

A dataset of future daily weather data for crop modelling over Europe derived from climate change scenarios

G. Duveiller¹  · M. Donatelli² · D. Fumagalli¹ · A. Zucchini¹ · R. Nelson³ · B. Baruth¹

Received: 11 January 2015 / Accepted: 24 September 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Coupled atmosphere-ocean general circulation models (GCMs) simulate different realizations of possible future climates at global scale under contrasting scenarios of land-use and greenhouse gas emissions. Such data require several additional processing steps before it can be used to drive impact models. Spatial downscaling, typically by regional climate models (RCM), and bias-correction are two such steps that have already been addressed for Europe. Yet, the errors in resulting daily meteorological variables may be too large for specific model applications. Crop simulation models are particularly sensitive to these inconsistencies and thus require further processing of GCM-RCM outputs. Moreover, crop models are often run in a stochastic manner by using various plausible weather time series (often generated using stochastic weather generators) to represent climate time scale for a period of interest (e.g. 2000 ± 15 years), while GCM simulations typically provide a single time series for a given emission scenario. To inform agricultural policy-making, data on near- and medium-term decadal time scale is mostly requested, e.g. 2020 or 2030. Taking a sample of multiple years from these unique time series to represent time horizons in the near future is particularly problematic because

selecting overlapping years may lead to spurious trends, creating artefacts in the results of the impact model simulations. This paper presents a database of consolidated and coherent future daily weather data for Europe that addresses these problems. Input data consist of daily temperature and precipitation from three dynamically downscaled and bias-corrected regional climate simulations of the IPCC A1B emission scenario created within the ENSEMBLES project. Solar radiation is estimated from temperature based on an auto-calibration procedure. Wind speed and relative air humidity are collected from historical series. From these variables, reference evapotranspiration and vapour pressure deficit are estimated ensuring consistency within daily records. The weather generator ClimGen is then used to create 30 synthetic years of all variables to characterize the time horizons of 2000, 2020 and 2030, which can readily be used for crop modelling studies.

1 Introduction

There is a general consensus in the scientific community that the change in the climatic system of the Earth is unequivocal (IPCC 2013). Agriculture is one of the sectors most vulnerable to climate change. In Europe, future impacts of climate change on agriculture can be generalized by a northward movement of crop suitability, with increased productivity in Northern Europe and a decline in both productivity and suitability in Southern Europe (Olesen et al. 2002; Maracchi et al. 2005; Falloon and Betts 2010). However, it is also foreseen that there will be an increase in extreme events, such as the heat waves over Europe of 2003 and 2010 (Schär et al. 2004; Barriopedro et al. 2011; Russo et al. 2014). These shifts and changes will offer opportunities and challenges requiring adaptation of European agriculture to the changing environment. In order to implement appropriate policies, appropriate tools

M. Donatelli was at the European Commission Joint Research Centre while developing this study.

✉ G. Duveiller
gregory.duveiller@jrc.ec.europa.eu

¹ Institute for Environment and Sustainability, European Commission Joint Research Centre, Viale E. Fermi, 2479 Ispra, Italy

² Consiglio per la Ricerca e sperimentazione in Agricoltura, CRA-CIN, Bologna, Italy

³ Washington State University, Biological Systems Engineering, Pullman, WA, USA

are required to characterize spatially the vulnerability of its agriculture based on future climate predictions.

Deterministic crop growth modelling is a major tool for analysing the impacts of climate change on agricultural production. Daily weather data is necessary to drive these models. The climate information required to estimate daily weather in the future typically originates from coupled atmosphere-ocean general circulation models (AOGCMs, but onwards referred to more simply as GCMs). These models can be used to generate projections of future climate based on scenarios, i.e. a coherent, internally consistent and plausible description of a possible future state of the world.

Some of the most commonly used climate change scenarios are those provided by the Special Report on Emissions Scenarios (SRES) (Nakićenović et al. 2000) of the Intergovernmental Panel on Climate Change (IPCC). These scenarios consider four main alternative storylines classified along two axes representing conditions where the world is either more globalized or regionalised, and whether it will be more environmentally or economically centred. More recently, the SRES have been replaced by representative concentration pathways (RCP) (van Vuuren et al. 2011) and with associated shared socio-economic pathways (SSPs) (O'Neill et al. 2013).

Climate change projections realized by running GCMs under different emission scenarios are intrinsically subject to a considerable amount of uncertainty. This has led the community to use various GCMs together, such as the Coupled Model Inter-comparison Project (CMIP) which provides an ensemble of model runs that are examined by the IPCC to produce their reports. The latest IPCC Fifth Assessment Report (AR5) is based on CMIP5.

There are several major obstacles to use information from GCMs to generate daily weather for crop models. There is first a serious challenge due to the significant differences in spatial and temporal scales between GCMs and crop growth models (Hansen et al. 2006). Despite an increasing ability of GCMs to successfully model present-day climate and provide realistic quantitative predictions of climate change at a continental scale, GCMs still have serious difficulties in reproducing accurate daily estimates at local scale. This is especially true for rainfall variability: while a GCM may estimate monthly precipitation correctly, the daily precipitation may be spread throughout the month in a highly unrealistic way (e.g. raining a little every day). Such distortions of daily weather variability can seriously bias crop model simulations (Semenov and Porter 1995; Mearns et al. 1996; Hansen and Jones 2000; Baron et al. 2005). The need for bias correcting GCM-RCM projections before using them to drive impact models is well known, e.g. Christensen et al. (2008), and the influence of such biases on hydrological and crop modelling has been extensively investigated by, e.g. Teutschbein and Seibert (2010),

who claimed that unless climate model outputs are corrected, their application in impact modelling may be unrealistic. Furthermore, it is important to note that the choice of the bias-correction method itself can influence the results as much as differences between GCMs (Hawkins et al. 2013).

There is a need to bridge this gap between the data provided by the climate community and the requirements of crop growth models. The objective of this paper is to present a dataset of future weather data covering Europe that (1) is suitable to use as input data for crop growth simulation models and (2) can be used for impact studies targeting the short-term time horizons (such as 2020 and 2030) that are highly relevant for policy-making. The application of the dataset for crop modelling is described in another paper (Donatelli et al. 2015).

2 Material and methods

2.1 Future weather data

The future weather dataset presented in this paper is based on simulations generated by different GCMs that have been dynamically downscaled using regional climate models (RCM). The original GCMs simulations are publicly available from the European ENSEMBLES project (Van der Linden and Mitchell 2009; <http://ensemblesrt3.dmi.dk/>). This data was bias-corrected (Dosio and Paruolo 2011; Dosio et al. 2012) by applying an extension of the technique proposed by earlier studies (Piani et al. 2009; Piani et al. 2010) in which both precipitation and temperature values are corrected independently with respect to the E-OBS dataset (Haylock et al. 2008). From this bias-corrected dataset, simulations of the A1B scenario by different coupled GCM-RCM models were selected (from now onwards, when the notation GCM is used in this paper, it refers to the bias-corrected dataset). In the present work, only the A1B SRES is considered because short-term time horizons (2020 and 2030) are targeted, and at such time scales, differences with other scenarios are modest with respect to temperature. A1B is also a scenario that has been widely used in simulation studies representing one of the possible high impact scenarios at later time scales deriving from GHG emissions. The first realization of the A1B scenario consists of simulations of the ECHAM5 GCM coupled with the HIRHAM5 RCM (Christensen et al. 2006) and run by the Danish Meteorological Institute (henceforth denoted DMI-HIRHAM5-ECHAM5). The second model, run by the Swiss Federal Institute of Technology, nests CLM (Jaeger et al. 2008) in the HadCM3 GCM (denoted ETHZ-CLM-HadCM3Q0). The third realization also uses the same HadCM3 but this time coupled with the HadRM3Q0 (Collins et al. 2011) by the UK Met Office Hadley Centre for Climate Prediction and Research (denoted as METO-

HC-HadRM3Q0-HadCM3Q0).¹ The rationale behind this choice of models was to (1) select widely used GCMs and (2) maximize diversity in terms of a weather variable of interest. Within the A1B realizations in the ENSEMBLES project that have been bias-corrected (Dosio et al. 2012), the DMI-HIRHAM5-ECHAM5 and METO-HC-HadRM3Q0-HadCM3Q0 simulations are respectively the coldest and the warmest in terms of surface air temperature. ETHZ-CLM-HadCM3Q0 is added to provide an intermediate realization with markedly different precipitations patterns. Daily data for these three A1B realizations are selected for the time period running from 1993 up to 2037, in order to group them into three time horizons of 15 years: (1) a baseline period around 2000 (1993–2007), (2) horizon 2020 (2013–2027), and horizon 2030 (2023–2037). The choice of time horizons of 15 years, rather than the more conventional 30 years typically used in climate modelling, stems from the practical request from agricultural policy-makers to deliver information on the near-to-medium decadal time scales, i.e. 2020 and 2030. The weather generator, described in section 3.2, is subsequently used to recreate weather for a longer time period. All data were also resampled using a nearest-neighbour interpolation over a common 25 by 25 km grid, which is the grid of the reference observed weather dataset described in the next section.

