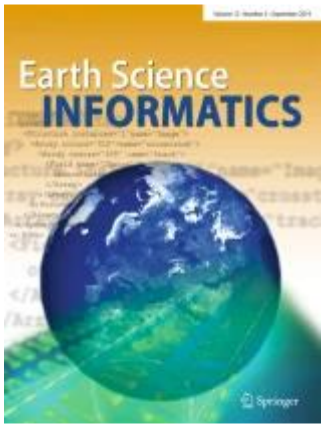
Accepted Manuscript



Calibration and spatial modelling of daily ET_0 in semiarid areas using Hargreaves equation

Francisco Gomariz-Castillo, Francisco Alonso-Sarría & Cabezas-Calvo-Rubio, Francisco

DOI:	https://doi.org/10.1007/s12145-017-0327-1
Reference:	TJSS-2017-0057
To appear in:	*Earth Science Informatics*
Received date:	5 June 2017
Revised date:	13 September 2017
Accepted date:	18 October 2017
Published online:	10 November 2017
Final publication:	1 September 2018

Please cite this article as: Francisco Gomariz-Castillo, Francisco Alonso-Sarría, & Cabezas-Calvo-Rubio, Francisco (2018) Calibration and spatial modelling of daily ET0 in semiarid areas using Hargreaves equation, Earth Science Informatics, 11, 325-340, DOI: 10.1007/s12145-017-0327-1

Correspondence to Francisco Gomariz-Castillo

Earth Science Informatics manuscript No.
(will be inserted by the editor)

Calibration and spatial modelling of daily ET_0 in semiarid areas using Hargreaves equation

Francisco Gomariz-Castillo · Francisco Alonso-Sarría · Francisco Cabezas-Calvo-Rubio

Received: date / Accepted: date

Abstract Evapotranspiration is difficult to measure and, when measured, its spatial variability is not usually taken into account. The recommended method to estimate evapotranspiration, Penman-Monteith FAO, requires variables not available in most weather stations. Simplified but less accurate methods, as Hargreaves equation, are normally used. Several approaches have been proposed to improve Hargreaves equation accuracy. In this work, 14 calibrations of the Hargreaves equation are compared. Three goodness of fit statistics were used to select the optimal, in terms of simplicity and accuracy. The best option was an annual linear regression. Its parameters were interpolated using regression-kriging combining Random Forest and Ordinary Kriging. Twelve easy to obtain ancillary variables were used as predictors. The same approach was used to interpolate Hargreaves and Penman-Monteith-FAO ET_0 on a daily basis; the Hargreaves ET_0 layers and the parameter layers were used to obtain calibrated ET_0 estimations. To compare the spatial patterns of the three estimations the daily layers were integrated into annual layers. The results of the proposed calibration are much more similar to Penman-Monteith FAO results than those obtained with Hargreaves equation. The research was conducted in south-east Spain with 79 weather stations with data from 01/01/2003 to 31/12/2014.

Keywords Evapotranspiration · Hargreaves equation · Allen calibration · Spatial interpolation · Random Forest

Francisco Gomariz-Castillo
Euro-mediterranean Water Institute, Campus de Espinardo s/n 30100 Murcia, Spain
Institute for Water and Environment, University of Murcia, Campus de Espinardo s/n 30100 Murcia, Spain
Tel.: +34-968-899851
Fax: +34-968-832510
E-mail: fjgomariz@um.es

Francisco Alonso-Sarría
Institute for Water and Environment, University of Murcia, Campus de Espinardo s/n 30100 Murcia, Spain

Francisco Cabezas-Calvo-Rubio
Euro-mediterranean Water Institute, Campus de Espinardo s/n 30100 Murcia, Spain

1 Introduction

Evapotranspiration is one of the most important processes in the water cycle; its knowledge is essential to water resources management, planning and design, especially in semiarid regions. However, its direct measurement is very expensive, time consuming and mainly carried out in agricultural plots; it is then almost impossible to have enough observations to obtain a regional interpolation. Several models have been proposed to estimate evapotranspiration from meteorological variables (Xu and Singh 2002), specifically a reference evapotranspiration (ET_0) defined by Allen et al (1998) as “the rate of evapotranspiration from an hypothetical crop with an assumed crop height (0.12 m), a fixed canopy resistance (70 s m-1) and albedo (0.23), such crop would closely resemble the evapotranspiration from a widespread surface of green grass of homogeneous height, fully growing, wholly shading the ground and not short of water”.

The Penman-Monteith FAO (PMFAO) equation (Allen et al 1989; Monteith 1965) is considered to provide the most accurate estimation of daily and monthly ET_0 in all climates (Subburayan et al 2011) and has been recommended by FAO (Allen et al 1994). It requires several weather variables that are not available in many weather stations and are also difficult to interpolate. Alternative methods using only temperature or radiation as independent variables have been proposed to overcome this issue (Xu and Singh 2000, 2001). These methods can be applied in a substantially greater number of weather stations. Their main drawback is a decrease in the accuracy of the estimations.

1.1 Hargreaves equation

One of the most widely used equations to estimate ET_0 was proposed by Hargreaves and Allen (2003):

$$ET_0 = 0.0135 \cdot R_S \cdot \left(\frac{T_{max} + T_{min}}{2} + C_T \right) \tag{1}$$

where ET_0 is the reference evapotranspiration (mm day-1), R_S is the incident radiation measured as ET_0 equivalent (mm day-1), T_{max} and T_{min} are the daily maximum and minimum temperature (C), and C_T is an empirical coefficient whose value is 17.8 (Hargreaves 1994a). Samani (2000) proposed an approximation using extraterrestrial radiation R_a to be used when real R_S data are not available:

$$R_S = KT \cdot R_a \cdot \Delta T^{EH} \tag{2}$$

where EH is a coefficient which usually equals 0.5, $\Delta T = T_{max} - T_{min}$, R_a is obtained from tables or calculated from temperature, and KT is an empirical coefficient whose default value is 0.17 (Hargreaves and Samani 1985). Hargreaves (1994b), cited by Sendanayake and Miguntanna (2014), recommended KT to be equal to 0.162 in continental areas and 0.19 near the coast. This variation takes into account the influence of cloudiness and humidity on the relation between ΔT and the proportion of R_a reaching the surface. However, other factors might influence this relation: latitude, elevation, topography, storm patterns, advection and proximity to a large body of water (Sendanayake and Miguntanna 2014). Using Eq. 2, Hargreaves equation becomes a temperature-based method (Xu and Singh 2000). If KT=0.17, Eq. 2 becomes:

$$ET_0 = 0.0023 \cdot R_a \cdot \sqrt{\Delta T} \cdot \left(\frac{T_{max} + T_{min}}{2} + C_T \right) \quad (3)$$

Di Stefano and Ferro (1997) obtained more accurate results with Hargreaves equation than with any other model, concluding that, for European climatic conditions, it produces the closest results to the PMFAO equation. Other successful attempts to use Hargreaves equation in arid and semiarid environments were made by López-Urrea et al (2006), Bautista et al (2009), Er-Raki et al (2010).

However, it has been demonstrated that Hargreaves equation over-estimates ET_0 in humid regions and under-estimates it in dry regions (Droogers and Allen 2002; Xu and Singh 2002). The main reason is probably that it does not include air humidity. Moreover, Hargreaves equation tends to over-estimate ET_0 in low ET areas and to under-estimate it in high ET areas (Droogers and Allen 2002; Xu and Singh 2002). These discrepancies may be explained because about 80% of ET_0 is explained by temperature and solar radiation (Samani 2000), while the remaining 20% is due to other factors. For example, when advection is severe, Hargreaves equation under-estimates ET_0 by up to 25% for daily periods (Berengena and Gavilán 2005). However, Hargreaves (1989) claims that incorporating ΔT into the equation compensates for the advective energy factor.

1.2 Hargreaves equation modification

The above mentioned problems with Hargreaves equation explain why local calibration or even modification is necessary before it can be applied to estimate ET_0 (Jensen et al 1997; Xu and Singh 2001, 2002).

Several attempts have been made to modify the equation by adding other weather factors. Allen (1995) proposes calculating KT as a function of the ratio between atmospheric pressure and the sea level atmospheric pressure; however, Samani (2000) argues that not all coastal weather stations would have the same coefficient and that an increase in elevation does not always imply a lower value of the coefficient. Attempts to adjust the equation to semiarid regions by including a wind function Jensen et al (1997) have resulted inconclusive (Droogers and Allen 2002; Martínez-Cob and Tejero-Juste 2004). Droogers and Allen (2002) proposed to include monthly rainfall as explanatory variable. However, due to the difficulties of including additional weather variables, some form of calibration seems a more convenient approach (Droogers and Allen 2002; Bautista et al 2009; Gavilán et al 2006; Jensen et al 1997; Martínez-Cob and Tejero-Juste 2004; Samani 2000; Vanderlinden et al 2004; Xu and Singh 2001). In addition, the need to use other meteorological variables than temperature would undermine the objective of having a temperature based method that could be used in most of the weather stations.

