Crop impact projections in Iberian Peninsula for mid and end of C21 improved by bias correction of RCM outputs

Supplemental Material

Ruiz-Ramos M.[1*], Rodríguez A.[1], Dosio A.[2], Goodess C. M.[3]., Harpham C.[3], Mínguez M.I.[1], Sánchez E.[4]

1 CEIGRAM-Agricultural systems group from Technical University of Madrid, Spain
margarita.ruiz.ramos@upm.es
2 European ComissionJoint Research Centre, Institute for Environment and Sustainability, Ispra, Italy
3 Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, UK
4 Facultad de Ciencias del Medio Ambiente, University of Castilla-La Mancha, Avda. Carlos III s/n, 45071 Toledo, Spain

* corresponding author
margarita.ruiz.ramos@upm.es
Phone number: +34 914524900 ext 1683

1. Data and Methods

1.1. Crop Modelling

Based on previous field experiments CERES-maize from DSSAT v. 3.5 calibrations (Iglesias and Mínguez, 1995), data from field experiments (Maturano, 2002; Gabriel and Quemada, 2011) and bulletins from technical services (the Instituto Técnico Agronómico Provincial de Albacete, ITAP), site-specific recalibration and validation of the version 4.5 of CERES-DSSAT was performed for both locations. Five and four years were used for recalibration for Aranjuez and Albacete respectively, and three and six years for validation. The recalibration was done for potential yield of irrigated maize. Calibration and validation were evaluated by the Root Mean Square Error (RMSE) and root mean square percentage error (RMSPE, i.e., RMSE normalized by the average of observed values).

$$RMSPE = \sqrt{\frac{\sum_{i=1}^{N}\left(\frac{O_i - S_i}{O_i}\right)^2}{N}} * 100 \qquad eq.(1)$$

where Oi represents the observed data, Si represents the simulated data and N is the number of data.

For all crop simulations, maize evapotranspiration (ET) was calculated by the Priestly-Taylor method (Priestly and Taylor, 1972).

1.2. Observed and simulated climate datasets

E-OBS v. 3.0 (Haylock et al., 2008) is a gridded data set, with the same spatial resolution as ENS, consisting of daily observations of temperature and precipitation from 2316 stations (*ca*. 50 for Spain) covering the whole of Europe for the period 1950–2006 (last version spans up to 2014). The improvement introduced by bias correction in ENS-EOBS has been evaluated in terms of the probability distribution functions (PDFs) of temperature and precipitation in Dosio and Paruolo (2011) and an application of ENS-EOBS for hydrological impact assessment is shown in Rojas et al. (2012).

Spain02 is an observational gridded data set for Spain for temperature and precipitation at similar resolution to E-OBS (0.2º for Spain02, 25 km for E-OBS) (Herrera et al., 2012), built with a network of ~2500 quality-controlled stations (~250 for temperature, *vs. ca*. 50 stations of E-OBS in Spain; see Fig. 1 in Haylock et al., 2008) from the Spanish Meteorological Agency (AEMET) spanning the period 1950 to 2008.Spain02 has been validated against station data using cross-validation, obtaining a good performance for precipitation occurrence, accumulated amounts, variability and seasonality, and shows good performance in the reproduction of the intensity and spatial variability of extremes (Herrera et al., 2012).

2.3. Techniques of bias correction and reduction

Nonparametric estimation of the CDF requires the fitting of individual observed and modelled empirical distribution functions (Piani et al., 2010). To reduce the computational cost of this procedure, Dosio and Paruolo (2011) used a parametric transfer function (TF) as a function of up to four parameters. They compared both methods of estimating TF, concluding that parametric estimation of the TF gives satisfactory results. Also, it is much cheaper computationally than nonparametric estimation, which makes it more suitable for extensive applications.

