Article

Spatial Downscaling of 2-Meter Air Temperature Using Operational Forecast Data

Thomas Huld * and Irene Pinedo Pascua

European Commission, Joint Research Centre, Via Fermi 2749, 21027 Ispra, Italy;
E-Mail: irene.pinedo@jrc.ec.europa.eu

* Author to whom correspondence should be addressed; E-Mail: thomas.huld@jrc.ec.europa.eu;
  Tel.: +39-033-278-5273; Fax +39-033-278-9268.

Abstract: We present a method for enhancing the spatial resolution of 2 m temperature (T_{2m}) estimates. The method is based on operational forecast data supplied by the European Centre for Medium Range Weather Forecast. From the hourly and monthly average 2-meter temperatures a vertical gradient is determined by linear fitting to the temperature data in larger areas of $1° \times 1°$ or $2° \times 2°$. Validation against data from more than 8000 meteorological stations worldwide shows that the estimates of annual average temperature at these points becomes significantly more accurate when applying the vertical gradients to correct the local temperature estimates to the elevation of the stations. When the elevation difference between forecast and station is larger than 300 m, the overall mean absolute deviation of the individual stations bias values decreases from 3.44 to 1.02 °C and the root mean square deviation decreases from 4.11 to 1.42 °C. The gradients have also been applied to the ERA-Interim reanalysis data and the validation results are similar. The vertical temperature gradients will be useful for studies in many fields, including renewable energy and the study of energy performance of buildings.

Keywords: downscaling; vertical temperature gradients; 2-meter temperature; building energy performance

1. Introduction

Many technical applications need data on outdoor air temperature at near-ground level, often called 2-meter temperature from the standard height of meteorological measurements. Only considering energy applications, temperature data are needed for the design of building insulation and heating and cooling systems [1], as well as for estimates of the performance of solar heating and photovoltaic solar energy [2,3]. However, the availability of 2 m-temperature measurements varies strongly: while some regions have a rather dense network of measurement stations, other regions in the world have very few. If the 2 m-temperature varies strongly over short distances, which happens especially if there are large variations in elevation, even measurement stations a short distance away may not be representative of local conditions.

Meteorological models are a possible tool to overcome the problem of regions with very sparse data both through reanalysis and operational products, either on a regional or global scale. Reanalysis products such as European Centre for Medium Range Weather Forecast (ECMWF) ERA-Interim [4], National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR) dataset [5] or Modern-Era Retrospective Analysis For Research and Applications (MERRA) dataset [6] have global coverage. However, the spatial resolution is low: 45′, 2.5° and 1/2° latitude by 2/3° longitude respectively. Thus, there is a risk that a given location is not well represented by the corresponding value from the reanalysis products, either because of a large difference in elevation or because of strong horizontal temperature gradients, such as near coastlines. The use of data from numerical weather forecast models may improve the situation somewhat, but at the cost of increased complexity. Regional forecast models only cover limited areas and some areas may not be included in any regional models. Global numerical forecast models such as those from the European Centre for Medium Range Weather Forecast (ECMWF) have now reached a fairly high spatial resolution, but this is only the case for the last few years, and even there, the spatial resolution may be insufficient to capture local orography in mountainous areas.

An illustration of the problem is shown in Figure 1. Three different sources for the elevation data are used to construct the corresponding topographic profiles along a transect of constant latitude 45°48′, running from 8° E to 12° E. The red line in the map illustrates this transect, which runs through part of the Italian Alps and lake region. The topographic profile based on the SRTM-3 data set [7] with a resolution of 3″ (about 60 m in the east-west direction at this latitude), is shown in green. The orange line corresponds to the elevations from the DEM used by the ECMWF operational forecast (7.5′ resolution) while for the blue line the DEM comes from the ECMWF ERA-interim reanalysis (45′). It is clear that in such mountainous areas the ERA-interim DEM renders the elevation very poorly. The DEM used in the operational forecast is a better fit to the actual elevation but also shows large differences with the high-resolution DEM in some areas, such as for example the northern Lago di Garda, at around 220 km in Figure 1, where the divergence is more than 700 m. With a lapse rate of –0.6 ° C/100 m this would correspond to an underestimate of more than 4 ° C in the annual mean temperature.

For many studies, particularly in environmental, agricultural and biological sciences, temperature data at fine spatial and/or temporal resolutions are necessary to capture local phenomena otherwise lost at lower resolutions.
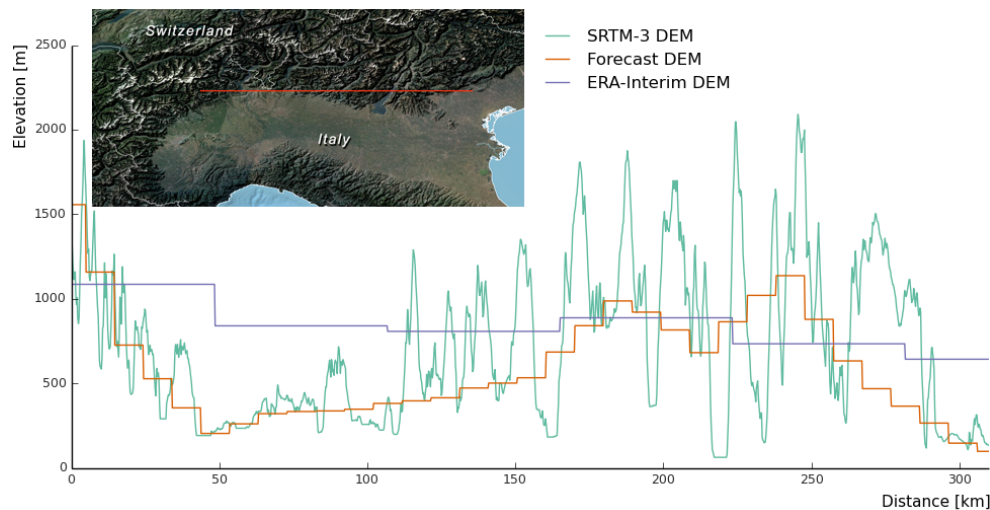
Figure 1. Comparison of topographic profiles along the transect with constant latitude 45 ° 48′N, running from 8° E to 12° E. The profile in green was calculated using SRTM-3 digital elevation model with a spatial resolution of 3″. The orange line using the DEM of the ECMWF operational forecast (7.5′ resolution) while the blue line represents the elevation as extracted from the DEM used by the ECMWF ERA-interim reanalysis (45′). These data can be downloaded from the ECMWF web site: www.ecmwf.int.

The literature shows that nearest neighbour methods, splines, regression, inverse distance weigthing and kriging have traditionally been the most common techniques used to interpolate ground stations data. Regarding the spatial and temporal enhancement, the improvements achieved in the last two decades are significant regarding both monthly and daily averages temperature. The first global grid datasets of monthly land surface temperature date back to the beginning of the 1990s [8,9], both at 0.5° resolution. In [10] the construction of a 10' latitude /longitude data set of 8 mean monthly surface climate elements is described. This dataset, based on 30 years of station data, builds on an earlier work by the same authors [11] of $30' \times 30'$ resolution. [12] used a thin-plate smoothing spline on a collection of public meteorological data sets of monthly records to produce global datasets at 30″ (called 1 km resolution) for the period from 1960 to 1990. A downscaling procedure was developed in [13], which used 3D spatial interpolation to produce a map of monthly average temperatures for Europe.

Average daily temperature (together with daily temperature range and daily precipitation) was developed for the first time on a global $1° \times 1°$ gridded data set by Piper and Stewart [14]. Haylock *et al.* [15] produced European coverage maps of daily mean, minimum, and maximum temperatures and precipitation with 0.25° and 0.5° resolution using the European Climate Assessment and Dataset Project. These maps were generated by first estimating monthly averages; daily anomalies from those averages were then interpolated using kriging and added back to the monthly estimates. [16] produced 0.5° global hourly land surface temperature datasets based on four reanalysis products (MERRA, ECMWF Re-Analysis ERA-40, ECMWF Interim Re-Analysis, and NCEP-NCAR reanalysis) by means of horizontal downscaling, temporal downscaling and monthly bias correction adjusting the final products to match the Climate Research Unit Time Series [17] monthly mean.

More recently, other authors suggest procedures to calculate temperature gridded datasets based on remotely detected thermal data. That is the case of [18], which describes a method to parameterized T_{2m} based on land surface temperature allowing the estimation of a continuous T_{2m} field with a temporal resolution of 30 min and spatial resolution of 1 km in Central Europe. At European scale, a similar approach, based on MODIS (moderate-resolution imaging spectroradiometer), was taken in [19], resulting in high resolution land surface temperature time series at European continental scale gaining 250 m spatial resolution and four daily values per pixel (each 8 h).

There are also regional examples evaluating spatial and seasonal variations of air temperature lapse rates. [20] used a series of linear equations to obtain the air temperature based on altitude, lapse rate and temperature at the sea level, concluding with the existance of gradual seasonal trends in monthly variations of lapse rates and the significant increased of the interpolation reliability after adding topographic information. [21] proved that the lapse rate can be calculated with a linear regression model, after examining the monthly, seasonal and annual characteristics of temperature lapse rate on the southern slope of central Himalayas based on 20 years record of surface air temperature at 56 stations in Nepal [22], in the process of interpolating and downscaling T_{2m} for the study of crop growth forecasting in Europe, proposes to used a fixed lapse rate of $-0.006\ °\mathrm{C} \cdot \mathrm{m}^{-1}$.

In this paper we suggest a rather simple procedure to estimate the vertical gradient of 2 m-temperature at global scale, and to the best of our knowledge, this is the first such attempt. Building on an idea proposed by Kunz *et al.* [23], who obtained a vertical gradient of surface air temperature for a region (Switzerland) by linear regression using data from a number of ground stations located at different elevations, we propose to calculate the time-varying temperature gradient using the high-resolution forecast data available from ECMWF by applying a linear regression procedure. The paper is structured as follows: Section 2 presents the data used for the study and the algorithm to calculate the temperature gradients. Section 3 describes the results of the validation of the temperature estimates using the vertical gradients against data from measurement stations across the world. Two examples illustrating the application of the downscaled temperature to solar energy and energy meteorology are discussed in Section 4. We present our conclusions in Section 5. The output data described in this paper will be made freely available (see link at the end of Section 5).

2. Data and Algorithm for the Surface Air Temperature Downscaling

In this section we will describe the methods we have used to downscale the 2-m temperature data so as to obtain a better resolution of the air temperature variations with elevation. Section 2.1 describes the data we have used from the ECMWF operational forecast and ERA-interim reanalysis. We then present the method to calculate vertical temperature gradients in Sections 2.2 and 2.3. A method to use these gradients together with reanalysis data is then described in Section 2.5.

