H2020-SC5-01-2017

MED-GOLD

Turning climate-related information into added value for traditional **MED**iterranean **G**rape, **OL**ive and **D**urum wheat food systems

D1.3

Assessment of quality of European climate observations and their appropriateness for use in climate services for each sector

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776467.

DOCUMENT STATUS SHEET

Deliverable Title	**Report assessing the quality of European climate observations and their appropriateness for use in climate services for each sector.**	
Brief Description		
WP number	1	Setting the scene
Lead Beneficiary	Met Office	
Contributors	*NOA, ENEA, SOGRAPE, DCOOP, JRC*	
Creation Date	12/09/2018	
Version Number	1.8	
Version Date	29/05/2019	
Deliverable Due Date	31/05/2019	
Actual Delivery Date	29/05/2019	
Nature of the Deliverable	R	*R - Report* *P - Prototype* *D - Demonstrator* *O - Other*
Dissemination Level/ Audience	PU	*PU - Public* *PP - Restricted to other programme participants, including the Commission services* *RE - Restricted to a group specified by the consortium, including the Commission services* *CO - Confidential, only for members of the consortium, including the Commission services*

REVISION HISTORY LOG

Version	Date	Created / Modified by	Pages	Comments
1.0	12-09-2018	Met Office	14	Initial Draft
1.1	01-10-2018	ENEA / Met Office	32	Assessment of data in Douro valley area added
1.2	02-10-2018	NOA / HORTA / Met Office	46	Assessment of data in Andalucía added. Table of weather stations in Ravenna added.
1.3	16-10-2018	ENEA / Met Office	52	Assessments of climatic indicators in Douro valley area added; conclusions added.
1.4	14-11-2018	Met Office	62	Assessment of datasets for Colombia added.
1.5	16-11-2018	JRC / Met Office	70	Assessment of climatic data for study regions in Italy added.
1.6	27-11-2018	Met Office	71	Edited following internal reviewers' comments.
1.7	20-05-2019	Met Office / ENEA	75	Further changes following review by EC and external experts
1.8	29-05-2019	Met Office	75	Added new section 7.4, plus final editing

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

EXECUTIVE SUMMARY

Temperatures and precipitation amounts from many different gridded datasets have been compared with weather observations in the key MED-GOLD study regions: Andalucía (southern Spain), Douro valley (Portugal), Foggia, Ancona and Ravenna (Italy), and Colombia. Monthly means, seasonal means of temperature, and the corresponding precipitation totals, plus climatic indices relevant for each crop type, were calculated from gridded datasets and local weather station data. Comparisons between these two datasets showed that, for Europe, the E-OBS gridded dataset consistently had the closest agreement with the local observations. It is recommended that this dataset could be used within MED-GOLD in areas lacking observations from weather stations, and evaluation of regional climate model simulations. For Colombia, an equivalent dataset to E-OBS doesn't exist. Instead, the AgMERRA dataset is recommended.

With this deliverable, the project has contributed to the achievement of the following objectives (DOA, Part B Table1.1):

No.	Objective	Yes
1	To co-design, co-develop, test, and assess the added value of proof-of-concept climate services for olive, grape, and durum wheat	X
2	To refine, validate, and upscale the three pilot services with the wider European and global user communities for olive, grape, and durum wheat	
3	To ensure replicability of MED-GOLD climate services in other crops/climates (e.g., coffee) and to establish links to policy making globally	X
4	To implement a comprehensive communication and commercialization plan for MED-GOLD climate services to enhance market uptake	
5	To build better informed and connected end-user communities for the global olive oil, wine, and pasta food systems and related policy making	

1. INTRODUCTION

Agriculture is primarily driven by weather and climate. Current projections of future climatic conditions indicate these crops will become increasingly vulnerable to failure and pest damage. The Mediterranean Basin is a hot-spot of global change, where higher than average projected climate change threatens an extremely rich and intertwined biological and cultural diversity, and will increase its vulnerability to natural hazards including biological invasions. A major challenge for this region is how to increase the resilience of this complex ecological, economic, and cultural heritage of global relevance in an era of decreasing resources and climate change.

The overall aim of MED-GOLD is to demonstrate climate services in agriculture by developing case studies for three staple parts of the Mediterranean food system: grapes, olives and durum wheat. High quality meteorological observations are a key requirement for the development of climate services. The response of these three crops (in terms of yields, quality, pest infestations, for example) to different weather patterns needs to be understood. Similarly, past severe climatic events, such as heat waves, droughts, or periods of heavy rain need to be understood and characterised. The impact of these events on crop growth, yield and quality can then be studied and understood. In some cases, models of the crops themselves have been developed to further understanding of how climatic events have affected the crops. These models can then be used to assess the impacts of future changes in climate on crop growth, yield and quality.

The objective of this report is to compare meteorological variables from weather stations with those in gridded datasets, and assess the suitability of these gridded datasets for use in development of climate services for the three crop types (grapes, olives and durum wheat). Values at coincident locations can be compared to assess how well the gridded data represent local conditions. If the agreement between the local observations and gridded data is satisfactory, it is reasonable to assume that the gridded data could be used to develop and assess climate service tools in regions where local weather observation are unavailable. This report focuses on temperature and rainfall, which were identified as key variables for all three crops (olives, grapes and durum wheat). These two variables are readily available from weather stations and gridded datasets.

The observations and gridded datasets used in this report are summarised in sections 5 and 6 respectively. Observations from the weather stations in each region are compared with the gridded datasets in section 7. A brief conclusion is given in section 8. The suitability of these gridded datasets for use in development of climate services for the three crop types (grapes, olives and durum wheat) is assessed. The most appropriate datasets are recommended.

2. PURPOSE

This report describes the comparison of many gridded meteorological datasets with local weather station observations in key study regions for MED-GOLD. Any gridded dataset(s) which have a satisfactory agreement with the local observations will be highlighted and recommended for further use within MED-GOLD, such as evaluation of seasonal forecasts and climate projections.

2.1. SCOPE

This report constitutes Deliverable D1.3 “Report assessing the quality of European climate observations and their appropriateness for use in climate services for each sector”.

2.2. DEFINITIONS AND ACRONYMS

2.2.1. DEFINITIONS

Concepts and terms used in this document and needing a definition are included in the following table:

Table 2-1: Definitions

Concept / Term	Definition
Climate services	The transformation of climate-related data and other information into customised products.
Gridding	The interpolation of irregularly distributed surface-based weather observations onto a regular grid.
Gridded data	Data produced by the gridding process. The spatial coverage of the gridded data may be global, continental or highly local. The temporal resolution can be monthly, daily or hourly.
Observations	A catch-all term for both surface meteorological observations and gridded data derived from them.
Reanalysis	A comprehensive meteorological dataset created by integrating a weather forecast model forward in time and ingesting all available observations every 6-12 hours, producing a dynamically consistent estimate of the state of the climate at each time step.
(Spatial) Resolution	The distance between points in a gridded dataset. The distances may be given in degrees, kilometres or both. “High resolution” refers to a dataset with small distances between points (of the order of a few kilometres). “Low resolution” means the distances between points are large, e.g., a few degrees or the order of hundreds of kilometres.
Surface observations	Meteorological observations made by a surface-based weather station. The number and type of meteorological variables measured, and the temporal frequency of the measurements, varies between stations.

2.2.2. ACRONYMS

Acronyms used in this document and needing a definition are listed in the following table:

Table 2-2: Acronyms

Acronym	Definition
ASOS	Automated Surface Observing Systems
ECA&D	European Climate Assessment and Data set
GHCN-D	Global Historical Climate Network - Daily
IBIMET	The Institute of Biometeorology at the National Research Council (CNR), Italy
IPMA	Instituto Português do Mar e da Atmosfera
MED-GOLD	The project entitled “Turning climate-related information into added value for traditional MEDiterranean Grape, OLive and Durum wheat food systems”
CMIP5	Coupled Model Intercomparison Project Phase 5.
RCM	Regional climate model
RMSE	Root mean square error
NRSME	Normalised Root Mean Square Error
MAE	Mean Absolute Error
RR	Precipitation (Rainfall Rate)

Acronym	Definition
JJA	Summer season
MAM	Spring season
DJF	Winter season
TN or Tmin	Minimum temperature
TX or Tmax	Maximum temperature
TM, Tmean or Tavg	Average temperature
WP	Work Package

3. REFERENCES

3.1. REFERENCE DOCUMENTS

The following documents, although not part of this document, amplify or clarify its contents. Reference documents are those not applicable and referenced within this document. They are referenced in this document in the form [RD.x]. Articles in scientific journals, books and other external publications are listed in the bibliography.

Table 3-1: Reference Documents

Ref.	Title	Code	Version	Date
[RD.1]	“Indices_Explained”	-	-	Not dated
[RD.2]	Report on the knowledge capitalization of the olive oil sector	D2.1	1.9	25-09-2018
[RD.3]	Report on the two case studies at seasonal- and long-term timescales for the wine sector	D3.1	1.5	23-08-2018
[RD.4]	Report on the identified specific needs and opportunities	D4.1	4.0	29-11-2018

4. DOCUMENT CONTENT

The surface observations are described in section 5. and the gridded datasets in section 6. The gridded data and surface observations are compared in section 7. to understand the representativeness of the gridded datasets in each MED-GOLD study region. The gridded data would be used to assess regional climate model simulations as well as providing raw data for development of the climate service tools. Recommended data for use within MED-GOLD is summarised in section 8.

5. SURFACE OBSERVATIONS

In this section, the sources of meteorological observations in each of the study regions are briefly described. The focus is on the key climatic variables identified in WPs 2, 3 and 4 ([RD.1] to [RD.4]), namely temperature and precipitation. The same variables were identified in all three scoping workshops ([RD.2] to [RD.4]), so focus is on sources of observations with those variables.

For the purposes of MED-GOLD, an ideal surface observing network would have the following attributes:

- Stations located amongst the crops in the areas of interest;
- a high density of observing stations;
- long series of observations which coincide with phenological and other relevant crop data;
- no missing data;
- observations made on at least a daily basis (preferably a higher frequency)

However, it would be highly unusual to find all of these characteristics in any network. In most observing networks, the coverage is sparse, and few stations are located amongst the crops. Equipment failures or other technical issues mean that observations on some days will be erroneous, not made or not recorded. The area in which a station is situated can change considerably, altering the local climatic environment (e.g., a change from a rural to urban environment at Uccle, Belgium, as described by van Weverberg et al. (2008)). Movement of weather stations and changes in the instruments used to make the observations can introduce artefacts into the data, which need to be identified and corrected.

Meteorological observations from several different sources have been assessed; these sources are summarised in sections 5.1 - 5.4 and Appendix A. The locations of the weather stations used within Europe are shown in Figure 5-1.

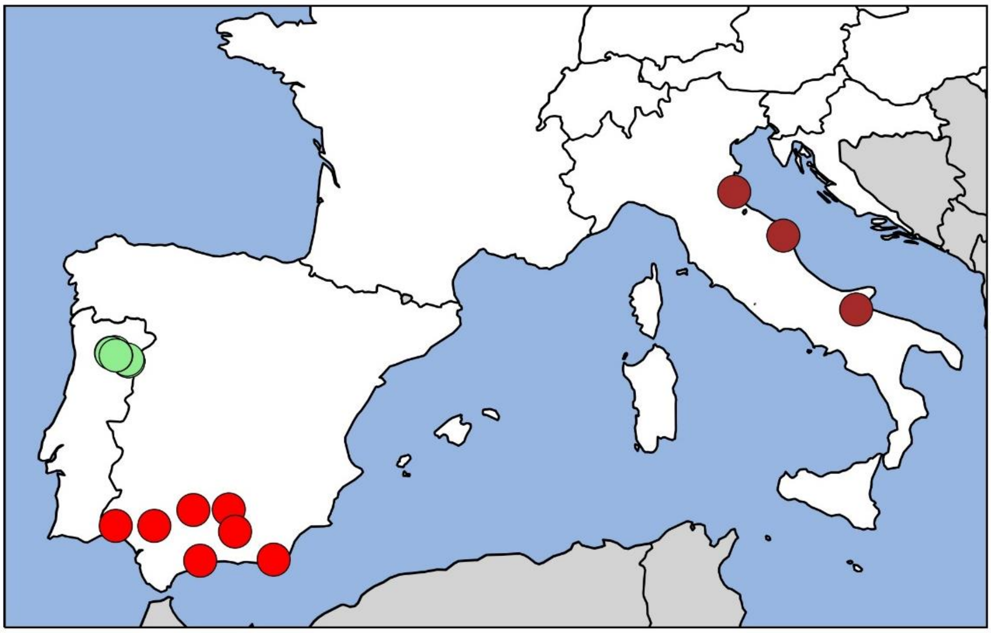


Figure 5-1: Map showing the locations of the weather stations in Europe used in the assessments. Red circles – Andalucía, Spain; Green circles – Douro valley, Portugal; Brown circles – three sites in Italy, from north to south: Ravenna, Ancona and Foggia.

5.1. SPAIN - ANDALUCÍA

Meteorological data from seven stations in the study area were available (Table 5-1). Temperature, precipitation and other data from these seven weather stations were available for 2001-2017 (see also Appendix A.1). The locations of these stations within Spain together with the altitude of the land areas are shown in Figure 5-2.

Table 5-1: Weather stations in Andalucía

Station	Latitude / °	Longitude / °	Altitude / m
Sevilla	37.457	-5.925	37
Malaga	36.756	-4.537	68
Jaen	37.891	-3.771	299
Huelva	37.412	-7.060	169
Granada	37.416	-3.551	935
Cordoba	37.857	-4.803	117
Almeria	36.835	-2.402	22

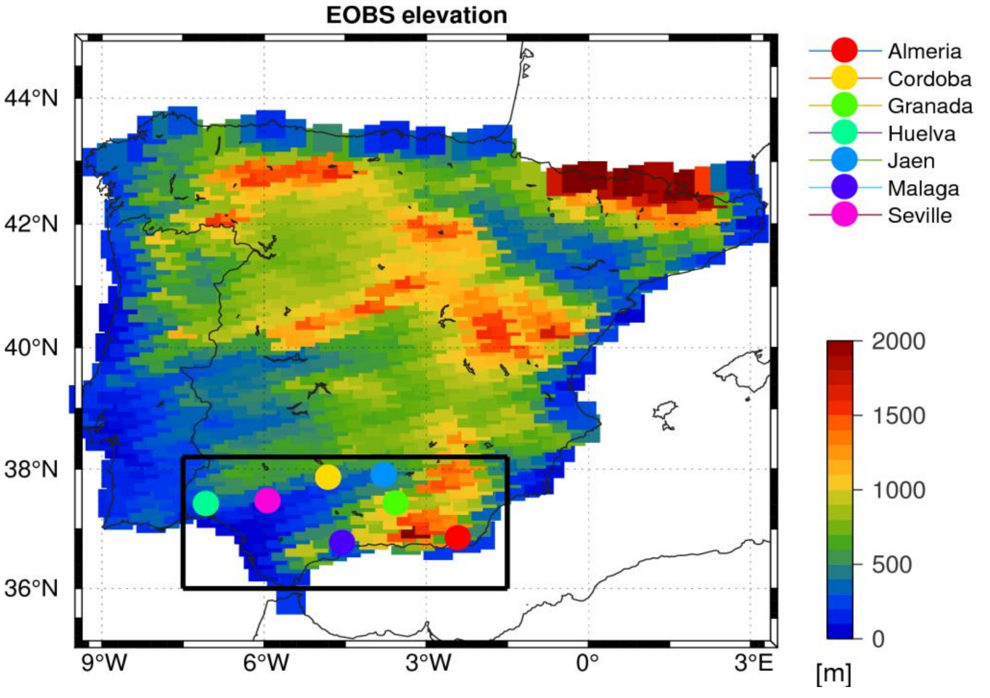


Figure 5-2: Map showing elevation in E-OBS (in m) for the Iberian Peninsula. The locations of the seven weather stations in Andalucía (Table 5-1) are shown by the coloured circles within the highlighted box.

Previous work has shown that two key variables for modelling activity of the olive fruit fly are temperature (Kalamatianos et al., 2016) and relative humidity (Broufas et al., 2009). The WP2 olive oil workshop held in Antequera [RD.1] identified temperature and precipitation as key variables of interest for olive grove management. These latter two variables are related to decisions concerning the olive sector production chain which can be divided into three categories:

- Main decisions: pest management, fertilization, harvest
- Secondary decisions: irrigation, tillage, pruning
- Other decisions: olive pests, olive diseases, vegetation cover, lipogenesis

(lipogenesis is the process by which oil is generated within the olives).

Temperature and precipitation data from the seven weather stations in Table 5-1 were post-processed to calculate climatological monthly means and totals of key meteorological variables (temperature and rainfall). These data were analysed and compared with the E-OBS v17 (Haylock et al., 2006) gridded daily observational dataset for the common period in all selected stations, 2001-2017. The seven weather stations are not part of the network used to create the E-OBS dataset, and so provide a useful independent test of E-OBS in Andalucía. The comparison is made in order to investigate whether the E-OBS dataset can be used for evaluation of regional climate model (RCM) data. The closest grid point in E-OBS to each weather station has been selected for comparison. The variables analysed are daily maximum and minimum temperature and daily total precipitation. Daily distributions and the mean annual cycles of both the observations and E-OBS data are analysed. The differences between the two datasets for a number of climatic indices related to the olive sector were also examined. These indices were provided by the DCOOP experts [RD.2] and are listed in Table 5-2.

Table 5-2: Climate indices relevant for the olive sector

Indices	Impact
Mean summer maximum temperature (TX)	Related to irrigation and fertilization-strategy for possible treatments against pests and diseases
Mean winter minimum temperature (TN)	Plant treatment and harvest
Number of days with TX > 30°C in Spring	Plant treatment and irrigation
Number of days with TX > 40°C in Summer	Plant treatment and irrigation-to reinforce plants against extreme temperatures
Total precipitation (Annual, Winter, Summer)	Influences all the operations performed during the year, from fertilization to the moment of pest treatment application
Dry days (Number of days with precipitation < 2 mm (Annual, Winter)	To know when to apply plant treatments and for crop management (pruning and harvesting)

5.2. PORTUGAL - DOURO VALLEY

Meteorological observations were available from several different sources (Appendix A.2). Key observed variables for the grape/wine sector are daily mean temperatures and daily rainfall totals. Five key indices for grapes are calculated from these variables, four from temperature and one from precipitation [RD.2]. The assessment of the gridded datasets focuses on comparisons with the data from weather stations operated by SOGRAPE and IPMA in the Douro valley (Table 5-3). Data from the weather stations operated by SOGRAPE are only available for 2011-2018, whereas longer data series are available from the IPMA stations. The positions of these stations are indicated by the solid red circles in Figure 5-1. The GHCN-D database was checked, but none of the weather stations were located in the Douro valley.

Table 5-3: Weather stations operated by SOGRAPE and IPMA in the Douro valley

Station	Latitude / °	Longitude / °	Altitude / m	Operator	Data Period
Pinhão - Caêdo	41.190	-7.458	210	SOGRAPE	2011-2018
Leda 1	41.022	-7.016	213	SOGRAPE	2011-2018
Leda 2	41.017	-7.018	318	SOGRAPE	2011-2018
Leda 3	41.047	-7.047	163	SOGRAPE	2011-2018
Pinhão - Porto	41.173	-7.570	193	SOGRAPE	2011-2018
Pinhão - Seixo	41.167	-7.553	199	SOGRAPE	2013-2018
Sairrão 1	41.126	-7.403	624	SOGRAPE	2011-2018
Sairrão 2	41.121	-7.411	548	SOGRAPE	2011-2018
Sairrão 3	41.119	-7.420	426	SOGRAPE	2011-2018
Santa Barbara	41.173	-7.549	130	IPMA	1941-2018
Folgosa	41.149	-7.684	100	IPMA	1964-2009
Guiaes	41.209	-7.660	560	IPMA	1964-2009

5.3. ITALY

Meteorological observations were obtained from three weather stations operated by IBIMET, one per study area (Table 5-4). These three weather stations have reasonably long data periods and very few missing data.

Table 5-4: Weather stations operated by IBIMET in the study regions in Italy

Station	Latitude / °	Longitude / °	Altitude / m	Data Period
Foggia	41.50	15.52	0	2013-2018

Station	Latitude / °	Longitude / °	Altitude / m	Data Period
Ancona	43.33	13.58	91	2007-2018
Ravenna	44.41	12.20	27	2007-2016

5.4. COLOMBIA

Meteorological observations from two sources were considered for Colombia, the global historical climate network – daily database (GHCN-D; Menne et al., 2012) and the Automated Surface Observing Systems (ASOS; National Oceanic and Atmospheric Administration, 1998). Data from many of the weather stations in Colombia are present in both databases. Daily data are archived in GHCN-D, whereas hourly data are available from ASOS. A map showing the locations of the stations in GHCN-D and ASOS is shown in Figure 5-3.

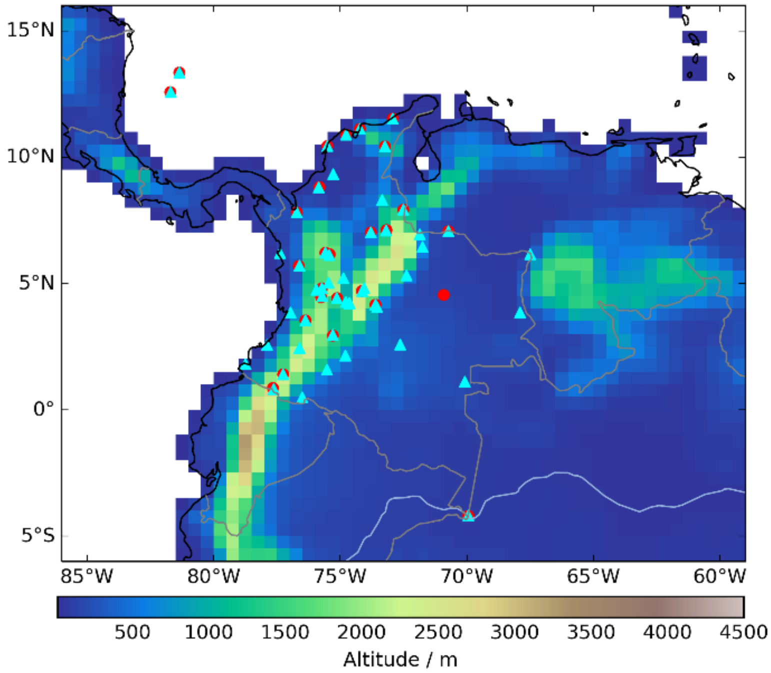


Figure 5-3: Map of Colombia and surrounding countries, showing locations of weather stations in the databases GHCN-D (red circles) and ASOS (cyan triangles). The altitude data are on a regular 0.5 degree grid.

5.4.1. GHCN-D

Data from 27 weather stations in Colombia were available from the GHCN-D (Table 5-5); their locations are shown in Figure 5-3. Some are located at higher altitudes, in or close to areas where coffee is grown. The only climatic variables available were daily maximum, minimum and mean temperatures, and daily total precipitation. Data are available from the 1960s for many stations. However, there were a large number of days with missing data at some stations. Initially, the percentages of days with valid data were calculated for each station and climate variable, which are listed in the final four columns of Table 5-5. Generally, daily mean temperatures (Tavg) have the highest data availability, whereas availability of daily maximum temperatures (Tmax) is often low.

The numbers of days with valid data in each month and year were also calculated (not shown). The periods with valid data varied considerably between the stations. For some, few data were missing between about 1973 and 2000, but outside of this period missing data were much more common, and the numbers of days per month with valid data were around 15 – 25. For other stations, data were present from 2010 with few missing days, whereas the coverage was sparse in earlier years.

Climatological monthly means and totals of climatic variables from the weather stations will be compared with the gridded datasets. Given the large amounts of missing data at some stations, the following approach was taken. First, for each climatic variable in Table 5-5, the months with complete data series (i.e. no days with missing data) were identified for each station. Monthly means of the three temperature variables and monthly precipitation totals were calculated for these months. Next, the minimum, mean and maximum values were calculated for each month, which produced the climatological values.

Table 5-5: Weather stations in Colombia from the GHCN-D database, ordered by decreasing latitude. The numbers in the final four columns (headed Tmax, Tmin, Tavg and Precip) are the percentages of days with valid data for the corresponding meteorological variable.

Station Code	Station Name	Lat / °	Lon / °	Altitude / m	First Date	Last Date	Tmax	Tmin	Tavg	Precip
COM00080002	EL EMBRUJO	13.357	-81.358	3.0	1979-04-01	2018-03-31	9.6	13.8	23.5	13.6
CO000080001	SAN ANDRES (ISLA)/S	12.583	-81.717	6.0	1962-01-01	2018-03-31	55.7	59.1	86.8	82.8
COM00080035	ALMIRANTE PADILLA	11.526	-72.926	13.1	1973-01-01	2018-03-31	12.2	23.1	33.0	13.0
COM00080009	SIMON BOLIVAR	11.12	-74.231	6.7	1964-09-01	2018-03-31	30.8	48.2	73.2	24.5
COM00080028	ERNESTO CORTISSOZ	10.89	-74.781	29.9	1941-05-01	2018-03-31	31.2	39.3	76.7	26.9
COM00080022	RAFAEL NUNEZ	10.442	-75.513	1.2	1963-11-01	2018-03-31	40.4	42.7	90.4	38.7
COM00080036	ALFONSO LOPEZ PUMAREJO	10.435	-73.25	147.2	1969-08-01	2018-03-31	11.8	29.2	38.3	16.6
COM00080063	LOS GARZONES	8.824	-75.826	11.0	1964-09-01	2018-03-31	13.0	22.1	34.2	10.3
COM00080097	CAMILO DAZA	7.928	-72.512	334.1	1964-09-01	2018-03-31	27.9	41.7	65.2	28.1
COM00080084	ANTONIO ROLDAN BETANCOURT	7.812	-76.716	14.0	1984-12-01	2018-03-31	6.1	25.4	34.2	19.8
COM00080094	PALONEGRO	7.127	-73.185	1187.8	1975-04-01	2018-03-31	45.8	54.0	83.8	49.5
COM00080099	SANTIAGO PEREZ	7.069	-70.737	128.0	1964-09-01	2018-03-31	14.3	25.4	40.4	17.0
COM00080091	YARIGUIES	7.024	-73.807	125.6	1964-09-01	2018-03-31	16.6	35.6	55.3	25.9
COM00080110	OLAYA HERRERA	6.22	-75.591	1505.7	1958-01-01	2018-03-31	20.4	41.0	63.3	30.6
COM00080112	JOSE MARIA CORDOVA	6.165	-75.423	2142.1	1985-12-01	2018-03-31	36.3	54.2	95.2	58.1
COM00080144	EL CARANO	5.691	-76.641	62.2	1963-01-01	2018-03-31	7.8	22.8	39.2	24.2
COM00080210	MATECANA	4.813	-75.74	1346.0	1963-03-01	2018-03-31	28.5	49.4	82.4	42.6
CO000080222	BOGOTA/ELDORADO	4.701	-74.15	2548.0	1941-03-01	2018-03-31	43.5	56.5	77.6	49.5
CO000080241	LAS GAVIOTAS	4.55	-70.917	167.0	1967-08-01	2002-05-31	59.5	59.0	14.9	86.0
COM00080211	EL EDEN	4.453	-75.766	1216.2	1964-09-01	2018-03-31	14.6	27.6	41.3	18.0
COM00080214	PERALES	4.422	-75.133	949.1	1973-01-01	2018-03-31	21.7	39.1	49.5	29.8
COM00080234	VANGUARDIA	4.168	-73.614	424.9	1964-09-01	2018-03-31	10.8	35.6	49.5	23.8
CO000080259	CALI/ALFONSO BONILL	3.55	-76.383	969.0	1961-01-01	2018-03-31	61.1	69.3	95.2	62.4
COM00080315	BENITO SALAS	2.95	-75.294	446.2	1964-09-01	2018-03-31	27.5	40.7	58.0	30.1
CO000080342	PASTO/ANTONIO NARIN	1.417	-77.267	1826.0	1957-01-01	2018-03-31	37.7	59.7	43.1	76.9

Station Code	Station Name	Lat / °	Lon / °	Altitude / m	First Date	Last Date	Tmax	Tmin	Tavg	Precip
COM00080370	SAN LUIS	0.862	-77.672	2976.4	1964-09-01	2018-03-31	4.3	33.8	41.2	22.7
COM00080398	ALFREDO VASQUEZ COBO	-4.193	-69.943	84.4	1967-08-01	2018-03-31	30.4	46.8	88.8	47.2

5.4.2. ASOS

Data from 51 stations were available from the ASOS database. However, data were only available from 2010-2016, depending on the station. The length of the data series varied considerably – for some stations, data for just one or two hours of a small number of days were available, whereas near-complete series were available from others. For some stations, data from only part of the day had been archived – e.g., 14:00 – 20:00. For these reasons, data from ASOS over Colombia were not used.

6. GRIDDED DATASETS

Gridded datasets consist of estimates of meteorological data at regular intervals, that is, on a regular grid. They are produced by applying an interpolation algorithm to irregularly-spaced surface observations to estimate values at regular intervals. Gridded datasets can be global, continental or local in scale, and have many useful features. They provide complete coverage of a region, including estimates of meteorological variables in places far from any observing stations, and (generally) do not contain points with missing data. They often contain long series of data, typically 50-60 years or more. Gridded datasets can be used to calculate averages or totals of meteorological variables over large areas, which are useful for understanding of larger-scale variations in climate. They are commonly used to evaluate global and regional climate models and drive other models.