The choice of the TF depends on the variable to be corrected. For temperature, the TF proposed by Piani et al. (2010) was a linear equation, with two parameters. For precipitation, the TF was a set of three equations (linear, logarithmical and exponential). that can be seen as an exponential tendency to an asymptote defined by the linear equation where one of the four parameters is the rate at which the asymptote is approached. Details of the parameters are provided in Dosio and Paruolo (2011). These parameters are estimated by least squares on a monthly basis, so that monthly TFs are produced, and then interpolated into daily TFs using a smoothing technique. For the case of temperature, Piani et al. (2010b) calculated the daily temperature range (Trng, where Trng=Tmax – Tmin) and the daily temperature skewness (Tsk, where Tsk, = (Tmean − Tmin)/Trng). Then they proposed to bias correct Tmean, Trng and Tsk, and subsequently to invert the calculation obtaining the bias-corrected fields for Tmean, Tmin, Tmax. However, as Spain02 provides Tmin and Tmax but not Tmean, in this studybias correction was applied to Tmax and Tmin.

Alternatively to bias correction, bias can be reduced by the use of a weather generator (WG); in our case the CRU WG (Kilsby et al., 2007). The WG is calibrated on observed station data and projection output is produced by perturbing the WG parameters with monthly change factors calculated from RCM present and future runs. In detail, the WG produces internally consistent series of “synthetic” meteorological variables including: rainfall, temperature, humidity, wind, sunshine, radiation (diffuse and direct), and a derivation of potential evapotranspiration. The system produces series at a daily time resolution, using two stochastic models in series: first, for rainfall (RainSim, Newcastle University, Kilsby et al., 2007) that produces an output series which is then used for a second model (CRU WG, Kilsby et al., 2007) generating the other variables dependent on rainfall (and for humidity and so on, dependent on rainfall and temperature). The steps required to produce rainfall are: 1) to calculate statistics from the observed time series, including mean, variance, probability of dry and auto-correlation; 2) to fit the model to the statistics, and 3) to generate the precipitation output. Then, the output from the RainSim model is read in by the CRU WG along with the available observed climate variables (at least temperature data is required to run). If projection output is required, then the statistics output file generated in (1) needs to have the change factors applied. The model can then be re-fitted to generate the projection precipitation output. To perturb temperature, the changes in mean and variance are used.

The coefficient of efficiency E (eq. 2) is the ratio of the mean square error between the measured and the simulated data to the variance in the observed data subtracted from unity. It varies from −∞ to 1. An E value of 1 corresponds to a perfect match of simulated data to the observations. An E value of zero (0) is obtained when simulation is as accurate as the mean of the observations. Negative E values indicate that the observed mean is a better predictor than the simulated data.

$$E = 1.0 - \frac{\sum_{i=1}^{N}(O_i - S_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2} \qquad eq.(2)$$

where *Oi* represents the observed data, *Si* represents the simulated data, $\bar{O}$ is the mean of the observed values and *N* is the number of data points of the compared datasets.

2. Results

2.1. Crop model calibration and validation

Crop model calibration and validation results were as follows: In Aranjuez, the RMSE was 569 kg ha-1 and the RMSPE was 5% for the calibration years, and 1181 kg ha^{-1} and 13% respectively for the validation years. In Albacete, the RMSE was 691 kg ha^{-1} and the RMSPE was 5% for the calibration, and 1577 kg ha-1 and 10% for the validation.

2.2. Datasets of RCM projections with reduced bias: Efficiency Coefficient and Probability density functions (PDFs)

The E values, calculated for daily Tmax and Tmin at both locations, of the corrected datasets were all positive and similar.

The E coefficients for precipitation (Table S-4) were always negative for all datasets, indicating that for this variable the residual variance of all datasets was larger than the variance of observations. All datasets obtained similar E values with ENS-WG showing the minimum value at both locations.

These improvement introduced by bias correction in crop projections can be partially quantified by comparing the E coefficients, which were much closer to 0 (and also to 1, see Table S-4) for the corrected datasets than for ENS. Improvement in E values was more evident for crop simulations than for climate data, and together with biases may help to select the more suitable method for phenology simulation at each location: ENS-SPAIN02 for Aranjuez (only method with positive E values) and ENS-WG for Albacete (method with phenological E values sum closest to 0).