Figure 2 presents a flowchart of the different steps we use to perform the vertical downscaling of the temperature data. The flowchart also shows where the different data sources are being used, and where the different steps are described in the text.
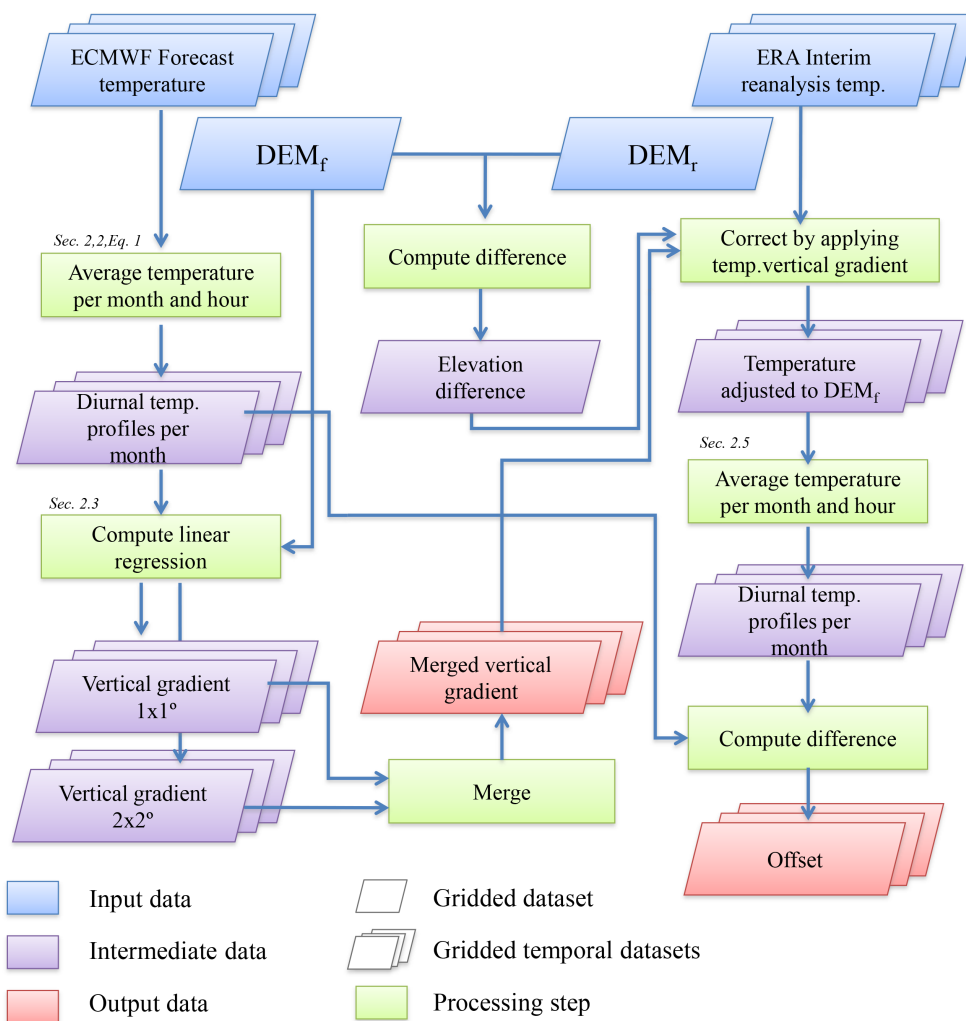
Figure 2. Flowchart of the proposed method to calculate the time-varying vertical gradients and its application to downscaling 2-m temperature. Annotations in italic refer to sections in which that particular processing step is explained.

2.1. Input Data

2.1.1. ECMWF ERA-Interim Data

The ECMWF ERA-Interim data set is a Reanalysis data set, obtained by running numerical weather predictions for the past and using actual station measurements to correct the results [4]. Data are available every 3 h from 1979 to the present. The spatial resolution is $0.75°$ ($45'$) on a regular latitude/longitude grid, corresponding to about 83 km at the equator.

2.1.2. ECMWF Operational Forecast Data

The ECMWF operational forecast data [24] consist of gridded data on a regular latitude/longitude grid. From the period January 2010 onwards, the spatial resolution is $0.125°$ ($7.5'$), or 6 times higher than ERA-Interim. The temporal resolution is 3-hourly in the earlier part of this time interval, later hourly.

Unlike the ERA-Interim, these data are outputs of the forecast models and have not been corrected using actual measurement data as is the case with the reanalysis.

Together with the temperature (and other) data of the forecast, ECMWF also specifies a DEM for both the forecast and the reanalysis.

The ECMWF forecast is run every 12 h (starting at 0:00 GMT and 12:00 GMT) and the results of the forecast are available with a 3-h time step. Each forecast run is preceded by an analysis based on conditions at that time. In principle, we could use the analysis results instead of the forecast for the 0:00 GMT and 12:00 GMT time slots. However, we have chosen the latter for greater consistency with the other time slots. Slots between 12:00 and 21:00 GMT are taken from the forecast starting at 12:00.

In the following, the ECMWF operational forecast data will be denoted as *forecast* data, and the associated DEM will be called the *forecast DEM*, Z_f . Similarly, the ECMWF ERA-Interim data will be referred to as *reanalysis* data, and the associated DEM will be called the *reanalysis DEM*, Z_r.

In contrast to the ERA-Interim reanalysis data, the operational forecast data are not publicly available. This may limit the usefulness of the data to researchers, a point that will be discussed in more detail below.

2.2. Calculation of Monthly Averages of Diurnal Temperature Profiles

The first step in the calculation of the vertical 2-m temperature gradients consists of constructing maps of monthly average temperatures for each 3-hourly time slot during the day. Four years of ECMWF operational forecast data were used for this, covering the period 2010–2013. The average temperature map $< T_{f:m,h} >$ for month m and hour h is calculated from the 3-hourly forecast temperature maps $T_{f:y,m,d,h}$ for year y, month m, day d and hour h as:

$$\langle T_{f:m,h} \rangle = \frac{1}{N_y N_{d,m}} \sum_{y=Y_s}^{y=Y_e} \sum_{d=1}^{N_{d,m}} T_{f:y,m,d,h} \tag{1}$$

Here, Y_s, Y_e are the start and end year in the data set, $N_{d,m}$ is the number of days in month m and N_y is the number of years used for the study. The result of this calculation is a set of 12×8 raster maps of temperature with the same spatial resolution as the original forecast data.
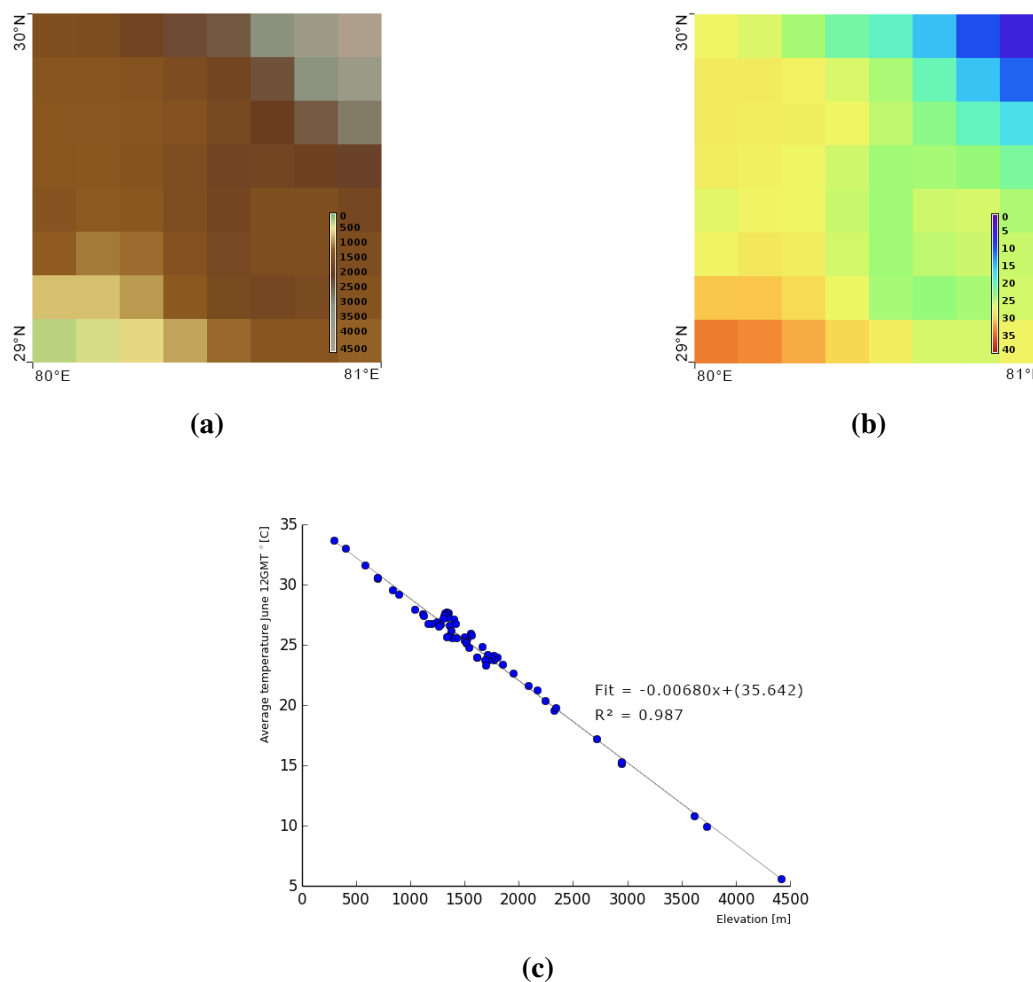
2.3. Calculation of vertical Temperature Gradients

In the next step, each of the 12×8 average temperature gridded surfaces is divided into square tiles on the latitude/longitude grid. If the tile size is $1° \times 1°$, there will be 64 grid points in each tile, while for a $2° \times 2°$ tile there are 256 grid points in each tile. Inside each tile we then have a number of average temperature values together with the corresponding elevation, Z_f, from the DEM of the ECMWF forecast grid. From pairs of values of Z_f and $< T_{m,h} >$, the gradient $\gamma_{m,h}$ is then estimated using linear regression. Two limitations have been imposed on the calculation:

- If the elevation variation within the tile is small, the calculation of the gradient becomes very uncertain. For this reason, the calculation is not performed if the range of elevation values in the tile is less than 200 m.

- Forecast grid points over the sea have been excluded from the calculation. This is done to minimize the effects of the horizontal temperature variations near the coast, which otherwise may introduce spurious effects in the calculation of the gradients. As a result of this procedure, the number of data points in a given tile may be less than stated above, which in turn may affect the reliability of the linear regression procedure.

As an illustration of the process, Figure 3 shows the DEM for a $1° \times 1°$ tile together with the average temperature at 12:00 GMT in June calculated using Equation 1. Plotting the 64 temperature values against the 64 elevation values produces the graph in Figure 3c. It is clear that at least for this tile the relationship between elevation and average temperature is very nearly linear and the linear regression finds the gradient with high accuracy. Of course, this particular tile lies in a mountain area with large elevation differences. Not all tiles will have such a smooth dependence of temperature on elevation.

(a)

(b)

(c)

Figure 3. Example of the calculation of the vertical gradient. (**a**) shows the *forecast DEM* for a $1° \times 1°$ tile located at $29°$ N to $30°$ N and $80°$ E to $81°$ E; while (**b**) shows the corresponding average temperatures at 12:00 GMT in June; (**c**) shows a graph of temperature values against the corresponding elevations, together with a least-squares linear fit to the data. The slope of the fit gives the local vertical temperature gradient.

The linear regression calculation has been performed using the Python SciPy library [25], using the *stats.linregress* function. The calculation returns the linear slope (gradient) and offset, as well as the statistical p-value and R^2. If $p > 0.05$ for a given tile, the calculation of the gradient is discarded (and hence the gradient will implicitly be zero for that hour and month).

