GHz, as a result of cloud contamination, means and standard deviations of ϵ_{as} and ϵ_{cs} often diverge from those computed using TELSEM values, generally becoming too low.

An issue unique to the all-sky approach is that the presence of cloud in the model profiles can also cause problems. This is the opposite problem to the one just illustrated; here the model has much more cloud than is present in the observation. In this situation in Eq. (7), the weight of the cloudy column in the emissivity calculation is very high (C close to or equal 1) and Γ_{clr} is relatively small (much smaller than Γ_{clr}). This leads to a small denominator in Eq. (7) and, if the observed brightness temperature is not consistent with the cloud and precipitation in the model, this generates unphysically high emissivities, often larger than 1. Once again, the number of these failures increases at higher frequencies, where scattering is the dominant effect and Γ_{clr} is much smaller than Γ_{clr} . Overall, this means that the all-sky method can generate more poor-quality retrievals in strongly scattering conditions than the clear-sky method, because there are roughly double the opportunities for problems: cases in which there is intense scattering in either the model or the observations. However, if we want to push the methodology into cloud- and precipitation-affected regimes, it still seems desirable to perform a retrieval that is consistent with the radiative transfer model that is being used. Clearly, we need to be able to identify those situations in which an emissivity retrieval is still feasible and to discard those retrievals affected by large errors.

3.3.3. Surface emissivity for the 183 GHz channels

It appears to be useful to perform an instantaneous emissivity retrieval in order to capture local changes in surface conditions that can affect the emissivity. We have chosen to use the all-sky emissivity retrieval, rather than the clear-sky one, for consistency with the way the simulations are performed and with a view to the future. However, to eliminate poor-quality retrievals, a number of screening stages have been applied. First, we decided to reject observations where the value of the symmetric scattering index is greater than 20 K. As a reference, for one month of SSMIS data, this quality control discards roughly 1.2% of observations. We are avoiding attempting to assimilate profiles where scattering is most intense, either in the model or the observation. Secondly, we chose to reject emissivity retrievals (but still potentially use the observations) where they are outside the range of expected physical behaviour, i.e. outside the range 0.55–1.0. Thirdly, we compared the retrieval with the emissivity atlas and rejected retrievals that were more than a certain threshold away from the atlas value. These thresholds have been estimated from the standard deviation of difference considering retrievals obtained only from cloud-free observations (e.g. observed SI ≤ 0 K as in Figure 4) and the corresponding atlas values, calculated globally for every frequency in summer and winter time (e.g. June and January). An average of the two seasons' standard deviations is used and can be labelled σ . The final threshold was chosen as roughly twice σ , in order to increase the chances of using an emissivity retrieval rather than the atlas. Thresholds used in every frequency are as follows: 0.04 (19H), 0.03 (19V), 0.04 (37H), 0.03 (37V), 0.06 (50H), 0.09 (91H), 0.07 (91V). The final sample of emissivities at 91H that has been used to simulate the brightness temperatures for the 183 GHz channels is most relevant to our assimilation system. Mean land emissivity maps at 91H for the month of June 2013 are shown in Figure 6, which compares the retrieved emissivities (but where the 10% failing quality checks have been replaced by TELSEM atlas values) with the estimates provided by the climatology. The statistics are based on a total of 405 577 SSMIS observations. The general offset between the retrievals and the atlas at 91 GHz is clearly visible. The mean map of emissivity difference (e.g. retrievals Figure 6(b) minus TELSEM Figure 6(a), not shown) highlights the fact that this offset is relatively featureless at around -0.02 over most land areas, excluding high orography. However, globally the retrieval does a good job in capturing the main features that characterize different surface types (e.g. desert, high orography, densely vegetated areas).

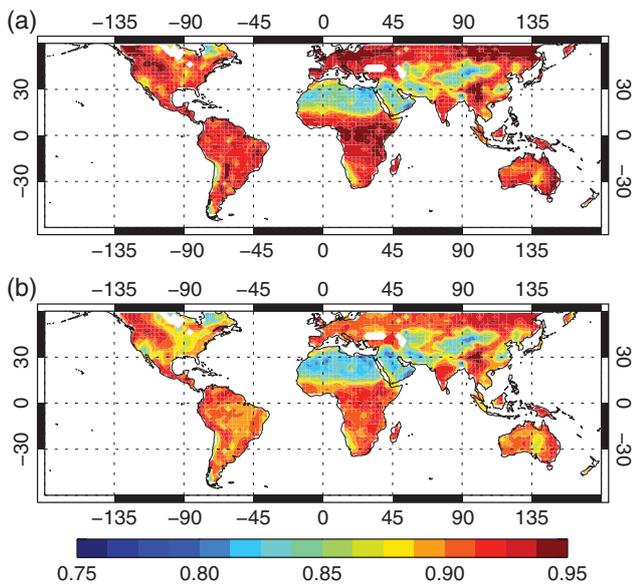


Figure 6. Mean land emissivity maps at 91H for the month of June 2013 for (a) corresponding TELSEM atlas values and (b) the mix of generated retrieval/atlas emissivities. The grid spacing is 2.5° .

Figure 7 gives a final illustration of the strengths and weaknesses of emissivity-retrieval methods in challenging conditions. It shows all-sky and clear-sky surface emissivity retrievals during the monsoon season over an area on the border between Bangladesh and India, where land usage is dominated by rice cultivation. The observed SI (Figure 7(a)) is regularly larger than 5 K, indicating frequent episodes of precipitation. The simulated SI is not shown; though it is often in reasonable agreement with the observed SI, it can become unreliable in the presence of strong convection, because to simulate 91 and 150 GHz brightness temperatures requires use of emissivity estimates (and we have chosen to use retrievals at 37 and 91 GHz, respectively) that themselves can become erroneous through the presence of scattering. The 19 GHz emissivity retrievals are largely unaffected by scattering and clear-sky and all-sky retrievals are generally very similar through the period shown in Figure 7(b). Retrieved emissivity peaks at around 0.9 in late June and declines rapidly to 0.7 in mid-July; we could speculate that this is associated with irrigation of the rice fields. The atlas emissivities show a similar pattern, while missing the shorter-term variability. There are a few cases, particularly in mid-August, where the all-sky retrieval generates unphysically low emissivity retrievals and the clear-sky retrieval is acceptable: these are likely to do with the presence of heavy precipitation in the model that is not seen in the observations.