Orang et al (1995) proposed a regression-based correction for different periods of the year. After analysing results obtained by Trajkovic (2005) from a linear regression approach, Subburayan et al (2011) claimed that the parameter to be optimised is the exponent EH in Eq. 1. Shahidian et al (2013) analysed the seven most used calibration methods, concluding that neither distance to the coast nor temperature or a monthly calibration using solar radiation improve accuracy.

In Spain, Aguilar and Polo (2011) proposed a regionalization of KT in the river Guadalfeo basin using a linear regression to estimate PMFAO ET_0 separating the dry and wet seasons. Vanderlinden et al (2004), working with 16 stations in Andalusia (south Spain), carried out a regional calibration resulting in a correction based on the ratio ΔT, where T is the long-term annual average of air temperature. They found a linear relation between Hargreaves equation coefficients and the ratio ΔT, which may be used as a first fit of the Hargreaves coefficient. The same procedure was applied by Mendicino and Senatore (2013) in southern Italy with 137 stations, obtaining accurate results both in mountainous and coastal areas.

Gavilán et al (2006) calibrated the Hargreaves equation considering only temperature and wind conditions using data from a regional meteorological network installed in 2000 in Andalusia (south Spain). The procedure was then validated using additional weather data from other locations in the country. Results showed that, at coastal areas, the Hargreaves equation generally under-estimated ET_0, although in some specific locations it was more accurate or even over-estimated it. In general, the accuracy of the equation was less predictable in inland areas, ranging from moderate under-estimation to high overestimation.

1.3 ET_0 interpolation

Site estimations of ET_0 based on meteorological data may be useful for local irrigation management but are not representative over large areas (Kidron and Zohar 2010; Vicente-Serrano et al 2007). Heterogeneous regions would require detailed mapping to accurately describe the spatial variability of evapotranspiration. The feasibility of such approach depends on how difficult is to obtain distributed estimations of the variables needed for the ET_0 model used. Temperature is the only variable whose accurate interpolation is somehow straightforward, being another advantage of the temperature based methods. A few studies have tried to map ET_0 (Dalezios et al 2002; Häntzschel et al 2005; Mardikis et al 2005; Ray and Dadhwal 2001; Vicente-Serrano et al 2007) most of them without using topographical or other relevant geographic variables (Vicente-Serrano et al 2007). However, Mardikis et al (2005) showed how the inclusion of elevation improves the accuracy of the results. In fact, geomorphometric variables may be used as proxies for climatic processes (Böhner and Antonić 2009). Geostatistical methods using the effect of elevation and distance to the sea as ancillary information were compared by these authors with ordinary kriging for interpolating ET_0. The cross-validation results for annual ET_0 showed a significant increase in accuracy when such variables were taken into account.

1.4 Objectives

The objective of this paper is to present and evaluate a methodology to obtain distributed daily ET_0 series using temperature and other easy to obtain ancillary data. Such a methodology could be very useful for estimating ET_0 using historical temperature series or climate change scenarios; in addition, it would be helpful to obtain ET_0 maps in developing countries where complete weather stations are rare, at least in semiarid environments similar to the study area.

This methodology can be summarised in two steps: 1) To compare 14 daily ET_0 estimations based on different calibrations of the Hargreaves equation. Eight of them are annual and six are monthly calibrations. Ten are temperature based and four use both temperature and solar radiation. Such methods were included to quantify the reduction in accuracy when real solar radiation values are not used. The optimal calibration method in terms of both simplicity and accuracy, using PMFAO as reference, was chosen to 2) interpolate its parameters with a regression-kriging approach using Random Forest in the regression step calibrated with a set of ancillary variables that might act as proxies of the climatic variables included in PMFAO model but not in Hargreaves equation, selecting among them the most relevant. This way, part of the variability of ET_0 not imputable to temperature was incorporated into a flexible machine learning model instead of using a more classical linear model. The residuals of the Random Forest regression model were interpolated using ordinary kriging. Finally, the obtained parameter layers were used to produce improved ET_0 layers.

The same approach was used to interpolate daily Hargreaves and PMFAO ET_0 to compare the spatial patterns of the three ET_0 estimations.

2 Material and methods

2.1 Study area

The research was conducted in south-eastern Spain including both the River Segura Water Authority (DHS) controlled area (Fig. 1) that includes the Segura river basin (19000 km^2), and the Vinalopó river basin (3000 km^2). It is a semiarid area with scarce and irregular precipitation, high temperatures, and a large number of hours of sun that cause high potential evapotranspiration. Climate is closely related with topography; there is a NW-SE precipitation gradient from 1000 mm/year in the NW mountain area to less than 300 mm/year in the coastal areas. A similar, but inverse, gradient appears in temperature. The high population density and intensive irrigated agriculture represent a significant water demand.

2.2 Data-set

The 79 weather stations used in this study (Fig. 1, Tables 1(a) and 1(b)) are integrated in the Spanish Agroclimatological Information System for Irrigation (*Sistema de Información Agroclimática para el Regadío*, SIAR), maintained by the spanish Min-

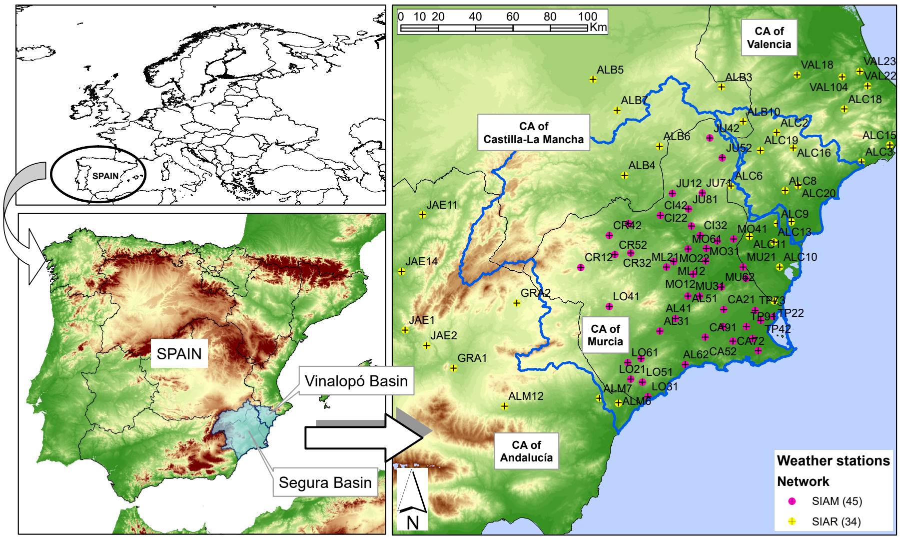


Fig. 1 Study area including the 45 weather stations of the SIAM network and the 34 stations of the SIAR network used in this study

istry of Agriculture, Food and Environment. Forty-five of them belong to the Murcia sub-network (SIAM) and the other 34 to the neighbouring regional sub-networks.

In these stations, ET_0 is not directly measured but estimated using the PMFAO equation, as all the stations include sensors to measure the variables needed, and those sensors are located in standard conditions following the recommendations of the World Meteorological Organisation (WMO), the Spanish Weather Agency (*Agencia Estatal de Meteorología*, AEMET), FAO and American Society of Agricultural Engineers (ASAE). The stations of both networks have similar characteristics as defined by the Spanish Ministry of Agriculture's Food and Environment. The stations are mounted on 2 m high tripods and installed on plots measuring 10mx10m that are enclosed by a metallic fence. Table 3 summarises the most commonly used sensors, although they can change from one station to the other. The thermometers are of platinum resistance, with a range of $-39.2-60^oC$ and a precision of $\pm0.2^oC$; The solid state capacitive humidity sensors have a range of $0.8-100\%$ and a precision of $\pm0.2\%$; The tipping bucket rain gauges have a precision of $0.2mm$; the anemometer consists of a four-blade propeller with a speed range of $1-100ms^{-1}$ and a precision of $\pm0.3ms^{-1}$ from 1 to 60 ms^{-1}; the radiation sensor is a pyranometer with sensitive silicon photocells working in a $350-1100$ nm range with a precision of 3%. A more detailed description of both stations and instruments can be consulted in the web sites of the SIAM [1] and SIAR [2] projects.

