Daily minimum and maximum surface air temperatures from geostationary satellite data

Elizabeth Good

Met Office Hadley Centre, FitzRoy Road, Exeter, Devon, EX1 3PB, United Kingdom

Tel: +44 (0)1392 884870, Fax: +44 (0)1392 885681

Email: elizabeth.good@metoffice.gov.uk

Key Points

- Geostationary satellite data are used to estimate daily Tmin/Tmax over Europe
- A dynamic empirical multiple-linear regression model is used
- The majority of estimated Tmin/Tmax are within 3-4 deg C of station data

Abstract

Observations of daily minimum and maximum land air temperatures, Tmin and Tmax, have traditionally been obtained through in situ observations at meteorological stations. While the station network is extensive, many land masses are poorly observed. Moreover, observations at stations are ‘point’ observations and may not be representative of air temperatures at neighbouring locations. Satellites provide the means to observe surface skin temperatures at spatial scales of tens of metres to kilometres. But although skin and near-surface air temperatures may be strongly coupled, the two quantities can differ by several degrees over land, where the magnitude of the difference is variable in both space and time. This study describes a method for estimating daily Tmin and Tmax at the pixel scale using geostationary satellite data, providing spatially-detailed observations for areas unobserved in situ. A dynamic multiple linear regression model is developed using daily minimum and maximum land surface temperature (LSTmin and LSTmax), fraction of vegetation, distance from coast, latitude, urban fraction and elevation as predictors. The method is demonstrated over Europe for 2012-2013; evaluation with collocated station observations indicates a mean satellite-minus-station bias of 0.0 to 0.5 °C with root-mean-square difference of 2.3 to 2.7 °C. The data derived here are not designed to replace traditional gridded station air temperature data sets, but to augment them.

Satellite surface temperature data usually have larger uncertainties than in situ data sets, but they can offer spatial detail and coverage that the latter may not provide.

Index Terms

3360 Remote sensing (4337)

0350 Pressure, density, and temperature

3322 Land/atmosphere interactions (1218, 1631, 1843, 4301)

0394 Instruments and techniques

Keywords

Land surface temperature, Near-surface air temperature, SEVIRI, Europe, multiple-linear regression model

1 Introduction

Near-surface air temperature (NSAT) is one of the key variables required by meteorologists and climate scientists, and is listed as one of the 50 Global Climate Observing System (GCOS) Essential Climate Variables (ECV). Over land, NSAT is observed in situ at meteorological stations. These data are typically aggregated onto grids that are widely used by the climate science community for model evaluation, monitoring and assessing climate change. Well-known examples include the CRUTEM data set (*Jones et al.*, 2012), which provides global land monthly temperature anomalies from 1850 to the present, the Goddard Institute for Space Studies Surface Temperature Analysis (GISTEMP), a global monthly anomaly time series from 1880 to the present (*Hansen et al.*, 2010), and the National Climatic Data Center (NCDC) gridded mean temperature anomaly data set that also extends back to 1880 (*Smith et al.*, 2008). Data sets at daily resolution include E-OBS, a gridded analysis for European land that includes actual mean, maximum and minimum temperatures (Tmean, Tmax and Tmin) from 1950 (*Haylock et al.*, 2008) and HadGHCND (*Caesar et al.*, 2006), which is a global land anomaly data set for Tmax and Tmin that also extends back to 1950.

These data sets provide invaluable information about Earth surface temperatures but suffer from gaps and/or high uncertainties where station density is low. Many such areas are in regions where the effects of climate change are predicted to be strongest, such as Africa and the high latitudes, and therefore a current research challenge is to seek ways to fill these gaps and reduce uncertainties.

Using surface temperature estimates from satellites is one possible solution to this problem and this approach is already well established for sea surface temperature (SST) analyses (*Donlon et al.*, 2012; *Rayner et al.*, 2003; *Reynolds et al.*, 2002). Efforts to produce similar analyses over land are not well developed for two main reasons. Firstly, satellite land surface temperature (LST) data sets are immature compared with SST. Land is heterogeneous in terms of land cover type, elevation and the overlying atmosphere and this makes estimating satellite LSTs very challenging as these effects must be accurately taken into account in any retrieval scheme [*Becker and Li*, 1990; *Dash et al.*, 2001; *Hulley and Hook*, 2010, 2012; *Hulley et al.*, 2012; *Inamdar et al.*, 2008; *Li et al.*, 2013; *Trigo et al.*, 2008; *Sun and Pinker*, 2007]. Secondly, satellites observe the Earth's 'skin' temperature, which is fundamentally different from the temperature usually measured in situ. This necessitates that an observational alignment is performed before the two data types can be combined. Over the ocean, satellite skin SSTs are adjusted to a SST at depth before they are blended with subsurface data from buoys and ships (*Fairall et al.*, 1996). Over the land, the analogous conversion of LST to NSAT is not as well understood. The NSATs measured at ground stations can differ by several degrees with respect to the underlying skin temperature, where the magnitude and sign of the difference varies in both space and time (Figures 1 and 2). The variation is a function of several factors, including time of day, meteorology, surface type, geographical location and elevation, and is therefore difficult to quantify. Additionally, land temperatures are spatially heterogeneous, which makes combining the point observations from meteorological stations with the areal averages observed by satellites inherently difficult.

Recent years have seen significant improvements in satellite LST retrieval techniques and a variety of LST products are now available to users, with many updated in real time. LST products from polar orbiting sensors include the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra (1999-present) and Aqua (2002-present) satellite platforms [*Wan and Dozier*, 1996; *Wan*, 2014] and the Along-Track Scanning Radiometer onboard the ERS and Envisat platforms (1991-2012) [*Prata*, 1993, 1994, 2002]. These products provide 'snapshot' LST images once or twice daily at 1 km or higher spatial resolution. LST products are also available from geostationary sensors including the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard Meteosat Second Generation (MSG; 2004-present) [*Freitas et al.*, 2010, 2013; *Trigo et al.*, 2008] and Geostationary Operational Environmental Satellites (GOES) [*Freitas et al.*, 2013; *Heidinger et al.*, 2013; *Sun et al.*, 2012]. These data provide LST images every 15-60 minutes but at a lower spatial resolution compared with polar-orbiters; for example, SEVIRI has 3 km spatial resolution at the sub-satellite point. Efforts are also underway by several data providers to process historical satellite data to LST, which will lead to multi-decadal data sets that could be suitable for climate applications. This includes the European Space Agency (ESA) GlobTemperature project (http://www.globtemperature.info/), the German Aerospace Center (DLR) TIMELINE project (http://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-9035/15754_read-38904/), and the Eumetsat Satellite Application Facility on Climate Monitoring (CM SAF; http://www.meteoswiss.admin.ch/web/en/research/current_projects/climate/cmsaf.html). Long-term LST data at 0.1° latitude/longitude are also available over N. America from PATMOS-x (http://cimss.ssec.wisc.edu/patmosx/data/) [*Heidinger et al.*, 2013]. These developments have led to increasing interest in satellite LSTs for weather and climate applications and as a result, efforts to make use of these data sets have intensified.

A number of studies have recently been published describing methods for estimating NSAT over land from satellite data, which is also useful as a first step toward achieving a blended satellite-station temperature analysis over land [*Basist et al.*, 1998]. The majority of these studies utilise empirical models, where NSAT is predicted from satellite LST and other observed variables, most commonly a vegetation metric such as the Normalised Difference Vegetation index (NDVI), and auxiliary data.

For example, *Nieto et al.* [2011] adopt the Temperature Vegetation Index (TVX) method to estimate day time NSAT over Spain during 2005 using data from the MSG, achieving accuracies of around 3-5 °C. The TVX method assumes NSAT approaches LST with increasing NDVI, such that the two are equivalent at full canopy cover. An issue with this method is establishing the value of NDVI for full canopy cover: *Nieto et al.* [2011] offer improvements on previous studies that use the TVX method by defining different NDVI values for different vegetation types. In another study, *Kawashima et al.* [2000] use spatially-averaged LSTs and NDVI derived from Landsat to predict NSAT over Japan for four winter nights between 1984 and 1986 with a reported accuracy of better than 1 °C.

Other studies present regressions that use predictors in addition to LST and vegetation. For example, *Vancutsem et al.* [2010] use LSTs from the MODIS together with the NDVI and solar zenith angle (SZA) to estimate Tmin and Tmax for 28 stations in Africa for various overpasses occurring between 2002 and 2008. They report a mean absolute error (MAE) of 1.73 °C and a standard deviation of 2.4 °C in their study. *Cristobal et al.* [2008] use latitude, distance from coast, altitude and solar radiation in addition to LST and NDVI over the Iberian Peninsula to estimate daily, monthly and annual mean, minimum and maximum NSAT (Tmean, Tmin and Tmax) between 2000-2005 using Landsat, MODIS and National Oceanic and Atmospheric Administration (NOAA) satellite data. They achieve an averaged root mean square error (RMSE) of 1.75 °C for daily NSATs and 1.00 °C for monthly and annual temperatures.

Several studies report success in estimating NSAT without using vegetation parameters. *Zhang et al.* [2011] use MODIS LST data to calculate daily Tmean, Tmax and Tmin over China during 2003. They find that improved results are obtained by including solar declination in their regression, and report uncertainties of less than 3 °C for 60-85% of their estimated NSATs.

Several studies have utilised kriging to predict NSAT. *Florio et al.* [2004] compare the traditional regression-based approaches with three kriging methods over N. America using data from the Advanced Very High Resolution Radiometer (AVHRR) for six days during 2000. They conclude that the kriging methods are superior to regression, achieving average errors of 0.9 °C and 1.4 °C, respectively. Recently, *Kilibarda et al.* [2014] used spatio-temporal regression kriging with in situ data using 8-day LST composites from MODIS together with topography to predict global daily NSAT during 2011. *Kilibarda et al.* [2014] report RMSEs in their method of around 2 °C for station-dense regions, 2-4 °C for station sparse regions, with RMSE of up to 6 °C for Antarctica. *Chen et al.* [2014] compare conventional regression- and kriging-based estimates of NSAT with a new geographically-weighted regression approach over China using MODIS LSTs for 2010. In this study, Tmean, Tmin and Tmax are estimated with 8-day and monthly temporal resolution using all three methods. *Chen et al.* [2014] find their new approach achieves an RMSE of 0.8-1.5 °C, and outperforms the conventional regression and kriging methods. They also find that errors are lower for monthly than for 8-day estimates.

An interesting physically-based method has also been proposed by *Sun et al.* [2005], who derive NSAT using thermodynamics. They define an equation for NSAT that requires inputs of LST, NDVI, net radiation, aerodynamic resistance and crop water stress index (CWSI). *Sun et al.* [2005] trial their method for two MODIS overpasses in 2002 over a 3x9 degree region on the North China Plain to estimate instantaneous NSAT and report an accuracy of better than 3 °C for 80% of the estimated NSATs.

Most studies reported in the literature use data from infrared sensors, but there are also examples of studies that use microwave data. Unlike infrared sensors, microwave sensors have the advantage of providing observations during cloudy conditions. However, the spatial resolution is much lower (several tens of km vs a few km or better), which is a disadvantage for many applications. Early in the satellite era, *Davis and Tarpley* [1983] used a multiple linear regression model to estimate daily NSAT from base atmospheric temperature, skin temperature and microwave channel-2 (from the Microwave Sounding Unit, MSU) top of atmosphere (TOA) brightness temperatures derived from the TIROS Operational Vertical Sounder (TOVS) system. They obtained uncertainties of around 1.6-2.6 °C under cloud-free and partially cloudy conditions and 2.9-4.0 °C under cloud conditions (microwave data only). Later, *Basist et al.* [1998] used data from the seven channels of the Special Sensor Microwave/Imager (SSM/I) to estimate NSAT with a reported standard error of approximately 2.0 °C. The authors also trial a blended station-satellite gridded temperature product and demonstrate this offers improved performance over one derived from station data alone.

Jones et al. [2010] use the Advanced Microwave Scanning Radiometer (AMSR-E/Aqua) channels at 18.7 and 23.8 (both polarisations) to estimate daily Tmin and Tmax at 25 km spatial resolution over ice and snow-free Northern hemisphere land masses. *Jones et al.* [2010] obtain an RMSE of 1-4 °C, although find uncertainties can exceed this over sparsely vegetated, desert, and mountainous regions. *Köhn and Royer* [2012] also use AMSR-E to estimate instantaneous NSAT over Northern Canada over two winters using an empirical model and obtain an RMSE of 4.9 °C.