2.2 Historical weather data

Observed weather data over the baseline period (centred around 2000) is required in this work for two reasons: (1) to complement the future dataset with some variables, as is explained further on, and (2) to assess how well the time series generated by the GCM represents the historical climate. The rationale behind selecting 2000 as a baseline period, instead of a more common period, such as 1961–1990, is partly to ensure a higher reliability and completeness of meteorological records. However, since the dataset in this study is intended for studies in agriculture, the choice of the baseline also aims at having a benchmark period in which agricultural practices are similar to those employed currently and in the near future. The weather database selected for this task was provided by the MARS Crop Yield Forecasting System (MCYFS) (Micale and Genovese 2004). The MCYFS is an operational decision support system run by the Joint Research Centre of the European Commission for the past 20 years to provide in-season crop yield forecasts at a European level. The estimates from the system are used in the decision-making process on market intervention and for policy support. The MYCFS weather database is based on daily weather observations from meteorological stations that are spatially interpolated into a 25

by 25 km grid. A base of circa 4000 weather stations over Europe is available from which only those satisfying reliability criteria for near-real-time delivery are used as data sources. The MYCFS data is used, as detailed in the following paragraph and with reference to time horizon centred on year 2000, to

- Evaluate GCM-RCM estimates of surface air temperature and rainfall,
- Evaluate solar radiation estimates, and
- Use wind and air relative humidity data.

The data is also used as support to evaluate the capability a weather generator to reproduce time series starting from those simulated by the GCMs (in section 3.3).

3 Methods

The methodology to construct the future weather dataset for Europe is described in the following sections and summarized in Fig. 1.

3.1 Complementing the weather dataset with biophysical variables necessary for crop simulation

Despite the bias-corrections, the future weather dataset proposed by Dosio and Paruolo (2011) is still inadequate to properly run process-based crop simulation models to assess climate change impacts on crop growth and yield. The main issue is the lack of consistency of weather variables resulting from the fact that the bias-correction is done only on a subset of the necessary variables: surface air temperature and rainfall. Other required variables, such as global solar radiation and wind speed, may still have unrealistic distributions when compared to observed data. Other input variables for crop growth models, such as evapotranspiration, are not directly available and must thus be calculated. The solutions below have been adopted to consolidate the weather dataset.

3.1.1 Global solar radiation

To ensure global solar radiation is coherent with the bias-corrected temperature values, it has been estimated using the Bristow-Campbell model (Bristow and Campbell 1984). Such methods to estimate global solar radiation using daily surface air temperature amplitude are based on the assumption that the site is not significantly affected by advection. This assumption does not necessarily hold when estimating the solar radiation pattern of a specific site, but when working with abstractions such as interpolated time series associated to a spatial grid and when site-specific information is lacking, the assumption can be considered non-limiting. This is because the range-based

¹ More details on the models and their institutions can be found in (Christensen et al. 2010).

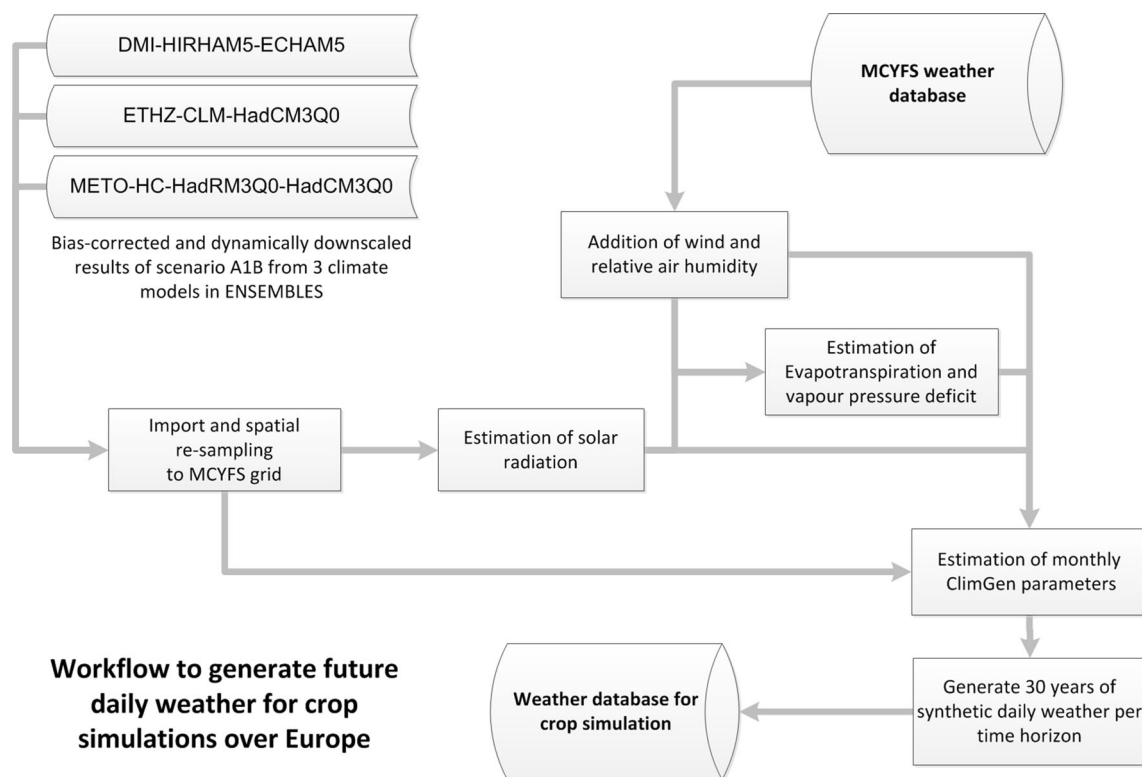


Fig. 1 Outline of the data processing workflow to generate the future daily weather for crop simulation presented in this paper

method is physically consistent: clear days show a greater range of temperature because solar irradiance is not filtered by clouds during the daytime, while the long-wave emission from soil surface is more rapidly lost in the atmosphere during the night. Seasonality is accounted for in the Bristow-Campbell model. The Bristow-Campbell model requires a continuous observed global solar radiation spanning at least 2 years for proper calibration; this is not available for all of Europe. Consequently, the auto-calibration procedure proposed by Bojanowski et al. (2013), which does not require reference data, was used. The auto-calibration method provides robust estimates of solar radiation which are consistent with temperature data. Given that GCM climate change scenarios do estimate changes in temperature, the Bristow-Campbell *b* parameter can be estimated for each scenario; solar radiation data can be derived accordingly. Clear sky transmissivity (CST) is estimated for each grid cell from remote sensing data as in Bojanowski et al. (2013). After CST was estimated, the *b* parameter is estimated keeping the value of the *c* parameter constant as $c = 2$. The uncertainties in the temperature estimates of GCM and RCMs can be propagated to solar radiation; however, it is beyond the scope of this study to articulate about the exogenous variables.