At 91 GHz (Figure 7(c)), there is a larger scatter in the emissivity estimates from both clear-sky and all-sky cases; 2% of clear-sky and 27% of all-sky emissivity retrievals are outside the physical bounds (0–1) and are not shown on the figure. However, the majority of retrievals are around 0.75–0.9, similar to the 19 GHz retrievals and much higher than the atlas emissivity of 0.6–0.7. It is possible that the atlas itself is suffering from cloud contamination here. Figure 7(d) shows the estimated emissivities given to the all-sky simulations at 183 GHz: because the quality check is based on atlas values, the retrieved emissivity is almost always rejected and the atlas is used instead, even if it likely has problems in this location. However, the emissivity estimate is almost immaterial for the 183 GHz assimilation here: throughout the period, the clear-sky surface-to-space transmittance in the 183 ± 7 GHz channel is around 10^{-8} . Overall, this example illustrates that atlas, clear-sky and all-sky retrievals can all have their problems but, at least for 183 GHz assimilation, the worst cases for emissivity retrieval are those where either the surface visibility is irrelevant anyway or strongly convective situations will be given a large observation error through the symmetric error model. In these cases, using an emissivity retrieval or a value from climatology

is a secondary problem to merely achieving the assimilation of strongly scattering-affected humidity-sounding channels.

4. Assimilation experiments

The all-sky microwave brightness temperatures are used alongside many other satellite and conventional observations within a 12 h assimilation window in the operational ECMWF 4D-Var assimilation system (Rabier *et al.*, 2000), which produces global analyses and forecasts. The observation minus model differences (the so called first-guess departures), which drive the data assimilation system, are computed through the 12 h assimilation window using the nonlinear forecast model at the highest available resolution to propagate the background atmospheric state forward in time. The incremental 4D-Var, at lower resolution (about 80 km), finds the 12 h forecast evolution that fits the available observations optimally using a linearized forecast model. The atmospheric control variables, which consist of transforms of surface pressure (but not the surface temperature) and 3D fields of humidity, temperature and the two horizontal wind components, are adjusted during the minimization. Cloud and precipitation are not part of the control vector, but during the minimization they are diagnosed from the dynamical and humidity fields every time step. Hence, by adjusting the temperature and moisture profile at the observation location, it is possible to modify cloud and precipitation to allow the analysis to fit the all-sky observations.

The next sections examine results from the all-sky and clear-sky assimilation of SSMIS humidity-sounding channels. The full assimilation system was run from 1 June–31 August 2013 using 137 levels in the vertical and a horizontal resolution of approximately 40 km ($T511$). Although this resolution is lower than the operational system (which uses $T1279$, about 16 km), it is sufficient to generate realistic cloud and precipitation fields and at the time this work was carried out it was the normal resolution for testing at ECMWF. The experiments are based on the research cycle 39r1. The 39r1 cycle contained two upgrades particularly relevant to the assimilation of microwave humidity sounding observations: (i) extension of clear-sky MHS assimilation to sea-ice surfaces and cold oceans (e.g. sea-surface temperature less than 278 K; Di Tomaso *et al.*, 2013) and (ii) assimilation of SSMIS humidity channels over the ocean in all-sky conditions (Geer, 2013). On top of this, we will test the addition of the SSMIS humidity channels over land in either the all-sky or clear-sky approach.

4.1. All-sky and clear-sky assimilation

We run two experiments, which both make use of the all-sky path of the ECMWF system: (i) an all-sky experiment that implements the methodology documented in the previous sections and (ii) a clear-sky experiment that emulates the operational assimilation of MHS 183 GHz channels in clear-sky conditions. The unchanged version of cycle 39r1 represents our control experiment, which assimilates the same satellite and conventional observations as would the operational system. On top of cycle 39r1, the two experimental configurations use the following settings for assimilation.

(1) All-sky experiment

- Observation operator includes cloud and precipitation effects, using discrete dipole calculations for a sector snowflake along with a C_{\max} approach to effective cloud fraction;
- no cloud screening is applied;
- emissivity retrievals are made through Eq. (7);
- observation-error formulation using the symmetric scattering index.

