

The impact of satellite derived land surface temperatures on numerical weather predication analyses and forecasts

Article

Published Version

Candy, B., Saunders, R. W., Ghent, D. and Bulgin, C. E. ORCID: <https://orcid.org/0000-0003-4368-7386> (2017) The impact of satellite derived land surface temperatures on numerical weather predication analyses and forecasts. *Journal of Geophysical Research: Atmospheres*, 122 (18). pp. 9783-9802. ISSN 2169-8996 doi: [10.1002/2016JD026417](https://doi.org/10.1002/2016JD026417) Available at <https://centaur.reading.ac.uk/72267/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1002/2016JD026417>

Publisher: American Geophysical Union

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

RESEARCH ARTICLE

10.1002/2016JD026417

Key Points:

- Satellite measurements of land surface temperature from the GlobTemperature project are compared to estimates from the Met Office numerical weather prediction model and to an in situ radiometer from a site in the UK
- A Kalman filter for land surface analysis has been adapted to ingest these satellite land surface temperature observations, initially for those made during nighttime
- Satellite nighttime observations of land surface temperature are found to improve near-surface air temperature forecasts, particularly over Africa

Correspondence to:

B. Candy,
brett.candy@metoffice.gov.uk

Citation:

Candy, B., Saunders, R. W., Ghent, D. & Bulgin, C. E. (2017). The impact of satellite-derived land surface temperatures on numerical weather prediction analyses and forecasts. *Journal of Geophysical Research: Atmospheres*, 122, 9783–9802. <https://doi.org/10.1002/2016JD026417>

Received 22 DEC 2016

Accepted 30 AUG 2017

Accepted article online 4 SEP 2017

Published online 30 SEP 2017

The Impact of Satellite-Derived Land Surface Temperatures on Numerical Weather Prediction Analyses and Forecasts

B. Candy¹ , R. W. Saunders¹, D. Ghent² , and C. E. Bulgin³ 

¹Met Office, Exeter, UK, ²Department of Physics and Astronomy, University of Leicester, Leicester, UK, ³Department of Meteorology, University of Reading, Reading, UK

Abstract Land surface temperature (LST) observations from a variety of satellite instruments operating in the infrared have been compared to estimates of surface temperature from the Met Office operational numerical weather prediction (NWP) model. The comparisons show that during the day the NWP model can underpredict the surface temperature by up to 10 K in certain regions such as the Sahel and southern Africa. By contrast at night the differences are generally smaller. Matchups have also been performed between satellite LSTs and observations from an in situ radiometer located in Southern England within a region of mixed land use. These matchups demonstrate good agreement at night and suggest that the satellite uncertainties in LST are less than 2 K. The Met Office surface analysis scheme has been adapted to utilize nighttime LST observations. Experiments using these analyses in an NWP model have shown a benefit to the resulting forecasts of near-surface air temperature, particularly over Africa.

Plain Language Summary An accurate analysis of the state of the land surface, particularly the temperature and moisture content of the soil has been shown to improve forecasts from weather models. The temperature within the soil can be analysed using conventional observations from meteorological stations. However, the stations are not distributed uniformly across the earth's surface. In this work we investigate what happens if we use remote sensing measurements of surface temperature to estimate the soil conditions. The resulting soil conditions are then used to initialize a weather forecast model and we find that the air temperature forecasts are improved, especially over Africa.

1. Introduction

Numerical weather prediction (NWP) models require accurate estimates of the fluxes of heat and moisture between the land surface and the atmosphere. Consequently, many current NWP models contain (or couple) to a land surface model (LSM). Examples include H-TESEL at the European Centre for Medium-Range Weather Forecasts (ECMWF) (Balsamo et al., 2009), JULES within the Met Office Unified Model (UM) system (Best et al., 2011), and Noah LSM at the National Centers for Environmental Prediction (NCEP) (Chen & Dudhia, 2001). Improvements to LSMs have been shown to lead to improved near-surface forecasts of atmospheric parameters such as near-surface air temperature. Examples which can make large changes to near-surface atmospheric forecasts over continental scales include improvements to hydrology modeling and the specification of surface albedo (Martin et al., 2010) and a more accurate representation of soil thermal and hydraulic parameters (Dharssi et al., 2009). At the regional-scale changes to soil temperature and soil moisture initialization methods (e.g., Di Giuseppe et al., 2011) have been shown to improve the representation of boundary layer parameters. In particular, a study by Fan (2009) shows that the use of observed soil temperatures can influence convective development and therefore precipitation forecasts.

While some regions, notably the continental United States and parts of Europe and Australia, contain networks of sensors that measure soil moisture and temperature directly, most regions of the globe are poorly covered with such in situ observations of the land state (Dorigo et al., 2011). Consequently, land data assimilation schemes for use in NWP models have been developed to use observations of the near-surface air temperature and humidity from conventional meteorological stations. In recent years these assimilation schemes have been adapted to also include estimates of soil moisture from satellite-based active and passive microwave instruments. Typical data assimilation methods include optimal interpolation (e.g., Douville et al., 2000) and the extended Kalman filter (EKF) technique (e.g., de Rosnay et al., 2013). Operational soil analyses at the Met Office are produced via an EKF that uses both observations of air temperature and humidity from land stations (typically, following WMO (World Meteorological Organization) recommendations, at a height

of between 1.2 and 2 m above ground level) and surface soil wetness from the advanced scatterometer (ASCAT) instrument. ASCAT is a radar scatterometer on board the MetOp series of satellites. The EKF was implemented in the Met Office operational NWP suite in 2013. An attractive aspect of this method is that the Jacobian or \mathbf{H} matrix is estimated as part of each EKF run through the embedded LSM. It is this matrix that contains the sensitivities between the parameters to be analyzed, in this case soil moisture and temperature, and each observation. Potentially, within this method it is possible to exploit satellite observations which have more indirect sensitivity to the analysis parameters, such as measurements of leaf area index and land surface temperature (LST). Measurements of LST, which represent the interface temperature between the land surface and the overlying atmosphere, are potentially challenging to assimilate. LST values can vary significantly between differing surface types and rapidly change magnitude in response to atmospheric effects such as changing daytime cloud cover. Nevertheless, there are several reasons to investigate the use of LST within NWP assimilation. First, there are now several satellite products available in near real time with good quality estimates of LST. One example of this is a product over Europe, Africa, and parts of South America that is generated every 15 min from observations made by the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument on board the Meteosat Second Generation satellites. These observations have been shown to agree with in situ estimates of LST to within 2 K at an arid desert site (Freitas et al., 2010). Second, several studies have demonstrated improvements when LST estimates have been assimilated into LSMs. For example, a recent study (Ghent et al., 2010) used a similar approach to the EKF to assimilate SEVIRI data into a LSM over a domain covering Africa. Improvements were shown in both modeled soil moisture and heat fluxes. Such benefits, if realized within an NWP system, are very likely to benefit near-surface atmospheric forecasts. A final motivation for examining LST assimilation concerns the use of satellite sounding channels over land within NWP. Modeling of the sounding radiances from the previous forecast is a critical step in using this data and for channels that are sensitive to the surface a good estimate of the LST is required. For instance, Pavelin and Candy (2014) reject infrared sounding data over land for some arid surfaces where the global model NWP forecast LST appears to have a large cold bias, particularly during daytime in the Sahara. This is also an issue for limited area models; Guedj et al. (2011) report a similar cold bias for the ALADIN/France NWP model simulations (when compared to geostationary LSTs) over the Spanish Peninsular. In this study we use LSTs acquired via the GlobTemperature project (Ghent et al., 2016) to perform impact studies in a state of the art NWP system via a land surface data assimilation scheme. GlobTemperature is part of the European Space Agency Data User Element program and has the aim to increase user uptake of Earth observation data over a wide range of disciplines concerning both research and operational applications. The GlobTemperature data include LST data sets from both geostationary and polar-orbiting satellites, and although the project does not provide the majority of these in near real time, the data sets are generated using the latest retrieval methods. To enhance user accessibility, there is a harmonized data format and each observation includes a quality indicator and an estimate of uncertainty. Although the work described in this paper concerns the impact of LSTs on numerical weather prediction for short-term weather forecasting, the GlobTemperature data record will span over 20 years of data and so will also be a useful source for the initialization of the land component of future NWP reanalysis projects. While there have been several reported studies of assimilating LSTs into land data assimilation schemes, to our knowledge the work reported here is the first attempt to assimilate LSTs into such a scheme which is then used to initialize an NWP model. This paper is structured as follows. In section 2 the satellite LST data sets are described and compared to LST forecasts from the Met Office operational NWP global model and also in situ measurements. In section 3 the Met Office land data assimilation system is described with a focus on how the LSTs are used within it. We demonstrate how the additional information within the satellite data sets, in particular, quality information and supplied uncertainties, can be used within the assimilation process. The impact of the LST assimilation is examined through a series of experiments using the trial version of the Met Office NWP system. Results from these experiments are explored in section 4, examining the impact on model variables in both the land and atmosphere components. Conclusions and recommendations for further work are made in section 5.

2. Comparisons Between Satellite, In Situ, and NWP Land Surface Temperature

We assess the impact on NWP of assimilating LST retrieved from satellite data into the land surface scheme. Improvements in LST retrieval schemes and data accuracy are now making this a viable option, but we first compare these LSTs to NWP forecasts and in situ data to assess data quality and their suitability for

Table 1*GlobTemperature LST Products Used in This Study*

Instrument	Platform	Orbit	Spatial resolution (km)	Product reference
AATSR	Envisat	Sun synchronous 10:00 A.M. equator crossing time	1	Ghent (2012)
SEVIRI	Meteosat Second Generation	Geosynchronous	5	Freitas et al. (2010)
MODIS	Aqua	Sun synchronous 1:30 P.M. equator crossing time	1	Ghent et al. (2016)

assimilation and determine any constraints on their use. We describe here the GlobTemperature LST, NWP forecast and in situ data sets used, and their comparisons.

2.1. GlobTemperature LSTs

GlobTemperature is an ESA funded project with the aim of facilitating the use of satellite-derived LSTs to the wider user community (Ghent et al., 2016). This project aims to enhance data set accessibility across a wide range of applications through the generation of data sets in a harmonized format from a variety of sensors operating in both the infrared and the microwave. The data sets include merged products from infrared sensors that have been intercalibrated in LST space. In this work we use LSTs that have been derived from each sensor at the orbit (or segment of orbit) level; these are commonly referred to as “Level 2” products. Table 1 summarizes each infrared sensor used in the study; both geostationary and polar instruments are represented. For Advanced Along Track Scanning Radiometer (AATSR) and SEVIRI LSTs the Globtemperature data sets utilize existing retrieval algorithms developed for these instruments. In the case of Moderate Resolution Imaging Spectroradiometer (MODIS) a new retrieval has been derived for Globtemperature (Ghent et al., 2016).