[1] http://siam.imida.es/apex/f?p=101:1000:5736042974903220

[2] http://www.mapama.gob.es/es/desarrollo-rural/temas/gestion-sostenible-regadios/sistema-informacion-agroclimatica-regadio/presentacion.aspx

Table 1 (a) Characteristics of the 79 weather stations considered in the study

Code	X(m)	Y(m)	Altitude(m)	Dist.Sea(km)	Avg.T(^{o}C)	Prec.(mm)	ET_0(mm)	Start-End year
AL31	631134	4177380	234	24.6	16.94	271.0	1294.2	2003-2014
AL41	639493	4184116	167	26.7	17.06	250.0	1383.0	2003-2014
AL51	646202	4196165	164	36.7	17.47	294.2	1284.0	2003-2014
AL62	644794	4159467	110	2.2	18.30	173.5	1163.0	2003-2014
ALB10	675774	4289496	531	61.2	15.17	340.1	1386.0	2003-2014
ALB3	664184	4307954	698	75.7	13.98	352.6	1316.3	2003-2014
ALB4	612300	4260690	579	99.6	14.56	338.5	1193.4	2003-2013
ALB5	595314	4311962	677	136.4	13.35	351.6	1330.0	2003-2014
ALB6	630940	4276220	682	90.9	14.82	322.2	1351.9	2003-2014
ALB7	608291	4295449	872	120.2	13.74	368.6	1301.7	2003-2014
ALC10	695317	4211642	58	3.3	18.25	271.8	1193.2	2003-2014
ALC101	701609	4235864	62	9.4	17.78	236.7	1174.5	2003-2014
ALC11	679120	4227949	94	26	17.39	264.6	1102.0	2003-2014
ALC12	692405	4193488	50	4.8	18.00	349.0	1185.7	2003-2014
ALC13	692274	4225086	4	14.7	17.79	294.5	1176.8	2003-2014
ALC15	754321	4276842	76	2.9	17.88	404.8	1154.5	2003-2014
ALC16	702610	4275463	664	31.6	14.48	396.2	1232.1	2003-2012
ALC18	729979	4296196	447	28	15.68	737.1	1058.2	2003-2014
ALC19	685024	4274032	488	43.8	14.72	307.5	1329.9	2003-2014
ALC2	693722	4283544	592	43.5	14.39	363.0	1207.1	2003-2014
ALC20	705164	4255200	282	16.2	16.55	272.6	1283.2	2003-2014
ALC3	739182	4267965	74	3.4	18.00	332.8	1175.1	2003-2014
ALC6	669278	4255019	629	47.1	15.09	320.7	1305.9	2003-2014
ALC8	698210	4252499	259	21	16.73	308.3	1237.0	2003-2014
ALC9	694012	4234860	73	16.2	18.69	266.9	1165.6	2003-2014
ALM12	547943	4137207	796	60.2	14.74	316.0	1279.5	2003-2014
ALM6	608951	4138960	185	9.2	17.47	257.0	1418.8	2003-2014
ALM7	598844	4141406	317	19.3	16.49	297.4	1246.8	2003-2014
CA12	680785	4173450	30	9	18.17	311.5	1257.8	2005-2014
CA21	665320	4188975	225	26.3	17.29	240.8	1391.2	2003-2014
CA42	664924	4179770	136	19	17.48	312.3	1227.7	2003-2014
CA52	670233	4171939	84	10.4	17.67	284.4	1385.2	2003-2014
CA72	683796	4166811	63	5.8	17.13	359.2	1185.7	2003-2014
CA91	655462	4174084	174	13.4	17.53	260.5	1199.3	2003-2014
CI22	648038	4233496	281	55.2	17.64	232.0	1301.9	2003-2014
CI32	652671	4228482	236	49	17.08	247.0	1513.9	2003-2014
CI42	631394	4239282	253	72.9	16.77	265.8	1246.9	2003-2014
CI52	614311	4234953	274	83.2	16.38	329.0	1203.7	2003-2014
CR12	588796	4211407	866	77.7	12.20	309.7	1296.3	2003-2014
CR32	615553	4219246	432	68.6	15.20	318.7	1237.3	2003-2014
CR42	604030	4228626	454	82.4	15.93	337.0	1374.3	2003-2014
CR52	607082	4218310	506	72	15.91	333.5	1147.8	2003-2014
GRA1	520628	4157712	814	84.8	14.55	390.3	1288.1	2003-2014
GRA2	554482	4192456	1110	85.2	12.87	336.4	1380.7	2003-2014
JAE1	494672	4177995	793	110.5	15.34	385.3	1674.4	2003-2014
JAE11	504003	4239630	510	154.3	15.21	615.1	1180.5	2003-2014
JAE14	492827	4209275	571	141.8	16.41	568.3	1077.0	2003-2014
JAE2	506267	4169627	893	100.6	13.92	472.4	1171.8	2003-2014
JU12	637803	4251007	394	72.9	16.02	281.8	1296.9	2003-2014
JU42	657918	4280624	657	69.8	13.99	323.2	1259.4	2003-2014
JU52	664558	4270147	565	59	14.49	251.8	1241.3	2003-2014
JU71	653858	4251259	405	59.5	16.59	275.5	1040.3	2003-2014
JU81	646480	4242623	341	61.1	16.53	237.3	1217.5	2003-2014
LO11	621083	4162736	323	19.8	16.73	269.6	1283.3	2003-2014
LO21	615537	4151777	356	14.5	16.09	253.4	1328.8	2003-2014
LO31	624681	4142445	30	1.8	18.50	220.5	1319.9	2003-2014
LO41	604060	4190613	692	51.9	14.46	300.0	1269.2	2003-2014
LO51	621756	4150090	329	10	17.75	237.3	1375.5	2003-2014
LO61	613917	4160518	310	22.7	16.76	241.5	1319.3	2003-2014
ML12	638577	4214494	263	56.5	14.55	369.3	1294.3	2003-2014
ML21	634664	4211679	274	54.5	18.68	288.2	1224.5	2003-2014
MO12	648990	4208022	157	47.9	17.79	269.8	1280.3	2003-2014
MO22	656063	4221660	142	43.5	17.98	259.5	1353.7	2003-2014
MO31	655664	4214906	80	42.5	15.91	283.0	1236.1	2003-2014
MO41	670577	4226616	138	32.2	17.56	234.9	1287.1	2003-2014
MO51	661952	4225502	196	39.3	17.75	250.5	1295.5	2003-2014
MO61	646290	4221343	161	52.9	18.42	309.9	1435.0	2003-2011
MU21	675661	4211733	27	22.2	17.59	298.8	1203.6	2003-2014
MU31	652374	4196142	138	35.6	19.69	307.3	1286.1	2003-2014
MU52	677199	4205450	134	20.7	18.28	306.3	1396.0	2003-2014
MU62	664029	4201022	53	33.2	18.21	329.3	1204.1	2003-2014

Table 1 (b) Characteristics of the 79 weather stations considered in the study

Code	X(m)	Y(m)	Altitude(m)	Dist.Sea(km)	Avg.T(^{o}C)	Prec.(mm)	ET_0(mm)	Start-End year
TP22	692087	4185147	6	1.8	18.45	309.7	1352.4	2003-2009
TP42	685178	4182992	31	5.7	16.93	359.1	1296.6	2003-2014
TP73	682155	4188493	88	11.5	17.92	286.9	1444.9	2003-2014
TP91	677479	4179933	53	11.7	17.94	170.1	1302.6	2005-2014
VAL104	728810	4313267	295	19.2	15.68	518.5	1147.3	2003-2014
VAL18	704644	4314386	247	40.5	17.06	512.7	1098.5	2003-2014
VAL22	742471	4308516	86	10.4	17.68	761.6	1051.7	2003-2014
VAL23	738202	4316412	104	9.3	16.90	609.5	1064.9	2003-2014

Table 2 Characteristics of the sensor in the weather stations.

Sensor	Brand	Model	Height (m)	Data capture freq.(s)	Measured parameter
Datalogger	Campbell	CR10X	-	-	-
Anemometer	Young	05103-5	2	5	Wind speed (ms^{-1}) and direction (^{o})
Pyranometer	Kipp & Zonen	CMP7	2	5	Solar radiation (W/m^2)
Rain gauge	Thies-Clima	ARG100	1	5	Precipitation (mm)
Termo-higrometer	Vaisala	HMP45C	1.5	120	Temperature (^{o}C) and relative humidity (%)

The data were validated by the Ministry of Agriculture, Food and Environment, according to the Spanish UNE 500540:2004 norm (AENOR 2004), testing internal, temporal and spatial consistency.

In this research, we used the period 2003-2014; odd years were used to calibrate the models and even years to validate.

