ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Guidance document for users of the climate indices

Issued by: UEA / Phil Jones

Date: 01/10/2018

Ref: C3S_D311a_Lot4.3.1.2_201809_user_guidance_indices_v1

Official reference number service contract: 2017/C3S_311a_Lot4_KNMI/SC1

This document has been produced in the context of the Copernicus Climate Change Service (C3S).
The activities leading to these results have been contracted by the European Centre for Medium-Range Weather Forecasts, operator of C3S on behalf of the European Union (Delegation Agreement signed on 11/11/2014). All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission and the European Centre for Medium-Range Weather Forecasts has no liability in respect of this document, which is merely representing the authors view.

Contributors

UEA
Phil Jones
Colin Harpham

KNMI
Else van den Besselaar
Gerard van der Schrier

Table of Contents

Introduction

An R-based script was developed to calculate the indices defined by the Expert Team on Climate Change Detection and Indices (ETCCDI) based on daily precipitation and temperature gridded fields from the ensemble version of the E-OBS dataset. We have taken the ETCCDI set of indices and added some others. A list of the indices together with the available time scales which can currently be calculated is given in Table 1. The current software includes more indices, but only those using daily maximum and minimum temperatures and precipitation totals can currently be calculated. The history of ETCCDI and the development of the 27 indices is described in Zhang et al. (2011). The indices were designed to determine whether changes have occurred in extremes around the world. They have been widely used (e.g. references within Zhang et al., 2011 and also Donat et al., 2013) and have been referred to by the IPCC when discussing both historic long-term changes in extremes, and also in climate change projections for the future. A key aspect of the software is that it relies on both an R-based software package and Fortran code, both of which are freely available and produce the same answers given the same data.

Here, the novel aspect is the use of the software with the ensemble version of E-OBS (described in Cornes et al. 2018), requiring both their efficient calculation and an understanding of the uncertainties that are incorporated into this ensemble version. The purpose of this document is to provide guidance for users on how to use the indices and their associated uncertainties. The later examples, here, use E-OBS v16.0e, but the indices will be updated with each new version. The first piece of user guidance is, therefore, that it is essential to state which version of the indices is being used (i.e. 16.0e, 17.0e etc.). Also, exactly the same indices software will be used with the regional gridded datasets across Europe (e.g. CARPATCLIM and NGCD) although none of these datasets currently involves ensembles.

The next sections discuss: (1) changes to the method of production and (2) guidance given through a discussion of graphical results from a selection of the indices at a selection of locations across Europe.

1. Changes to the Method of Production

The overarching aim of the ETCCDI software has been to ensure conformity of the calculation of the indices, and also to use indices that can be used for specified thresholds (e.g. 0˚C for frost days or 25˚C for summer days). In order for the indices to be comparable across the whole European region, the most widespread usages have come from the percentile-based thresholds for temperature. Here the thresholds to provide the number of days above the 90th percentile or below the 10th percentile are determined by the data themselves based on a common or reference period (set here to 1961-1990 but this could be altered in the future). The other choices for a reference period are 1971-2000 and 1981-2010, given the length of the E-OBS dataset (1951-2018), but 1961-1990 is used as this is the period when the grids are most reliable as station numbers used are greatest.

In deliverable C3S_D311a_Lot4.3.1.1 we indicated three possibilities for calculating the temperature threshold percentiles:

1. Calculate a different set of thresholds for each of the ensembles, but with this, differences would be responding to both differences between the members when calculating the indices, but also in the calculation of the thresholds.
2. Related to the first, a simpler approach is one that users of the dataset might adopt. This would base the thresholds on the best guess values (e.g. ensemble mean) and then calculate the indices based on this single grid of best guess values.
3. The final and better choice is to calculate the indices based on all the data from all the ensemble members. So instead of using daily data for the 30 years in the base period, the threshold calculation would be based on 30*[the number of available members] years. This one set of percentile-based thresholds would then be used to calculate 100 sets of indices, 1 for each member. As the thresholds are the same, the spread would then provide a measure of uncertainty.

Option 3 was our ideal choice, but it proved impossible to use all 100 of the ensembles members with the software, because of the numerous checks made on the dates and not allowing the same dates to be used more than once. Instead we developed an approach that mixes options 2 and 3. The thresholds were determined by the best guess values (e.g. ensemble mean at each point and time step). Then sets of indices were calculated from each of the 100 ensemble members.

Only six of the ETCCDI indices are percentile-based. For the other ETCCDI and additional (non-ETCCDI) indices, the approach used is direct calculations using each ensemble member producing 100 sets of the indices.

The major two issues with the work here were (1) ensuring that the R software works efficiently across the grid taking into account the parallel processing environment of the computer hardware, and (2) that the bootstrapping adjustments for the temperature percentile thresholds are working properly. The reason for these bootstrapping adjustments is discussed in Zhang et al. (2005). With (1) an additional R routine (gridclimind) has been developed which calculates the indices from gridded data. This also includes a parameter which tells the software how to break up the grid into parts to efficiently use the number of processors available with the hardware. With (2) we didn't consider the bootstrapping adjustments of the temperature percentile thresholds to be working properly, so we used the thresholds based directly on the 1961-1990 period. The reason for Zhang et al. (2005) to use bootstrapping adjustments is that during the base period for a specific threshold (e.g. 10/90%) there ought to be exactly 10% of days below or above this threshold. Calculating the percentage of days for the four main percentile indices (Tx90p, Tn90p, Tx10p and Tn10p), the thresholds that gave the closest answers to the expected 10% were those using the direct calculation from the 1961-1990 period with no bootstrapping.

2. Guidance through discussion of examples

For E-OBS we have chosen the 2.5 and 97.5 percentile ranges to calculate the ensemble spread of each index. Our illustrations show the ensemble mean (50th percentile) and 2.5 and 97.5 percentile series. This choice means that ranges in our examples encompass 95% of the ensemble spread. Other less stringent choices would give smaller ranges. It was noted in deliverable C3S_D311a_Lot4.2.1.5 that some indices produce wider ranges for most locations, while some other indices produce smaller ranges for most locations. Thus we expect the ranges to be index dependent. We will illustrate this discussion using six temperature-based indices at a number of locations across Europe. Additionally, we will look at three precipitation-based indices for five different locations.

2.1 Temperature Indices

Here, the simplest indices are the absolute extremes. We choose annual TXx (annual maximum of the daily maximum temperature) and annual TNn (annual minimum of the daily minimum temperature), so the warmest day and coldest night of each year. Then we chose annual TX90p (the percentage of days in a year above the 90th percentile of maximum temperature for the time of year) and annual TN10p (percentage of days in a year below the 10th percentile for minimum temperature for the time of year). Finally, we look at spells of warm and cold days (WSDI and CSDI). They are defined as the period in days with maximum/minimum temperature more than 5 days above/below the $90^{th}/10^{th}$ percentile thresholds.

For both the TX- and TN-based time series, uncertainty ranges are largest for WSDI and CSDI, and smallest for the percentile series. For WSDI and CSDI the lowest number (other than zero) is 5, as the spells need to be this long to be classed as a spell. For many of the WSDI and CSDI series, average values of zero occur, but for most years the uncertainty ranges are from zero to a small positive number between 5 and 10. Occasionally, the uncertainty range is zero. When the upper uncertainty range is a positive number, but the average is zero, this means that only a few of the 100 ensemble members have positive values in a given year.

Figures 1 to 5 show examples for five locations: Linkoping (southern Sweden), Pontevedra (northwestern Spain), Brno (Czech Republic), Perugia (central Italy) and Kiev (Ukraine). For each location, each figure shows the annual values of TXx, TX90p and WSDI. Figures 6 to 10 show plots for the same five locations but for the indices based on minimum temperatures.

Trends are visible in many the graphs for TX90p. The average value for the 1961-1990 period is 10%, with all five locations showing values of 15 to 25% for years since 2000. Trends for WSDI are also positive, but series of WSDI are not well distributed with some years before 1990 recording zero. Trends based on minimum temperatures (Figures 6 to 10) indicate negative trends (i.e. locations have fewer cold extremes), but these are less clear than for maximum temperatures. With the expected value for TN10p being 10% during the 1961-1990 period, this index can only drop to zero (a drop of 10% of days), while those for TX90p can rise from 10% to a much higher percentage. CSDI for these five locations hardly show any trends, with the series being slightly more poorly distributed than WSDI.

The TXx and TNn series mostly indicate no or very weak trends. This is because they are each based on just a single value each year. Error ranges for these TXx and TNn series are related to the error ranges of individual days (i.e. the ranges given in Table 1 of Cornes et al., 2018). This paper gives mean absolute errors for TX of 0.92 and for TN of 1.22. In the examples shown here the error ranges for TNn are also slightly larger than for TXx. Finally, in the examples shown here for annual absolute temperature extremes, virtually the same plots would be produced for TXx if we restricted the season to summer, and for TNn if the season were restricted to winter.

In this document, we have only briefly alluded to the trends or lack of trends in some of the series. To calculate trends, the series of the 50th percentile should be used. Estimating the uncertainty of this trend should, in an ideal world, use all the ensemble members (see discussion of this in Morice et al. 2012 for HadCRUT4). As only 2.5, 50 and 97.5 percentile values will be available, a simpler approach is suggested that should be used for many of the indices. When the differences between the 2.5 and 50th percentiles and the 50 and 97.5th percentiles are approximately the same it will be possible to subsample the difference between the 2.5 and 97.5th percentiles randomly each year and estimate a range of trends and to calculate a two standard deviation range. However, for indices where the spread of the 2.5 and 97.5th percentiles is not similar (e.g. for the CSDI and WSDI series shown here), this process is not recommended.