More recently, *Jang et al.* [2014] estimate instantaneous all-sky NSAT by using data derived from MODIS and AMSR-E onboard Aqua through empirical linear regression. Rather than using MODIS LST retrievals directly, *Jang et al.* [2014] use temperature profiles from the MODIS atmospheric product to derive NSAT for clear-sky conditions and augment these with estimates of NSAT from AMSR-E under cloud. They obtain an RMSE of 4.5 to 5 °C on evaluating their data set with NSAT observations from forty US stations.

There have also been several studies that document methods to estimate ocean NSAT using satellite data [*Jackson et al.*, 2006; *Roberts et al.,* 2010; *Shi et al.*, 2012].

Of the land-based studies described above, only the *Jones et al.* [2010] data are publicly available in digital form and a global version of the data set covering June 2002 to September 2011 is available from http://nsidc.org/data/nsidc-0451. Most of these studies have been limited to discrete periods of time or specific satellite overpasses for a specific region or country. This paper describes a new study where daily Tmin and Tmax are estimated at the satellite native field of view (FOV) using a multiple linear regression model. The work carried out here differs from the other infrared regression-based studies described above in that the method is applied to the whole of Europe and that maximum and minimum LST (LSTmin and LSTmax) are estimated from geostationary data before a regression is performed. In addition, an urban fraction predictor is included in the regression for the first time and the model regression coefficients are calculated on a daily basis in order to improve the prediction of NSAT. Until now, MODIS has been the sensor of choice in nearly all studies of this nature. The benefits of using geostationary data include providing sub-daily observations that potentially resolve the diurnal cycle, required for many applications, and increasing the pool of available satellite observations. (Nearly all current satellite LST data sets utilise infrared observations, which are only available under cloud-free conditions, which can result in large data gaps.). Additionally, geostationary platforms provide some of the longest satellite data records: for example, viable Meteosat observations extend back to the early 1980s. This means that the method developed here could eventually be used to generate a global, multi-decadal satellite, sub-daily NSAT data set. Europe is selected as the study region as it is generally well observed in situ and so presents a good opportunity to test and develop the methodology. Data are presented here for the years 2012-2013 and are freely available to download in NetCDF format at http://www.metoffice.gov.uk/hadobs/msg_tmaxmin/. The data will be updated regularly.

2 Data Sets

2.1 Satellite data

The data used here are from the SEVIRI onboard the geostationary MSG platforms, which have been the operational 'weather' satellites for Europe since 2004. The MSG-1 was launched in August 2002, and was re-designated Meteosat-8 in January 2004 when it became operational. MSG-2 and MSG-3 then followed in December 2005 and July 2012, respectively. SEVIRI has twelve spectral channels covering visible and infrared wavelengths and provides full-disc images every 15 minutes with a spatial resolution at the sub-satellite point (0 degrees longitude/latitude) of 3 km (1 km for the high-resolution visible channel).

An operational LST product is provided by the Land Surface Analysis Satellite Applications Facility (LSA SAF) in near-real time at the full space/time resolution of the SEVIRI. The product is available from 2009 and can be obtained by users through the LSA SAF website (http://landsaf.meteo.pt/), or automatically in near-real-time through the EUMETCAST dissemination service. LSTs are calculated using a generalised split window (GSW) algorithm, following *Wan and Dozier* [1996], with estimated errors and quality information, including a cloud mask, provided in the product for each pixel. The GSW method utilises observations from the 'window' channels at approximately 11 and 12 μm to account for atmospheric effects, which vary with wavelength. Regression coefficients are sourced from a look up table based on satellite viewing angle, water vapour and NSAT. Surface emissivity is estimated using a geometrical model and Fraction of Vegetation Cover (FVC), another LSA SAF product, available at daily resolution. As for LST, FVC data are available at the full spatial resolution of the SEVIRI instrument from 2009 onwards from the LSA SAF website and EUMETCAST service. FVC is estimated from the weighted combination of the probability that an observation is bare soil or fully vegetated, where the probability is computed using a Bayesian model (see LSA-SAF Product User Manual for Vegetation Parameters, v2.1, 200, available from the LSA-SAF website). The data are expressed as values between 0 and 100%, with estimated errors and quality information given per pixel; the overall accuracy of the FVC product is expected to be within 10% for 70-75% of land pixels.

Pixel cloud information is sourced from the Satellite Application Facility to support NoWCasting and very-short-range forecasting (SAFNWC). The theoretical RMSE for the LST retrievals is typically less than 2 °C under cloud free conditions, with similar uncertainties obtained through in situ validation experiments [*Freitas et al.*, 2010; *Trigo et al.*, 2008]. However, undetected cloud can cause systematic errors in the LST retrievals that may exceed this.

2.2 Ancillary Data

The latitude, longitude and elevation for each SEVIRI pixel is available from the LSA-SAF static data products archive (http://landsaf.meteo.pt/). In this study, sub-pixel land use information was also considered using data from the European Space Agency's 300-m GlobCover data set [*Arino et al.,* 2008; http://due.esrin.esa.int/globcover/]. These data were used to estimate the urban and water fraction (UF and FoW) for each SEVIRI pixel by calculating the fraction of 300-m GlobCover pixels nominally contained within each SEVIRI pixel for each of these surface classes. The GlobCover data were also used to calculate the distance from coast (DfC) for each SEVIRI land pixel.

2.3 Station Data

In situ Tmin and Tmax were sourced from the European Climate Assessment and Dataset (ECA&D) [*Klein Tank et al.*, 2002]. This data set includes observations from 7862 meteorological stations throughout Europe, where more than half the station records are available for public download from the ECA&D website (http://eca.knmi.nl/). The data used here were from the archive of 'blend and update' station data. The blended station data were used in this analysis to ensure coverage during the analysis period used in this study (2012-2013) as some stations have a lag time of several months or more before they are updated by their contributing countries. In these cases, ECA&D 'blend and update' station records are in filled with data from the Global Telecommunication System (GTS) or other stations within 12.5 km distance and 25m elevation. This method means that some observations are duplicated in two or more station records. The source for each station observation is provided by ECA&D as metadata and this has been used in this analysis to remove any duplicate entries. Data from the public ECA&D archive were used to build the regression relationships in this study (example shown in Figure 3). Additional data from the non-public ECA&D archive were provided for this study by the Royal Netherlands Meteorological Institute (KNMI) for Germany and the UK. These additional data were used as independent validation data for the satellite-derived Tmin and Tmax produced in this study.

3 Estimation of near surface air temperatures from satellite data

The method selected for this study is a regression-based approach. This approach is attractive because it is fairly simple to implement and in theory, relationships can be derived in well-observed regions that could be applied to sparsely-observed regions. While kriging-based methods appear to give good results, they do not perform so well in poorly-observed regions (*Kilibarda et al.*, 2014). Physical models, such as that of *Sun et al.* [2005], are less practical as they require input data that are not readily available. LST data from the SEVIRI are used here to estimate daily LSTmin and LSTmax. These LST data are then regressed with FVC, latitude, UF, elevation and DfC against station Tmin and Tmax data collocated in both space and time. Separate regression models are produced for Tmin and Tmax; the derived regression coefficients are applied to every available SEVIRI LSTmin/max observation providing estimates of Tmin and Tmax where station data are absent. Tmin and Tmax are the dependent variables of choice in this study rather than instantaneous NSAT, for example, principally because the study area (-25 to +45°E) spans several time zones. Deriving a generalised regression model is therefore impractical as time-matched in situ and satellite temperature observations used to train the regression across the domain would be at different points in the diurnal cycle (in situ NSAT observations are typically at 00, 06, 12, 18 GMT, or hourly at best). Figure 4 shows a flow chart of the process used to estimate Tmin and Tmax from the SEVIRI data. Each stage of the process is explained in detail below.

3.1 Calculation of LSTmin and LSTmax

The 15-minute SEVIRI LST data from the LSA-SAF are processed to obtain the LSTmin and LSTmax for each pixel for each day. Observations are only considered between the local times of 23:00 and 08:00 for LSTmin, and 11:00 and 16:00 for LSTmax. These temporal windows are employed to ensure only realistic data are considered for each observation type, for example to avoid daily LSTmax data being sourced from night time hours when there are few or no day time SEVIRI observations owing to rejection because of cloud-contamination. Use of these temporal windows also achieves some consistency with conventional meteorological station observations of Tmin and Tmax. (Note: LSTmax occurs close to solar noon, which often occurs earlier than Tmax.) LST values are only recorded where the number of valid observations (cloudy and non-cloudy) exceeds 33 for the LSTmin window and 18 for the LSTmax window, which equates to approximately 90% valid observational coverage in each window. This is a necessary step as satellite data are prone to missing observations, for example due to data feed interruptions or corruption. For the LSA-SAF archive, 4.9% of slots (i.e. entire 15-minute full-disc images) are missing or corrupt for the years 2012-2013. The cloud mask provided within the LSA-SAF LST product is used to identify and exclude cloudy pixels from the analysis. No limit on the number of valid cloud-free observations in each window is employed in deriving LSTmin and LSTmax. However, the fraction of cloud-free observations in each window is recorded at this stage of the processing as additional user information.

3.2 Regression Formulation

The linear regression model is constructed by regressing the observed station Tmin/Tmax data (in °C) against temporally and spatially collocated LSTmin/LSTmax (in °C), fraction of green vegetation (FVC; in %), latitude (Lat; in degrees N), elevation (Z; in metres), urban fraction (UF; in %) and distance from coast (DfC; in metres):

$$T\min = \alpha_{\min} + \beta_{\min}.LST\min + \chi_{\min}.FVC + \delta_{\min}.Z + \phi_{\min}.UF + \varphi_{\min}.Lat + \gamma_{\min}.DfC \qquad (1)$$

$$T\max = \alpha_{\max} + \beta_{\max}.LST\max + \chi_{\max}.FVC + \delta_{\max}.Z + \phi_{\max}.UF + \varphi_{\max}.Lat + \gamma_{\max}.DfC \qquad (2)$$

These variables were selected as they are all strongly and significantly correlated with NSAT ($p << 0.05$; also see Figures 5b and 6b). Other variables, such as wind speed, solar zenith angle (SZA) and albedo were also considered but were not selected as they did not improve the regression model enough to warrant their inclusion. For SZA and albedo this is partly because these parameters are strongly co-linear with LST/latitude and FVC, respectively. For wind speed, the issue is also one of practicality as reanalysis or operational analyses provide the only viable source of wind speed data for locations without stations.

Satellite-station matchups are excluded from the regression where the DfC is less than 1 km, the fraction of water (FoW; see Section 2.2) within the SEVIRI pixel is more than 50%, or the elevation is more than 1500 m. These are excluded as the satellite LST retrievals are likely to be unreliable under these conditions and because satellite-station agreement is likely to be influenced by factors that cannot be accounted for easily (e.g. the moderating effect of water bodies on local air temperature, coastal winds and sharp changes in elevation and challenging landscape at very high elevations).

The regression analysis is performed for each day with a moving 11-day window such that each day has a different set of regression coefficients and offset. The moving window approach was selected for two reasons. Firstly, the relationship between NSAT and several of the predictor variables varies with time, both on seasonal and local timescales (Figures 5b and 6b). Secondly, the approach ensures that a large enough pool of data is available to perform a meaningful regression analysis, since the presence of cloud can significantly reduce the number of LST and FVC data available. The FVC data are less affected by cloud as the LSA-SAF product uses observations from previous days for cloudy pixels. Data from up to the preceding two days are used to fill gaps in the LSA-SAF FVC archive (10 daily files are missing between 2012 and 2013). A regression relationship is only formulated for situations with at least 400 valid matchups.

Figures 5 and 6 illustrate the regression analysis for 2012-2013 for Tmin and Tmax, respectively. The figures include the time series of regression coefficients using normalised (panel c; normalisation of variables achieved by subtracting the mean and dividing by the standard deviation) and non-normalised data. The coefficients for the normalised regression data allow the relative contribution of each predictor in the model to be assessed. The dominant predictor variable is found to be the satellite LST, which accounts for around 70-80% of the variance in air temperature. Latitude is the second strongest predictor. The other predictors have lower weighting but are nonetheless important in the regression model as they can have significant effect on the predicted air temperatures in certain regimes. For elevation, the mean correction, ΔT, is -0.4 ºC for both Tmin and Tmax (root mean square (RMS) 1.4 ºC). For DfC, the mean ΔT is -0.2 ºC for Tmin (RMS 0.7 ºC) and 0.0 ºC for Tmax (RMS 0.9 ºC). The UF correction ranges between -1.0 and 1.3 for Tmin and -2.5 and 1.3 ºC for Tmax in heavily urbanised areas (e.g. capital cities). The ΔT values for FVC range between approximately -5 and +6 ºC for Tmin and -2 and 8 ºC for Tmax. For latitude, the ΔT values are large in magnitude (40 ºC), but this is largely counteracted by the regression offset, which is typically between 10 and 30 ºC; if latitude is removed from the predictor variables used in the regression the offset reduces to a few degrees.