(2) Clear-sky experiment

This forces C to be zero, consequently

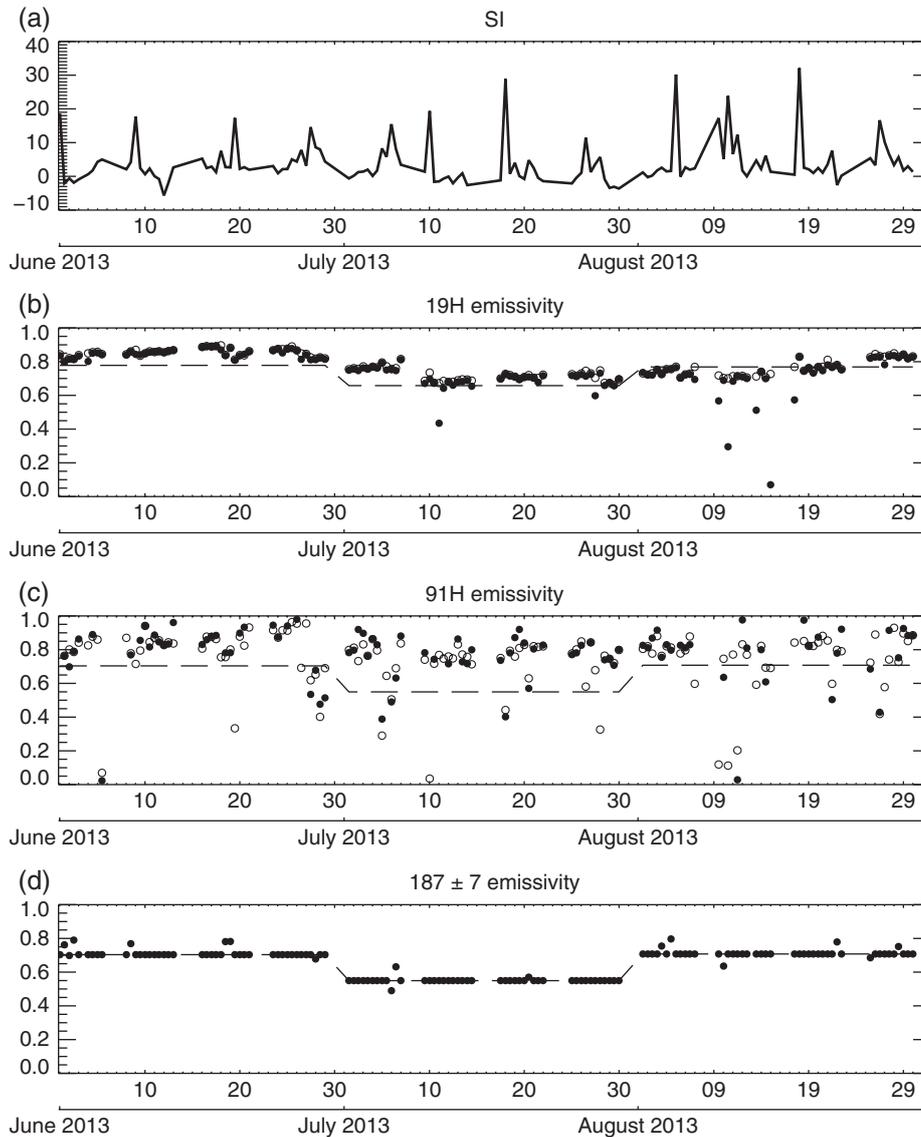


Figure 7. Time-series of SSMIS emissivities and scattering index at 25.61°N , 88.5°E , which is a rice-growing region on the India–Bangladesh border: (a) observed scattering index (91H–150H); (b) atlas (dashed line), clear-sky emissivity retrieval (empty circle) and all-sky emissivity retrieval at 19H; (c) as (b) but at 91H; and (d) the emissivity used for assimilation at 183 GHz.

- the simulated brightness temperature is determined only by the clear subcolumn (e.g. Eq. (1));
- cloud-affected observations are identified and rejected by looking at the size of FG departures at 150 GHz (an observation is considered cloud-free only if the difference between the observed and simulated radiance is less than 5 K);
- the emissivity retrievals are computed using Eq. (6);
- a constant observation error of 2 K is applied.

In both cases, as described earlier, the emissivity retrieval has been replaced by atlas values if it fails or if it is too far from the climatology estimate. As already mentioned, the land sample is restricted to latitudes less than 60° and to grid-points where land cover is nearly unbroken (land–sea mask greater than 0.95). In this study we also use a conservative approach and we follow the strategy adopted in the operational system to assimilate clear-sky MHS observations over land: observations are discarded according to screening criteria based on fixed values of model surface temperature and orography. To avoid possible snow-covered areas, which are prone to larger errors in the retrieval of the surface emissivity, we reject observations where the temperature is lower than 278 K. There is also an orography check to avoid situations where the surface signal becomes too large. The three 183 GHz channels have weighting functions that peak at different levels in the troposphere and consequently they can be classified as lower (183 ± 7), mid (183 ± 3) and upper

(183 ± 1) tropospheric humidity channels. Hence, observations are rejected where the surface altitude is higher than 800, 1000 or 1500 m for channels 9, 10 or 11. As an example, the surface-temperature check rejects about 5% (21 414) of the total number of SSMIS data available for the month of June 2013. Loss of data is not particularly high during this period, i.e. the summer period for the Northern Hemisphere. In winter time, due to colder conditions over land, the percentage of rejected observations is larger. In both cases, many of the observations rejected for cold temperatures would also have been rejected by the orography check. The orography check rejects 31, 24 and 12% of data in channels 9, 10 and 11 (percentages are always with respect to 1 month of SSMIS data).

4.2. Results

Figure 8 examines the normalized departures in the lower and upper tropospheric humidity channels for the all-sky and clear-sky experiments. The normalized departure is the FG departure divided by the observation error predicted by the symmetric error model in the all-sky case and by a constant observation error of 2 K in the clear-sky experiment. The normalized departure is strongly related to the data usage in the assimilation: the square of this quantity gives the observation's weight in the 4D-Var cost function. The sample of data in Figure 8 is one month of SSMIS observations considering all available data over land

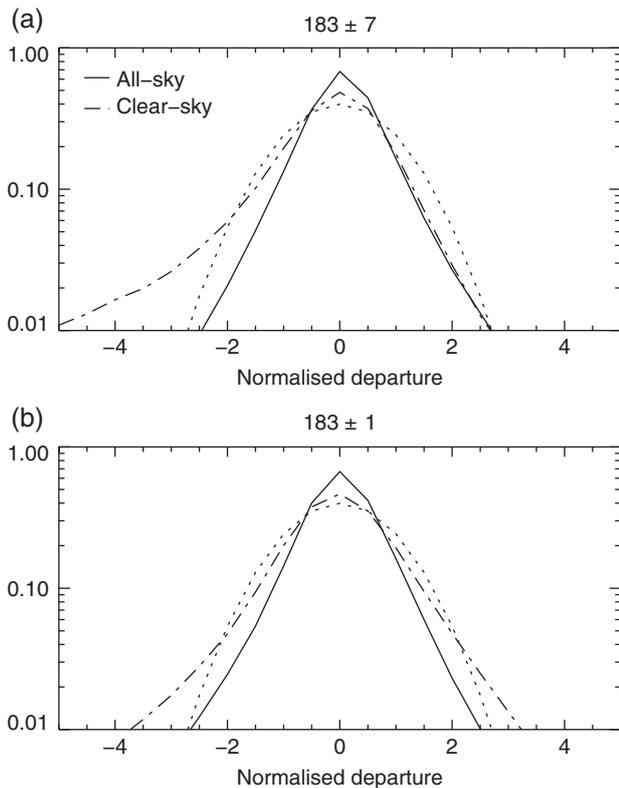


Figure 8. Log histograms of normalized departures (bias-corrected departure divided by the applied observation error) for SSMIS channels (a) 9 and (b) 11 for the month of June 2013 from the all-sky (solid line) and clear-sky (dash-dotted line) experiments. The dotted line shows a Gaussian with a standard deviation of 1. The bin size is 0.5.