3.1.2 Wind speed and relative air humidity

Wind speed and relative air humidity are of direct interest for plant disease models and for the estimation of reference

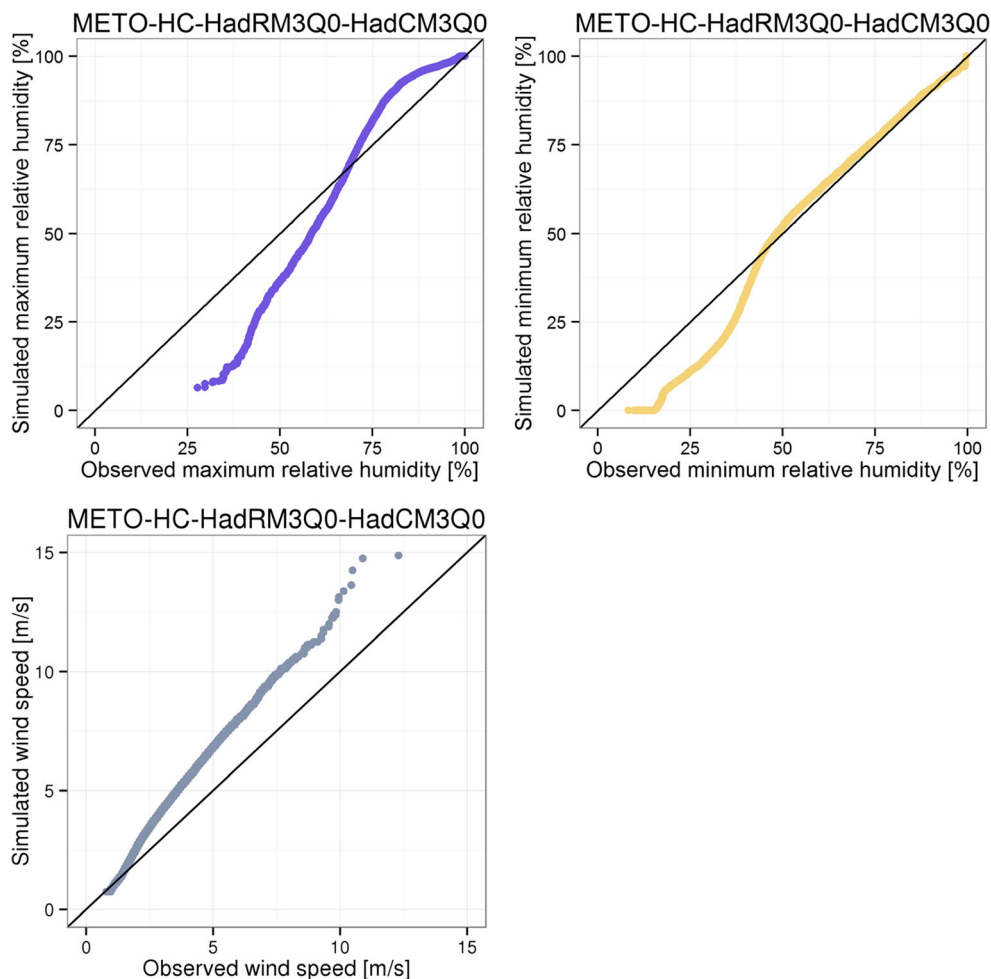
evapotranspiration based on models such as Penman-Monteith. As illustrated in Fig. 2, the distributions of these variables produced by GCMs are not realistic when compared to observations from MCYFS. The wind speed data contained some very excessive and unrealistic values in the METO-HC-HadRM3Q0-HadCM3Q0 model simulation. To avoid problems with the Q-Q plot, data above 15 m/s (less than 2 % of the total samples considered here) are discarded from the simulated data and the respective records are removed from the observed data. Even after this operation, wind speeds appear overestimated when compared to the observed data.

Some relationship between these variables and temperature and precipitation patterns exists, but it is much less straightforward to derive the former from the latter as it was done with global solar radiation. Here, it is conservatively assumed that patterns of wind speed and humidity should not change considerably in the near future. The historical observed data from MCYFS during the period 1993–2007 are thus used to represent all time horizons (the baseline, 2020 and 2030).

3.1.3 Evapotranspiration and vapour pressure deficit

Reference evapotranspiration and vapour pressure deficit were estimated from the GCMs and/or derived weather element variables using the FAO56 realization of the Penman-Monteith model (Allen et al. 1998) as implemented in the CLIMA libraries (Donatelli et al. 2006; Donatelli et al. 2009). While simpler temperature-driven empirical methods

Fig. 2 Comparison of maximum relative humidity (*top left*), minimum relative humidity (*top right*) and wind speed (*bottom left*) from gridded observations in the MCYFS weather database against those simulated by dynamically downscaled METO-HC-HadRM3Q0-HadCM3Q0 model. Data come from a subset of ten locations across Europe (indicated in Fig. 3), collected over the period going from 1 April till 30 September and over the years 1993–2007. For wind speed, simulated values above 15 m/s are discarded



can also be used, this physically based approach is preferred to ensure that the bio-meteorological values are consistent with the driving thermic, aerodynamic and radiative elements.

3.2 Processing the dataset for short-term time horizons

The data produced by GCMs (or RCMs) are often used to represent the general trend that climate variables such as temperature are expected to have. However, there is some variability around such a trend representing the weather patterns that are simulated by GCMs. For a given time horizon, climate studies will typically use datasets with 30 years or more to characterize a given variable or to derive biophysical variables from impact models (such as crop yields). Such sample size is deemed large enough so that the short-term random fluctuations—such as yearly weather pattern variations—do not significantly affect the characterization of the climate during the target time horizon.

Typically, climate studies distinguish time horizons that are well separated in time, e.g. 2020, 2050 and 2100. Using the larger sample size tempers statistical descriptors such that identified trends are not significantly influenced by short-

term anomalous fluctuations. If time horizons of interest are close in time, such as 2020 and 2030, taking windows of 30 years around these horizons results in an overlap that is too large, rendering the separation into two horizons not significant. Conversely, when considering only 10 years (thereby avoiding the overlap in the above-mentioned case of 2020 vs. 2030), the sample size becomes too small in order to assume that specific short-term weather fluctuations do not dominate the trend. Indeed, 3 or 4 years that are much warmer than the average during a period of 10 years will have stronger consequences on the average indicators derived by impact models, such as crop growth models, than if these 3–4 years occur within a period of 30 years.

A compromise was made to characterize the climate of the target time horizons by using a stochastic weather generator, ClimGen (Stöckle et al. 2001), to increase the sample size corresponding to each time horizon. We used 15 years of data around each time horizon to derive monthly parameters for the weather generator (WG); these parameters resume the distribution of each weather variable for each grid cell. These parameters are then used to generate a set of 30 synthetic years for every grid cell, which have the characteristics of the 15-

Fig. 3 Location of the selected 25 by 25 km grid cells from the MCYFS weather database which were used for the comparison with the generated future daily weather over the baseline period. The intensity of the green colour relates to the percentage of arable land in each grid cell according to the MCYFS database