4 Evaluation of predicted near surface air temperatures

4.1 Analysis of regression residuals

The bottom panels in Figures 5 and 6 show the residuals of the multiple-linear regression for the central analysis day (i.e. day 6 in the 11-day moving window). Statistics for the residuals are only shown where at least 50 valid matchups occur to ensure the results are statistically robust. The results are slightly better for Tmax than for Tmin, with the latter generally exhibiting more noise in the median satellite-minus-station difference. In general, the Tmax 10th/90th percentiles are close to ±2.5 degrees, indicated by the dotted line on the plots, with the 5th/95th percentiles occurring around the ±4 °C mark. For Tmin, these limits are about 0.5 °C or so larger in magnitude. A slight reduction in variance is evident during the summer months, which is likely to be a result of increased data availability during this time due to less cloud. A larger data pool will permit a more robust regression and reliable statistics. Other than this, no seasonality is apparent in the results, which is to be expected given the moving 11-day analysis window used in the regression. The mean satellite minus station bias is 0.0 °C for both Tmin and Tmax, with RMS differences of 2.7 °C and 2.5 °C, respectively.

4.2 Independent station evaluation

Figures 9 and 10 show a summary of the comparison between collocated non-public ECA&D German & UK station data, and satellite LSTs and NSATs. The correlation between LSTmin and the station Tmin is notably lower than for the satellite Tmin vs station Tmin (Figure 7b). The comparison shows the satellite LSTmin data are typically cold-biased by about 5 °C with respect to the station data (Figure 7c), whereas the mean satellite-minus-station Tmin bias is 0.5 °C (Figure 7d). The satellite Tmin minus station Tmin differences also have much lower variance compared with the LSTmin comparison. Overall, the satellite-station Tmin bias is slightly worse than that obtained from the analysis of the regression residuals discussed above (0.5 °C compared with 0.0 °C), but with similar RMS difference (2.5 °C compared with 2.7 °C in Section 4.1).

Similar results are obtained for Tmax: the correlation between satellite Tmax and station Tmax is higher than for satellite LSTmax vs station Tmax (Figure 8b), and the agreement between the modelled NSATs is better than between the LSTmax and station Tmax data

(Figures 8c,d). The LSTmax minus station Tmax differences show strong seasonal variability where the difference is negative in winter months and more positive during the summer (Figure 8c). This seasonal variability is considerably reduced (nearly absent) in the satellite vs station Tmax data, which is encouraging and suggests that the dynamic approach adopted for the regression model is valid. In fact, the performance of the Tmax regression is actually quite good, with high proportion of matchups falling within the ±2.5 °C limits marked on the plot (dotted line), particularly during the spring to autumn months. The mean bias for satellite-station Tmax is 0.0 °C with a RMS difference of 2.3 °C, which is nearly the same as the results obtained from the residuals analysis (0.0 °C and 2.5 °C, respectively).

The data used for the independent station evaluation are from the non-public ECA&D archive. Only non-public data for Germany and the UK were available for this study and therefore the analysis is geographically limited. The agreement between these results and those obtained for the residuals analysis in Section 4.1, which is for the whole of Europe, is encouraging and suggests that similar uncertainties could be expected in other regions. However, Figure 3 shows that a high proportion of the stations used for the regression model training are located in Scandinavia and Central Europe, which may result in the regression coefficients being weighted to favour conditions in these regions.

5 Example data: March 2012 northern European heat wave

Figures 9-11 show NSAT corresponding to a short heat wave experienced by parts of northern Europe during March 2012. These Figures include the relevant SEVIRI Tmin and Tmax estimates from this study, station data from the ECA&D public archive, specifically the Paris Parc Montsouris station, and the data from E-OBS, which is a gridded version of the ECA&D described by *Haylock et al.* [2008]. The E-OBS data used here correspond to the 0.25 degree regularly gridded product. The line graph (Figure 9) shows that all three data sets display a similar temperature evolution during the heat wave, while the maps (Figures 10 and 11) show that the three data sets also exhibit similar spatial patterns. Examining Figures 10 and 11 highlight some important points. Firstly, the satellite data suffer from significant gaps due to cloud (white areas: e.g. the Alps and Poland) and secondly, the satellite data offer spatial detail that is not available from the station-based data set; for example, the Paris urban heat island is clearly visible in the SEVIRI data (Figure 11). The largest differences between the in situ and satellite data sets are apparent when looking at the diurnal temperature ranges (Tmax minus Tmin). Particularly large differences occur in some regions such as northern England, and northern Germany, the Netherlands and Belgium (Figure 10); in both cases the SEVIRI diurnal temperature range is lower than for the equivalent in situ data. The cause appears to be a low-biased SEVIRI Tmax in both areas. This is indicated by the higher station Tmax values corresponding to the cluster of ECA&D stations in the Netherlands, and therefore probably corresponds to situations where the geographically generalised regression approach adopted in this study has not worked well.

6 Discussion

The satellite NSAT data may have lower uncertainties than the evaluation performed in this and other studies suggest. In situ temperatures can only be measured at discrete points and can therefore only be representative of the air immediately around the station at the time of the observation. In many cases, the observation can be extrapolated beyond the measurement location with confidence, depending on the location of the station. Indeed, meteorological station locations are chosen with this in mind but choices are limited by practicality and therefore the measured temperatures may not be representative of even close-neighbouring locations. Satellite sensors such as the SEVIRI, on the other hand, view the radiation emitted by every point within their field of view and therefore the observed surface temperatures they provide are truly area-averaged. For SEVIRI, this area is several km (3 km at the satellite point of 0 deg latitude/longitude and around 4-10 km over Europe) and therefore a SEVIRI pixel temperature observation may be inherently different from an in situ observation located within the same SEVIRI pixel. This is illustrated in Figure 12, which shows the NSAT differences between two neighbouring stations in Sweden which are separated by only 1.26 km and 1 m elevation, but are both located within the same SEVIRI pixel. The mean differences are 0.1 and 0.3 °C for Tmin and Tmax, respectively, but daily differences of several °C are frequently observed. This NSAT heterogeneity is documented elsewhere as it can result in non-climatic changes in station temperature records if a station is relocated [*Yan et al*, 2010].

The issue of satellite spatial scale has also been investigated recently by *Sohrabinia et al.* [2014] who find that the strength of the relationship between station-observed NSAT in New Zealand and LSTs from MODIS is strongest when the 1 km MODIS data are spatially averaged, for example, over 5x5 pixels or more. *Sohrabinia et al.* [2014] attribute this to the lower internal variability of LST at larger scales. However, they find that the station NSAT-satellite LST relationship begins to deteriorate at scales larger than 25 km. This could have implications for derivation and evaluation of microwave estimates of NSAT and may, at least in part, explain why the uncertainties of the *Jones et al.* [2010], *Köhn and Royer* [2012] and *Jang et al.* [2014] NSATs are higher than obtained in most of the infrared-based studies (e.g. this study, *Cristobal et al.* [2008], *Kilibarda et al.* [2014], *Vancutsem et al.* [2010], *Zhang et al.* [2011]). In this study, the issue of scale has also been investigated by repeating the evaluation in Section 4 with averaged SEVIRI NSATs over 3x3, 5x5, 7x7, 9x9, 11x11 and 13x13 pixel blocks. Changes to the RMS differences for these pixel blocks are negligible (≤0.1 °C), although the tendency is for the RMS difference to decrease up to 7x7, beyond which the tendency is an increase. Changes to the magnitude of the Tmax mean bias are also <0.1 °C while the Tmin mean bias becomes slightly more negative with increasing pixel block size (e.g. a reduction of 0.4 °C is observed for the 13x13 pixel block. This reduction could be a result of the inclusion of an increasing number of colder rural pixels around an urban station, for example.

Spatio-observational discrepancies between the satellite and in situ station may also occur because of parallax effects resulting from the interaction between the SEVIRI line of sight (view angle) and topography. In the extreme case, for example, SEVIRI might view only one side of a mountain when the station is located on the other side. In this study, the regression formulation is only performed using a large pool of station data, which should minimise the uncertainties on the regression parameters that occur as a result of these problems as these satellite-station spatial mismatch errors are expected to be random. For the evaluation component of the study, using a large pool of data will help to reduce any effects on the mean bias but not the variance, which is likely to be inflated as a result of the issues discussed here.

Figure 13 shows an example of the satellite temperature maps for 21 August 2013, showing both LSTmin and LSTmax, and predicted Tmin and Tmax for this day. The temperature difference between the LST and NSAT images is marked, particularly for LSTmax/Tmax, where a change of several degrees can be observed for some places – for example, Spain, where the LSTmax is more than 5 °C warmer. Also of note are the (grey) gaps in the satellite data, for example in north-eastern Europe. This is where SEVIRI observations of either LSTmin/LSTmax or FVC are not available owing to cloud obscuring the satellite view of the surface. Cloud contamination is a major disadvantage of using infrared and visible satellite data for surface temperature applications. Not only does it lead to gaps in the data coverage, but undetected cloud may lead to large errors in the surface temperature retrievals. The effects of the coverage problem are minimised in this study through the use of multi-hour temporal windows for estimating LSTmin and LSTmax. This increases the probability of obtaining a cloud-free observation for each temperature. In the case of LSTmax, the chance of using an observation contaminated with undetected cloud is also reduced as the maximum LST in the window is the observation most likely to be cloud-free as cloud almost always results in a cold-bias in the retrieved LST. For Tmin, however, the effect is the opposite with the minimum LST being that most likely to be contaminated with undetected cloud. This may, at least in part, explain why the evaluation results for satellite Tmin data are not as good as for Tmax. This sensitivity of Tmin to cloud-contamination biases could lead to some seasonal variation in bias at locations where there is a significant seasonal cycle in cloud fraction.

Figures 14 and 15 illustrate the cloud problem. The figures show that fewer day time observations are available due to presence of cloud compared with night (Figure 14 and 15b). This may be a real effect, or a result of the SEVIRI cloud detection methods, which include visible channel checks that cannot be carried out at night leading to increased incidences of missed cloud at night and larger errors in LST (cloud detection using infrared is carried out both at night and during the day). As expected, incidences of cloud-free observations are higher in southern Europe, dropping off to around 30% coverage above a latitude of 50 degrees (Figure 15a). The availability of cloud-free observations is highest during the summer (typically 40-70% depending on time of day) and lowest during northern hemisphere winter months (20-30%).

6.1 Extension of the method in both space and time

In addition to the SEVIRI, public geostationary LST data sets are currently available for GOES [*Freitas et al.*, 2013; *Heidinger et al.*, 2013; *Sun et al.*, 2012] and the Multi-Functional Transport Satellite (MTSAT) [*Freitas et al.*, 2013]. Although geostationary satellite records extend for several decades, these data sets are currently short in length and only cover the past 10 years or less, with the exception of the *Heidinger et al.*, [2013] data, which extends back to 1995 over N. America. Previous studies also document efforts to derive LST from other geostationary sensors. For example, *Prata and Cechet* [1999] estimate LST from the Visible and Infrared Spin Scan Radiometer (VISSR) onboard the Geostationary Meteorological Satellite 5 (GMS-5), which is the predecessor to the MTSAT. *Tang et al.*, [2008] use the generalised split window of *Wan and Dozier* [1996] to estimate LST from the geostationary FengYun Meteorological Satellite (FY-2C). Effort is also underway by the CM-SAF to develop multi-decadal LST from predecessors of the SEVIRI, onboard the Meteosat First Generation (MFG) platforms (http://www.meteoswiss.admin.ch/web/en/research/current_projects/climate/cmsaf.html).