The gradients for month m and hour h are termed $\gamma_{m,h}$ and consist of gridded data sets with spatial resolution equal to the size of the tiles used for the calculation of the gradients.

For the present study we have calculated the gradients only north of $56°$ S, which excludes Antarctica from the study area. This is done mainly due to the lack of measurement stations in that region.

For the actual calculation of the gradients we used the *forecast* data for four years (2010–2013) to calculate the vertical temperature gradient using the method outlined in Section 2.3. The calculation was made using both $1° \times 1°$ tiles and $2° \times 2°$ tiles. The tiles for which the gradients were calculated were chosen as those in which the *forecast* elevation Z_f varies by more than 200m within the tile.The gradients were calculated using the monthly averaged daily temperature profiles as described in Section 2.2. In this way the gradient is calculated for each of 8 daily slots for each of the 12 months, for a total of 96 gradient fields. For each field, the gradient in a tile was discarded if the output of the linear regression was not statistically significant from zero Section 2.3. Thus, a given tile may have a gradient in the temperature field for some months and hours but not for others.

The resulting gradient fields only cover some of the land surface of the earth. Considering the study area ($90°$ N to $56°$ S), the DEM provided with the *forecast* has 1,122,142 cells on land out of a total of 3,362,672 cells (each cell is $7.5' \times 7.5'$). Using $1° \times 1°$ tiles, the number of cells on land for which at least one gradient is present is 638,782 or 57% of cells, while for the $2° \times 2°$ tiles the number of cells with at least one gradient is 864,767 or 77%.

As will be shown in the validation results below, the gradients calculated on$1° \times 1°$ tiles generally yield slightly more accurate downscaling results than those using the $2° \times 2°$ tiles. On the other hand, the $2° \times 2°$ tiles have a wider coverage. We have therefore also constructed a merged gradient data set using the following rules:

- For a given $1° \times 1°$ tile, if the gradient has been calculated using the $1° \times 1°$ subdivision, use that gradient,
- Otherwise, if the gradient has been calculated using the $2° \times 2°$ subdivision, use that gradient,
- If a given tile has no calculated gradients, there is no correction to the temperature for that month and hour.

The coverage of this merged data set is equal to that of the $2° \times 2°$ gradient data sets.

2.4. Calculation of the Downscaled Temperature Field

Using the vertical temperature gradient calculated for each month and 3-hourly time slot, the corrected temperature values for a given location with elevation z for the time t can be calculated as:

$$T_{corr}(t) = T_f(t) + \gamma_{m,h}\left(z - Z_f\right) \tag{2}$$

Here m and h denote the month and hour corresponding to time t.

2.5. Calculation of a Downscaled Temperature Field Based on ERA-Interim Temperature Data

The above calculation makes it possible to construct a temperature field with higher spatial resolution taking into account the dependence of temperature on elevation. However, the model depends both on the average vertical gradients derived from the *forecast* data and on the 3-hourly *forecast* data. The *forecast* data are not freely available, and therefore the gradient fields will have limited utility to users without access to these data.

To overcome this limitation we also investigate the possibility of applying the gradient fields to the *reanalysis* data. Equation (2) will be used for this, only substituting the *forecast* temperature T_f by the *reanalysis* temperature T_r:

$$T_{corr,r}(t) = T_r(t) + \gamma_{m,h}(z - Z_r) \tag{3}$$

Here, Z_r is the elevation from the ERA-interim data set.

Due to the coarse resolution of the *reanalysis* data, some areas may be poorly represented by these data, in particular areas such as coastal zones where there are large horizontal temperature variations even at similar elevations. It will therefore also be investigated whether the temperature correction in Equation (3) can be improved by adding an offset calculated from the average difference between the *forecast* and *reanalysis* data with the *reanalysis* data first corrected to the forecast elevation using the calculated temperature gradient:

$$\langle \delta T_{m,h} \rangle = \frac{1}{N_y N_{d,m}} \sum_{y=Y_s}^{y=Y_e} \sum_{d=1}^{N_{d,m}} (T_{f:y,m,d,h} - T_{r:y,m,d,h}) + \gamma_{m,h}(Z_r - Z_f) \tag{4}$$

Here, $T_{r:y,m,d,h}$ is the *reanalysis* temperature for year y, month m, day d and hour h. Equation (3) can then be modified as:

$$T_{corr,r}(t) = T_r(t) + \langle \delta T_{m,h} \rangle + \gamma_{m,h}(z - Z_r) \tag{5}$$

2.6. Limitations of the Model

The temperature downscaling model as presented has a number of limitations and the results should not be used uncritically. Among these are:

- The calculation of the vertical temperature gradients is based on 4 years of data. This time interval is far too short to note any trend in the gradients. Indeed, since we have used the full 4 years of data to produce a set of constant gradients for each month and 3-hourly slot, any trend has been automatically removed. The gradient data are therefore not intended to be used with temperature data from far in the past (more than, say, 20 years).
- The vertical gradient is not the only source of spatial variation in T_{2m}. A number of geographical features may cause the average temperature to vary significantly. This is especially the case for land surfaces close to the sea or to large lakes. Mountains may also show large variations in temperature at the same elevation due to shadowing, where the side of the mountain facing the equator would tend to be warmer than the side facing the opposite way.

- The vertical gradients are long-term averages, but the instantaneous situation may be very different, especially in the case of temperature inversions. It should therefore not be expected that using the gradients together with a low-resolution data set would faithfully reproduce the temperatures at a given instant in time. The improvement may well be seen only in a statistical sense.

3. Validation of the Downscaling Method

3.1. Source and Selection of Suitable Ground Station Data

The National Oceanic and Atmospheric Administration (NOAA) offers free access to a variety of meteorological and climatological information through their data portal, called the National Climatic Data Center (NCDC). One of the accessible databases is The Integrated Surface Database (ISD), formerly called Integrated Surface Hourly database (ISH). It is the result of a project started in 1998 which goal was to merge numerous surface hourly data sets produced by over 100 different sources into a common format and data model, reducing inconsistencies and applying quality controls [26]. ISD comprises over 2 billion surface observations from more than 20,000 stations worldwide, with some having data as far as 1901. At present there are more than 11,000 active stations that are updated daily in the database [26]. ISD includes some of the most common meteorological parameters such as wind speed and direction, wind gust, dew point, cloud data, pressure, visibility, precipitation, air temperature and many others.

A routine written in Python-Pandas [27] was used to download the ground station data from NOAA's FTP server and to select the ones suitable to our purposes. The following steps were followed:

- Temporal coverage: Data files are organized in folders by years. The folders corresponding to the period 2010–2013 were checked, obtaining over 50,000 files from more than 15,000 stations.
- Correct geographical location attributes: The stations code was then used in combination with the metadata to select those whose coordinates latitude/longitude were not null or equal to 0. Using a GIS software, each pair of coordinates was projected into a map. 14,671 stations were found to contain measures for the period 2010–2013 and admisible locality (coordinates).
- Minimum number of measurements: In order to assure the statistical representativeness of a station, a minimum number of measurements restriction was imposed. Firstly, all the files corresponding to the same station were merged into a single file. Using this file, the number of valid measurements (not null and in the range of –200 to 200 °C) for a particular hour per month were summed up. A station was kept if that number was greater than 30. As a result of the former criteria, 11,752 stations were selected. Excluding stations south of 56° S leaves 11,555 stations.

For the present study it is of course important that the elevation of the station is correct. As the latitude and longitude of the stations are given with a precision of only 1′ there is no way to directly compare the stated elevation with the high-resolution DEM. In order to exclude at least some stations with wrong elevation (or location) the following method was applied:

- The 3″ DEM was divided into square regions of size 1′.
- For each 1′ square the minimum and maximum elevation was found.

- For each station, the station elevation was compared with the minimum/maximum values for the corresponding square. If the station elevation was outside the minimum/maximum interval by more than 100 m, the station was excluded.

Since the SRTM3 DEM contains data only for the area between 60° N and 60° S, the check of station elevation can only be done for stations within these boundaries (90%; 1269 stations are located outside those limits). Using this criterion, 240 stations were eliminated, leaving 11,315 stations.

The lists of stations used in the validation can be found in the supplementary material online [28].

The geographic distribution (see Figure 4) is neither random nor homogeneous. There are large gaps in areas with lowest population densities such as the Amazon Basin, Siberia, Greenland or the Sahara, contrasting with higher number per km^2 of stations in United States, Japan and Europe. Investments in meteorological observation networks coupled with extreme conditions seems the most plausible cause to explain this spatial pattern.

Almost 10% of the meteorological stations are located above 1000 m, 1300 (11.5%) lie between 500 and 1000 m, and 3719 (32.9%) between 300 and 500 m. 68.7% (7732) are below 300 m. Of the total, 7450 (66%) are located on flat terrain (less than 2° slope); 3630 (32%) are on a 2°–15° slope, and 2% are located in slopes steeper than 15°. Regarding aspect, 17.9% (2030) of the stations are on slopes face the Poles, while 59.4% (6719) face the Equator.
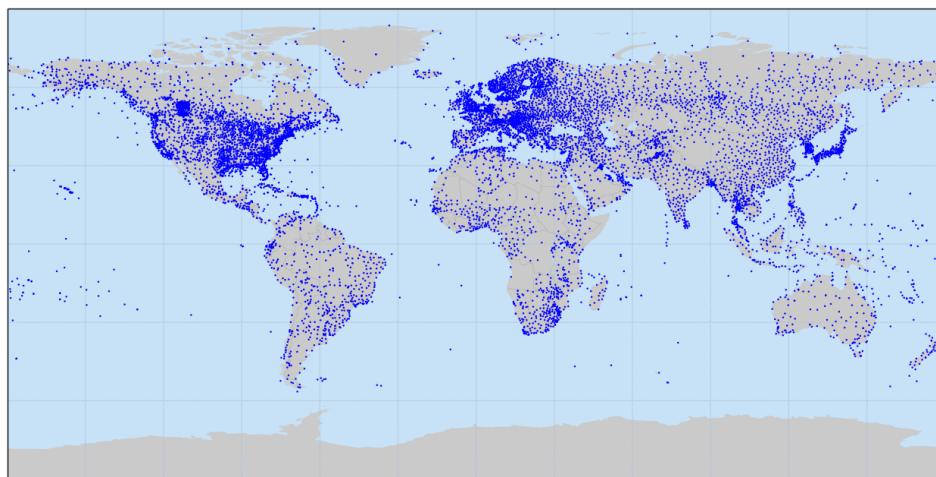
Figure 4. Meterological stations used for the validation.

A potential problem should be noted before presenting the results of the validation. The ECMWF *forecast* and *reanalysis* procedures both use ground station data as part of the analysis. According to the information on the ECMWF web site, the stations used are the SYNOP (Surface Synoptic Observations) stations of the World Meteorological Organization (WMO) [29], many of which are also present in the data set we have used in the validation. This could lead to an underestimate of the uncertainty the user will face when using the forecast data, especially in areas with few stations. However, it should still be possible to assess the *improvement* in the estimates that results from the downscaling procedure. If the deviation of the estimate from the measurements is less when the downscaling is applied, it will be an indication that the *forecast* in itself does not adequately represent the conditions at the site of the station.

Nevertheless, we will try to perform some of the validation also using a reduced set of stations which are not present in the WMO list of SYNOP stations. Out of the 11,315 stations used, 3730 are not in the SYNOP list. Most of these stations are in the USA so this subset is not sufficient for a detailed validation. However, it may serve to see if the validation results at these sites are significantly different from the overall results.