ENS-EOBS, ENS-SPAIN02, and ENS-WG datasets produced PDFs of summer Tmax closer than ENS to that of the AEMET data, as expected, for both locations. ENS showed an underestimation of the probability of median and an overestimation of the probability of lowest and highest temperatures (distribution tails). In both locations, ENS-SPAIN02 produced a PDF very close to that of AEMET, closer than that of ENS-EOBS, which can be explained by the fact that SPAIN02 matched AEMET better than

E-OBS, especially in Albacete (Figure S-2). At this location, the spread and tails of the Tmax corrected distributions matched those of AEMET, indicating a good representation of both the standard deviation and extreme events, respectively.

ENS matched the AEMET PDF of winter Tmin in Aranjuez, and the ENS-SPAIN02 PDF was the closest to AEMET of all the bias reduced datasets. ENS-EOBS, ENS-SPAIN02 and ENS-WG had improved PDFs of winter Tmin in Albacete, but a displacement towards warmer temperatures remained. In both locations, ENS-SPAIN02 presented a small overestimation of the probability of the median temperatures. The observational gridded data sets showed an overestimation of temperatures and a higher probability peak, especially in Albacete; this overestimation was more important for E-OBS (Figure S-2).

The PDFs of spring and autumn precipitation showed similar features for both locations (Figure S-3). Both ENS-EOBS and ENS-SPAIN02 improved the simulation of the left tail of the distribution (for events equal or below 12 mm, Figure S-3), with ENS-SPAIN02 closer to the AEMET curve in Albacete in autumn. ENS-WG also showed good agreement with observations, but with a displacement of the left tail to lower precipitation values at Aranjuez in spring (Figure S-3).

References

Dosio A, Paruolo P (2011) Bias correction of the ENSEMBLES high-resolution climate change projectionsfor use by impact models: Evaluation on the present climate. J Geophys Res 116, D16106, doi:10.1029/2011JD015934

Gabriel JL, Quemada M (2011) Replacing bare fallow with cover crops in a maize cropping system: Yield, N uptake and fertiliser fate. Eur J Agron 34:133-143

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. J Geophys Res Atmos 113, D20119, doi:10.1029/2008JD10201

Herrera S et al (2012) Development and Analysis of a 50 year high-resolution daily gridded precipitation dataset over Spain (Spain02). Int J Climatol 32:74-85, doi:10.1002/joc.2256

Iglesias A, Mínguez MI (1995) Perspectives for maize production in Spain under climate change. In: Rosenzweig C, Allen Jr LH, Harper LA, Hollinger SE, Jones J (eds) Climate change and agriculture: analysis of potential international impacts. Am SocAgron, 259–273

Kilsby CG, Jones PD, Burton A, Ford AC, Fowler HJ, Harpham C, James P, Smith A, Wilby RL (2007) A daily weather generator for use in climate change studies. EnvironModellSoftw, 22:1705-1719

Maturano M (2002) Estudio del uso del agua y del nitrógeno dentro del marco de una agricultura sostenible en las regiones Castellano-Manchega y Argentina. PhD Thesis, 246 pp. Universidad de Castilla-La Mancha

Piani C, Weedon GP, Best M, Gomes SM, Viterbo P, Hagemann S, Haerter JO (2010) Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. J Hydrol 395:199–215, doi:10.1016/j.jhydrol.2010.10.024

Priestley CHB, Taylor RJ (1972) On the assessment of surface heat flux and evaporation using large-scale parameters. MonWeatherRev 100:81-92

Rojas R, Feyen L, Bianchi A, Dosio A (2012) Assessment of future flood hazard in Europe using a large ensemble of bias corrected regional climate simulations. J

Geophys Res 117:D17109

Supplemental Tables

Table S-1. Cultivar name and phenology, management, soil type and depth used in the crop simulations conducted at each location. GDD, growing degree days, Tb, base temperature. References used for recalibration at each location.