2.3 Hargreaves equation calibration

Fourteen calibration approaches of the Hargreaves equation were used to generate daily ET_0 series (Table 3). *H* uses just Eq. 3. *HRs* uses Eq. 2 with real incident solar radiation (R_S) available at the weather stations. *Al1* is obtained with a linear model to estimate PMFAO results using Hargreaves equation as independent variable. However, as ET_0 is defined in the range $[0,+\infty]$, it is assumed that the intercept (b_0) should be zero, so this estimation uses just the slope (b_1) coefficient. *Al2* is obtained as the previous one but using the two parameters (b_0 and b_1). *Al1m* is similar to Al1, but calculating a different slope parameter for each month. *Al2m* is similar to Al2, but calculating both parameters for each month. *K1* includes the one parameter linear model and a clarity index to calibrate EH and KT using real R_S data. Clarity index, initially defined by Liu and Jordan (1960), estimates the atmospheric transmittance by means of the relation between the measured solar radiation and the extraterrestrial radiation (R_S/R_a). Following (Subburayan et al 2011), if the exponent 0.5 in Eq. 2) is considered a parameter that should be optimized, then it can be linearised as:

$$\log\left(\frac{R_S}{R_a}\right) = EH \cdot \log\left(\Delta T\right) \cdot \log\left(KT\right) \quad (4)$$

Table 3 Models proposed in this study; the number of parameters is included between parentheses

Abbreviation	Description
H	Hargreaves equation. with R_a
HRs	Hargreaves equation with R_S
Al1	Linear calibration of H with $b_0 = 0$ (1 parameter)
Al2	Linear calibration of H (2 parameters)
Al1m	Monthly linear calibration of H with $b_0 = 0$ (12 parameters)
Al2m	Monthly linear calibration of H (24 parameters)
ARs1	Linear calibration of HRs with $b_0 = 0$ (1 parameter)
ARs2	Linear calibration of HRs (2 parameters)
ARs1m	Monthly linear calibration of HRs with $b_0 = 0$ (12 parameters)
Ars2m	Monthly linear calibration of HRs (24 parameters)
K1	Clarity index calibration with $b_0 = 0$ (3 parameters)
K2	Clarity index calibration (4 parameters)
K1m	Monthly clarity index calibration with $b_0 = 0$ (24 parameters)
K2m	Monthly clarity index calibration (36 parameters)

K2 includes the two parameter linear model and a clarity index to calibrate EH and KT using real R_S data. *K1m* and *K2m* are similar to K1 and K2, but all the parameters are calibrated monthly. *ARs1* is similar to Al1 but using real R_S data. *ARs2* is similar to Al2 but using real R_S data. *ARs1m* is similar to Al1m but using real R_S data. Finally, *ARs2m* is similar to Al2m but using real R_S data.

K1, K2, K1m and K2m use real R_S data only to calibrate the model in the available weather stations but it is not needed in prediction. The last four methods use real R_S data both to calibrate and to predict.

2.4 Validation

Bennett et al (2013) state that a single goodness of fit statistic is not enough to select the most accurate model as such statistics measure different performance aspects. Three statistics were used to evaluate calibration and validation error, a detailed description of these statistics and their use with hydrological series, including criteria for their interpretation, can be found in (Legates and McCabe 1999) or (Bennett et al 2013).

Root mean square deviation (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(O_i - E_i)^2}{n}} \tag{5}$$

The modified Nash-Sutcliffe efficiency (nse) determines the relative magnitude of the residual variance compared to the measured data variance, it is less sensitive than R^2 to extreme values (Legates and McCabe 1999):

$$nse = 1 - \frac{\sum_{i=1}^{n}(O_i - E_i)^2}{\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2} \tag{6}$$

Percent bias (PBIAS) measures the average tendency of the estimated values to be greater or smaller than the observations. It is not as sensitive as RMSE to extreme values or to the magnitude of the variables:

$$PBIAS = \frac{\sum_{i=1}^{n} (O_i - E_i)}{\sum_{i=1}^{n} O_i} \tag{7}$$

The three error estimation statistics were calculated in each weather station. As normality and homocedasticity could not be assumed, a Kruskal Wallis contrast was used to test if there were significant differences among the methods. When significative differences among models were observed, a post-hoc contrast between pairs of models based on Mann-Whitney and using Holm method to correct p-values among classes, was performed to discover groups of non-significantly different methods.

The results of these three goodness of fit statistics may be contradictory; it would also be important to take into account not only means but also the dispersion of the statistics. Thus, we summarised both aspects of the three statistics in a distance to ideal point measurement taking into account both the median and the median absolute deviation (MAD) of each goodness of fit statistic. We transformed the medians of nse (mnse) as 1 - mnse and the medians of PBIAS as their absolute values. In this way, the optimum values of the six statistics is 0 and the larger their values the further from the ideal point. In a space defined by these six statistics, the best model may be identified as the closest to the point of origin (the six statistics equal to 0) using euclidean distance in this 6-statistics feature space.

2.5 Parameter interpolation

Regression-kriging (Hengl et al 2004) is an interpolation framework that uses a regression method to predict the spatial distribution of the variable, and the interpolation of its residuals if the semivariogram shows a clear spatial structure. Both layers, regression prediction and interpolated residuals, are finally added. Usually, the regression part is achieved by a Generalized Lineal Model (GLM) (McCullagh and Nelder 1989), but in this case we used Random Forest (RF) (Breiman 2001) as it obtained more accurate results than GLM in previous tests. RF is an ensemble method based on decision trees used both in classification and in regression problems. It provides a measurement of the importance of the variables in the prediction. It is also possible, when used in regression, to use dependence plots to discover the effect of each predictor on the dependent variable.

Twelve variables were used as independent predictors: 1) Distance to the coastline or to the coastal lagoons as a continentality indicator (DIST), 2) Height derived from the official DEM of the *Instituto Geográfico Nacional* (Spanish National Geographic Institute) with a resolution of 25 m, 3) Daily potential irradiation (RPOT) obtained from the DEM with the methodology proposed by Hofierka and Súri (2002), 4) Terrain slope in degrees (SLP) obtained from the DEM, 5-6) Sine and cosine of terrain aspect (SASP y CASP) obtained from the DEM. CASP represents the north (CASP=1) and south (-1) hillslopes, whereas SASP represents the east (1) west (-1) orientation, 7-8) Profile (in the direction of the slope) and transverse terrain curvatures (PCURV and TCURV) obtained form the DEM. Both variables convey information concerning land concavity or convexity. Positive values indicate concavity and negative values indicate convexity. Finally, four easy to interpolate climatic variables were added: maximum, minimum and average temperature, and precipitation layers. Such layers were estimated with 200 m resolution using regression-kriging (Gomariz-Castillo and Alonso-Sarría 2013).

The first eight variables are important as they might be proxies for the meteorological variables influential on ET_0 but not included in the Hargreaves equation. Daily potential irradiation is a clear proxy for incident radiation; height is a proxy for the ratio between atmospheric pressure and sea level pressure and also for wind intensity, higher in higher altitudes; distance to the sea is a proxy for both wind intensity, higher in the coast than in the interior, and for the relation $T/\Delta T$; finally, geomorphometric variables might alter the amount of radiation reaching the surface.

Before using RF, variables with a large Variance Inflation Factor (Fox and Monette 1992) were iteratively eliminated until the VIF of all features were below a threshold of 10. A cross-validation approach described in Kuhn and Johnson (2013) was used to select the number of variables that optimise accuracy in the RF model, a procedure very similar to the step-wise regression. The residuals of the RF model were finally interpolated using ordinary kriging.

We used the same approach to obtain daily layers of Hargreaves and PMFAO ET_0 estimations. The Hargreaves estimation was then used to obtain Al2 ET_0 estimations using a simple map algebra equation: $Al2 = b_0 + b_1 \cdot H$. The three sets of 365 layers (H, PMFAO and Al2) were finally added into three annual ET_0 estimations to compare results.

3 Results and discussion

3.1 Hargreaves equation calibration and validation

Fig. 2 summarises the values of the three goodness of fit statistics and the euclidean distance of each ET_0 calibration method to the optimal point in the 6-statistics feature space. The groups obtained with the post-hoc contrast are represented by letters. The clearest results are that the more complex the model (in terms of numbers of parameters), the more accurate. In addition, models using R_S perform better than the corresponding using R_a, but the differences are only substantial when no calibration is done (H vs. HRs) and in models with a large number of parameters (Al2m vs. ARs2m).

The best option, in terms of RMSE (Fig. 2 at top left), is AlRs2m, one of the most complex approaches with 24 parameters and using R_S. In second place are the other calibrated methods based on R_S (AlRs1, AlRs2, AlRs1m and AlRs2m). Calibrated methods based on R_a appear in third place and H shows the highest RMSE median and MAD with several cases of very high error, the results derived from nse (Fig. 2 at top right), are quite similar to those obtained for RMSE.