A number of gridded datasets were assessed by comparison with surface-based observations (section 6.1). Some issues with gridded datasets are discussed in section 6.2.

6.1. GRIDDED DATASETS ASSESSED

The gridded datasets assessed in this report consist of three main types. Many gridded datasets are derived by use of a statistical model to estimate values of meteorological variables, such as temperature and precipitation, on a regular grid from irregularly-spaced surface weather stations. A wide range of methods to interpolate the observations vary widely, including area-averaging, kriging, or other statistical models that use additional information such as topography, proximity to urban areas, and the underlying climatology of the area of interest (e.g., Perry and Hollis, 2005). Examples of this first type of gridded data are E-OBS (Table 6-1) and others not assessed here but which could still be useful (e.g., Spain02, PTHRES; see Appendix B).

Table 6-1: Gridded datasets assessed in this report. Further details of each dataset are given in Appendix B. The resolutions given are valid for mid-latitudes and are approximate.

Dataset name	Type[1]	Period	Variables[2]	Resolution / km
E-OBS v17	Interpolated	1961-2017	Precip, Tmax, Tmin	25
JRC gridded (MarsMet)	Interpolated	1975 - present	Precip, Tmax, Tmin, Tmean, RH, + others	25
NCEP2	Reanalysis	1979-present	Precip, Tmax, Tmin	250
ERA-Interim [EIN75]	Reanalysis	1979-present	Precip, Tmax, Tmin	80
ERA5	Reanalysis	2000-present	Precip, Tmax, Tmin	30
JRA-55	Reanalysis	1979-present	Precip, Tmax, Tmin	55
MERRA2	Reanalysis	1980-present	Precip, Tmax, Tmin	50
AgMERRA[3]	Reanalysis	1980-2010	Precip, Tmax, Tmin	50
NEX	Climate Projection	1979-2016	Precip, Tmax, Tmin	25
CHIRPS	Satellite	1981-2017	Precip	5.5
SM2RAIN	Satellite	2007-2017	Precip	12.5
UERRA[4] (eswi)	Reg. Reanalysis	1979-2017	Precip, Tmax, Tmin	5.5 – 35
UERRA[4] (egrr)	Reg. Reanalysis	1980-2017	Precip, Tmax, Tmin	5.5 – 35
UERRA[4] (lfpw)	Reg. Reanalysis	1979-2015	Precip, Tmax, Tmin	5.5 – 35

[1]Interpolated – Derived by interpolation of surface weather station data; Reanalysis – gridded data from a reanalysis dataset; Climate Projection – statistically downscaled and bias-corrected data from global climate model simulations (CMIP5); Satellite – gridded data derived from satellite-based observations; Reg. Reanalysis – gridded data from a regional climate model used to downscale a global reanalysis.

[2]Precip – daily total precipitation; Tmax – Daily maximum temperatures; Tmin – Daily minimum temperatures; Tmean – Daily mean temperatures; RH – Relative humidity.

[3]AgMERRA was created by bias-correcting climate variables relevant for agriculture as simulated by the older MERRA reanalysis (Rienecker et al., 2011) with surface observations. See Ruane et al. (2015) for details.

[4]RCMs used: eggr - Met Office Unified Model (UM)); eswi - HARMONIE; lfpw - MESCAN-SURFEX.

The second type of gridded datasets are reanalyses. Briefly, a reanalysis dataset is created by integrating a high resolution global weather forecast model forward in time, in intervals of 6-12 hours, and tightly constraining the model with a wide variety of observations. Reanalyses are designed to provide a best estimate of actual meteorological conditions over a long time period. Some reanalysis have been further downscaled using regional climate models.

The third type of gridded datasets are derived from satellite data, and often consist of a single meteorological variable, such as rainfall amounts and lightning flash rates. Some satellite datasets have very high spatial resolutions, of the order of a few km. However, some of these datasets only cover the tropics or lower latitudes, owing to the orbits of the satellites.

The gridded dataset of most interest for MED-GOLD is E-OBS (Haylock et al., 2008; Cornes et al., 2018). E-OBS[1] v17 provides the key variables of temperature and rainfall for Europe at 25 km resolution for 1950-2017. The first version of E-OBS was published over 10 years ago. The latest versions are based on many more stations (approximately 3 times more). However, the majority of the extra stations are located in Germany, Poland, and Scandinavia, which are far from the MED-GOLD study regions (Figure 5-1). A few extra stations have been used in E-OBS around the area of the Douro valley (Cornes et al., 2018). The numbers of weather stations was at a maximum during the 1990s, but has fallen since 2000, especially for precipitation. In some countries, only a subset of the available weather stations are made available to ECA&D and hence for use within E-OBS.

6.2. LIMITATIONS OF GRIDDED DATASETS

Gridded datasets are essential for assessing regional climates and evaluating simulations by climate models. They are often used to drive crop models and rainfall-runoff models in areas where high-quality surface observations are few in number or unavailable. For example, Ledesma and Futter (2017) compared rainfall-runoff simulations using gridded temperature and precipitation data from E-OBS and local observations. The simulations of runoff using gridded data were closer to the measured values than when the observed data were used in almost all cases. However, gridded datasets have some limitations.

As stated above, many gridded datasets are derived through interpolation of surface-based station data. The quality of the gridded data is dependent on the ability of the interpolation method to estimate values on the regular grid from the underlying station network (Hofstra et al., 2009). The interpolation algorithm will be loosely constrained in areas with few stations, leading to larger errors in the interpolated values. Some stations are located in or close to airports, and so will not be representative of agricultural areas.

Most weather stations record temperature and rainfall, but a smaller number record other additional variables such as solar radiation, humidity and wind speeds. These latter variables can vary considerably over small distances, meaning it is difficult to interpolate the measurements and estimate values at other locations. For meteorological variables with limited measurements, or those that are harder to interpolate, the gridded data are limited to monthly averages or totals over large spatial areas. Some gridded datasets are updated fairly frequently; e.g., E-OBS data for Europe are updated on an annual basis. However, others (e.g., AgMERRA) have not been updated since they were created, and so become increasingly out-of-date with time.

All of the variables required for crop modelling are available from reanalyses. However, the accuracy of the simulations some variables, such as humidity and solar radiation, is generally low (although newer reanalyses show significant improvement over older ones). Solar radiation from reanalyses is most accurate for clear skies but notably poorer under cloudy conditions (Huld et al., 2017). High resolution satellite-based solar radiation products are available, but are derived from geostationary satellites which only cover certain regions of the Earth.

To summarise, a range of different gridded datasets are available which will be of use within MED-GOLD. Most of these datasets do not contain all of the agriculturally-relevant variables. They often cover different (but overlapping) time periods and have a range of spatial resolutions. Combining data from these different gridded datasets is not straightforward. For example, E-OBS could contain warm temperatures for a given location, but the solar radiation data from another dataset might be low owing to high cloud cover, and so be more consistent with cooler temperatures. A thorough intercomparison of these gridded datasets is beyond the scope of MED-GOLD.

[1] It is noted that version 19 of E-OBS, released during March 2019, has a higher spatial resolution (about 12 km) and contains data up to the end of 2018. However, this newer version was not available in time for use within this deliverable. For compatibility reasons, the following deliverable (D1.4) will also use v17 of E-OBS.

7. COMPARISON OF SURFACE OBSERVATIONS WITH GRIDDED DATA

In this section, climatic variables calculated from weather station and gridded datasets are compared for the three study regions in Europe of MED-GOLD: Andalucía (Spain), Douro Valley (Portugal), and Foggia, Ancona and Ravenna (Italy). A similar analysis is performed for Colombia in section 7.5. The objective is to assess the gridded datasets, to see if they are suitable for use in development of the climate service tools and evaluation of climate models.

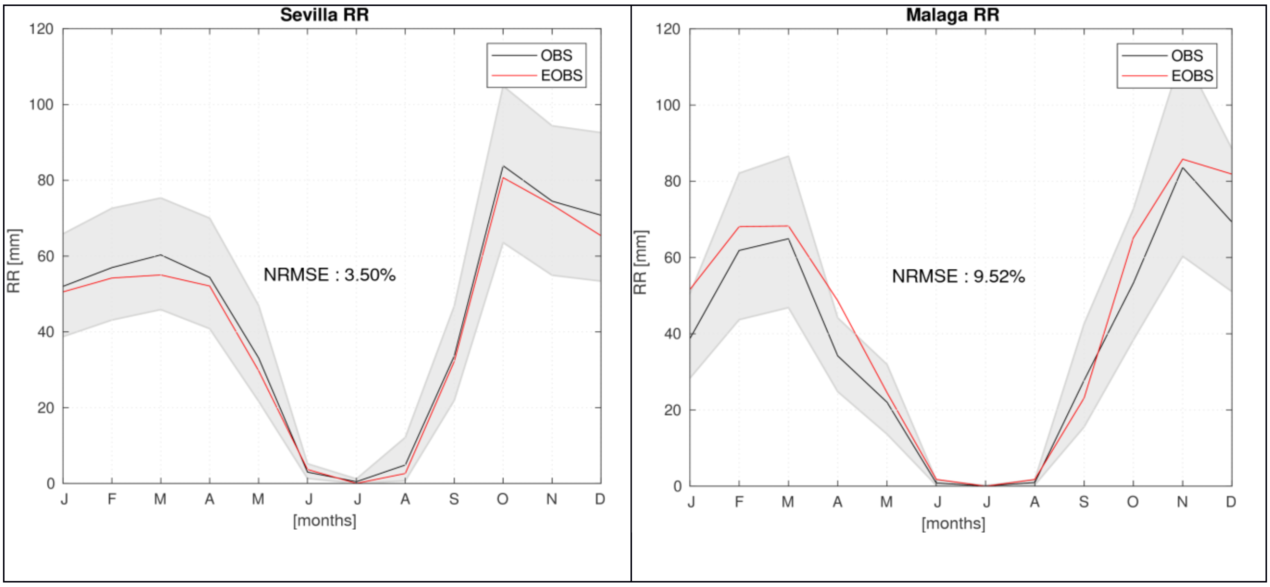
Care needs to be taken when comparing gridded data with observations from weather stations. Many gridded datasets, contain average or total values over an area, whereas weather station data are 'point data', valid for a specific location. Discrepancies in temperature between the two data types may be caused by differences in altitude between the weather station and the gridded data. Precipitation amounts in gridded data can be 'smeared out' over wider areas, when in reality the precipitation might have occurred over a small area. Some comparisons show that gridded and observed data agree reasonably well for low to moderate daily precipitation amounts, but diverge for very high amounts.

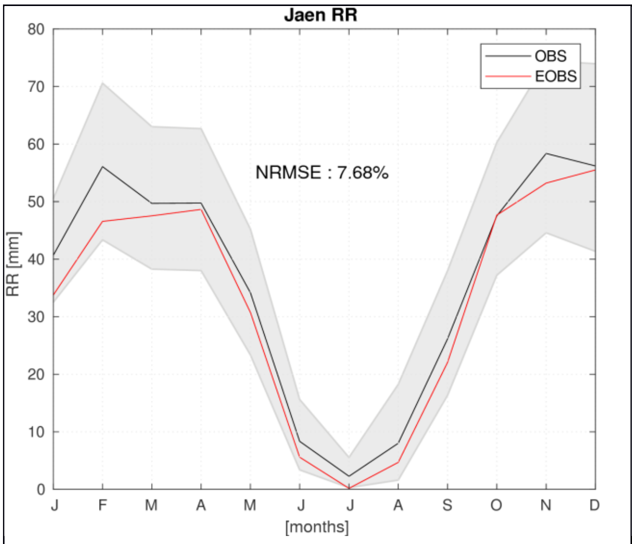
7.1. ASSESSMENT FOR ANDALUCÍA (SPAIN)

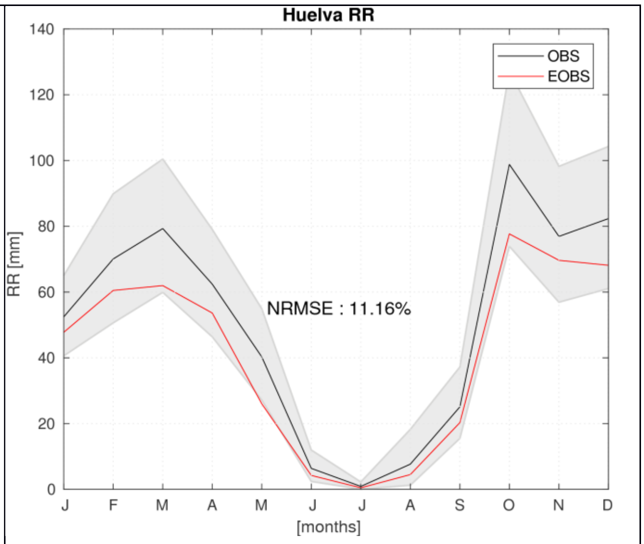
Olive groves in Andalucía, southern Spain, are a key study region for WP2 on olives and olive oil. This assessment focuses on the observations in Andalucía and a comparison with the E-OBS gridded dataset. The seven weather stations used are listed in Table 5-1.

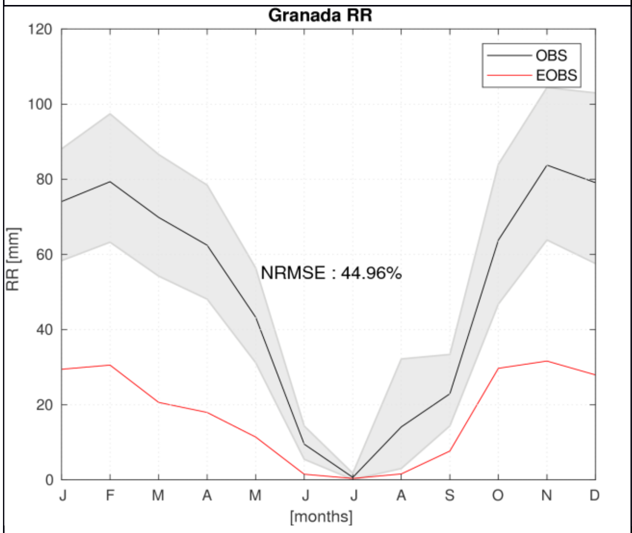
7.1.1. DAILY AND MONTHLY PRECIPITATION

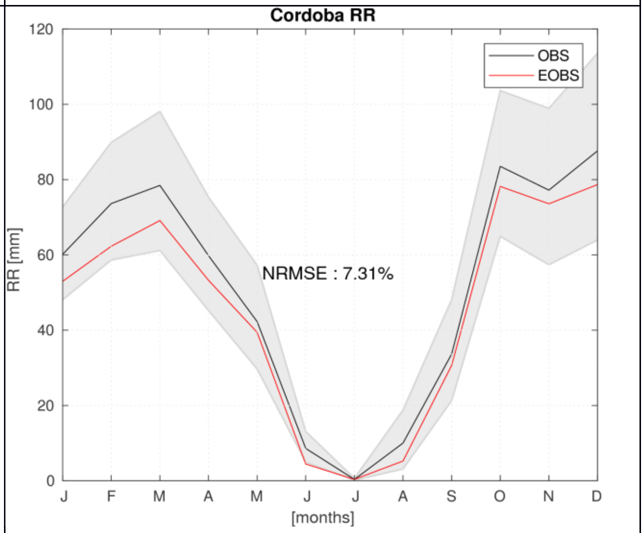
The mean annual cycle of mean monthly precipitation (RR) from the weather stations and E-OBS are shown in Figure 7-1. The E-OBS data reproduce satisfactorily the observed cycles at the majority of the stations. The only exception is the Granada station, where the precipitation in E-OBS is notably smaller. The normalised root mean square error (NRMSE) for all stations ranges between 3.5 and 11.2%, apart from Granada, where it is almost 45%. In addition, the E-OBS annual cycle lies within the observed 95th percentile confidence intervals for each station (except Granada) as obtained by bootstrapping.











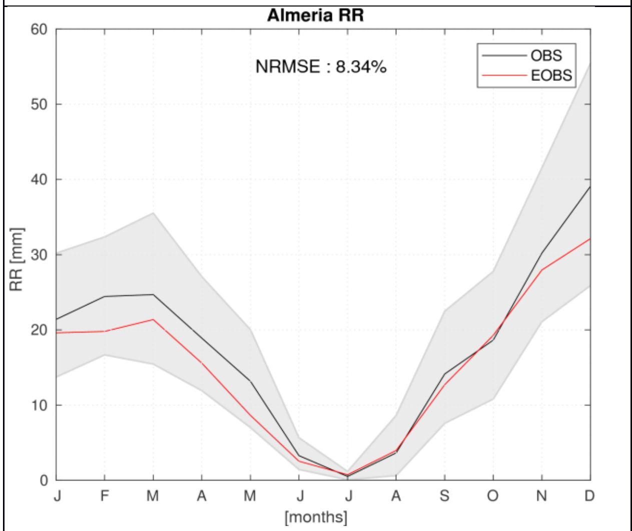
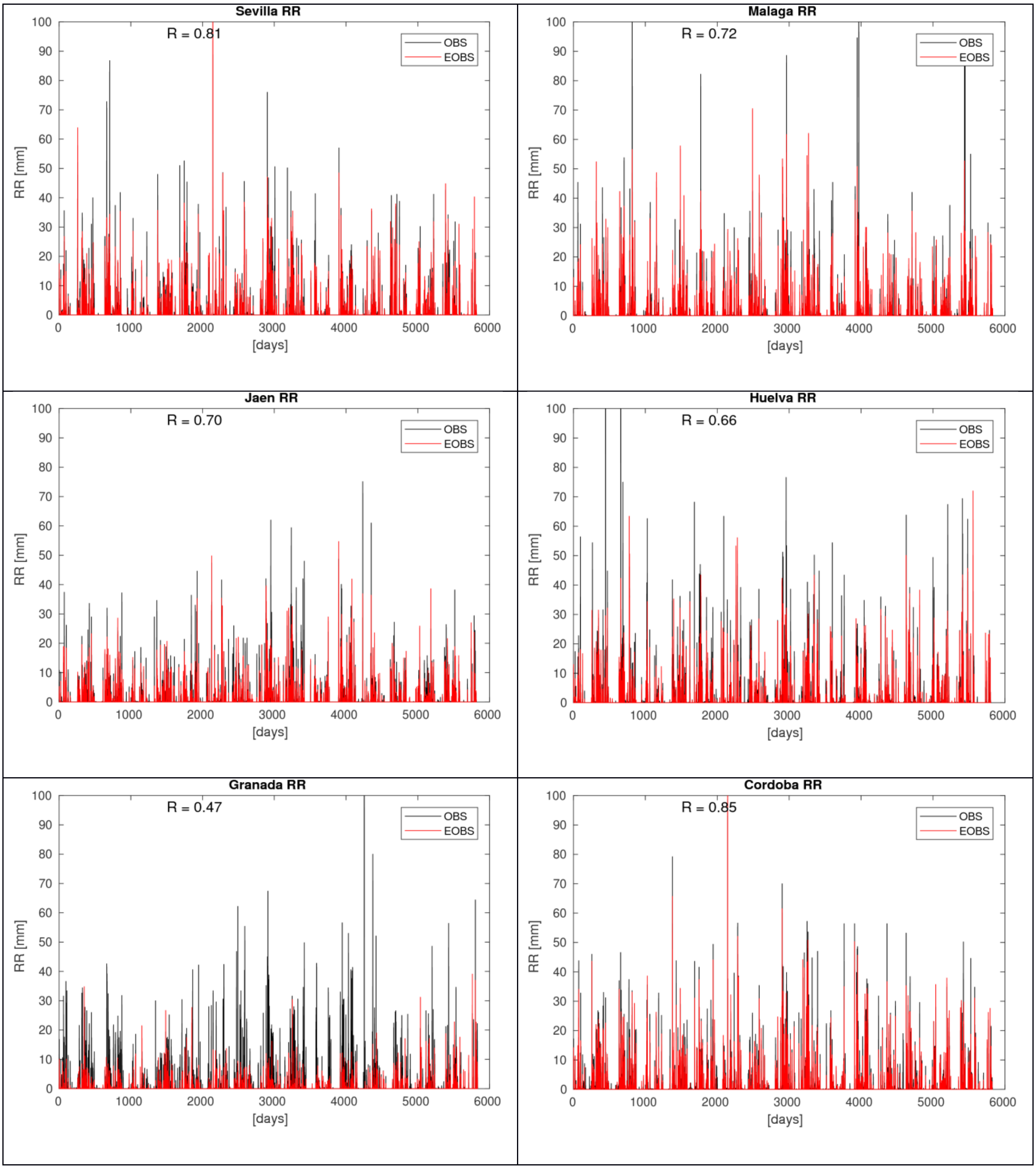


Figure 7-1: Annual cycle of total monthly precipitation for the period 2001-2017 from the weather stations (black curve) and the nearest point in the E-OBS dataset (red curve). The Normalized Root Mean Square Error (NRMSE) for each of the stations is indicated in each panel.

Time series of daily precipitation totals are shown in Figure 7-2 for each station and the nearest points in the E-OBS gridded data for 2001-2017. The rainfall totals in the two datasets are comparable, except for Granada, where E-OBS clearly underestimates the observed rainfall amounts. The correlation coefficients (R) between the weather station and E-OBS data range between 0.68 and 0.85 for all stations, apart from Granada, where the R value is only 0.48.



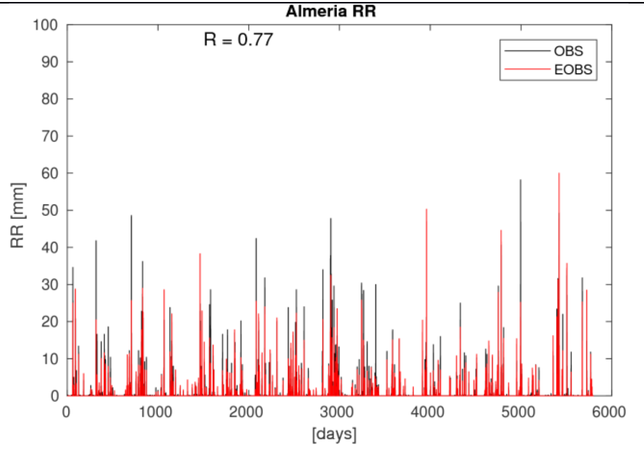


Figure 7-2: Time series of daily precipitation totals (RR) for the period 2001-2017 for the weather stations ("OBS"; black lines) and E-OBS data (red lines) for the stations in Andalucía. The linear correlation coefficient (R) between the two datasets is indicated at the top of each panel.

The average differences in the precipitation indices listed in Table 5-2 (relative or absolute differences, depending on the index) calculated from two datasets (E-OBS and the weather stations) are shown in Figure 7-3. Regarding the annual, winter and summer precipitation relative differences (top three panels), the highest deviations between the two datasets are found in Granada where E-OBS underestimates the observed precipitation by up to 70%. For the rest of stations smaller differences are found. More specifically, for the annual total precipitation the differences lie in the range -40 to +40 %. The lowest and the highest deviations from the observed precipitation totals are found in Granada and Almeria respectively. For the rest of the stations, the differences are smaller than 16%.

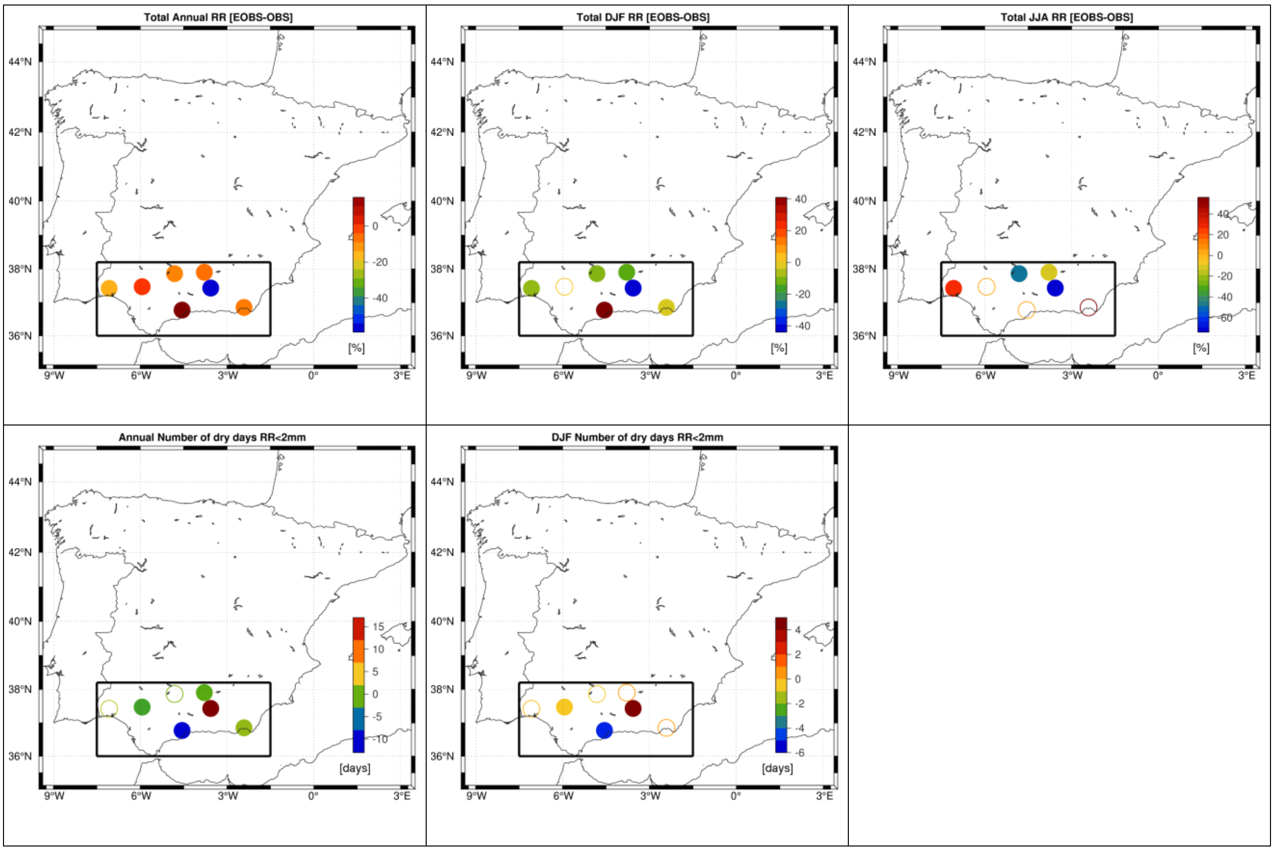


Figure 7-3: Average differences in the precipitation indices listed in Table 5-2 between E-OBS and observations at the seven selected stations in Andalucía. Positive values indicate the index has a larger value in E-OBS than the weather station. Statistically significant differences are shown with filled coloured circles, whereas open circles indicate differences which were not statistically significant.

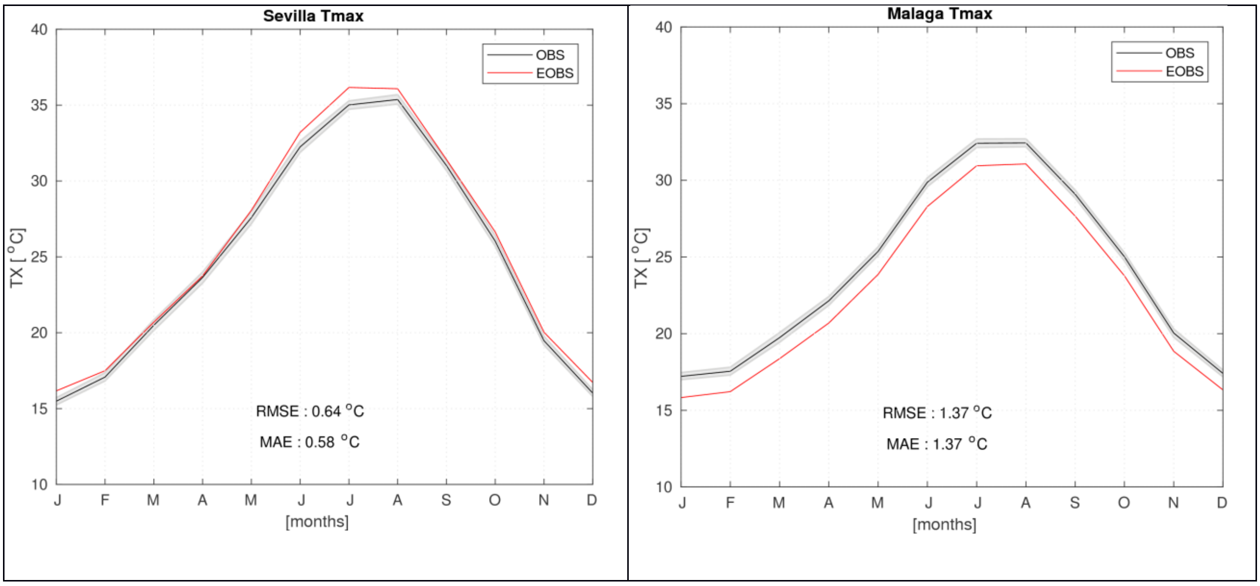
For the winter total precipitation totals, the largest differences (both negative and positive) are found in Almeria (~50%), Cordoba (~-51%) and Granada (~-70%). For Malaga and Seville no statistically significant changes are found whereas for the other two stations, Huelva and Jaen the differences are lower than 24%. For the summer total precipitation, the differences for all stations are lower than 17%, with the exception of Granada.