The method described in this study could, in theory, be applied to LST records derived from other geostationary sensors to create a global, multi-decadal satellite, sub-daily NSAT data set. However, generating such a data set would pose several challenges. Firstly, the intercalibration and homogeneity of different LST data sources would need to be addressed as spectral characteristics and calibration vary between instruments and also with time. For example, GOES I-L images were equipped with split-window channels where as the more recent GOES M-P do not have a 12 μm channel, prohibiting the use of a split-window retrieval to generate long-term LST [*Heidinger et al.*, 2013]. Changes in instrument spectral characteristics will also affect the consistency of cloud detection techniques. LST also varies with viewing geometry such that different geostationary sensors may observe a different LST for the same Earth location [*Freitas et al.*, 2013; *Vinnikov et al.*, 2012]. Secondly, the long-term homogeneity and spatial consistency of some of the other predictors used in the regression described in this study will be an issue. In particular, the FVC data are derived from SEVIRI and are not global, thus alternative sources of data would be required for other sensors. The GlobCover land use map adopted here to derive UF and DfC is global but derived from satellite data acquired during 2005/6 and may not accurately represent urban areas during the early part of the satellite era. Thirdly, the characteristics of the station network will impact the application of this method to other regions. Derivation of the regression coefficients requires the availability of training station data. While it may be possible to construct the regressions in well-observed areas that could be applied to less-well observed areas with similar meteorology and land use, this is yet untested and will be investigated in future work. It should be noted that the dynamic formulation of the regression presented here does not require stations to be consistently available over time. However, inhomogeneities in station data records may propagate through the regression process and result in inhomogeneities in the derived satellite NSATs.

6.2 Data set applications

A global satellite NSAT data set, such as the one proposed above, is unlikely to replace conventional in situ data sets. The primary application for these data will be in providing new observations of Tmin and Tmax for regions that are poorly, or unobserved in situ, e.g. to compare with model-estimated NSAT, as in independent data source, or incorporated into in situ-based NSAT analyses to reduce uncertainties. The impact of incorporating additional data sources in data-sparse regions has recently be demonstrated by *Jones et al.* [2012], who find that including new Russian stations in the CRUTEM4 data set results in changes in estimates of recent high-latitude warming. Using satellite estimates of NSAT in this way will be explored further in a new Horizon2020 project, 'EUSTACE' (EU Surface Temperatures for All Corners of Earth; http://www.theclimatesymposium2014.com/ClimateSymposiumNickRaynerEUSTACEfinal.pdf), which aims to produce a globally complete daily NSAT analysis back to 1850 using a combination of satellite and in situ observations. A component of EUSTACE is to develop the method presented in this study further and apply it to a larger space and time domain; these data will then be statistically combined with in situ observations to produce the analysis over land.

As demonstrated in Section 5, and Figures 10 and 11, satellite NSATs can also provide spatial detail that cannot be observed in situ, but is required for some applications, e.g. urban temperature studies [*Dousset et al.*, 2011; *De Ridder et al.*, 2012]. Existing Tmin and Tmax observational analyses are currently available at comparatively coarse resolution, for example, E-OBS (0.25° latitude/longitude) and HadGHCND (2.5° latitude by 3.75° longitude); satellite NSAT data, such as those presented in this study will therefore provide an alternative data source that may better meet the needs of some users.

7 Conclusions

This study describes a method for estimating daily Tmin and Tmax from geostationary satellite data, based on dynamic, empirical regression of satellite-observed and auxiliary parameters against collocated station observations. The method, which has been tested over Europe for two years of data, allows the prediction of Tmin and Tmax for locations without station data at the spatial resolution of the satellite data (3 km at the sub-satellite point). Analysis of the model residuals and evaluation with independent station data from the UK and Germany not used in the regression process suggests that for most days at least 50% of the estimated NSATs are within 3 °C of collocated station observations, with around 80% within 4 °C and 90% within 5 °C. Results for Tmax are slightly better than for Tmin, with an overall mean bias of 0.0 °C compared with 0.0-0.5 °C for Tmin, and RMS differences of 2.3-2.5 °C for Tmax and 2.5-2.7 °C for Tmin. The mean bias of the satellite-estimated NSATs shows no seasonal variation, although the variance is noted to be higher during winter months owing to the lower number of station-satellite matchups used to build the regression model because of increased cloud frequency. The variance results are likely inflated through the inherent, random discrepancies that arise from comparing satellite area-averaged with in situ point temperature observations. Errors in the satellite data due to undetected cloud may cause significant biases in the estimated NSATs, particularly for Tmin.

The results presented in this study are consistent with previous studies noted in this article, where typical uncertainties of around 2-5 °C are obtained for predictions of daily or sub-daily NSAT over timescales of more than a few days [*Cristobal et al.*, 2008; *Davis and Tarpley*, 1983; *Jang et al.*, 2014; *Jones et al.,* 2010; *Kilibarda et al.*, 2014; *Köhn and Royer,* 2012; *Nieto et al.*, 2011; *Vancutsem et al.*, 2010; *Zhang et al.,* 2011]. The approach adopted for this study provides some enhancements over these previous studies. In particular, high-frequency geostationary observations are used to improve the probability of obtaining cloud-free daily IR observations of Tmin and Tmax, which is a significant problem for methods that use daily polar-orbiting IR data. A method that uses geostationary observations is also beneficial as these data could provide the long-term records (>30 years) suitable for climate applications. The use of IR LST data is advantageous as these data have higher spatial resolution compared with microwave observations; the NSAT estimates produced here at the satellite native FOV. The regression developed in this study also uses urban fraction as a predictor for the first time. Finally, the data set presented here is available at http://www.metoffice.gov.uk/hadobs/msg_tmaxmin/ and will be updated regularly; at the time of writing, only the data of *Jones et al.,* [2010] (daily microwave, 25-km spatial resolution for 2002-2011) are currently in the public domain. A real-time version of the data set, updated daily, will also be implemented in 2015.

The satellite NSATs presented here are not designed to replace conventional gridded station NSAT data sets, but to augment them. Future work includes application of the method described in this study to other land masses in the new Horizon2020 project, 'EUSTACE', which aims to produce a globally complete daily NSAT analysis back to 1850 using a combination of satellite and in situ observations.

8 Acknowledgements

This work was supported by the EU Framework 7 project, EURO4M (http://www.euro4m.eu/). The SEVIRI daily LSTmin/max and Tmin/max data produced in this study are available at http://www.metoffice.gov.uk/hadobs/msg_tmaxmin/. The underlying SEVIRI satellite products were provided by the LSA-SAF (http://landsaf.meteo.pt/). The author would like to acknowledges the ESA and ESA GlobCover project, led by MEDIAS-France, for the provision of GlobCover data (http://due.esrin.esa.int/globcover/), the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). The non-public ECA&D data for the UK and Germany used in this study for independent evaluation of the new SEVIRI products were provided courtesy of KNMI and ECA&D project; the author would also like to thank DWD for provision of these extra German station data. The HadISD sub-daily station 2m temperature data used in this study were provided by Robert Dunn, Met Office Hadley Centre, and are available from http://www.metoffice.gov.uk/hadobs/hadisd/. Please contact the author, Elizabeth Good (email elizabeth.good@metoffice.gov.uk), for the IDL code used to produce the analysis described in this article. All the underlying data used in this study are freely available online at the specified web addresses, with the exception of the non-public ECA&D data, which was obtained through private agreement with the ECA&D team (please see http://www.ecad.eu or email eca@knmi.nl for further information). The author would also like to thank the two anonymous reviewers whose valuable comments have helped improve this article.

9 References

Arino O., P. Bicheron, F. Achard, J. Latham, R. Witt and J.-L. Weber (2008), GlobCover the most detailed portrait of Earth, ESA Bulletin, 136, 25-31, November 2008.

Basist, A., N. C. Grody, T. C. Peterson, and C. N. Williams (1998), Using the special sensor microwave/imager to monitor land surface temperatures, wetness, and snow cover, Journal of Applied Meteorology, 37(9), 888–911.

Becker, F., and Z. L. Li, (1990), Towards a local split window method over land surfaces. Int. J. Remote Sensing, 11, 369-393.

Caesar, J., L. Alexander, and R. Vose (2006), Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set, J. Geophys. Res., 111, D05101, doi:10.1029/2005JD006280

Chen, F., Y. Liu, Q. Liu and F. Qin (2014), A statistical method based on remote sensing for the estimation of air temperature in China. Int. J. Climatol.. doi: 10.1002/joc.4113

Cristobal, J., M. Ninyerola, and X. Pons (2008), Modeling air temperature through a combination of remote sensing and GIS data, J. Geophys. Res., 113, D13106, doi:10.1029/2007JD009318.

Davis, F. A., and J. D. Tarpley (1983), Estimation of Shelter Temperatures from Operational Satellite Sounder Data, J. Climate Appl. Meteor., 22, 369–376.

Dash, P., F. M. Goettsche, F. S. Olesen, H Fischer (2002), Land surface temperature and emissivity estimation from passive sensor data: Theory and practice-current trends, International Journal of Remote Sensing, 23(13), 2563-2594.

Donlon, C. J., M. Martin, J. D. Stark, J. Roberts-Jones, E. Fiedler and W. Wimmer, (2012), The Operational Sea Surface Temperature and Sea Ice analysis (OSTIA). Remote Sensing of the Environment, 116, 140-158, doi: 10.1016/j.rse.2010.10.017.

Dousset, B., F. Gourmelon, K. Laaidi, K. Zeghnoun, E. Giraudet, P. Bretin, E. Mauri, and S. Vandentorren (2011), Satellite monitoring of summer heat waves in the Paris metropolitan area, Int. J. Climatol., 31, 313–323, doi:10.1002/joc.2222.

Dunn, R. J. H., K. M. Willett, P. W. Thorne, E. V. Woolley, I. Durre, A. Dai, D. E. Parker, and R. S. Vose (2012), HadISD: A Quality Controlled global synoptic report database for selected variables at long-term stations from 1973-2011, Climate of the Past, 8, 1649-1679.

Fairall, C. W., E. F. Bradley, J. S. Godfrey, G. A. Wick, J. B. Edson, and G. S. Young (1996), Cool-skin and warm-layer effects on sea surface temperature. J. Geophys., 101 (C1), 1295–1308.

Florio, E. N., S. R. Lele, Y. C. Chang, R. Sterner, and G. E. Glass (2004), Integrating AVHRR satellite data and NOAA ground observations to predict surface air temperature: A statistical approach. International Journal of Remote Sensing, 25(15), 2979–2994.

Freitas, S. C., I. F. Trigo, J. M. Bioucas-Dias, and F. M. Göttsche (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., I.F. Trigo, J. Macedo, C. Barroso, R. Silva and R. Perdigão (2013), Land surface temperature from multiple geostationary satellites. International Journal of Remote Sensing, 2013. 34(9-10): p. 3051-3068. Doi: 10.1080/01431161.2012.716925

Hansen, J., R. Ruedy, M. Sato, and K. Lo (2010), Global surface temperature change. Rev. Geophys., 48, RG4004, doi:10.1029/2010RG000345.

Haylock, M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones and M. New (2008), A European daily high-resolution gridded dataset of surface temperature and precipitation. J. Geophys. Res (Atmospheres), 113, D20119, doi:10.1029/2008JD10201.

Heidinger, A., I. Laszlo, C.C. Molling, and D. Tarpley (2013), Using SURFRAD to Verify the NOAA Single-Channel Land Surface Temperature, J. Atmos. Oceanic Technol., 30, 2868–2884, doi: http://dx.doi.org/10.1175/JTECH-D-13-00051.1 .

Hulley, G. C., C. G. Hughes, and S. J. Hook (2012), Quantifying uncertainties in land surface temperature and emissivity retrievals from ASTER and MODIS thermal infrared data, J. Geophys. Res., 117, D23113, doi:10.1029/2012JD018506.

Hulley, G. C., and S. J. Hook (2010), Generating Consistent Land Surface Temperature and Emissivity Products Between ASTER and MODIS Data for Earth Science Research, IEEE Transactions on Geoscience and Remote Sensing. 49 (4), 1304 – 1315, DOI: 10.1109/TGRS.2010.2063034.

Hulley, G. C., and S. J. Hook (2012), A radiance-based method for estimating uncertainties in the Atmospheric Infrared Sounder (AIRS) land surface temperature product, J. Geophys. Res., 117, D20117, doi:10.1029/2012JD018102.

Inamdar, A. K., A. French, S. Hook, G. Vaughan, and W. Luckett (2008), Land surface temperature retrieval at high spatial and temporal resolutions over the southwestern United States, J. Geophys. Res., 113, D07107, doi:10.1029/2007JD009048.

Jackson, D.L., G.A. Wick, and J.J. Bates (2006), Near-surface retrieval of air temperature and specific humidity using multisensor microwave satellite observations, J. Geophys. Res., 111, D10306, doi:10.1029/2005JD006431.

Jang, K., S. Kang , J.S. Kimball, and S.Y. Hong (2014), Retrievals of All-Weather Daily Air Temperature Using MODIS and AMSR-E Data. Remote Sens., 6, 8387-8404; doi:10.3390/rs6098387.