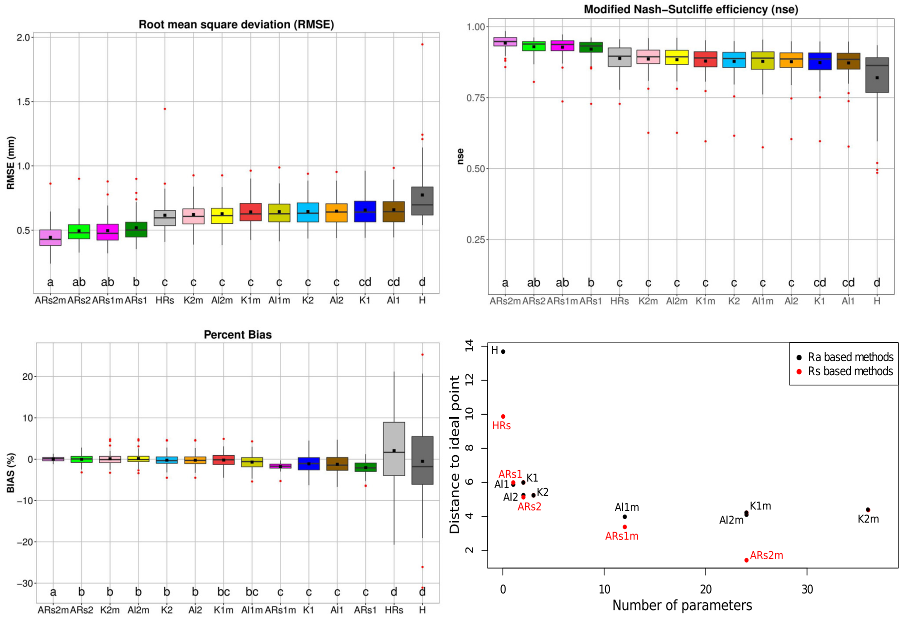


Fig. 2 RMSE, nse, PBIAS and distance to the ideal point of the different calibration models (see Table 3) in the validation dataset. The meaning of the different labels can be consulted in Sect. 2.3. The letters a to b indicate the obtained groups of significantly different models (Mann-Whitney using Holm method, $p < 0.05$).

However, the results obtained for PBIAS (Fig. 2 at down left) are different. The 24 parameter model ARs2m appears as the best model followed by a group of quite different models among which 2 simple models: K2 (3 parameters) and Al2 (2 parameters) perform similarly or even better than more complex models. Interestingly, a third group include all regression models without intercept (both yearly and monthly models and both R_S based and R_a based models); this models show a tendency to slightly infraestimate ET_0. Finally, the non-calibrated models H and HRs appear in a fourth group with a quite larger dispersion in the PBIAS values.

Fig. 2 at down right shows the distance to the ideal point. The increase in the number of parameters only produces a clear decrease in distance in those models that use R_S instead of R_a. In models that use R_a, a minimum distance around 4 is reached with 10 parameters. The two candidate models that offer a better trade off between accuracy and simplicity are ARs2 (slightly more accurate but using real R_S values) and Al2, that is a model using only temperature with an accuracy only slightly lower than ARs2. Interestingly, the use of R_S only improves accuracy when the number of parameters is large (monthly calibrations) or when no calibration is done; it does not improve accuracy in the more parsimonious annual calibrations.

After analysing all these results, we decided to use Al2 as the calibration approach to estimate ET_0 from the Hargreaves equation. Its main advantages are simplicity, only 2 parameters, and accuracy. A2Rsm gives more accurate values; however, it uses 24 parameters, its use of R_S is a real problem because this variable is not available in most weather stations, and its estimation would need complex calibration. In addition, it could be argued that the accuracy of Al1m is a bit higher without using R_S, but at the cost of multiply six-fold the number of parameters.

As Shahidian et al (2012) points out, H has, in general, produced good results because at least 80 percent of ET_0 can be explained by temperature and solar radiation (Jensen et al 1997) and ΔT is related to humidity and cloudiness (Samani and Pessarakli 1986). The proportion of ET_0 variability not explained by H can be mainly explained by other climatic factors not included in the equation. We analysed the relationship between the ratio $PMFAO/H$ and several monthly climatic variables (mean, maximum and minimum temperatures; mean, maximum and minimum relative humidity; mean and maximum wind speed, radiation, total precipitation and the ratio between the number of sunshine hours and the total amount of hours, to infer cloudiness). $R^2 < 0.1$ in all cases except in mean wind speed ($R^2 = 0.672$), maximum wind speed ($R^2 = 0.293$) and mean relative humidity ($R^2 = 0.215$); Fig. 3 shows the relationship between the ratio $PMFAO/H$ vs. the mean wind speed (a)) and the maximum relative humidity (b).

With regard to the wind effect (Fig. 3a), a high non-linear positive relation is observed, indicating an underestimation of ET_0 in observatories located in areas with high relief (northwest of the study area and in the narrow areas of river valleys). The same effect has been detected in other studies (Trajkovic 2005; Raziei and Pereira 2013) that examine the wind effect on a wide variety of climates in Iran, concluding that extreme values and temporal variability in wind speed cause discrepancies between H and PMFAO, especially in arid and hyper-arid climates. Wind speed is considered to be an important variable in arid climates, whereas the number of sunshine hours is considered to be the more dominant variable in sub-humid and humid climates (Shahidian et al 2012). The reason is that wind removes saturated air from the boundary layer and thus increases evapotranspiration (Brutsaert 1991).

The effect of maximum relative humidity is lower although still significant. This variable can overestimate ET_0 in humid regions (Lu et al 2005; Trajkovic 2005) because the model was calibrated for the semi-arid conditions of California, and does not explicitly account for relative humidity (Shahidian et al 2012). This effect is shown in the negative trend (Fig. 3b) indicating that ET_0 overestimation increases with maximum relative humidity. The higher values of this variable usually occur in observatories near the coast.

3.2 Interpolation of calibration parameters

Once Al2 is selected as the best calibration approach, its 2 parameters were interpolated using a RF based regression-kriging approach. The VIF procedure removed elevation and average temperature because of their high correlation with the other temperature variables. The cross-validation approach selected six variables, in order of importance, for b_0: Distance to the coastline, Annual precipitation, Minimum temperature, Slope, Potential radiation and Maximum temperature; and four variables, in order of importance, for b_1: Minimum temperature, Distance to the coastline, Maximum temperature and Annual precipitation. Fig. 4 and Fig. 5 show the observed data,

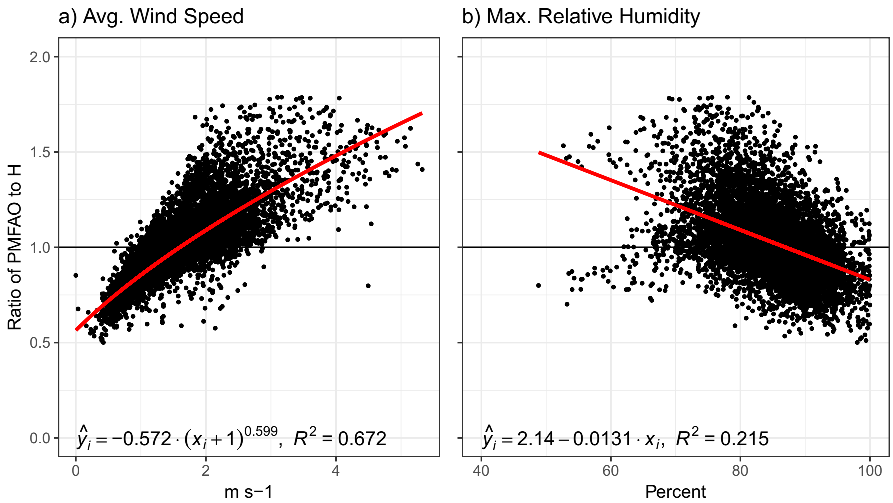


Fig. 3 Effects of a) wind and b) maximum relative humidity on the ratio $PMFAO/H$. The fitting lines were estimated using Weighted Least Squares regression to reduce the effect of heterocedasticity on the residuals. Values higher than 1 in the ratio $PMFAO/H$ indicate a subestimation of ET_0 in H with respect to PMFAO. Each dot corresponds to a monthly average in one of the observatories

the effects of these variables, and a 95 % confidence interval, on b_0 and b_1 respectively.

A high b_0 coefficient indicates a tendency of the Hargreaves model to underestimate ET_0. Distance to the coastline might be, in the b_0 estimation, a proxy variable to humidity, with an increase in b_0 until 45 km where it stays steady. Annual precipitation might be a proxy for cloudiness, as b_0 is higher until 320 mm/year and drops substantially for higher values.