	Aranjuez	Albacete
Cultivar	PR31G98 (FAO 700)	Prisma (FAO 700)
GDD from emergence to flowering (ºCd), Tb=8ºC	390	280
GDD from flowering to maturity (ºCd), Tb=8ºC	770	789
Sowing date	EarlyApril	Secondhalf of April
Harvest date	End of September/Early October	November
Soiltype	TypicCalcixerept , 120 cm	Xerochrepts, 70 cm
Soildepth (cm)	120	70
References	Gabriel and Quemada (2011)	Iglesias and Mínguez(1995) Maturano (2002)

Table S-2. Institution, driving GCM and name of each RCM used for ENS and Delta. The corrected column indicates the 12 RCMs used for ENS-EOBS, ENS-SPAIN02, ENS-WG, ENS-EOBS-WG and ENS-SPAIN02-WG.

Institution	RCM	Driving GCM	Corrected
C4I	C4IRCA3	ECHAM5	No
C4I	C4IRCA3	HadCM3Q16	Yes
CNRM	CNRM-RM4.5	ARPEGE	Yes
DMI	DMI-HIRHAM5	ARPEGE	Yes
DMI	DMI-HIRHAM5	ECHAM5-r3	Yes
DMI	DMI-HIRHAM5	BCM	Yes
ETHZ	ETHZ-CLM	HadCM3Q0	Yes
HC	METO-HC_HadRM3Q0	HadCM3Q0	Yes
HC	METO-HC_HadRM3Q3	HadCM3Q3	No
KNMI	KNMI-RACMO2	ECHAM5-r3	Yes
METNO	METNOHIRHAM	BCM	No
METNO	METNOHIRHAM	HadCM3Q0	No
MPI	MPI-M-REMO	ECHAM5-r3	Yes
SMHI	SMHIRCA	BCM	Yes
SMHI	SMHIRCA	ECHAM5-r3	Yes
SMHI	SMHIRCA	HadCM3Q3	Yes
UCLM	UCLM-PROMES	HadCM3Q0	No

Table S-3. Description of the ensembles of RCM projections built for this study: Name, number of members, post-processing technique used to build the ensemble and observational data set used as reference in the post-processing.

Ensemble Name	Number of members (RCMs)	Post-processing technique	Observational data set
ENS	17	None	-
ENS-EOBS	12	Biascorrection	E-OBS
ENS-SPAIN02	12	Biascorrection	Spain02
ENS-WG	12	WeatherGenerator	AEMET
ENS-EOBS-WG	12	Biascorrection +WeatherGenerator	E-OBS, AEMET
ENS-SPAIN02-WG	12	Biascorrection +WeatherGenerator	Spain02, AEMET
DELTA	17	Delta method	AEMET

Table S-4. Evaluation of corrected ensembles: Coefficient of efficiency (E) of the ensembles ENS, ENS-EOBS, ENS-SPAIN02 and ENS-WG, when compared to 1) AEMET, for daily Tmax, Tmin and precipitation and 2) AEMET-driven crop simulations, for anthesis and maturity dates and crop yield.

		E coefficient					
Location	Ensemble/Method	Tmax	Tmin	Precipitation	Anthesis date	Maturity date	Yield
Aranjuez	ENS	0,77	0,77	-0,11	-16,60	-6,12	-0,71
	ENS-EOBS	0,79	0,75	-0,10	-2,03	-0,37	-1,24
	ENS-SPAIN02	0,79	0,77	-0,15	0,28	0,43	-0,96
	ENS-WG	0,65	0,54	-0,94	-1,39	-1,08	-1,11
Albacete	ENS	0,78	0,69	-0,06	-12,47	-4,07	-1,06
	ENS-EOBS	0,79	0,72	-0,08	-2,42	-0,90	-0,25
	ENS-SPAIN02	0,78	0,73	-0,08	-3,34	-1,50	-0,45
	ENS-WG	0,61	0,49	-0,80	-0,96	-0,65	-0,75

Supplemental Figures

Figure S-1. Locations where comparison of ensembles of climate and crop simulations were conducted. Altitude of Aranjuez (-3.716ºW, 40.30ºN) is around 500 m a.s.l. and altitude of Albacete(-1.85ºW, 38.95ºN) is around 700 m a.s.l.