LSTs are estimated from each of these sensors for both day and night scenes using cloud-free observations from channels with low atmospheric absorption in the 10–12 μm region. A good estimate of LST from top of the atmosphere measurements needs to take into account several factors including atmospheric effects, particularly water vapor absorption, and the variation in surface emissivity. The split window technique, in which a linear combination of measurements from several infrared channels is used to derive LST, is a common method for infrared sensors. For instance, the SEVIRI LSTs used in this paper are derived using observations from the 10.8 and 12 μm channels with coefficients computed as a function of the surface emissivity (Freitas et al., 2010). Accounting for the variation in emissivity for land scenes is important as there is a wide variation in the infrared region; typically, depending on the surface type, values are found in the range of 0.9 to 1. Also, for regions of the Earth with a distinct growing cycle, there is a seasonal variation in emissivity as the fraction of vegetation increases, peaks, and then reduces in the field of view. Consequently, relationships can be determined between the surface emissivity and satellite-derived fraction of vegetation. An alternative approach is to use climatological emissivity maps which typically are derived from multispectral instruments (e.g., Seemann et al., 2008) and interpolate in space and time to obtain an emissivity estimate at the satellite observation location. The former approach is adopted for retrievals using the AATSR instrument, the latter by MODIS. Typically, the uncertainties in LST arising from the split window technique over land are in the range 1 to 3 K, with the largest errors associated with desert regions where the surface emissivity represents the largest component of the uncertainty budget (Freitas et al., 2010). LST retrievals for scenes with a combination of large satellite zenith angles in regions of high water vapor will also contain large uncertainties. Consequently, these uncertainties are very specific to the location and manner in which the observation is made; an attractive aspect of the GlobTemperature data sets in studies such as the one reported in this paper is that they include an uncertainty estimate with each LST observation. Provision of LST uncertainty information on a per pixel basis builds on the methodology recently developed within the sea surface temperature community (Bulgin et al., 2016), increasing consistency between remote sensing Earth surface temperature products. Uncertainty estimates are made within the retrieval process itself, propagating uncertainties from instrument noise, use of auxiliary data and retrieval process into the Level 2 product, where they can be independently validated using other data sources such as ground-based observations. As will be shown in section 3 the uncertainties can be used to correctly weight the observations in the data assimilation scheme. The following results are a summary of comparison results between GlobTemperature LSTs and Met Office NWP fields/in situ data.

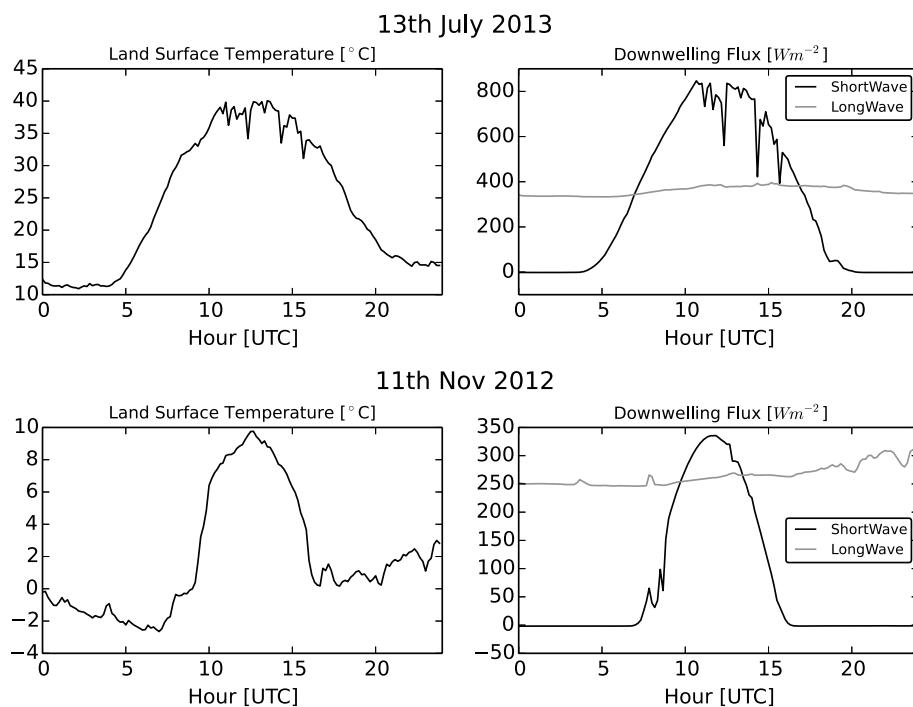


Figure 1. Time series of LST and downwelling radiative fluxes made at the Cardington observing site for (top row) 13 July 2013 and (bottom row) 11 November 2012.

2.2. NWP Forecast Data

The Met Office NWP system includes a land component based on the JULES model (Best et al., 2011), with land usage split into five vegetation types (broadleaf trees, needleleaf trees, temperate (C3) grass, tropical (C4) grass, and shrubs) and four nonvegetated types (urban areas, inland water, bare soil, and land ice). The fraction of each of these types at each land grid point is mapped from a global land use data set provided by the International Geosphere-Biosphere Programme. Model forecast data are provided via the operational Met Office archive, and forecasts are stored in 3-hourly intervals, so the NWP data are matched to the satellite observations via linear interpolation of the model fields in space and time. For the period of data used in the comparison (2011–2013) the NWP horizontal grid was made up of 1024 (east-west) by 769 (north-south) points which represents a resolution of 25 km at midlatitudes. Model forecasts of cloud cover were also extracted and matched to the satellite location and time. The model cloud information was used to ensure that the model state represents an estimate of LST under clear skies, in order to be consistent with the satellite observation. This was tested via a simple cutoff in cloud; any matchup with an NWP cloud cover in excess of 20% was ignored in the following comparisons. It should also be borne in mind that the satellite cloud masks are specific to each instrument and are not expected to be perfect.

2.3. In Situ Measurements of LST

Although not routinely recorded at meteorological stations, LST measurements are collected at several sites across the globe using fixed infrared instruments pointing at the ground. For instance at sites in Portugal (woodland) and Namibia (arid gravel plain). Both these locations have been compared to SEVIRI LST measurements to gain insight into their uncertainties (e.g., Trigo et al., 2008). We use data from the Met Office observation site at Cardington, UK (52.10°N, -0.42°E). This site is based at a former airship station and is surrounded by open farmland and several villages in a relatively cloudy environment and so is a useful addition to the existing in situ comparisons made at the other sites. At Cardington an estimate of the surface temperature is made via a Heitronics KT15D infrared thermometer pointed at the ground from a height of 2.5 m leading to a target size over a grass surface of 1 m² in area. The instrument measures in the range of 8 to 14 μm wavelength with an uncertainty of 0.5 K, according to the manufacturer. Examples of LST time series made during summer and winter are shown in Figure 1.

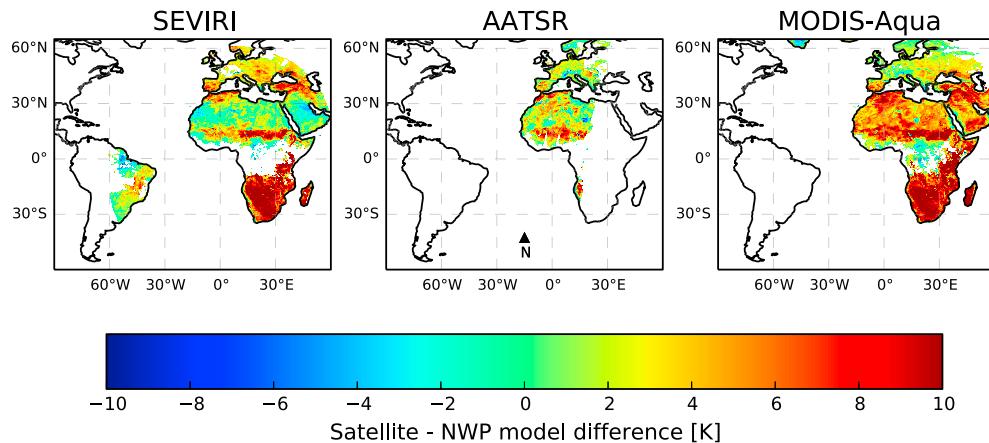


Figure 2. Daytime satellite-NWP model difference in LST for (left) SEVIRI, (middle) AATSR, and (right) MODIS. Accumulation periods are October 2013 for SEVIRI/MODIS and October 2011 for AATSR.

It is apparent in Figure 1 that the diurnal range of LST during summer is much larger than during winter, in part due to the higher magnitudes of the short wave fluxes. Both these cases are predominantly cloud free, and it can also be seen that the LST responds rapidly to overhead cloud cover, for example, at around midday on 13 July. This highlights that it is important to ensure that LST matchups between differing sources are made as close as possible in time and represent similar conditions of cloud cover.

Since the infrared thermometer actually measures the radiance arriving at the aperture of the instrument, and as most real surfaces are unlikely to behave as a black body, then before comparisons can be made with satellite LSTs, a correction is required in order to obtain the true skin temperature of the surface. The correction follows a similar approach used in comparisons with AATSR data (Noyes et al., 2007). The radiance arriving at the aperture of the instrument can be expressed in the following way:

$$B(T_{\text{measure}}) = \varepsilon B(\text{LST}) + (1 - \varepsilon)B(T_{\text{sky}}) \quad (1)$$

where $B(T_{\text{measure}})$ represents the measured radiance, $B(\text{LST})$ the upwelling radiance emitted by the surface, $B(T_{\text{sky}})$ the downwelling radiance from the atmosphere, and ε the surface emissivity. B represents the Planck Function. Equation (1) leads to a conversion to LST from the observed radiance of the form:

$$\text{B(LST)} = \frac{1}{\varepsilon} (B(T_{\text{measure}}) - (1 - \varepsilon)B(T_{\text{sky}})). \quad (2)$$

Use of equation (2) requires an estimate of the surface emissivity, and this is taken from Infrared Atmospheric Sounding Interferometer (IASI) retrievals collocated to the Cardington site (see Pavelin & Candy, 2014). This was determined from the IASI spectral emissivity retrieval by applying a weighting using the response function of the Cardington radiometer in the interval 8–14 μm . In this way the annual average emissivity for the in situ instrument was found to be 0.973. An estimate of T_{sky} is also required and as there is no dedicated upward looking radiometer at the Cardington site this value has been set to 210 K which was also used by Noyes et al. (2007). The value appears appropriate as it lies within the typical clear sky range of 200–250 K determined using radiative transfer modeling for representative atmospheric profiles at the site. Using the above values of emissivity and T_{sky} yields a typical correction to the thermometer measurements of +1.3 K.

2.4. Comparisons of Satellite LST With NWP Forecasts

Here we present comparisons between the Met Office operational short-range forecasts of LST with the equivalent values from the GlobTemperature data sets. Figure 2 shows the mean daytime satellite NWP model LST difference using SEVIRI, AATSR, and MODIS observations. The plots show an average of the comparisons over 1 month of data for daytime overpasses. In each case as the viewing overpass time is different for each instrument (see Table 1), we use the following 2 h periods from each day to construct the monthly average; 10–12 UTC for SEVIRI, 10–12 UTC for AATSR, and 12–14 UTC for MODIS.

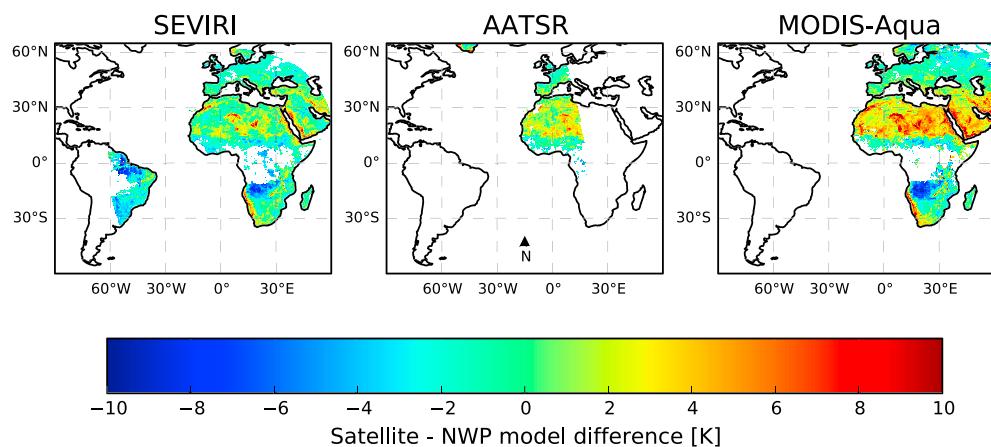


Figure 3. Nighttime satellite-NWP model difference in LST for (left) SEVIRI, (middle) AATSR, and (right) MODIS. Accumulation periods are October 2013 for SEVIRI/MODIS and October 2011 for AATSR.