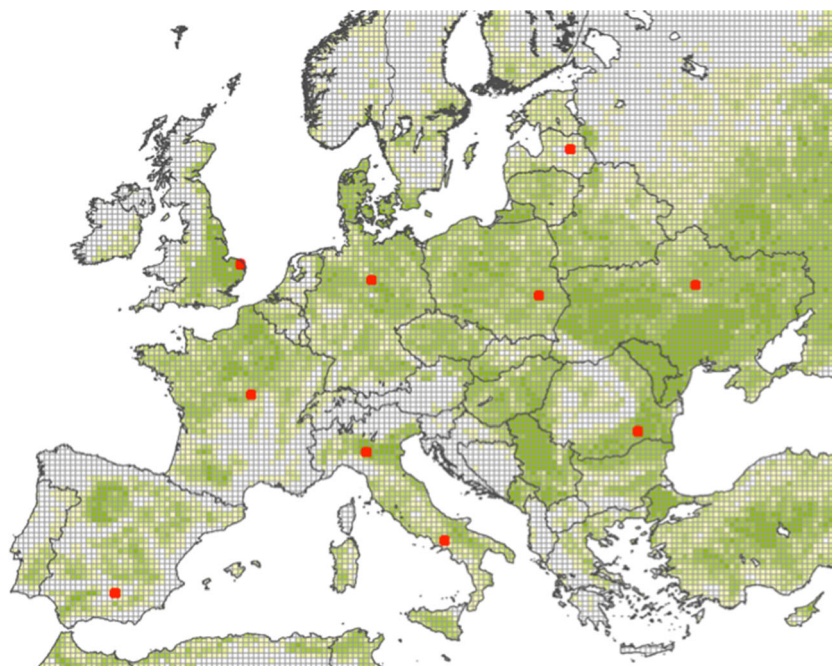
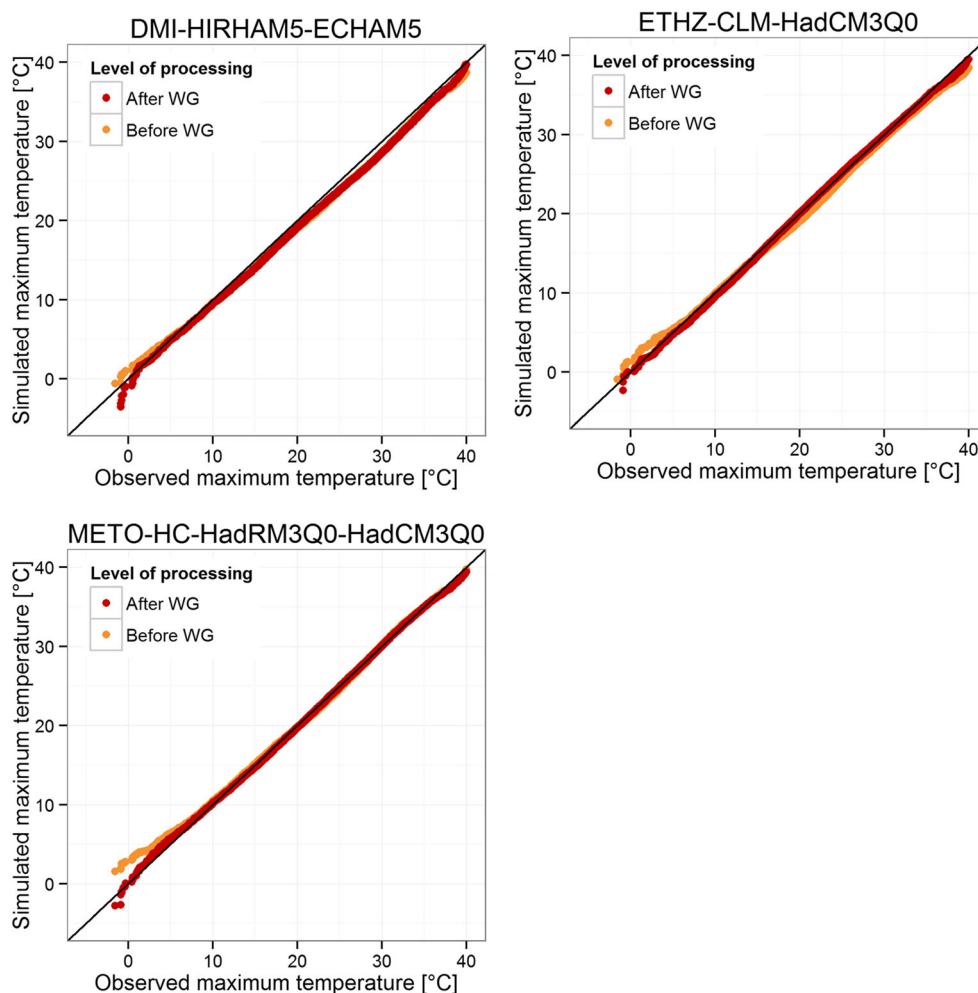


Fig. 4 Comparison of maximum temperature for the baseline period (centred on 2000) between observations from the MCYFS weather database and the simulated values from the three different bias-corrected and dynamically downscaled climate models. Two levels of processing are shown: before and after applying the weather generation (*WG*) step. Both data based on observation (the MCYFS database) and the WG data ranges from 1993 to 2012. WG years were randomly sampled from the 30 years generated



year period. Although the 15-year periods used to generate parameters do overlap by 4 years, the new synthetic years within each time horizon are distinct since they are regenerated randomly.

It must be acknowledged that due to the stochastic nature of the weather generation process, and to the fact that the generation is applied independently on every grid cell, there is a lack of spatial consistency: a generated value for a variable in any given cell on any given day will not necessarily be similar in value to variable's values in adjacent cells. In reality, there would generally be a continuum of the values between cells, if not throughout the region. This also applies to any biophysical variables derived from this synthetic weather dataset. The weather dataset built here is consequently targeted to be used with impact models in which runs are spatially independent. Spatially-continuous results are obtained after averaging all the simulated variables at each grid cells. Hence, results can only be analysed in terms of statistic properties, and not for investigating patterns of individual years.

3.3 Comparison against gridded observed weather data

A comparison between the observed MCYFS and the generated weather datasets for the overlapping period of the baseline serves to assess how well the dataset correlates to the past climate. Three evaluations were devised to ensure the cogency of the assumptions: (1) an assessment of the bias-correction process by analysing potential differences between the GCM-RCMs (the reference database MCYFS is indeed different from that of E-OBS used for the bias-correction); (2) an analysis of how the duration, or sample size, of generated weather time horizons (section 3.2) may affect results and (3) an evaluation of how well global solar radiation estimation is improved by the auto-calibration of the temperature-based model (section 3.1.1).

The evaluations were conducted on a subset of grid cells encompassing key sites representing regions in Europe with contrasting agro-ecological conditions (as shown in Fig. 3). In each case, every year in the baseline time horizon of 1993 to 2007 (2000 ± 7 years), was included, but only during a typical growing season ranging from 1 April up to 30 September.

Fig. 5 Same as Fig. 4 but for minimum temperature

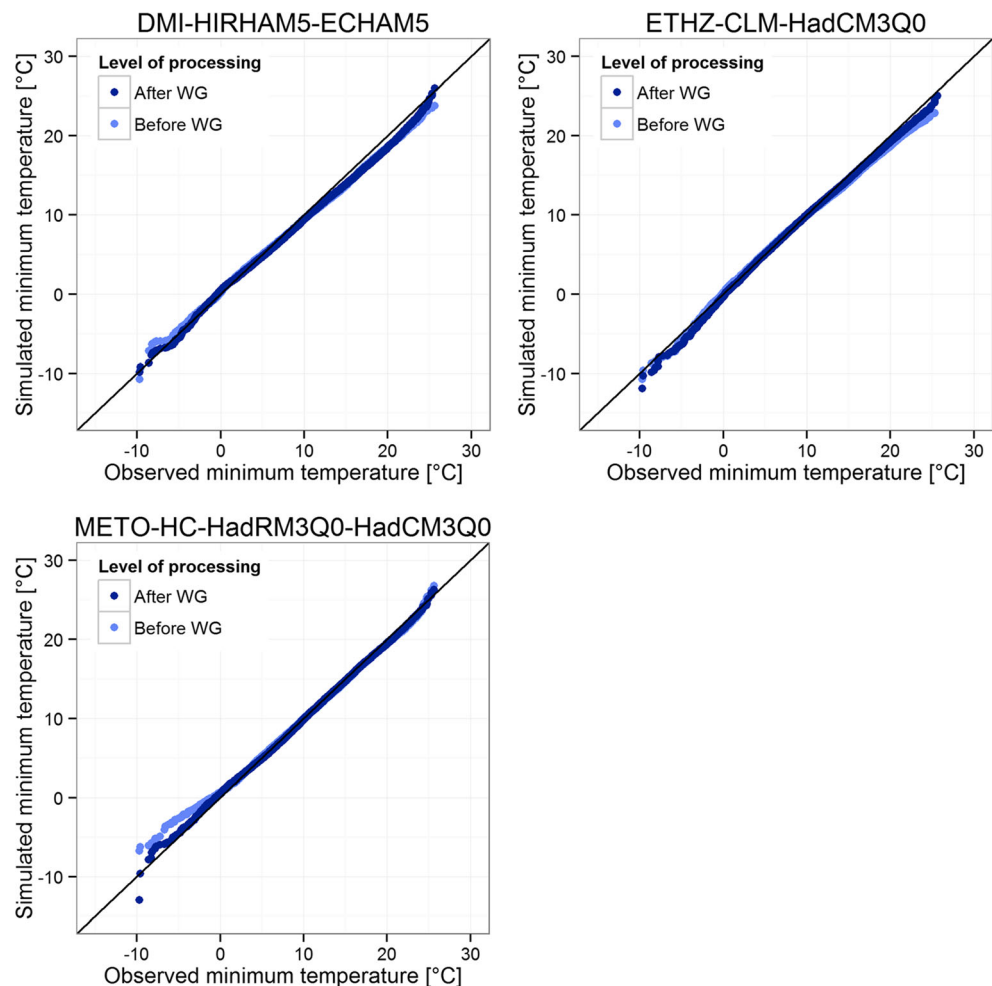
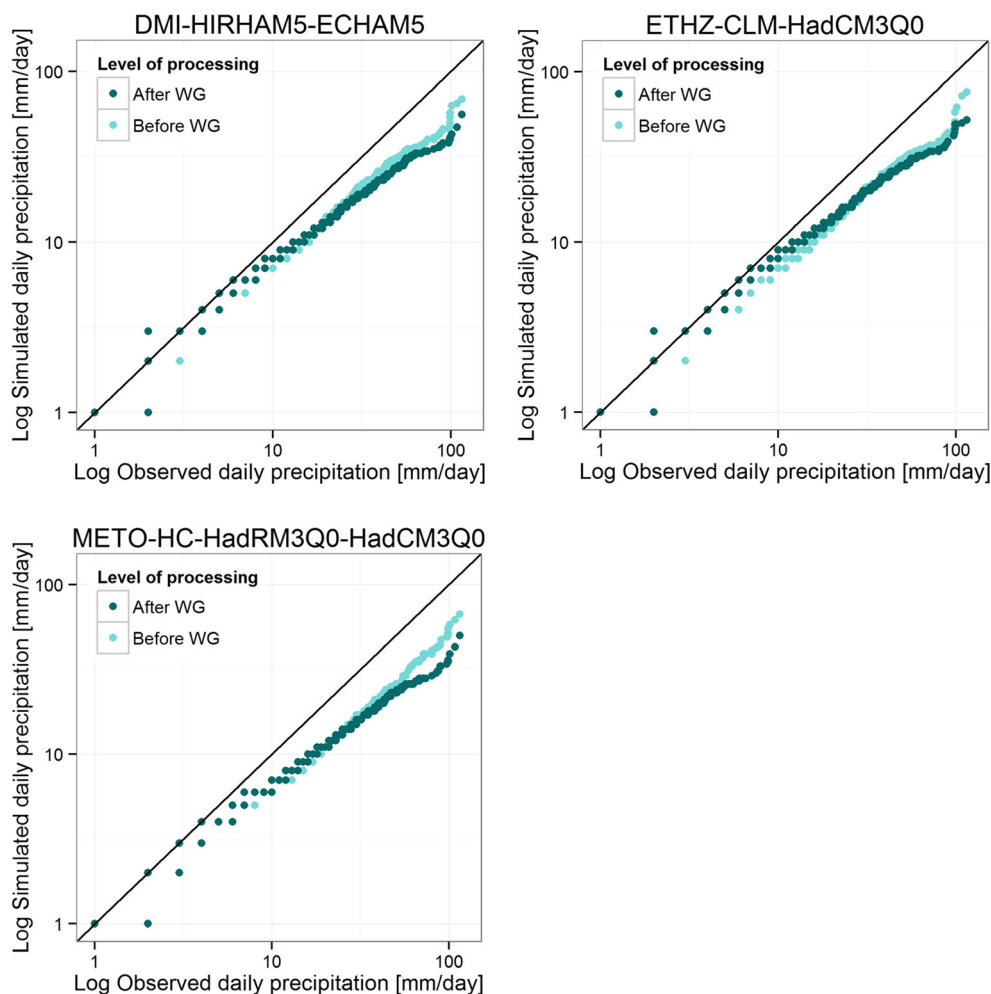