3.2. Validation Criteria

A number of statistical quantities are used to assess the quality of the *forecast* and the downscaled *forecast* data. These quantities are widely used and well known. Appendix A contains a brief overview of the methods we have employed.

In the following sections we will first attempt a validation of the *forecast* data without downscaling, choosing stations with little elevation difference to the forecast DEM. This is done in order to get a baseline for the accuracy of the forecast data. This will then be followed by the validation of the downscaling procedure.

3.3. Validation of Forecast Data without Vertical Gradient Correction

Before assessing the accuracy of the downscaled *forecast* data against station measurements it is useful to consider the accuracy of the *forecast* data when there is no elevation effect, *i.e.*, when the elevation difference between the elevation of the station and elevation of the corresponding pixel in the *forecast* data set is small.

For this purpose we selected a subset of the stations for which $|Z_s - Z_f| < 100$ m. Out of the original 11,315 stations selected for the validation, 8656 stations fulfill this criterion. For each station we then calculated MBD Equation (9), MAB Equation (10) and RMSD Equation (11).

To get an overview of these results we calculated also gMBD, gMAB and gRMSD Equations (12–14). Here we find that the overall bias of the *forecast* for all the stations is gMBD = -0.12 °C, while gMAB = 0.53 °C and gRMSD = 0.71 °C.

Out of the 8656 stations 7446 or 86% have a *forecast* bias between -1 °C and $+1$ °C. Only 109 stations are more than 2 °C away from the *forecast* data. These stations are shown in Figure 5.

The locations of these stations with high bias does not show a clear pattern. In particular, it does not seem to be the case that coastlines or mountain areas are overrepresented among the stations with high absolute bias. It is of course also possible that the data from the stations may not be correct, for instance if the supplied data do not have the correct timestamp.

The gRMSD value gives an estimate of the standard error of the annual average *forecast* temperature. A 95% confidence interval would be approximately twice this value, so in 95% of locations the error in the annual average temperature will be less than ~1.4 °C. With a lapse rate of 0.6 °C/100 m this uncertainty would correspond to an elevation difference of 200 m. This therefore provides a useful cutoff for how small elevation differences can be used to determine the vertical temperature gradient.

Overall, it can be concluded that for many practical purposes, the uncertainty in the *forecast* data is low enough to make the data useful. This uncertainty also gives us an idea of what an acceptable uncertainty would be for the downscaling procedure.

As noted in the previous section, if a part of the stations have been used to produce the *forecast* data, this would of course tend to lower gRMSD and so it is possible that the actual uncertainty in areas with few stations will be higher. The statistical measures were therefore also calculated for the subset of stations that are not part of the SYNOP stations. Using the same criteria $|Z_s - Z_f| < 100$ m, the number of non-SYNOP stations is 3133. For these stations we found gMBD = $+0.006$ °C, while gMAB = 0.54 °C and gRMSD = 0.72 °C. These results are practically identical to the results from the larger set of stations.
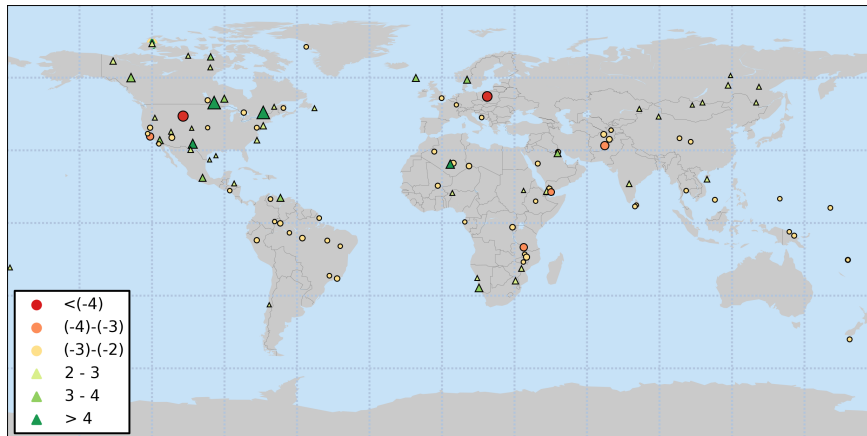
Figure 5. Stations for which the absolute annual mean bias deviation exceeds 2 °C. The color scale indicates the MBD for the *forecast* at each station. The circle represents negative values and the triangle positive ones.

3.4. Validation of the Downscaling Procedure

Out of the stations selected according to the method described in Section 3.1, we selected subsets corresponding to all the stations lying within the area covered by the $1° \times 1°$ tiles and the $2° \times 2°$ tiles, respectively. In the $1° \times 1°$ tiles are 6263 stations, while the $2° \times 2°$ tiles contain 8087 stations (Figure 6) .
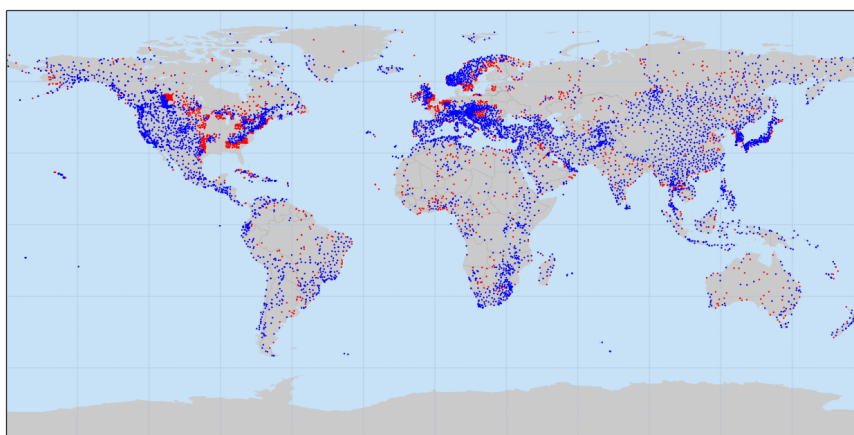
Figure 6. Location of weather stations from which data was used in the validation. In blue the 6263 stations lying in the area covered by $1° \times 1°$ tiles. In red the stations lying in the area covered by the $2° \times 2°$ tiles not included in the previous subset (1824).

For each station, the 3-hourly *forecast* data were compared with the available station measurements. Then the downscaled temperature was calculated for the station elevation using Equation (2). For each station, MBD, MAB and RMSD were calculated. From these individual station values we then calculated the following global quantities:

- gMBD Equation (12)
- gMAB Equation (13)
- gRMSD Equation (14)
- <MAB>, the mean of the individual station MAB values
- <RMSD>, the mean of the individual station RMSD values

These values were calculated for the *forecast* against station values, and for the downscaled values against the station values. This was done both for the $1° \times 1°$ gradients and the $2° \times 2°$ gradients. The results are shown in Table 1.

In order to make a direct comparison between the downscaling using $1° \times 1°$ gradients and the $2° \times 2°$ gradients, the same statistical measures were also calculated using the $2° \times 2°$ gradients but choosing only the 6263 stations present in the $1° \times 1°$ tiles. This is shown in Table 1 in the column denoted $2° \times 2°$ C (C as in common). The statistical values for the rest of the stations (*i.e.*, stations present in the $2° \times 2°$ tiles but not in the $1° \times 1°$ tiles), are shown in the column denoted $2° \times 2°$ E (E as exclusively in $2° \times 2°$ tiles).

Table 1. Statistical measures for the *forecast* and downscaling procedures, as compared against station measurements. All values except the elevations are in degrees celsius. Columns named "F" denote the validation of the *forecast* data while the columns named D_f give the validation results for the downscaling procedure. $< z_s >$ in the *forecast* columns gives the average elevation for the *forecast* DEM at the stations used, while $< z_s >$ in the downscaling columns give the average elevation of the stations used for the validation in the given column.

Statistical measures	$1° \times 1°$		$2° \times 2°$		$2° \times 2°$ C		$2° \times 2°$ E		Combined	
N_s	6263		8087		6263		1824		8087	
	F	D_f	F	D_f	F	D_f	F	D_f	F	D_f
$< z_s >$ (m)	598	510	510	442	598	510	206	207	510	442
gMBD	–0.62	–0.11	–0.49	–0.12	–0.62	–0.14	–0.06	–0.07	–0.49	–0.10
gMAB	1.13	0.67	0.99	0.67	1.13	0.71	0.52	0.52	0.99	0.64
gRMSD	1.79	0.94	1.61	0.97	1.79	1.02	0.72	0.72	1.61	0.89
<MAB>	2.10	1.77	1.97	1.76	2.10	1.82	1.55	1.54	1.97	1.72
<RMSD>	2.63	2.34	2.50	2.32	2.63	2.40	2.07	2.06	2.50	2.28

The results in Table 1 are obtained using all the available station data in the regions where the gradients are calculated, whether or not the elevation difference is large enough that the downscaling procedure has any noticeable effect. We therefore repeated the analysis, but restricted to stations for which $|Z_s - Z_f| > 300$ m. These results are shown in Table 2.

Table 2. Statistical measures for the *forecast* and downscaling procedures, as compared against station measurements, using stations for which the altitude difference between station and *forecast* has an absolute value greater than 300 m. All values except the elevations are in degrees celsius.

Statistical measures	1° × 1°		2° × 2°		2° × 2° C		2° × 2° E		Combined	
N_s	854		859		854		5		859	
	F	D_f	F	D_f	F	D_f	F	D_f	F	D_f
$< z_s >$(m)	1272	933	1265	931	1272	933	102	624	1265	931
gMBD	–2.28	–0.20	–2.26	–0.33	–2.28	–0.32	1.67	–1.46	–2.26	–0.21
gMAB	3.45	1.02	3.44	1.21	3.45	1.21	1.86	1.46	3.44	1.02
gRMSD	4.11	1.42	4.11	1.70	4.11	1.71	2.33	1.67	4.11	1.42
<MAB>	4.06	2.33	4.05	2.57	4.06	2.57	3.29	2.32	4.05	2.33
<RMSD>	4.62	3.02	4.62	3.27	4.62	3.27	3.94	3.16	4.62	3.02

The results show that the downscaling procedure yields somewhat better results in all the stastical measures. In particular, when the elevation difference is large, the improvement is noticeable. The values gMAB and gRMSD give an measure of the overall uncertainty in the long-term average temperature estimates. Here we see a reduction from 3–4 °C to 1–1.5 °C.

A comparison of the 1° × 1° columns to the 2° × 2 °C columns shows that the gradients calculated using the 1° × 1 ° tiles give slightly better results. For the areas not covered by the 1° × 1° tiles (the rightmost columns in Tables 1 and 2) there is no overall improvement when considering all the stations in these areas. However, if only the (few) stations with large elevation difference are considered, there is a moderate improvement.

3.5. Validation of the Downscaling Using ERA-Interim Temperature Data

Section 2.5 described a proposed method to use the ERA-Interim reanalysis data with the downscaling procedure. The validation exercise described above for the *forecast* data can equally well be used with the results of applying Equations (3) and (5).

Before applying the downscaling procedure we first investigated the accuracy of the ERA-Interim data in areas with little elevation variation, using the same approach as Section 3.3 above. In this case the number of stations with elevation within 100 m of the corresponding ERA-Interim surface elevation was 6850 stations. For these stations the overall statistical quantities are: gMBD = +0.19 °C, while gMAB = 0.81 °C and gRMSD = 1.06 °C. These values are somewhat larger than the corresponding values for the validation of the *forecast* data in Section 3.3. One reason for this may be that the stations typically are further away from the points at which the *reanalysis* data are sampled.