SEVIRI versus NWP comparisons show a large range in the difference and are generally positive (satellite warmer than NWP), apart from some regions of desert and tropical forest. The Sahel and southern African regions contain a particularly large bias. AATSR daytime overpasses show a similar spatial pattern in the bias, for instance, in the Sahel region. MODIS results are also quite consistent with this apart from warmer retrievals over the Sahara. Notice that some regions during the day, such as the Congo, contain little or no matchups because of the requirement of cloud-free conditions in both the NWP model and satellite data sets. The equivalent plots for nighttime biases are shown in Figure 3. In this case monthly averages are constructed using observations made during 22–00 UTC for SEVIRI, 22–00 UTC for AATSR, and 00–02 UTC for MODIS. The total number of observations used to construct the monthly averages is in excess of 1 million per day in the case of SEVIRI and MODIS while AATSR numbers ~0.5 million per day.

Nighttime biases are more mixed in sign across regions and, while magnitudes are relatively small across Europe, very large negative biases are present over tropical forests and the arid regions of southern Africa. Of the principal surface types in the Africa/Europe region (forests, grassland, and bare soil) the smallest yearly average differences between model and satellite are found for scenes containing a majority of bare soil.

The large positive daytime bias in the Sahel region has also been observed in comparisons using alternative LST data sets. For example, using LSTs derived from the IASI hyperspectral infrared sounder as discussed in Pavelin and Candy (2014). This suggests that the bias originates through a modeling deficiency in the Met Office NWP model. Figure 4 shows the variation in satellite and model LSTs at 12 UTC for several regions computed from 2 years of matchups. It demonstrates that the seasonal variation in the bias between satellite and model can be complex and is dependent on location. For example, in the central Sahel region it can be seen that the bias is consistent in sign throughout the year, with a magnitude rising to almost +10 K as in late 2013. In contrast, at the Saharan location there is a clear summer time peak of similar magnitude but opposite sign, indicating that the model is warmer than the satellite, with good agreement at other times. The strong seasonal change in LST from winter to summer is evident over the British Isles, and this region exhibits a smaller range of daytime biases, with peak values of around +5 K during spring 2013.

2.5. Comparisons of Satellite LST With In Situ Measurements of LST

The corrected Cardington observations were compared to the SEVIRI/NWP model matchups discussed in the previous section, for those satellite pixels that contain the Cardington site. Table 2 summarizes these comparisons; there is very good agreement between the two observing systems for nighttime measurements. For instance, the standard deviations of the differences are less than 2 K and are of a similar magnitude of 1–2 K reported by Trigo et al. (2008) at the site in Portugal. This range gives an indication of the LST uncertainty and is the same range of values that are assigned to uncertainties for surface air temperature observations within the Met Office operational assimilation scheme; it suggests therefore that satellite LSTs are now of sufficient quality for use in data assimilation to support NWP. However, during the day the fits are not as good for the daytime observations, both in terms of bias and standard deviation. This is also seen at the

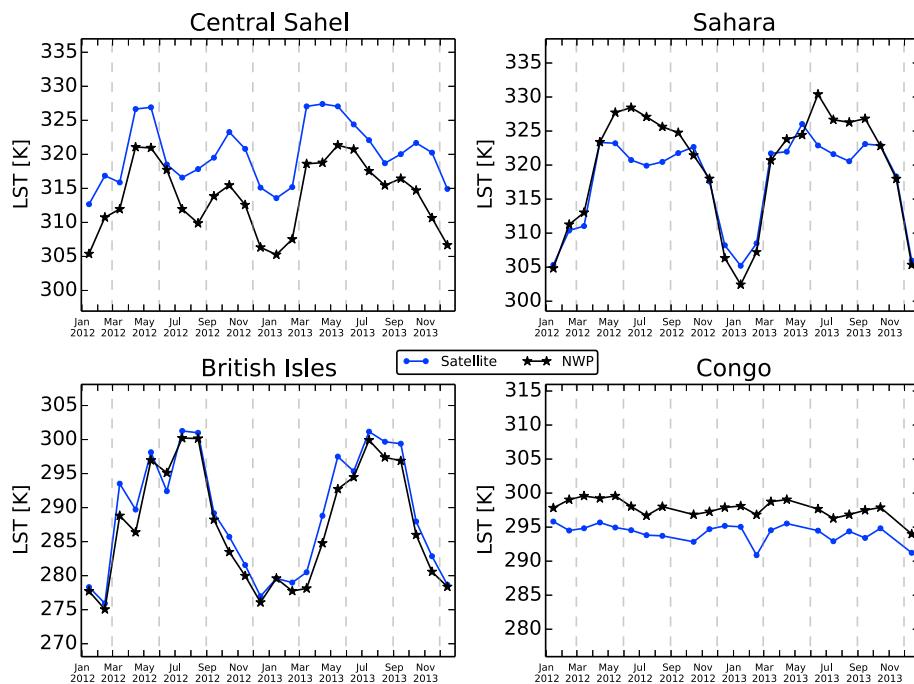


Figure 4. Time series of daytime (12 UTC) monthly mean LSTs from SEVIRI and the Met Office NWP model for four regions. Values are computed from a 2.5×2.5 box, centered on the following locations: central Sahel (12°N , 12°E), Sahara (20°N , 0°E), British Isles (51°N , 1°W), and Congo (0°N , 20°E). Dashed lines represent seasonal boundaries.

Portugal site and reported by Noyes et al. (2007) for AATSR comparisons. The poorer fit is most likely due to an increase in heterogeneity in the LST field due to differing responses to diurnal warming for each type of surface. The in situ radiometer views radiances from a single (grass) surface, whereas the satellite will measure a mixture of radiances from the various surfaces contained within the satellite pixel, such as trees, grass, crops, ponds, and buildings. The effect of this is explored in Figure 5 which shows the variability in LST at each grid point for UM forecasts made over the British Isles at high (1.5 km) resolution. At each land grid point the standard deviation of LSTs for each active tile (trees, grass, bare soil, etc.) is displayed for midnight and midday conditions for the summer case in Figure 5. It can be seen that the variability is much higher, typically in excess of factor 2, at midday for most of the domain.

Because of the difference in scale the satellite observations will contain some of this increased variability, whereas the in situ measurements will not and this leads to an increase in the measured spread in the computed differences between the two observing types. It suggests that at this site the nighttime data set gives a better indication of the effective accuracy of the satellite product.

3. GlobTemperature LSTs and the Met Office Land Surface Scheme

3.1. Land Surface Analysis Scheme Overview

In common with other NWP centers the Met Office NWP system includes real-time analyses of the atmospheric and land states that are used to initialize subsequent atmospheric forecasts. For the global suite analyses are provided every 6 h, and for the case of the land component these are created using a scheme based on the extended Kalman filter (EKF) methodology. Other centers, e.g., ECMWF, also use this approach (de Rosnay et al., 2013) with the aim of combining information from meteorological stations and satellite

Table 2

Day and Night Matchups of SEVIRI LST and the Cardington Radiometer During 2012 and 2013

Day/night	Number	Mn sat-in situ (K)	SD sat-in situ (K)	Correlation coefficient	Mn NWP-in situ (K)
Night	152	0.07	1.52	0.97	2.51
Day	86	-2.25	3.63	0.97	-2.94

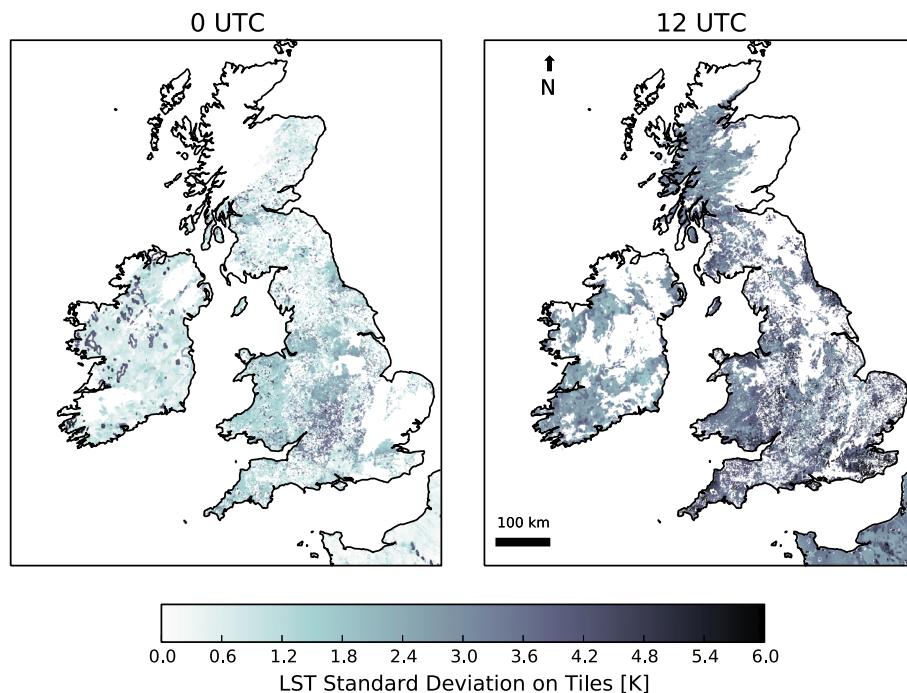


Figure 5. Variability of LST across each surface tile for cloud-free grid points from the Unified Model high-resolution domain: (left) 0 UTC and (right) 12 UTC on 13 July 2013.

observations to form an accurate representation of the land variables. In the case of the scheme implemented at the Met Office the soil moisture and temperature variables are in columns of soil at each land grid point. The soil columns correspond to those used in the operational forecast model and are split into four layers with the following thicknesses (from surface to depth): 0.10, 0.25, 0.65, and 2 m. Analysis increments are applied to each soil column in turn using the following equations:

$$\Delta \mathbf{x} = \mathbf{K}(\mathbf{y} - H(\mathbf{x}_b)). \quad (3)$$

The symbols in equation (3) represent the following: \mathbf{x} represents the analysis vector (soil moisture and soil temperature at each soil level), \mathbf{y} a vector of observations, \mathbf{x}_b is the background vector or short-range forecast from the previous cycle, and H is the observation operator which provides the conversion from model to observation space.

\mathbf{K} represents the Kalman gain and is computed in the following way:

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

where \mathbf{B} represents the background error covariance matrix, \mathbf{R} the observation error covariance matrix, and \mathbf{H} the linear version of the observation operator. \mathbf{H} contains the sensitivity of the observation to a change in an analysis variable. It is the key link between the indirect observations and the variables we wish to estimate within the soil layers. \mathbf{H} is derived through short duration runs of the JULES land surface model, and these are performed as part of the analysis cycle, to ensure that they are appropriate to the current land surface conditions at each grid point. Operationally, meteorological reports from land stations of near-surface air temperature and humidity are used alongside satellite observations of soil wetness within the Met Office scheme.