(i.e. applying restrictions on latitude and land–sea mask, but no additional screening is applied even for cloud or orography). Figure 8(a) and (b) indicate that in the all-sky experiment the symmetric observation error is doing its job of providing a more Gaussian behaviour of FG departures globally. This is one of the necessary requirements of data assimilation algorithms. As demonstrated earlier, in situations where convection is present in the model and absent in the observation (or vice versa), FG departures can be very large. However, the observation-error model can assign errors of up to 57 K (e.g. in channel 9), making the distribution of normalized departures relatively Gaussian. Compared with an actual Gaussian PDF, the all-sky normalized departures are smaller and the peak of the distribution is higher. This suggests that the use of our data may be too cautious: the observation error might be too large and consequently the weight of the observation during the minimization is smaller than it could be. Similar results were found for the all-sky assimilation of SSMIS humidity sounding over the ocean (Geer, 2013), but it was still possible to obtain improvements in forecasts.

The clear-sky experiment shows a tail of negative departures (where the observation is colder than the simulated brightness temperature), which represents those cases where the observations are cloud-affected and the clear-sky radiative transfer, because it is not simulating cloud absorption or scattering, is not able to generate radiances cold enough. For assimilation, of course these situations are mostly removed by the cloud screening. The negative tail is more obvious in the lower peaking channel, which overall observes clouds and precipitation more frequently than the higher-peaking channels. In addition to the quality controls already described, SSMIS humidity-sounding observations in both all-sky and clear-sky tests are quality-controlled by a first-guess departure check, which rejects normalized first-guess departures greater than 2.5. From Figure 8, it can be inferred that a larger number of observations are discarded by the first-guess check in the clear-sky case.

We can now compare the all-sky and clear-sky samples of assimilated observations in terms of maps of mean normalized FG departure for the month of June 2013 (Figure 9). The general impression is that, in both lower and upper tropospheric channels, the all-sky experiment shows less residual bias in the sample of assimilated observations. Disagreement in normalized FG departures between the all-sky and clear-sky experiments might be explained taking into account three main differences in modelling: (i) radiative transfer; (ii) observation error; (iii) land-surface emissivity. The clear-sky experiment reduces complexity in radiative transfer computations and screens out cloud-affected observations which, as shown in Figure 8(a), would otherwise systematically generate negative FG departures. Broadly, the cloud-screening at 150 GHz is doing its job, considering that Figure 9(b) and (d) do not show large negative features. However, at 183 ± 1 GHz in the clear-sky experiment there are some positive normalized FG departures larger than those in the all-sky case (e.g. Northern South America and Central Africa; Figure 9(d)). These cases most likely represent situations where the model has cloud and precipitation and the observation is cloud-free. In the clear-sky assimilation, even though the cloud and precipitation radiative transfer is ignored, the model humidity profile will likely be saturated, leading to colder simulated brightness temperatures compared with the observations. Overall, the all-sky approach is doing well: the use of the scattering radiative transfer and the observation-error formulation is effective in handling ‘representivity’ errors of the assimilation system and this brings smaller FG departure bias in the sample of assimilated observations.

The total number of assimilated observations is larger in the all-sky assimilation than in the clear-sky one: globally, the all-sky approach brings an additional 8, 10 and 13% (relative to the total available) of assimilated observations respectively in channels 9, 10 and 11. Table 4 gives percentages of assimilated observations in the all-sky and clear-sky experiments in different parts of the globe: Northern Hemisphere (NH), Tropics (T) and Southern Hemisphere (SH). The NH gains most observations when the all-sky approach is used. Globally, in channel 9, the lowest peaking channel, data usage goes from 41 to 49%, the other half of the data still being lost due to high orography or potential snow cover. Baordo *et al.* (2013) describe some of the early development of the all-sky assimilation of SSMIS humidity-sounding channels over land, including an investigation of the use of the surface-to-space transmittance to guide the choice of which observations to discard. Instead of using thresholds based on the surface orography and skin temperature, a fixed transmittance threshold could be used to reject observations with too great a sensitivity to the surface. This might allow more observations to be assimilated. However, forecast scores of assimilation experiments were not improved compared with those shown here (for the summer period, June–August 2012, results showed no significant difference, while for the winter period, January–March 2012, forecast scores appeared degraded). It is clearly important to do more work to allow the use of data over high-altitude and colder land surfaces, as they are still responsible for the rejection of half of all data over land in the lowest peaking channel.

The impact on analysis and forecasts of the all-sky and clear-sky experiments is examined in Figure 10, looking at the normalized change in root-mean-square (RMS) forecast error for vector wind at 850 hPa for the Northern Hemisphere. Assimilation of additional humidity observations over land in all-sky conditions is likely to have its greatest impact on NH winds, through the 4D-Var model-tracing effect that helps to infer winds directly in the data assimilation system. Both the all-sky and clear-sky experiments are compared with the control (experiment minus control), so that negative values indicate reduced RMS forecast errors and consequently an improvement with respect to the control. No statistically significant reductions in RMS forecast error are observed in Figure 10. This is not surprising, considering that the number of observations we are adding