2.2 Precipitation indices

Here the indices investigated are annual RX5day (the maximum 5-day precipitation for each year), annual R95pTOT (the total amount of rain per year from days above the 95th percentile, i.e. the total annual rain amount on very wet days) and annual CDD (consecutive number of dry days during the year). Here a dry day is defined as a day with less than 1mm of rain.

Figures 11 to 15 show these three indices for five locations: Norwich (eastern England), Murcia (southeastern Spain), Munich (southern Germany), Grosseto (central Italy) and Alexandroupoli (northeastern Greece). For each location, the figure shows three plots: annual values of RX5day (in mm), R95pTOT (in mm) and CDD (in days). Error ranges tend to be larger for R95pTOT than for RX5day for the same locations. At dry locations (e.g. Murcia and Alexandroupoli), average values for R95pTOT are zero in a few years with occasionally a zero range. This implies that there were not enough wet days in some years to calculate the 95th percentile threshold for rainfall. With this occurring at the annual timescale, this will almost certainly happen for most months and most seasons at these dry locations with strongly seasonal precipitation cycles. CDD values are much higher in regions with long dry seasons (e.g. in the Mediterranean). Average values (the 50th percentile) for CDD occasionally sit at the lower or upper limit of the error range. This implies that some ensemble members have a day in excess of 1mm that breaks the dry spell, while other members have this day with zero or less than 1mm of rainfall. If CDD were to be calculated for individual months or seasons it is likely that values would not add up to the annual values, as dry spells could extend across months and seasons. Plots of CDD for the dry season in the Mediterranean region will have little meaning.

References

Cornes, R.C., G.van der Schrier,E.J.M. van den Besselaar and P.D. Jones., 2018: Ensemble version of the E-OBS temperature and precipitation dataset. J. Geophys. Res. 123, 9391-9409. doi:10.1029/2017JD028200.

*Donat, M. G., L. V. Alexander, H. Yang, I. Durre, R. Vose, R. Dunn, K. Willett, E. Aguilar, M. Brunet, J. Caesar, B. Hewitson, C. Jacks, A. M. G. Klein Tank, A. Kruger, J. Marengo, T. C. Peterson, M. Renom, C. Oria Rojas, M. Rusticucci, J. Salinger, S. Sekele, A. Srivastava, B. Trewin, C. Villarroel, L. Vincent, P. Zhai, X. Zhang, and S. Kitching, 2013: Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. Journal of Geophysical Research - Atmospheres, **118**, 2098-2118. http://dx.doi.org/10.1002/jgrd.50150 (open access)*

Morice, C.P., Kennedy, J.J., Rayner, N.A. and Jones, P.D., 2012: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 dataset. Journal of Geophysical Research, 117, D08101, doi:10.1029/2011JD017187.

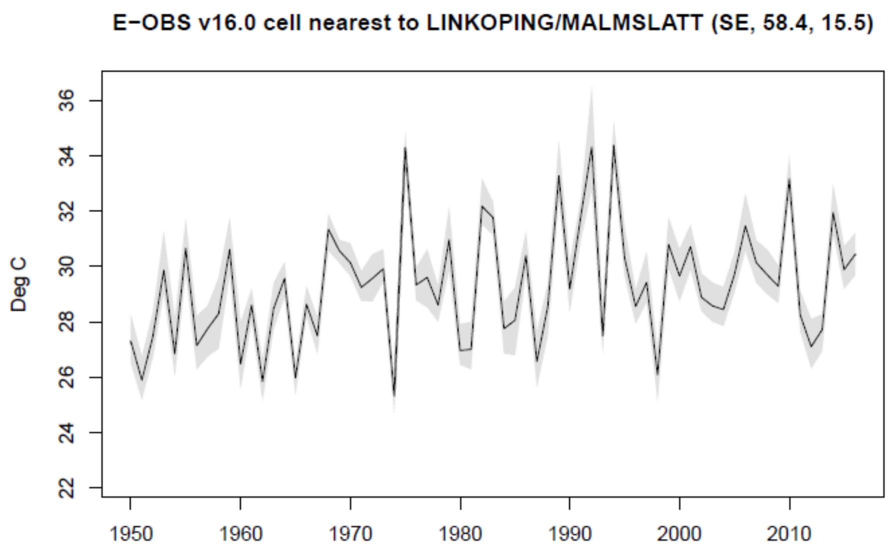
*Zhang, X., Hegerl, G.C. and Zwiers, F.W., 2005: Avoiding inhomogeneity in percentile-based indices of temperature extremes. J Climate **118**, 1641–1651.*

*Zhang, X., Alexander, L., Hegerl, G.C., Jones, P.D., Klein Tank, A., Peterson, T.C., Trewin, B. and Zwiers, F.W., 2011: Indices for monitoring changes in extremes based on daily temperature and precipitation data. WIREs Climate Change, **2**, 851-870, doi:10.1002/wcc/147.*

Table 1: Indices that are available, and the timescales for which they have been calculated. 27 from ETCCDI are first, then additional temperature, then additional precipitation ones as well.

Index	Monthly	Seasonal	Half Yearly	Annual
TXx	Y	Y	Y	Y
TNx	Y	Y	Y	Y
TXn	Y	Y	Y	Y
TNn	Y	Y	Y	Y
TN10p	Y	Y	Y	Y
TX10p	Y	Y	Y	Y
TN90p	Y	Y	Y	Y
TX90p	Y	Y	Y	Y
DTR	Y	Y	Y	Y
GSL	N	N	N	Y
FD0	Y	Y	Y	Y
SU25	Y	Y	Y	Y
TR20	Y	Y	Y	Y
WSDI	N	N	N	Y
CSDI	N	N	N	Y
RX1day	Y	Y	Y	Y
RX5day	Y	Y	Y	Y
SDII	Y	Y	Y	Y
R10	Y	Y	Y	Y
R20	Y	Y	Y	Y
CDD	N	N	N	Y

CWD	N	N	N	Y
R95pTOT	Y	Y	Y	Y
R99pTOT	Y	Y	Y	Y
PRCPTOT	Y	Y	Y	Y
ID	Y	Y	Y	Y
R1mm	Y	Y	Y	Y
AltCSDI	N	N	N	Y
AltWSDI	N	N	N	Y
CFD	Y	Y	Y	Y
CSU	Y	Y	Y	Y
HD17	Y	Y	Y	Y
HI	N	N	N	Y
Alt CDD	N	N	N	Y
Alt CWD	N	N	N	Y
R75p	Y	Y	Y	Y
R75pFRAC	Y	Y	Y	Y
R95pFRAC	Y	Y	Y	Y
R99pFRAC	Y	Y	Y	Y
SPI3	Y	N	N	N
SPI6	Y	N	N	N



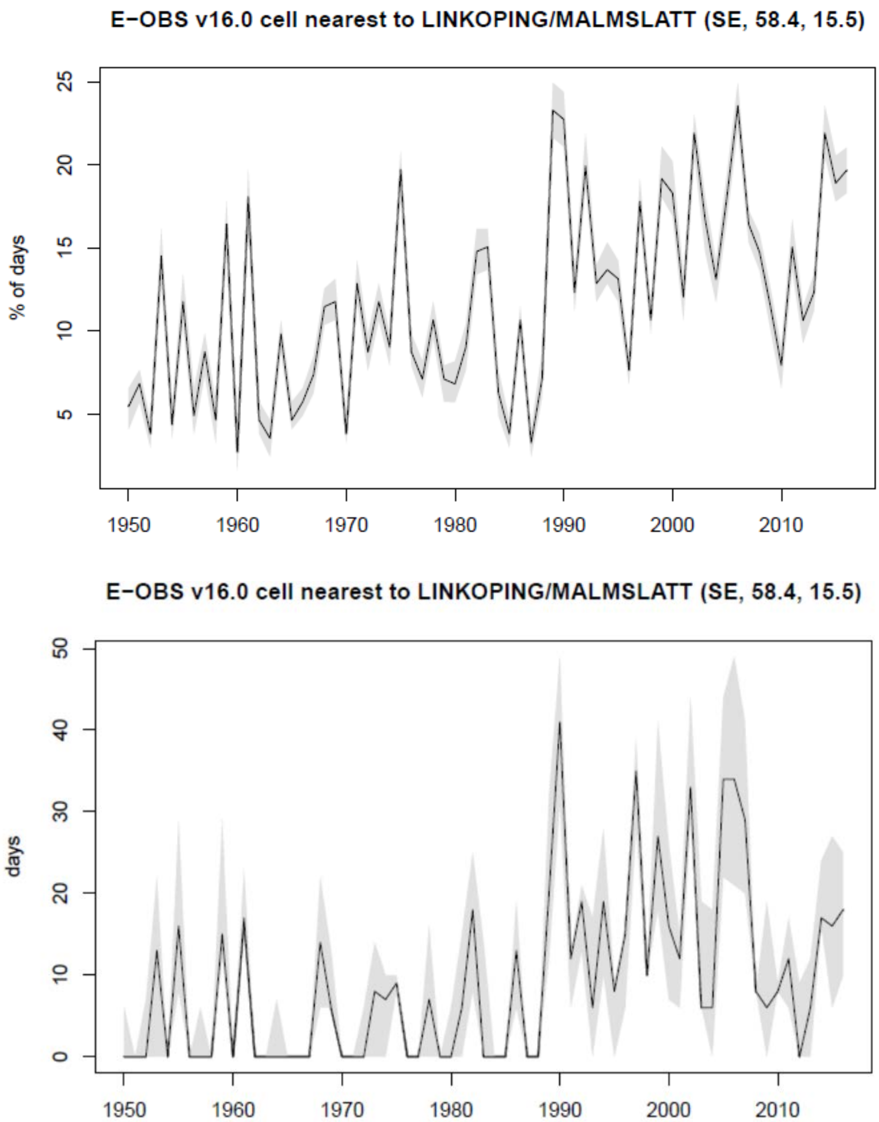


Figure 1: Annual values of TXx (˚C), TX90p (% of days) and WSDI (days) for Linkoping.