The downscaling procedures of Equations (3) and (5) were then applied to the 8087 stations lying within the 2° × 2° tiles. The results are shown in Tables 3 and 4, the latter for stations with elevation difference more than 300 m from the corresponding *reanalysis* data point.

Table 3. Statistical measures for the *reanalysis* and *reanalysis* downscaled temperature using the ERA-Interim *reanalysis* data, as compared against station measurements, using both the downscaling and the downscaling + offset methods.

Statistical measures	Correlation Coefficient (R)	Downscaling	Offset and Downscaling
$< z_s >$	552	442	442
gMBD	–0.47	0.12	0.02
gMAB	1.43	0.85	0.72
gRMSD	2.16	1.14	1.01
<MAB>	2.44	2.02	1.97
<RMSD>	3.04	2.65	2.59

Table 4. Statistical measures as in Table 3, but considering only stations with an elevation difference between *reanalysis* and station greater than 300 m.

Statistical measures	Correlation Coefficient (R)	Downscaling	Offset and Downscaling
$< z_s >$	1032	720	720
gMBD	–1.51	–0.04	–0.002
gMAB	2.94	0.99	0.92
gRMSD	3.80	1.38	1.31
<MAB>	3.74	2.33	2.29
<RMSD>	4.35	3.03	2.97

The downscaling approach combined with the temperature offset gives slightly better results than the downscaling alone, as seen in Table 3. For the stations with significant elevation difference the two methods give almost identical results, possibly because there are few stations in areas such as coastal zones where the offset would be significant.

Overall, the statistical measures are a bit higher with the *reanalysis* than the results in Tables 1 and 2, when comparing with the downscale results with $1° \times 1°$ tiles. However, when the downscaling is combined with the offset calculation, most of the difference disappears.

*3.6. Monthly Variation in the Statistical Measures*

The validation results up to here have been concerned only with the annual averages of the statistical quantities. However, also the monthly variation of the statistics may be of interest. The results of the validation of the annual averages showed that there is little difference between using the gradients with the *forecast* data or with the *reanlysis* data with offsets. We will therefore only present the results here for the latter combination.

Figures 7 and 8 show a box-and-whisker plot of MBD values for the stations, with results shown for each month for the *reanalysis* alone, *reanalysis* + gradients and finally with also the offset applied.

From both figures it is clear that the extent of the boxes are considerably reduced when the gradients or gradients + offset are applied. Also the mean values of MBD are closer to zero for nearly all months,

especially for stations with large elevation differences (Figure 8). The large reduction in overall annual MBD seen in Table 4 is also evident Figure 8 for nearly all months.
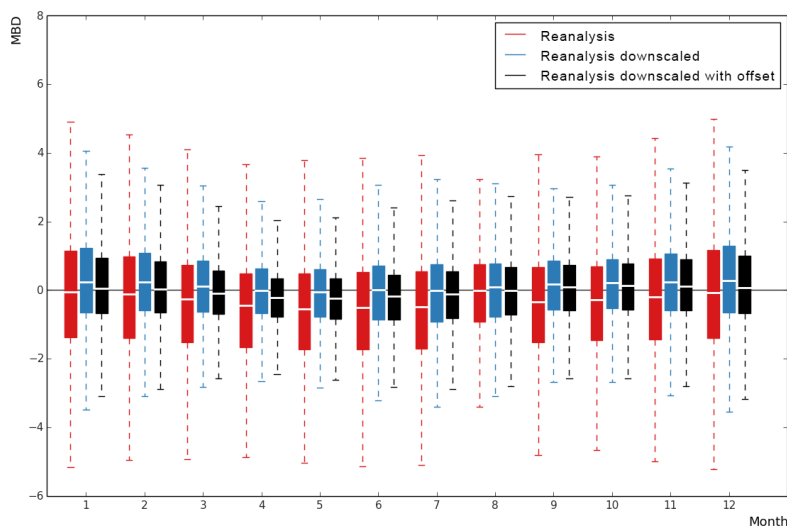
Figure 7. Box plot comparing MBD reanalysis and downscaled, with and without offset.The white line in the boxes denote the overall MBD value for that month, the extent of the boxes show the extent of 95% the MBD values for the stations, while the whiskers indicate the largest outliers.
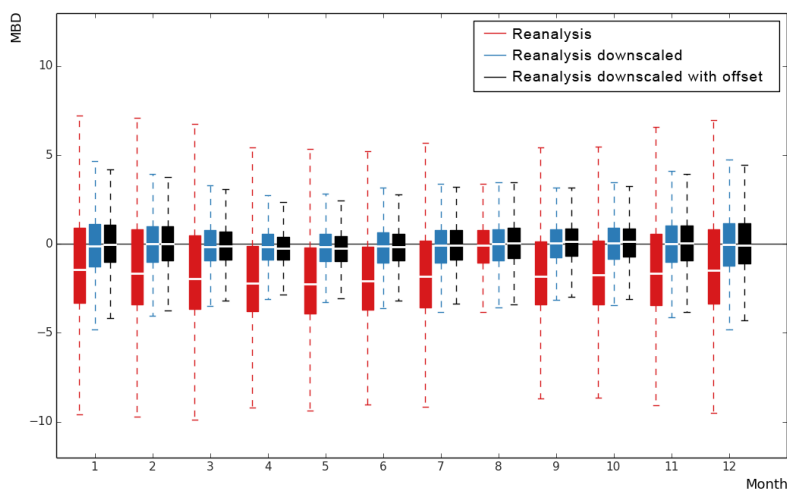
Figure 8. Box plot comparing MBD reanalysis and downscaled as in Figure 7, with and without offset, but considering only stations with an elevation difference between reanalysis and station greater than 300 m.

The preceding discussion has considered results from all stations regardless of geographical location. To investigate the performance of the downscaling in different climatic zones we divided the stations into 5 groups according to the 5 main climatic divisions of the Köppen–Geiger climate zones.

In this, we have considered only stations with an elevation difference > 300 m, and restricted the area investigated to north of 20° S, to avoid mixing different seasons in the two hemispheres.The number of stations in each zone is given in Table 5.

Table 5. Köppen-Geiger main climatic zones and the number of temperature stations in each zone.

Climatic Zone	Main Zone Code	N_s
Equatorial	A	160
Arid	B	230
Warm temperature	C	680
Snow	D	561
Polar	E	123

The monthly values of gRMSD are shown in Figure 9 for each climatic zone for the reanalysis only and for the reanalysis combined with downscaling and offset. From the results it is clear that for zones A, B, and C there is not a strong seasonality in the accuracy for the downscaled values, though interestingly, there is some seasonal variation in arid areas in the original reanalysis data. In contrast, in the colder zones D and E the downscaling procedure works well in the summer half of the year but produces poorer results in winter. However, except for the winter months in the polar areas (E) there is still a considerable improvement in accuracy.
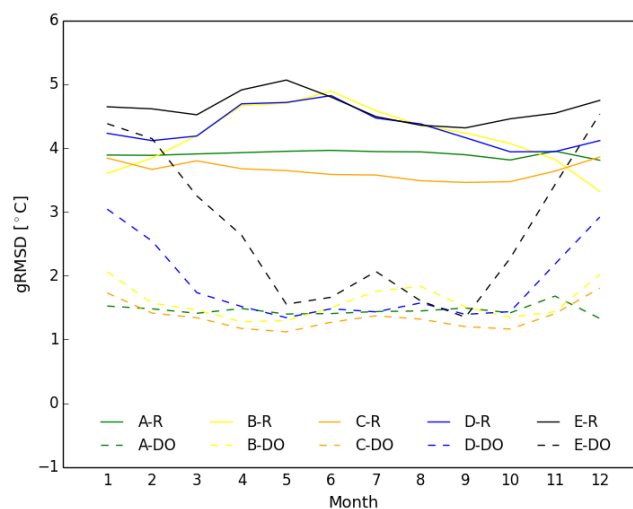

Figure 9. Graph comparing gRMSD reanalysis and downscaled with offset, dividing stations according to the climatic zone (Koeppen-Geiger main zones), considering only stations with an elevation difference between the reanalysis DEM and station greater than 300m. Solid lines are the values for the reanalysis only, dashed lines denote the values using downscaling and offset.

3.7. Variation in Results with Elevation

The primary aim of this study is to improve the estimate of 2-meter temperature at different elevations. It therefore makes sense to see how the deviations from station values depend on the elevation difference used for the downscaling. To study this, we have used the annual MBD values for all stations used to validate the *reanalysis* with downscaling and offset as in the previous section.

Figure 10 shows scatter plots of the annual MBD values as a function of the elevation difference between the station locations and the corresponding *reanalysis* DEM. The graph shows MBD values for both the *reanalysis* alone and with the downscaling. For both sets of points we have also performed a linear least-squares fit which is also shown.

It is clear that the higher the difference in elevation, the more the *reanalysis* tends to overestimate the temperature (positive MBD). The linear fit has a slope of 0.45 °C/100 m. With the downscaling the trend with elevation difference is much lower. There is however a slight negative trend with slope –0.036 °C/100 m which is statistically significant (standard error of the slope is about 10%). This indicates that the downscaling method slightly overcompensates for elevation.
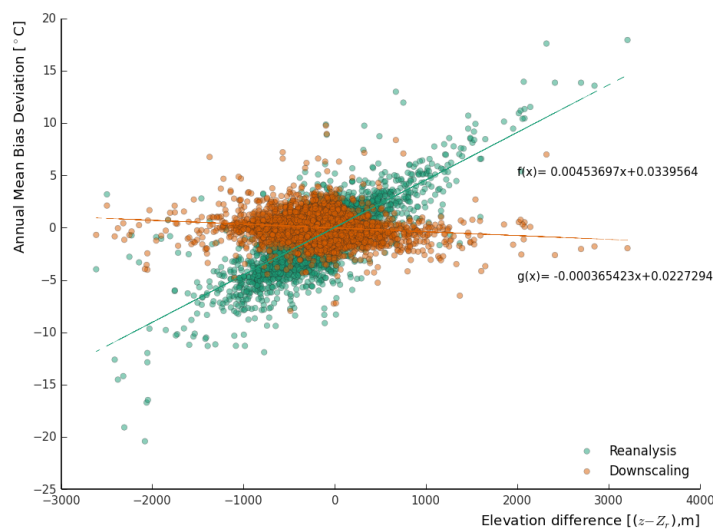

Figure 10. MBD values for each station location as a function of elevation difference between stations and reanalysis DEM. Red points are for the *reanalysis* alone, green points with the downscaling procedure with offset. Linear fits to the two data sets are also shown.

### 3.8. Dependence of Bias Deviation on the Magnitude of the Gradient

Often, the vertical gradient of the 2-meter temperature has a value rather close to the "traditional" value of about –0.6 °C/100 m. However, as will be shown in the next section, there are strong geographical and temporal variation in the gradient. Here we will look at how the magnitude of the gradient affects the downscaling results.