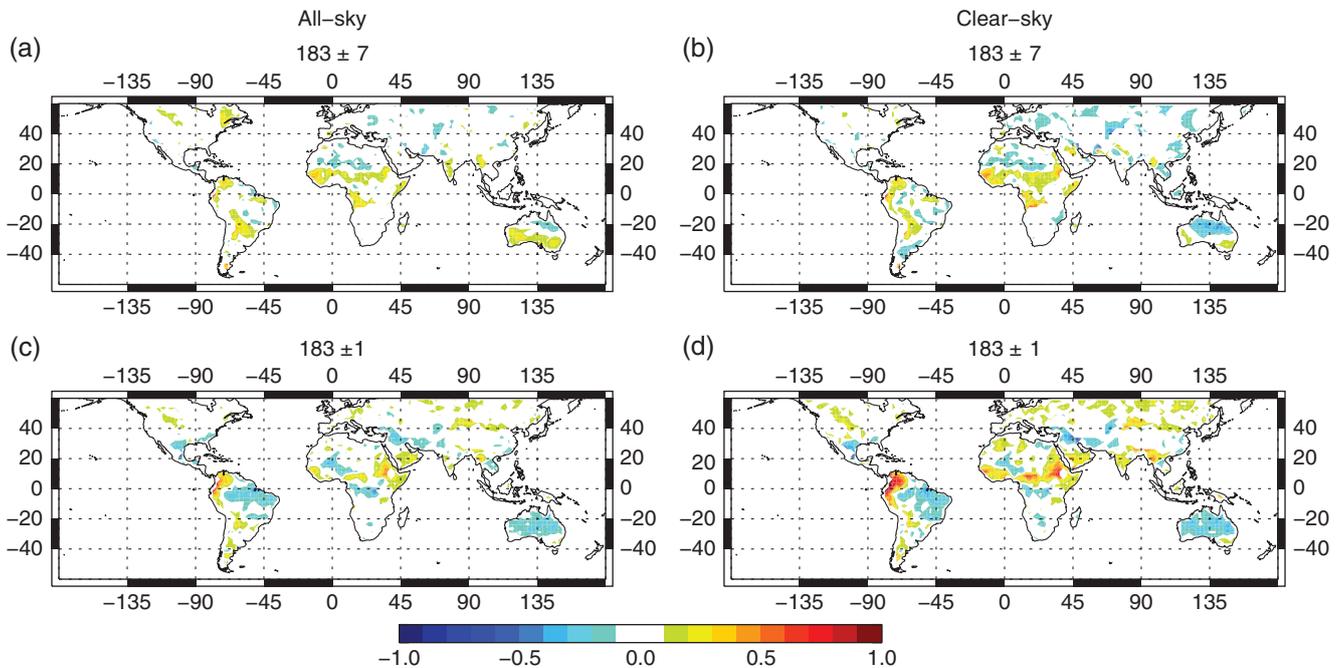


Figure 9. Maps of mean normalized FG departure (bias-corrected departure divided by the applied observation error) for assimilated SSMIS humidity-sounding observations for the month of June 2013: channels 9 and 11 from (a,c) all-sky and (b,d) clear-sky experiments. The grid spacing is 2.5° .

Table 4. Percentage of assimilated SSMIS humidity-sounding observations over land for the month of June 2013 from the all-sky and clear-sky experiments.

Channel	All-sky (%)	Clear-sky (%)
Global (60°S – 60°N) Total number of observations 405 577		
9	49.2	40.8
10	55.1	44.9
11	65.2	52.2
NH (20°N – 60°N) Total number of observations 243 006		
9	45.2	36.3
10	51.8	40.7
11	62.9	48.0
T (20°S – 20°N) Total number of observations 111 167		
9	58.7	50.3
10	63.5	54.0
11	72.9	61.7
SH (60°S – 20°S) Total number of observations 51 404		
9	47.4	41.7
10	52.1	45.3
11	59.1	51.7

Assimilation is restricted to latitudes equatorward of 60° . The percentage is computed with respect to the total number of observations.

to the system, no matter which assimilation approach is used (clear-sky or all-sky), is small compared with the number already available. The control experiment already benefits from microwave humidity sounding observations from four MHS instruments (in clear sky over ocean, land and sea ice), one Advanced Technology Microwave Sounder (ATMS) (clear sky over ocean and land) and one SSMIS (all-sky over ocean). This means that it is difficult to obtain clear improvements in forecast scores. Nevertheless, it is clear that forecasts are not degraded when the new observations are assimilated and the RMS forecast error moves in the right direction (towards negative values), suggesting that the new assimilated observations are not harmful for the system. A generally neutral impact is also seen in relative humidity, temperature and geopotential (not shown here).

Another way to identify the effect of all-sky assimilation is to examine the fits to other assimilated observations. A reduction in standard deviations of FG departures is a sign that the short-range forecast (and the analysis that initialized the forecast) is improving. As an example, Figure 11 examines the normalized change in standard deviation of FG departures for

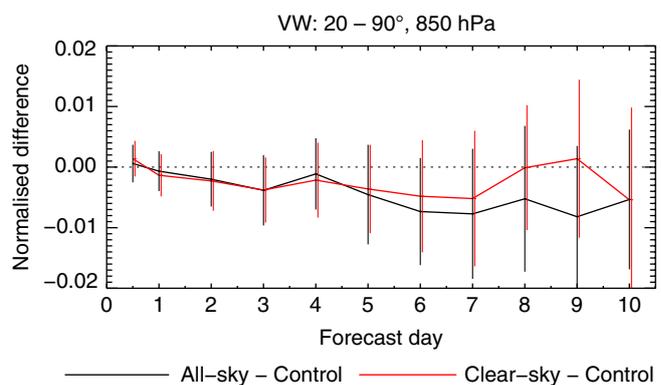


Figure 10. Normalized change in RMS vector wind error at 850 hPa in the northern extratropics. Verification is against own analysis. Scores are based on a maximum of 77 forecasts from summer 2013 (16 June–31 August). A two-week spin-up is excluded from the beginning of each experiment. Error bars indicate the 95% confidence level.