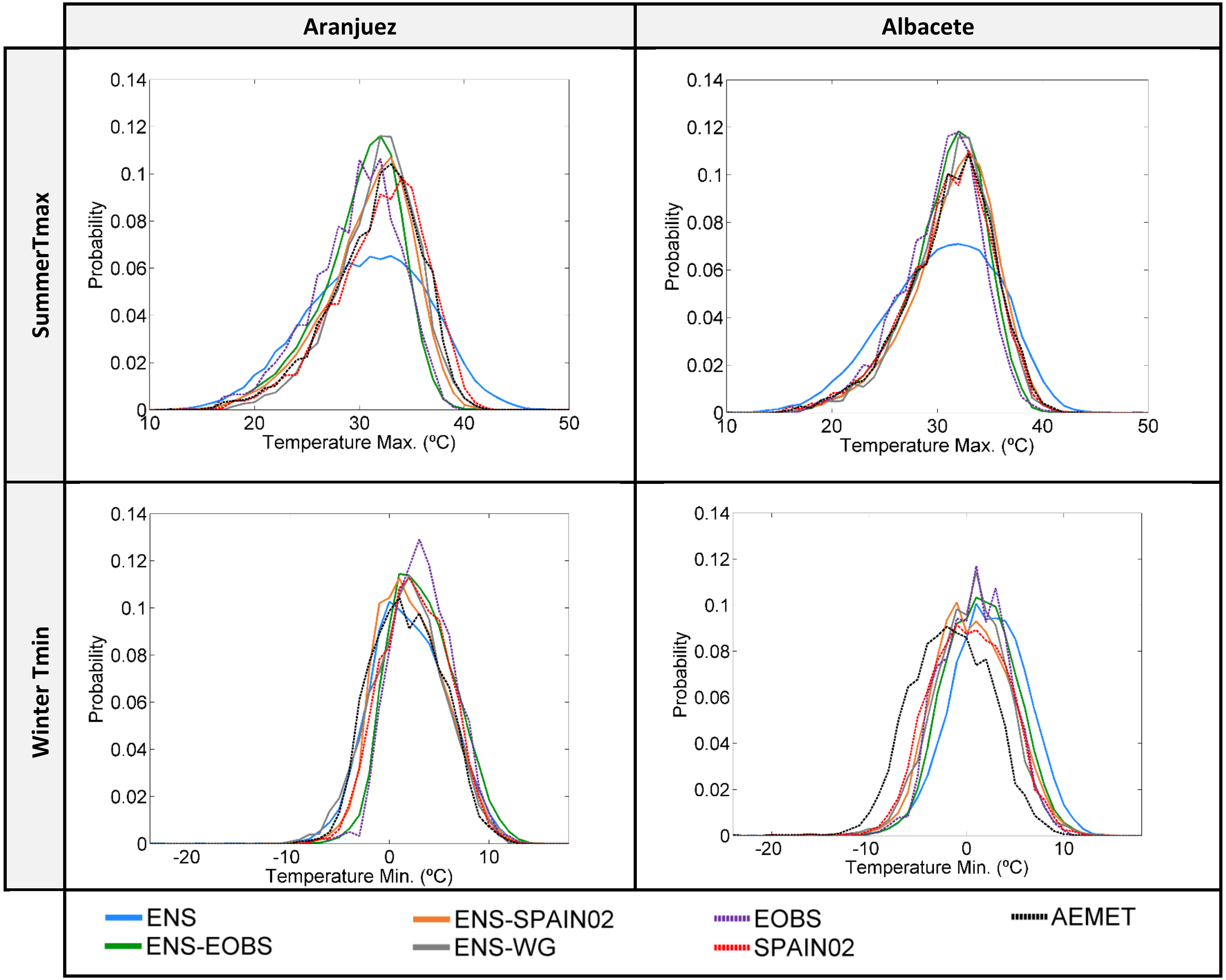

Figure S-2. Probability distribution function (PDF) of summer Tmax and winter Tmin for Aranjuez and Albacete for the period 1971-2000, for the observational datasets AEMET, E-OBS and SPAIN02, the uncorrected ensemble ENS, and the corrected ensembles ENS-EOBS, ENS-SPAIN02 and ENS-WG.