The combined effect of maximum and minimum temperature shows that the higher ΔT the higher b_0. This way, we are indirectly introducing the corrections to EH and KT coefficients proposed by Hargreaves (1994a), Allen (1995) or Gavilán et al (2006). Potential irradiation, an easy to obtain proxy for global radiation, is correcting the absence of R_S in Hargreaves equation as b_0 clearly increases where this variable increases. Finally, the effect of the slope is not very clear. Depending on the aspect, incident radiation might increase with slope; however, aspect related variables were not included in the RF model, moreover, the slope range of the weather stations is quite limited.

While b_0 allows to correct a global bias in ET_0 estimation, b_1 is related with how the under(over)estimation of PMFAO ET_0 increases when Hargreaves estimation increases. The effect and explanation of annual precipitation is similar to b_0. The effect of the distance to coastline is the opposite to the b_0 case. It is related with the underestimation of ET_0 using Hargreaves equation (Gavilán et al 2006). It is probably related with the higher wind velocities in the coastal sector of the study area. The combined effect of maximum and minimum temperature has also an opposite effect to the observed in b_0; however, we think that it is also indirectly correcting KT and

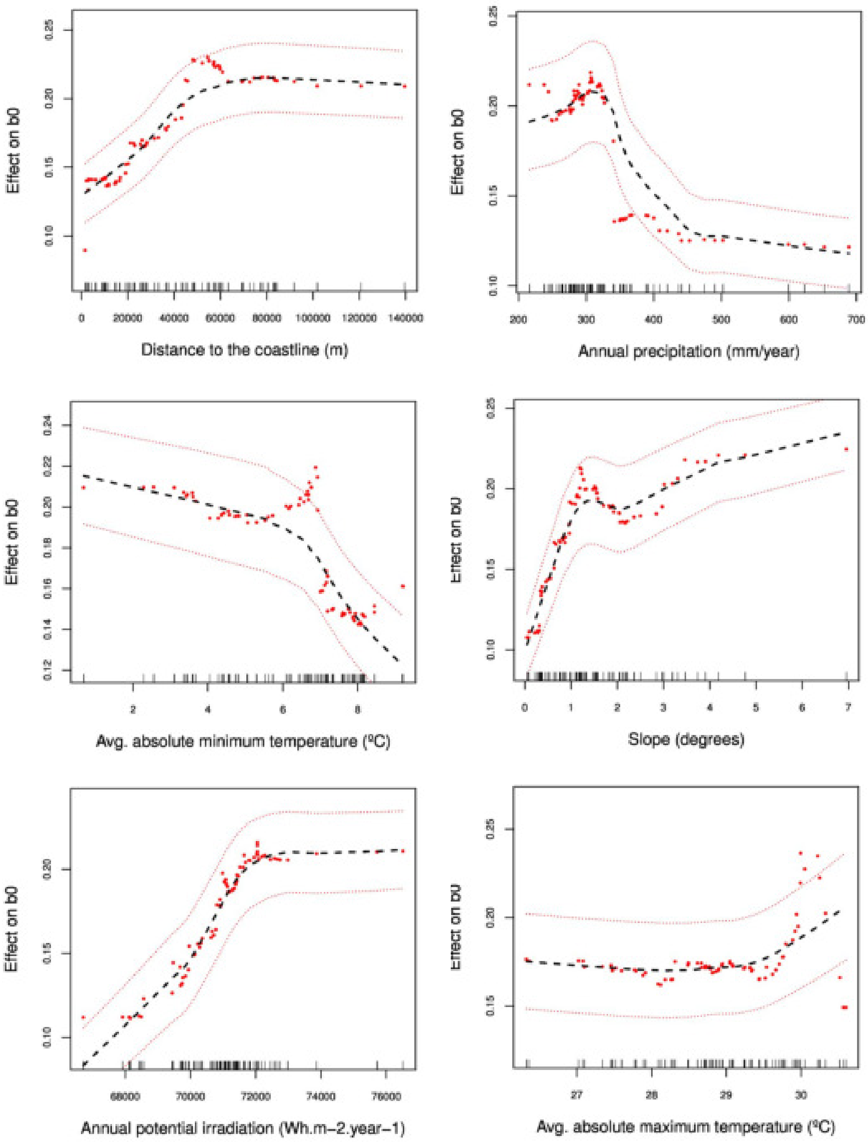


Fig. 4 Effects on b_0 of the ancillary variables according to the Random Forest model. The black dashed line represent the estimated effect, the dotted red lines the 95 % confidence intervals and the red dots the observed values.

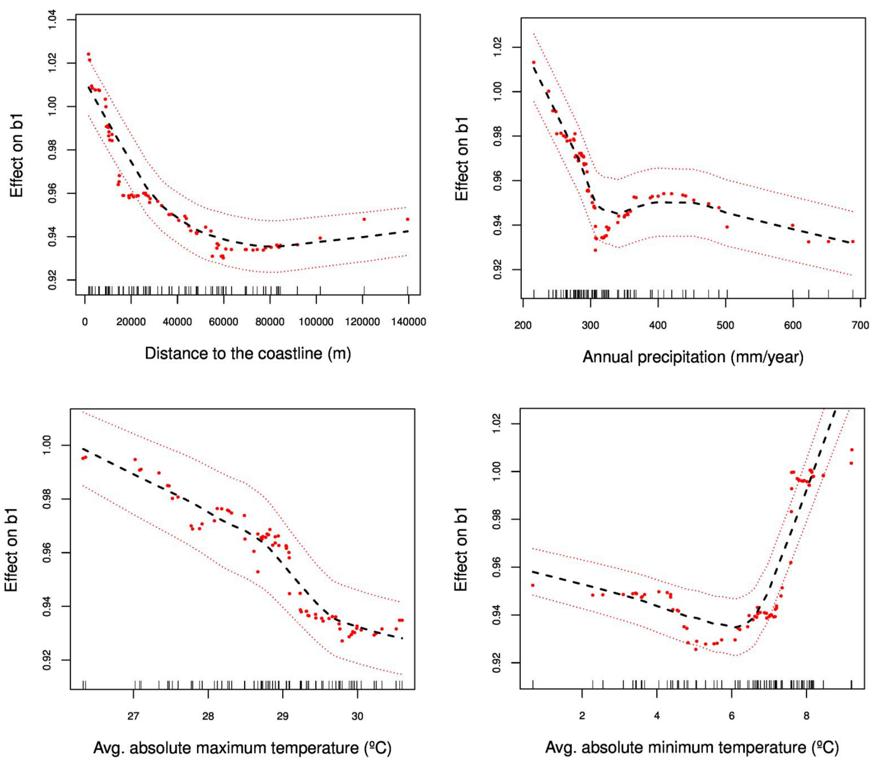


Fig. 5 Effects on b_1 of the ancillary variables according to the Random Forest model. The black dashed line represent the estimated effect, the dotted red lines the 95 % confidence intervals and the red dots the observed values.

EH coefficients. In addition, as MDE was removed from the variable set because of its high correlation with temperature variables, they might also be indirectly including in the model variables related with elevation, such as wind velocity.

RMSE of the Random Forest model is 0.2157 for b_0 and 0.0831 for b_1, lower than the standard deviation of the values in the weather stations (0.238 and 0.112 respectively). When the residuals are interpolated, accuracy improves slightly to 0.2148 and 0.0825 respectively. Fig. 6 shows the final b_0 and b_1 layers, after interpolating the residuals with ordinary kriging.

Using RF is more flexible than linear modelsm including Generalized Linear Models (McCullagh and Nelder 1989) or Generalized Additive Models (Wood 2006), in taking into account non-linear relations and interactions among predictive variables. So, instead of trying to discover an explicative functional relationship (an equation) among these variables and ET_0, RF produces a purely predictive model. It is noteworthy that in the north-west sector of the study area, where there is al-