Differences in numbers of dry days (days with less than 2 mm of precipitation) annually and for winter are shown in the lower row of Figure 7-3. For the annual total number of dry days, the highest absolute differences are found in Granada and in Malaga, which are 18 and -12 days respectively. For the rest of the stations, the differences are less than 4 days, with non-statistically significant differences at Huelva and Cordoba. For the number of dry days during winter, the absolute differences for all stations are less than 5 days, indicating the E-OBS data are in good agreement with the observations. In four out of the seven stations, no statistically significant differences were found.

7.1.2. DAILY AND MONTHLY MAXIMUM TEMPERATURES

The annual cycle of mean monthly maximum temperatures for the seven stations and the closest E-OBS grid points are shown in Figure 7-4. It is evident that E-OBS captures the observed mean annual cycle at the majority of the stations. A significant deviation is only present at Granada (RMSE = 3.50°C), which is attributed to the mountainous characteristics of the area. It is known that the interpolation method used within E-OBS degrades with altitude (Haylock et al., 2008; Kotlarski et al., 2014 and references therein). For the rest of the stations the RMSEs range between 0.64°C and 0.88°C for Sevilla, Jaen, Huelva, Cordoba and Almeria, while for Malaga the RMSE is larger, at 1.37°C.

In Figure 7-5, time series of the maximum daily temperature records from the seven weather stations in Andalucía are presented and used for evaluation of the corresponding E-OBS data series. Overall, the daily maximum temperatures in E-OBS are in good agreement with the observations; the correlation coefficients between the two datasets is greater than 0.90 in all cases, reaching 0.99 in Sevilla, Jaen, Huelva and Cordoba. In Granada station, the E-OBS dataset seems to underestimate the Tmax during summer months and overestimate Tmax during winter, compared with the station observations.



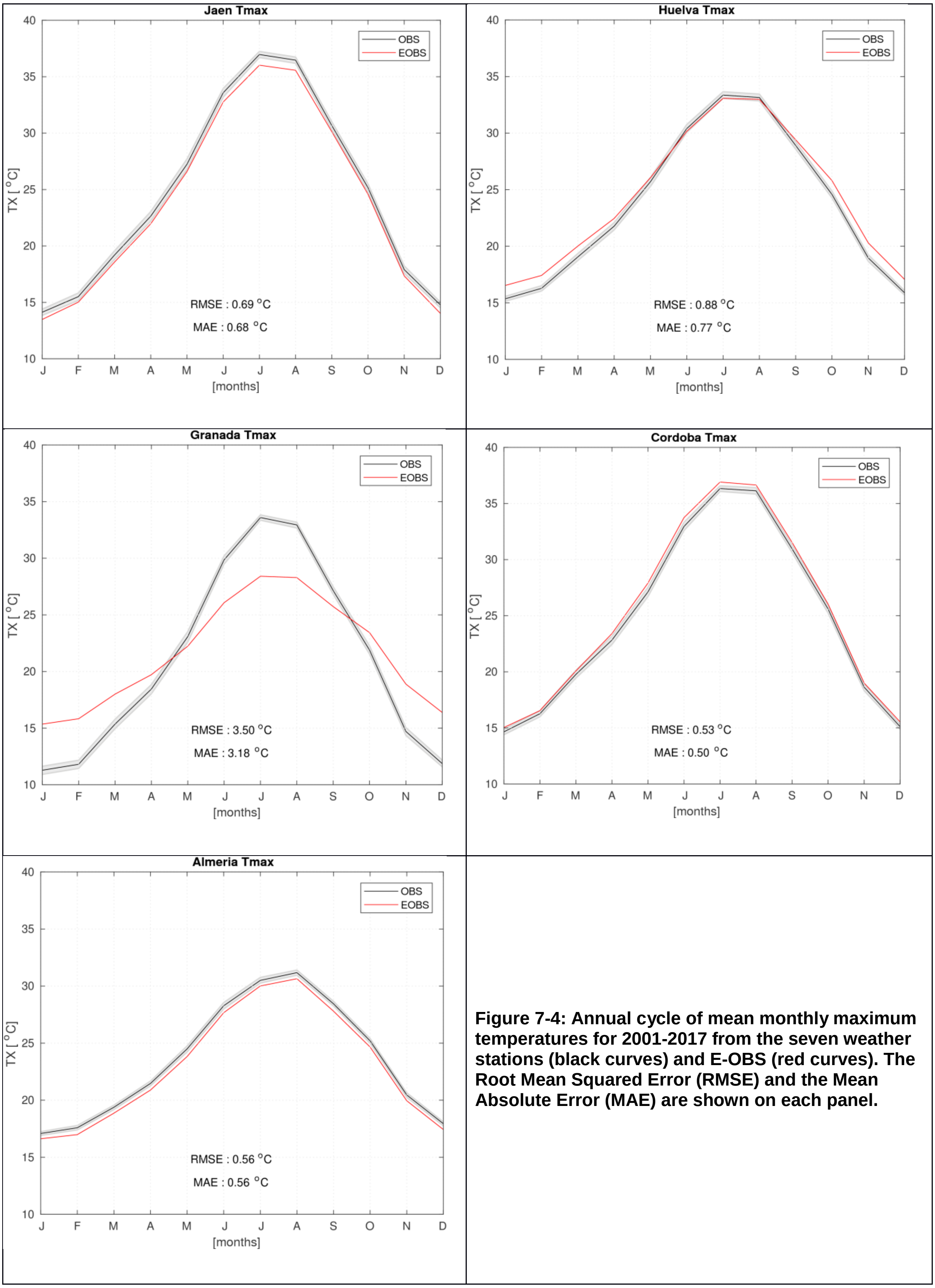
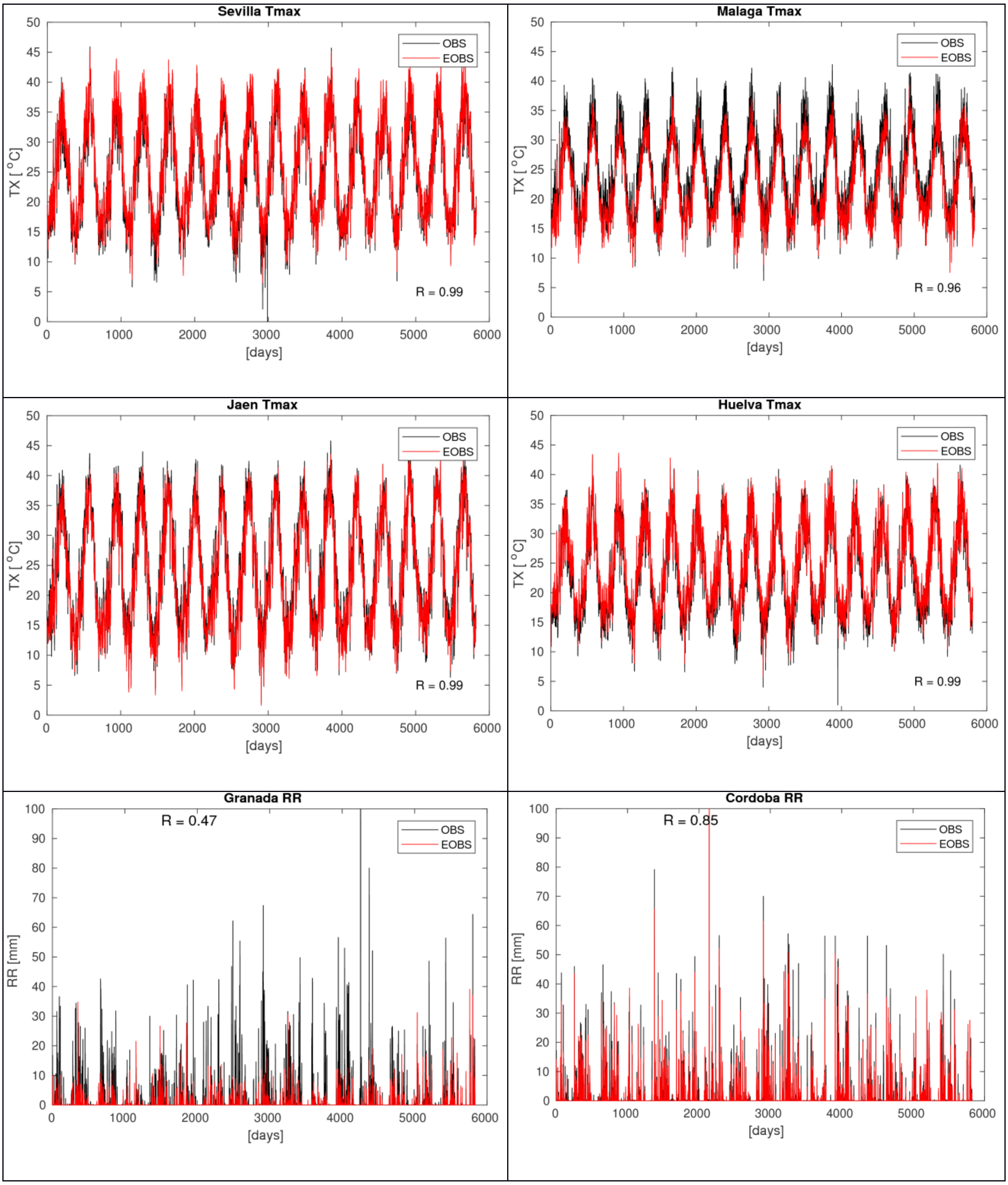


Figure 7-4: Annual cycle of mean monthly maximum temperatures for 2001-2017 from the seven weather stations (black curves) and E-OBS (red curves). The Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) are shown on each panel.


Sevilla Tmax
OBS
EOBS
R = 0.99
TX [°C]
[days]
Malaga Tmax
OBS
EOBS
R = 0.96
TX [°C]
[days]
Jaen Tmax
OBS
EOBS
R = 0.99
TX [°C]
[days]
Huelva Tmax
OBS
EOBS
R = 0.99
TX [°C]
[days]
Granada RR
R = 0.47
OBS
EOBS
RR [mm]
[days]
Cordoba RR
R = 0.85
OBS
EOBS
RR [mm]
[days]

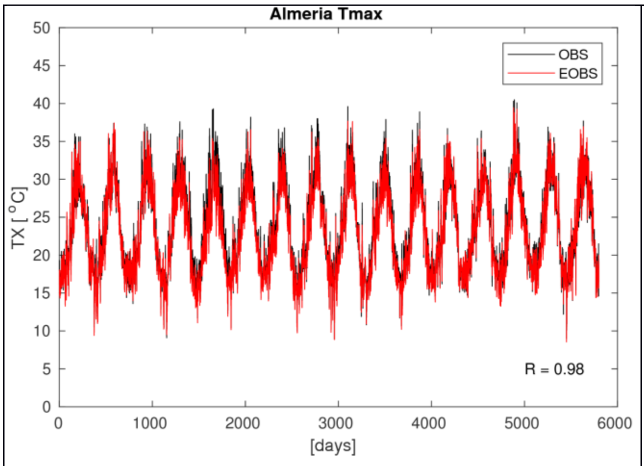
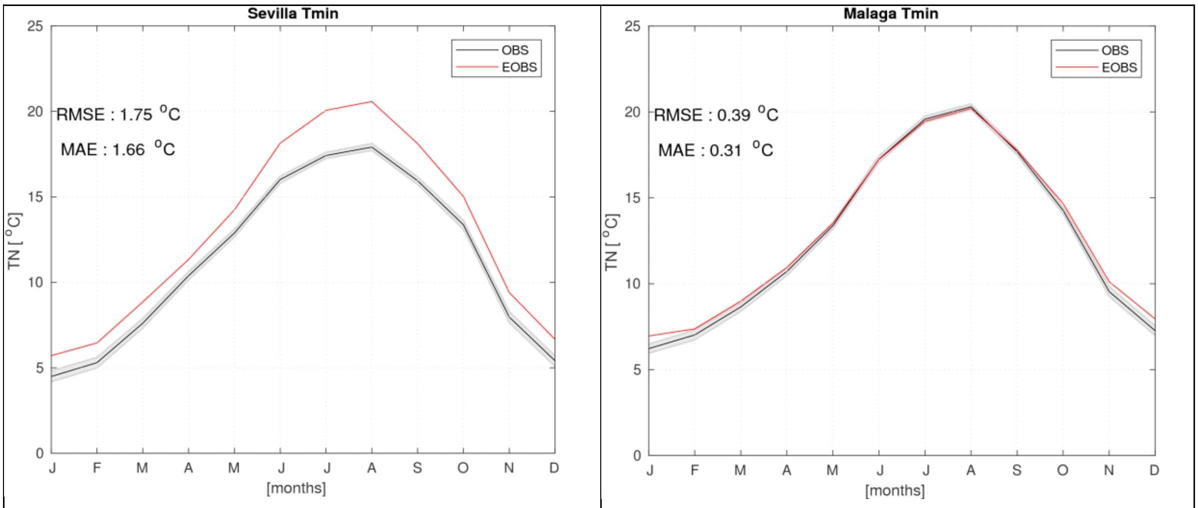


Figure 7-5: Time series of daily maximum temperatures (TX) for the period 2001-2017. Observed vales from the weather stations are shown in black, and E-OBS data are indicated in red. The linear correlation coefficient (R) between the two datasets for each station is shown on each panel.

7.1.3. DAILY AND MONTHLY MINIMUM TEMPERATURES

An assessment was also made of daily minimum temperatures from the seven weather stations and E-OBS. In Figure 7-6 the mean annual cycle of the mean monthly minimum temperature for the seven stations and the closest E-OBS grid point to each station is shown. The results indicate that the E-OBS datasets give a consistent demonstration of the mean annual cycle compared with observations. The expected summer maximum and winter minimum is exhibited in all cases with the E-OBS minima and maxima levels being in accordance with the observation levels at four of the stations. These stations are Malaga, Huelva, Cordoba and Almeria, where the RMSE values are 0.39°C, 0.64°C, 0.50°C and 0.15°C respectively. In the remaining three stations, E-OBS overestimate the observed minimum temperatures, with RMSE values ranging from 1.75°C in Seville to 4.70°C in Granada. The two error metrics, RMSE and MAE, have similar values at all stations, which indicates the errors (the differences between E-OBS and the station data) at each station have similar values.

Time series of daily minimum temperatures are shown in Figure 7-7. These comparisons confirm that the daily minimum temperatures at each station and the corresponding nearest points in E-OBS are similar. The linear correlations coefficients are 0.95 or larger, except for Granada, where E-OBS underestimates the magnitude of the annual cycle.



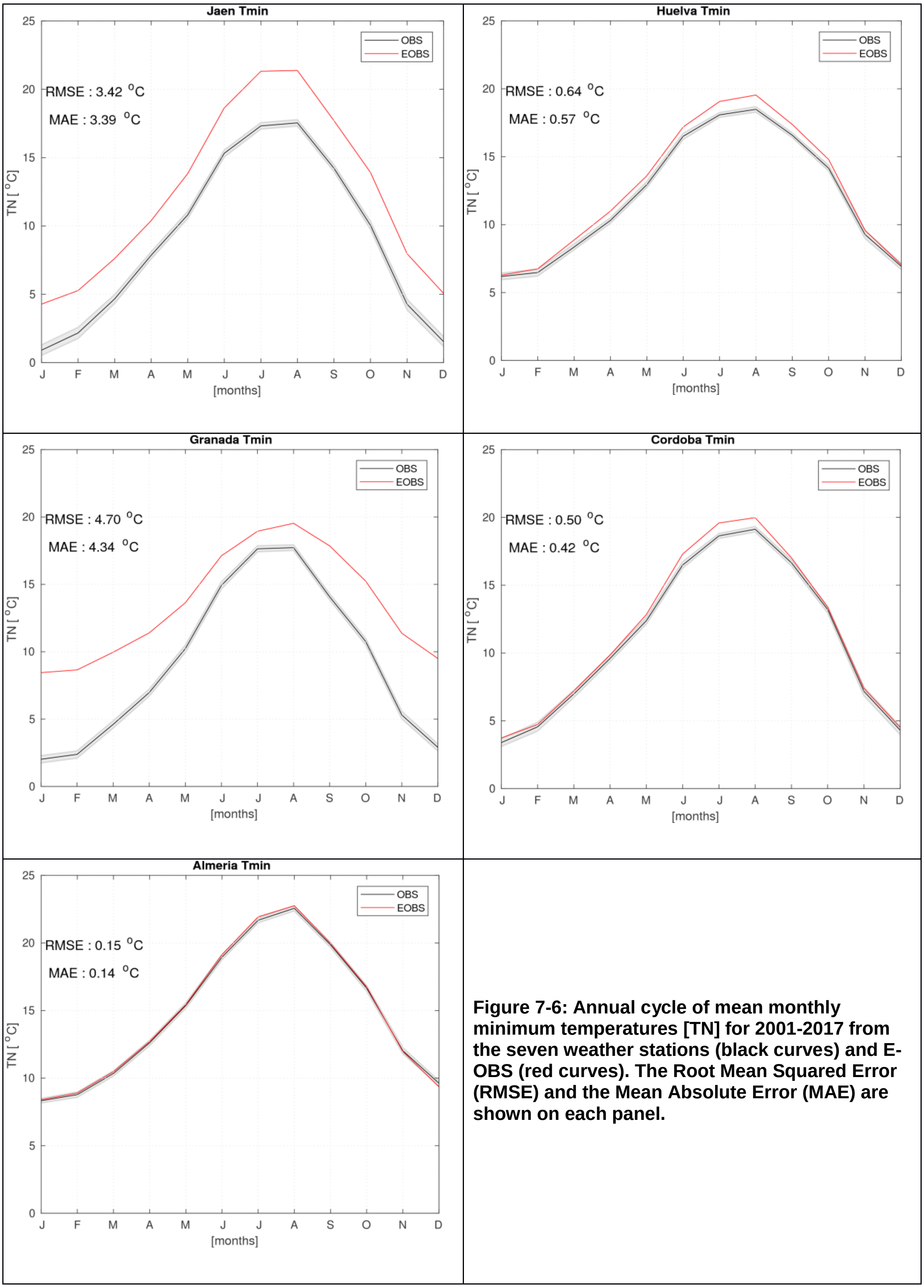
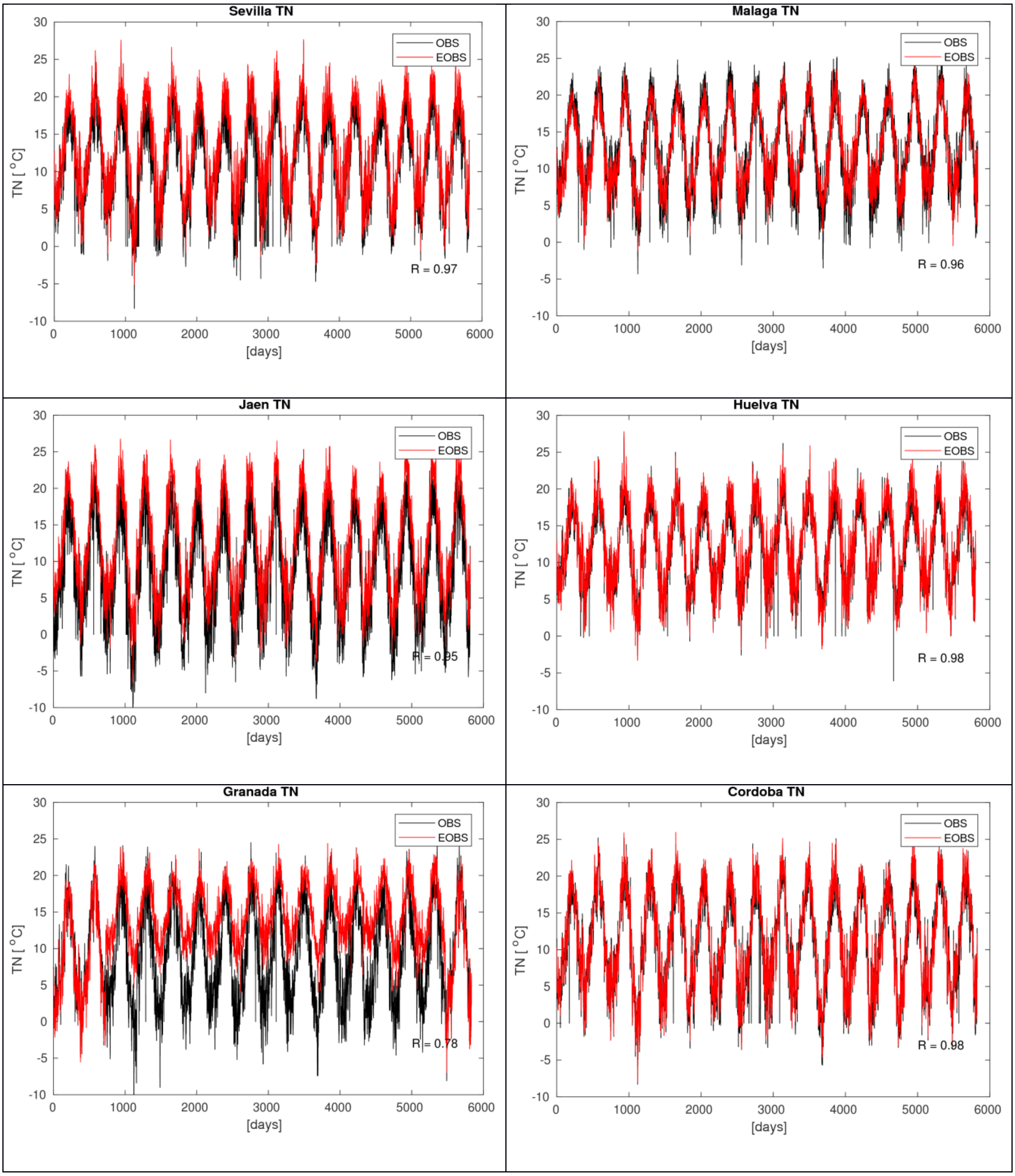


Figure 7-6: Annual cycle of mean monthly minimum temperatures [TN] for 2001-2017 from the seven weather stations (black curves) and E-OBS (red curves). The Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) are shown on each panel.


Sevilla TN
R = 0.97
Malaga TN
R = 0.96
Jaen TN
R = 0.95
Huelva TN
R = 0.98
Granada TN
R = 0.78
Cordoba TN
R = 0.98
OBS
EOBS
TN [°C]
[days]

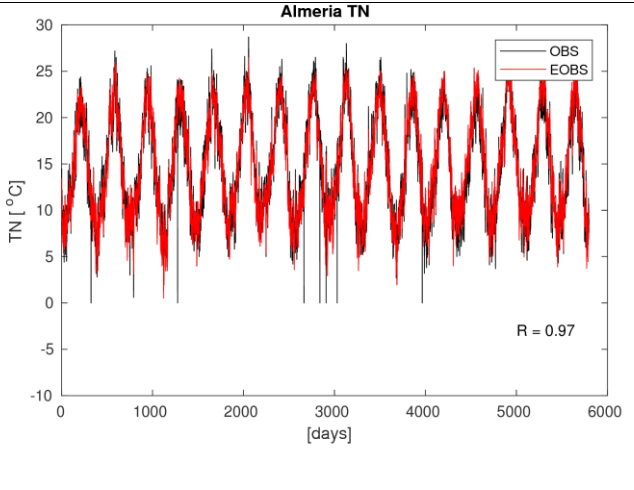
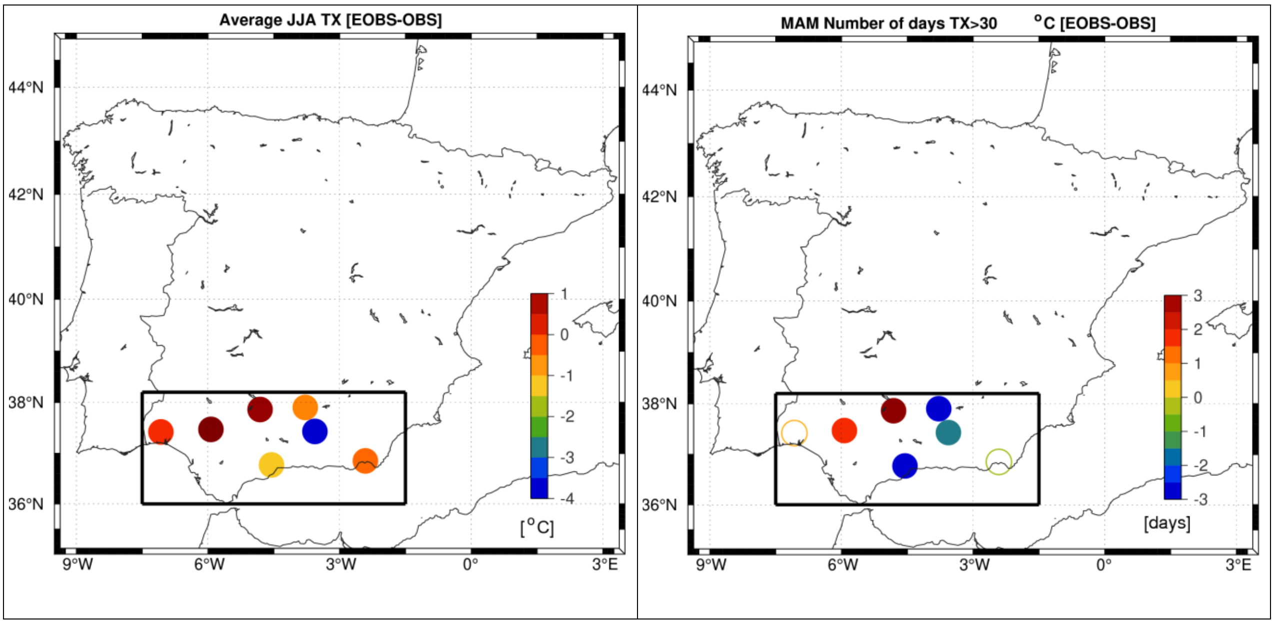


Figure 7-7: Time series of daily minimum temperatures [TN] for 2001-2017 from the seven weather stations (black lines) and E-OBS (red lines). The linear correlation coefficients (R) are shown on each panel.

7.1.4. DERIVED TEMPERATURE INDICES

In this section, four different temperature indices (Table 5-2) derived from both TX and TN are calculated at each station and from E-OBS. The average differences (calculated as E-OBS - OBS) between the indices at each station are shown in Figure 7-8.The statistical significance of these differences is examined using the 95th percentile bootstrap confidence intervals. The differences in mean summer temperature ('Average JJA') lie in the range -1.5°C to 1°C. Only in Granada does E-OBS underestimate mean summer maximum temperatures by 4°C. This large difference is attributed to the same reason mentioned in the previous sections. For the average number of days with TX > 30°C in spring, the differences are in the range of -3 to +3 days. E-OBS underestimates the numbers of days in Malaga, Jaen and Cordoba and overestimates them in Sevilla and Cordoba. No statistically significant differences were found for Almeria and Huelva. The average number of days with TX > 40°C in summer in Almeria, Huelva and Granada were very similar in both datasets, and the differences were not statistically significant. For the remainder of the stations, the differences range from -4 to +4 days. Finally, for the average winter minimum temperatures, no statistically significant changes were found in Almeria and Huelva. The highest difference was calculated for Granada, where E-OBS overestimated winter minimum temperatures by about 6°C. For the rest of the stations, the differences were lower, with values of about 3.5°C. As before, the large difference at Granada may be partly due to the high altitude of this station (which is not well represented within E-OBS, thereby introducing a bias in the temperatures), or sparse coverage of weather stations in the area.



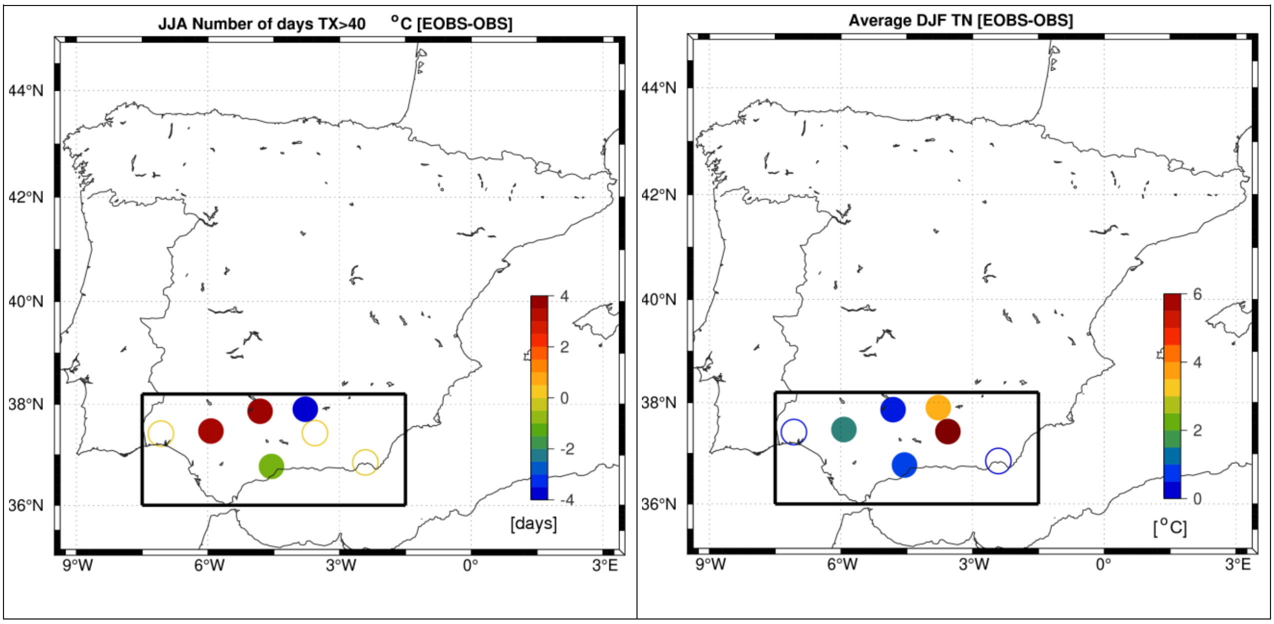


Figure 7-8: Average differences in the temperature indices in Table 5-2 calculated with E-OBS and the observations for the selected stations in Andalucía. Statistically significant differences are shown by filled circles whereas the non-significant differences are indicated by open circles.