ATMS observations in the Northern Hemisphere. The clear-sky and all-sky experiments behave roughly the same with a few exceptions: the fit to ATMS temperature channel 9 is in favour of the clear-sky experiment, counterbalanced by a larger reduction in standard deviations in ATMS channels 18 and 19 for the all-sky experiment (such changes are also statistically significant, as shown by the error bars, which indicate 95% confidence level). Channels 18 and 19 (respectively 183 ± 7 and 183 ± 4.5 GHz) are two of the ATMS humidity-sounding channels and would be expected to be most sensitive to the assimilation of similar water-vapour-sensitive channels on SSMIS. The improvements in the all-sky experiment are statistically significant and are likely to be the result of the larger number of assimilated observations in the NH, compared with clear-sky assimilation. This is one confirmation that the all-sky approach is bringing additional and beneficial information to the assimilation system.

As mentioned, Baordo *et al.* (2013) carried out a number of different assimilation experiments, initially to investigate the all-sky assimilation of SSMIS humidity-sounding channels over land, using an earlier version of the ECMWF system. These experiments were run covering different seasons and used a variety of screening configurations (e.g. cloud screening; rejecting observations for symmetric SI values greater than 10 K; blacklisting channel 9;

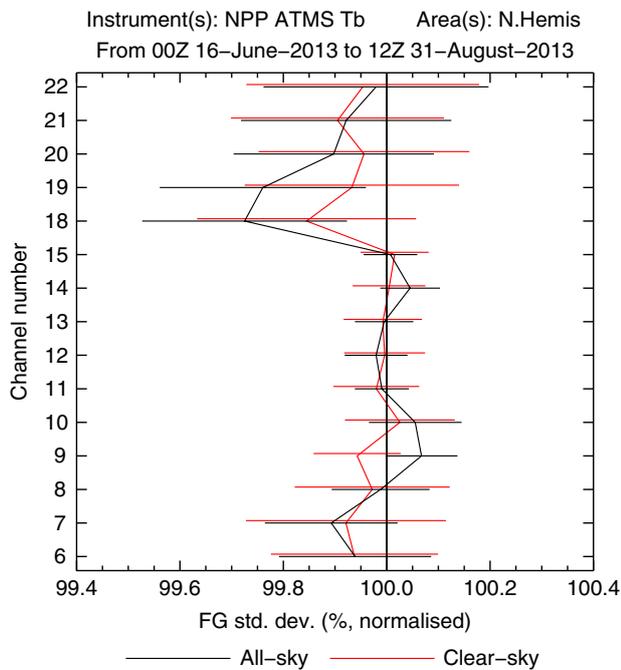


Figure 11. Normalized standard deviation of FG departures in temperature (channels 6–15) and humidity (channels 18–22) for ATMS observations in the Northern Hemisphere. The departures are normalized by the results of the control experiment, so that 100% corresponds to that experiment. A two-week spin-up is excluded from the beginning of each experiment. Error bars indicate the 95% confidence level.

transmittance screening). The all-sky configuration as presented in this article provided the best results. However, the majority of configurations showed similar beneficial (if insignificant) changes in the forecast quality, as in Figure 10, and beneficial, statistically significant improvements in observation fits, as in Figure 11.

5. Conclusions

Emissivity retrieval techniques have allowed clear-sky microwave observations to be assimilated over land surfaces even when there is significant sensitivity to the surface. Developments in all-sky assimilation have allowed the assimilation of cloudy and precipitating scenes, but only over ocean surfaces. This study combines the two new developments to try assimilating cloudy and precipitating observations over land surfaces. Particular issues have been the development of a new observation-error model and the estimation of land-surface emissivity in cloudy and precipitating conditions. These developments are documented and tested in this work using SSMIS 183 GHz channels as an example.

To handle the non-Gaussian behaviour of first-guess departures due to errors in the forecast model in predicting cloud and precipitation both in the right place and with the correct intensity, we formulated an observation-error modelling based on the symmetric (i.e. the average of observed and simulated) scattering index given by the difference between SSMIS channel 18 (91 GHz) and channel 8 (150 GHz). This formulation, mainly driven by scattering from deep convective cloud, is able to assign small errors in cloud-free scenes (2.8 K) and increase the observation error linearly where the influence of scattering, cloud and precipitation increases. The adaptive error modelling can generate observation errors as large as 57 K, allowing us to handle situations where deep convection is present in the observations and not in the model (or vice versa) and where first-guess departures can be as large as 90 K.

In the all-sky framework, we attempt emissivity retrievals for all SSMIS observations (e.g. clear, cloudy or precipitating scenes) over land. In order to be consistent with the way the observations are used in the all-sky assimilation, we implemented an emissivity retrieval taking into account the presence of cloud

and precipitation in the model. This follows the two independent column approach (one clear and one cloudy) used for all-sky radiative transfer simulations; these are weighted according to the effective transfer cloud fraction. We call this the all-sky emissivity retrieval, although it is not intended to be valid in situations with significant scattering. It has been assessed using estimates from a monthly-mean emissivity database (TELSEM) and the more usual retrieval approach, which considers the clear column only (what we call clear-sky retrieval). In most cases, the all-sky and clear-sky retrievals are in agreement. In cloud-free scenes, the use of instantaneous emissivity retrievals might be helpful in capturing real adjustments in local surface conditions, as well as adapting better to the specific instrument characteristics, including the slightly varying fields of view associated with each observation. In these cloud-free cases, the use of emissivity estimates at 91 GHz is good enough to simulate observations in the 183 GHz channels. We observed that a 5% change in the emissivity climatology leads to 1 K error in the simulation at 183 ± 7 but an almost identical simulation at 183 ± 3 and 183 ± 1 .