From the values of the gradient at each pixel for each hour and month we have calculated the probability density function (PDF) of the gradient values. This was also done for the yearly average gradient values at each pixel. Figure 11 shows the resulting PDFs. The figure also shows the PDF of the yearly average gradient values at the locations of the 6263 stations used for the validation of the downscaled *forecast* values in Table 1. Most of the pixels have values close to –0.6 °C/100 m, but a significant fraction have values below –0.8 °C/100 m or above –0.4 °C/100 m. This is especially the case for the hourly and monthly gradients. The PDF of the gradient values at the stations shows that the station locations are representative of the variation in temperature gradient.
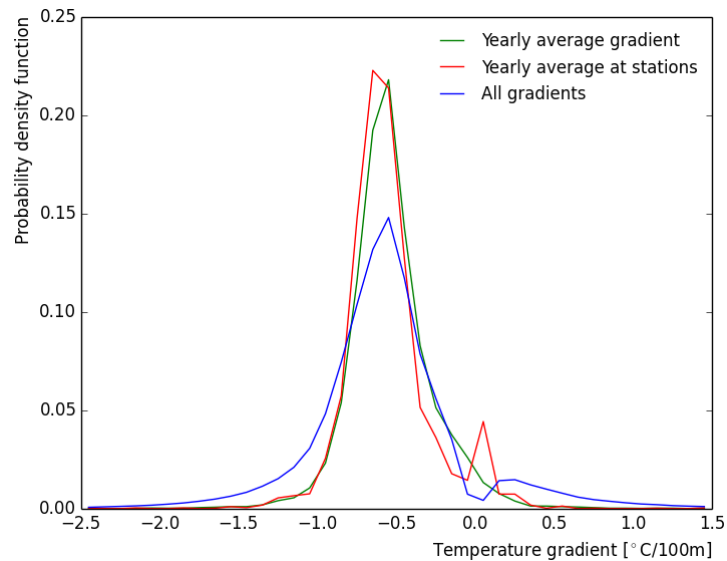
Figure 11. Probability density function (PDF) of the values of the gradients in the gradient maps. The three curves show the PDF for the values of all 8×12 gradient maps (blue), PDF for the values of a map of the yearly average gradients, and the PDF of the yearly average gradients at 6263 stations.

Figure 12 shows how the absolute value of the yearly MBD depends on the value of the yearly average gradient. The graph shows values for the *forecast* only and for the *forecast* with downscaling. To reduce clutter, we only show stations where $|Z_s - Z_f| > 300$ m. As expected, |MBD| is reduced at most stations. Interestingly, this is not only the case when the gradient value is close to –0.6 °C/100 m but is seen also further away from this value, in the range –1.0 to –0.2 °C/100 m.
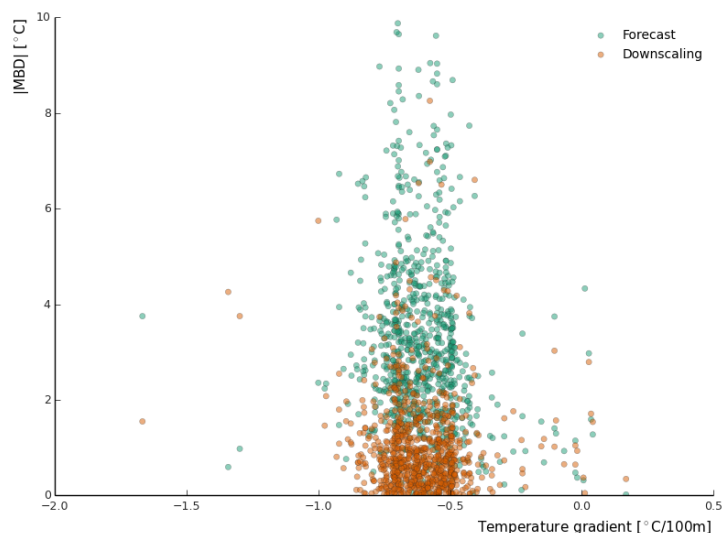

Figure 12. Scatter plot of the absolute values of the yearly MBD values, plotted against the yearly average gradient at the station location. Results are shown for the *forecast* only (red) and the downscaling results (green).

4. Results and Discussion

4.1. Geographical Variation of the Temperature Gradient

The temperature gradient fields have been calculated for each 3-hourly time slot and month, using the corresponding average temperatures as calculated by Equation (1). There are thus a total of 96 such gradient fields. The geographical areas for which the gradients have been calculated varies slightly. Out of the 52,560 $1° × 1°$ cells most will be over sea or flat terrain and have no calculated gradient. The number of cells with gradient varies from 14,144 (February at 03:00 GMT) to 14,472 (September at 00:00 GMT). This range is small (<3% variation) but there is a tendency that the smallest number of cells with significant gradients is in February/March (10 of the 20 gradients with lowest number of significant cells) and highest in September/October (15 of the 20 gradients with the highest number of significant cells).

Examples of the gradient fields are shown in Figure 13 for two time slots in September: 00:00 GMT and 12:00 GMT. The colour scale shows the gradients in units of °C/100 m.

(a)

(b)
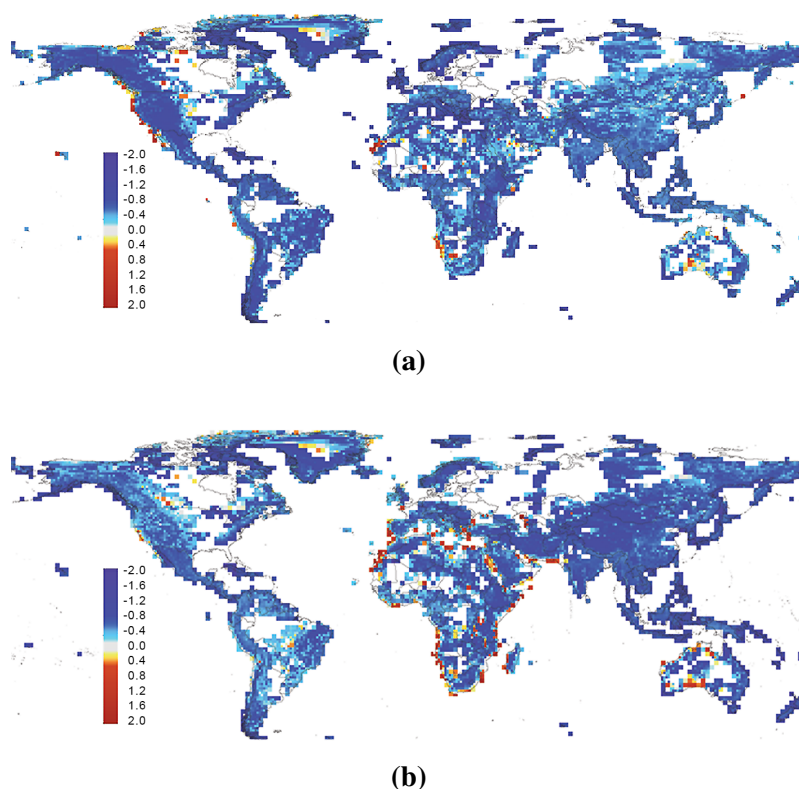
Figure 13. Temperature gradients, in °C/100 m for two hours in September: (**a**) 00:00 GMT; (**b**) 12:00 GMT. In the areas coloured white there is no significant gradient calculated for the given month and time slot.

The overall pattern is that in most areas the gradients are negative, with typical values 0.4–0.8 °C temperature drop per 100 m of elevation rise. However, some areas lie outside this range and in some cases the gradient may become positive. This is especially the case near the coastline and

in this case the effect is more marked during the day than at night. For example, there are positive gradients along the west coast of North America at 00:00 GMT which is local afternoon. At 12:00 GMT the positive gradients are more marked around the coasts of Africa and Southern Europe. The effect seems to be more pronounced at the coast of arid land areas such as Western Sahara. This is perhaps not surprising, as these areas are characterized by large daily temperature variation, and during day higher-elevation areas inland will be hotter than the sea surface near the shore and the low-lying areas right next to the shore. This will give rise to a positive temperature gradient.
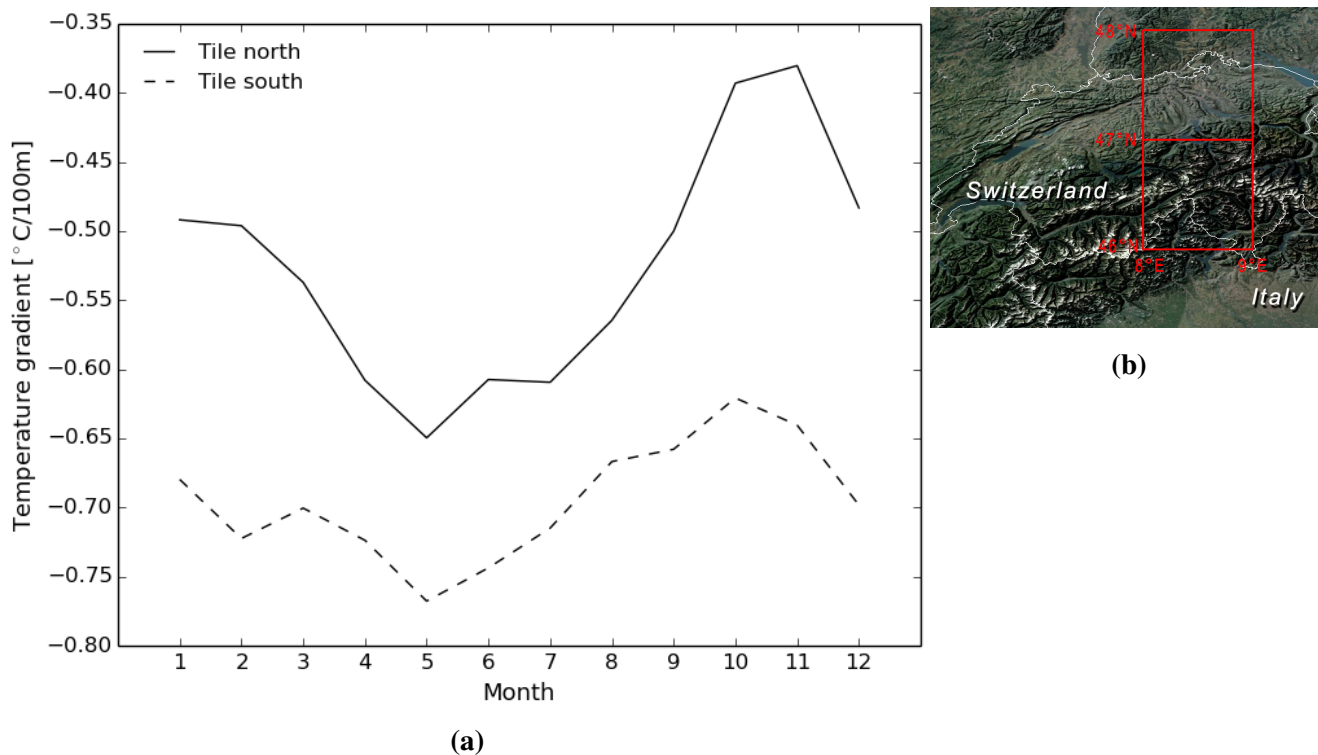
4.2. Seasonal Variation

The data set of gradients covers a large part of the earth on both hemispheres. It is therefore to be expected that any seasonal patterns will vary widely between different climates and it is difficult to show any general trend. However, it is possible to look at smaller areas to see if there are clear seasonal variations in the temperature gradient. Kunz *et al*. [23] calculated the temperature gradient in Switzerland over the seasons, so we have chosen two $1° \times 1°$ tiles covering parts of Switzerland. Both tiles range from $8°$ E to $9°$ E. The northern tile extends from $47°$ N to $48°$N, covering parts of the northern Swiss alps, while the southern tile goes from $46°$ N to $47°$ N and includes both high mountains up to 4000 m elevation and parts of the Italian Alpine lakes where the climate is influenced by the Po plain to the south.