7.1.5. SUMMARY

A comparison between data from seven weather stations in Andalucía and gridded data from E-OBS has shown the two datasets to be in satisfactory agreement. However, the comparisons for the station at Granada were notably worse than for the other stations. It is noted that the station at Granada has the highest altitude (939 m). One reason for the disagreement may be the average altitude of the corresponding E-OBS grid square is quite different. The coverage of weather stations around Granada that were used in the creation of E-OBS may be sparse, so that there were fewer data to constrain the interpolation algorithm.

7.2. ASSESSMENT IN THE DOURO VALLEY (PORTUGAL)

The Douro valley is a key study area for WP3, whose focus is vines and wine grapes. An overall assessment of the quality of gridded datasets is reported, by comparing them with data from weather stations operated by SOGRAPE and IPMA in the Douro Valley. These stations are listed in Table 5-3, and their locations are shown in Figure 5-1 by the solid red circles. The focus is on the variables of main interest for the wine service, namely precipitation and surface air temperatures. Additionally, climatic indices related to growth of the vines, grape yields and dates of harvesting (Fontes et al., 2016) are also calculated and compared.

For each variable in the gridded datasets, the average value over the box which covers the Douro valley is used (8°W - 7°W, 41°N - 41.5°N), owing to the close proximity of the weather stations in this area. For gridded datasets with lower resolutions a larger box was used (8.5°W - 6.5°W, 40.5°N - 42°N). For the sake of simplicity, both types of gridded data (i.e. interpolated station data and reanalysis data) will henceforth be referred to as “gridded data”.

In some of the figures in this section, correlations between the gridded data and the weather station observations at Seixo are shown. This station was chosen as an example. Correlations between the gridded data and remaining stations were very similar (data not shown).

7.2.1. SPRING PRECIPITATION

Spring precipitation is important for the growth of the vines and occurrence of fungal diseases. A dry spring would delay growth of the vines but reduce the chance of fungal diseases, whereas a wet spring could promote growth and increase the chance of fungal disease occurring. Spring total precipitation from the different gridded datasets over the Douro valley is shown in Figure 7-9. The shaded areas indicate the maximum and minimum spring precipitation totals from the SOGRAPE and IPMA weather stations (the mean value is represented by the thick pink line). The large spread of the weather station data reflects the different locations of the stations and the complex orography of the area.

An overall reasonable agreement between the different sources of information can be recognized. Generally, data from E-OBS agrees well with the weather station data. The gridded datasets have a clear positive bias with respect to E-OBS between 1979 and 2005, and NCEP2 has the largest bias. After 2006, the agreement between the different gridded datasets is much better. The wet springs of 1988, 1993, 1994 and 1997 are evident in all of the datasets and the long term behaviour is generally fairly similar.

The root mean square errors (RMSE) of the gridded datasets against the weather station data for spring precipitation over 1981-2017 (2011-2017 for the SOGRAPE stations) are reported in Figure 7-10. The smallest RMSE values are seen for E-OBS and subsequently for UERRA_egrr and ERA5. The agreement is poorest between the stations at Porto, Sairrão 3 and Seixo, and the gridded datasets.

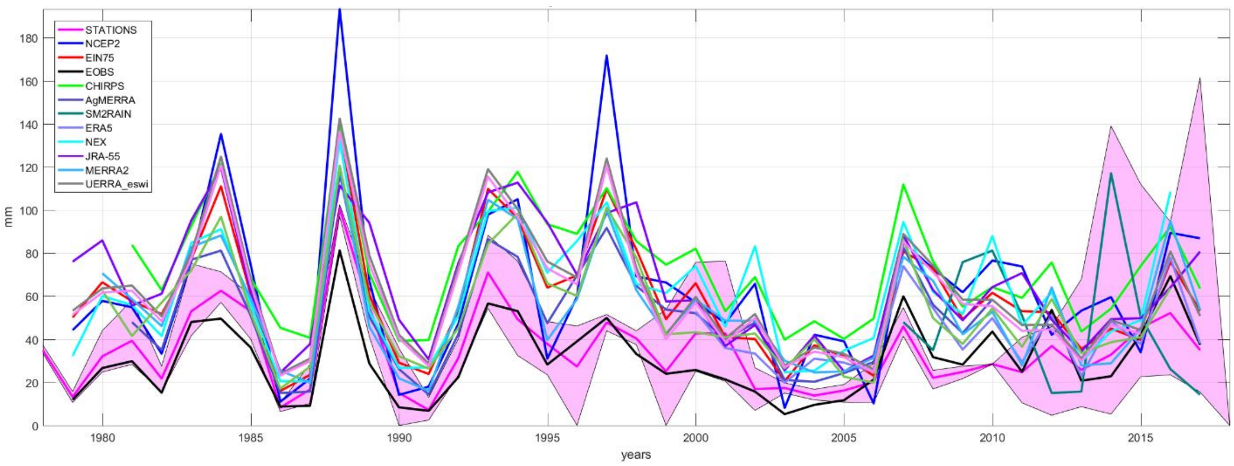


Figure 7-9: Spring total precipitation in the Douro valley area from different gridded datasets. The shaded area and thick pink line indicate the range of values from the SOGRAPE+IPMA weather stations, and the mean value from all of the stations respectively.

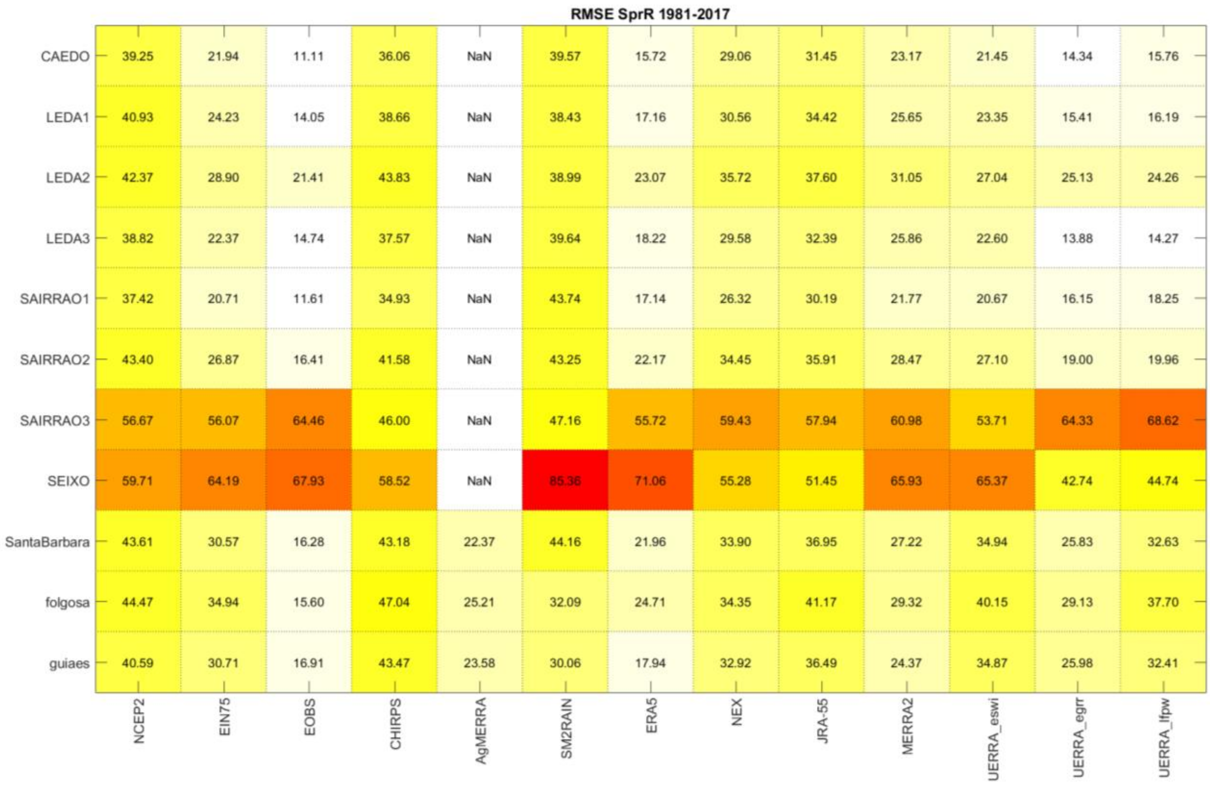


Figure 7-10: Root Mean Square Error (RMSE) of spring precipitation from the gridded datasets against the SOGRAPE weather stations for the overlapping period 2011-2017.

Linear correlation coefficients between total spring rainfall calculated from the three IPMA stations and each of the gridded datasets are shown in Figure 7-11. The correlations are reasonable in most cases, having values of 0.70 or larger. However, poor agreement is seen between the SM2RAIN dataset and all three stations, and at Folgosa for ERA5, as highlighted in red in Figure 7-11. The best agreement is seen for rainfall recorded at Guiaes, where the correlations are greater than 0.80 for several of the gridded datasets, including E-OBS.

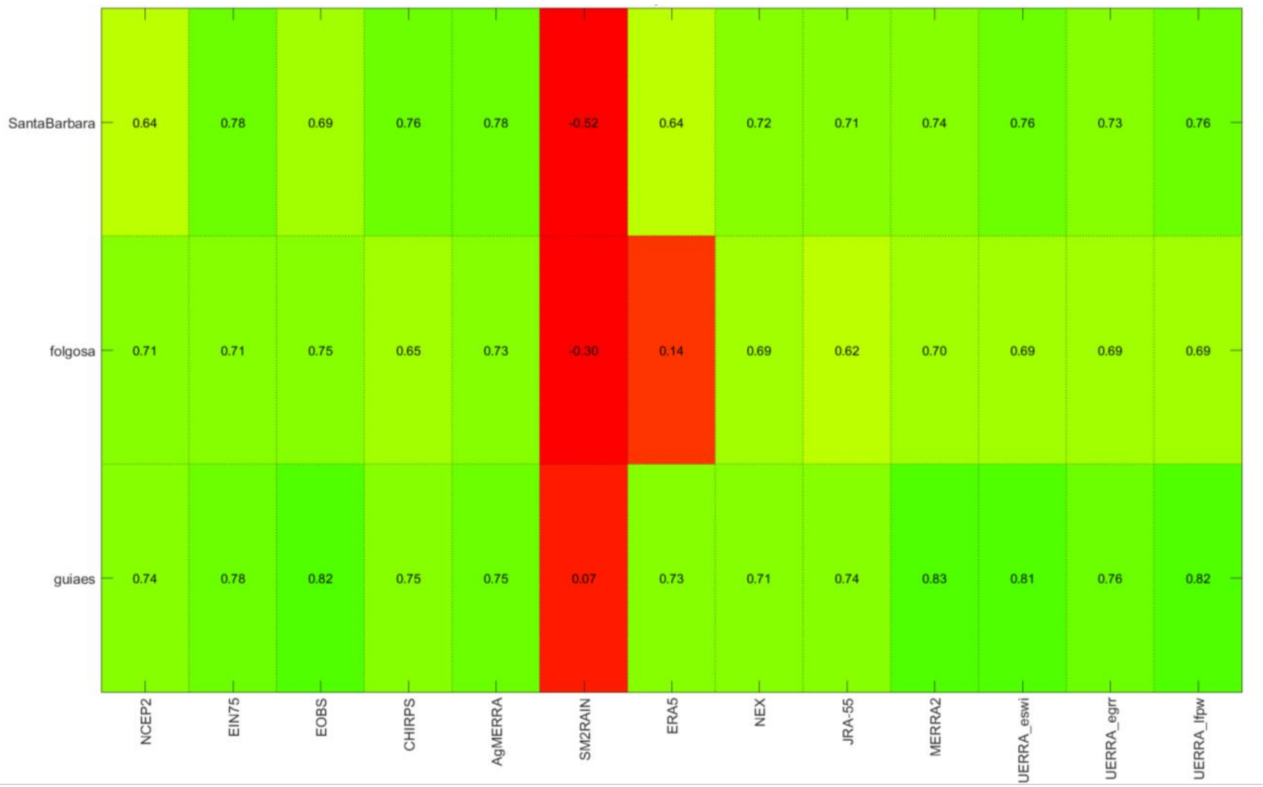


Figure 7-11: Linear correlations of spring precipitation between the gridded datasets (named on the x-axis) and the IPMA weather stations for the overlapping period 1981-2017.

7.2.2. SUMMER PRECIPITATION

Total summer precipitation calculated from the various gridded datasets (Table 6-1) is shown in Figure 7-12 for the years 1979 to 2017. The shaded area indicates the maximum and minimum rainfall amounts calculated from the SOGRAPE weather stations, and the mean value is represented by the thick purple line). The large spread in the weather station data reflects their different locations and the complex orography of the Douro valley area. Overall, a good agreement is seen between the different gridded datasets and the surface weather stations. All of the gridded datasets have a positive bias with respect to E-OBS data. The wetter summers in the second part of the 1980s and 1990s, and the local maximum during the second part of the 2000s are also well represented. All of the gridded datasets suggest there has been a small but steady decline in summer rainfall in the Douro valley region.

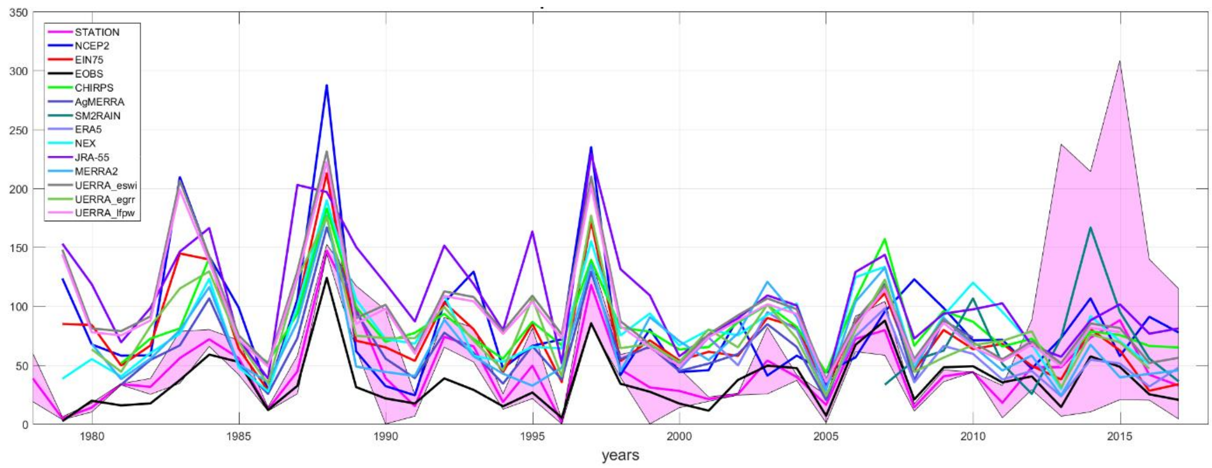


Figure 7-12: Summer (JJA) total precipitation over the Douro valley area from different gridded datasets. The shaded areas indicates the range of values from the SOGRAPE and IPMA weather stations, and the mean value across all stations is shown by the thick pink line.

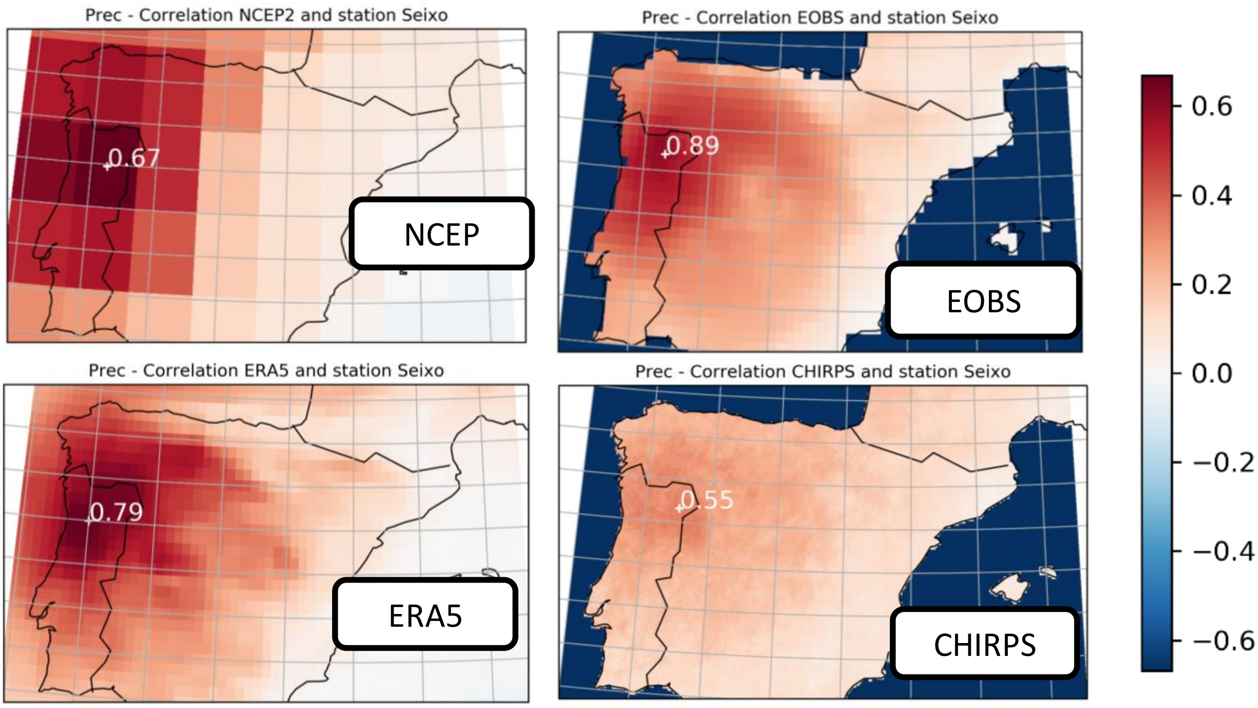


Figure 7-13: Temporal correlation coefficients of precipitation from four different gridded datasets with the SOGRAPE weather station at Seixo (Table 5-3).

The temporal correlations of summer precipitation as represented in the different gridded datasets against one of the SOGRAPE weather stations (Seixo) are reported in Figure 7-13. Correlations were calculated for the other stations (not shown) and were found to be very similar. The highest correlation is seen for the E-OBS dataset, and a high correlation is also found for ERA5. The root mean square error (RMSE) of the gridded datasets against the weather station data for summer precipitation in the common period 2011-2017 are reported in Figure 7-14. The smallest RMSE values (indicating the closest agreement) can be found for E-OBS and subsequently for ERA5 and MERRA2. All of the datasets have a large bias at stations Sairrão 3 and Seixo, as indicated by the relatively large RMSE values.

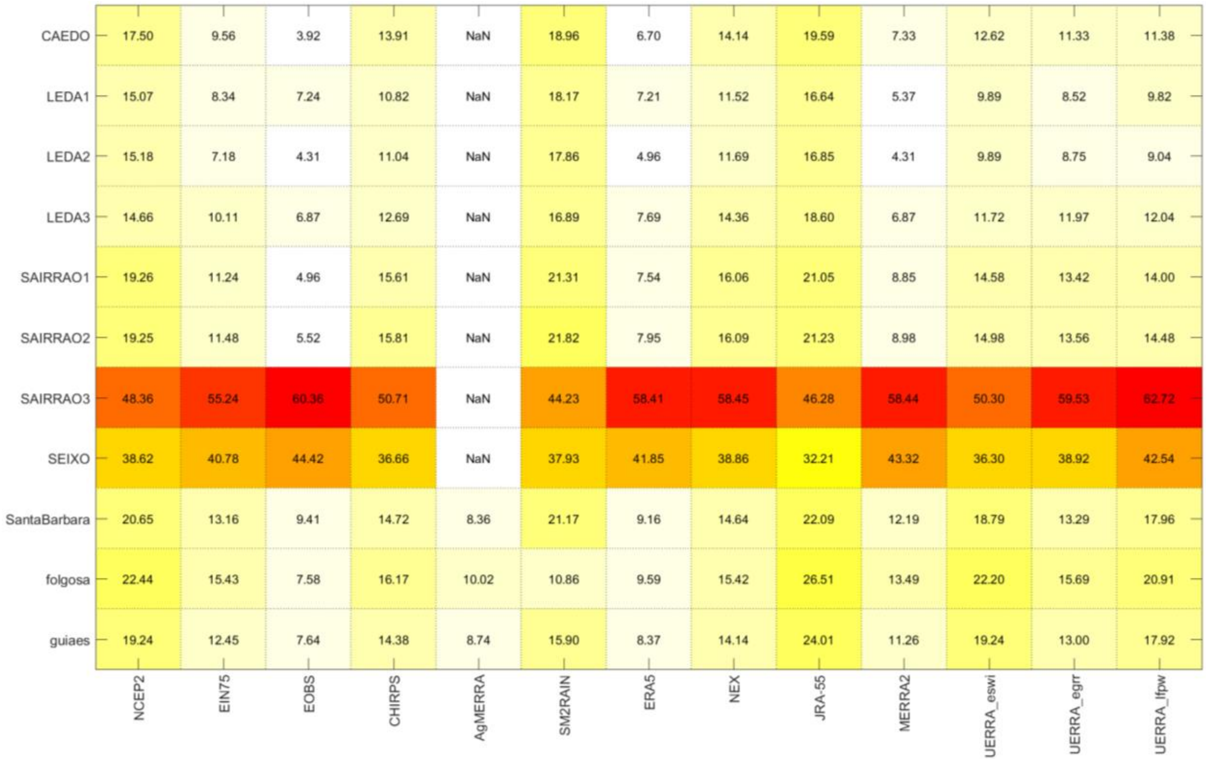


Figure 7-14: Root Mean Square Error (RMSE) of summer precipitation in the gridded datasets (x-axis) against the weather stations (ordinate), using data for 2011-2017 (SOGRAPE stations) and 1981-2017 (IPMA stations). The shading emphasises the accuracy of each dataset; white and pale yellow colours indicate the closest agreement, whereas dark yellow, orange and red indicate progressively poorer agreement.

7.2.3. MAXIMUM TEMPERATURES

A similar analysis was undertaken for surface air daily maximum temperatures (henceforth “Tmax”). Once again, average values over the area 8°W - 7°W; 41°N - 41.5°N were calculated (for gridded datasets with lower spatial resolutions, a larger area 8.5°W - 6.5°W; 40.5°N - 42.0°N was used). These areas cover the Douro valley. First, mean summer maximum temperatures from the gridded datasets and weather stations were compared. Two climate indices, the number of summer days with Tmax > 25°C (SU) and percentage of summer days with Tmax < 10th percentile (TX10p) were also calculated and compared. Finally, several additional indices of relevance to choice of grape varieties and grape quality are calculated.

Time series of summer mean daily maximum temperatures from the gridded datasets are shown in Figure 7-15 (upper panel). The shaded areas indicates the range of mean values from the SOGRAPE weather stations. A reasonable agreement is seen between the gridded datasets and the weather station data. A systematic cold bias is apparent between E-OBS and the remainder of the gridded datasets (with the partial exception of MERRA-2). This cold bias might be a result of the differences in resolution; E-OBS has a resolution of 25 km whereas most of the other gridded datasets have coarser resolutions, between 25 km and 250 km (Table 6-1). A large spread exists between the Tmax values between the SOGRAPE stations, which will be partly caused by the range of altitudes (from 160 m to over 600 m). If the long term mean of each dataset is subtracted from the individual values for each summer, a notably closer agreement is seen between the resulting anomalies for almost all of the gridded datasets (Figure 7-15(b)).

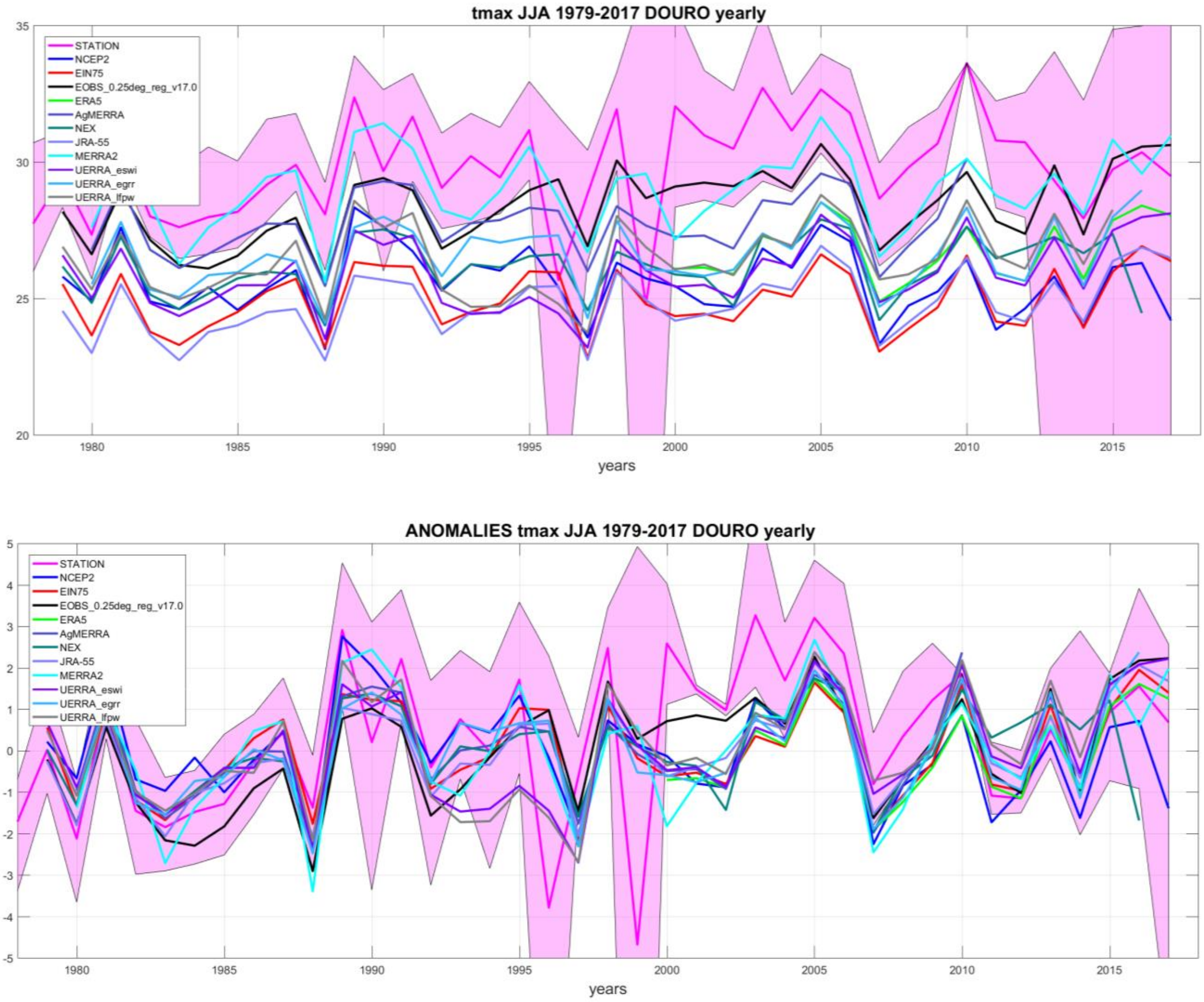
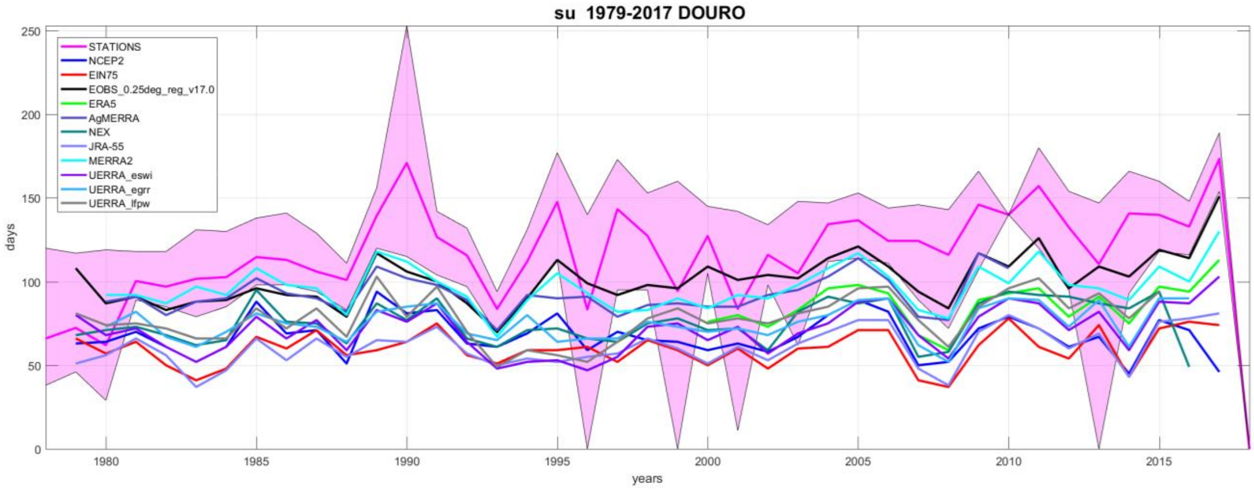


Figure 7-15: Upper panel, Time series of mean summer Tmax (°C) from the gridded datasets. Lower panel, Temperature anomalies with respect to the long-term average values. In both panels, the shaded areas show the range of values from the SOGRAPE and IPMA weather stations.