In the presence of clouds, it is feasible to retrieve emissivities in the lower window channels (e.g. 19 and 37 GHz), but at higher microwave frequencies (e.g. 50 and 91 GHz) all-sky and clear-sky retrievals are both affected by cloud contamination. As the amount of scattering increases in the observations, emissivity estimates can become unphysically low. This generally happens when the observation has much more cloud than the model. In the opposite case, when the model has much more cloud than the observation, the all-sky retrieval can generate emissivities greater than 1. This is a problem unique to the all-sky approach, as a retrieval under the clear-sky approach will generate a good retrieval. However, in these situations the atlas emissivity is used instead and the symmetric error model will provide large observation errors (greater than 10 K) due to the presence of cloud in the model. This means that a 1–2 K error due to uncertainty in the atlas emissivity estimate is not important. However, in the case of moderate scattering and light rain and where observation and model have roughly same amount of cloud, retrievals appear possible and the emissivity estimates look plausible compared with the TELSEM atlas values. Overall, the all-sky emissivity retrieval has been demonstrated to work well in most cases and TELSEM atlas values can be used for quality control and to substitute an emissivity estimate when the retrieval is failed or considered erroneous (about 10% of cases considering 1 month of SSMIS observations, compared with 6% with a clear-sky retrieval). The combination of an all-sky retrieval and all-sky assimilation does seem to reduce the size of FG departure biases over land compared with the clear-sky approach. We would argue that treating the retrieval and assimilation problem symmetrically and working with a complete, all-sky sample of observations (rather than just selected clear-sky observations) is likely to be less biased overall. Further, this approach looks to the future, where short-range model forecasts of clouds and precipitation will become increasingly reliable and ‘mislocation’ errors between model and observation will cause fewer problems.

The reliability of the all-sky land framework has also been investigated by running assimilation experiments and comparing results of the all-sky assimilation of SSMIS 183 GHz channels with those provided by the equivalent clear-sky approach. Overall, the all-sky experiment not only reduces the magnitude of the bias in the assimilated sample of normalized FG departure, but also increases the total number of assimilated observations globally. In channel 9, the lowest peaking 183 GHz channel, data usage over land is increased from 41 to 49%; in channel 11, the highest peaking, it increases from 52 to 65%. The remaining data rejections are due to high orography or possible snow cover. The assimilation of either clear-sky or all-sky SSMIS humidity-sounding observations over land has a generally neutral impact on forecasts. However, fits to other assimilated observations are improved (standard deviations of first-guess departures decrease with respect to the control experiment) in both cases,

confirming that the assimilated humidity-sounding observations are not harmful for the system. In the all-sky experiment, fits to assimilated ATMS humidity-sounding observations (channels 18 and 19) in the Northern Hemisphere are significantly improved. This is likely linked either to the larger number of assimilated observations or the less-biased FG departures. This is vicarious evidence that the all-sky approach is bringing additional useful information into the assimilation system.

Further developments in the all-sky assimilation of humidity sounders are examined by Geer *et al.* (2014b), who extended the all-sky emissivity retrieval to sea-ice areas following the approach of Di Tomaso *et al.* (2013) and finally transferred the four MHS instruments into the all-sky framework. With five humidity-sounding instruments and considering ocean, land and sea-ice surfaces, that work was able to demonstrate clear and statistically significant benefits from all-sky humidity-sounder assimilation, with roughly double the impact of the clear-sky approach (these changes became operational at ECMWF in May 2015). The work described in this article was essential to making that possible.

Acknowledgements

Thanks to many colleagues across ECMWF for technical support and important scientific discussions, in particular Peter Bauer and Stephen English, who helped to develop and guide this project. Fatima Karbou is acknowledged for initial development work on emissivity retrievals in the all-sky system at ECMWF. Fabrizio Baordo was funded through the EUMETSAT Research Fellowship Programme.