Jones, L. A., C. R. Ferguson, J. S. Kimball, K. Zhang, S. Chan, K. C. McDonald, E. G. Njoku, and E. F. Wood (2010), Satellite Microwave Remote Sensing of Daily Land Surface Air Temperature Minima and Maxima From AMSR-E. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 3(1), 111-123, doi: 10.1109/JSTARS.2010.2041530.

Jones, P. D., D. H. Lister, T. J. Osborn, C. Harpham, M. Salmon, and C. P. Morice (2012), Hemispheric and large-scale land surface air temperature variations: An extensive revision and an update to 2010, J. Geophys. Res., 117, D05127, doi:10.1029/2011JD017139.

Kawashima, S., T. Ishida, M. Minomura, T. Miwa (2000), Relations between Surface Temperature and Air Temperature on a Local Scale during Winter Nights, J. Appl. Meteor., 39, 1570–1579. doi: http://dx.doi.org/10.1175/1520-0450(2000)039<1570:RBSTAA>2.0.CO;2.

Kilibarda, M., T. Hengl, G. B. M. Heuvelink, B. Gräler, E. Pebesma, M. Perčec Tadić, and B. Bajat (2014), Spatio-temporal interpolation of daily temperatures for global land areas at 1km resolution, J. Geophys. Res. Atmos., 119, 2294–2313, doi:10.1002/2013JD020803.

Klein Tank, A. M. G., J. B. Wijngaard, G. P. Können, R. Böhm, G. Demarée, A. Gocheva, M. Mileta, S. Pashiardis, L. Hejkrlik, C. Kern-Hansen, R. Heino, P. Bessemoulin, G. Müller-Westermeier, M. Tzanakou, S. Szalai, T. Pálsdóttir, D. Fitzgerald, S. Rubin, M. Capaldo, M. Maugeri, A. Leitass, A. Bukantis, R. Aberfeld, A. F. V. van Engelen, E. Førland, M. Mietus, F. Coelho, C. Mares, V. Razuvaev, E. Nieplova, T. Cegnar, J. Antonio López, B. Dahlström, A. Moberg, W. Kirchhofer, A. Ceylan, O. Pachaliuk, L. V. Alexander, and P. Petrovic (2002), Daily dataset of 20th-century surface air temperature and precipitation series for the

European Climate Assessment, Int. J. Climatol., 22, 1441–1453, doi: 10.1002/joc.773.

Köhn, J., and A. Royer (2012), Microwave Brightness Temperature as an Indicator of Near–Surface Air Temperature Over Snow in Canadian Northern Regions, International J. of Remote Sensing 33(4), 1126-1138. doi:10.1080/01431161.2010.550643.

Li, Z.-L., B.-H. Tang, H. Wu, H. Ren, G. Yan, Z. Wan, I.F. Trigo and J.A. Sobrino (2013), Satellite-derived land surface temperature: Current status and perspectives, Remote Sens. Environ., 131, 14–37, doi: 10.1016/j.rse.2012.12.008.

Nieto, H., I. Sandholt, I. Aguado, E. Chuvieco, S. Stisen (2011), Air temperature estimation with MSG-SEVIRI data: Calibration and validation of the TVX algorithm for the Iberian Peninsula, Remote Sensing of Environment, 115(1), 107–116.

Prata, A. J. (1993), Land surface temperatures derived from the advanced very high resolution radiometer and the along-track scanning radiometer: 1. Theory, J. Geophys. Res., 98(D9), 16689–16702, doi:10.1029/93JD01206.

Prata, A. J. (1994), Land surface temperatures derived from the advanced very high resolution radiometer and the along-track scanning radiometer: 2. Experimental results and validation of AVHRR algorithms, J. Geophys. Res., 99(D6), 13025–13058, doi:10.1029/94JD00409.

Prata, A.J, and R.P. Cechet (1999), An Assessment of the Accuracy of Land Surface Temperature Determination from the GMS-5 VISSR, Remote Sensing of Environment, Volume 67, Issue 1, pp 1–14, doi: 10.1016/S0034-4257(98)00055-8.

Prata, F. (2002), Land Surface Temperature Measurement from Space: AATSR Algorithm Theoretical Basis Document, contract report to ESA, https://earth.esa.int/web/guest/-/land-surface-temperature-measurement-from-space-aatsr-algorithm-theoretical-basis-document-4514.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V.Alexander, D. P. Rowell, E. C. Kent, A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century J. Geophys. Res., 108, No. D14, 4407 10.1029/2002JD002670.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes and W. Wang (2002), An improved in situ and satellite SST analysis for climate. J. Climate, 15, 1609–1625, doi: http://dx.doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2.

De Ridder, K., C. Bertrand, G. Casanova, and W. Lefebvre (2012), Exploring a new method for the retrieval of urban thermophysical properties using thermal infrared remote sensing and deterministic modeling, J. Geophys. Res., 117, D17108, doi:10.1029/2011JD017194.

Roberts, J.B., C.A. Clayson, F.R. Robertson, and D.L. Jackson (2010), Predicting near-surface characteristics from SSM/I using neural networks with a first guess approach. J. Geophys. Res., 115, D19113, doi:10.1029/2009JD013099

Shi, L., G. Peng, and J. J. Bates (2012), Surface Air Temperature and Humidity from Intersatellite-Calibrated HIRS Measurements in High Latitudes, J. Atmos. Oceanic Technol., 29, 3-13, doi:10.1175/JTECH-D-11-00024.1

Smith, T. M., R. W. Reynolds, T. C. Peterson and J. Lawrimore (2008), Improvements to NOAA's Historical Merged Land–Ocean Surface Temperature Analysis (1880–2006). J. Climate, 21, 2283–2296, doi: http://dx.doi.org/10.1175/2007JCLI2100.1

Sohrabinia, M. , P. Zawar-Reza and W. Rack (2014), Spatio-temporal analysis of the relationship between LST from MODIS and air temperature in New Zealand, Theor. Appl. Climatol., DOI 10.1007/s00704-014-1106-2.

Sun, D., L. Fang, Y. Yu, 2012, GOES LST Algorithm Theoretical Basis Document v3, NOAA NESDIS STAR publication, http://www.ospo.noaa.gov/Products/land/glst/algo.html.

Sun, D. And R. T. Pinker, (2007), Retrieval of surface temperature from the MSG-SEVIRI observations: Part I. Methodology, International Journal of Remote Sensing, 28 (23), 5255-5272, DOI: 10.1080/01431160701253246.

Sun, Y. J., J. F. Wang, R. H. Zhang, R. R. Gillies, Y. Xue, Y. C. Bo (2005), Air temperature retrieval from remote sensing data based on thermodynamics, Theor. Appl. Climatol., 80(1), 37–48, doi:10.1007/s00704-004-0079-y.

Tang, B., Y. Bi, Z.-L. Li, J. Xia (2008), Generalized Split-Window Algorithm for Estimate of Land Surface Temperature from Chinese Geostationary FengYun Meteorological Satellite (FY-2C) Data, Sensors 8, no. 2: 933-951.

Trigo, I. F., I. T. Monteiro, F. Olesen, and E. Kabsch (2008), An assessment of remotely sensed land surface temperature, J. Geophys. Res., 113, D17108, doi:10.1029/2008JD010035.

Vancutsem, C., P. Ceccato, T. Dinku, S. J. Connor (2010), Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa, Remote Sens Environ, 114(2), 449–465, doi:10.1016/j.rse.2009.10.002, http://www.sciencedirect.com/science/article/pii/S0034425709003113.

Vinnikov, K. Y., Y. Yu, M.D. Goldberg, D. Tarpley, P. Romanov, I. Laszlo, and M. Chen (2012), Angular anisotropy of satellite observations of land surface temperature, Geophys. Res. Lett., 39, L23802, doi:10.1029/2012GL054059.

Wan, Z., and J. Dozier 1996, A generalised split-window algorithm for retrieving land-surface temperature from space, IEEE Trans. Geosci. Remote Sens., 34(34), 892-905.

Wan, Z. (2014), New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product, Remote Sensing of Environment, 140, 36-45, doi: 10.1016/j.rse.2013.08.027

Yan, Z., Z. Li, Q. Li and P. Jones (2010), Effects of site change and urbanisation in the Beijing temperature series 1977–2006, Int. J. Climatol. 30: 1226–1234, DOI: 10.1002/joc.1971

Zhang, W., Y. Huang, Y. Yu, W. Sun (2011), Empirical models for estimating daily maximum, minimum and mean air temperatures with MODIS land surface temperatures, Int. J. Remote Sens., 32(1), 1–26, doi:10.1080/431161.2011.560622.

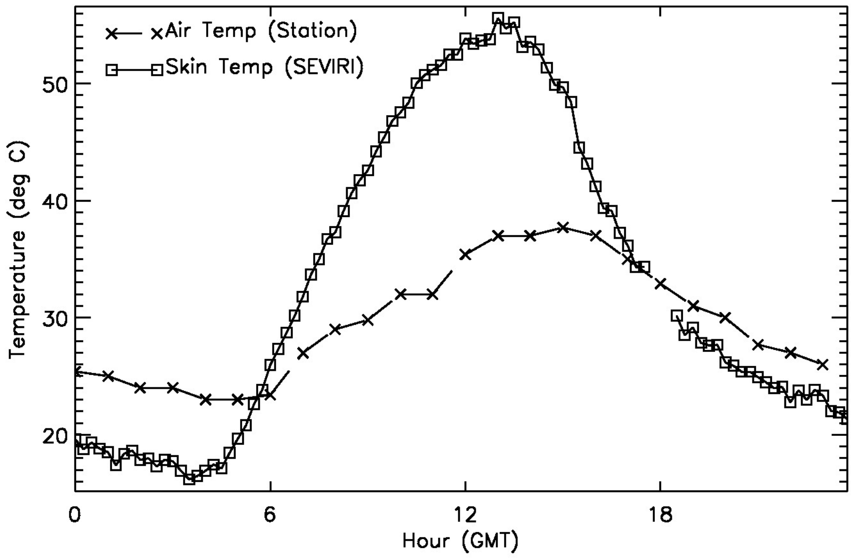


Figure 1: Diurnal cycle on 2 July 2011 for land surface temperature (LST) and near surface air temperature (NSAT) at meteorological station 607350 (35 40N, 10 06E; Kairouan, Tunisia). The skin temperature data are from the Spinning Enhanced Visible and Infrared Imager (SEVIRI; see Section 2.1). Missing SEVIRI data correspond to cloudy observations. The in situ air temperatures are from the HadISD data set [*Dunn et al.*, 2012].

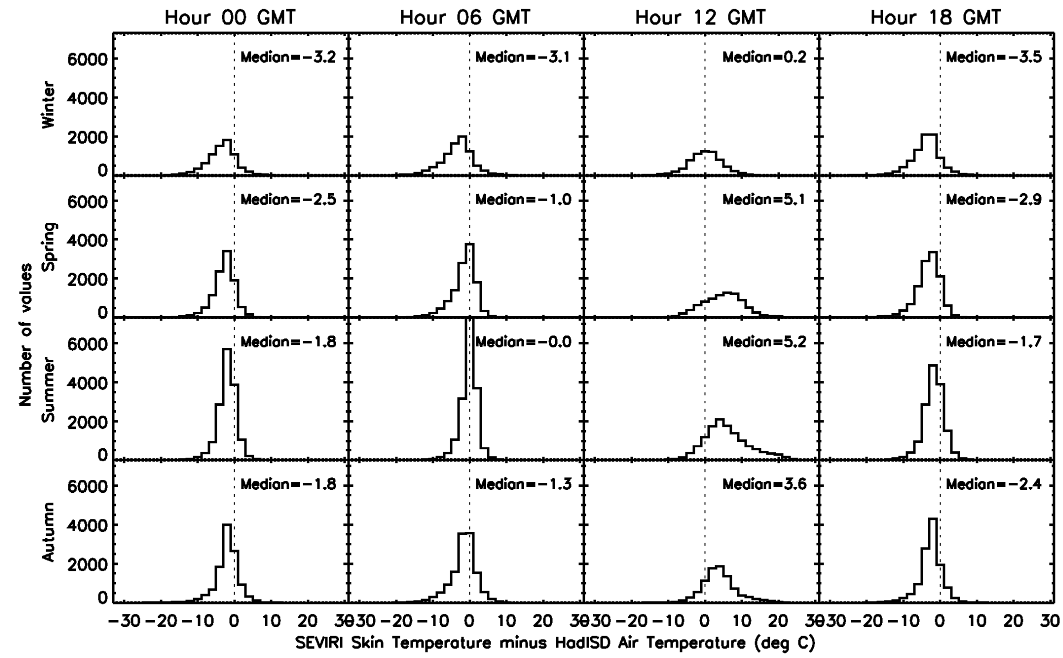


Figure 2: Histograms showing the difference between LST from the SEVIRI onboard MSG and temporally and spatially collocated HadISD [*Dunn et al.*, 2012] station temperatures over Europe between 0 and 20 degrees longitude for December 2009-February 2010 (winter), March-May 2010 (spring), June-August 2010 (summer) and September-November 2010.