A reasonable agreement is found between the climatic indices SU (summer days) and TX10p (number of summer days with Tmax below the 10th percentile) as shown in Figure 7-16. However, fewer summer days are calculated from the gridded datasets than from the weather stations. A steady increase in the numbers of summer days can be seen in some of the gridded datasets (e.g., E-OBS, ERA5). In the other datasets the trends are smaller, or no clear increase is seen. The daily maximum temperatures in E-OBS tend to be higher than those in the other gridded datasets, which leads to a higher number of summer days. For the same reason, values of Tx10p in E-OBS are smaller than in the other datasets, especially between 1995 and 2010 (Figure 7-16, lower panel). A small downward trend in TX10p derived from E-OBS can be seen, but no trend is apparent in many of the other datasets.



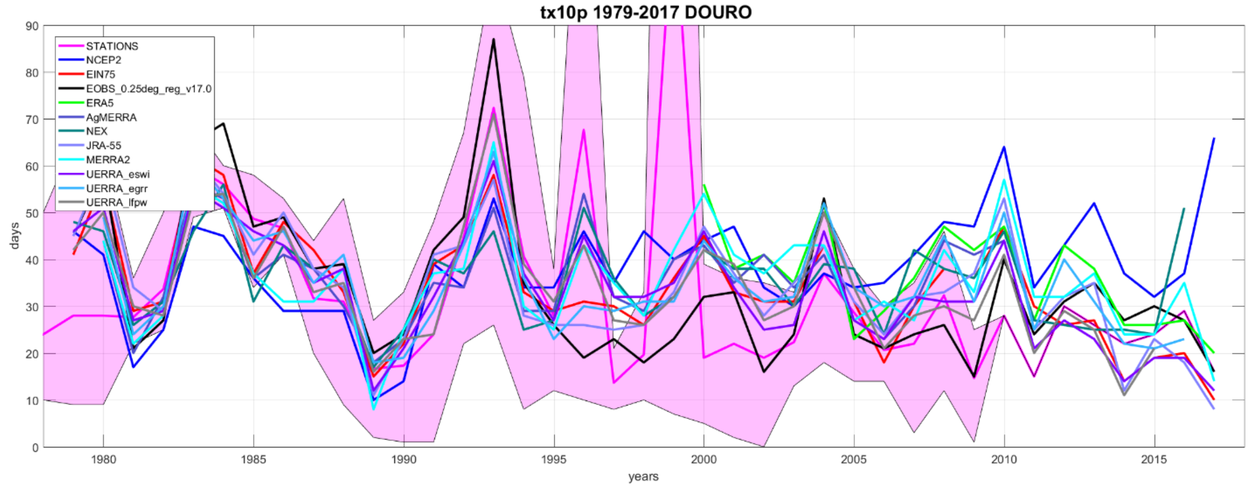


Figure 7-16: Climate indices based on daily Tmax values for the Douro Valley region. Upper panel: Numbers of summer days (days with Tmax > 35°C); Lower panel: Tx10p (percentage of days with Tmax below the 10th percentile). In both panels, the shaded areas show the range of values from the SOGRAPE and IPMA weather stations.

Root mean square errors (RMSEs) of the gridded datasets against the weather station data for numbers of summer days are reported in Figure 7-17. The smallest RMSE values (indicating the closest agreements) can be found for E-OBS and MERRA2. The RMSE values for AgMERRA with the three IPMA stations are similar to those for E=OBS and MERRA2. The poorest agreements are seen with the NCEP2, EIN75, JRA55 and UERRA_eswi datasets.

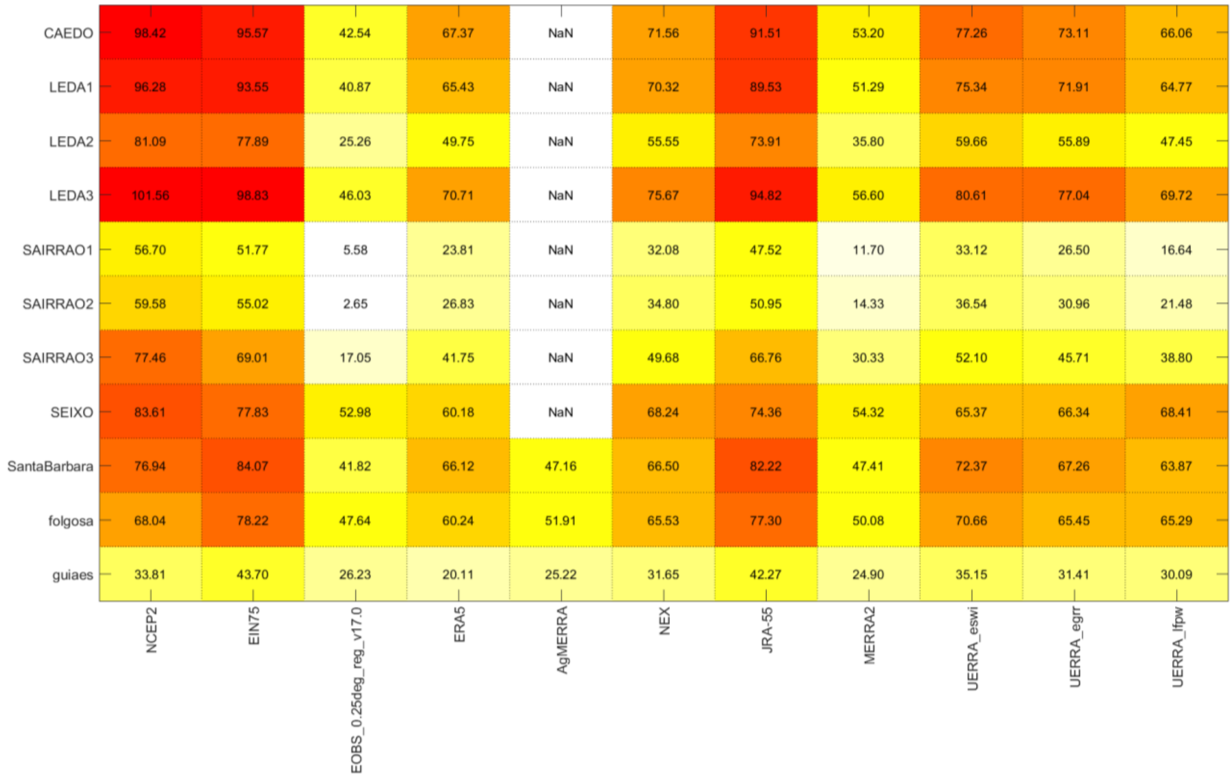
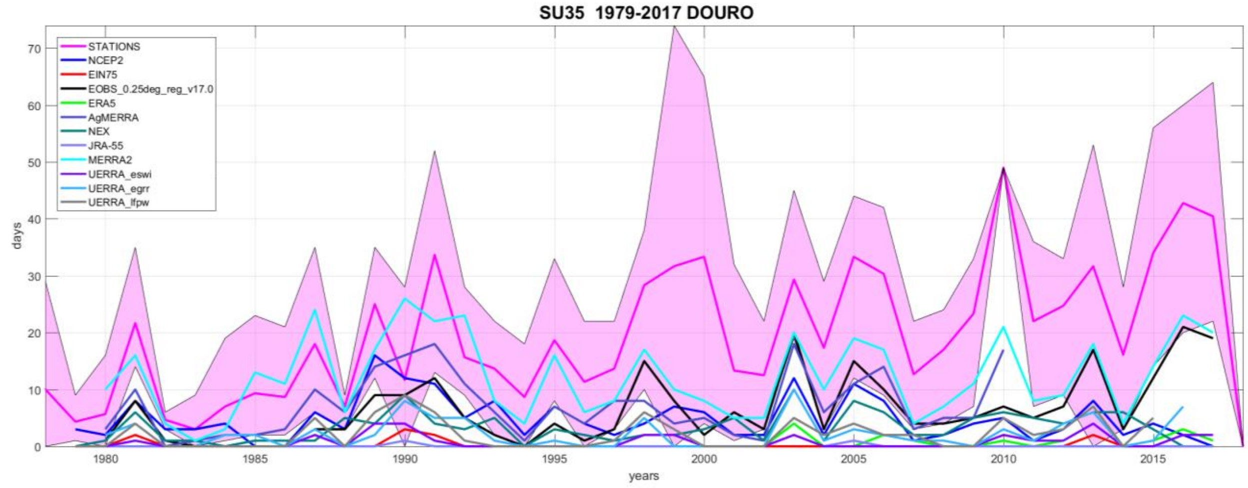


Figure 7-17: Root mean square errors (RMSE) of the gridded datasets (x-axis) against SOGRAPE weather stations (2011-2017) and IPMA weather stations (1981-2017) for summer days (SU). The shading emphasises the accuracy of each dataset; white and pale yellow colours indicate the closest agreement, whereas dark yellow, orange and red indicate progressively poorer agreement. AgMERRA data are only available for 1980-2010, hence no comparison with the SOGRAPE stations was possible.

Further climatic indices of relevance to grape growth and quality were calculated from daily maximum temperatures. The number of heat stress days (SU35; days with Tmax > 35°C) is a good indicator of grape quality. When temperatures exceed 35°C, the vines stop photosynthesising and maturation of the grapes will stop, decreasing sugar, polyphenol and aroma precursor levels, all of which are essential for grape and wine quality. Time series of SU35 are shown in Figure 7-18 (upper panel). There is a large variation in numbers of heat stress days (SU35) between the gridded datasets, with the largest values present in MERRA2. SU35 values are notably larger in the weather stations than those calculated from the gridded datasets. An increasing trend in SU35 is seen in the weather station data, but any trends in the gridded data are much weaker or not present. The differences in SU35 between individual datasets appears to be partly related to their resolutions. The largest numbers of heat stress days are seen in E-OBS and MERRA-2, which have the highest spatial resolutions. In contrast, very few SU35 days are seen in either of the ERA reanalyses (EIN75 and ERA5).

The warm spell duration index (WSDI) is defined as the annual count of days with at least six consecutive days when the daily maximum temperatures exceed their 90th percentiles. This index is a measure of when temperatures become too extreme, so that additional grape losses can be incurred owing to flowering disruption (too early in the season) or extreme berry and leaf dehydration and scalding. The values of WSDI are small and similar in each of the gridded datasets, and are generally smaller than the values calculated from the weather stations. For this index, WSDI from E-OBS data are in closest agreement with WSDI from the weather stations.

Very large values of WSDI are seen in 2017 in many of the datasets, partly a consequence of the severe heat wave which affected southern Europe in this year (Kew and Philip, 2019). Although small increasing trends were seen in mean summer maximum temperatures (Figure 7-15) and summer days (Figure 7-16) in some of the gridded datasets, no such trends are apparent in either SU35 or WSDI.



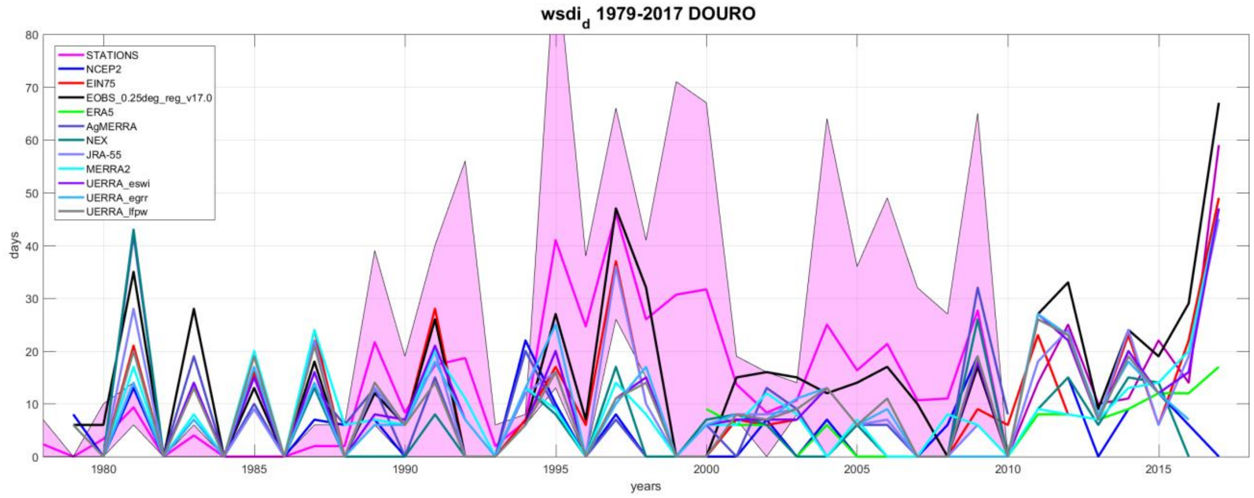


Figure 7-18: Upper panel, numbers of heat stress days (SU35); lower panel, warm spell duration index (WSDI) in the Douro valley. The shading shows the range of values from the SOGRAPE and IPMA stations. WSDI from the SOGRAPE stations was not calculated owing to the short data series.

Root mean square errors (RMSEs) of the gridded datasets against the weather station data for numbers of heat stress days (SU35) are shown in Figure 7-19. The smallest RMSE values (indicating the closest agreement) can be found for E-OBS and MERRA2. The poorest agreements are seen at the stations Caedo, Leda 1 and Leda 3 in all of the gridded datasets. For the IPMA stations, the agreement with the gridded data becomes poorer in the order Guiaes, Folgosa and Santa Barbara.

Linear correlation coefficients for SU35 between the gridded datasets and the IPMA weather stations are shown in Figure 7-20. The closest agreement for all of the datasets is with Guiaes; the coefficients for the other two stations are much smaller, indicating poor agreement. The overall agreement between E-OBS and AgMERRA and the three weather stations is notably better than for any of the other gridded datasets.

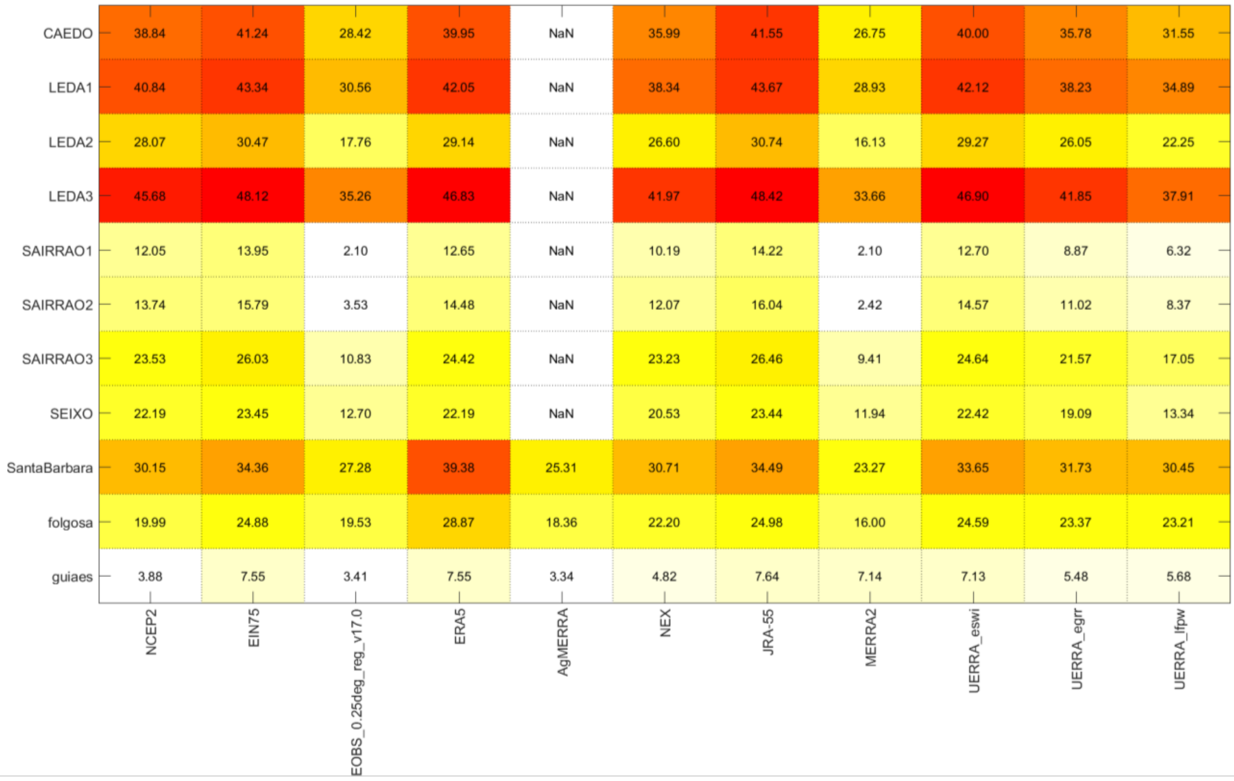


Figure 7-19: Root mean square errors (RMSE) of gridded climate datasets (x-axis) against SOGRAPE and IPMA weather stations (ordinate) for summer heat stress days (SU35). The shading emphasises the accuracy of each dataset; white and pale yellow colours indicate the closest agreement, whereas dark yellow, orange and red indicate progressively poorer agreement. AgMERRA data are only available for 1980-2010, hence no comparison with the SOGRAPE stations was possible.

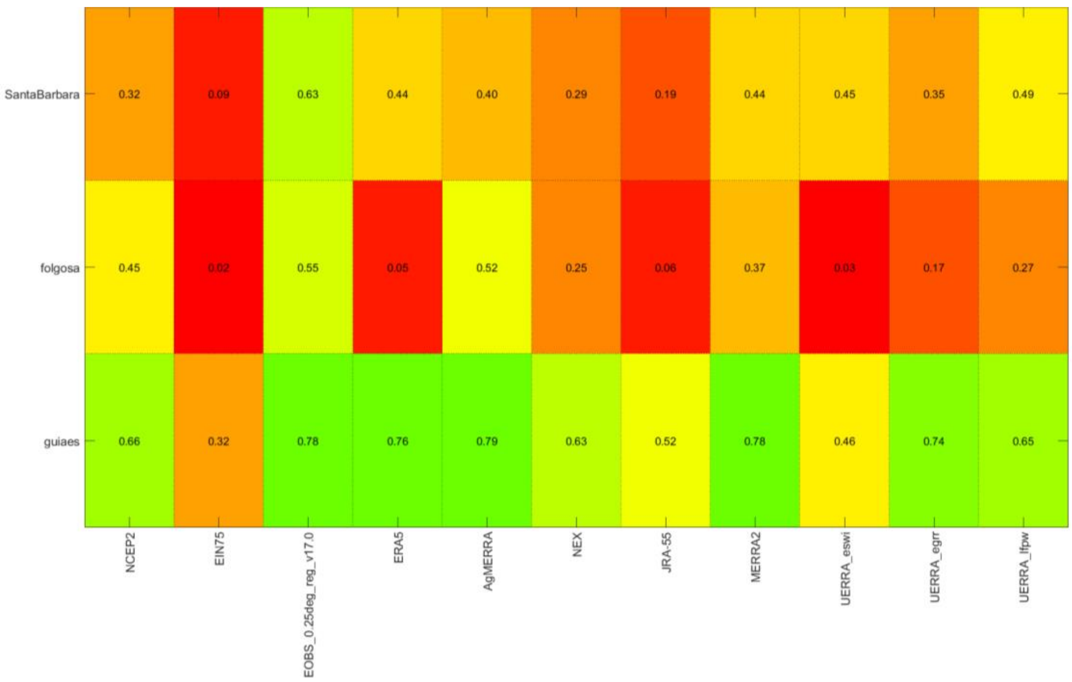


Figure 7-20: Linear correlation coefficients of SU35 from the gridded datasets against the IPMA weather stations for the overlapping period 1981-2017.

7.2.4. INDICES DERIVED FROM DAILY MEAN TEMPERATURES

Two further climatic indices relevant for vines are derived from daily mean temperatures (Tmean). The growing season average temperature (GST) is the average of daily mean temperatures recorded between April 1st and October 31st (in the Northern Hemisphere). This index can provide useful information on which grape varieties are best suited to that location. Growing degree days (GDD) is the sum of differences between daily mean temperatures and 10°C when the daily mean temperatures exceed 10°C, over the same period as GST (i.e. April 1st to October 31st). GDD is used for strategic decision-making in the wine business.

Time series of GST and GDD are shown in Figure 7-21. It can be seen that the two indices calculated from the weather station data are closely correlated. Both indices have low values in 1996 and 1999. The values of each index calculated from the various gridded datasets agree reasonably well, although low values are not seen in 1996 or 1999. Both indices have small upward trends, consistent with a warming climate. Two of the coolest summers seen in Figure 7-15 (1988 and 2007) can also be identified, although the indices in those years are not especially small. The behaviour of the indices calculated from the NEX data from 2010 (low values and a downward trend) is very different to the other gridded datasets.

Figure 7-21: Time series of two indices derived from daily mean temperatures. Upper panel: Growing degree days (GDD); lower panel: Growing season temperature (GST). Both indices are calculated over the period April 1st to October 31st.

Root mean square errors (RMSEs) for the two indices GST and GDD from the weather stations and gridded datasets in the common period 2011-2017 are shown in Figure 7-22. The smallest RMSE values (indicating the closest agreements) are seen for the E-OBS and UERRA_eswi datasets, and the NEX dataset generally has the largest RMSE values and hence poorest agreement across all stations. For the GST (Figure 7-25(a)), the RMSE values for the remaining gridded datasets for each climate index are similar. None of the gridded datasets agree well with the station Sairrao 3. The agreements between the gridded and station data for GDD (Figure 7-25(b)) is poorer than GST.

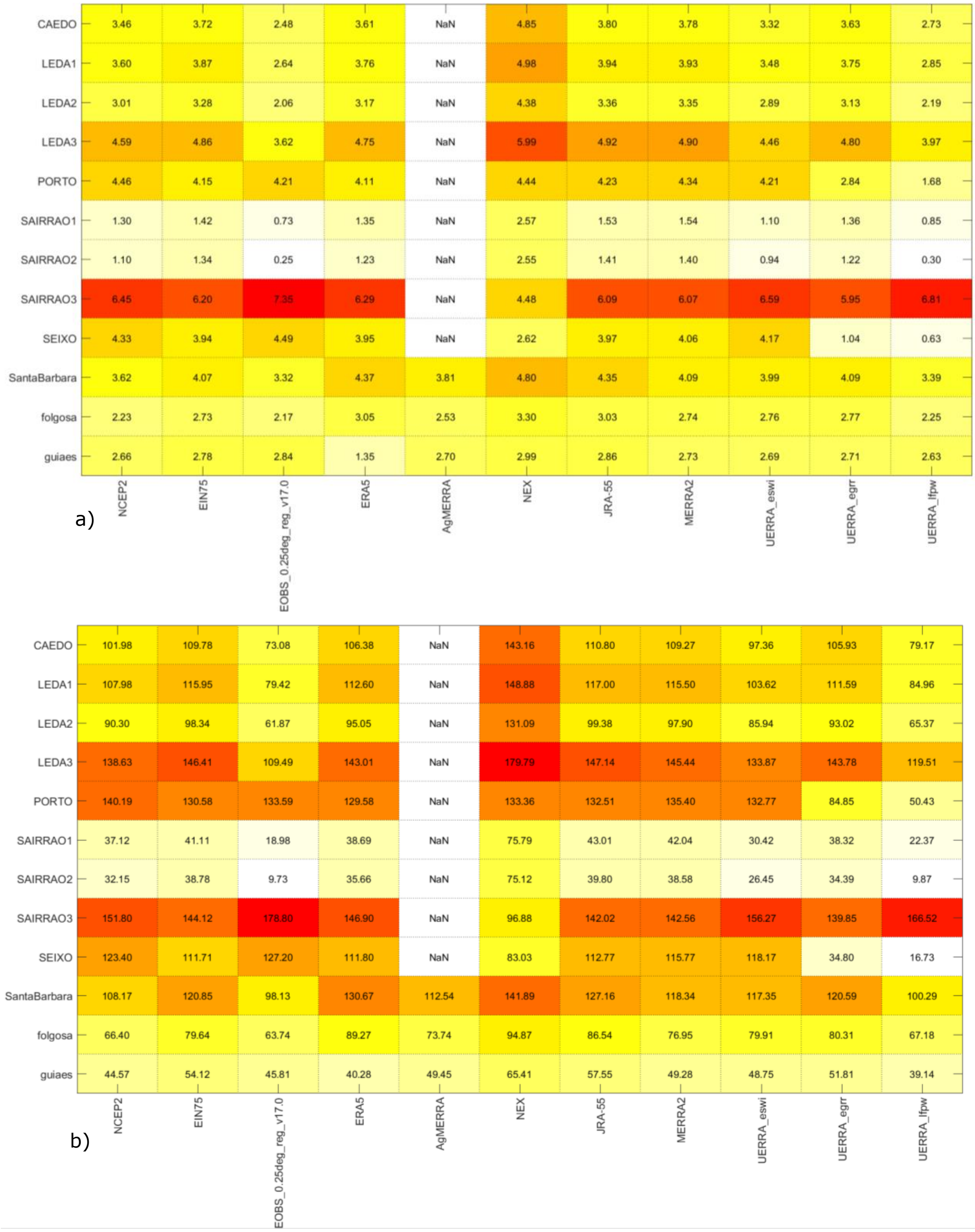


Figure 7-22: Root mean square errors (RMSE) of gridded climate datasets (x-axes) against SOGRAPE and IPMA weather stations (ordinate) for (a) growing season temperatures (GST) and (b) growing degree days (GDD). The shading emphasises the level of agreement between the gridded datasets and weather stations. White and pale yellow colours indicate the closest agreement, whereas dark yellow, orange and red indicate progressively poorer agreement.

Linear correlation coefficients for growing season temperatures calculated from the gridded datasets and the three IPMA weather stations are shown in Figure 7-26. The agreement with the station at Santa Barbara is reasonable, but notably poorer for the other two stations, especially Guiaes. This change in agreement between the three stations is opposite to that seen for SU35 (Figure 7-20), where the best agreement was found for Guiaes.

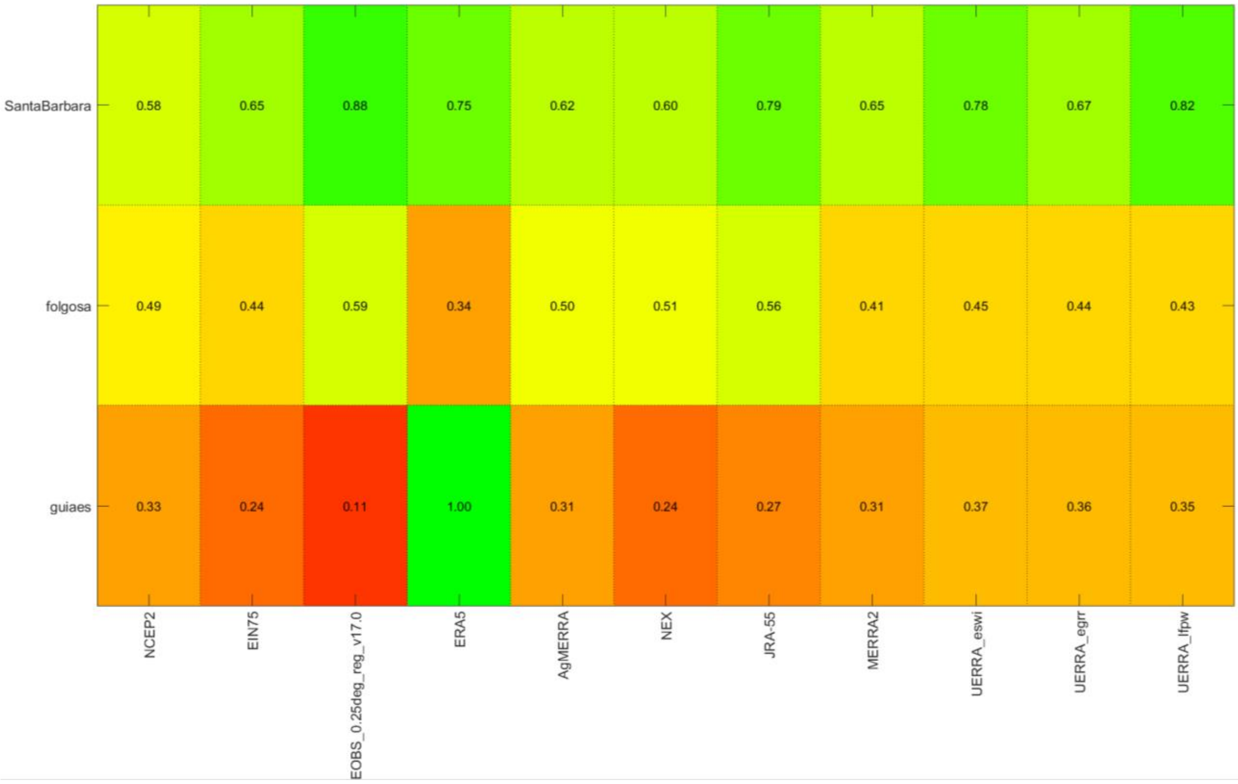


Figure 7-23: Linear correlation coefficients of GST from the gridded datasets with the IPMA weather stations for the period 1981-2017.