Fig. 6 Same as Fig. 4 but for daily precipitation



This period covering much of the spring and summer is considered most relevant for crop growth in Europe.

The bias-correction process evaluation was performed for each meteorological variable. Each variable was sorted by value irrespectively of spatial location for every GCM-RCM dataset separately. Each series was matched with a series obtained with the same procedure from data of the same cells in the MCYFS dataset. Both GCM and WG series were compared; in the case of WG data, 15 years were sampled from the 30 generated to have the same number of items in each series. This analysis quickly identifies if biases persist and if extreme values are correctly represented.

4 Results

4.1 Evaluation against observed data

The capacity of the generated weather to represent the climate in the baseline time horizon of 2000 is summarized by the Q-Q plots for maximum temperature, minimum temperature and daily precipitation in Figs. 4, 5 and 6, respectively.

Minimum and maximum temperature appears to be in line with the observed gridded MCYFS data. The weather

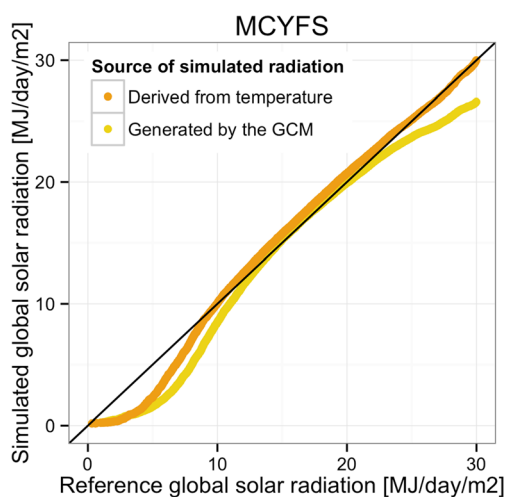


Fig. 7 Comparison of simulated solar radiation, either directly generated by the GCM or derived from temperature, against the reference source: MCYFS observed weather database

generation step even seems to slightly improve the representation of the lower temperature values in some cases.

For daily precipitation, the distributions of the uncorrected RCM and WG-simulated data do not quite match those of observed data. Over the selected grid cells, there is an overestimation of days with no precipitation (not visible on Fig. 6. because of the logarithmic scale) and an under-estimation of heavy precipitation.

For global solar radiation, the Q-Q plot presented in Fig. 7 shows how the values generated using the auto-calibrated Bristow-Campbell model are generally closer to measured data. Furthermore, the graphs in Fig. 7 do not show the improved consistency with the other variables. Figure 8 complements the picture by showing how the WG calculated global solar radiation is more coherent with respect to the daily temperature amplitude. This is expected as days with larger thermic amplitude are typically cloudless and thus allow more radiation to reach the land surface. This does not hold in situations in which substantial advection or convection phenomena occur, but in the absence of reference data as for future weather scenarios, the physically based assumption that the correlation between air temperature amplitude and radiation flux exists is more acceptable than the opposite.

4.2 Overview of the generated dataset

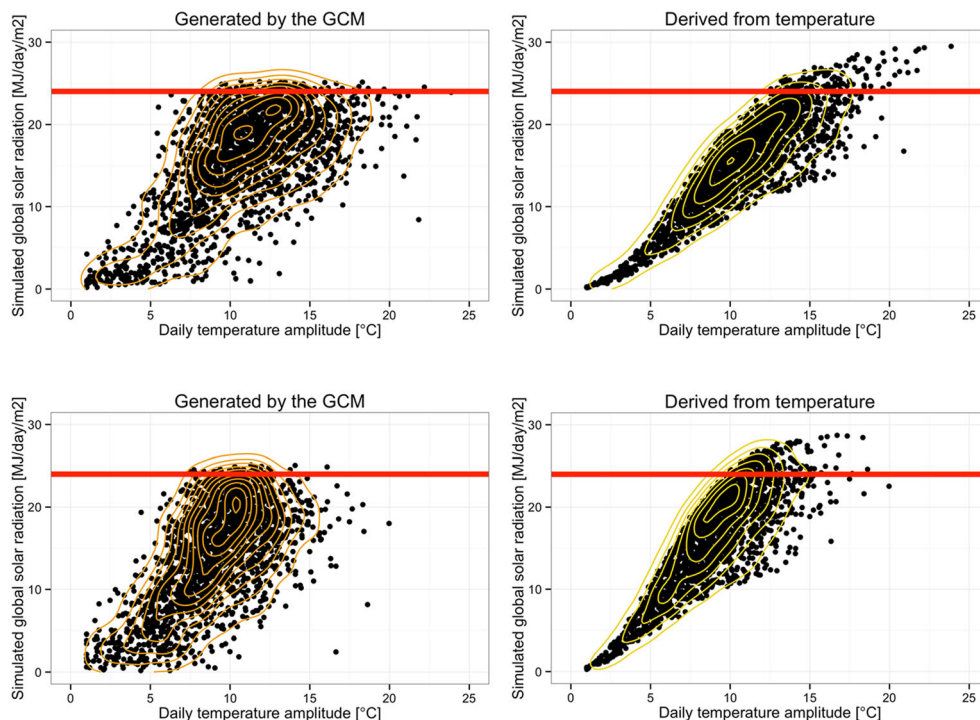
The general patterns of cumulated daily precipitation and average maximum temperature during the period comprising the months of April up to September (included) are presented to illustrate the generated dataset. The weather data variables are

respectively illustrated in Figs. 9 and 10, and consist of the average of 30 individual synthetic years with an additional spatial smoothing, using a 3 by 3 averaging filter to further increase the spatial consistency and facilitate the interpretation of the maps.