Figure 14 shows the hourly and monthly gradient for the two tiles. For each month, the graphs show the 8 3-hourly time slots. There is a clear diurnal variation with the largest (negative) gradients during daytime. The figure also shows a seasonal variation for the northern tile while this is less clear for the southern area. Figure 15 shows the seasonal variation of the daily average of the gradients, obtained by calculating the average of the 3-hourly gradients for each month. In the northern tile, gradients are about twice as large in summer than in winter. For the southern tile the pattern is similar, but the amplitude is much smaller. The curve for the northern tile is very similar to that found in [23] (Figure 7 in that paper), where the gradient was calculated from ground station data at varying elevation.
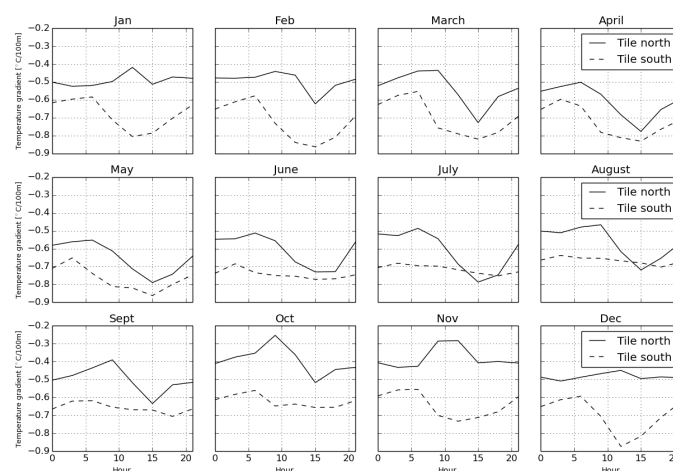
Figure 14. Diurnal and seasonal variation in the temperature gradients for two areas in the Alps, between $8°$ E and $9°$ E. The southern tile has latitude extension $46°$ N to $47°$ N and the northern $47°$ N to $48°$ N.

Figure 15. (**a**) Seasonal variation in the temperature gradients for two areas in the Alps. Here, the diurnal variation has been averaged to produce a single value for each month. (**b**) Areas with coordinates to which data on the left refers to (also the same areas as in Figure 14).

4.3. Applications to Solar Energy and Energy Meteorology

The main motivation for the present work has been to have accurate 2-m temperature estimates for studies in solar energy and energy consumption in buildings. We will therefore highlight a couple of applications in these fields.

4.3.1. Heating Degree Days

The number of *degree days* for a given location is a rough metric for the need for heating or cooling during a period. There is no universally agreed standard for how to calculate degree days, but in this study we have used the following expression for the number of heating degree days N_{hdd} in a year:

$$
\begin{aligned}
N_{\mathrm{hdd}} &= \sum_{d=1}^{365} H_d \\
H_d &= \begin{cases} H_d = (18 - T_d) & T_d < 15. \\ H_d = 0 & T_d >= 15. \end{cases}
\end{aligned}
\tag{6}
$$

Here, H_d is the number of heating degree days for a given day d, and T_d is the average temperature of that day. The sum is done over a period of one year.

We performed the calculation of heating degree days for one year (2010) in an area in the Alps covering $2° \times 2°$: from $45°30'$ N to $47°30'$ N and from $7°$ E to $9°$ E. N_{hdd} was calculated using the 3-hourly *forecast* temperatures averaged to daily values. The calculation was then repeated with the hourly temperatures corrected by the downscaling procedure of Equation (2). The elevation data used came from the SRTM-3 digital elevation model [7] and has a spatial resolution of 3 arc-seconds. Maps of N_{hdd} using the two methods are shown in Figure 16.

It should be noted that at such high resolution other effects will become important, such as the difference between north- and south-facing slopes where the temperature may differ considerably though the elevation is the same. The method presented here is not able to account for such effects.
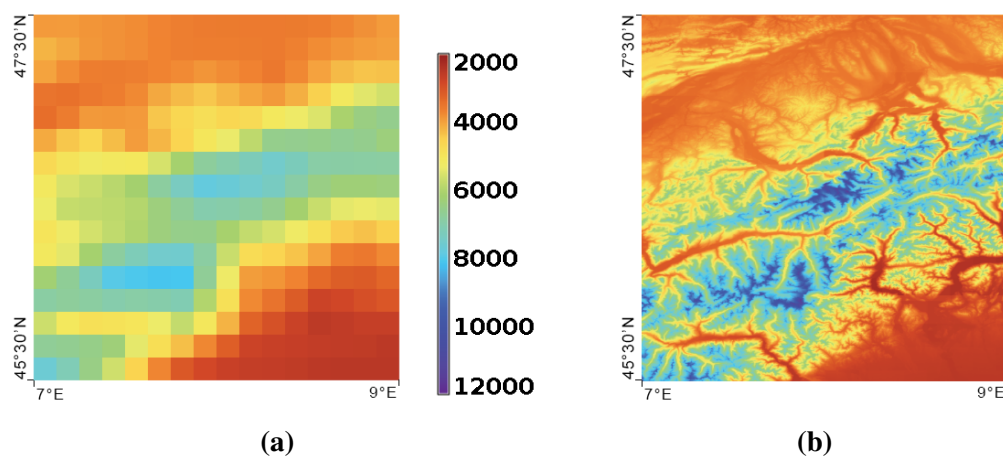
Figure 16. Number of degree days for 2010 for an area in the Alps ($45°30'$ N to $47°30'$ N and from $7°$ E to $9°$ E). Red/yellow colours indicate a low number of heating degree days, blue a high number. (**a**) Calculation using daily averages of the ECMWF forecast temperatures; (**b**) Calculation using downscaled temperatures with a 3 arc-second DEM.

4.3.2. Estimating Performance of Photovoltaic Modules

The energy conversion efficiency of photovoltaic (PV) modules depends on a number of external influences, chief among which are the intensity of the solar radiation and the temperature of the modules. The module temperature in turn depends (among other things) on the temperature of the surrounding air. To make an estimate of PV performance for a given location, it is therefore necessary to know the 2-m temperature at that location.

A measure of the influence of 2-m temperature on PV performance is given by the *Module Performance Ratio* (MPR) [3]. This is the ratio of the actual PV energy output to that which would have been obtained if the PV module efficiency were always constant at the nominal efficiency. MPR is defined as:

$$\mathrm{MPR} = \frac{1000 \cdot E_{\mathrm{PV}}}{P_{\mathrm{nom}} H} \tag{7}$$

Here E_{PV} is the energy produced by the module during a certain time period, H is the total irradiation impinging on the module during the same time, and P_{nom} is the nominal power of the module,

In order to investigate the effect of the downscaled temperatures on PV performance we performed a simulation of PV energy production for the same time period and the same area as that used in Figure 16. Solar radiation data were obtained from the CM SAF collaboration (www.cmsaf.eu) [30,31], while the model for PV efficiency is taken from [32] assuming crystalline silicon PV modules. As with the example of Section 4.3.1, the simulation was performed using two different resolutions: one with a resolution of $1'30''$ (the native resolution of the solar radiation data) with the ECMWF forecast temperatures; and one with a resolution of $3''$ with downscaled temperatures. The resulting maps of MPR are shown in Figure 17. As in the example with degree-days we see here that also the PV performance estimates are significantly different when using the downscaled temperature information. The differences in MPR may reach more than 5%, which is similar to the uncertainty in the estimate of the solar irradiation [33]. In mountain areas the downscaling procedure thus makes it possible to reduce the overall uncertainty in estimates of PV performance.
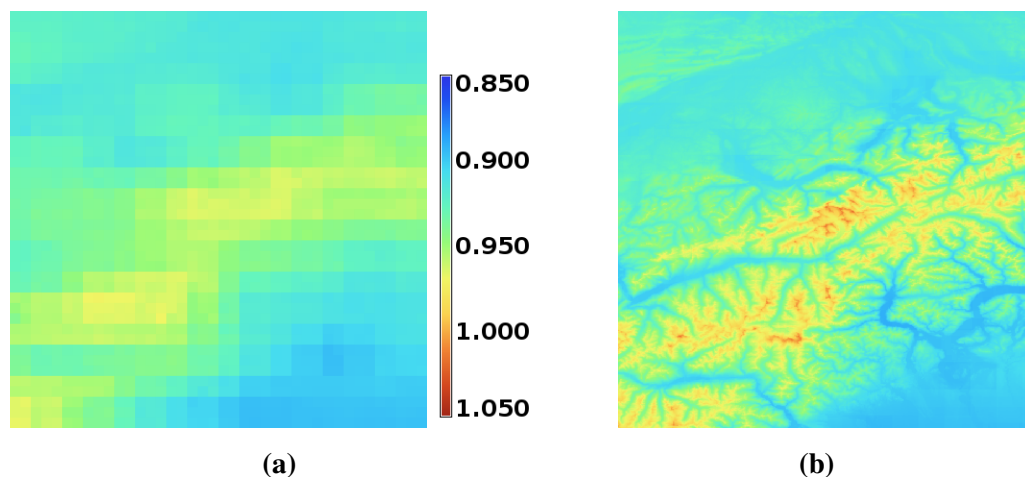
Figure 17. Module performance ratio for crystalline silicon PV modules, for an area in the Alps ($45°30'$ N to $47°30'$ N and from $7°$ E to $9°$ E). (**a**) Calculation using ECMWF forecast temperatures and a resolution of $1'30''$; (**b**) Calculation using downscaled temperatures with a 3 arc-second DEM.

5. Conclusions

We have constructed digital maps of the average vertical gradient in 2-m temperature, using data from the ECMWF operational weather forecast and applying a linear fit to the values in areas of size $1° \times 1°$. The gradients have been calculated for each month and for every three hours during the day using data from 2010 to 2013.

Validation of the procedure against measurement station data shows that using these gradients it is possible to significantly improve the estimates of 2-m temperature at arbitrary points using the forecast data together with these gradients and knowledge of the local elevation, relative to using just the forecast data. Comparing with 8087 stations worldwide, the overall yearly mean bias deviation <MBE> reduces from -0.49 to -0.10 °C, while gMAB Equation (13) goes from 0.99 to 0.64 °C and gRMSD Equation (14) goes from 1.61 °C to 0.89 °C. Considering only stations with an elevation difference

>300 m from the forecast elevation, the improvement is more dramatic: from –2.26 °C to –0.21 °C for <MBE>, from 3.44 to 1.02 °C for gMAB and from 4.11 to 1.42 °C for gRMSD.

We have also investigated applying the gradients to the publicly available ECMWF ERA-Interim reanalysis data set. If a suitable offset is applied to account for the difference between the forecast and the reanalysis, the validation results yield similar results as when the gradients are applied to the forecast data.

Two weaknesses of the present approach are apparent: the data used cover only four years, and these data are not publicly available. As reanalysis data with a sufficiently high spatial resolution become available (about the same 1/8° resolution as for the forecast data), this procedure should be applied also to these data. If the gradients can be calculated using reanalysis data there will no longer be a need for calculating the offset between forecast and reanalysis. It will also be possible to investigate whether it is shorter time periods can be used to calculate the gradients, maybe even using instantaneous values.