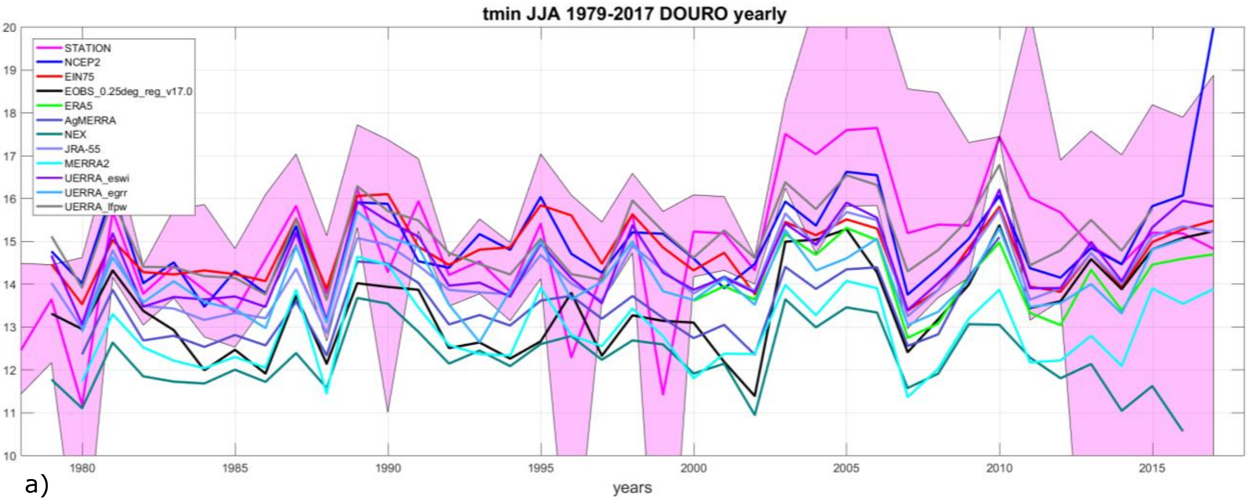
7.2.5. MEAN SUMMER MINIMUM TEMPERATURES

Time series of mean summer daily minimum temperatures (Tmin) and the corresponding anomalies for the Douro valley are shown in Figure 7-24. As for daily maximum temperatures (section 7.2.3), there is an overall good agreement between the long-term trends and interannual variability in the gridded datasets. All of the gridded datasets underestimate summer mean Tmin between 2003 and 2010. Temperatures in E-OBS (solid black lines) are colder than most of the other gridded datasets (Figure 7-24(a)), but ERA-Interim (EIN75) agrees reasonably well with the range of values from the weather stations.

The numbers of days with summer mean Tmin values below the 10th percentile (TN10p) are shown in Figure 7-28. The interannual variability of the TN10p values calculated from the gridded data agrees reasonably well with TN10p from the weather stations. The TN10p values from the gridded data are also larger, owing the colder Tmin values (Figure 7-27(a)). The weather station data show that TN10p values fell between the 1990s and 2000s, a trend which is less apparent in the gridded datasets.

The summer mean Tmin calculated from the NEX dataset has a notable downward trend, especially after 2005, which is not seen in any of the other gridded datasets (Figure 7-24(a)). A corresponding increase in the TN10p index from the NEX data can be seen in Figure 7-25. A sharp increase in summer mean Tmin from the NCEP2 data between 2016 and 2017 (Figure 7-27(a), dark blue line) is apparent, which has a corresponding fall in TN10p over the same period (Figure 7-28).

The summer mean Tmin anomaly series (Figure 7-24(b)) indicates a slight positive trend in the E-OBS dataset (black line). This trend leads to decreasing values of TN10p in E-OBS after 2000, as shown in Figure 7-25. This downward trend in TN10p is not apparent in the other gridded datasets.



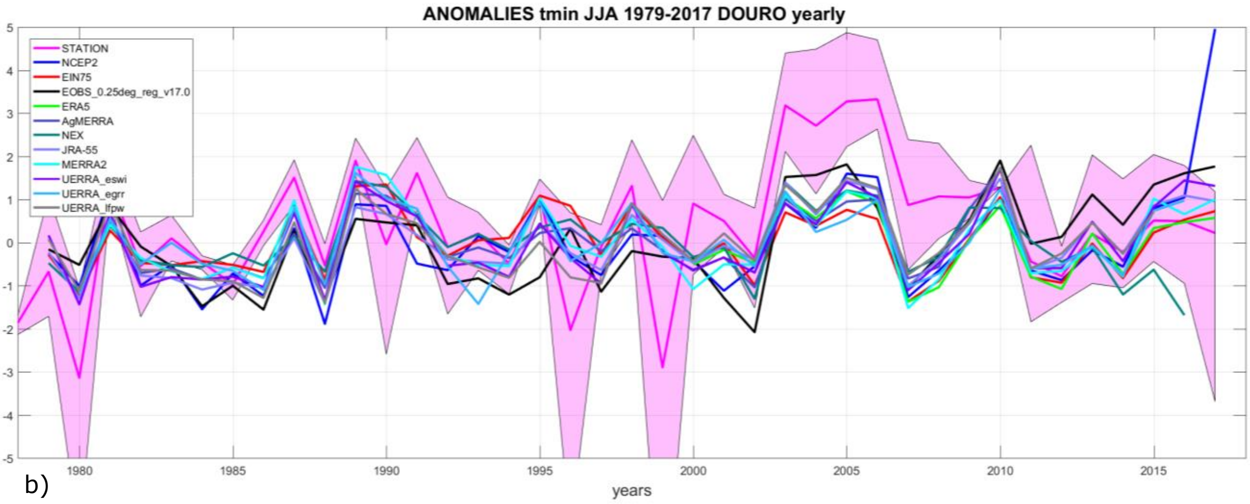


Figure 7-24: a) Time series of mean summer daily minimum temperatures (Tmin) for the Douro valley. The shaded area indicates the range of values from the IPMA and (from 2011) the SOGRAPE weather stations. b) Time series of the corresponding anomalies, where the long-term mean has been removed from each dataset.

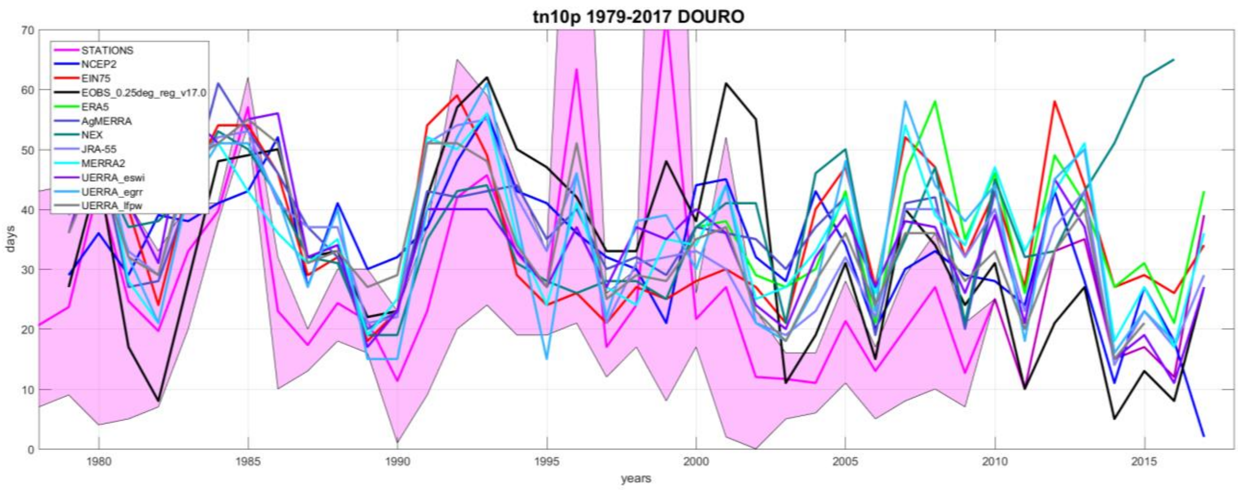


Figure 7-25: Time series of numbers of days with summer mean Tmin values below the 10th percentile (TN10p) for the Douro valley region. The shaded areas show the range of values from the IPMA weather stations.

7.2.6. SUMMARY

Climatic variables and indices calculated from weather station data and gridded datasets have been compared for the Douro valley region of Portugal. These comparisons have shown that the E-OBS dataset consistently had the closest agreement with the weather stations. One of the other gridded datasets, NEX, was created by statistically downscaling global climate model simulations from CMIP5 using gridded data derived from surface observations. This dataset often had the poorest agreement with the gridded data. Another gridded dataset, AgMERRA, was created by bias-correcting the MERRA reanalysis using surface observations. Interestingly, despite the fact that similar surface observations were used in the creation of E-OBS, NEX and AgMERRA, the values of many climatic variables and derived indices from these three datasets are notably different over the Douro valley. These three datasets have similar horizontal resolutions, so the differences are probably caused by the underlying data used (i.e. CMIP5 projections versus MERRA reanalysis), interpolation and bias-correction methods.

7.3. ASSESSMENT FOR STUDY REGIONS IN ITALY

Three study regions in Italy (Foggia, Ancona and Ravenna) are used for WP4, whose focus is durum wheat. Two sources of meteorological data are compared. First, IBIMET operate weather stations in each region; data from one station per region, chosen to represent a specific site, were retrieved. Secondly, a gridded dataset of meteorological variables (MarsMet) has been created via interpolation of weather station data (Toreti, 2014; Toreti et al., 2018), as part of the Crop Growth Monitoring System (Biavetti et al., 2014). This gridded dataset begins in 1975 and is regularly updated.

The analysis of the meteorological variables is conducted by comparing weather station data with the gridded dataset in their overlapping time periods. Trends and seasonal mean statistics are also calculated and compared. Results are shown for a selection of meteorological variables (precipitation, temperature and relative humidity).

7.3.1. FOGGIA

A comparison of daily mean temperatures from a weather station at Foggia and the JRC gridded dataset is shown in Figure 7-26. The two datasets are qualitatively in agreement, except during winter where the temperatures in the JRC dataset are warmer than those recorded at the weather station. Summary statistics for the two datasets are presented in Table 7-1.

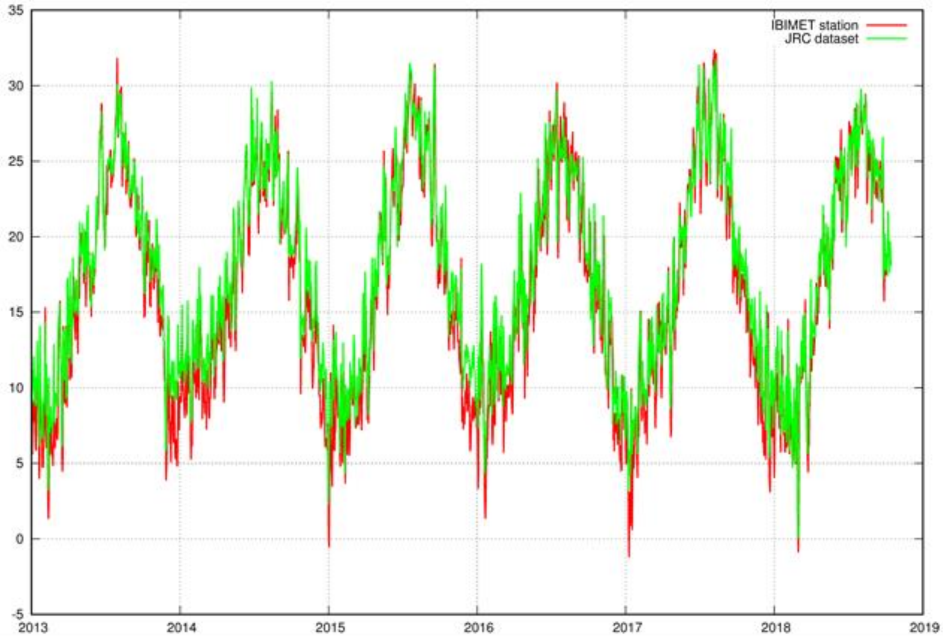


Figure 7-26: Daily mean temperatures from the IBIMET station 02B, located close to Foggia (41.50028°N, 15.51612°E) and the corresponding grid point from the JRC gridded dataset (centred at 41.43448°N, 15.56456°E). The temperature scale on the y-axis is in °C. In the case of the IBIMET station, the daily mean was obtained by averaging hourly values.

Time series of daily precipitation totals from the weather station and JRC dataset are shown in Figure 7-27. The differences between these two datasets are larger than for temperature, owing to the shorter decorrelation distance. Summary statistics for the two datasets are presented in Table 7-1.

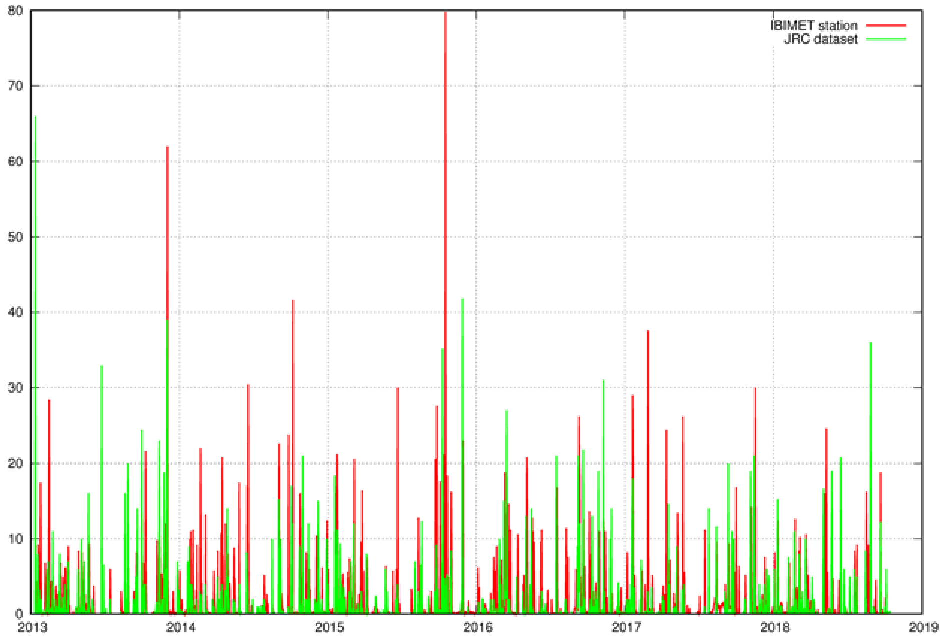


Figure 7-27: Daily total precipitation (mm) from the IBIMET station 02B (red lines), located close to Foggia (41.50028°N, 15.51612°E) and the corresponding grid point from the JRC gridded dataset (green) centred at 41.43448°N, 15.56456°E.

Table 7-1: Summary statistics for seasonal means of five meteorological variables from the IBIMET station and JRC dataset for Foggia.

IBIMET	Tmean / °C	Tmin / °C	Tmax / °C	Precip / mm	Rel Hum / %
Winter	8.2 ± 2.9	4.3 ± 3.6	13.5 ± 3.5	1.5 ± 4.6	75 ± 8
Spring	14.4 ± 4.1	9.0 ± 4.0	20.7 ± 5.4	1.5 ± 3.7	68 ± 8
Summer	25.1 ± 2.9	18.5 ± 2.8	32.2 ± 4.0	0.7 ± 2.7	54 ± 9
Autumn	17.5 ± 4.6	12.8 ± 4.4	23.5 ± 5.5	2.1 ± 6.1	70 ± 11
JRC	**Tmean / °C**	**Tmin / °C**	**Tmax / °C**	**Precip / mm**	**Rel Hum / %**
Winter	10.1 ± 2.7	6.6 ± 3.0	13.5 ± 3.0	1.1 ± 4.2	69 ± 14
Spring	15.4 ± 3.7	11.3 ± 3.6	19.5 ± 4.5	0.9 ± 2.7	69 ± 14
Summer	25.3 ± 2.6	20.8 ± 2.6	29.6 ± 3.2	0.6 ± 2.9	65 ± 10
Autumn	18.6 ± 4.2	14.8 ± 3.8	22.3 ± 4.8	1.6 ± 4.6	67 ± 13

The JRC gridded dataset contains data from 1975. Time series of the seasonal means are shown in Figure 7-28. The very warm summer of 2003 can be seen (‘JJA’; red line with asterisks). The trends in the seasonal mean temperatures and integrated rainfall totals, plus the corresponding trends in maximum and minimum temperatures, are shown in Table 7-2, together with their significance levels. The positive trend in minimum and mean temperatures was highly significant in all seasons, and was largest in spring and summer, mainly due to the pronounced warming of night-time temperatures. Rainfall amounts were highly variable; drying trends were found for all four seasons, but was most significant for spring precipitation.

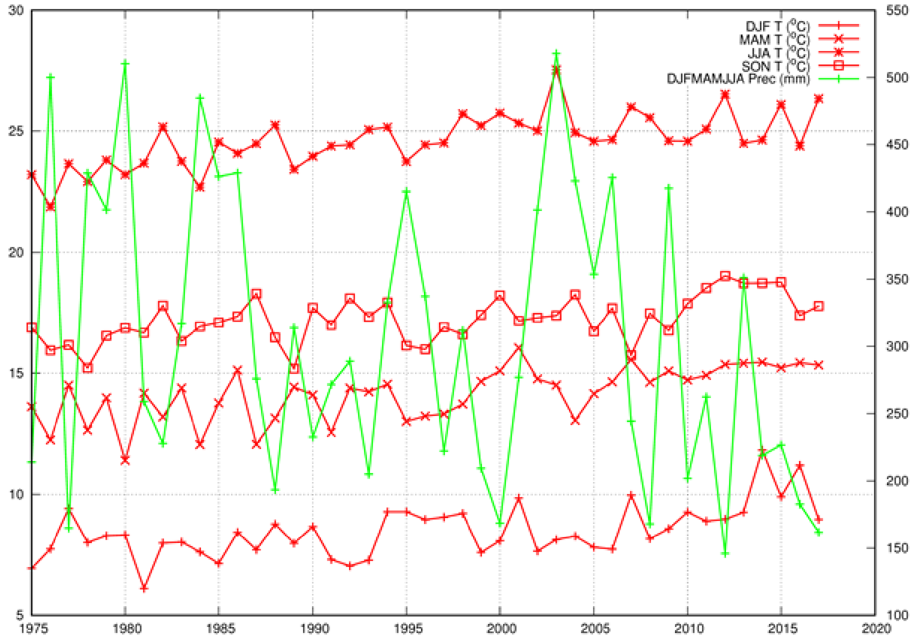


Figure 7-28: Time-series of seasonal mean temperatures and integrated precipitation totals from the JRC dataset for Foggia. Temperatures (in °C) are shown on the left-hand axis, and precipitation totals (in mm) are shown on the right-hand axis. Precipitation is expressed in mm accumulated per long season (December to August).

Table 7-2: Trends in the seasonal means (shown in Figure 7-36), which were computed from the JRC dataset for Foggia. The trends have units of °C, mm or % per year. The statistical significance codes are: '*' 0.001; '**' 0.01; '*' 0.05; '.' 0.1; ' ' 1.0**

Season	Tmean	Tmin	Tmax	Precip	Rel Hum
Winter	0.04***	0.08***	0.02	-0.02	-0.09
Spring	0.06***	0.10***	0.02	-1.4*	0.27**
Summer	0.06***	0.10***	0.01	-1.1 .	0.49***
Autumn	0.04***	0.08***	-0.001	-0.5	0.12

7.3.2. ANCONA

Errore. L'origine riferimento non è stata trovata. shows the comparison between daily mean temperatures from a weather station and the JRC dataset. A longer series of data is available from this station (from 2007) compared with the one located at Foggia. The two datasets are in good agreement. A similar time series of daily precipitation totals is shown in Figure 7-30. Summary statistics of the two datasets are listed in Table 7-3. Compared with Foggia, the Ancona site display a more moderate seasonal cycle i.e. warmer autumn and winter and cooler spring and summer. In this case, the comparison between the JRC dataset and the IBIMET station is in good agreement in winter as well. Daily mean precipitation is slightly larger than for Foggia, especially in spring and summer.

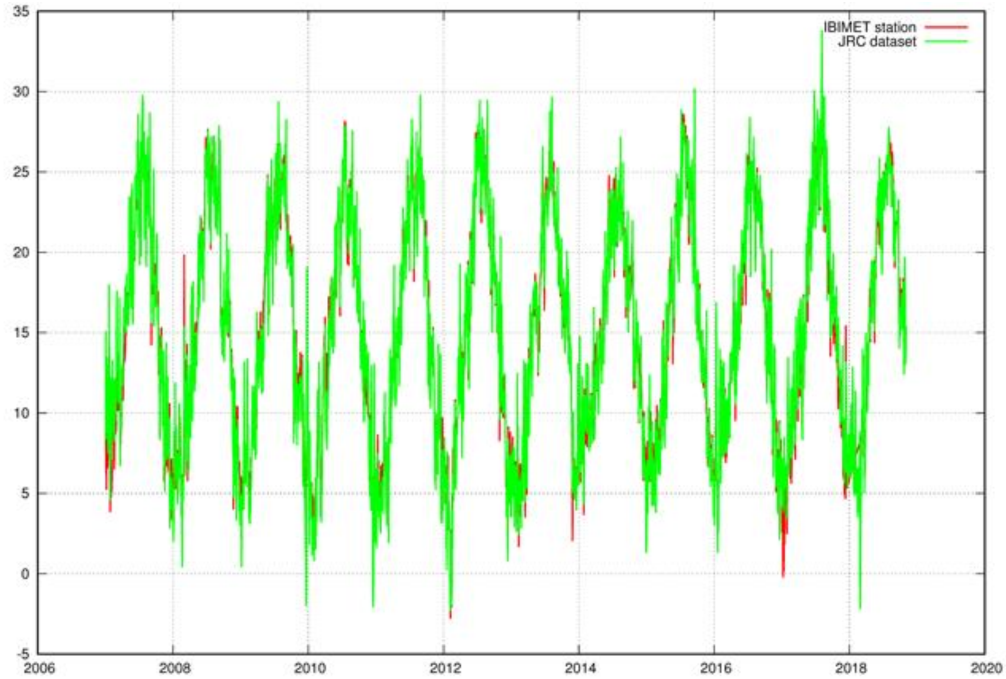


Figure 7-29: Daily mean temperature series from a weather station located in Ancona (Porto S. Elpidio Lon 13° 46', Lat 43° 13'), red lines, and the corresponding grid point from the JRC gridded dataset (centred at 43.32718°N, 13.57282°E), green lines. The temperature scale on the y-axis is in °C. The daily mean temperatures from the weather station were obtained by averaging hourly values.

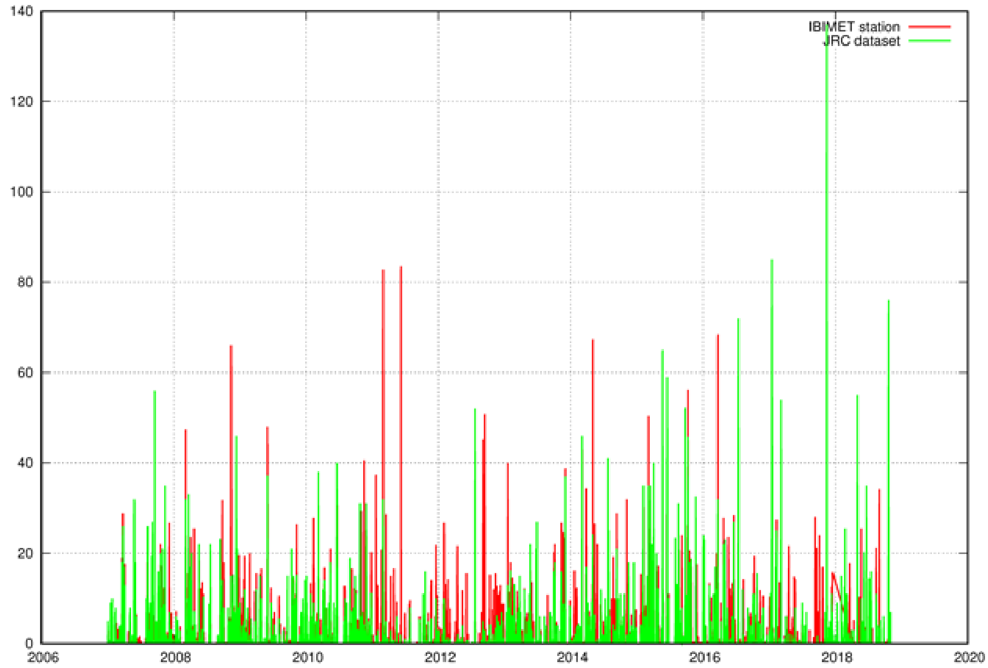


Figure 7-30: Daily total precipitation series (mm) from a weather station (red lines) and the corresponding grid point from the JRC gridded dataset (green), centred at 43.32718°N, 13.57282°E.

Table 7-3: Summary statistics for seasonal means of five meteorological variables from the IBIMET station and JRC dataset for Ancona.

IBIMET	Tmean / °C	Tmin / °C	Tmax / °C	Precip / mm	Rel Hum / %
DJF	6.8 ± 2.5	3.4 ± 2.8	10.9 ± 3.0	1.7 ± 4.5	75 ± 8
MAM	13.6 ± 3.8	9.1 ± 3.8	18.9 ± 4.3	2.2 ± 6.7	68 ± 8
JJA	23.5 ± 2.5	18.4 ± 2.0	28.0 ± 2.9	1.0 ± 4.2	54 ± 9
SON	15.7 ± 4.4	11.7 ± 4.2	20.0 ± 4.7	2.2 ± 6.0	70 ± 11
JRC	**Tmean / °C**	**Tmin / °C**	**Tmax / °C**	**Precip / mm**	**Rel Hum / %**
DJF	6.7 ± 2.9	2.9 ± 3.4	10.6 ± 3.5	1.5 ± 5.3	78 ± 14
MAM	13.6 ± 4.1	8.9 ± 4.2	18.2 ± 4.6	1.7 ± 5.7	68 ± 14
JJA	23.4 ± 2.7	18.4 ± 2.8	28.4 ± 3.0	1.1 ± 4.9	61 ± 11
SON	15.8 ± 4.5	11.7 ± 4.5	20.0 ± 5.1	1.7 ± 6.8	74 ± 13

Time series of seasonal mean temperatures and accumulated precipitation for Ancona are shown in Figure 7-31. Seasonal means from both the JRC dataset (from 1975) and the IBIMET station are shown. The seasonal mean temperatures are in good agreement, but larger differences are seen between the two precipitation totals, as could be expected when comparing a point measurement with an area-averaged value. Once again, the very warm summer of 2003 can be seen (‘JJA’; red line with asterisks), although 2017 was slightly warmer.

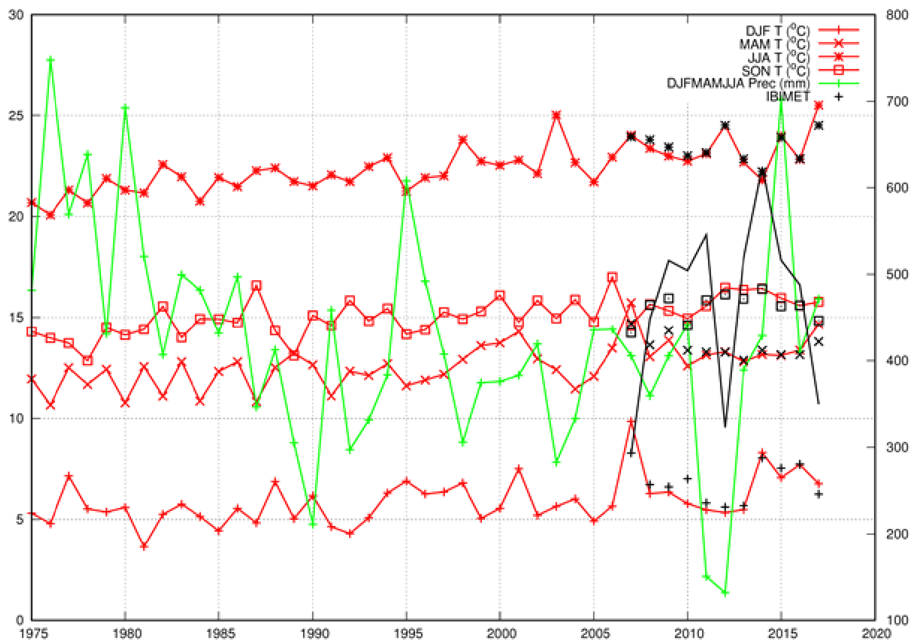


Figure 7-31: Time-series of seasonal mean temperatures and integrated precipitation totals for Ancona from the JRC dataset. Temperatures (in °C) are shown on the left-hand axis, and precipitation totals (in mm) are shown on the right-hand axis. Precipitation is expressed in mm accumulated per long season (December to August). The black symbols represent seasonal means calculated from the IBIMET station, and the solid black line shows accumulated precipitation from the IBIMET station data.