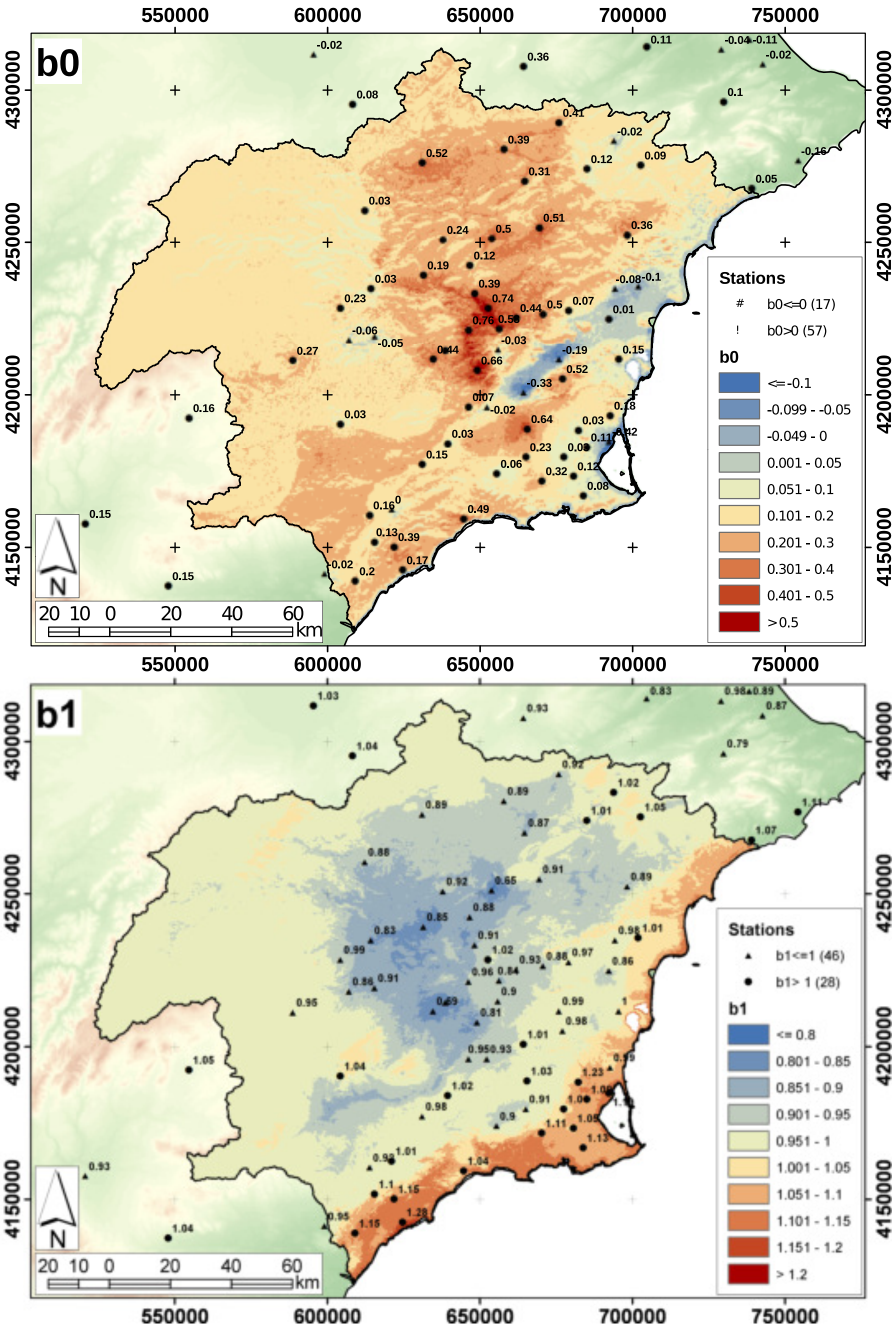


Fig. 6 b_0 and b_1 spatial distribution according to the RF-OK model

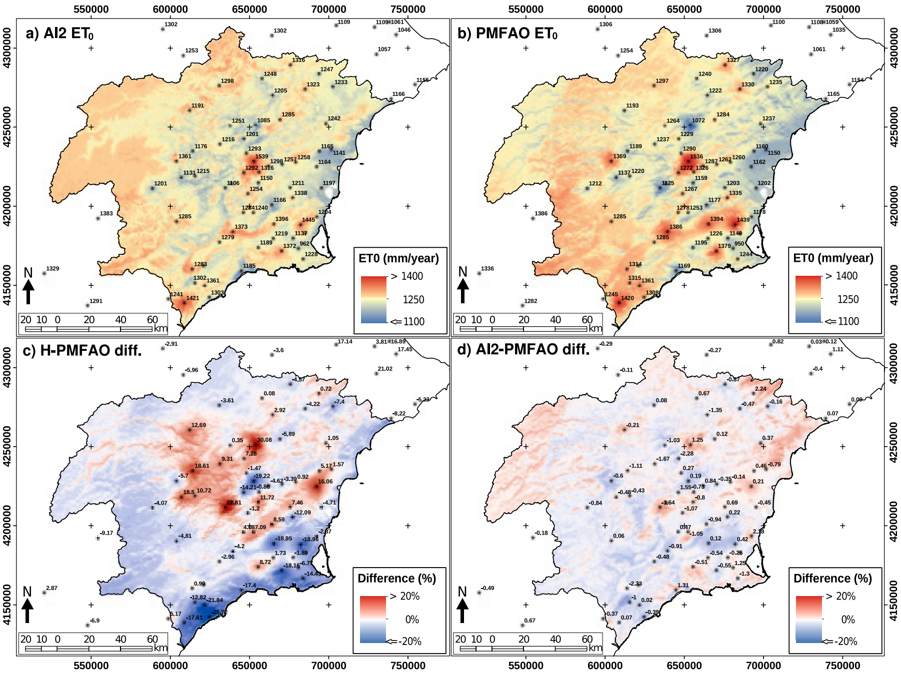


Fig. 7 Spatial distribution of Al2 (a) and PMFAO (b) annual ET_0 estimations; and differences between Hargreaves and PMFAO annual ET_0 estimations (c), and Al2 and PMFAO annual ET_0 estimations (d).

most no weather stations, the extrapolation is quite more conservative that would be expected using linear models.

3.3 Mapping ET_0

Daily Hargreaves and PMFAO ET_0 interpolations were also interpolated using RF based regression-kriging. The selected variables in the Hargreaves case were minimum absolute temperature (96.2% of the days), distance to the coast (87.4%) and maximum absolute temperature (77%). For the PMFAO estimations, the most often selected variables were total precipitation (92.9%), potential radiation (98%), minimum absolute temperature (76.2%), maximum absolute temperature (52.6%) and distance to the coast (55.9%). These results reflect that Hargreaves uses only temperature and ΔT (related with the distance to the coast) but PMFAO uses other variables, mainly radiation which is approximated by potential radiation and total precipitation (cloudiness).

The annual maps were obtained by adding the daily layers, they appear in Fig. 7 a and b. The differences between Hargreaves and PMFAO and between Al2 and PMFAO appear in Fig. 7, c and d respectively.

The spatial patterns of Al2 and PMFAO are very similar. The statistics of the differences between Al2 and PMFAO (M=-0.03, SD=2.98, Min=-11.7, Max=16.4) show a quite higher accuracy that those of the difference between H and PMFAO (M=-1.39, SD=5.93, Min=-29.8, Max=40.3).

The spatial pattern of Al2 also reproduces PMFAO pattern, whereas the map of differences between Hargreaves and PMFAO reflects how ET_0 is underestimated in the coastal, more arid, area and infraestimated in the interior. These results coincide with those obtained by Droogers and Allen (2002), Xu and Singh (2002) or Berengena and Gavilán (2005).

4 Conclusions

Any type of calibration increases the accuracy of the Hargreaves equation to estimate PMFAO ET_0. However, more complex calibrations, including a large number of parameters and real solar radiation values, are not really needed. Allen calibration of the Hargreaves equation using just a linear regression model produces quite accurate results. The accuracy of the methods that use real R_S and a high number of parameters was not much higher.

Random Forest provides predictive models for both b_0 and b_1 using environmental variables whose effects can be represented in dependence plots. The variables selected by the model as more important are Distance to the coastline, Annual precipitation, Minimum temperature, Slope, Potential radiation and Maximum temperature for b_0 and Minimum temperature, Distance to the coastline, Maximum temperature and annual precipitation for b_1.

The annual ET_0 spatial distribution and individual values estimated with our proposed model are quite close to those of PMFAO; however, our method uses only temperature and some easy to obtain ancillary variables.

Interpolating the parameters of a simple calibration using a flexible predictive model based on easy to obtain ancillary variables, instead of producing a complex calibration of Hargreaves equation, might be a useful approach to obtain ET_0 maps in data scarcity scenarios: underdeveloped countries, past climate reconstruction of climate change projections.

After the results obtained in this paper we think there is room for the research of more sophisticated predictive models to estimate the spatial distribution of ET_0 or, in general, any environmental variables. Acquiring more complete data sets is a very expensive and long process while exploring these methods is almost costless. In addition, there are several scenarios (developing countries, historical studies and climate change projections) in which it would be very difficult or just impossible to acquire new data.

Acknowledgements This study was funded by the Segura Hydrographic Confederation, MAPAMA (Government of Spain), within the framework of the Research Project *Scientific Support Activities for the Hydrological Planning and European Cooperation Process*.

This study was funded in part by the Seneca Foundation, Science and Technology Agency of the Murcia Region (Ref. 19325 / PI / 2015), project *The Role of Water Markets in Integrated Water Resources Management in Watersheds*.