The soil wetness data originate from the ASCAT scatterometer instrument on board the MetOp-A and B satellites, and the processing of this data set is described in Dharssi et al. (2011). Station reports are not uniformly distributed across the continents, and in order to provide observations at each grid point preprocessing is applied before use in the EKF. A simplified version of the atmospheric analysis scheme is used to provide a two-dimensional field of surface humidity and temperature. Table 3

Table 3
Observation and Background Errors Used in the Met Office EKF Analysis

Variable	Value
Observation Error: Near-Surface Air Temperature	1.5 K
Observation Error: Near-Surface Air Humidity	8% (relative humidity)
Observation Error: Scatterometer Soil Moisture	0.035 m ³ m ⁻³
Background Error: Soil Moisture Level 1	0.030 m ³ m ⁻³
Background Error: Soil Moisture Levels 2–4	0.026 m ³ m ⁻³
Background Error: Soil Temperature All Levels	2 K

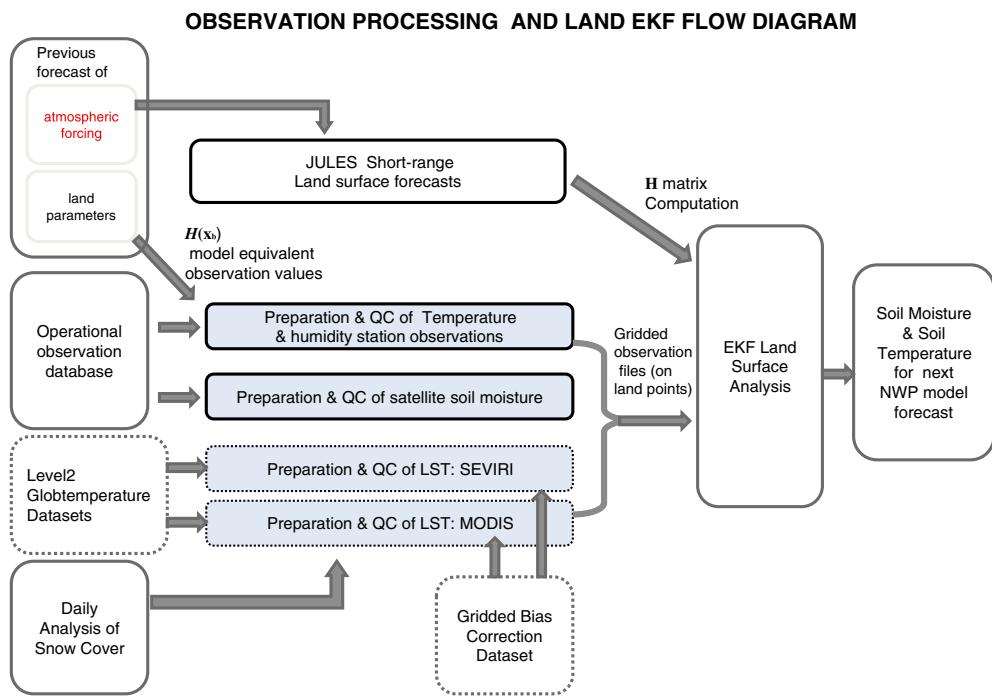


Figure 6. Flow diagram showing preprocessing of land observations and the EKF. The additions for LST processing are highlighted by dashed boxes.

summarizes the background and observation errors used in the EKF computations. These variables are used to construct \mathbf{B} and \mathbf{R} , which are assumed to be diagonal. Observation errors for the surface air data refer to the two-dimensional data sets.

3.2. Inclusion of GlobTemperature LSTs

LST data sets are integrated into the Met Office EKF in order to assess their impact on NWP forecasts. Since the EKF operates on each land point separately, the aim of the LST preprocessing for each instrument (SEVIRI, MODIS, etc.) is to provide a quality-controlled gridded data set covering the Met Office land grid points. The first step is to read in the LST observations made within the analysis window (typically set to 6 h centered on the analysis time) and the corresponding model forecasts of LST and cloud cover (which are available at an hourly frequency). Observations then pass through a series of quality control tests, and observations are individually rejected if any of the following are true:

1. The Observation is located on model sea point.
2. The LST value is outside of the reasonable range taken as 220–350 K.
3. Probability in excess of 50% that the observation contains a gross error through comparison with the model forecast LST using Bayesian theory (Lorenc & Hammon, 1988). This test helps to remove any residual cloud contamination.
4. Snow cover detected at observation location. The Met Office daily snow analysis is used as the source of snow information (Pullen et al., 2011), with a value in excess of 0.05 kg m^{-2} denoting snow.
5. The Model cloud fraction is in excess of 0.2, denoting cloud cover at location of observation.

The GlobTemperature-harmonized data sets include quality control flags for each observation to identify cloud contamination or low confidence in the retrieval and observations are also removed if either conditions are true. For this initial study of NWP impact an additional constraint is applied, whereby all observations made during daylight conditions are removed. This decision is influenced from the comparisons in section 2 where it was found that daytime LST observations appear to have very large positive biases in some regions, such as the Sahel, and the evidence from the in situ study that the fit between sources of LST representing different spatial scales is not as close during daytime conditions.

The final step in the preprocessing concerns the removal of the large biases between the LST and the model. As has been shown in section 2 even for the nighttime LST data there are many regions where the mean bias

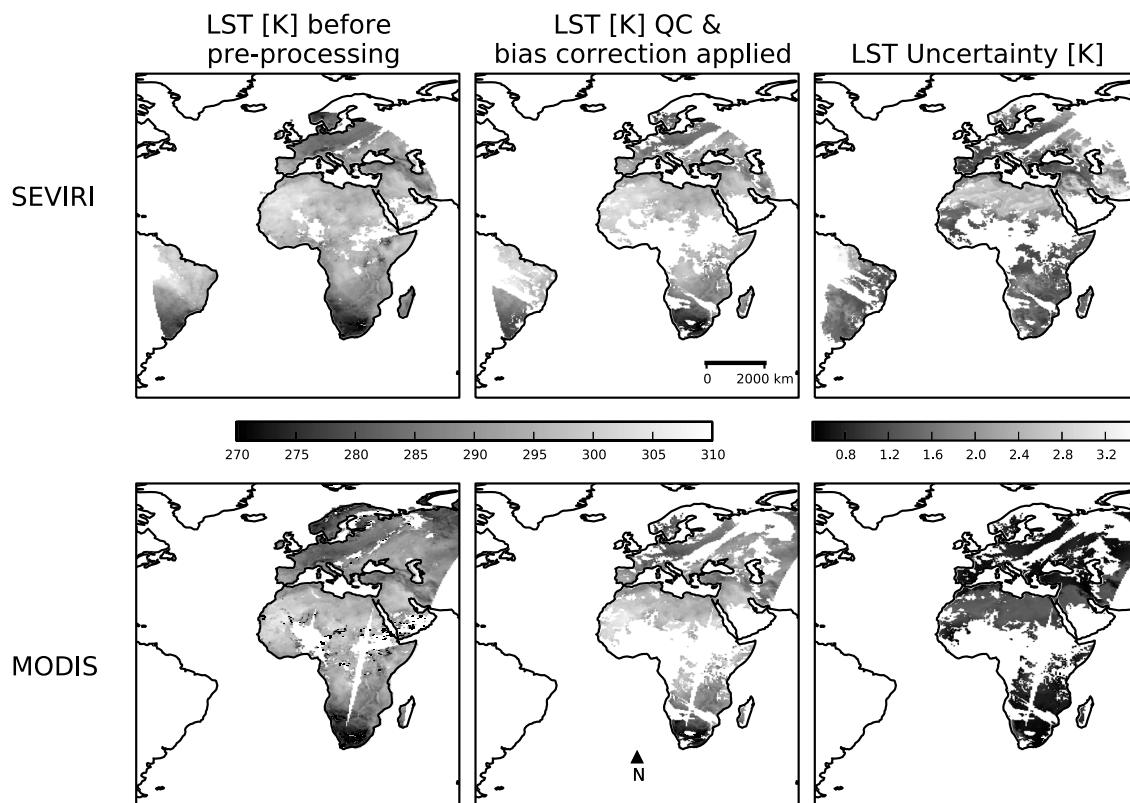


Figure 7. Gridded LST observations before and after preprocessing for (top row) SEVIRI and (bottom row) Aqua MODIS. Date: 0 UTC 15 August 2013. In each case Figure 7 (right column) shows the supplied uncertainties.

between the satellite and model estimates of LST is much greater than the typical uncertainties of around 2 K that are associated with the observations. The approach used in other studies involving satellite measurements of LST and also other parameters, such as soil moisture, is to remove this mean bias prior to assimilation, e.g., as discussed in Ghent et al. (2010). A similar strategy is employed here in which the monthly mean bias is computed between GlobTemperature LST and the NWP model at each land grid point from several years of data, in this case 2012–2013. Biases are computed for each satellite instrument individually to capture effects that are likely to influence the bias such as overpass time, assumptions in the retrieval method and instrument systematic errors. It should be noted that for polar-orbiting instruments, such as MODIS, this treatment does not currently remove any bias that is dependent on satellite viewing angle. Figure 6 summarizes the satellite preprocessing in the form of a flow diagram, while Figure 7 shows LST observations before and after preprocessing for a single analysis time in 2013 for the SEVIRI and MODIS instruments (note that there were no AATSR data available for this time as the satellite failed in April 2012). Typically, about 30% of the observations are removed, with the requirement that the model cloud is clear at the observation location responsible for the majority of rejections. By way of comparison with SEVIRI, Figure 8 shows the typical distribution of in situ reports over Africa, again for a single analysis time. These are currently the main source of soil temperature information within the EKF, and it can be seen that large regions of the continent are unobserved, e.g., south of the Congo River. It is in these regions with conventional observing gaps that the satellite data have real potential to provide information about the land state.

An important change to the system required for the inclusion of LSTs is to include a new row in the EKF \mathbf{H} matrix to account for the sensitivity of LST observations to a change in each of the analysis variables. This is achieved by storing the LST forecast (in addition to the existing variables such as air temperature) from the JULES short duration runs. Figure 9 shows the resulting element in the \mathbf{H} matrix representing the sensitivity of LST at each grid point with respect to a change, or small perturbation, in Level 1 soil temperature (representing a thickness of 10 cm below the surface). The size of the soil temperature perturbation, which

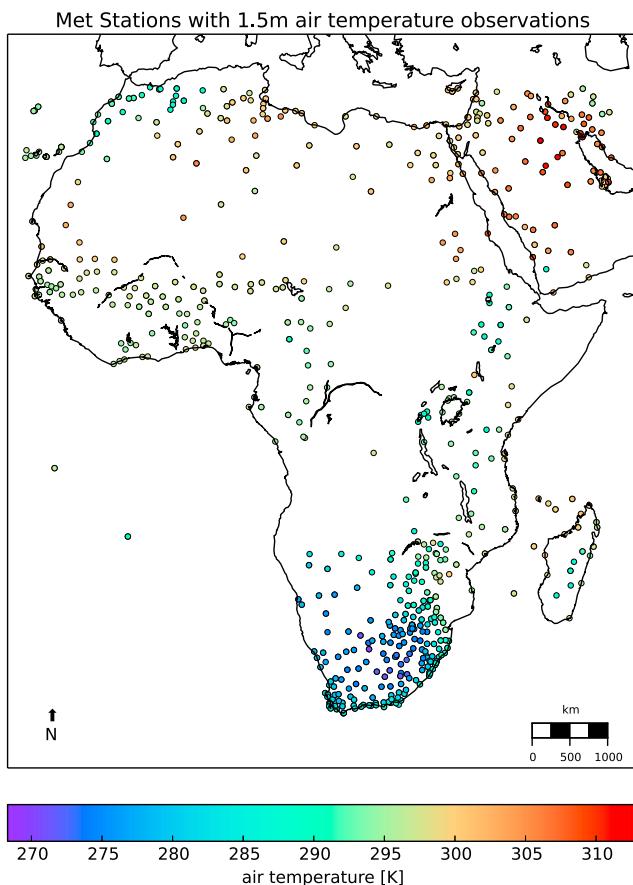


Figure 8. Observations of air temperature from African meteorological stations. Date: 31 August 2013 at 0 UTC.

is applied at the start of the JULES run, is set to -0.05 K. Results show a positive value at all locations, and the strongest coupling between LST and soil temperature is found to be during the night. This confirms studies by Dharssi et al. (2013) for an experimental version of the EKF applied to a different season, where a similar strong nighttime signal was observed. There is also a similar pattern of variation found for \mathbf{H} matrix values with respect to Level 2 soil layers (not shown), albeit with about half the magnitude.