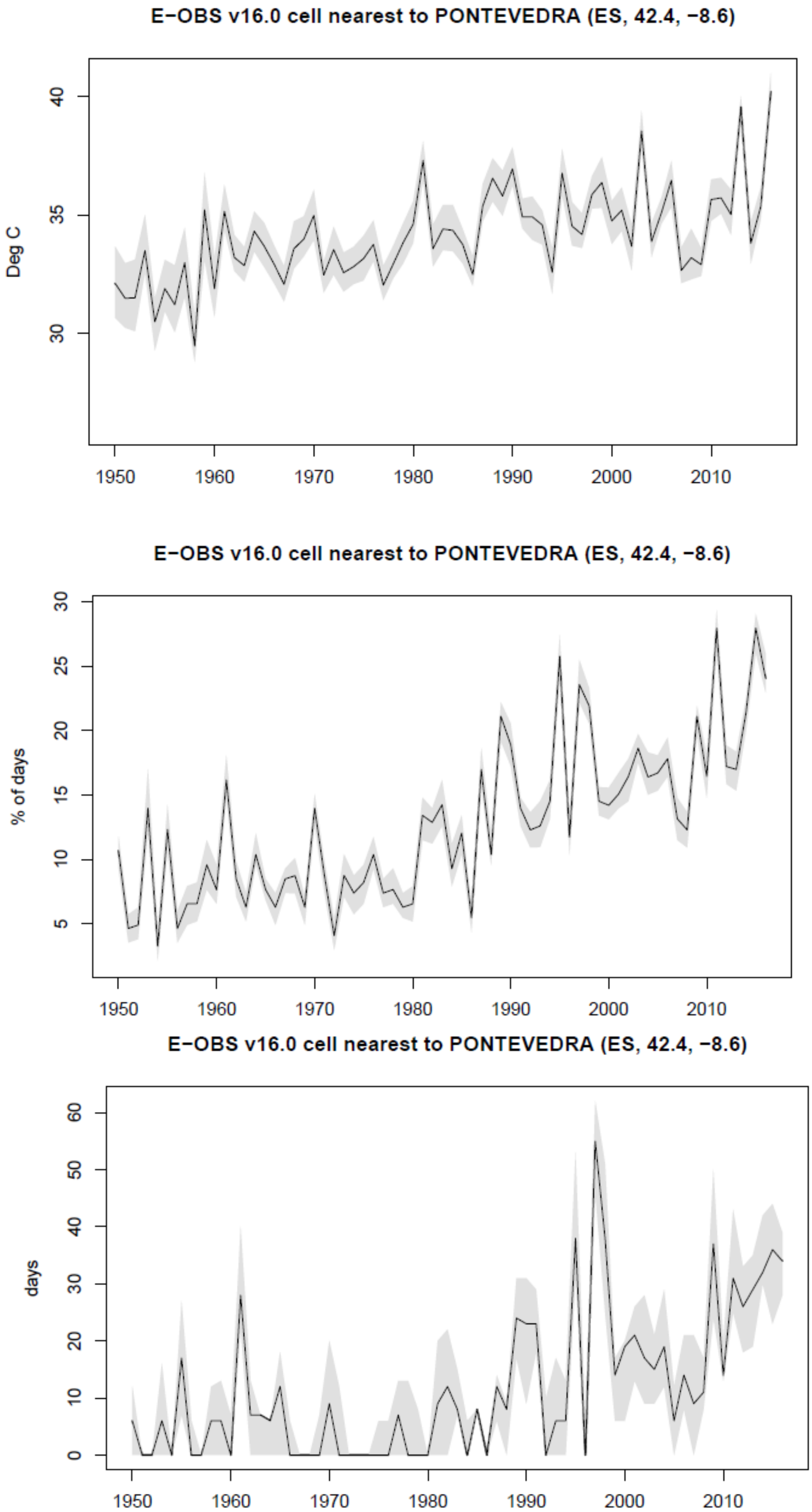


Figure 2: Annual values of TXx (˚C), TX90p (% of days) and WSDI (days) for Pontevedra.

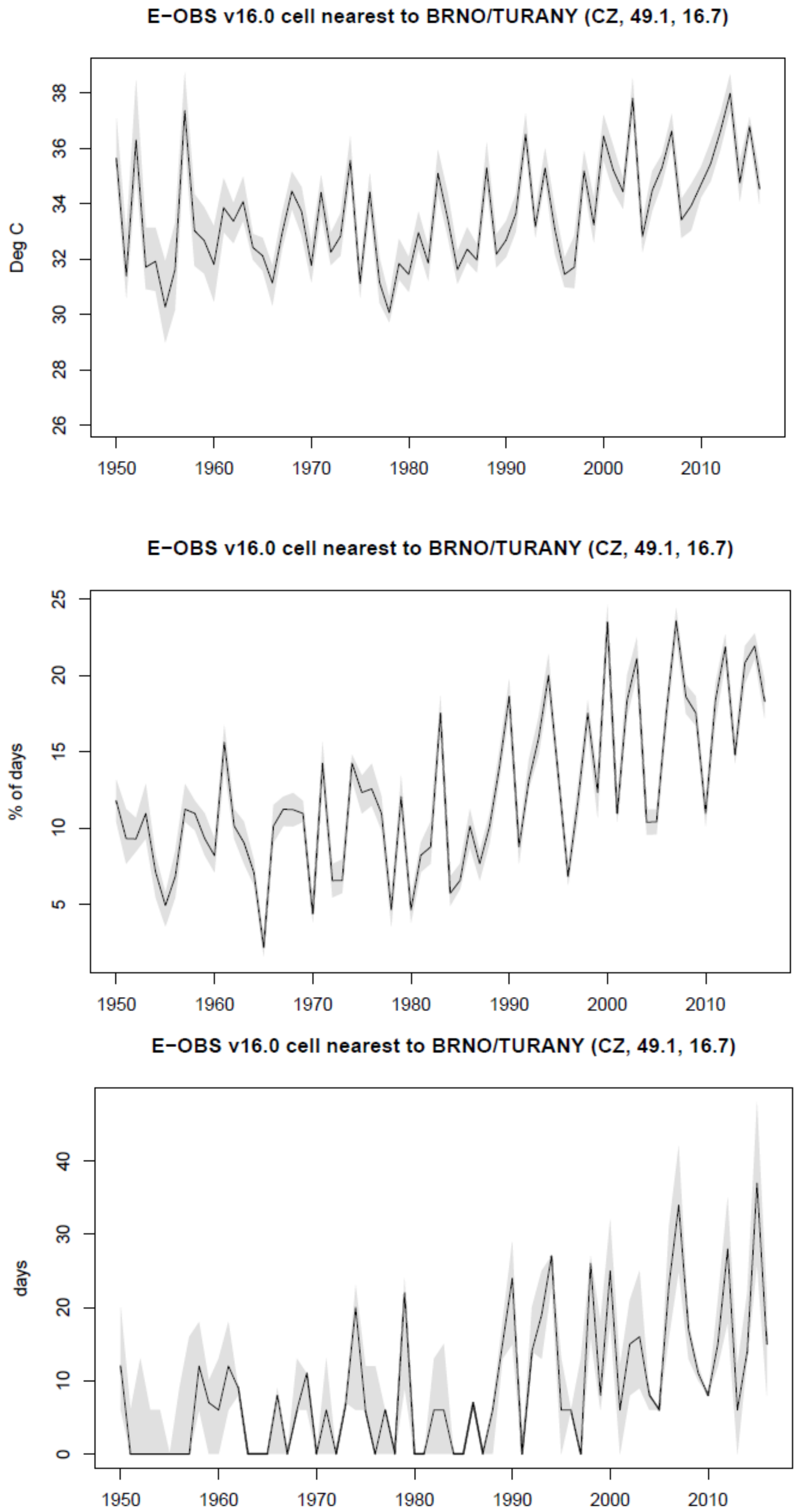


Figure 3: Annual values of TXx (˚C), TX90p (% of days) and WSDI (days) for Brno.