The present study has only considered the effect of elevation on 2-meter temperature. However, there are other effects that influence temperature locally. If this effect can be quantified with a high spatial resolution it may be used as a basis for downscaling the temperature data to account for this effect. One possibility is to use land cover or population density data to account for the urban heat island effect, to the extent that this effect is not already present in the forecast or reanalysis data. This is at present under investigation. Shadowing by mountains or hills may also influence the local temperature, and this effect could be studied given the availability of very-high-resolution DEM data.

The data set of gradients and the offset maps will be made freely available as GIS raster maps at the web page with supplementary material [28]. At the same web site we have also made the validation data available.

As a first application, the gradient data are being used by the PVGIS web application for producing estimates of PV performance (http://re.jrc.ec.europa.eu/pvgis/).

Author Contributions

Thomas Huld wrote the software to calculate the temperature gradients, part of the code for the validation, and produced Figures 3, 13, 16 and 17. Irene Pinedo Pascua wrote the software to collect and select the ground station data, as well as most of the code to perform the validation, and produced all the remaining figures. It is not possible to separate the contributions to the manuscript itself.

Nomenclature

Acronyms

CM SAF	Climate Monitoring Satellite Application Facility
DEM	Digital Elevation Model
ECMWF	European Centre for Medium-range Weather forecast
ERA	ECMWF reanalysis products, for instance ERA-Interim
NCEP	National Centers for Environmental Protection
NCDC	National Climatic Data Center
NOAA	National Oceanic and Atmospheric Administration
PV	Photo-Voltaic
PVGIS	PhotoVoltaic Geographical Information System
SRTM	Shuttle Radar Topography Mission

Symbols

$\langle \delta T_{m,h} \rangle$	Average temperature offset between *reanalysis* and *forecast* data (°C)
E_{PV}	Annual energy output of a PV module/array (kWh)
$\gamma_{m,h}$	Vertical temperature gradient for month m and hour h (°C/m)
H	Annual solar irradiation (kWh/m^2)
H_d	Number of degree days for a given day (-)
MAB	Mean Absolute Bias
gMAB	global Mean Absolute Bias
MBD	Mean Bias Deviation
gMBD	global Mean Bias Deviation
MPR	PV module performance ratio (-)
N	Number of data points used for the validation (-)
N_s	Number of ground stations used for the validation (-)
N_{hdd}	Number of degree days in a year (-)
P_{nom}	Nominal power of a PV module (W)
RMSD	Root Mean Square Deviation
gRMSD	global Root Mean Square Deviation
$T_{f:y,m,d,h}$	*Forecast* temperature for year y, month m, day d and hour h. (°C)
$T_{corr}(t)$	Temperature corrected for elevation. (°C)
$T_{corr,r}(t)$	*Reanalysis* temperature corrected for elevation. (°C)
$\langle T_{f:m,h} \rangle$	Average *forecast* temperature for month m and hour h. (°C)
Z_f, Z_r	Elevation of the *forecast* DEM and the *reanalysis* DEM, respectively (m)
z	Elevation at at given point (m)

A. Statistical Measures Used for the Validation

A number of different statistical measures have been used to estimate the uncertainty in the forecast and downscaled. These are defined briefly here.

For a given set of data (station or forecast/downscaled data), the annual average is calculated from the N values:

$$\langle G \rangle = \frac{1}{N} \sum_{i=1}^{N} (T_i) \tag{8}$$

The *Mean Bias Deviation* (MBD) is calculated as shown in Equation (9):

$$\text{MBD} = \frac{1}{N} \sum_{i=1}^{N} \left(T_i^f - T_i^s \right) \tag{9}$$

Here T_i^s is the measured temperature at the ith time point while T_i^f is the forecast (or downscaled) temperature at the same time point. N is the total number of time points for which both measured and estimated data are available.

The *Mean Absolute Bias* (MAB) is calculated as Equation (10):

$$\text{MAB} = \frac{1}{N} \sum_{i=1}^{N} \left| T_i^f - T_i^s \right| \tag{10}$$

The *Root Mean Square Deviation* (RMSD) is given as:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{N} \left(T_i^f - T_i^s \right)^2}{N}} \tag{11}$$

To get an overall estimate of the bias in yearly average temperature, we can calculate the average of the MBD values at all the station locations. This is termed the *global* MBD, gMBD:

$$\text{gMBD} = \frac{1}{N_s} \sum_{s=1}^{N_s} \text{MBD}_s \tag{12}$$

Here, MBD_s is the MBD of station number s. N_s is the number of stations.

A measure of the uncertainty of the yearly average temperature in a given point can be given as the *global* MAB and RMSD:

$$\text{gMAB} = \frac{1}{N_s} \sum_{s=1}^{N_s} |\text{MBD}_s| \tag{13}$$

$$\text{gRMSD} = \sqrt{\frac{\sum_{s=1}^{N_s} (\text{MBD}_s)^2}{N_s}} \tag{14}$$

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Clarke, J.A. *Energy Simulation in Building Design*, 2nd ed.; Routledge: London, UK, 2001.
2. Koehl, M.; Heck, M.; Wiesmeier, S.; Wirth, J. Modeling of the nominal operating cell temperature based on outdoor weathering. *Sol. Energy Mater. Sol. Cells* **2011**, *95*, 1638–1646.
3. Huld, T.; Gottschalg, R.; Beyer, H.; Topič, M. Mapping the performance of PV modules, effects of module type and data averaging. *Sol. Energy* **2010**, *84*, 324–338.
4. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.A.; Balsamo, G.; Bauer, P.; *et al.* The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 553–597.
5. Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; *et al.* The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 437–471.
6. Rienecker, M.M.; Suarez, M.J.; Gelaro, R.; Todling, R.; Bacmeister, J.; Liu, E.; Bosilovich, M.G.; Schubert, S.D.; Takacs, L.; Kim, G.K.; *et al.* MERRA: NASA's modern-era retrospective analysis for research and applications. *J. Clim.* **2011**, *24*, 3624–3648.
7. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; *et al.* The shuttle radar topography mission. *Rev. Geophys.* **2007**, *45*, doi:10.1029/2005RG000183.
8. Legates, D.; Willmott, C. Mean seasonal and spatial variability in global surface air temperature. *Theor. Appl. Climatol.* **1990**, *41*, 11–21.
9. Leemans, R.; Cramer, W.P. *The IIASA Database for Mean Monthly Values of Temperature, Precipitation and Cloudiness of a Global Terrestrial Grid*; Technical Report; International Institute for Applied Systems Analysis (IIASA): Laxenburg, Austria, 1991.
10. New, M.; Hulme, M.; Jones, P. Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *J. Clim.* **1999**, *12*, 829–856.
11. New, M.; Lister, D.; Hulme, M.; Makin, I. A high-resolution data set of surface climate over global land areas. *Clim. Res.* **2002**, *21*, 1–25.
12. Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **2005**, *25*, 1965–1978.
13. Huld, T.A.; Šúri, M.; Dunlop, E.D.; Micale, F. Estimating average daytime and daily temperature profiles within Europe. *Environ. Model. Softw.* **2006**, *21*, 1650–1661.
14. Piper, S.C.; Stewart, E.F. A gridded global data set of daily temperature and precipitation for terrestrial biospheric modeling. *Glob. Biogeochem. Cycles* **1996**, *10*, 757–782.
15. Haylock, M.R.; Hofstra, N.; Klein Tank, A.M.G.; Klok, E.J.; Jones, P.D.; New, M. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.* **2008**, *113*, doi:10.1029/2008JD010201.
16. Wang, A.; Zeng, X. Development of global hourly 0.5° land surface air temperature datasets. *J. Clim.* **2013**, *26*, 7676–7691.

17. Harris, I.; Jones, P.; Osborn, T.; Lister, D. Updated high-resolution grids of monthly climatic observations—The CRU TS3.10 dataset. *Int. J. Climatol.* **2014**, *34*, 623–642.

18. Zakšek, K.; Schroedter-Homscheidt, M. Parameterization of air temperature in high temporal and spatial resolution from a combination of the SEVIRI and MODIS instruments. *ISPRS J. Photogram. Remote Sens.* **2009**, *64*, 414–421.

19. Metz, M.; Rocchini, D.; Neteler, M. Surface temperatures at the continental scale: Tracking changes with remote sensing at unprecedented detail. *Remote Sens.* **2014**, *6*, 3822–3840.

20. Rolland, C. Spatial and seasonal variations of air temperature lapse rates in alpine regions. *J. Clim.* **2003**, *16*, 1032–1046.

21. Kattel, D.; Yao, T.; Yang, K.; Tian, L.; Yang, G.; Joswiak, D. Temperature lapse rate in complex mountain terrain on the southern slope of the central Himalayas. *Theor. Appl. Climatol.* **2013**, *113*, 671–682.

22. Van Diepen, K.; Boogaard, H.L.; Supit, I.; Lazar, C.; Orlandi, S.; van der Goot, E.; Schapendonk, A.H.C.M. Methodology of the MARS crop yield forecasting system. In *Agrometeorological Data Collection, Processing and Analysis*, 1st ed.; European Commission: Luxembourg, 2004; Volume 2, p. 105.

23. Kunz, H.; Scherrer, S.C.; Liniger, M.A.; Appenzeller, C. The evolution of ERA-40 surface temperatures and total ozone compared to observed Swiss time series. *Meteorol. Z.* **2007**, *16*, 171–181.

24. European Centre for Medium-Range Weather Forecasts. ECMWF Operational Analysis data, NCAS British Atmospheric Data Centre, 2006. Available online: http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_ECMWF-OP (accessed on 15 February 2014).

25. Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open source scientific tools for Python. Available online: http://www.scipy.org/ (accessed on 26 July 2014).

26. Smith, A.; Lott, N.; Vose, R. The integrated surface database: Recent developments and partnerships. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 704–708.

27. Pandas. Python Data Analysis Library. Available online: http://pandas.pydata.org/ (accessed on 15 July 2014).

28. Huld, T.A.; Pinedo Pascua, I. Supplementary material for the present paper. Available online: http://re.jrc.ec.europa.eu/pvgis/papers_supplementary/downscaling/index.html (accessed on 20 March 2015).

29. World Meteorological Organization. List of SYNOP Stations. Available online: ftp://ftp.wmo.int/wmo-ddbs/VolA_New/Pub9volA141117x.flatfile (accessed on 11 May 2014).

30. Müller, R.; Matsoukas, C.; Gratzki, A.; Behr, H.; Hollmann, R. The CM-SAF operational scheme for the satellite based retrieval of solar surface irradiance-A LUT based eigenvector hybrid approach. *Remote Sens. Environ.* **2009**, *113*, 1012–1024.

31. Huld, T.; Müller, R.A.G. A new database of global and direct solar radiation using the eastern meteosat satellite, models and validation. *Sol. Energy* **2012**, *86*, 1803–1815.

32. Huld, T.A.; Friesen, G.; Skoczek, A.; Kenny, R.A.; Sample, T.; Field, M.; Dunlop, E.D. A power-rating model for crystalline silicon PV modules. *Sol. Energy Mater. Sol. Cells* **2011**, *95*, 3359–3369.

33. Gracia Amillo, A.; Huld, T.; Müller, R. A new database of global and direct solar radiation using the eastern meteosat satellite, models and validation. *Remote Sens.* **2014**, *6*, 8165–8189.