Until this work the EKF has used fixed uncertainties for the existing observation types (i.e., the values in Table 3 were applied at each land point, irrespective of surface type). For the introduction of LST into the system we have chosen to use the GlobTemperature uncertainty estimates to construct the relevant term in the \mathbf{R} matrix for each grid point which then feeds into the Kalman gain matrix (matrix \mathbf{K} in equation (4)). \mathbf{K} controls the size of the updates to the analysis variables (soil temperature and moisture) arising from the latest batch of observations and its magnitude can be considered the effective weight for a given observation type. For SEVIRI (Figure 7) a large variation in LST uncertainty can be observed, from the relatively low values of 1 K to extremes in excess of 3 K. As discussed before these are dependent on the surface type and viewing angle; for SEVIRI the highest errors are found in desert regions and at the edge of the viewing disc. This variation has a large effect on \mathbf{K} and consequently is an important influence on the weight of the satellite data within the analysis. The elements of \mathbf{K} that link SEVIRI LST to Level 1 soil temperature range from 0.1 (low weight over desert) to nearly 1.0 (high weight over grassland). As a comparison the equivalent elements for near-surface air temperature are in the range 0.1 to 0.5, suggesting that the meteorological station data have overall a lower weight in the analysis for updating soil temperature.

4. Impact of LST Observations in NWP

We investigate the impact of assimilating the LST data in the land scheme on the resulting analyses and subsequent atmospheric forecasts. This is carried out through comparisons of a series of runs of the trial version of the Met Office NWP suite. This contains the following components in addition to the land surface EKF; each has the purpose to provide initial conditions for the subsequent atmospheric forecasts: (1) atmospheric analyses using a hybrid 4D-Var system (Clayton et al., 2012) and (2) daily Sea ice and SST analyses (Donlon et al., 2011).

The atmospheric forecasts are generated via the Unified Model (UM) with land surface exchange modeled by the JULES scheme. The current configuration, including land surface settings, of the UM is described in detail by Walters et al. (2014). After each analysis atmospheric forecasts are generated using the UM with a range out to 48 h.

The trial version of the NWP suite differs from the operational setup in two main respects. First, the horizontal grid is reduced from 1536 by 1152 points to 640 by 480. This equates to a degradation in spatial resolution from 16 km to 40 km. Second, the background error matrix for the 4D-Var analyses is derived from climatology, rather than containing an additional flow-dependent component estimated from an ensemble of short-range atmospheric forecasts. Both these changes are made in order to reduce the computational cost. A monthlong run of the trial NWP suite from 1 August to 31 August 2013 was performed using the components as listed above and represents the control run. In the control run the land EKF assimilated the current data sets used in operations; meteorological station reports of temperature and humidity, alongside soil wetness observations from the ASCAT scatterometer. A second run was performed in an identical manner with the addition of the use of satellite LSTs within the land EKF. This is referred to as the experiment run in the following sections. LST observations from SEVIRI on board the Meteosat-10 satellite and MODIS on board the Aqua satellite were used in the experiment run. These instruments were chosen from those available in the

Hmatrix element for LST sensitivity with respect to Soil Temp

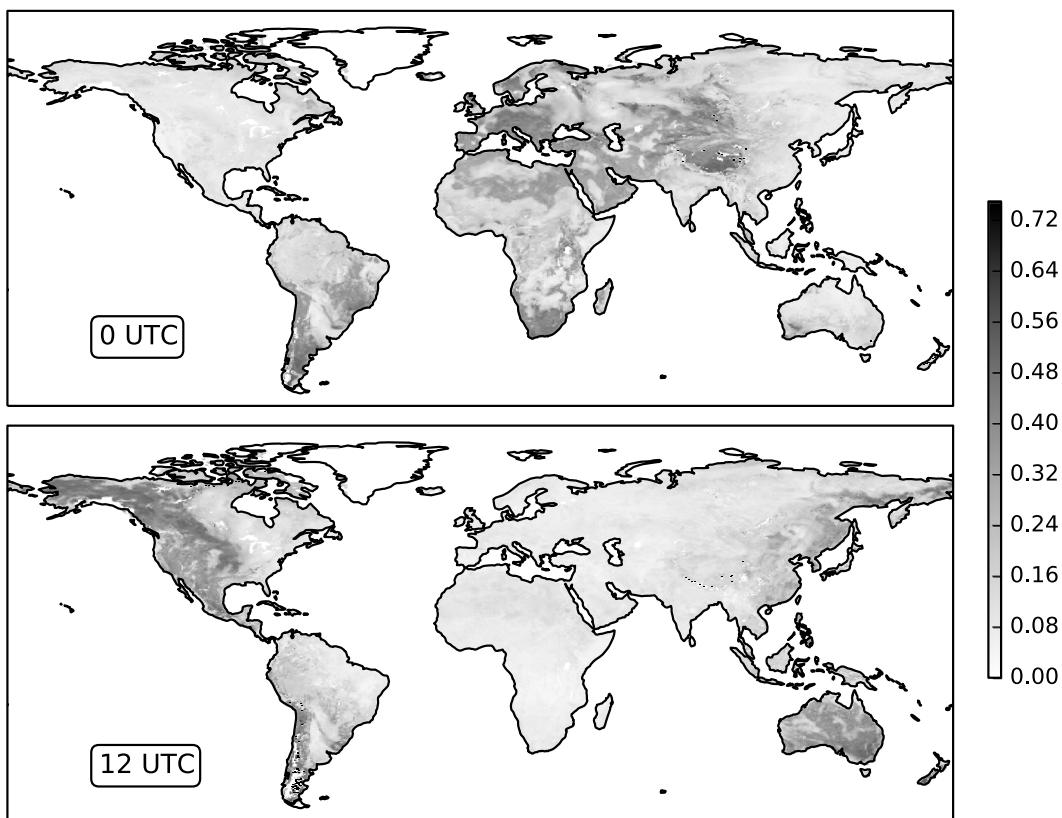


Figure 9. The element of the EKF \mathbf{H} matrix showing the sensitivity of LST with respect to a change in Level 1 soil temperature at 0 UTC and 12 UTC.

GlobTemperature catalogue as they represent instruments in both geostationary and polar orbits. As mentioned earlier the AATSR instrument was no longer operational during the trial period and so could not be assimilated. A second important point is that both instruments have the potential to be available in near real time, making the results seen here of particular interest to other operational NWP centers.

4.1. Land Surface Analyses

Figure 10 shows a time series of analysis increments; these represent the adjustment made by the EKF to create the analysis from the previous forecast. At each cycle the values represent mean and standard deviations taken over all valid grid points in which the EKF is active, typically for the period analyzed this totals 90% of all land points north of Antarctica. Results are shown for the top two soil levels (down to a depth of 35 cm) for soil temperature. It can be seen that in general the mean increment into the soil is slightly more positive in the experiment run when LST observations are assimilated. The standard deviations are significantly higher (by a factor of around 2) in the experiment run suggesting that generally when the LST data are assimilated the size of the changes to create the new temperature analysis at each cycle are much larger. At the two lowest soil levels the change in temperature increments arising from the LST data is negligible and so it can be concluded that the information arising from the LST data is only changing the temperature analyses in the first two layers. To adjust the deeper soil layers, further experiments could be performed in which the forecast error matrix (\mathbf{B} in equation (4)) contains some estimation of the error correlations between soil layers, but this is beyond the scope of this study. Nevertheless, it is very encouraging that the EKF is providing updates into the soil from this new data source and confirms the expected results given the LST/soil temperature coupling suggested from the \mathbf{H} matrix in Figure 9. Notice also that the changes to the soil moisture increments are much smaller, an increase of around 5–10% in the experiment run when compared to the control. Recalling that only nighttime LST is assimilated, this is in line with studies elsewhere (e.g., Balsamo et al., 2007) that suggest information on soil moisture from LST is mainly from daytime observations when the rate of surface temperature rise is affected by the amount of moisture in the soil. Accordingly, subsequent analysis

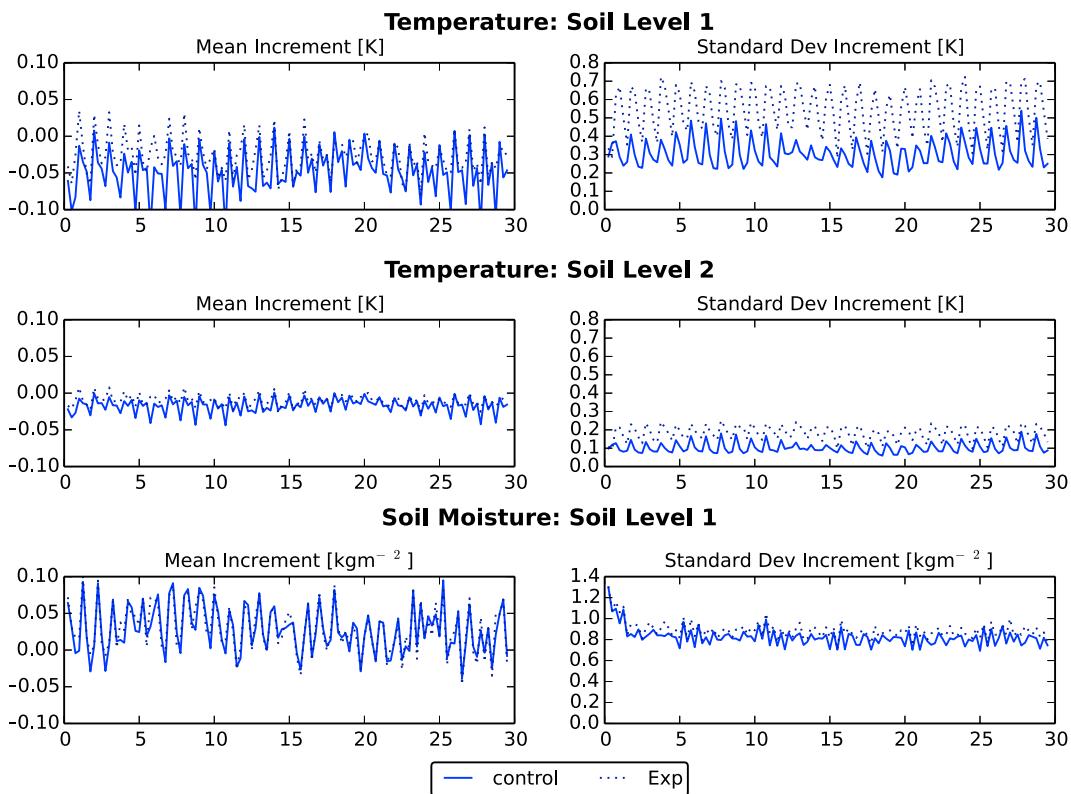


Figure 10. Global time series of analysis increments into the soil arising from the two NWP trials. (top row) Soil temperature Level 1, (middle row) soil temperature Level 2, and (bottom row) soil moisture Level 1. Figure 10 (left column) represents mean increments, and Figure 10 (right column) represents the standard deviations. Statistics are computed from all active grid points in the EKF.

in this paper focuses on impacts related to soil and near-surface temperature, rather than the hydrological cycle.