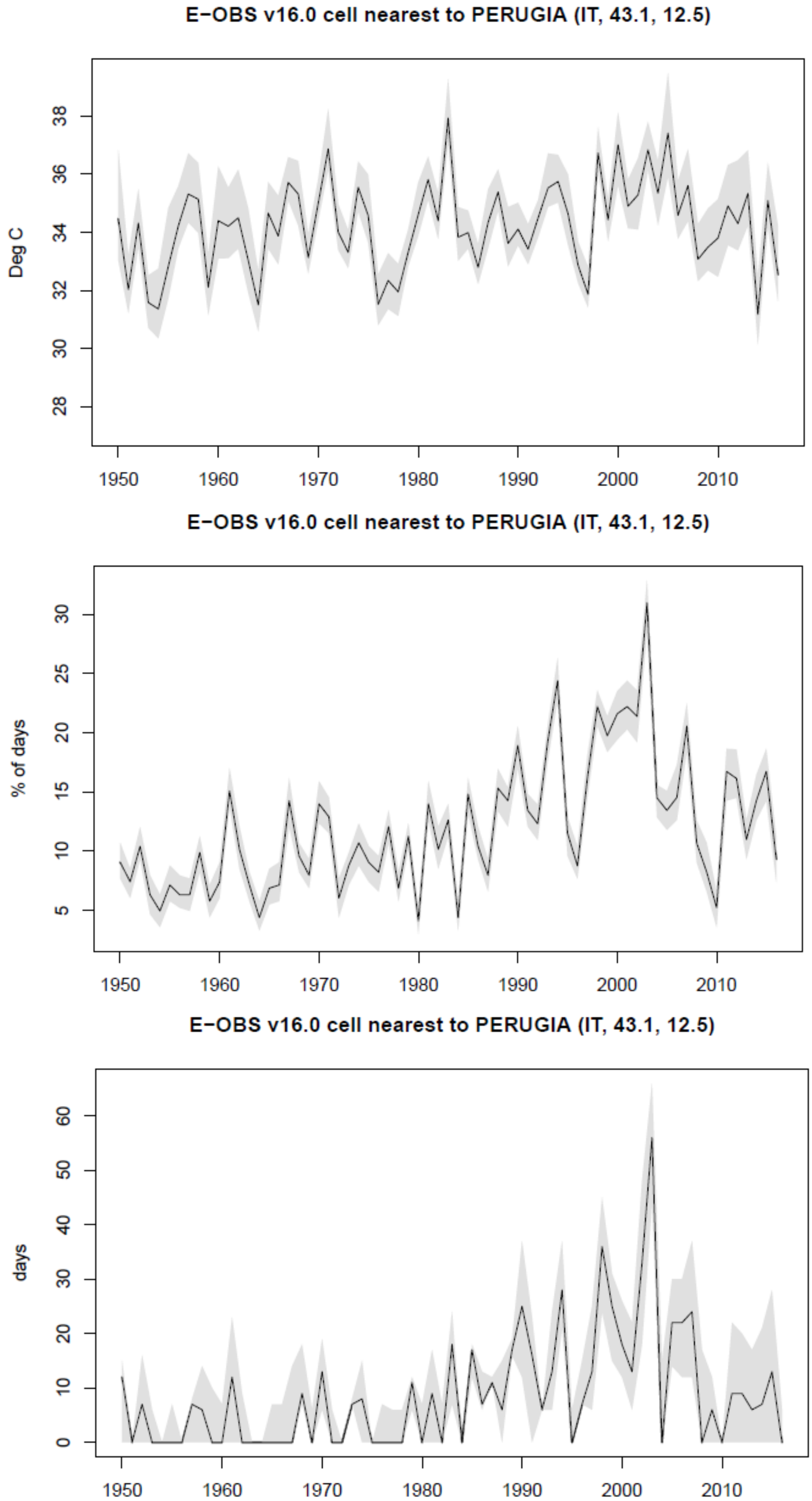


Figure 4: Annual values of TXx (˚C), TX90p (% of days) and WSDI (days) for Perugia.

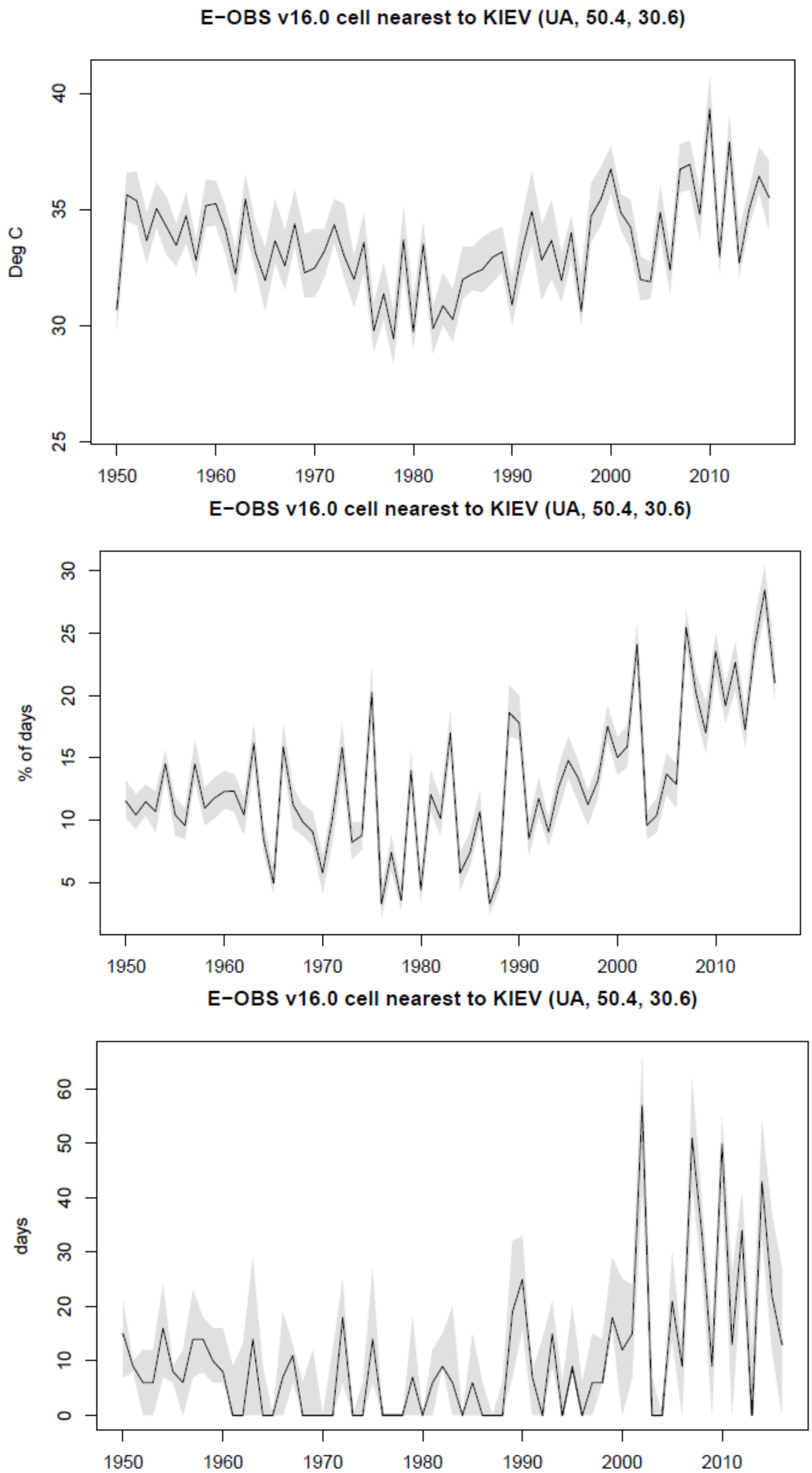


Figure 5: Annual values of TXx (˚C), TX90p (% of days) and WSDI (days) for Kiev.

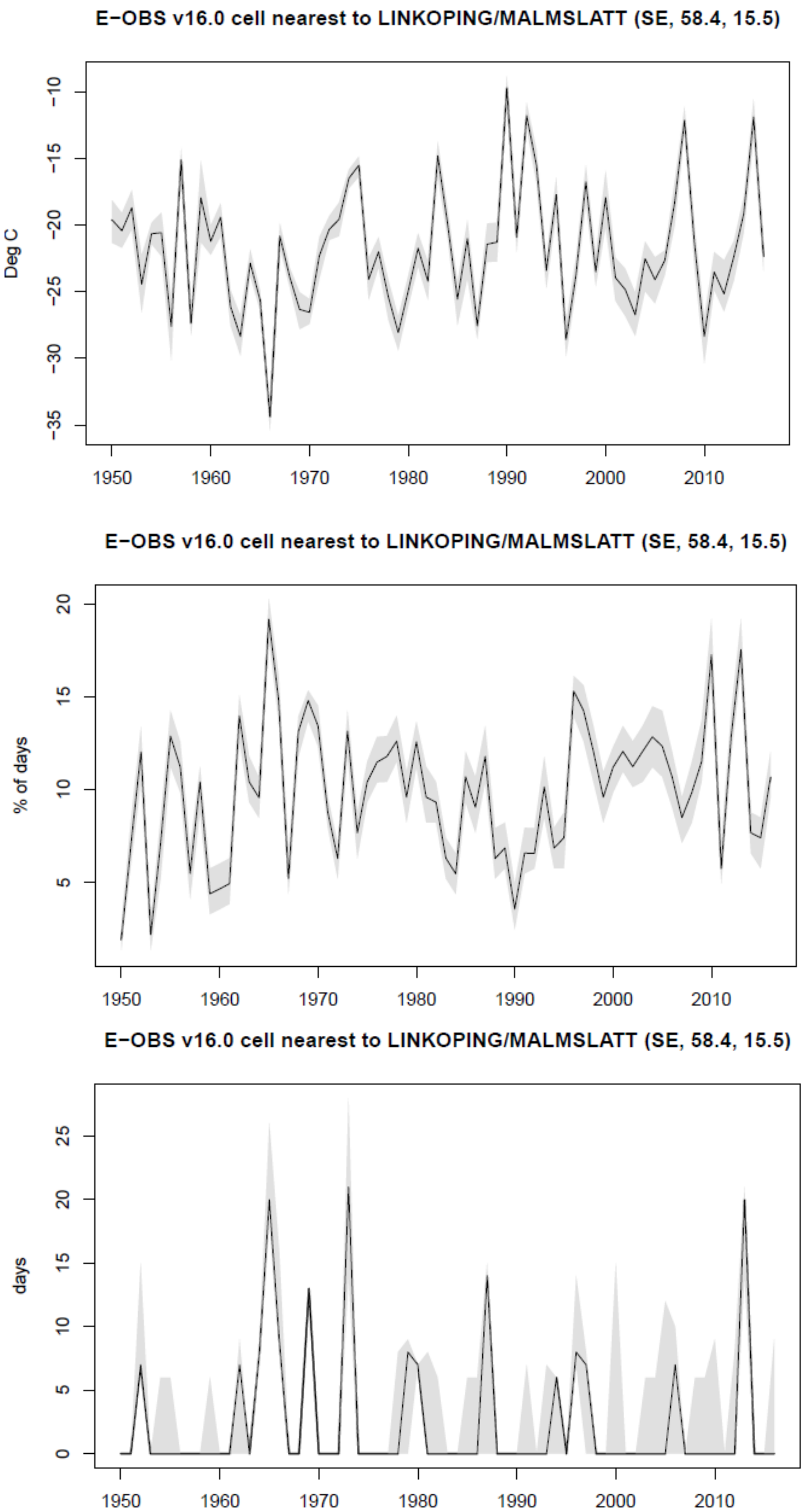


Figure 6: Annual values of TNn (˚C), TN10p (% of days) and CSDI (days) for Linkoping.