Table 7-4: Trends in the seasonal means (shown in Figure 7-31), which were computed from the JRC dataset for Ancona. The trends have units of °C, mm or % per year. The statistical significance codes are: '*' 0.001; '**' 0.01; '*' 0.05; '.' 0.1; ' ' 1.0**

JRC	Mean T	Min T	Max T	Precip	Rel.Hum.
Winter	0.03**	0.05***	0.03*	-0.42	0.11
Spring	0.05***	0.07***	0.03*	-0.9	-0.02
Summer	0.06***	0.09***	0.05***	-2.5**	-0.27**
Autumn	0.04***	0.07***	0.03**	-2.1*	-0.01

Trends in the seasonal means (shown in Figure 7-31), which were computed from the JRC dataset for Ancona, are listed in Table 7-4. The temperature trends are highly significant and have similar values to those computed for Foggia. One major difference in Ancona is that significant warming is also found in maximum temperatures. Ancona exhibits a larger negative trend in precipitation in summer and autumn than Foggia. In these two seasons, the maximum temperature and the precipitation trends are possibly related. A large negative trend in precipitation is seen in winter, which is notably larger than the trend at Foggia, but is not statistically significant.

7.3.3. RAVENNA

Daily mean temperatures for the common period 2007-2016 from the IBIMET station and corresponding grid box of the JRC dataset are shown in Figure 7-32. The temperatures are very similar, although those in the JRC dataset tend to be slightly cooler than those recorded at the weather station. Time series of daily precipitation totals from the weather station and JRC dataset are shown in Figure 7-33. Similar totals and temporal behaviour is seen in both datasets. The agreement between the summary statistics calculated from the station and JRC data at Ravenna is reasonable (Table 7-5). The summary statistics for Ravenna are closer in value to those for Ancona (Table 7-3) than Foggia (Table 7-1).

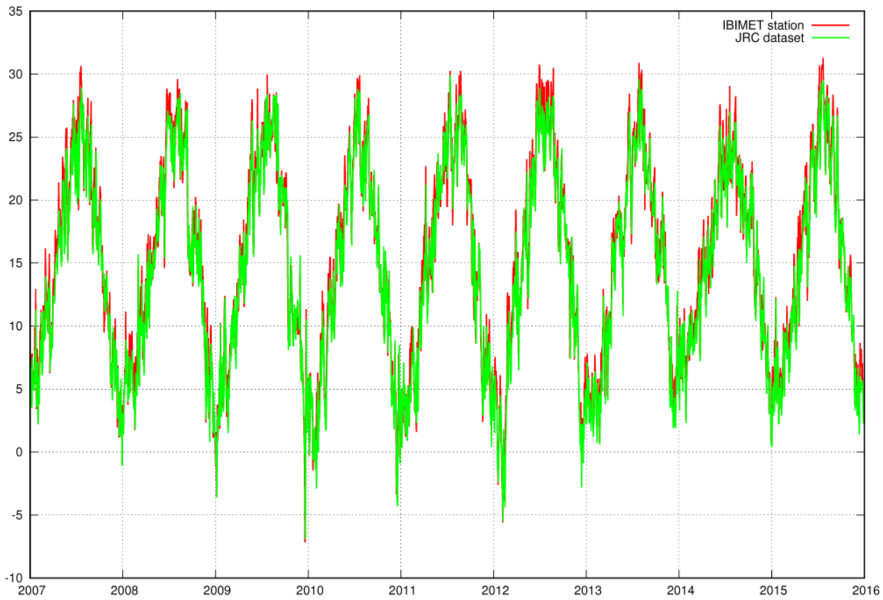


Figure 7-32: Daily mean temperatures in °C from the IBIMET station located in Ravenna Porto S. Elpidio (12.200°N, 44.414°E) and the corresponding grid point from the JRC gridded dataset (centred at 44.50°N, 12.09°E). For the IBIMET station, the daily mean was obtained by averaging hourly values.

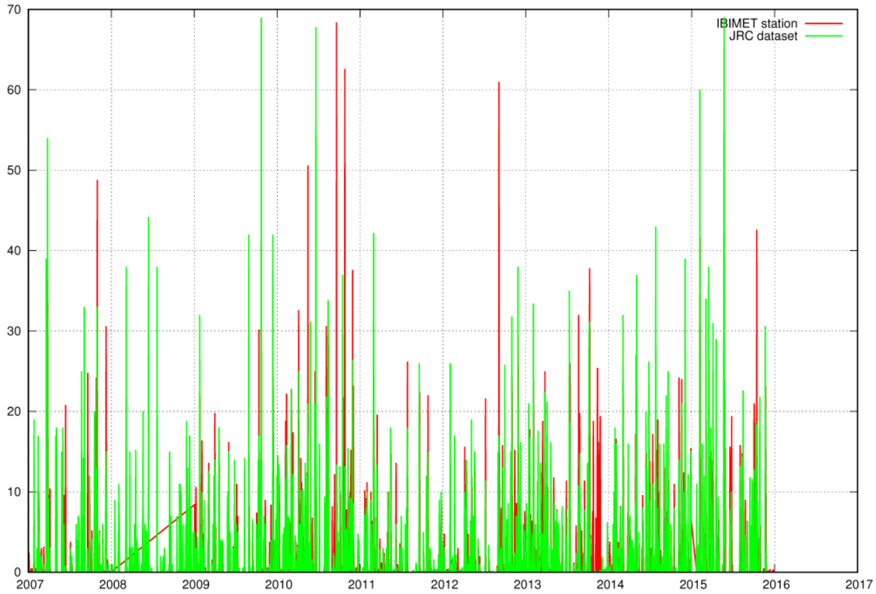


Figure 7-33: Daily total precipitation series (mm) from a weather station (red lines) and the corresponding grid point from the JRC gridded dataset (green), for Ravenna. No precipitation data were available from the weather station for 2008.

Table 7-5: Summary statistics for seasonal means of five meteorological variables from the IBIMET station and JRC dataset for Ravenna.

IBIMET	Tmean / °C	Tmin / °C	Tmax / °C	Precip / mm	Rel Hum / %
DJF	5.3 ± 3.1	2.2 ± 3.0	8.4 ± 3.7	1.5 ± 4.1	82 ± 13
MAM	14.6 ± 4.5	10.3 ± 4.2	18.9 ± 5.2	1.9 ± 5.0	66 ± 15
JJA	24.5 ± 2.9	19.7 ± 2.7	29.3 ± 3.7	1.1 ± 3.9	57 ± 11
SON	15.4 ± 5.2	11.6 ± 4.8	19.2 ± 5.8	2.2 ± 6.8	74 ± 14
JRC	**Tmean / °C**	**Tmin / °C**	**Tmax / °C**	**Precip / mm**	**Rel Hum / %**
DJF	6.7 ± 2.9	1.5 ± 3.3	8.0 ± 3.6	1.7 ± 4.7	82 ± 10
MAM	13.6 ± 4.1	8.6 ± 4.2	19.0 ± 5.2	2.3 ± 6.4	72 ± 14
JJA	23.4 ± 2.7	18.0 ± 2.5	29.8 ± 3.3	1.3 ± 5.2	62 ± 13
SON	15.8 ± 4.5	11.1 ± 4.7	19.6 ± 6.1	1.9 ± 5.4	75 ± 11

Time series of seasonal mean temperatures and accumulated precipitation for Ravenna are shown in Figure 7-34. Temperatures in all four seasons exhibit an upward trend. The very warm summer of 2003 can be seen in the maximum temperature series ('JJA'; red line with asterisks). The precipitation data (green line) exhibit very high interannual variability, with wet years (1987, 2005, 2010 and 2015) separated by several dry years.

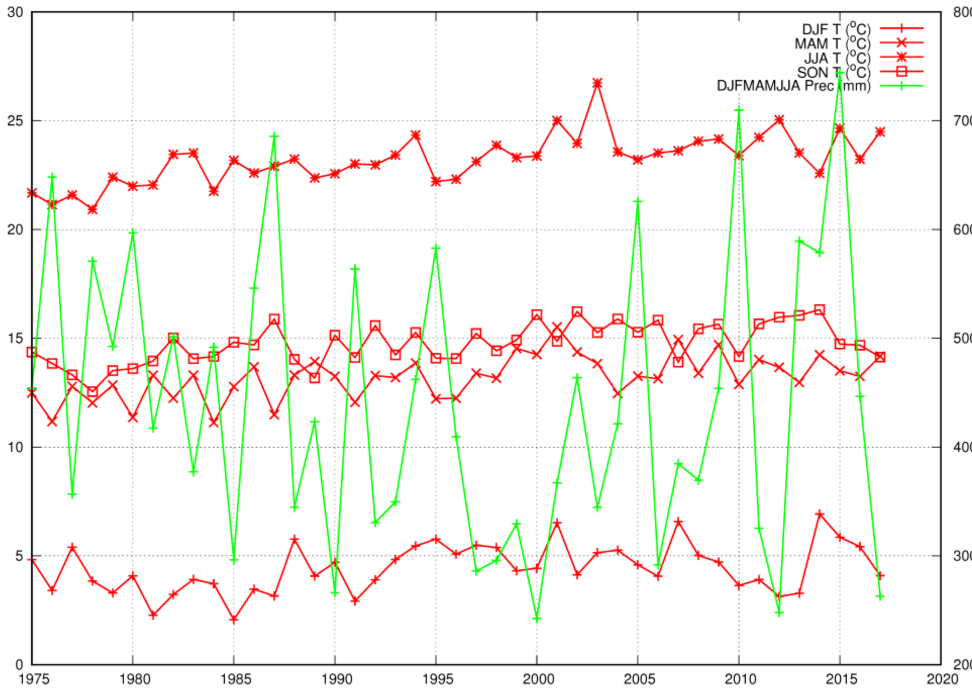


Figure 7-34: Time-series of seasonal mean temperatures and integrated precipitation totals for Ravenna from the JRC dataset. Temperatures (in °C) are shown on the left-hand axis, and precipitation totals (in mm) are shown on the right-hand axis. Precipitation is expressed in mm accumulated per long season (December to August).

Trends calculated from the data in Figure 7-34 are summarised in Table 7-6, together with their levels of statistical significance. As for Foggia and Ancona, the trends in temperatures are significant in all four seasons. However, the precipitation trends in winter and spring at Ravenna are positive, whereas the corresponding trends at Foggia and Ancona are negative (although these trends are not statistically significant). Ravenna, like Ancona, is characterized by significant warming trends in maximum temperatures. The warming trends in minimum temperatures are generally larger than those in maximum temperatures.

Table 7-6: Trends in the seasonal means (shown in Figure 7-34), which were computed from the JRC dataset for Ravenna. The trends have units of °C, mm or % per year. The statistical significance codes are: '*' 0.001; '**' 0.01; '*' 0.05; '.' 0.1; ' ' 1.0**

JRC	Tmean	Tmin	Tmax	Precip	Rel Hum
DJF	0.03*	0.04**	0.02	0.34	0.14
MAM	0.04***	0.05***	0.04**	0.96	0.12
JJA	0.06***	0.05***	0.07***	-2.25*	-0.14
SON	0.04***	0.05***	0.03**	-0.2	-0.06

7.3.4. ROOT MEAN SQUARE ERRORS

RMSEs have been calculated for the three stations in Italy, and are shown in Table 7-7. The lowest values are seen for Ancona, except for relative humidity.

Table 7-7: RMSE values for the three stations in Italy, which were computed using the MarsMet gridded dataset.

Station	Tmean / °C	Tmin / °C	Tmax / °C	Precip / mm	Rel Hum/ %
Foggia	2.21	3.08	3.05	6.10	15.4
Ancona	1.34	2.04	1.67	4.56	12.7
Ravenna	2.02	2.78	2.72	6.51	9.70

7.3.5. SUMMARY

Data from a single weather station in each of the three regions Foggia, Ravenna and Ancona have been compared with a gridded dataset created by JRC as part of the crop growth monitoring system (MarsMet). The two datasets were found to agree satisfactorily. Using the gridded data, significant upward trends in temperature were found at all sites between 1975 and 2018. A downward trend in summer rainfall over the same period was noted for all three sites, although the level of statistical significance varied between the sites.

The RMSE values for Tmax in Italy (Table 7-7) are larger than those for Andalucía (section 7.1.2). In the latter region, the RMSE values for Tmax were mostly in the range 0.64 – 1.37, with one station higher at 3.50. For Tmin, the values in Andalucía were in the range 0.15 – 4.70 (section 7.1.3), and so the RMSE values in Italy are comparable. For precipitation, the RMSEs in Italy are slightly smaller than those for Andalucía (section 7.1.1).

7.4. SUMMARY FOR EUROPE

RMSE values for three key meteorological variables (daily maximum temperature, daily minimum temperature and daily total precipitation) are shown in Table 7-8 for stations in Spain, Portugal and Italy. All of these RMSEs were calculated using daily data from E-OBS, and include comparisons for additional stations in Spain and Italy. Overall, the RMSE values show that the errors in daily maximum and minimum temperatures are similar, although they are smaller for minimum temperatures at stations in Portugal than for maximum temperatures. The RMSE values for precipitation are almost always higher than those for temperature. This latter result is expected, as rainfall amounts tend to vary over shorter distances than temperatures. Additionally, the extreme precipitation totals seen in Italy for stations in the Marche region (e.g., Mozzano, Porto St. Elpidio), when, for example, daily precipitation totals exceeding 90 mm were recorded in November 2013, are not reproduced in E-OBS, which may partially explain the high RMSE values.

Table 7-8: RMSE values for the European stations. These RMSE values were calculated using all valid daily data for each station. The data periods will differ between the stations, and the RMSE values will be slightly different to those shown earlier. No precipitation data were available for Ancona, Ravenna or Villa Fastiggi.

Country	Station	Max Temperature	Min Temperature	Precipitation
Spain	Almeria	4.45	4.31	5.45
	Basurta	0.93	2.07	3.70
	Cordoba	0.82	1.16	3.10
	Huelva	1.10	1.08	4.55
	Granada	3.32	4.54	4.82
	Jaen	1.55	4.16	3.14
	Malaga	2.56	1.97	4.18
	Sevilla	1.13	2.42	3.18
Portugal	Caêdo	2.79	1.94	2.54
	Leda 1	2.37	1.54	3.10
	Leda 2	1.31	1.38	2.73

Country	Station	Max Temperature	Min Temperature	Precipitation
	Leda 3	2.71	2.26	2.61
	Porto	3.02	2.37	3.68
	Sairrão 1	1.22	1.23	2.63
	Sairrão 2	0.96	0.73	2.24
	Sairrão 3	0.77	0.83	1.65
	Seixo	2.26	2.16	2.92
	Folgosa	4.06	2.71	2.78
	Guiaes	3.60	2.36	2.89
	Santa Barbara	4.45	2.88	4.15
Italy	Ancona	2.55	2.19	-
	Classe	2.80	1.80	7.21
	Colle	3.12	2.53	6.03
	Foggia	3.88	3.58	5.08
	Mozzano	3.61	3.93	9.64
	Porto St Elpidio	1.96	2.41	7.08
	Ravenna	3.07	2.47	-
	San Benedetto Tronto	2.08	2.73	8.19
	Villa Fastiggi	2.63	2.71	-

7.5. ASSESSMENT FOR COLOMBIA

Colombia is a major producer of coffee, which is generally grown at altitudes over 1200 m. An equivalent dataset to E-OBS or CGMS is not available for Colombia. Additionally, local surface observations located within coffee plantations could not be obtained. The decision was made to assess gridded data from AgMERRA (Ruane et al., 2015) by comparison with observations from the GHCN-D dataset (section 5.4.1). AgMERRA was derived by bias-correcting climatic data from the MERRA reanalysis (Rienecker et al., 2011) using surface observations, and so is the closest to a gridded dataset derived from surface observations available for Colombia. This assessment of AgMERRA is less rigorous than the assessments of gridded datasets in sections 7.1, 7.2 and 7.3. Many of the observations in GHCN-D were used to produce the AgMERRA dataset. Daily mean temperatures are not available from AgMERRA.

For some stations, there were long gaps where few or no data were available. The numbers of days with missing data varies between different stations, and with the climatic variable. Hence, the numbers of stations used in the comparisons of AgMERRA data with each climatic variable is not the same. Additionally, the years with high data availability vary between stations. In some cases, data are mostly available from the 1960s to the 1980s, whereas for other stations, the most complete coverage begins around 2005. The assessment of observations and AgMERRA in Colombia will focus on climatological monthly mean values of minimum and maximum temperatures, and monthly total precipitation amounts. The climatological monthly values for AgMERRA were calculated using all available years (1980 – 2010). Some of the differences between the observations and AgMERRA may reflect the different time periods over which the observations were available. In all of the figures in this section, black lines and solid circles bars represent observed values. Purple shaded areas and solid lines show the minimum and maximum values for each month, and the climatological mean value from AgMERRA, respectively.

7.5.1. PRECIPITATION

Precipitation totals from the observations and AgMERRA are compared in Figure 7-35**: Climatological monthly total precipitation (Prcp) for stations in Colombia. Observed values are shown by the black lines and solid circles. The mean values from AgMERRA are shown by the solid purple line, and the shading represents the minimum and maximum values in each month. The precipitation data have units of mm per month.**. The paucity of precipitation observations for some stations makes an assessment difficult. Nevertheless, some conclusions can be made. For those stations where observed values are available for most or all months, the agreement between the station and AgMERRA data is quite good. The agreement is poor at station 80112, where AgMERRA overestimates the rainfall totals in most months. Little can be said for the stations 80036, 80063, 80084, 80211, 80214 and 80315, as there are very few monthly data available, although the AgMERRA data are generally consistent with the few observed values available.

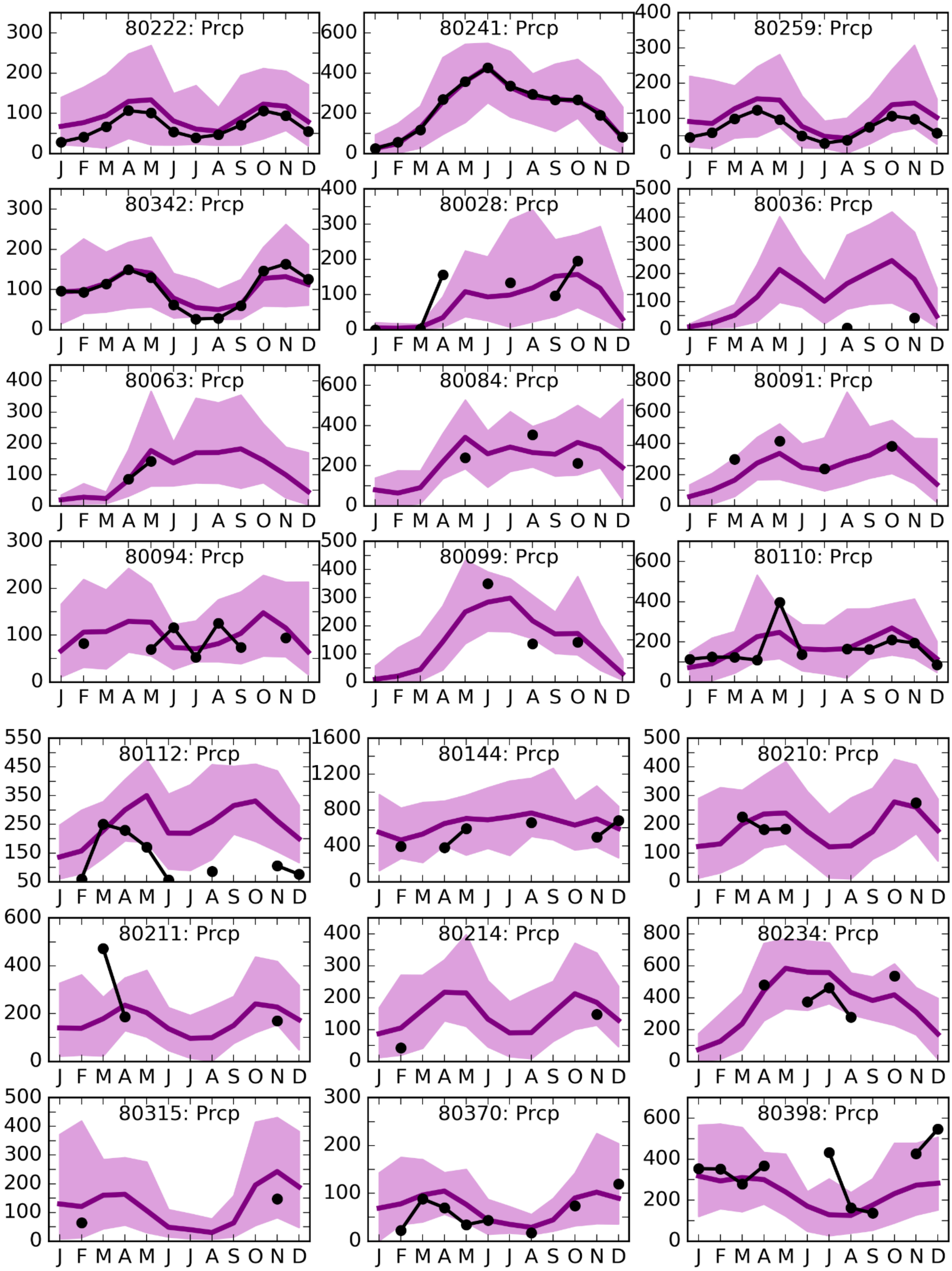


Figure 7-35: Climatological monthly total precipitation (Prcp) for stations in Colombia. Observed values are shown by the black lines and solid circles. The mean values from AgMERRA are shown by the solid purple line, and the shading represents the minimum and maximum values in each month. The precipitation data have units of mm per month.

7.5.2. MAXIMUM TEMPERATURES

The least amount of data was available for maximum temperatures across all stations in the GHCN-D in Colombia. Sufficient data to calculate monthly means and climatological values were only available for eight stations, which are shown in Figure 7-36. There is little variation in maximum temperatures during the year at all stations. The agreement between the observed and AgMERRA values is generally quite good, although the AgMERRA data for stations 80110 and 80342 tend to be lower than the observed values by 2-3°C, whereas at station 80094, the AgMERRA maximum temperatures are about 3°C higher in each month.

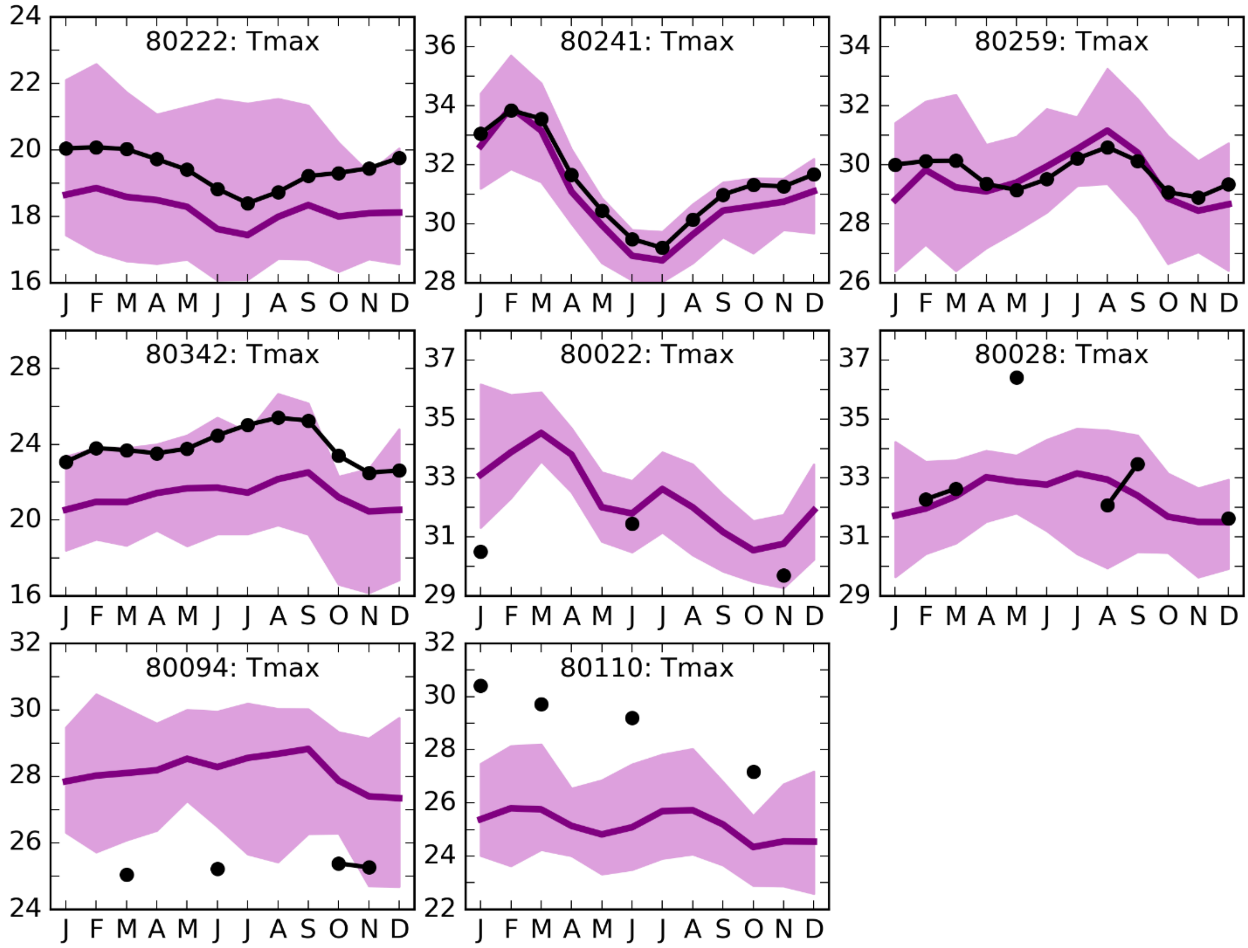


Figure 7-36: Climatological monthly means of maximum temperatures (Tmax) for stations in Colombia. Observed values are shown by the black lines and solid circles. The mean values from AgMERRA are shown by the solid purple line, and the shading represents the minimum and maximum values in each month. The temperatures have units of °C.

7.5.3. MINIMUM TEMPERATURES

Climatological monthly means of minimum temperatures are shown in **Errore. L'origine riferimento non è stata trovata.** and Figure 7-37. The minimum temperatures in AgMERRA are often lower or higher than the observed values. The observed values lie wholly or mostly within the range from AgMERRA at just 7 stations.

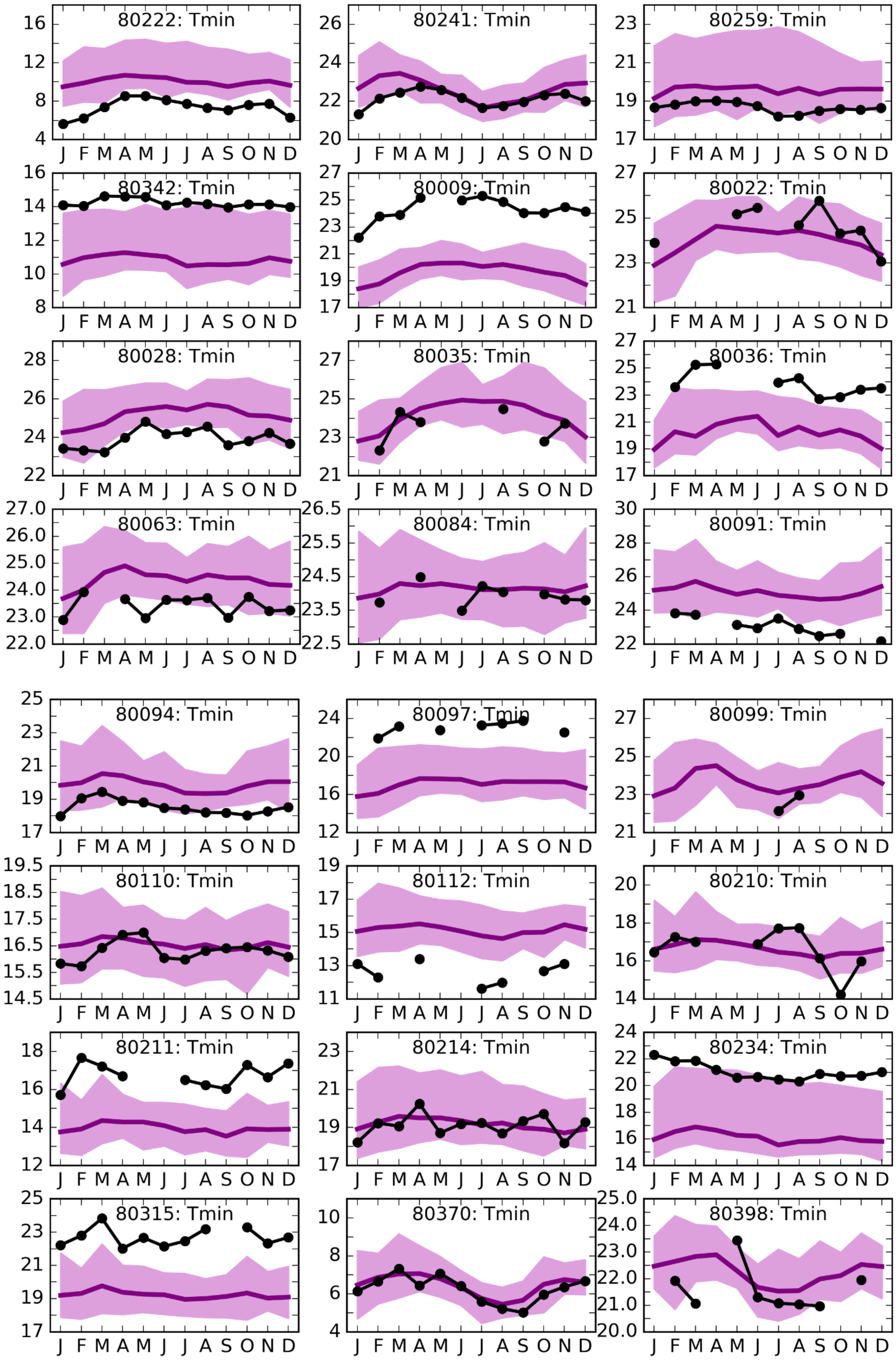


Figure 7-37: Climatological monthly means of minimum temperatures (Tmin) for stations in Colombia. Observed values are shown by the black lines and solid circles. The mean values from AgMERRA are shown by the solid purple line, and the shading represents the minimum and maximum values in each month. The temperatures have units of °C.