The impact on the soil temperature analyses from the experiment run can be assessed using independent in situ measurements, and several networks around the globe provide these via sensors embedded directly in the soil. Although not routinely reported by meteorological stations, the data are available from the International Soil Moisture Network and a description of the contributing networks can be found in Dorigo et al. (2011). Figure 11 shows examples of these observations over a full year for the Met Office Cardington site and five other networks in Europe that were reporting during the period of the NWP trial runs.

These other networks are as follows: SMOSMANIA, comprising 20 stations across Southern France (Calvet et al., 2007); HOBE, 29 stations in Denmark (Bircher et al., 2011); TERENO, 5 stations in Germany (Zacharias et al., 2011); REMEDHUS, 20 stations in Spain (University of Salamanca); and WEGENERNET, 12 stations in Austria (University of Graz).

The in situ sensors are generally located at a depth of 5 cm and so are comparable to the first soil level of the NWP model. Figure 11 also shows the equivalent 6 h forecast values for this level at several sites from the operational NWP model, and it can be seen that there is a good agreement with the in situ data; the range of differences is within 5 K, suggesting that the background error of 2 K used within the land EKF is an appropriate estimate. Correlation coefficients are in excess of 0.9. Notice that during the summer period at El Coto there is clear indication that the NWP forecast is too low compared to in situ. This is also seen in other stations of the Spanish network and is consistent with the positive bias seen by the LST observations over the Spanish peninsular seen earlier in this paper (Figures 2 and 3). Similar comparisons were made for the short-range forecasts of soil temperature from the control and experiment runs for each station in the networks and also at the Met Office observing site. Quality control has been applied to remove sites which, for a variety of reasons, such as poor sensor calibration or the site is unrepresentative of the local area, have a poor fit to the NWP model. To do this, the method used by Albergel et al. (2015) is followed; sites are removed with

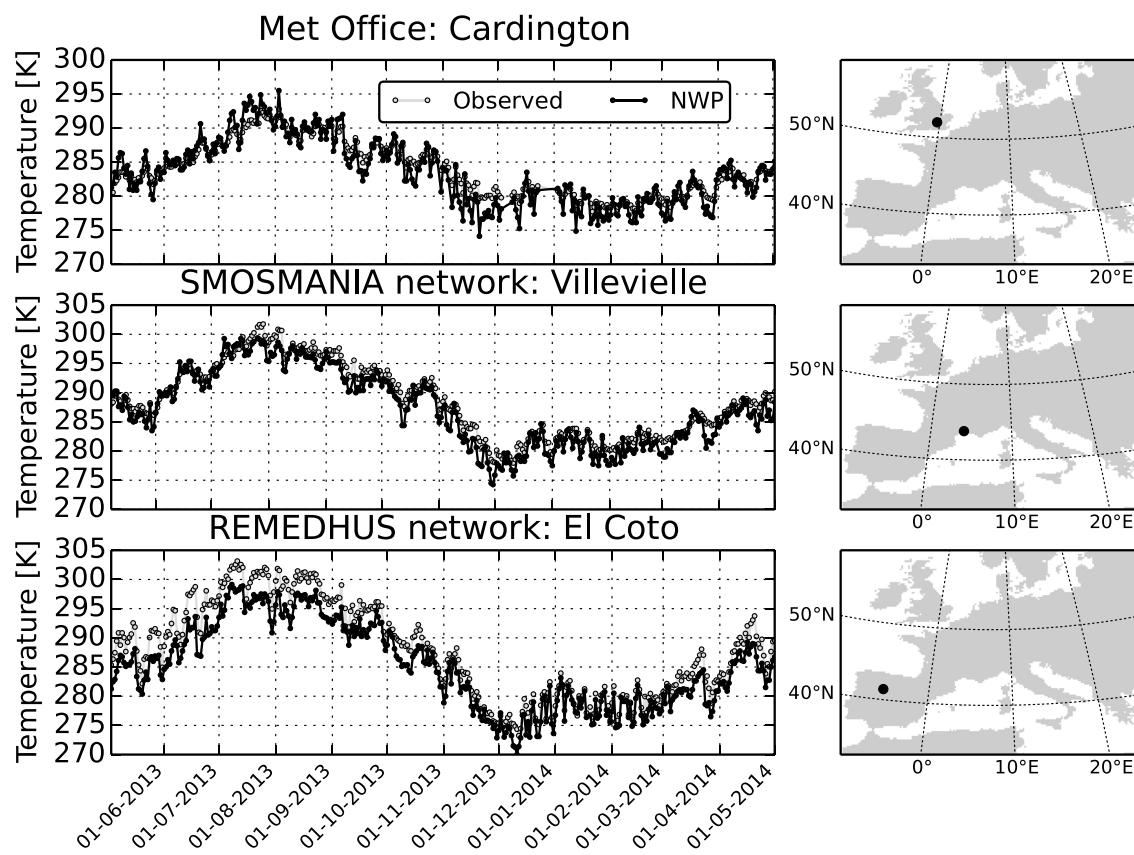


Figure 11. Time series of in situ measurements of soil temperature and short range forecasts of operational NWP model soil temperature at three sites in Europe. (top row) Cardington, UK, (middle row) Villevielle, France, and (bottom row) El Coto, Spain.

poor correlations against the model over 1 year of data as specified by negative correlation coefficient or p value statistic in excess of 0.05. Several networks had stations that were not reporting during the period of the NWP trial, but despite this, 51 sites contributed to the statistics below.

Table 4 shows the mean results for each network, and generally, the changes are small. There is, however, indication for the REMEDHUS network that the effect of the LST assimilation is to reduce the large positive bias between in situ and the NWP model. A similar change is also seen at the Cardington site. Here both the mean and standard deviations of the differences have been reduced by more than 10%. HOBE and WEGENERNET sites also show a reduction in standard deviation.

Another aspect of assessing the impact of the new observations is to look at how the fit of the model LST forecasts changes with respect to the observations. Figure 12 shows the standard deviation of the difference between SEVIRI LST and the short-range model forecast for both the control and experiment runs. The

Table 4

Comparison Statistics of Level 1 Soil Temperature From Each NWP Trial Run Against In Situ Measurements From Five Networks

Network	Number of stations passing QC	Mean Station-NWP model (K)		Standard deviation Station-Model (K)	
		Control	Experiment	Control	Experiment
Met Office: Cardington	1/1	1.14	1.02	0.87	0.75
SMOSMANIA	19/20	2.93	2.90	0.91	0.96
REMEDHUS	15/20	5.70	5.38	0.92	0.90
HOBE	10/14	0.27	0.47	1.17	1.13
WEGENERNET	1/1	-0.62	-0.71	1.72	1.61
TERENO	5/5	0.69	0.71	1.10	1.27

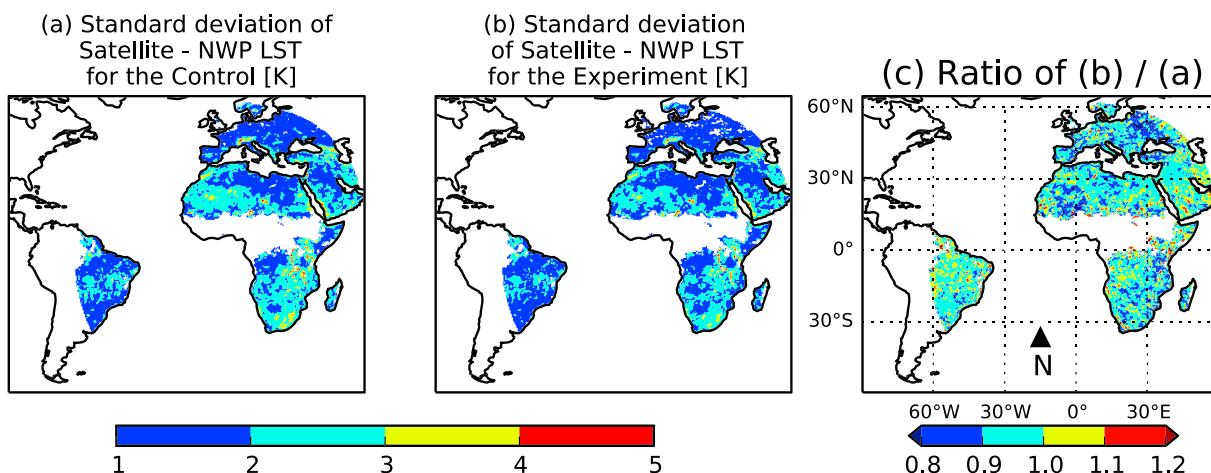


Figure 12. The standard deviation of the difference between SEVIRI and model LSTs for (a) control and (b) experiment runs. The satellite observing time is 3 UTC, and the NWP model forecasts originate from analyses at 0 UTC. (c) The ratio of the standard deviations.

forecasts and observations are valid at 3 UTC. Figure 12c shows the ratio of these differences with values less than 1 indicating that the fit between satellite and model has improved due to the LST assimilation. The majority of grid points have ratios less than 1, suggesting that for most of the region covered by SEVIRI the LST assimilation is resulting in an improved short-range forecast of LST. Splitting statistics by land cover the fit is improved by 5% or more for bare soil regions and grasslands. There is less improvement over forested regions, ~2%, which may be due to reduced coupling between the tree canopy top and the soil layer. This weaker sensitivity is confirmed by looking at the nighttime \mathbf{H} matrix (Figure 9) which contains smaller values over the Amazon and Congo rainforest regions.

4.2. Atmospheric Forecasts

It is common practice at NWP centers to measure forecast performance through comparisons with in situ observations, allowing the calculation of forecast error for a particular parameter between the forecast and the observations. In the case of surface atmospheric temperature this is evaluated using differences between NWP forecasts (generated at each analysis time) representing air temperature at 1.5 m height above the surface and corresponding observations from meteorological stations. In Figure 13 the change in the root-mean-square (RMS) forecast error for surface air temperature between the experiment run and the control is shown for three standard verification regions: Europe (70°N – 25°N , 10°W – 28°E), Africa (40°N – 40°S , 20°W – 52°E) and North America (60°N – 25°N , 145°W – 50°W). The number of forecasts used is one per analysis time and totals 120 across the trial period. Positive values represent a reduced forecast error, or improvement, in the experiment run compared to the control and it can be seen that there is a small but consistent improvement for the European and Africa regions over the full forecast range. The improvements are largest over Africa with reductions in the forecast error in excess of 1% at a forecast range of 24 h. Over North America the results are smaller and of mixed sign representing an overall neutral impact. As the NWP runs examined in this work total 1 month of results and taken together with the number of verifying observations collected, changes in excess of 1% for temperature errors are considered statistically significant. This is based on random sampling from a synthetic distribution of forecast errors following the method described by Chai and Draxler (2014). Consequently, for the trials described here it is the African region in which the LST assimilation results in a statistically significant improvement.

The neutral impact over North America is striking and was investigated further by performing a second experiment with fixed observation uncertainties for LST, as it was noted that the MODIS supplied uncertainties (range 0.5–1.5 K) were much smaller than the equivalent values for SEVIRI (range 1–3 K). An indication of this can be seen in Figure 7. The observation uncertainties were fixed in this second experiment at 2 K, a value consistent with the standard deviation of the MODIS innovations. The surface air temperature impacts in this second experiment were also found to be neutral over the North America region in the range of +0.1 to –0.3%, which is very similar to the results in Figure 13. It was therefore concluded that the neutral impact over North America was not sensitive to the lower observation uncertainties supplied with MODIS.