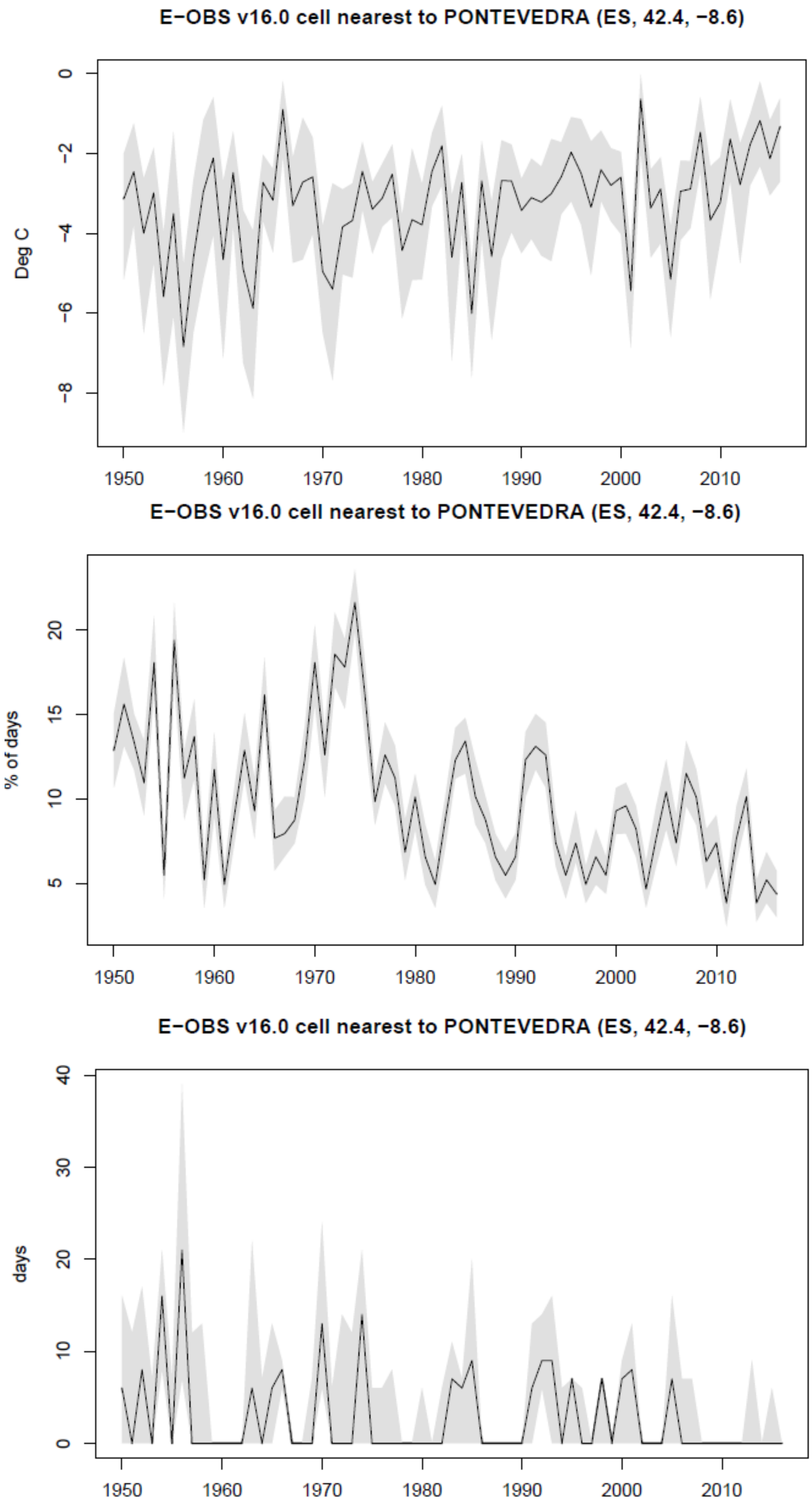


Figure 7: Annual values of TNn (˚C), TN10p (% of days) and CSDI (days) for Pontevedra.

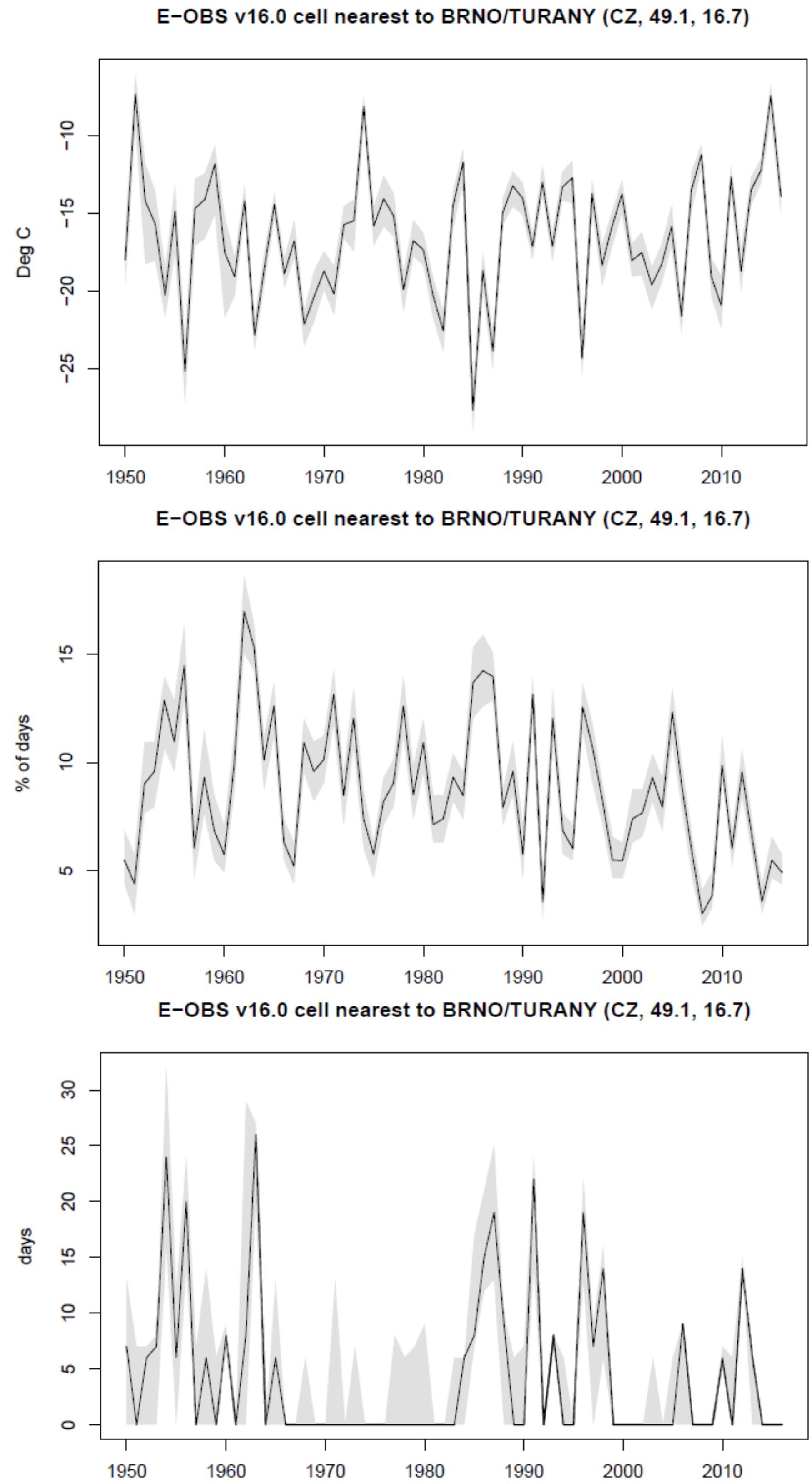


Figure 8: Annual values of TNn (˚C), TN10p (% of days) and CSDI (days) for Brno.