According to the three GCMs, under the A1B scenario, temperature is foreseen to rise steadily over most of Europe from the baseline in 2000 up to 2020 and then 2030. However, within the general pattern of surface air temperature increase, the areas with the estimated highest increase by 2030 differ according to GCMs: (i) Eastern Europe (DMI-HIRHAM5-ECHAM5), (ii) Southwestern Europe (ETHZ-CLM-HadCM3Q0), (iii) most of Europe excepting Italy, Southeastern Europe and the British Isles (METO-HC-HadRM3Q0-HadCM3Q0). The rise in minimum temperatures in April–September follows roughly the same patterns (not shown).

Precipitation patterns vary spatially, even to a noticeable extent, according to both the A1B realizations and the time horizons. With an exception of more humid zone spanning Ireland, Scotland, Scandinavia and Northern Russia, Europe is generally expected to become considerably drier over the period of April–September in the time horizon 2030. However, before reaching this rather dry situation in 2030, the data show that Europe is estimated to see higher precipitations in horizon 2020 compared to the situation in 2000, in particular for METO-HC-HadRM3Q0-HadCM3Q0. Diversifying from the general trend, all models present areas in Southern Europe which are estimated to become wetter in the 2030 time horizon compared to the 2000 one. This zone is

Fig. 8 Simulated global solar radiation against daily temperature amplitude for two grid cells. Simulated global solar radiation is either (*left column*) directly generated from the GCM or (*right column*) derived from temperature values as proposed in this paper. The *red solid line* highlights difference in maximum values



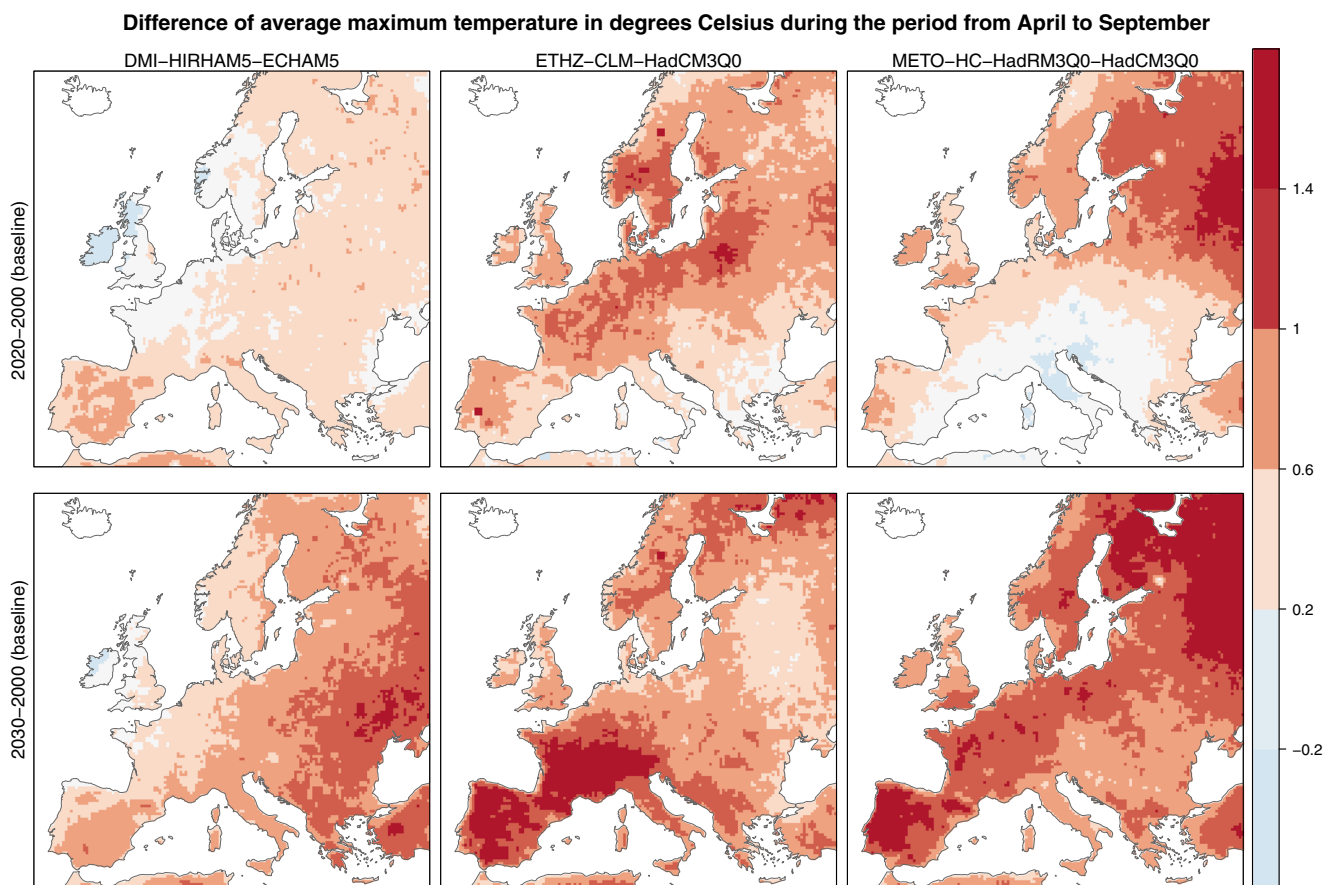


Fig. 9 Difference in mean maximum temperature, averaged over the 30 synthetic years and for the period going from 1 April until 30 September, for three different bias-corrected model runs (*columns*) and time horizons (*rows*)

shifted for each model: in METO-HC-HadRM3Q0-HadCM3Q0, it spans all along the Mediterranean and includes the western part of the Black Sea; in ETHZ-CLM-HadCM3Q0, it is localized in the Iberian and Italian peninsulas; while in DMI-HIRHAM5-ECHAM5, it encompasses central France, the Alps and Northern Italy, leaving the Iberian peninsula much drier than the other two models.

5 Discussion

The presented database provides a coherent characterization of three realizations of the A1B scenario in the short-term time horizons of 2020 and 2030. The target of the database are users of impacts models, and more specifically crop/crop diseases and pests models, who aim at providing insight to policy-makers into how agriculture will evolve in Europe under changing climate conditions in the near future. This dataset, for which a preliminary version had been announced in a short communication (Donatelli et al. 2012b), has been used in a series of studies of impact modelling (Bregaglio et al. 2013; Maiorano et al. 2013; Manici et al. 2014; Donatelli et al. 2015). It has been used to make policy reports for the

European Union through the AVEMAC (Donatelli et al. 2012a) and PESETA II projects (Ciscar et al. 2014), and it was also used to generate crop simulations that contributed to the Fifth Assessment Report of the IPCC (Kovats et al. 2014).

The considerable differences in precipitation patterns among the three ENSEMBLES models lead to different possibilities regarding what strategy to take when defining adequate policy for adaptation and/or mitigation. Depending on what model is considered, the outlooks for France and Spain are reversed, for instance. Policy can be conservative, trying to tackle both diverging possibilities, or it can be tailored to address only the more costly scenario. The fact that in many places there is a change in trajectory between the 2020 and 2030 horizons further stresses the importance in looking at short-term time horizons for decision-making and policy planning. Measures taken to address the situation in 2020 may quickly become obsolete if the situation changes considerably by 2030. Inversely, policy addressing 2030 (or beyond) may be strongly criticized if the events in 2020 are portraying a contrasting picture.