7.5.4. SUMMARY

Monthly minimum and maximum temperatures and monthly precipitation totals from weather stations in Colombia from the GHCN-D database have been compared with corresponding values from the AgMERRA gridded dataset. Large periods with missing data were found at most stations, and the periods with complete series of data also varied between stations and with the variable in question. Daily mean temperatures and daily rainfall totals had the highest proportions of data available.

The agreement between the weather stations and AgMERRA was reasonable at some stations, but biases were evident at others. This apparent bias may be partly caused by the periods for which weather station data were available. The monthly means from AgMERRA were calculated using data for 1980-2010, but data from some of the stations were largely unavailable until around 2010. An apparent warm bias in the weather station temperatures might therefore be expected.

8. APPROPRIATENESS OF DATASETS FOR USE WITHIN MED-GOLD

Local weather station data have been compared with gridded datasets, principally E-OBS, in the three MED-GOLD study regions. For Andalucía, temperature and precipitation data from E-OBS were compared against observations provided by the project partner DCOOP for the period 2001-2017. The analysis revealed that E-OBS satisfactorily captured the observed patterns of temperature and precipitation at the majority of stations. Similar conclusions were reached for the comparison in the Douro valley. The E-OBS gridded dataset consistently had close agreement with the local weather observations.

In Italy, temperatures and precipitation totals from the MarsMet gridded data agreed satisfactorily with the weather station data, although some differences in the precipitation amounts was noted. No published report or paper exists comparing E-OBS with MarsMet, but an earlier comparison led to satisfactory results (Zampieri 2018, personal communication). A study by Toreti et al. (2018) showed that crop yield monitoring and forecasting systems driven by AgMERRA showed similar performance to those driven by MarsMet. A direct comparison of AgMERRA with E-OBS could not be found. For this report, it has to be assumed that similar results would be obtained if the comparisons in Italy had used E-OBS instead of MarsMet.

RMSE values were calculated for the data in all three study regions. For Andalucía and Italy, the RMSE values for maximum and minimum temperatures were lower for Andalucía, but slightly higher for precipitation. Overall, the errors in E-OBS compared with the station data are similar in both regions. The RMSEs for the Douro valley were larger, as they were calculated using seasonally averaged values.

However, E-OBS does have its limitations, which are apparent when data from E-OBS were compared with higher resolution datasets created for individual countries that are based on larger numbers of weather stations. Hofstra et al. (2009) compared the first version of E-OBS (Haylock et al., 2008) with other gridded datasets that were based on higher density station networks. They found that precipitation values were usually biased toward lower values in E-OBS. They concluded that E-OBS was fundamentally limited by its underlying station network (i.e. the low density of stations in many areas). Kyselý and Plavcová (2010) compared v2.0 of E-OBS with a high resolution gridded dataset for the Czech Republic, which was based on a much larger number of stations. They concluded that further increases in the amount and quality of station data available within ECA&D and used in the E-OBS data set were needed to reduce the biases.

Similar conclusions were reached by Cornes et al. (2018), who compared the latest versions of E-OBS (v18 and v18e) with several different regional gridded data sets produced by a selection of National Meteorological Services. E-OBS generally had a dry bias in many areas, although a wet bias was apparent across most mountainous regions. The newer versions of E-OBS (v16 onwards) use roughly three times as many stations as the earliest versions (Cornes et al., 2018), so it would be expected that some of the biases in E-OBS will have been reduced.

Although temperatures and precipitation amounts over Europe in many of the gridded datasets exhibit similar features (e.g., warm years, wet years), they do not exhibit the same long-term trends. In the E-OBS data, a gradual increase in mean summer maximum temperatures was noted in the Douro valley. This small increase was not seen, or was notably smaller, in many of the other gridded datasets. The reasons behind these differences are not clear.

One objective of this report is to recommend datasets for evaluation and bias-correction of climate models. In Europe, two approaches are possible. First, high-resolution gridded datasets produced by National Meteorological Services (NMS) could be used, as they are based on large numbers of stations and will have smaller biases than E-OBS. Examples include Spain02 and PTHRES. However, an equivalent dataset for Italy does not appear to exist. One disadvantage of this approach is that only model data over the same region as the NMS dataset could be evaluated and bias-corrected. If it were desired to expand the studies to other regions (e.g., olive groves outside of Andalucía) the required NMS data may not exist to bias-correct the model. The use of different gridded datasets in different parts of Europe may also introduce inconsistencies in the corrected data.

An alternative dataset could be AgMERRA, for which daily values of agriculturally-relevant variables are readily available for a 30-year period at a reasonably high resolution, although the data period ends in 2010. The comparison in the Douro valley showed that the time series of values in E-OBS and AgMERRA were often quite different. E-OBS almost always had the closest agreement with the local weather station data, compared with other gridded data. These results suggest that biases in AgMERRA could be higher than those in E-OBS.

Overall, the E-OBS dataset is suitable for use within MED-GOLD for the evaluation and bias-correction of regional climate model simulations. Its wide coverage means a consistent dataset can be used across Europe for this purpose. It is noted that several published studies have used E-OBS to bias-correct regional climate model projections. E-OBS also has the advantage of being available on several different grids, of which two are the rotated grids used by the EURO- and MED-CORDEX regional climate model simulations. From a practical point of view, the interpolation of the model data to the observed grid or vice-versa is avoided, which is in itself a source of uncertainty.

For Colombia, climatic indices (monthly means of precipitation totals and maximum and minimum temperatures) derived from AgMERRA and observations in the GHCN-D database were compared. Although data were sparse at some stations, the daily maximum temperatures and precipitation totals were found to agree satisfactorily, whereas larger differences were seen with the daily minimum temperatures. Nevertheless, AgMERRA is the dataset which provides meteorological data that are closest in value to surface observations over Colombia. It is noted that a comparison of AgMERRA in Afghanistan with available measurements concluded that AgMERRA could be used to fill in gaps in the meteorological records (Razavi et al., 2018).

9. CONCLUSIONS

Climatic variables from weather stations and a range of gridded datasets have been compared in the key study regions of MED-GOLD. Generally, the agreement between the two datasets was satisfactory, indicating that the climate service tools could be developed and tested with the gridded data. However, in some areas, the comparisons were limited owing to the short data periods of the weather stations. Similarly, the time period of some of the gridded datasets ends around 2010, which limits their usefulness, as they cannot be used with recent phenological and pest data.

Overall, E-OBS is the most useful dataset for bias-correcting the regional climate model data for MED-GOLD.

10. ACKNOWLEDGEMENTS

Climate scenarios from the NEX-GDP dataset were prepared by the Climate Analytics Group and NASA Ames Research Center using the NASA Earth Exchange and distributed by the NASA Center for Climate Simulation (NCCS). The authors acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and they thank the climate modeling groups for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and leads development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

The authors acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu).

11. BIBLIOGRAPHY

Biavetti I, Karetsos S, Ceglar A, Toreti A, Panagos P. 2014. European meteorological data: contribution to research, development, and policy support, in: Hadjimitsis DG, Themistocleous K, Michaelides S, Papadavid G. (Eds.), Second International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2014). International Society for Optics and Photonics, p. 922907. doi:10.1117/12.2066286.

Blanco-Ward D, Monteiro A, Lopes M, Borrego C, Silveira C, Viceto C, Castro J. 2017. Analysis of climate change indices in relation to wine production: A case study in the Douro region (Portugal). In BIO Web of Conferences (Vol. 9, p. 01011). EDP Sciences.

Borsche M, Kaiser-Weiss AK, Undén P, Kaspar F. 2015. Methodologies to characterize uncertainties in regional reanalyses. Advances in Scientific Research, 12, 207-218.

Brinckmann S, Krähenmann S, Bissolli P. 2016. High-resolution daily gridded data sets of air temperature and wind speed for Europe. Earth System Science Data, 8, 491-516.

Brocca L, Ciabatta L, Massari C, Moramarco T, Hahn S, Hasenauer S, Kidd R, Dorigo W, Wagner W, Levizzani V. 2014. Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. Journal of Geophysical Research (Atmospheres), 119, 5128-5141.

Broufas GD, Pappas ML, Koveos DS. 2009. Effect of relative humidity on longevity, ovarian maturation, and egg production in the olive fruit fly (Diptera: Tephritidae). Annals of the Entomological Society of America, 102, 70-75.

Cornes RC, van der Schrier G, van den Besselaar EJM, Jones PD. 2018. An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. Journal of Geophysical Research (Atmospheres), 123, 9391-9409.

Dee DP et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137, 553-597.

Fonseca AR, Santos JA. 2018. High-resolution temperature datasets in Portugal from a geostatistical approach: variability and extremes. Journal of Applied Meteorology and Climatology, 57, 627-644.

Fontes N, Martins J, Graça A. 2016. Study of agrometeorological measurements on “Terroirs” of Alentejo wine region: Impact on grape yield and harvest date variation. Acta 10° Simpósio de Vitivinicultura do Alentejo (Vol.2), At Évora, Portugal, 4-6 May.

Frick C, Steiner H, Mazurkiewicz A, Riediger U, Rauthe M, Reich T, Gratzki A. 2014. Central European high-resolution gridded daily data sets (HYRAS): Mean temperature and relative humidity. Meteorologische Zeitschrift, 23, 15-32.

Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A, Michaelsen J. 2015a. The climate hazards infrared precipitation with stations – a new environmental record for monitoring extremes. Scientific Data 2:150066, doi:10.1038/sdata.2015.66.

Funk C, Verdin A, Michaelsen J, Peterson P, Pedreros D, Husak G. 2015b. A global satellite assisted precipitation climatology. Earth System Scientific Data 7, 275-287.

García 2003.

Gelaro R. et al. 2017. MERRA-2 Overview: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Journal of Climate, 30, 5419-5454.

Harada Y, Kamahori H, Kobayashi C, Endo H, Kobayashi S, Ota Y, Onoda H, Onogi K, Miyaoka K, Takahashi K. 2016. The JRA-55 Reanalysis: Representation of atmospheric circulation and climate variability. Journal of the Meteorological Society of Japan, 94, 269-302.

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high-resolution gridded dataset of surface temperature and precipitation. Journal of Geophysical Research (Atmospheres), 113, D20119, doi:10.1029/2008JD10201.

Herrera S, Fernández J, Gutiérrez JM. 2016. Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: assessing the effect of the interpolation methodology. International Journal of Climatology, 36, 900-908.

Hofstra N, Haylock M, New M, Jones PD. 2009. Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. Journal of Geophysical Research D: Atmospheres, 114 (21), doi: 10.1029/2009JD011799.

Huld T, Urraca R, Gracia Amillo A, Trentmann J. 2017. A global hourly solar radiation data set using satellite and reanalysis data. Proceedings of the 33rd European Photovoltaic Solar Energy Conference and Exhibition, pp.2458-2462.

Jones PD, Lister DH, Harpham C, Rusticucci M, Penalba O. 2012. Construction of a daily precipitation grid for southeastern South America for the period 1961-2000. International Journal of Climatology, 33, 2508-2519.

Kalamantios R, Kermanidis K, Avlonitis M, Karydis I. 2016. Environmental impact on predicting olive fruit fly population using trap measurements. IFIP International Conference on Artificial Intelligence Applications and Innovations, Springer, pp. 180-190.

Kanamitsu M, Ebisuzaki W, Woollen J, Yang S-K, Hnilo JJ, Fiorino M, Potter GL. 2002. NCEP-DOE AMIP-II Reanalysis (R-2). Bulletin of the American Meteorological Society, 83, 631-1643.

Kew SF, Philip SY. 2019. The Exceptional Summer Heat Wave in Southern Europe 2017, in: “Explaining Extremes of 2017 from a Climate Perspective”, Bulletin of the American Meteorological Society, 100(1), S49–S54, doi:10.1175/BAMS-D-18-0135.1.

Klok EJ, Klein Tank AMG. 2009. Updated and extended European dataset of daily climate observations. International Journal of Climatology, 29, 1182-1191.

Kobayashi S, Ota Y, Harada Y, Ebita A, Moriya M, Onoda H, Onogi K, Kamahori H, Kobayashi C, Endo H, Miyaoka K, Takahashi K., 2015. The JRA-55 Reanalysis: General specifications and basic characteristics. Journal of the Meteorological Society of Japan, 93, 5-48.

Kotlarski S, Keuler K, Christensen OB, Colette A, Déqué M, Gobiet A, Goergen K, Jacob D, Lüthi D, van Meijgaard E, Nikulin G, Schär C, Teichmann C, Vautard R, Warrach-Sagi K, Wulfmeyer V. 2014. Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. Geoscientific Model Development, 7, 1297-1333.

Kyselý J, Plavcová E. 2010. A critical remark on the applicability of E-OBS European gridded temperature data set for validating control climate simulations. Journal of Geophysical Research (Atmospheres), 115, D23, doi: 10.1029/2010JD014123.

Ledesma JLJ, Futter MN. 2017. Gridded climate data products are an alternative to instrumental measurements as inputs to rainfall–runoff models. Hydrological Processes, 31, 3283-3293.

Ma Y, Guttorp P. 2012. Estimating daily mean temperature from synoptic climate observations. International Journal of Climatology, 33, 1264-1269.

Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG. 2012. An overview of the Global Historical Climatology Network-Daily database. Journal of Atmospheric and Oceanic Technology, 29, 897-910.

Müller R, Pfeifroth U, Träger-Chatterjee C, Cremer R, Trentmann J, Hollmann R. 2015. Surface Solar Radiation Data Set - Heliosat (SARAH) - Edition 1, Satellite Application Facility on Climate Monitoring, doi:10.5676/EUM_SAF_CM/SARAH/V001. https://doi.org/10.5676/EUM_SAF_CM/SARAH/V001.

National Oceanic and Atmospheric Administration. 1998. Automated Surface Observing System (ASOS) User’s Guide. https://www.weather.gov/asos/

Perry M, Hollis D. 2005. The generation of monthly gridded datasets for a range of climatic variables over the United Kingdom. International Journal of Climatology, 25, 1041-1054.

Razavi AR, Nassiri-Mahallati M, Koocheki A, Beheshti A. 2018. Applicability of AgMERRA for gap-filling of Afghanistan in-situ temperature and precipitation data. Journal of Water and Soil, 32, 601-616.

Rienecker MM et al. 2011. MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications. Journal of Climate, 24, 3624-3648.

Ruane AC, Goldberg R, Chryssanthacopoulos J. 2015. AgMIP climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. Agricultural and Forest Meteorology, 200, 233-248.

Thrasher B, Maurer EP, McKellar C, Duffy PB. 2012. Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. Hydrology and Earth System Sciences, 16, 3309-3314.

Toreti A. 2014. Gridded agro-meteorological data in Europe. European Commission, Joint Research Centre (JRC) [Dataset] PID: http://data.europa.eu/89h/jrc-marsop4-7-weather_obs_grid_2015.

Toreti A, Maiorano A, De Sanctis G, Webber H, Ruane AC, Fumagalli D, Ceglar A, Niemeyer S, Zampier M. 2018. Using reanalysis in crop monitoring and forecasting systems. Agricultural Systems, doi:10.1016/j.agsy.2018.07.001.

Van Weverberg K, De Ridder K, Van Rompaey A. 2008. Modeling the contribution of the Brussels heat island to a long temperature time series. Journal of Applied Meteorology and Climatology, 47, 976-990.

Willett KM, Dunn RJH, Thorne PW, Bell S, de Podesta M, Parker DE, Jones PD, Williams Jr, CN. 2014. HadISDH land surface multi-variable humidity and temperature record for climate monitoring. Climate of the Past, 10, 1983-2006.

ANNEX A. WEATHER STATION OBSERVATIONS

The sources of weather stations in each of the three study regions are described in sections A.1 – A.3. Other sources of weather station data are summarized in section A.4.

A.1. WEATHER STATIONS IN ANDALUCÍA, SPAIN

DCOOP supplied observations from seven stations in Andalucía, southern Spain (Table A-1). The data series start between December 1999 and April 2001, and continue to the present day (June 2018 at the time of writing). The series are very complete, with few if any missing data. A useful range of meteorological variables are available:

- Daily Precipitation (mm)
- Daily maximum, mean, and minimum temperatures (plus the times at which the maxima and minima occurred).
- Daily maximum, mean, and minimum relative humidity
- Wind speed and direction
- Solar radiation
- (Potential?) Evapotranspiration

Table A-1: Names and locations of weather stations operated by DCOOP

Station	Latitude	Longitude	Altitude / m
Sevilla	37.457	-5.925	37
Malaga	36.756	-4.537	68
Jaen	37.891	-3.771	299
Huelva	37.412	-7.060	169
Granada	37.416	-3.551	935
Cordoba	37.857	-4.803	117
Almeria	36.835	-2.402	22

A.2. WEATHER STATIONS IN THE DOURO VALLEY, PORTUGAL

SOGRAPE were able to supply weather observations from several different sources in the Douro valley region. These sources are summarised in Table A-2.

Table A-2: Summary of weather station data supplied by SOGRAPE

Source	Number of Stations	Data Period	Comments
SOGRAPE	10	2011 - 2017	Very complete data series; there are only a small number of days with missing data.
DRAPN (Manual)[1]	21	1964 - 2011	Many missing data between 1964 and 1980 for almost all stations.
DRAPN (Automatic)[1]	48	2006 - 2017	Data from some stations only available for 5 years or less.
IPMA[2]	9	1981 - 2010	Data periods vary between stations. Few missing data.
EDP[3]	31	1961 - 2018	Start dates vary between 1961 and 2014. Very few missing data. Some stations may only record rainfall.

1. DRAPN - Regional Direction of Agriculture and Fisheries – North.
2. Portuguese Institute of Sea and Atmosphere. Many more stations are present in this network, but the data must be purchased.
3. EDP-Produção (next to dams)

A.3. WEATHER STATIONS IN ITALY

Daily observations from 3 stations were supplied by IBIMET. The data series are very complete with few missing values. Additional stations (section 7.4) were analysed for the Europe-wide comparison.

Table A-3: Summary of weather stations operated by IBIMET in the study regions in Italy

Station	Latitude	Longitude	Data Period
Foggia	41.50	15.52	2013-2018
Ancona	43.33	13.58	2007-2018
Ravenna	44.41	12.20	2007-2016

A.4. OTHER SOURCES OF WEATHER OBSERVATIONS

ECA&D - The ECA&D dataset consists of daily observations from meteorological stations throughout Europe and the Mediterranean (Klok and Klein Tank, 2009). The numbers of meteorological variables recorded and data lengths vary considerably between stations. For some countries, data from only a small proportion of possible weather stations have been supplied to ECA&D.

ISD – The Integrated Surface Database consists of global hourly and synoptic observations compiled from numerous sources. The database includes over 35,000 stations worldwide, with some having data as far back as 1901, though the data show a substantial increase in volume in the 1940s and again in the early 1970s. Currently, there are over 14,000 "active" stations that are updated on a daily basis in the database.

GHCN - The Global Historical Climatology Network is an integrated database of climate summaries from land surface stations across the globe (Menne et al., 2012). Daily and monthly data are available. The climatic variables available varies considerably between stations. In Colombia, only daily minimum, mean and maximum temperatures, and daily rainfall totals were available. A wider range of variables are available from stations in some other countries.

ASOS - Data from Automated Surface Observing Systems (ASOS) worldwide are archived by the Iowa Environmental Mesonet (IEM), which is part of Iowa State University. ASOS networks can be selected at the following link, and the data may be downloaded via the web interface or by running a short program in Python or R:

https://mesonet.agron.iastate.edu/request/download.phtml

The ASOS data contain values of a wide range of climatic variables – surface air temperature, dew point temperature, relative humidity, wind speed and direction, gust speed, sea level pressure, visibility, and cloud cover at four different altitudes. The temperatures are listed in Fahrenheit, but appear to have been converted from Centigrade as the values are not continuous, but have discrete values. A daily mean temperature could be estimated from the hourly values, although a number of different methods are possible (Ma and Guttorp, 2012). Estimates of daily maximum and minimum temperatures from hourly data are likely to be erroneous, as the true maxima and minima need not occur at the time the temperature was recorded and archived.

ANNEX B. GRIDDED DATASETS

A brief summary of the gridded datasets assessed is given below. The datasets in section B.1 were created by interpolation of surface-based weather observations. Gridded rainfall datasets derived from satellite data are summarised in B.2. The remainder of the gridded datasets in section B.3 are reanalyses, created by integrating a weather forecast model forwards in time and tightly constrained it with many different types of observations. All of the reanalyses have global coverage, except for the regional reanalyses of UERRA.

B.1. GRIDDED DATASETS DERIVED FROM SURFACE OBSERVATIONS

E-OBS

E-OBS is a European land-only gridded dataset derived from over 2,000 weather stations (Haylock et al., 2008). Data are available (at the time of writing) from 1950 to 2018. E-OBS consists of daily values of mean temperature, minimum temperature, maximum temperature, total precipitation and averaged sea level pressure. E-OBS data are available over the land area within the region 25°N-75°N, 40°W-75°E. The E-OBS data are available on four different resolutions; for MED-GOLD, data on the same rotated grid used by the CORDEX simulations, at a resolution of 0.22°, were used. A version of E-OBS at 0.11º (about 11 - 12 km) does not exist, at the time of writing. In this report, data from version 17.0 of E-OBS were used.

Spain02

Spain02 is a high-resolution (0.1º, ~10 km) gridded dataset developed for Spain and the Balearic islands (Herrera et al., 2016). It contains daily precipitation totals and daily maximum and minimum temperatures for the period 1951-2015. Spain02 was derived from ~2500 quality-controlled stations (~250 stations for temperatures). There is also a EURO-CORDEX compliant version on a 0.11º rotated grid for the purpose of comparison with regional climate model results.

CGMS / MARSMET

CGMS / MarsMet was created by JRC. It has similarities to E-OBS, but has a larger domain. It is a land-only gridded dataset derived from around 4500 stations within Europe, western Russia and northern Africa (the number of stations has generally increased since 1975). CGMS has a resolution of 25 km, and is briefly described in Toreti (2014) and Toreti et al. (2018).

DOUROZONE

DOUROZONE is an EU project developed by the University of Aveiro by DAO and CESAM, with the objective of assessing the risk of vines to ozone exposure in the Douro Wine Region, for the present climate and for a future scenario. In order to carry out the project, Marta-Almeida et al. (2016) developed a set of high-resolution climate simulations, using the WRF model forced by ERA-Interim reanalysis data and a global climate model (MPI-ESM-LR). The future climate scenario used in DOUROZONE was RCP8.5. With the WRF model, high resolution climate simulations for three different intervals were created: 1986-2005, 2046-2065 and 2081-2100. The horizontal resolution was 9 km for the demarcated region of the Douro valley. Temperature and precipitation data were bias-corrected using E-OBS data. In order to study specific years characterized by heat waves, additional dynamic downscaling down to 1 km was carried out with WRF, for the following years:

- WRF with ERA- Interim: 2003 to 2005, 2017;
- WRF with MPI-ESM-LR: 2000, 2049, 2064, 2096, 2097

PTHRES

PTHRES was produced during the INNOVINE & WINE project for Portugal, with a very high horizontal resolution of 1 km. PTHRES contains average, minimum and maximum temperatures for the domain, based on data from E-OBS for 1950 to 2015. PT.HRES was based on additional observational data to those used within E-OBS, thus filling in gaps in areas where there are no meteorological stations. At the time of writing, a similar dataset for daily precipitation amounts was not available.

B.2. SATELLITE-BASED GRIDDED DATASETS

CHIRPS

Climate Hazards group Infrared Precipitation with Stations (CHIRPS) is a high resolution precipitation dataset derived from TRMM rainfall estimates further calibrated with surface rainfall gauges (Funk et al., 2015a,b). Data are available from 1981 onwards, on daily, 5-day and monthly timescales. CHIRPS lags real time by 1-2 months. The rainfall data have a spatial resolution of 0.05° (~5.5 km). CHIRPS data are only available between latitudes 50°N - 50°S.

SM2RAIN

SM2RAIN is a satellite-based rainfall product (Brocca et al., 2014). Most satellite-based rainfall data are retrieved from the atmospheric signals reflected or radiated by atmospheric hydrometeors. SM2RAIN, however, is derived from satellite-based soil moisture measurements.

SARAH

The Surface Solar Radiation Data Set - Heliosat (SARAH) is a satellite-based climatology of the solar surface irradiance, the surface direct normalized irradiance and the effective cloud albedo derived from satellite observations of the visible channels of the MVIRI and SEVIRI instruments onboard the geostationary Meteosat satellites. The data are available from 1983 to 2013 and cover the region ±65° longitude and ±65° latitude. The products are available as monthly, daily, and hourly averages on a regular latitude/longitude grid with a spatial resolution of 0.05° x 0.05°.

See: https://wui.cmsaf.eu/safira/action/viewDoiDetails?acronym=SARAH_V001

B.3. REANALYSIS DATASETS

AgMERRA

AgMERRA is a climate forcing dataset for agricultural modelling (Ruane et al., 2015). AgMERRA was produced by calibrating a reanalyses dataset (NASA's Modern-Era Retrospective analysis for Research and Applications, MERRA) with observational datasets from in-situ observational networks and satellites. AgMERRA has a resolution of 0.25° × 0.25° (~25km), with global coverage. Daily values are available for the period 1980-2010.

NCEP2

NCEP-DOE Reanalysis 2 (NCEP2; Kanamitsu et al., 2002) is an improved version of an earlier reanalysis (NCEP Reanalysis I). NCEP2 is available from 1979, and is continually updated in near real-time. It has global coverage at low resolution (2.5° × 2.5°). The data are available four times a day, as well as daily and monthly values.

EIN75

EIN75 refers to data from the ERA-Interim reanalysis (Dee et al., 2011) at 0.75° resolution. ERA-Interim is a global atmospheric reanalysis which begins in 1979, and is continuously updated in real time. It will eventually be replaced by a newer reanalysis, ERA5.

ERA5

ECMWF Reanalysis 5 (ERA5) is a newer reanalysis which will eventually replace ERA-Interim. Data from ERA5 are currently available from 2008 and are updated in near real time. ERA5 will also be extended back in time to 1950 during 2019.

NEX

NASA Earth Exchange (NEX) Global Daily Downscaled Projections (NEX-GDDP). This dataset is comprised of downscaled and bias-corrected climate projections derived from the General Circulation Model (GCM) runs conducted under the Coupled Model Intercomparison Project Phase 5. More details are provided in Thrasher et al. (2012) and:

https://nex.nasa.gov/nex/projects/1356/

JRA55

The Japanese 55-year Reanalysis (JRA55; Kobayashi et al., 2015; Harada et al., 2016) is available from 1958. JRA55 has a longer time period than the other reanalyses. However, it is known to contain discontinuities when certain datasets became available (e.g. around 1979 when satellite data first became available).

MERRA2

The Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2; Gelaro et al., 2017) provides data beginning in 1980. It was introduced to replace the original MERRA dataset owing to advances made in the data assimilation system that enabled use of newer observations. It is not known whether an updated version of AgMERRA using the newer MERRA2 data will be created.

UERRA

The aim of the Uncertainties in Ensembles of Regional Reanalyses (UERRA) project is to produce ensembles of European regional meteorological reanalyses of Essential Climate Variables for several decades and to estimate the associated uncertainties in the data sets (Borsche et al., 2015). Three different regional climate models were used within UERRA. Data are available for 1961-2013, but the exact period depends on the model. In the figure captions in section 5.1, the four letters following ‘UERRA’ refer to the regional climate models: eggr -> UM; eswi -> HARMONIE; lfpw -> MESCAN-SURFEX.

END OF DOCUMENT