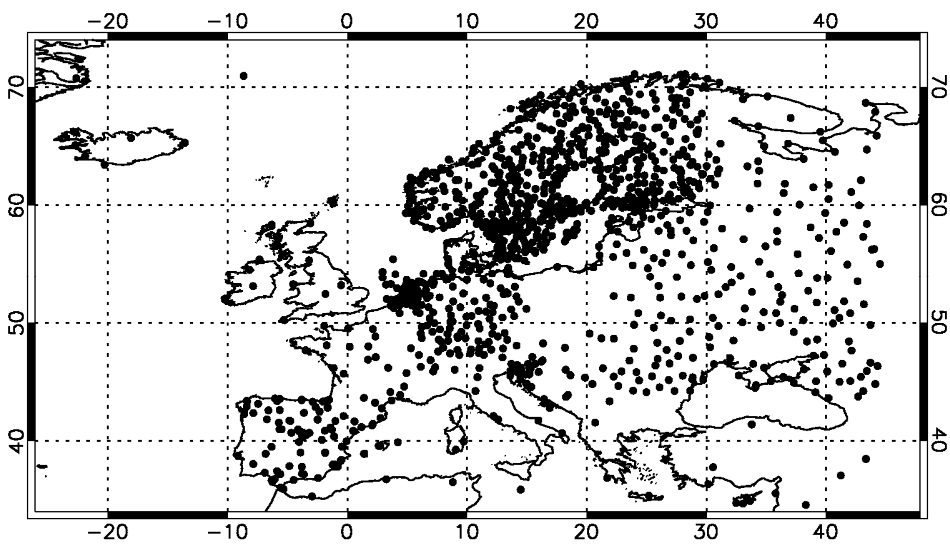


Figure 3: ECA&D stations (publicly available) with Tmin/Tmax data for 2012/2013.

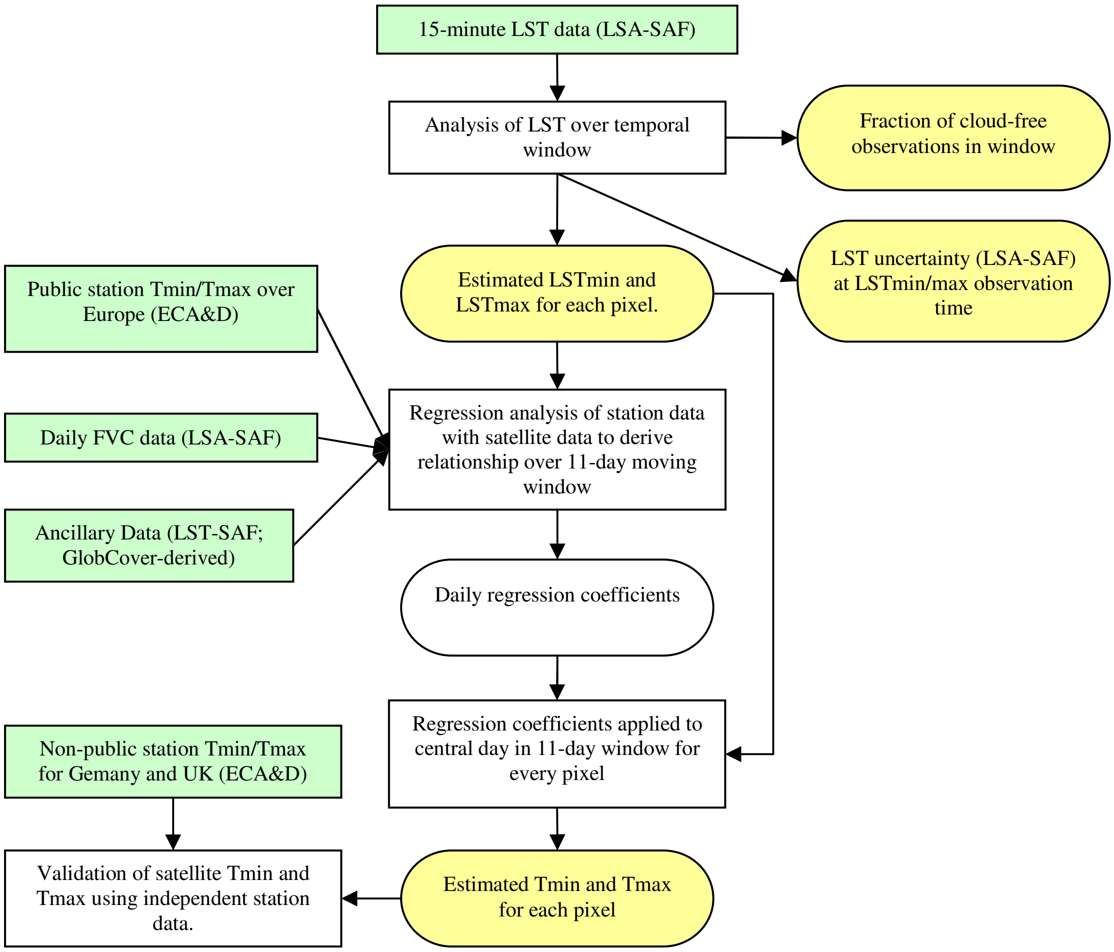


Figure 4: Process of deriving minimum and maximum land temperatures. Input data sets are in green and output data sets are in yellow.

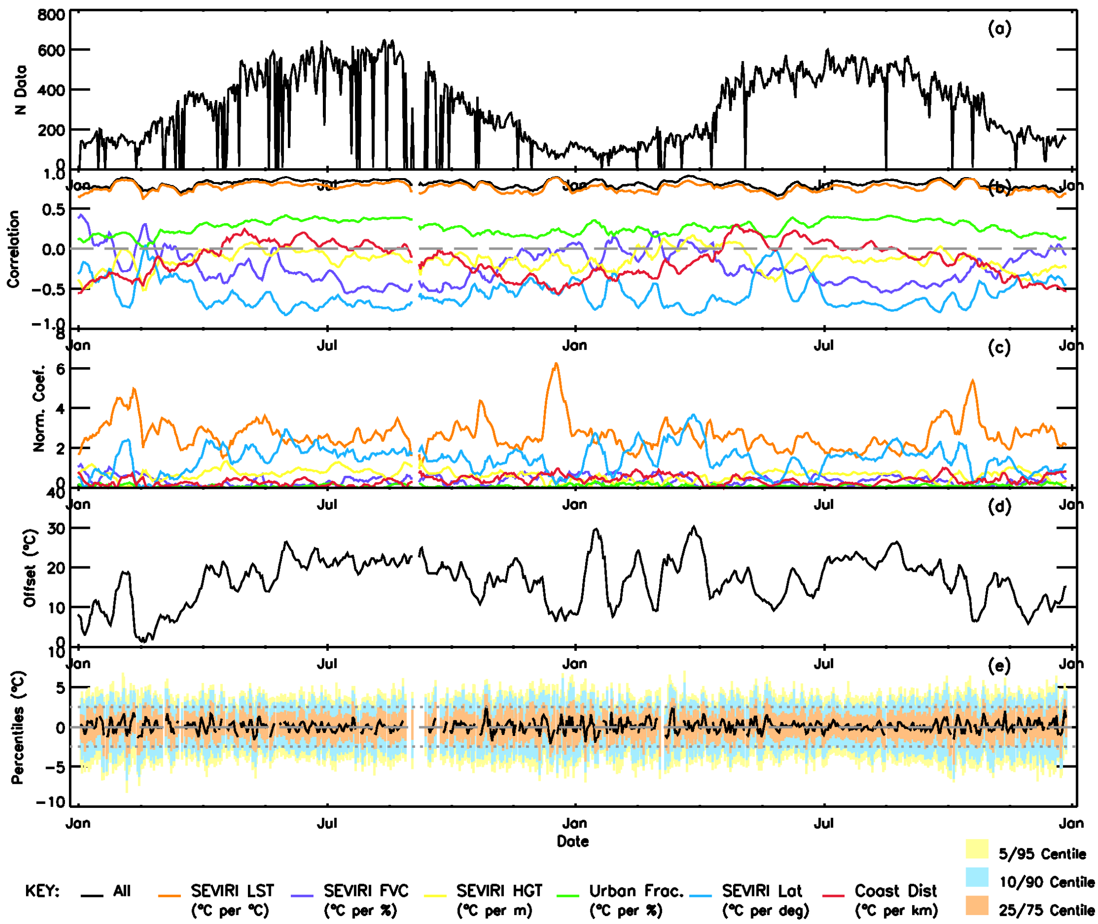


Figure 5: Analysis of station air temperatures vs. empirical model predictor variables for Tmin for 2012/2013. Regression coefficients formed from moving 11-day window centred on day in question. Panels show (a) number of cloud-free satellite-station matchups on day of observation (i.e. central day in window), (b) correlation coefficient between station air temperature and collocated predictor variables (black line indicates multiple linear regression coefficient), (c) normalised multiple linear regression coefficients, (d) regression offset, and (e) residuals (i.e. satellite minus station NSAT) on day of observation (i.e. central day in window), the solid black line indicates the median.

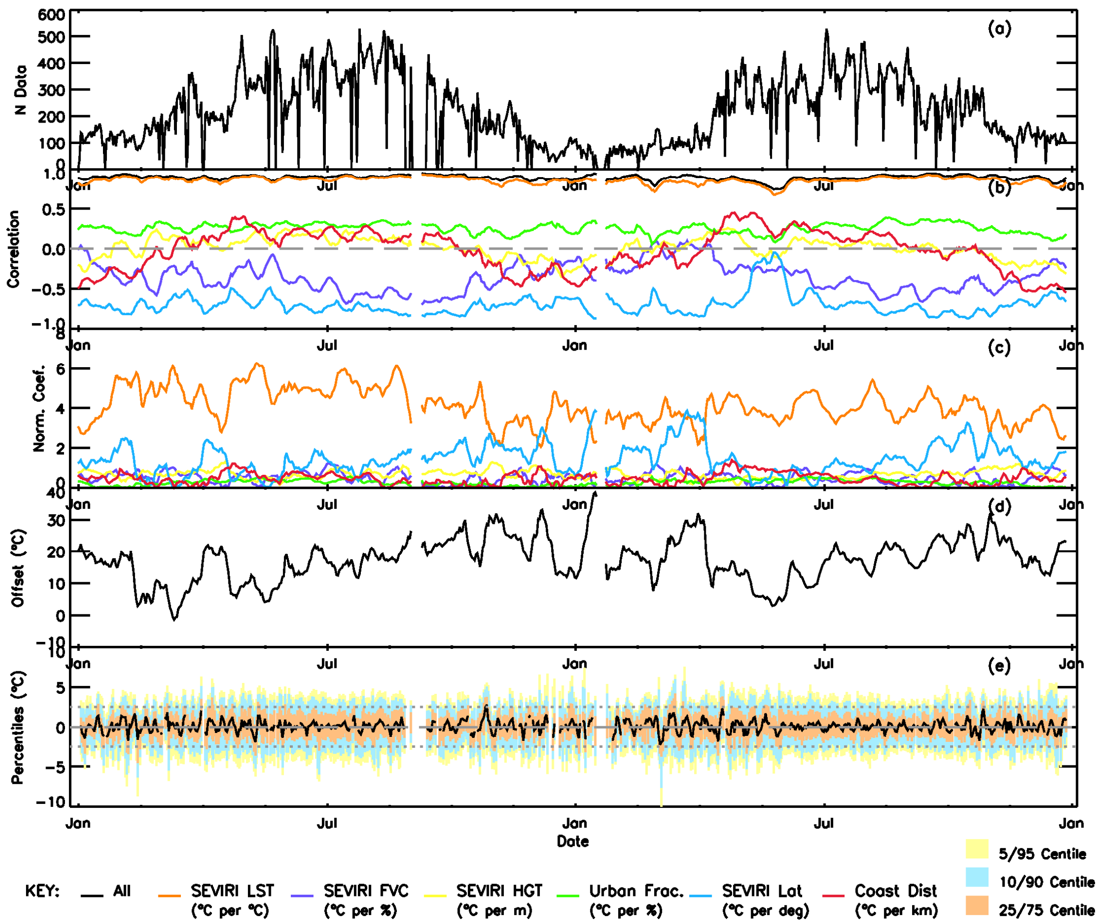


Figure 6: As for Figure 5 but for Tmax.