References

- Aires F, Prigent C, Bernardo F, Jiménez C, Saunders R, Brunel P. 2011. A Tool to Estimate Land-Surface Emissivities at Microwave frequencies (TELSEM) for use in numerical weather prediction. *Q. J. R. Meteorol. Soc.* **137**: 690–699.
- Auligné T, McNally AP, Dee DP. 2007. Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. R. Meteorol. Soc.* **133**: 631–642.
- Baordo F, Geer AJ, English S. 2012. 'SSMIS radiances over land in the all-sky framework: One year EUMETSAT fellowship report', EUMETSAT/ECMWF Fellowship Programme Research Report 27. ECMWF. <http://www.ecmwf.int> (accessed 29 July 2016).
- Baordo F, Geer AJ, English S. 2013. 'All-sky assimilation of SSMIS humidity-sounding channels over land: Second year EUMETSAT fellowship report', EUMETSAT/ECMWF Fellowship Programme Research Report 30. ECMWF. <http://www.ecmwf.int> (accessed 29 July 2016).
- Bauer P, Moreau E, Chevallier F, O'Keefe U. 2006. Multiple-scattering microwave radiative transfer for data assimilation applications. *Q. J. R. Meteorol. Soc.* **132**: 1259–1281.
- Bauer P, Geer AJ, Lopez P, Salmon D. 2010. Direct 4D-Var assimilation of all-sky radiances: Part I. Implementation. *Q. J. R. Meteorol. Soc.* **136**: 1868–1885.
- Bell W, Candy B, Atkinson N, Hilton F, Baker N, Bormann N, Kelly G, Kazumori M, Campbell W, Swadley S. 2008. The assimilation of SSMIS radiances in numerical weather prediction models. *IEEE Trans. Geosci. Remote Sens.* **46**: 884–900.
- Bennartz R, Bauer P. 2003. Sensitivity of microwave radiances at 85–183 GHz to precipitating ice particles. *Radio Sci.* **38**: 8075.
- Bennartz R, Thoss A, Dybbroe A, Michelson B. 2002. Precipitation analysis using the Advanced Microwave Sounding Unit in support of nowcasting applications. *Meteorol. Appl.* **9**: 177–189.
- Bormann N, Geer A, English S. 2012. Evaluation of the Microwave Ocean Surface Emissivity Model FASTEM-5 in the IFS, ECMWF Technical Memorandum 667: ECMWF. <http://www.ecmwf.int> (accessed 29 July 2016).
- Di Tomaso E, Bormann N, English S. 2013. 'Assimilation of ATOVS radiances at ECMWF: Third year EUMETSAT fellowship report', EUMETSAT/ECMWF Fellowship Programme Research Report 29. ECMWF. <http://www.ecmwf.int> (accessed 29 July 2016).
- English SJ. 2008. The importance of accurate skin temperature in assimilating radiances from satellite sounding instruments. *Q. J. R. Meteorol. Soc.* **126**: 2911–2931.
- Fabry F, Sun J. 2010. For how long should what data be assimilated for the mesoscale forecasting of convection and why? Part I: On the propagation of initial condition errors and their implications for data assimilation. *Mon. Weather Rev.* **138**: 242–255.
- Field PR, Heymsfield AJ, Bansemir A. 2007. Snow size distribution parameterization for midlatitude and tropical ice clouds. *J. Atmos. Sci.* **64**: 4346–4365.
- Geer AJ. 2013. All-sky assimilation: better snow-scattering radiative transfer and addition of SSMIS humidity-sounding channels, ECMWF technical memorandum 706: ECMWF. <http://www.ecmwf.int> (accessed 29 July 2016).
- Geer AJ, Baordo F. 2014. Improved scattering radiative transfer for frozen hydrometeors at microwave frequencies. *Atmos. Meas. Tech.* **7**: 1839–1860.
- Geer AJ, Bauer P. 2010. 'Enhanced use of all-sky microwave observations sensitive to water vapour, cloud and precipitation', Published simultaneously as ECMWF Technical Memorandum 620 and ECMWF/EUMETSAT Fellowship Report 20. ECMWF. <http://www.ecmwf.int> (accessed 29 July 2016).
- Geer AJ, Bauer P. 2011. Observation errors in all-sky data assimilation. *Q. J. R. Meteorol. Soc.* **137**: 2024–2037.
- Geer AJ, Bauer P, O'Dell CW. 2009a. A revised cloud overlap scheme for fast microwave radiative transfer. *J. Appl. Meteorol. Climatol.* **48**: 2257–2270.
- Geer AJ, Forbes R, Bauer P. 2009b. 'Cloud and precipitation overlap in simplified scattering radiative transfer', ECMWF/EUMETSAT Fellowship Programme Research Report 18. ECMWF. <http://www.ecmwf.int> (accessed 29 July 2016).
- Geer AJ, Bauer P, Lopez P. 2010. Direct 4D-Var assimilation of all-sky radiances: Part II. Assessment. *Q. J. R. Meteorol. Soc.* **136**: 1886–1905.
- Geer AJ, Bauer P, English SJ. 2012. 'AMSU-A temperature sounding channels in presence of cloud and precipitation', Published simultaneously as ECMWF Technical Memorandum 670 and ECMWF/EUMETSAT Fellowship Report 24. ECMWF. <http://www.ecmwf.int> (accessed 29 July 2016).
- Geer AJ, Baordo F, Bormann N, English S. 2014. All-sky assimilation of microwave humidity sounders, ECMWF technical memorandum 741: ECMWF. <http://www.ecmwf.int> (accessed 29 July 2016).
- Joseph J, Wiscombe WJ, Weinman JA. 1976. The delta-Eddington approximation for radiative flux transfer. *J. Atmos. Sci.* **33**: 2452–2459.
- Karbou F, Prigent C, Eymard L, Pardo J. 2005. Microwave land emissivity calculations using AMSU-A and AMSU-B measurements. *IEEE Trans. Geosci. Remote Sens.* **46**: 863–883.
- Karbou F, Gérard E, Rabier F. 2010a. Global 4D-Var assimilation and forecast experiments using AMSU observations over land. Part-I: Impact of various land-surface emissivity parameterizations. *Weather and Forecasting* **25**: 5–19.
- Karbou F, Rabier F, Lafore J-P, Redelsperger J-L, Bock O. 2010b. Global 4D-Var assimilation and forecast experiments using AMSU observations over land. Part II: Impact of assimilating surface sensitive channels on the African Monsoon during AMMA. *Weather and Forecasting* **25**: 20–36.
- Kazumori M, English S. 2014. Use of the ocean surface wind direction signal in microwave radiance assimilation. *Q. J. R. Meteorol. Soc.* **141**: 1354–1375.
- Krzeminski B, Bormann N, Karbou F, Bauer P. 2009. 'Improved use of surface-sensitive microwave radiances over land at ECMWF'. In *Proceedings of the EUMETSAT Meteorological Satellite Conference, EUMETSAT*. Darmstadt, Germany.
- Kunke D, Poe G, Boucher D, Swadley S, Hong Y, Wessel J, Uliana E. 2008. Design and evaluation of the first special sensor microwave imager/sounder. *IEEE Trans. Geosci. Remote Sens.* **46**: 863–883.
- Liu G. 2008. A database of microwave single-scattering properties for nonspherical ice particles. *Bull. Am. Meteorol. Soc.* **111**: 1563–1570.
- Liu Q, Weng F, English S. 2011. Water emissivity model. *IEEE Trans. Geosci. Remote Sens.* **49**: 1238–1250.
- Marshall JS, Palmer WMK. 1948. The distribution of raindrops with size. *J. Meteorol.* **5**: 165–166.
- Petty GW, Huang W. 2011. The modified gamma size distribution applied to inhomogeneous and nonspherical particles: Key relationships and conversions. *J. Atmos. Sci.* **68**: 1460–1473.
- Peubey C, McNally AP. 2009. Characterization of the impact of geostationary clear-sky radiances on wind analyses in a 4D-Var context. *Q. J. R. Meteorol. Soc.* **135**: 1863–1876.
- Prigent C, Chevallier F, Karbou F, Bauer P, Kelly G. 2005. AMSU-A surface emissivities for numerical weather prediction assimilation schemes. *J. Appl. Meteorol.* **44**: 416–426.
- Rabier F, Järvinen H, Klinker E, Mahfouf J-F, Simmons A. 2000. The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.* **126**: 1148–1170.
- Ruston B, Weng F, Yan B. 2008. Use of one-dimensional variational retrieval to diagnose estimates of infrared and microwave surface emissivity over land for ATOVS sounding instruments. *IEEE Trans. Geosci. Remote Sens.* **46**: 376–384.