Targeting these short-term horizons is challenging because they are more prone to be influenced by the inter-annual

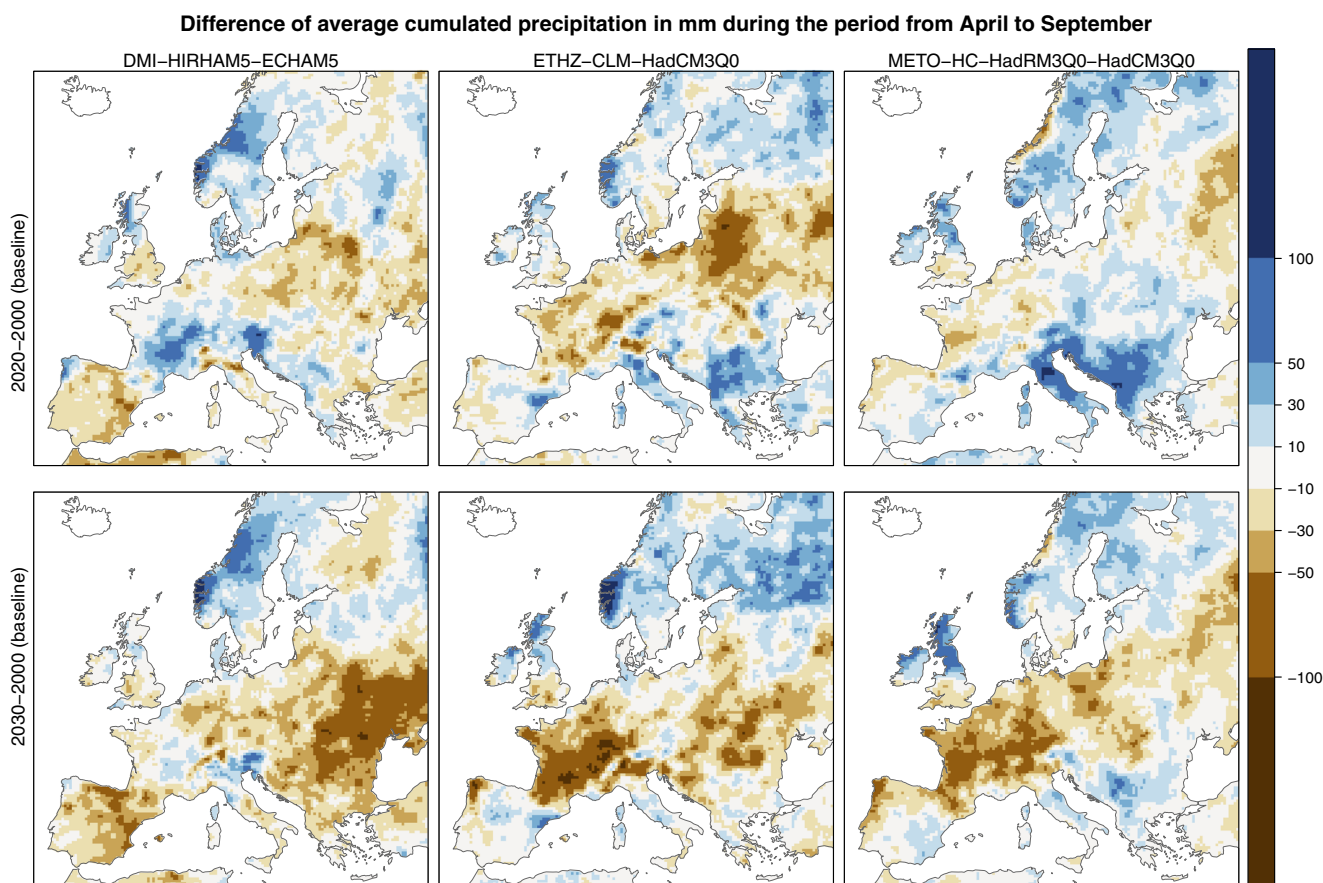


Fig. 10 Difference in the cumulated daily precipitation, averaged over the 30 synthetic years and for the period going from 1 April until 30 September, for three different bias-corrected model runs (*columns*) and time horizons (*rows*)

variability rather than by the long-term trends (Maraun 2013a). Multiplying the number of years over the 15-year time horizon windows using a weather generator to characterize the climatology is not necessarily a perfect solution. The 15-year windows might still be strongly influenced by some more extreme events that are more attributable to inter-annual variability than changes in the climate. However, this method proved to be a suitable compromise, since the comparison over the baseline period does not show large differences of the distributions of synthetic weather variables with respect to those of observed ones, even though the observed ones ranged over the period 1988–2012 while the synthetic dataset ranged from 1993 to 2007. One thing that could be improved is to generate a larger number of synthetic years (such as 100 instead of 30) so as to better represent the extreme events. Regarding the comparison with observed data, it must be acknowledged that the gridded observed weather data can also have problems. Some researchers have noted potential weaknesses of the E-OBS data set in data-sparse regions (e.g. Herrera et al. 2010). The MCYFS database has similar shortcomings with respect to the spatial density of weather stations and how these are interpolated to the grid.

The results overall indicate that the presented dataset is coherent with respect to observed weather data for the bias-corrected variables. For the other variables necessary for crop growth simulation, pragmatic solutions have been proposed to render the database usable. For global solar radiation, the proposed approach of deriving it from temperature using the auto-calibrated Bristow-Campbell model showed a higher consistency with respect to other weather variables and allowed rescaling to the maximum values expected in clear sky conditions. Patterns in wind and relative humidity had to be assumed constant over the time spans of interest to be able to derive evapotranspiration and vapour pressure deficit, given that currently there is neither evidence nor quantitative estimates of the future change for those variables. Even if the accuracy and robustness of GCMs estimates will likely improve with time, the need of using weather data to estimate impacts must be match with usable weather data products, and this is what this paper presents for the crop modelling community.

The methodology used here to derive a dataset dedicated to crop modelling is not limited to the input data currently used. The ENSEMBLES data are being replaced by those of the CORDEX project (<http://wcrp-cordex.ipsl.jussieu.fr/>), just as

ENSEMBLES replaced those of PRUDENCE (Christensen and Christensen 2007). Furthermore, as already mentioned, the SRES scenarios are being replaced by the representative concentration pathways (RCPs) and shared socio-economic pathways (SSPs). However, there is no reason the present methodology should not be similarly applied to these new products. The bias-correction step realized by Dosio and Paruolo (2011) will need to be repeated, bearing in mind caveats regarding the spatial scale (Maraun 2013b), since the European part of CORDEX data is at a finer spatial resolution (11 km) than ENSEMBLES.

6 Data availability

Data are available to public users from the MARS AGRI4CAST data portal of the European Commission Joint Research Centre, which can be accessed from the following weblink: <http://agri4cast.jrc.ec.europa.eu/DataPortal/>.

7 Conclusions

Weather data are the main driving forces of models used to make impact assessments of climate change scenarios on agriculture. Differences in the processing required to prepare these weather datasets may lead to different weather outputs, which in turn can have large repercussions on the resulting simulations of impact models. Sharing a common database across Europe and neighbouring countries, as the one proposed in this paper, removes a potential source of uncertainty in climate change and agriculture analyses across the region.

The three GCMs used to generate the dataset show a noticeable heterogeneity in the short-term future. This is particularly the case for precipitation patterns, which can be substantially different even for the same emission scenario and despite having all GCMs simulations showing similar increases in temperature values in the short term. Furthermore, the differences in precipitation patterns appear to be more pronounced in areas in which rainfall is already critical. Because rainfall is such a key driver of the performance of agricultural systems in the short term, such variability should be further investigated extending the analysis beyond by using more GCMs than the three considered here.

The process to build a dataset as the one developed here has requested considerable resources, domain-specific knowledge and substantial infrastructure and technological expertise. The data made available, which could be extended in the near future to other emission scenarios, provide a ready to use, cost-free resource to public institutions.