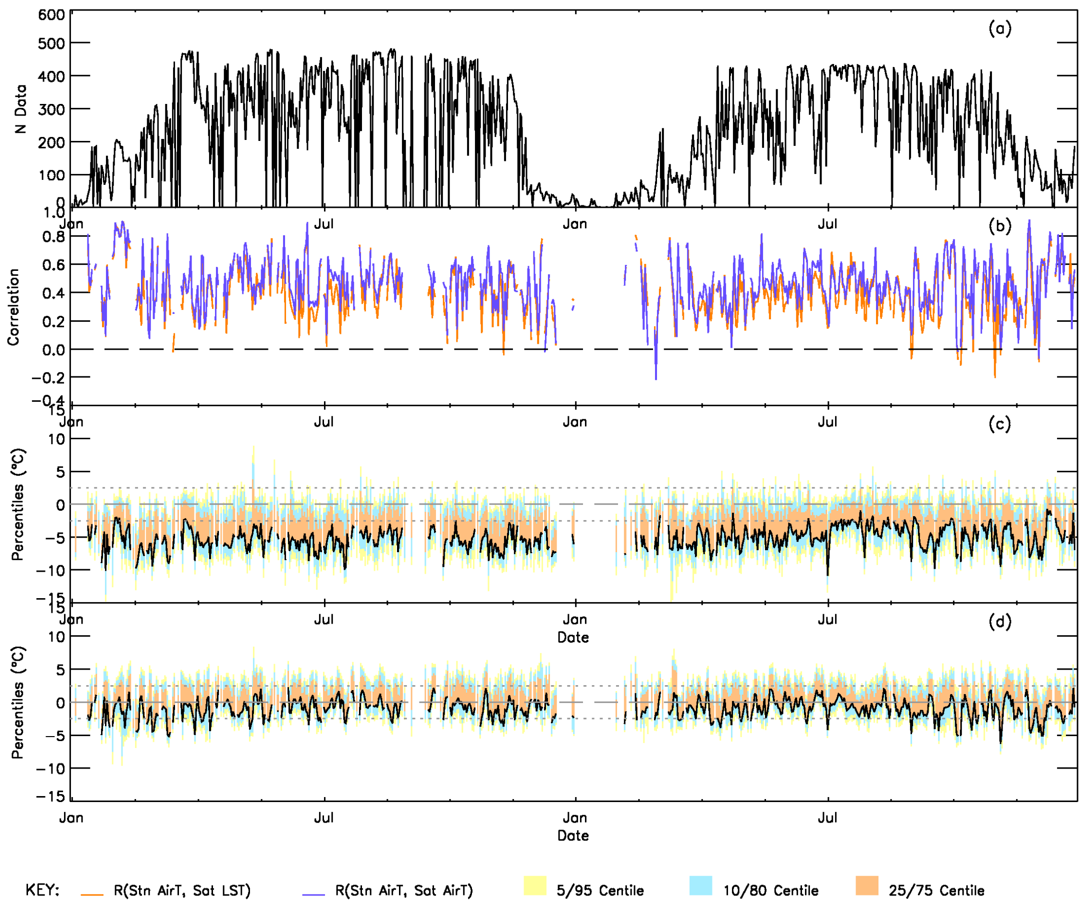
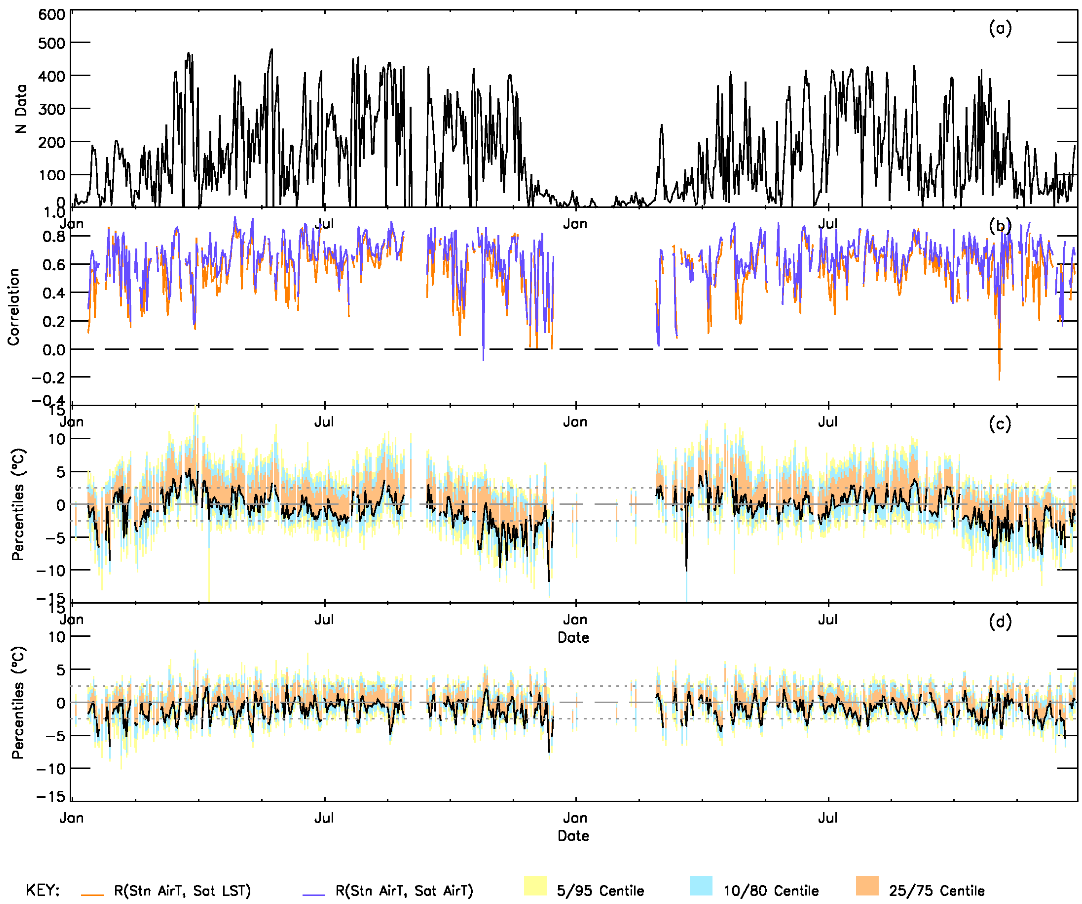


Figure 7: Validation of SEVIRI Tmin over Germany & UK for 2012-2013 showing (a) the number of SEVIRI-station matchups, (b) the correlation between station Tmin and SEVIRI LSTmin, and station Tmin and SEVIRI Tmin, (c) SEVIRI LSTmin minus station Tmin and (d) satellite Tmin minus station Tmin distributions for each day. 'SEVIRI LSTmin' is the daily minimum LST (i.e. 'skin' temperature), while 'SEVIRI Tmin' is the derived daily minimum NSAT (i.e. through the multiple-linear regression method described in this study).

950



951 Figure 8: As for Figure 7 but for Tmax.

952

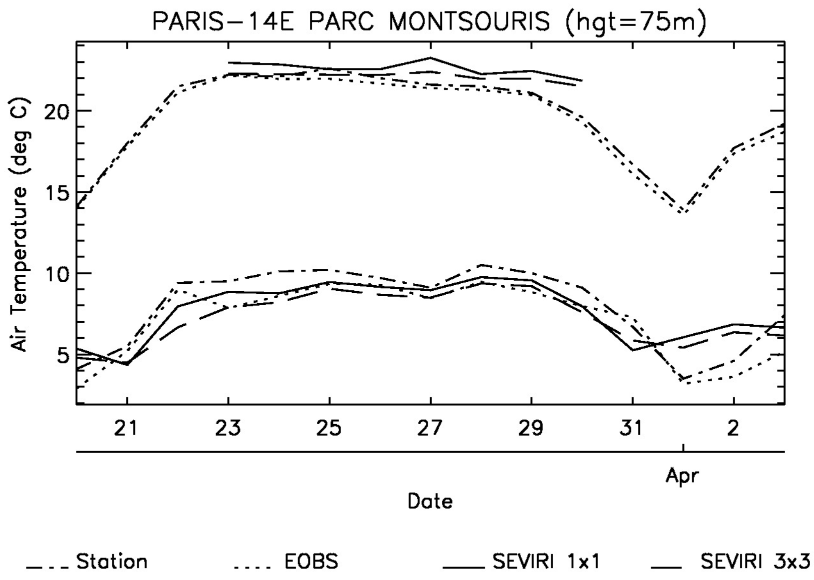


Figure 9: Daily Tmin (lower group of lines) and Tmax (upper group of lines) over Paris between 20 March and 3 April 2012 observed in situ (ECA&D), from EOBS and SEVIRI. The SEVIRI data correspond to a single pixel (1x1) and the average of a 3x3 block of pixels (3x3).

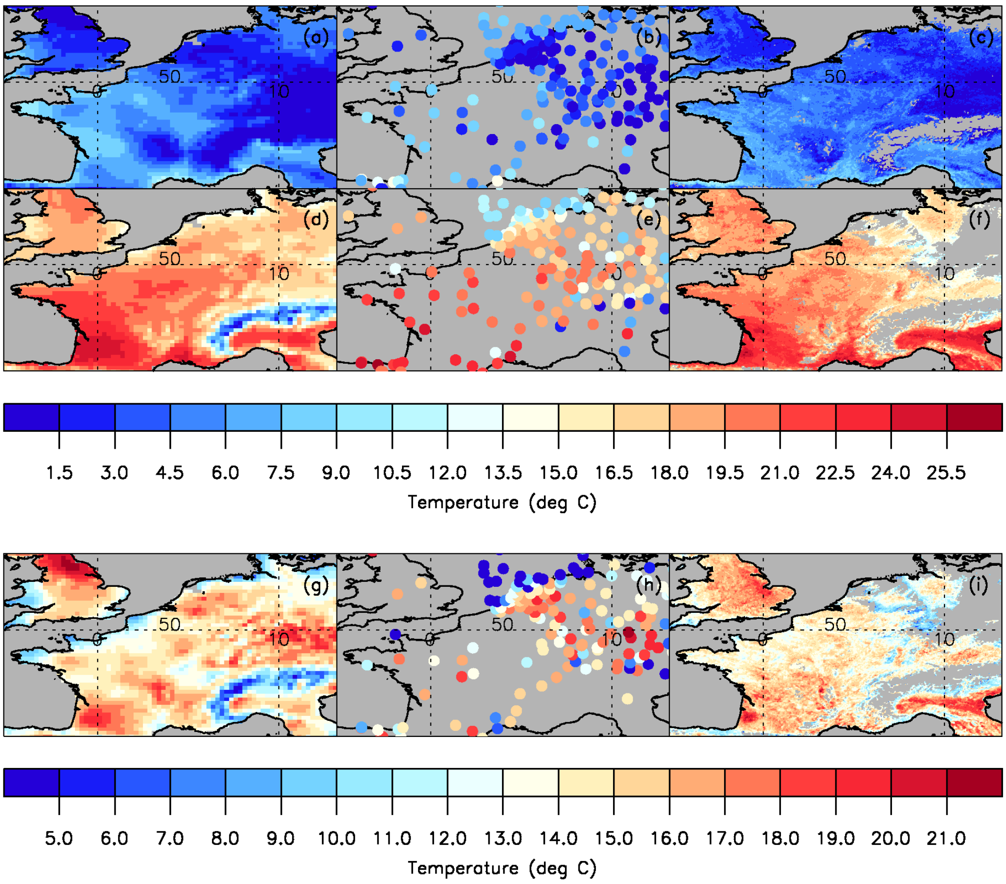


Figure 10: Tmin and Tmax observations for 27 March 2012. (a) EOBS Tmin, (b) ECA&D public station Tmin, (c) SEVIRI Tmin, (d) EOBS Tmax, (e) ECA&D public station Tmax, (f) SEVIRI Tmax. Panels (g) to (i) show the corresponding diurnal temperature ranges, i.e. Tmax minus Tmin. Locations with no data or missing data, e.g. due to cloud, are shown in grey.