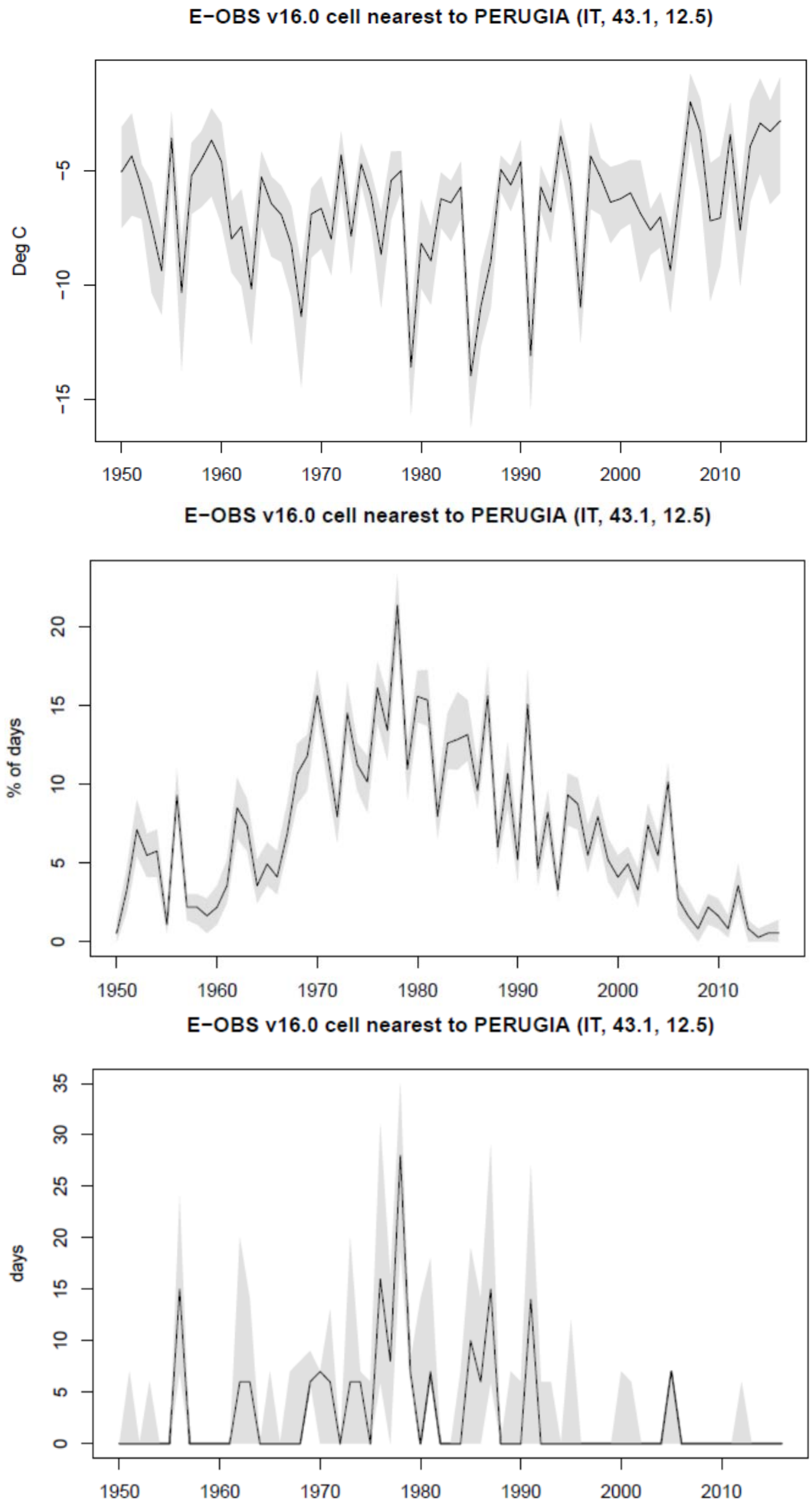


Figure 9: Annual values of TNn (˚C), TN10p (% of days) and CSDI (days) for Perugia.

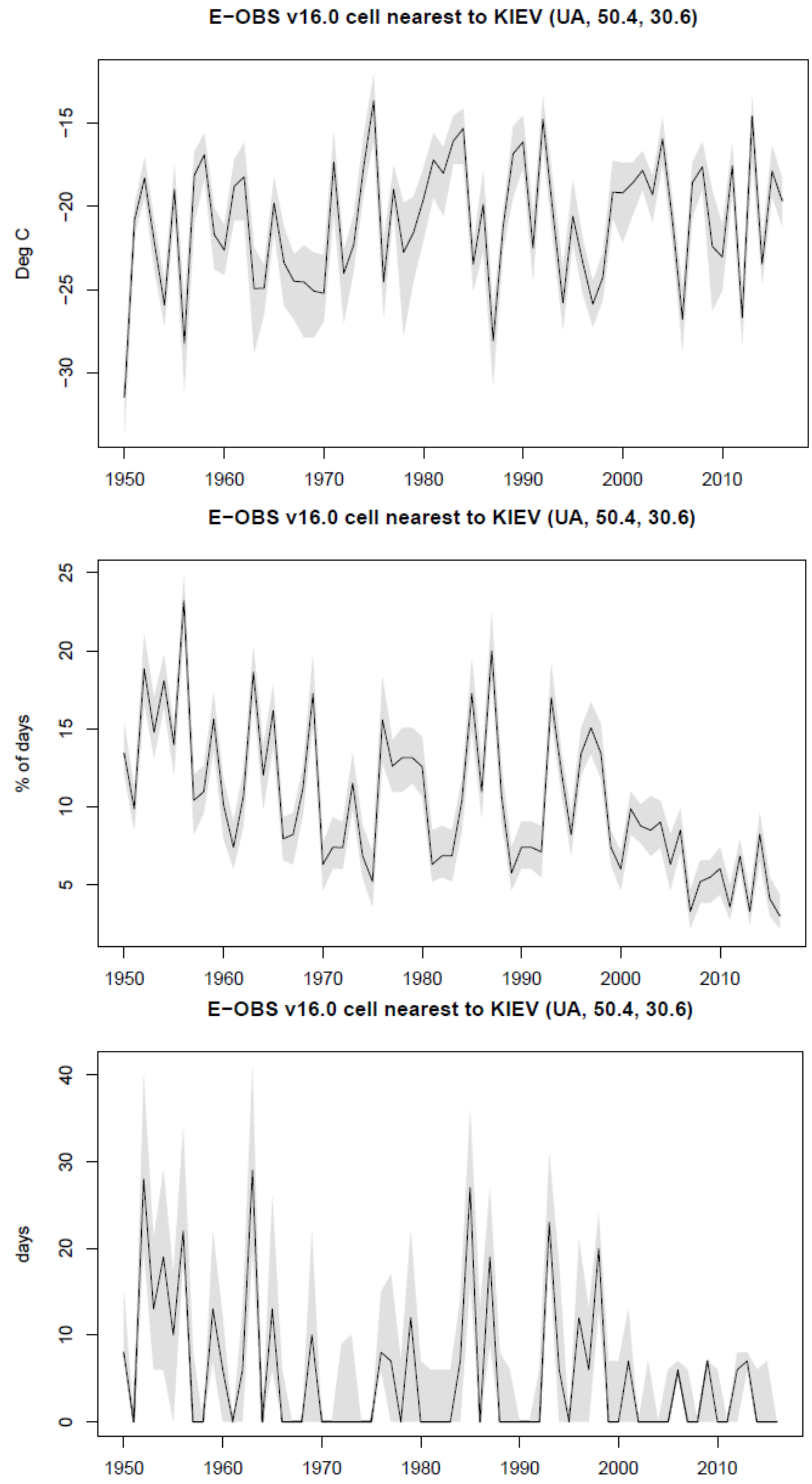


Figure 10: Annual values of TNn (˚C), TN10p (% of days) and CSDI (days) for Kiev.