Acknowledgments This work was carried out under the AVEMAC project, an administrative arrangement between DG-CLIMA and DG-JRC of the European Commission. It was also partially funded by the project AgroScenari of the Italian Ministry for Agriculture, Food and Forestry Policies. The authors thank A. Dosio and anonymous reviewers for providing valuable feedback on the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration—guidelines for computing crop water requirements—FAO irrigation and drainage paper 56.
- Baron C, Sultan B, Balme M, et al. (2005) From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. *Philos Trans R Soc B Biol Sci* 360:2095. doi:10.1098/rstb.2005.1741
- Barriopedro D, Fischer EM, Luterbacher J, et al. (2011) The hot summer of 2010: redrawing the temperature record map of Europe. *Science* 80(332):220–224. doi:10.1126/science.1201224
- Bojanowski JS, Donatelli M, Skidmore AK, Vrieling A (2013) An auto-calibration procedure for empirical solar radiation models. *Environ Model Softw* 49:118–128. doi:10.1016/j.envsoft.2013.08.002
- Bregaglio S, Donatelli M, Confalonieri R (2013) Fungal infections of rice, wheat, and grape in Europe in 2030–2050. *Agron Sustain Dev* 33:767–776. doi:10.1007/s13593-013-0149-6
- Bristow KL, Campbell GS (1984) On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric For Meteorol* 31:159–166. doi:10.1016/0168-1923(84)90017-0
- Christensen J, Kjellström E, Giorgi F, et al. (2010) Weight assignment in regional climate models. *Clim Res* 44:179–194. doi:10.3354/cr00916
- Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Clim Chang* 81:7–30. doi:10.1007/s10584-006-9210-7
- Christensen JH, Boberg F, Christensen OB, Lucas-Picher P (2008) On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophys Res Lett* 35(20). doi:10.1029/2008GL035694
- Christensen O, Drews M, Christensen J, et al. (2006) The HIRHAM regional climate model version 5 (β). Tech Rep 06–17. ISSN 1399-1388. Copenhagen
- Ciscar J, Feyen L, Soria A, et al (2014) Climate impacts in Europe. The JRC PESETA II project
- Collins M, Booth BBB, Bhaskaran B, et al. (2011) Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles. *Clim Dyn* 36:1737–1766. doi:10.1007/s00382-010-0808-0
- Donatelli M, Bellocchi G, Carlini L (2006) Sharing knowledge via software components: models on reference evapotranspiration. *Eur J Agron* 24:186–192. doi:10.1016/j.eja.2005.07.005
- Donatelli M, Bellocchi G, Habyarimana E, et al. (2009) CLIMA: a weather generator framework. In: Anderssen RS, Braddock RD, Newham LTH (eds) 18th World IMACS Congr. MODSIM09 Int. Congr. Model. Simul. Modelling and Simulation Society of Australia and

- New Zealand and International Association for Mathematics and Computers in Simulation, Cairns, Australia, pp 852–858
- Donatelli M, Duveiller G, Fumagalli D, et al. (2012a) Assessing agriculture vulnerabilities for the design of effective measures for adaptation to climate change (AVEMAC project). doi: 10.2788/16181
- Donatelli M, Fumagalli D, Zucchini A, et al. (2012b) A EU27 database of daily weather data derived from climate change scenarios for use with crop simulation models. *Int Environ Model Softw Soc 2012 Int Congr Environ Model Software Manag Resour a Ltd Planet Pathways Visions Under Uncertainty, Sixth Bienn Meet 1–5 July 2012*,
- Donatelli M, Srivastava AK, Duveiller G, et al. (2015) Climate change impact and potential adaptation strategies under alternate realizations of climate scenarios for three major crops in Europe. *Environ Res Lett* 10:075005. doi:10.1088/1748-9326/10/7/075005
- Dosio A, Paruolo P (2011) Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: evaluation on the present climate. *J Geophys Res* 116:D16106. doi:10.1029/2011JD015934
- Dosio A, Paruolo P, Rojas R (2012) Bias correction of the ENSEMBLES high resolution climate change projections for use by impact models: analysis of the climate change signal. *J Geophys Res* 117: D17110. doi:10.1029/2012JD017968
- Falloon P, Betts R (2010) Climate impacts on European agriculture and water management in the context of adaptation and mitigation—the importance of an integrated approach. *Sci Total Environ* 408:5667–5687
- Hansen JW, Challinor A, Ines A, et al. (2006) Translating climate forecasts into agricultural terms: advances and challenges. *Clim Res* 33: 27
- Hansen JW, Jones JW (2000) Scaling-up crop models for climate variability applications. *Agric Syst* 65:43–72
- Hawkins E, Osborne TM, Ho CK, Challinor AJ (2013) Calibration and bias correction of climate projections for crop modelling: an idealised case study over Europe. *Agric For Meteorol* 170:19–31. doi:10.1016/j.agrformet.2012.04.007
- Haylock MR, Hofstra N, Klein Tank a MG, et al. (2008) A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J Geophys Res* 113:D20119. doi:10.1029/2008JD010201
- Herrera S, Fita L, Fernández J, Gutiérrez JM (2010) Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel simulations over Spain. *J Geophys Res* 115:D21117. doi:10.1029/2010JD013936
- IPCC (2013) Summary for policymakers. *Clim. Chang. 2013 Phys. Sci. Basis. Contrib. Work. Gr. I to Fifth Assess. Rep. Intergov. Panel Clim. Chang.*
- Jaeger EB, Anders I, Lüthi D, et al. (2008) Analysis of ERA40-driven CLM simulations for Europe. *Meteorol Zeitschrift* 17:19. doi:10.1127/0941-2948/2008/0301
- Kovats RS, Valentini R, Bouwer LM, et al. (2014) Europe. In: Barros VR, Field CB, Dokken DJ, et al (eds) *Clim. Chang. 2014 Impacts, adapt. vulnerability. Part B Reg. Asp. Contrib. Work. Gr. II to Fifth Assess. Rep. Intergov. Panel Clim. Chang.* Cambridge University Press, Cambridge and New York, NY, pp. 1267–1326
- Maiorano A, Cerrani I, Fumagalli D, Donatelli M (2013) New biological model to manage the impact of climate warming on maize corn borers. *Agron Sustain Dev* 34:609–621. doi:10.1007/s13593-013-0185-2
- Manici LM, Bregaglio S, Fumagalli D, Donatelli M (2014) Modelling soil borne fungal pathogens of arable crops under climate change. *Int J Biometeorol* 58:2071–2083. doi:10.1007/s00484-014-0808-6
- Maracchi G, Sirotenko O, Bindi M (2005) Impacts of present and future climate variability on agriculture and forestry in the temperate regions: Europe. *Clim Chang* 70:117–135
- Maraun D (2013a) When will trends in European mean and heavy daily precipitation emerge? *Environ Res Lett* 8:014004. doi:10.1088/1748-9326/8/1/014004
- Maraun D (2013b) Bias correction, quantile mapping, and downscaling: revisiting the inflation issue. *J Clim* 26:2137–2143. doi:10.1175/JCLI-D-12-00821.1
- Mearns LO, Rosenzweig C, Goldberg R (1996) The effect of changes in daily and interannual climatic variability on CERES-wheat: a sensitivity study. *Clim Chang* 32:257–292
- Micale F, Genovesi G (2004) Vol 1. Meteorological data collection, processing and analysis. AgriFish unit. In: *Methodology of the MARS crop yield forecasting system.* Joint Research Centre of the European Commission, Ispra
- Nakićenović N, Alcamo J, Davis G, De Vries B, Fenhann J, Gaffin S, et al (2000) *Special Report on Emissions Scenarios (SRES), Working Group III, Intergovernmental Panel on Climate Change (IPCC).* Cambridge University Press, Cambridge, pp 595
- O'Neill BC, Krieglner E, Riahi K, et al. (2013) A new scenario framework for climate change research: the concept of shared socioeconomic pathways. *Clim Chang* 122:387–400. doi:10.1007/s10584-013-0905-2
- Olesen JE, Petersen BM, Berntsen J, et al. (2002) Comparison of methods for simulating effects of nitrogen on green area index and dry matter growth in winter wheat. *F Crop Res* 74:131–149
- Piani C, Haerter JO, Coppola E (2009) Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor Appl Climatol* 99:187–192. doi:10.1007/s00704-009-0134-9
- Piani C, Weedon GP, Best M, et al. (2010) Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *J Hydrol* 395:199–215
- Russo S, Dosio A, Graversen RG, et al. (2014) Magnitude of extreme heat waves in present climate and their projection in a warming world. *J Geophys Res Atmos* n/a–n/a. doi:10.1002/2014JD022098
- Schär C, Vidale PL, Lüthi D, et al. (2004) The role of increasing temperature variability in European summer heatwaves. *Nature* 427:332–336
- Semenov MA, Porter JR (1995) Climatic variability and the modelling of crop yields. *Agric For Meteorol* 73:265–283
- Stöckle CO, Nelson R, Donatelli M, Castellvi F (2001) *ClimGen: a flexible weather generation program.* 2nd Int. Symp. Model. Crop. Syst., Florence, pp. 16–18
- Teutschbein C, Seibert J (2010) Regional climate models for hydrological impact studies at the catchment scale: a review of recent modeling strategies. *Geogr Compass* 4:834–860. doi:10.1111/j.1749-8198.2010.00357.x
- Van der Linden P, Mitchell JFB (eds) (2009) *ENSEMBLES: climate change and its impacts: summary of research and results from the ENSEMBLES project.* Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK
- Van Vuuren DP, Edmonds J, Kainuma M, et al. (2011) The representative concentration pathways: an overview. *Clim Chang* 109:5–31. doi:10.1007/s10584-011-0148-z