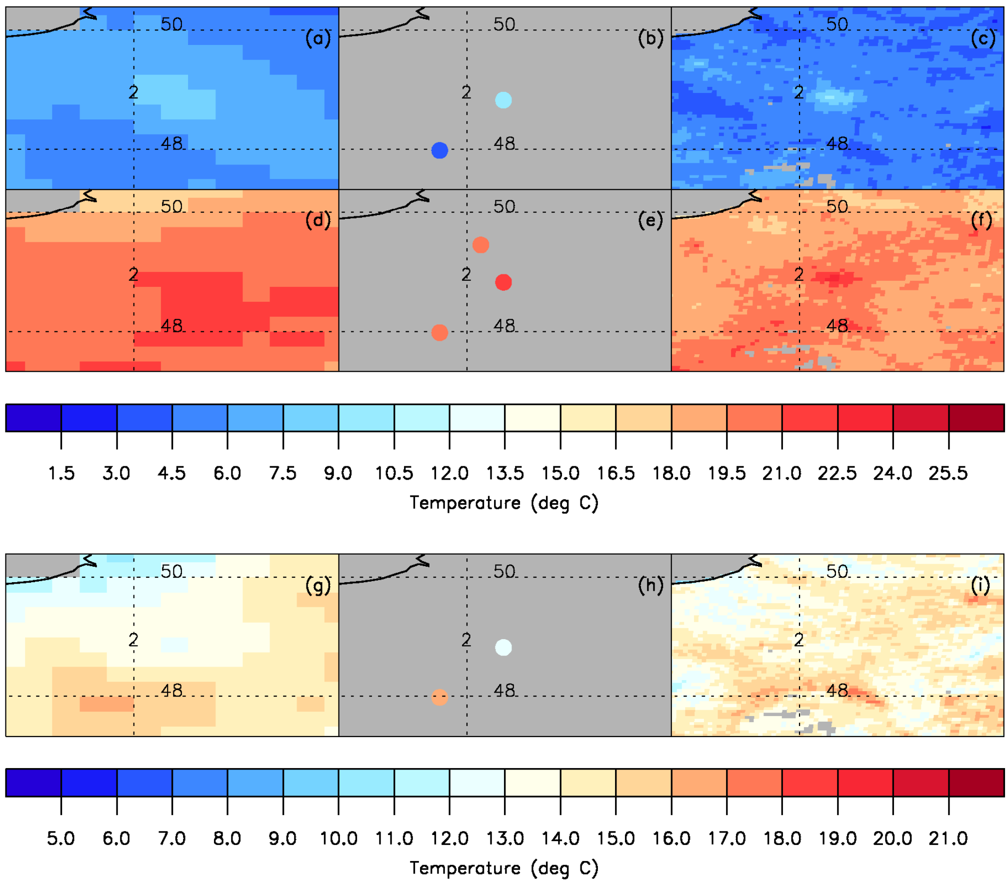


Figure 11: As for Figure 10 but centred over Paris.

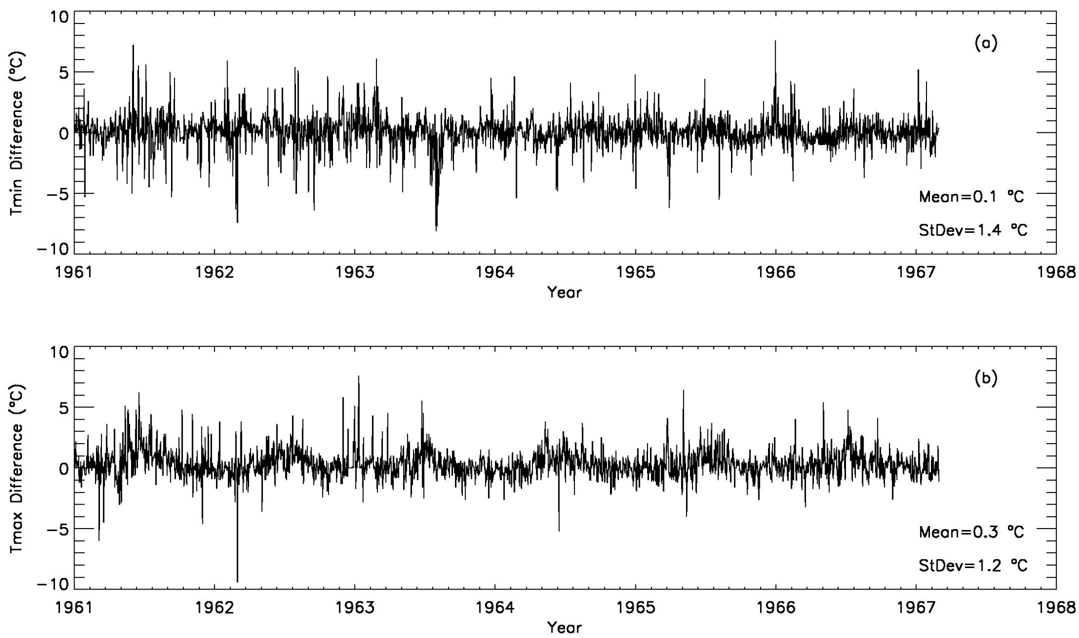


Figure 12: Example of NSAT differences between neighbouring stations.

Trollhattans-Flygpla (Sweden, ECA&D source 35772) minus Trollhattans (Sweden, ECA&D source 35769) for (a) Tmin and (b) Tmax. The distance between stations is 1.26 km with 1 m elevation difference. Both stations are located within the same SEVIRI pixel.

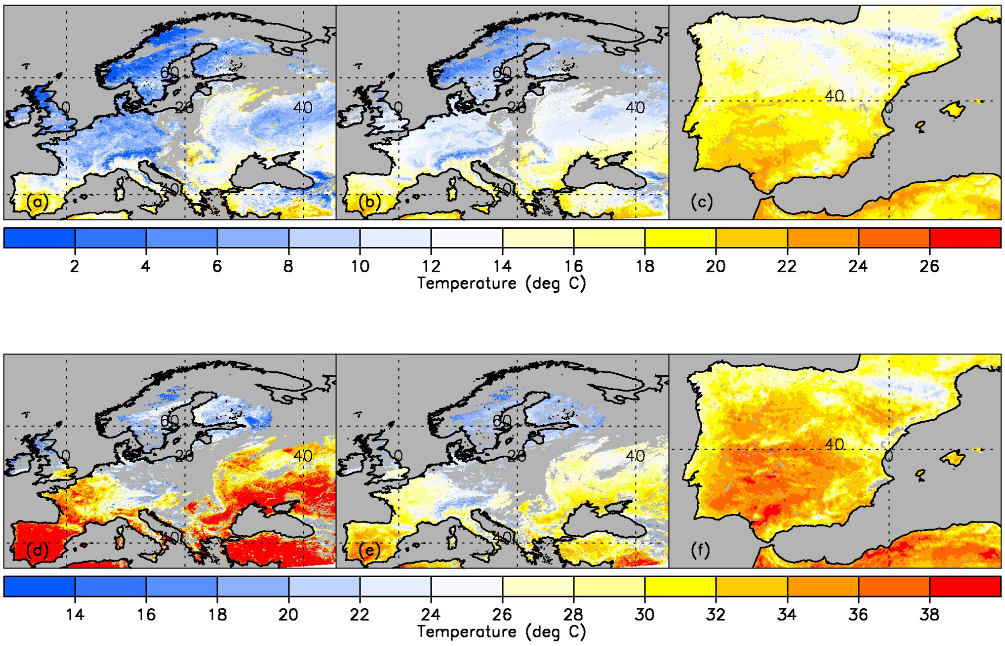


Figure 13: Example of satellite temperature data on 21 August 2013 showing (a) LSTmin, (b) Tmin, (c) Tmin for the Iberian Peninsula, (d) LSTmax, (e) Tmax and (f) Tmax for the Iberian Peninsula. Locations with no data or missing data, e.g. due to cloud, are shown in grey.

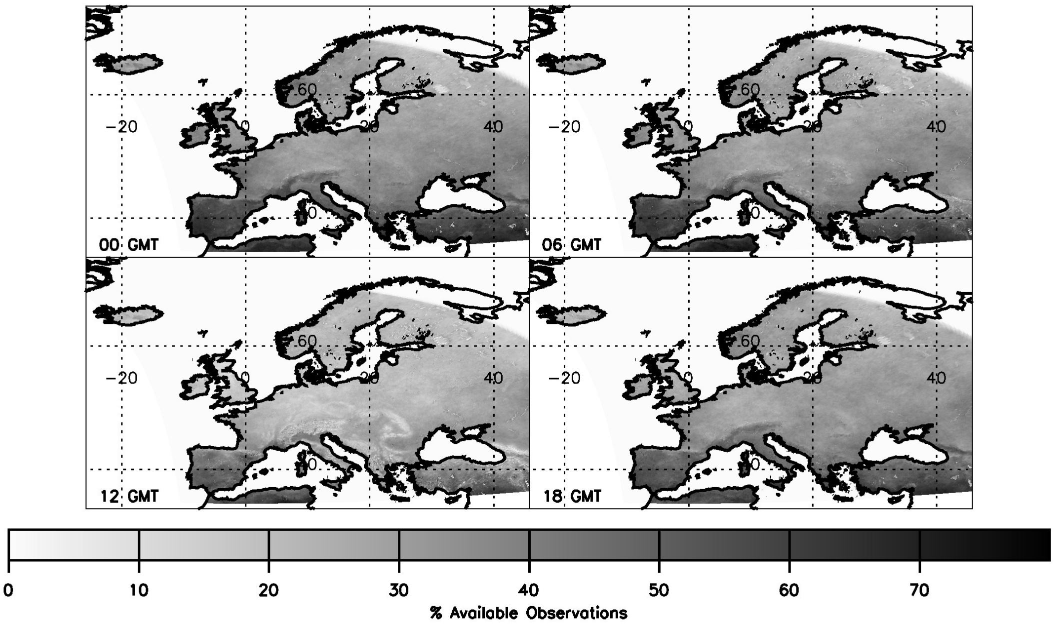


Figure 14: Percentage of days in 2010 with cloud-free SEVIRI observations at synoptic times (labelled on plots).

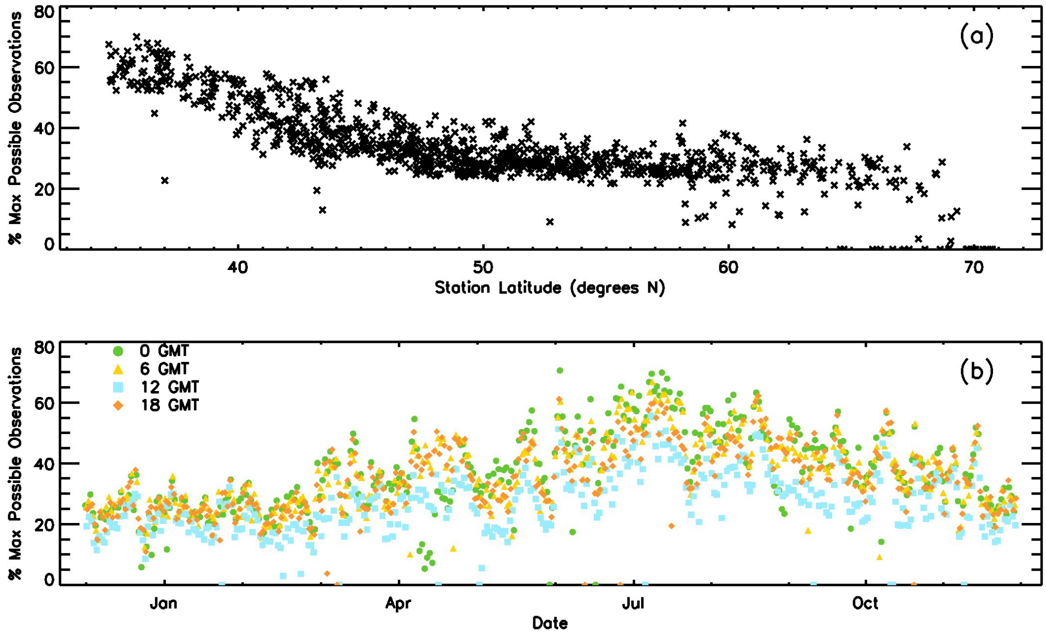


Figure 15: Percentage of days in 2010 with cloud-free observations at HadISD station locations [*Dunn et al.*, 2012] plotted as a function of (a) latitude and (b) date and coloured according to observation time.