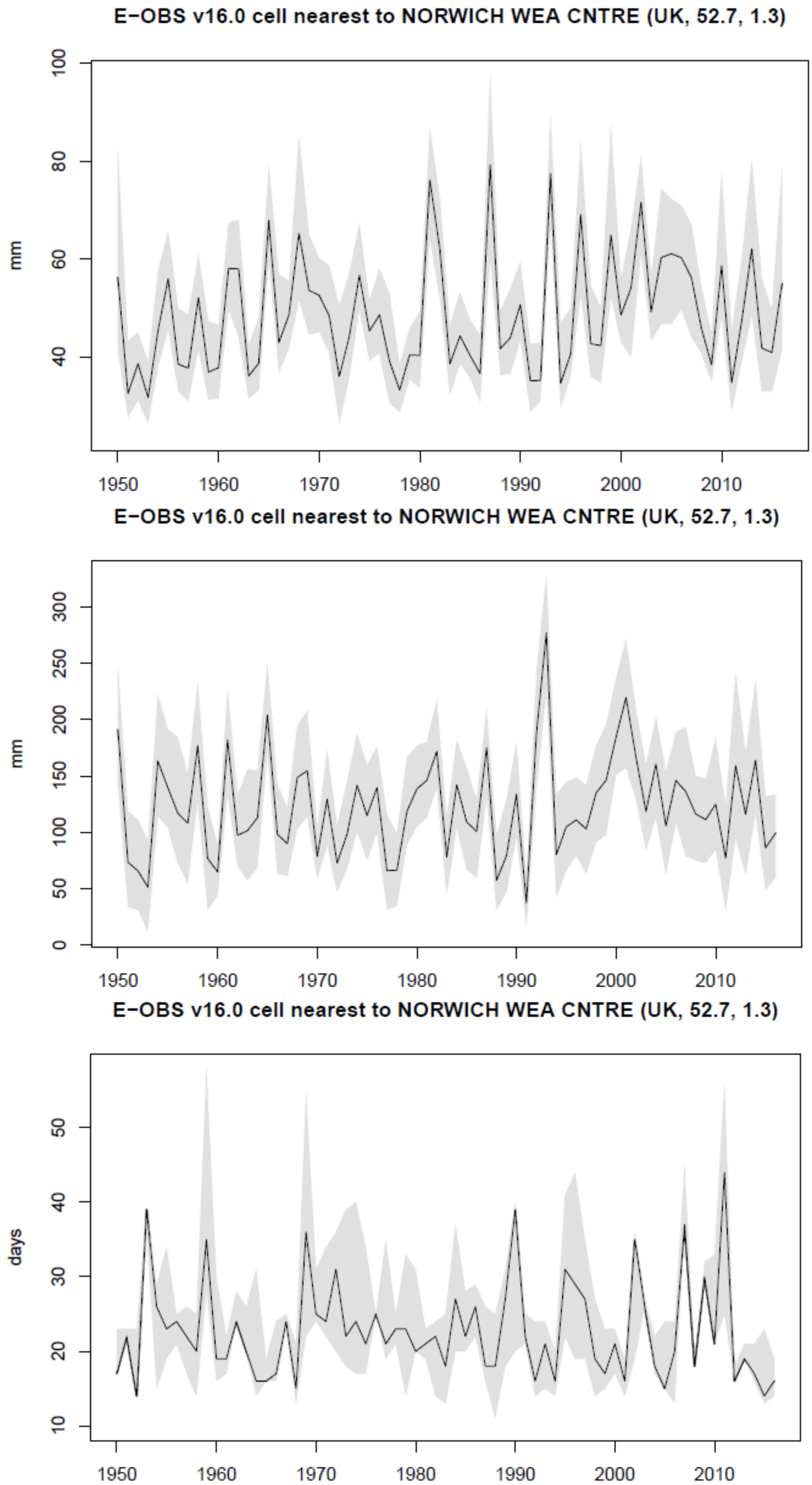


Figure 11: Annual values for RX5day (mm), R95pTOT (mm) and CDD (days) for Norwich.

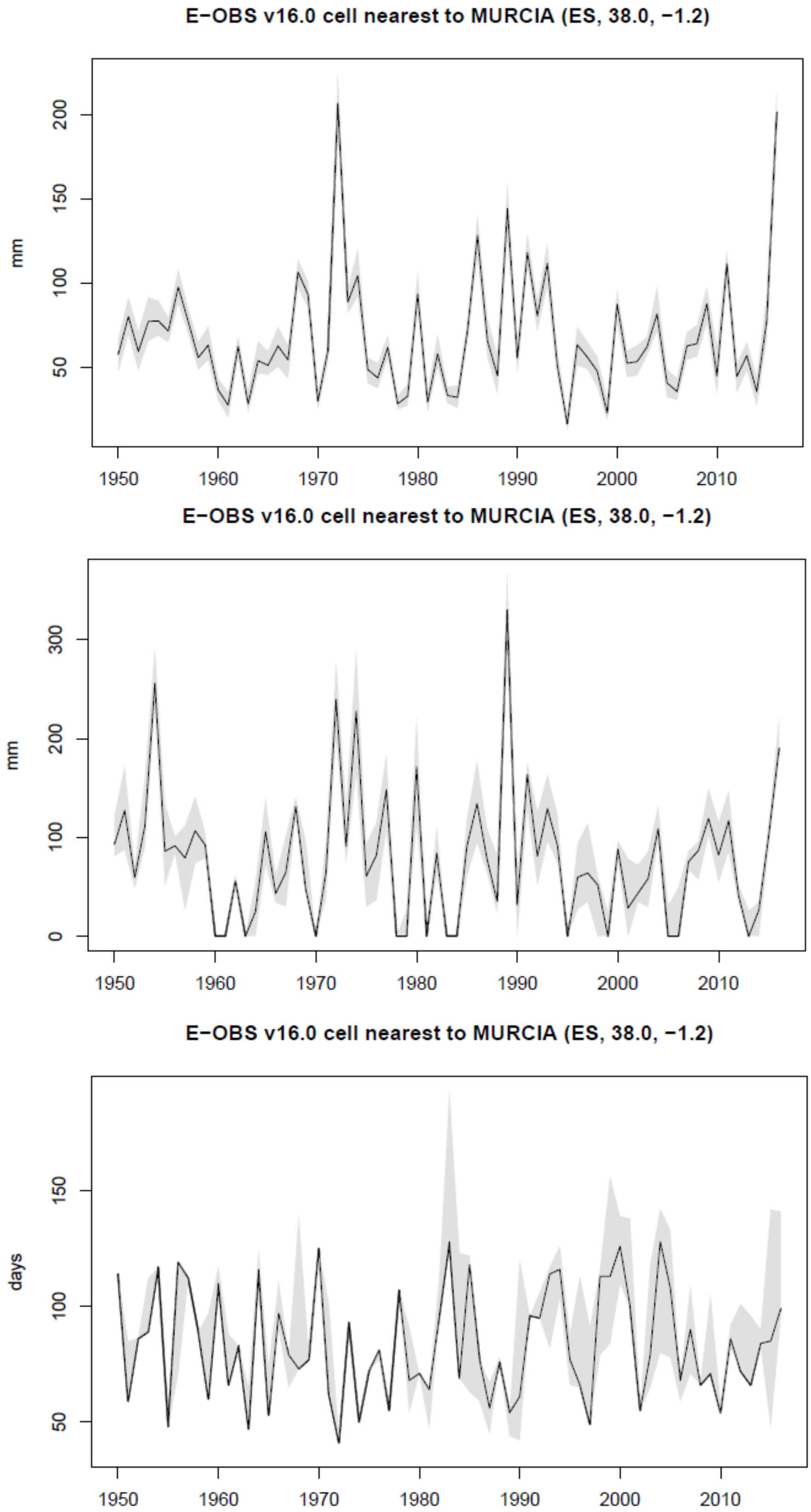


Figure 12: Annual values for RX5day (mm), R95pTOT (mm) and CDD (days) for Murcia.

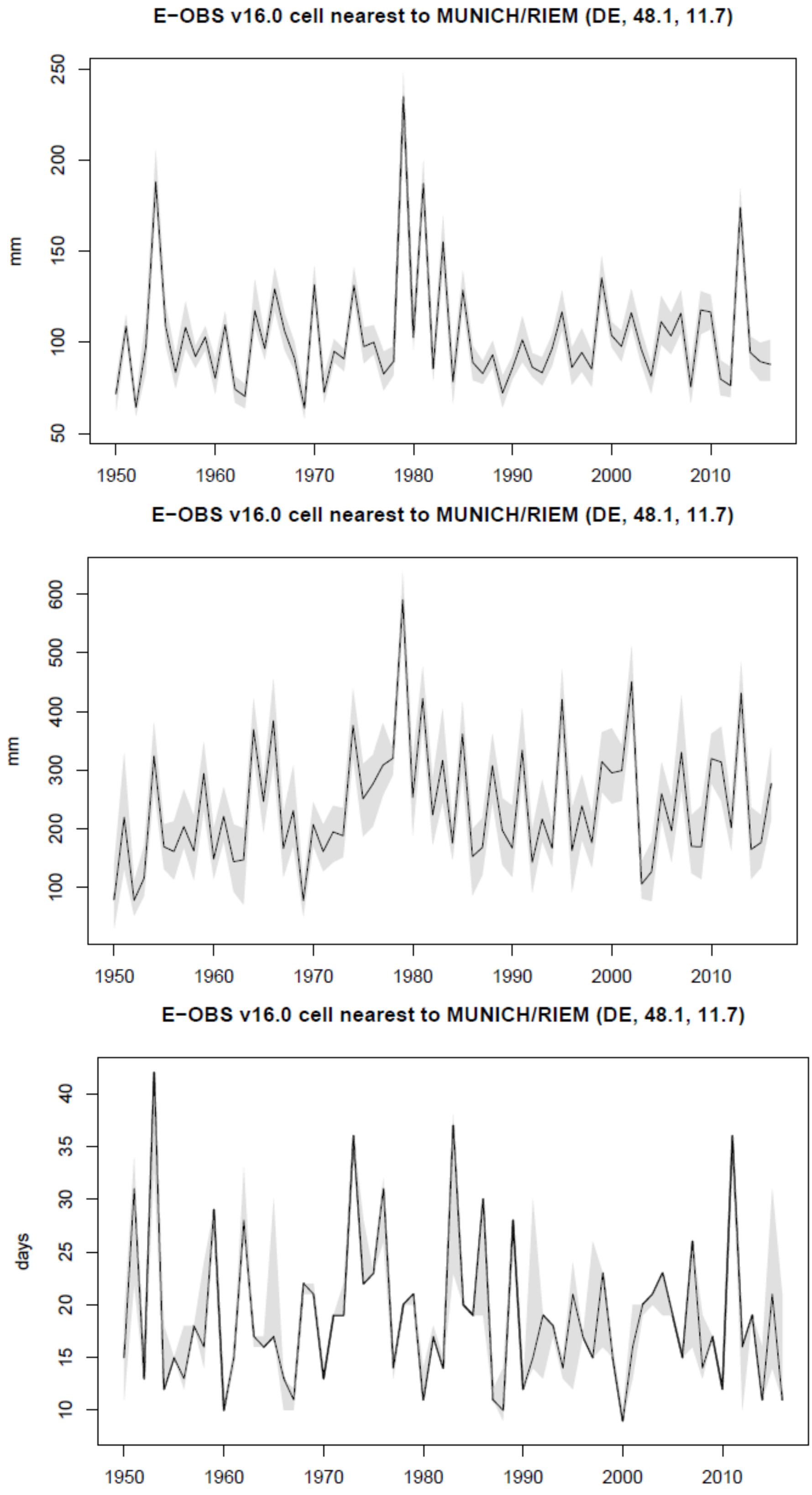


Figure 13: Annual values for RX5day (mm), R95pTOT (mm) and CDD (days) for Munich.