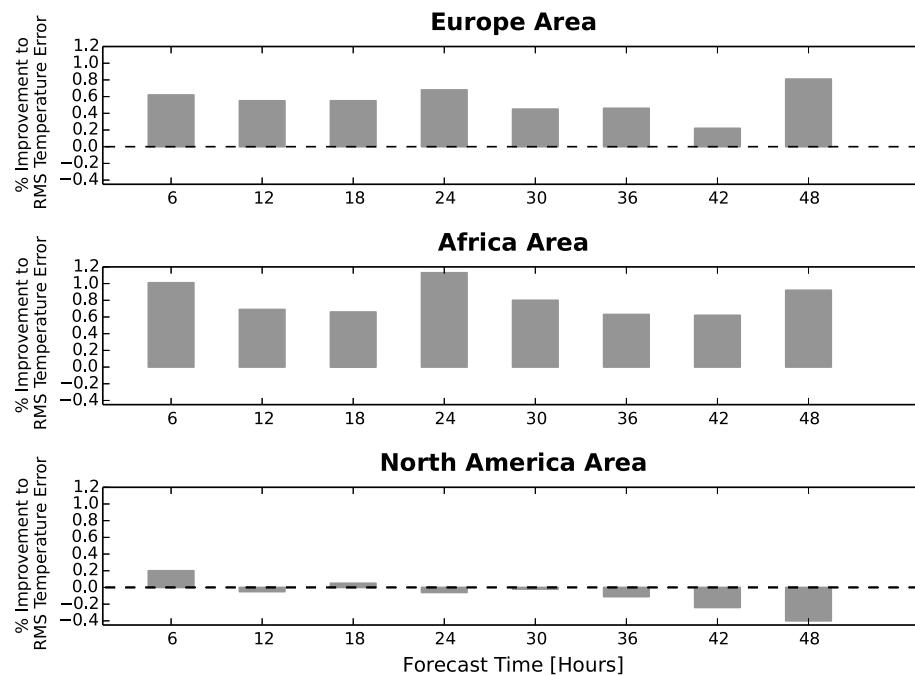


Figure 13. Bar chart showing the change in forecast error for surface air temperature between the experiment and control runs of the NWP trial suite. Results are shown for three geographical regions and as a function of forecast range.

Other forecast variables typically assessed during NWP trials include those representative of the general atmospheric circulation, such as pressure at mean sea level (PMSL), geopotential height, and upper level winds. Of these variables the largest changes were found in the PMSL field and are of relatively small and mixed impact. Figure 14 shows the changes to the PMSL forecast error for the African region, and it can be seen that overall the changes are beneficial. Figure 14 (top row) shows the change to the mean forecast error (verified by observations from meteorological stations as before), and in general, the NWP forecasts are smaller in magnitude than the verifying observations. The LST assimilation reduces this negative bias slightly across the entire forecast range out to 5 days. In addition, Figure 14 (bottom row) shows the RMS forecast error and in the forecast range interval of 1 to 3 days there is a slight reduction in this metric within the LST Experiment.

4.3. Discussion: Impact Over North America

The positive impact in the atmospheric forecasts over Africa is very encouraging; however, the neutral impact over the North American region requires further investigation. While we have noted that the impact may be smaller over certain surface types, due to the weaker coupling between LST and the soil layer, this region contains a mix of grasslands, forested regions, and desert. Consequently, we should expect a positive response to LST assimilation. However, unlike the African region, North America contains a very dense network of over 8,000 meteorological stations. The air temperature from these stations contributes via the EKF to soil temperature analyses in both the control and experiment runs. To examine the relative impact of the two observing systems, we undertook a case study at the end of the trial period for one analysis cycle. This looked at the effect of assimilating the meteorological stations and the satellite LSTs individually over North America. The time of the case was at 06 UTC; it is this analysis in the 6 h cycle when the region is covered by darkness and so the LST observations can contribute. Table 5 shows the resulting innovations and analysis increments in the first soil layer for the North American area (60°N – 25°N , 145°W – 50°W).

The table shows that both the station and satellite data have innovations (or fits to the NWP short range forecast) over North America that are similar to other regions (e.g., comparing with the SEVIRI innovations over Africa in Figure 12). The resulting soil temperature increments arising from each observing type are of a similar magnitude, noting that the effect of the meteorological stations is to cool the soil slightly. However, Figure 15 shows that for this assimilation cycle the LST data are confined to the eastern side of the continent and, for the most part, in regions that contain a high spatial density of meteorological stations.

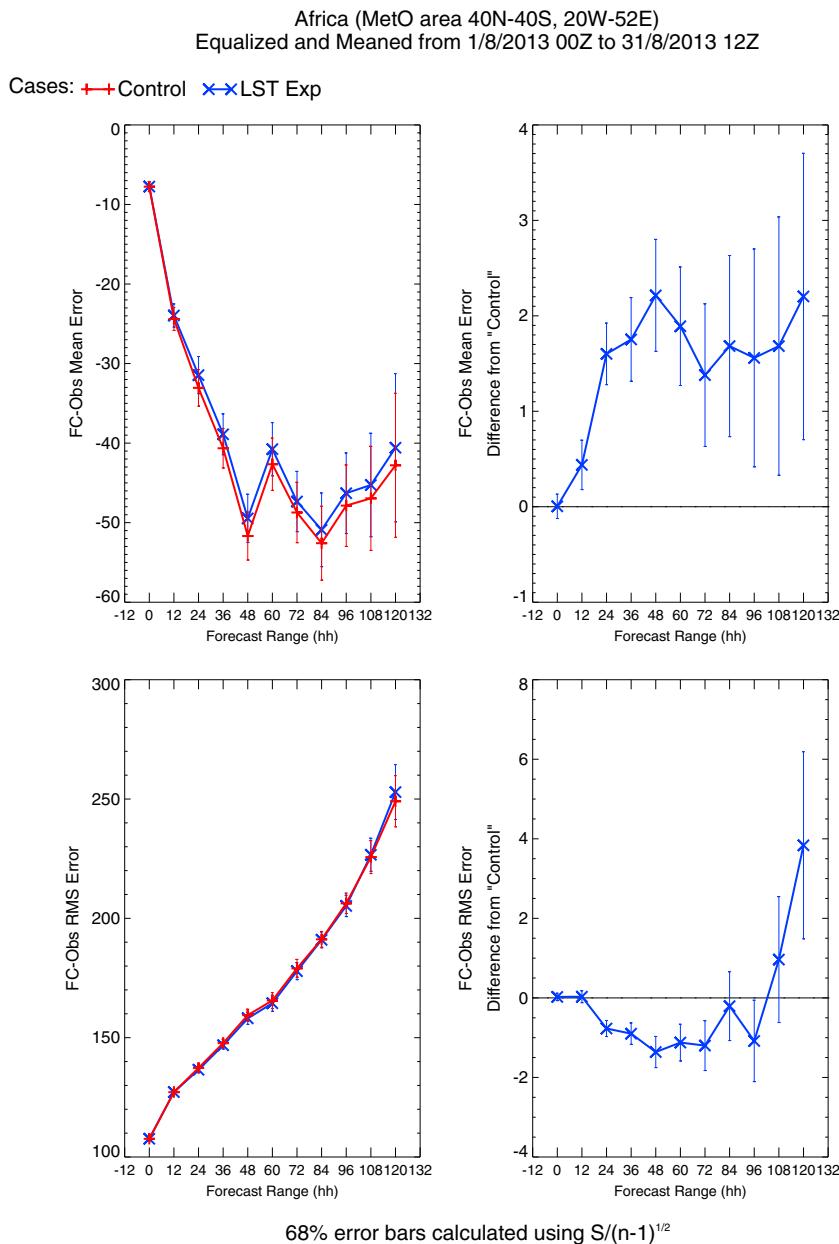


Figure 14. Forecast range plots of PMSL forecast errors (Pa) for the control and LST experiment using observations from meteorological stations across Africa. (top left) The bias and (top right) the difference in bias from the control. (bottom left) The RMS error and (bottom right) the difference in this from the control.

We suggest therefore that in order to obtain a improved forecast (compared to the control) for the near-surface air temperature over North America the spatial coverage of LST observations needs to be increased beyond what was used in the experiment run. This could be achieved in two ways: first, by increasing the length of the observation time window used in each analysis. Currently, this is set to 2 h centered on the

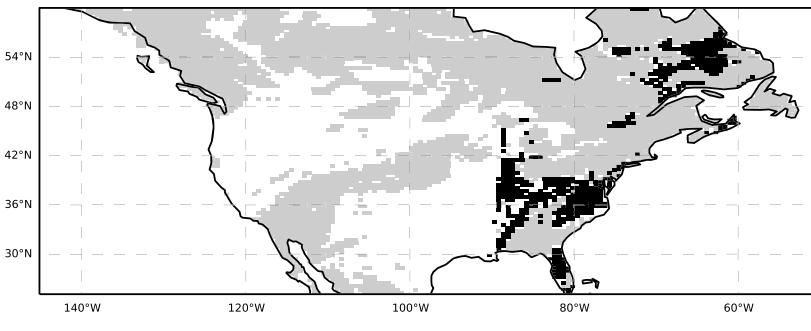
Table 5

Innovation and Soil Temperature Increments for the North American Case Study

Assimilation case	Meteorological stations	Aqua MODIS	Aqua MODIS + meteorological stations
Innovation Standard Deviation (K)	1.8	2.3	n/a
Soil Increment Mean (K)	-0.16	0.01	-0.14
Soil Increment Standard Deviation (K)	0.35	0.41	0.53

Note. Data time: 31 August 2013 06 UTC.

Locations of Gridded Satellite observations after Quality Control
Grey: model cloudy regions, Black: Satellite data after quality control



Locations of meteorological stations

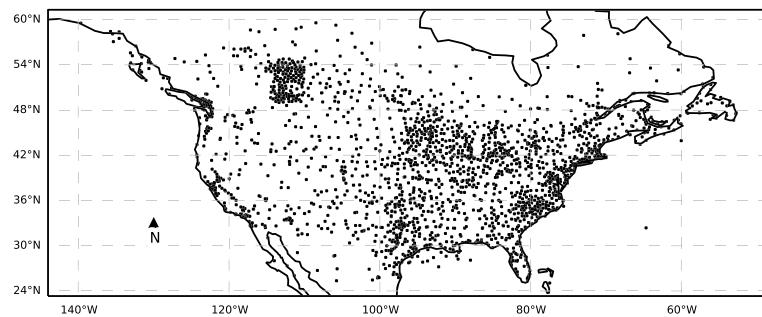


Figure 15. The location of the observations used in the North American case study. (top) Gridded Aqua MODIS LST observations and (bottom) meteorological stations.

analysis time. However, doing this may create representation issues with the EKF computation as the \mathbf{H} matrix sensitivity is currently computed at the analysis time for all observations, regardless of their actual observation time within the window. An alternative is to extend the amount of data through the addition of more geostationary observations of LST, in this case from a satellite centered over North America. In particular, this would give extended coverage into the region west of longitude 100°W, where there is a notable reduction in the spatial density of surface observations and consequently more potential for forecast impact in the atmospheric temperature.

5. Conclusions and Future Work

This work represents a pioneering study in the evaluation on the use of satellite land surface temperature (LST) within the Met Office land surface EKF assimilation system. To our knowledge it is also the first time that LSTs have been assimilated into a land assimilation system that is then used to provide the surface state for a NWP model. It allows us to investigate the impact in both the soil state and atmospheric forecasts. The approach to assimilate this new data source, which composes of observations from both geostationary and polar-orbiting satellites, has been guided through comparisons with short-range forecasts of LST from the operational global NWP model and an in situ radiometer. These showed that, in general, the smallest differences were during nighttime scenes, and comparisons with the in situ radiometer suggested at night the satellite LST uncertainties are less than 2 K.