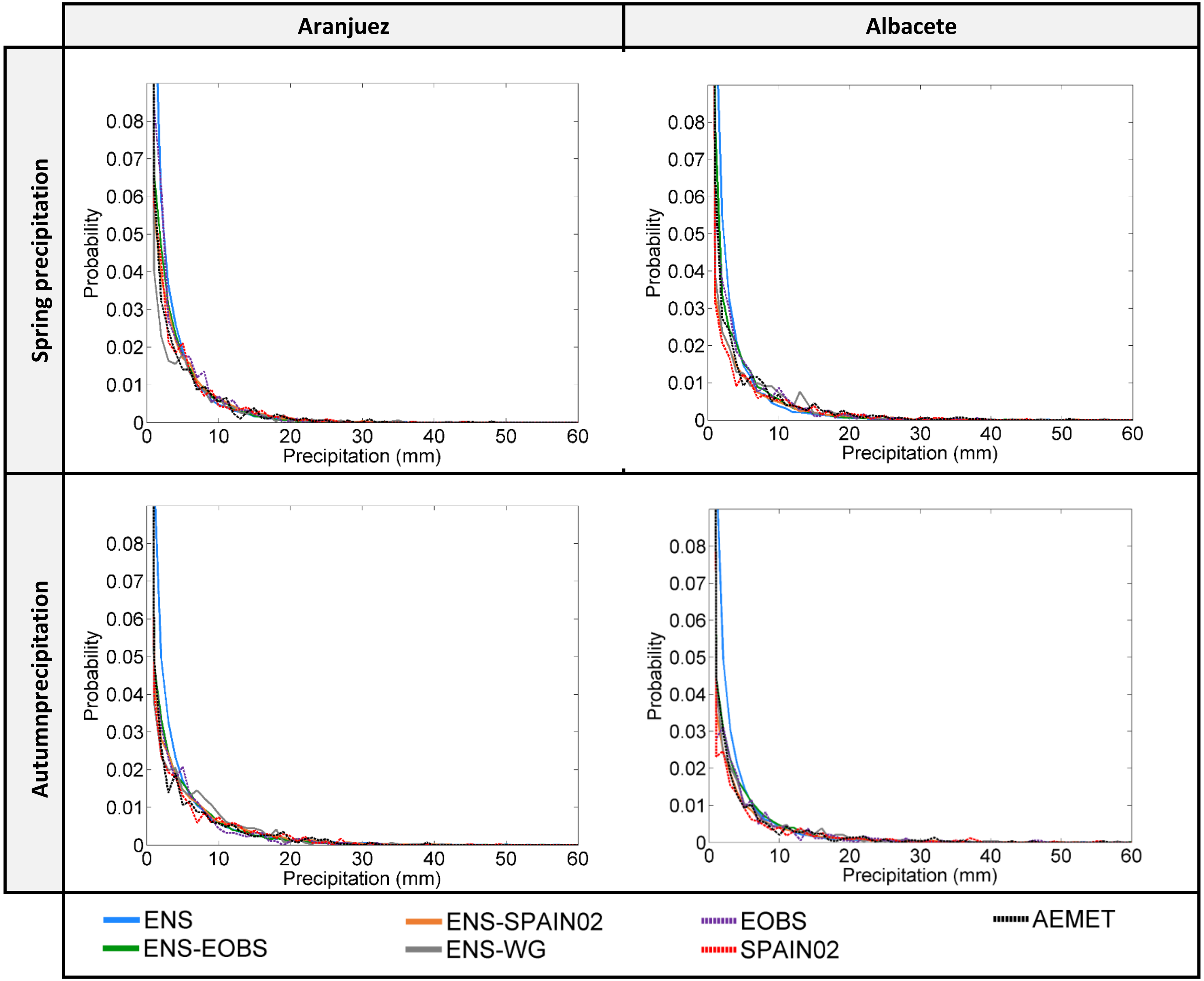

Figure S-3. As Figure S-2, but for spring and autumn precipitation.