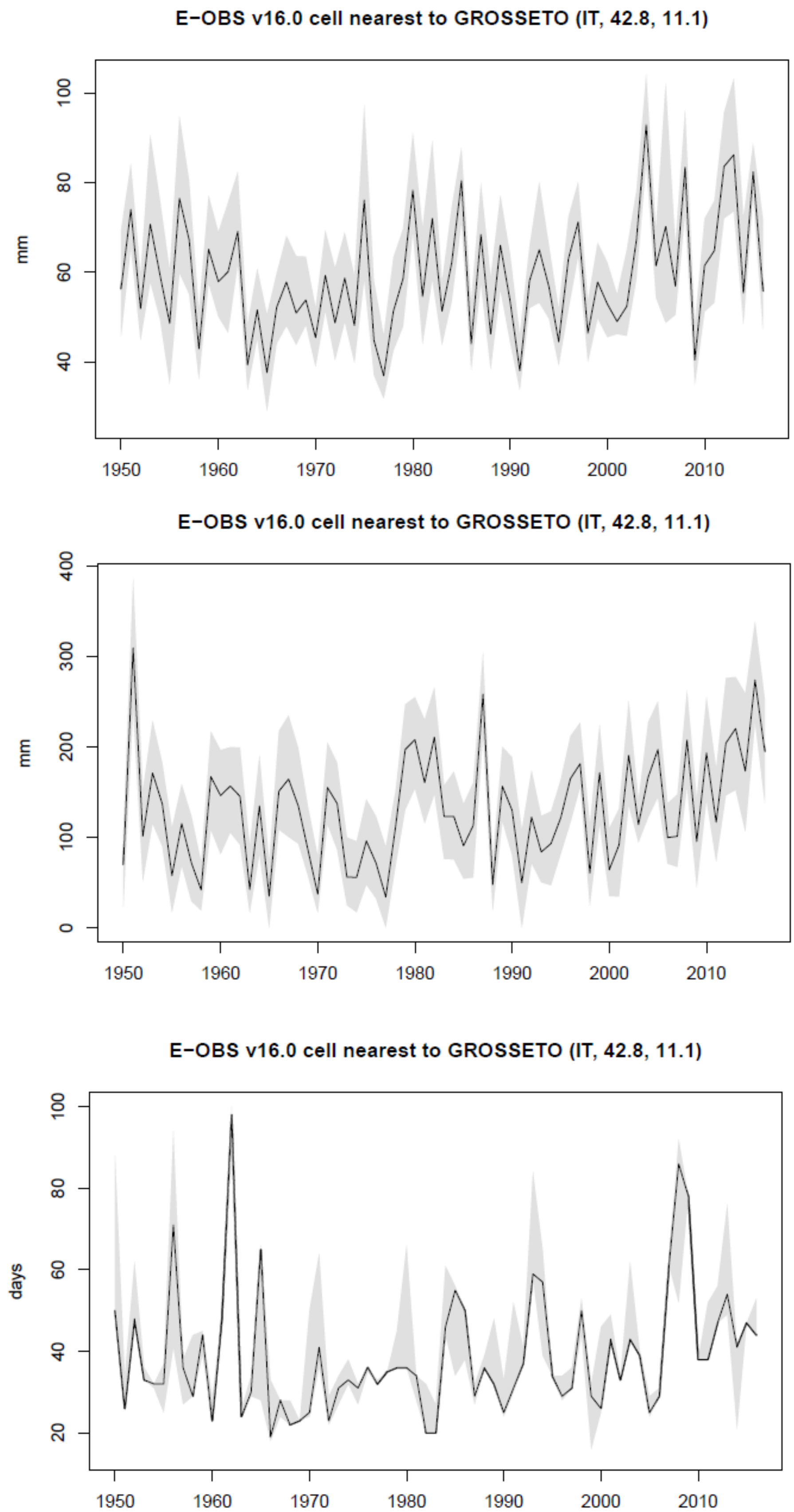
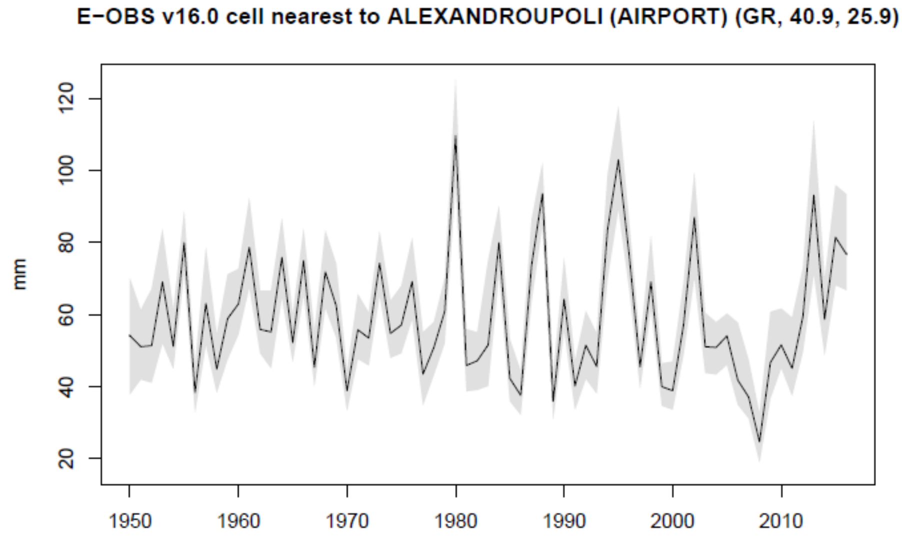
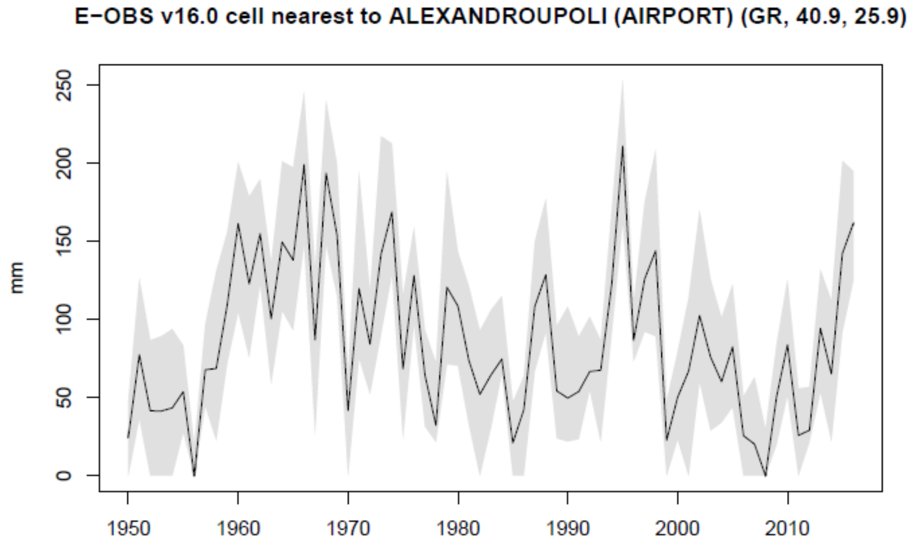


Figure 14: Annual values for RX5day (mm), R95pTOT (mm) and CDD (days) for Grosseto.





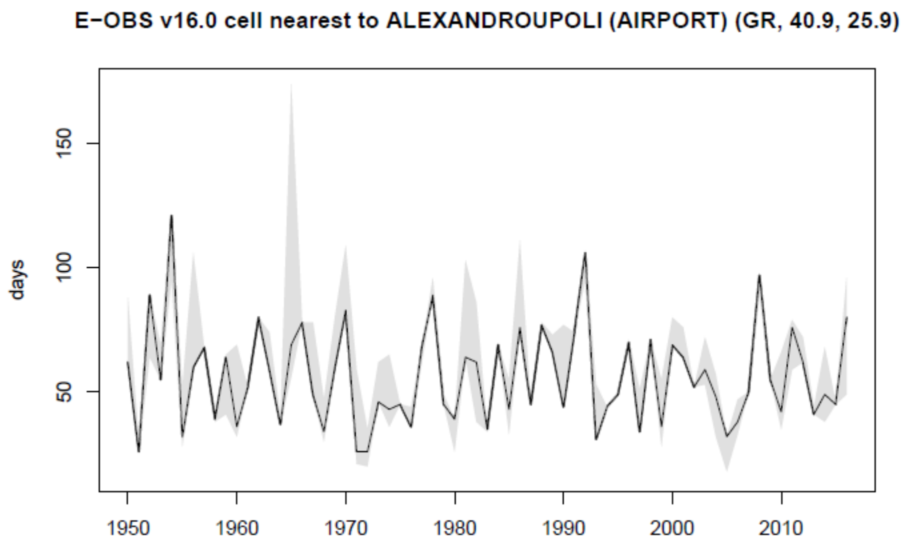


Figure 15: Annual values for RX5day (mm), R95pTOT (mm) and CDD (days) for Alexandroupoli.

Copernicus Climate Change Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-climate.eu