Experiments were run within the NWP system to test the impact of night time satellite LSTs within the EKF using observations from SEVIRI (geostationary) and MODIS (polar-orbiting) instruments. The GlobTemperature data sets include uncertainties, and a novel feature of the NWP experiments was to construct the observation error (\mathbf{R}) matrix on a scene-by-scene basis. The \mathbf{R} matrix directly affects the weight given to the satellite observations within the EKF compared to the other sources, particularly the observations from meteorological stations. Over Africa, for instance, the satellite LSTs will have very low weight, compared to the station data in desert and other bare soil regions, but a much higher weight in grassland regions. It is important therefore that future satellite LST products contain meaningful uncertainty estimates and their use should be encouraged in data assimilation systems.

Results from the experiments showed that assimilation of GlobTemperature LSTs provided updated information for soil temperature in the first two soil levels, as expressed by an increase in the analysis increments within the EKF. Validation of the soil temperature arising from the EKF was undertaken using in situ measurements from several networks, and over the Iberian Peninsula a cold bias in the model soil temperature was reduced. Over Africa the short-range forecast of LST was also improved by up to 20% in some locations, notably over the desert and grasslands. An assessment of the impact on the atmospheric forecasts was also made, and the main impact found was a small, but consistent, reduction in forecast error for surface air temperature over Europe and Africa. Of the three regions studied, Africa showed the largest improvement due to LST assimilation. We suggest that this is probably due to the fact that meteorological stations, which are an existing source of observations for the EKF soil and surface temperature analyses, are very sparse in parts of this continent.

These results are a very encouraging first step toward the direct use of satellite-derived LSTs within the Met Office scheme, and the intention is to explore this work further with additional experiments. Two potential studies which would be beneficial are first to include daytime LSTs. Although the daytime satellite/model biases are large, with a potential dependency on viewing angle which could be challenging to correct in an operational scheme, there is the likelihood that these data could provide additional information on near-surface soil moisture as suggested by several studies. Since the use of LST in the EKF has been designed to accommodate additional instruments, through running the preprocessing in parallel for each observation type, a second useful study then would be to examine the impact when LSTs from additional satellites are assimilated. As a priority, focus should be given to geostationary satellites as they should greatly extend data coverage. The hourly or higher observing frequency in this orbit is particularly attractive for assimilation as it increases the likelihood of clear observations at each analysis time. A suitable choice for such an experiment would be LSTs generated from the imager onboard the Geostationary Operational Environmental Satellite (GOES) over North and South America (Freitas et al., 2013). Extended use of polar orbiting data with differing overpass times to the Aqua satellite, for example, from the recently launched sea and land surface temperature radiometer on Sentinel-3, will also help to improve coverage at high latitudes.

Acknowledgments

The work described in this paper was funded by ESA as a user case study for the GlobTemperature project. The satellite LST observations are archived at <http://www.globtemperature.info>. James McGregor and Jeremy Price of the Met Research Unit provided access to the Met Office Cardington in situ data. Available from the CEDA archive at <http://catalogue.ceda.ac.uk>. Soil temperature from the European continental networks were made available via the International Soil Moisture Network website (<https://ismn.geo.tuwien.ac.at>), an initiative supported by ESA and the Global Energy and Water Cycle Experiment (GEWEX). Forecast data sets of air and soil temperature from the assimilation experiments are available upon request from the first author. These are stored in NetCDF format. The authors also wish to thank Richard Renshaw (Met Office) and the anonymous referees who gave valuable comments on the paper.

References

Albergel, C., Dutra, E., Munoz-Sabater, J., Haiden, T., Balsamo, G., Beljaars, A., ... Wedi, N. (2015). Soil temperature at ECMWF: An assessment using ground-based observations. *Journal of Geophysical Research: Atmospheres*, 120, 1361–1373. <https://doi.org/10.1002/2014JD022505>

Balsamo, G., Beljaars, A., Scipal, P., Viterbo, P., van den Hurk, B., Hirschi, M., & Betts, A. K. (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *Journal of Hydrometeorology*, 10(3), 623–643.

Balsamo, G., Mahfouf, J. F., Bélair, S., & Deblonde, G. (2007). A land data assimilation system for soil moisture and temperature: An information content study. *Journal of Hydrometeorology*, 8(6), 1225–1242.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R., Ménard, C. B., & Mercado, L. M. (2011). The Joint UK Land Environment Simulator (JULES), model description—Part 1: Energy and water fluxes. *Geoscientific Model Development*, 4(3), 677–699. <https://doi.org/10.5194/gmd-4-677-2>

Bircher, S., Skou, N., Jensen, K. H., Walker, J. P., & Rasmussen, L. (2011). A soil moisture and temperature network for SMOS validation in western Denmark. *Hydrology and Earth System Sciences Discussions*, 8, 9961–10006. <https://doi.org/10.5194/hessd-8-9961>

Bulgin, C. E., Embury, O., Corlett, G., & Merchant, C. J. (2016). Independent uncertainty estimates for coefficient based sea surface temperature retrieval from the Along-Track Scanning Radiometer Instruments. *Remote Sensing of Environment*, 178, 213–222.

Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., and Piquet, B. (2007). In situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network, International Geoscience and Remote Sensing Symposium, IGARSS, Barcelona, Spain, 23–28 July 2007, 1196–1199. <https://doi.org/10.1109/IGARSS.2007.4423019>

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>

Chen, F., & Dudhia, J. (2001). Coupling an advanced land-surface/hydrology model with the Penn State/NCAR MM5 modeling system. Part I: Model description and implementation. *Monthly Weather Review*, 129(4), 569–585.

Clayton, A. M., Lorenc, A. C., & Barker, D. M. (2012). Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quarterly Journal of the Royal Meteorological Society*, 139(675), 1445–1461.

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., & Isaksen, L. (2013). A simplified extended Kalman filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 139(674), 1199–1213. <https://doi.org/10.1002/qj.2023>

Dharssi, I., Bovis, K., Macpherson, B., & Jones, C. (2011). Operational assimilation of ASCAT soil wetness at the Met Office. *Hydrology and Earth System Sciences Discussions*, 15, 2729–2746. <https://doi.org/10.5194/hess-15-2729-2011>

Dharssi, I., Candy, B., Bovis, K., Steinle, P., and Macpherson, B. (2013). Analysis of the linearised observation operator in a land surface data assimilation scheme for numerical weather prediction, In J. Piantadosi, R. S. Anderssen, & J. Boland (eds.), MODSIM2013, 20th International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand (pp. 2868–2862). December 2013, ISBN: 978-0-9872143-3-1.

Dharssi, I., Vidale, P. L., Verhoef, A., Macpherson, B., Jones, C., & Best, M. (2009). New soil physical properties implemented in the Unified Model at PS18. Forecasting Research Technical Report 528, Met. Office, UK. [Available at <http://www.metoffice.gov.uk/learning/library/publications/>]

Di Giuseppe, F., Cesari, D., & Bonafe, G. (2011). Soil initialization strategy for use in limited-area weather prediction systems. *Monthly Weather Review*, 139(6), 1844–1860.

Donlon, C. J., Martin, M., Stark, J. D., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2011). The operational sea surface temperature and sea ice analysis (OSTIA). *Remote Sensing of Environment*. <https://doi.org/10.1016/j.rse.2010.10.017>

Dorigo, W. A., Wagner, W., Hohnsinn, R., Hahn, S., Paulik, C., Xaver, A., ... Jackson, T. (2011). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, 15(5), 1675–1698. <https://doi.org/10.5194/hess-15-1675-2011>

Douville, H., Viterbo, P., Mahfouf, J.-F., & Beljaars, A. C. M. (2000). Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using FIFE data. *Monthly Weather Review*, 128(6), 1733–1756.

Fan, X. (2009). Impacts of soil heating condition on precipitation simulations in the weather research and forecasting model. *Monthly Weather Review*, 137(7), 2263–2285.

Freitas, S. C., Trigo, I. F., Bioucas-Dias, J. M., & Gottsche, F.-M. (2010). Quantifying the uncertainty of land surface temperature retrievals from SEVIRI/Meteosat. *IEEE Transactions on Geoscience and Remote Sensing*, 48(1), 523–534.

Freitas, S. C., Trigo, I. F., Macedo, J., Barroso, C., Silva, R., & Perdigão, R. (2013). Land surface temperature from multiple geostationary satellites. *International Journal of Remote Sensing*, 34(9–10), 3051–3068.

Ghent, D. (2012). Land surface temperature validation and algorithm verification, Report to European Space Agency, UL-NILU-ESA-LST-VAV.

Ghent, D., Kaduk, J., Remedios, J., Ardö, J., & Balzter, H. (2010). Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. *Journal of Geophysical Research*, 115. D19112. <https://doi.org/10.1029/2010JD014392>

Ghent, D., Trigo, I., Pires, A., Sardou, O., Bruniquel, J., Gottsche, F., ... Remedios, J. (2016). ESA DUE GlobTemperature product user guide V2. [Available from <http://www.globtemperature.info>]

Guedj, S., Karbou, F., & Rabier, F. (2011). Land surface temperature estimation to improve the assimilation of SEVIRI radiances over land. *Journal of Geophysical Research*, 116. D14107. <https://doi.org/10.1029/2011JD015776>

Lorenc, A. C., & Hammon, O. (1988). Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. *Quarterly Journal of the Royal Meteorological Society*, 114(480), 515–543. <https://doi.org/10.1002/qj.49711448012>

Martin, G. M., Milton, S. F., Senior, C. A., Brooks, M. E., & Ineson, S. (2010). Analysis and reduction of systematic errors through a seamless approach to modeling weather and climate. *Journal of Climate*, 23(22), 5933–5957. <https://doi.org/10.1175/2010JCLI3541>

Noyes, E. J., Corlett, G. K., Remedios, J. J., Kong, X., Llewellyn-Jones, D. T. (2007). An accuracy assessment of AATSR LST data using theoretical and empirical methods, In Proceedings of the Second Envisat Symposium, 23–27 April 2007, Montreux, Switzerland, ESA Publication SP-636.

Pavelin, E. G., & Candy, B. (2014). Assimilation of surface-sensitive infrared radiances over land: Estimation of land surface temperature and emissivity. *Quarterly Journal of the Royal Meteorological Society*, 140(681), 1198–1208. <https://doi.org/10.1002/qj.2218>

Pullen, S., Jones, C., & Rooney, G. (2011). Using satellite-derived snow cover data to implement a snow analysis in the Met Office global NWP model. *Journal of Applied Meteorology and Climatology*, 50(5), 958–973.

Seemann, S. W., Borbas, E. E., Knuteson, R. O., Stephenson, G. R., & Huang, H.-L. (2008). Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multispectral satellite radiance measurements. *Journal of Applied Meteorology and Climatology*, 47(1), 108–123.

Trigo, I. F., Monteiro, I. T., Olesen, F., & Kabsch, E. (2008). An assessment of remotely sensed land surface temperature. *Journal of Geophysical Research*, 113. D17108. <https://doi.org/10.1029/2008JD010035>

Walters, D. N., Williams, K. D., Boulte, I. A., Bushell, A. C., Edwards, J. M., Field, P. R., ... Lock, A. P. (2014). The Met Office Unified Model global atmosphere 4.0 and JULES global land 4.0 configurations. *Geoscientific Model Development*, 7, 361–386. <https://doi.org/10.5194/gmd-7-361-2014>

Zacharias, S., Bogena, H. R., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., ... Frenzel, M. (2011). A network of terrestrial environmental observatories in Germany. *Vadose Zone Journal*, 10, 955–973. <https://doi.org/10.2136/vzj2010.013>