References

AENOR (2004) Automatic weather stations networks: Guidance for the validation of the weather data from the station networks. real time validation. Tech. rep., Asociación Española de Normalización y Certificación

Aguilar C, Polo MJ (2011) Generating reference evapotranspiration surfaces from the hargreaves equation at watershed scale. Hydrology and Earth System Sciences 15(8):2495–2508, doi:10.5194/hess-15-2495-2011

Allen RG (1995) Evaluation of procedures for estimating mean monthly solar radiation from air temperature. Tech. rep., Report submitted to the United Nations Food and Agricultural Organization (FAO), Rome Italy

Allen RG, Jensen ME, Wright JL, Burman RD (1989) Operational estimates of evapotranspiration. Agronomy Journal 81:650–662

Allen RG, Smith M, Perrie M, Pereira LS (1994) An update for the calculation of reference evapotranspiration. ICID Bulletin 43(2):35–92

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration: Guidelines for computing crop requirements. 56, FAO, doi:10.1016/j.eja.2010.12.001

Bautista F, Bautista D, Delgado-Carranza C (2009) Calibration of the equations of hargreaves and thornthwaite to estimate the potential evapotranspiration in semi-arid and subhumid tropical climates for regional applications. Atmosfera 22(4):331–348

Bennett ND, Croke BFW, Guariso G, Guillaume JHA, Hamilton SH, Jakeman AJ, Marsili-Libelli S, Newham LTH, Norton JP, Perrin C, Pierce SA, Robson B, Seppelt R, Voinov AA, Fath BD, Andreassian V (2013) Characterising performance of environmental models. Environmental Modelling & Software 40:1–20, doi:10.1016/j.envsoft.2012.09.011

Berengena J, Gavilán P (2005) Reference evapotranspiration estimation in a highly advective semiarid environment. Journal of irrigation and drainage engineering 131(2):147–163, doi:10.1061/(ASCE)0733-9437(2005)131:2(147)

Böhner J, Antonić O (2009) Chapter 8 land-surface parameters specific to topo-climatology. In: Hengl T, Reuter HI (eds) Developments in Soil Science, vol 33, Elsevier, pp 195–226, doi:10.1016/S0166-2481(08)00008-1

Breiman L (2001) Random forests. Machine Learning 45(1):5–32, doi:10.1023/A:1010933404324

Brutsaert W (1991) Evaporation into the Atmosphere. Theory, History and Applications. D. Reidel Publishing Company, Boston

Dalezios NR, Loukas A, Bampzelis D (2002) Spatial variability of reference evapotranspiration in greece. Physics and Chemistry of the Earth 27(23-24):1031–1038, doi:10.1016/S1474-7065(02)00139-0

Di Stefano C, Ferro V (1997) Estimation of evapotranspiration by hargreaves formula and remotely sensed data in semi-arid mediterranean areas. J agric Engng Res 68:189–199, doi:10.1006/jaer.1997.0166

Droogers P, Allen RG (2002) Estimating reference evapotranspiration under inaccurate data conditions. Irrigation and Drainage Systems 16:33–45, doi:10.1023/A:1015508322413

Er-Raki S, Chehbouni A, Khabba S, Simonneaux V, Jarlan L, Ouldbba A, Rodriguez JC, Allen RG (2010) Assessment of reference evapotranspiration methods in semi-arid regions: Can weather forecast data be used as alternate of ground meteorological parameters? Journal of Arid Environments 74(12):1587–1596, doi:10.1016/j.jaridenv.2010.07.002

Fox J, Monette G (1992) Generalized collinearity diagnostics. Journal of the American Statistical Association 87(417):178–183, doi:10.1080/01621459.1992.10475190

Gavilán P, Lorite IJ, Tornero S, Berengena J (2006) Regional calibration of hargreaves equation for estimating reference et in a semiarid environment. Agricultural Water Management 81(3):257–281, doi:10.1016/j.agwat.2005.05.001

Gomariz-Castillo F, Alonso-Sarría F (2013) An r script to model monthly climatic variables with glm to be used in hydrological modeling in river segura basin. In: The 9th International R User Conference, Albacete, Castilla-La Mancha (Spain)

Häntzschel J, Goldberg V, Bernhofer C (2005) Gis-based regionalisation of radiation, temperature and coupling measures in complex terrain for low mountain ranges. Meteorological Applications 12(1):33–42, doi:10.1017/S1350482705001489

Hargreaves GH (1989) Accuracy of estimated reference evapotranspiration. Journal of Irrigation Drainage Engineering 115(6):1000–1007, doi:10.1061/(ASCE)0733–9437(1989)115:6

Hargreaves GH (1994a) Defining and using reference evapotranspiration. Journal of Irrigation and Drainage Engineering 120(6):1132–1139, doi:10.1061/(ASCE)0733-9437(1994)120:6(1132)
Hargreaves GH (1994b) Simplified coefficients for estimating monthly solar radiation in north america and europe. Tech. rep., Utah State University, Logan
Hargreaves GH, Allen RG (2003) History and evaluation of hargreaves evapotranspiration equation. Journal of Irrigation and Drainage Engineering-Asce 129(1):53–63, doi:10.1061/(ASCE)0733-9437(2003)129:1(53)
Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture 1(2):96–99, doi:10.13031/2013.26773
Hengl T, Heuvelink GBM, Stein A (2004) A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120(1-2):75–93, doi:10.1016/j.geoderma.2003.08.018
Hofierka J, Súri M (2002) The solar radiation model for open source gis: implementation and applications. In: Proceedings of the Open source GIS - GRASS users conference, Trento, Italy, pp 1–19
Jensen DT, Hargreaves GH, Temesgen B, Allen RG (1997) Computation of eto under non-ideal conditions. Journal of Irrigation Drainage Engineering 123(5):394–400, doi:10.1061/(ASCE)0733-9437(1997)123:5(3)
Kidron GJ, Zohar M (2010) Spatial evaporation patterns within a small drainage basin in the negev desert. Journal of Hydrology 380(3-4):376–385, doi:10.1016/j.jhydrol.2009.11.012
Kuhn M, Johnson K (2013) Applied Predictive Modeling. Springer-Verlag, New York, doi:10.1007/978-1-4614-6849-3
Legates DR, McCabe GJ (1999) Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. Water Resources Research 35(1):233–241, doi:10.1029/1998WR900018
Liu BYH, Jordan RC (1960) The interrelationship and characteristic distribution of direct, diffuse, and total solar radiation. Solar Energy 4:1–19, doi:10.1016/0038-092X(60)90062-1
López-Urrea R, Martín de Santa Olalla F, Fabeiro C, Moratalla A (2006) Testing evapotranspiration equations using lysimeter observations in a semiarid climate. Agricultural Water Management 85(1-2):15–26, doi:10.1016/j.agwat.2006.03.014
Lu J, Sun G, McNulty SG, Amatya DM (2005) A comparison of six potential evapotranspiration methods for regional use in the southeastern united states. JAWRA Journal of the American Water Resources Association 41(3):621–633, doi:10.1111/j.1752-1688.2005.tb03759.x
Mardikis MG, Kalivas DP, Kollias VJ (2005) Comparison of interpolation methods for the prediction of reference evapotranspirationan application in greece. Water Resources Management 19(3):251–278, doi:10.1007/s11269-005-3179-2
Martínez-Cob A, Tejero-Juste M (2004) A wind-based qualitative calibration of the hargreaves et0 estimation equation in semiarid regions. Agricultural Water Management 64(3):251–264, doi:10.1016/S0378-3774(03)00199-9
McCullagh P, Nelder J (1989) Generalized Linear Models. Chapman & Hall/CRC, Boca Raton
Mendicino G, Senatore A (2013) Regionalization of the hargreaves coefficient for the assessment of distributed reference evapotranspiration in southern italy. Journal of Irrigation and Drainage Engineering 139(5):349–362, doi:10.1061/(ASCE)IR.1943-4774.0000547
Monteith JL (1965) Evaporation and the environment. In: XIXth Symposium Society for Experimental Biology, Cambridge University Press, Swansea, pp 205–234
Orang MN, Grismer ME, Ashktorab H (1995) New equations estimate evapotranspiration in delta. California Agriculture 49(3):19–21, doi:10.3733/ca.v049n03p19
Ray SS, Dadhwal VK (2001) Estimation of crop evapotranspiration of irrigation command area using remote sensing and gis. Agricultural Water Management 49:239–249, doi:10.1016/S0378-3774(00)00147-5
Raziei T, Pereira LS (2013) Estimation of eto with hargreavessamani and fao-pm temperature methods for a wide range of climates in iran. Agricultural Water Management 121(Supplement C):1 – 18, doi:10.1016/j.agwat.2012.12.019
Samani Z, Pessarakli M (1986) Estimating potential crop evapotranspiration with minimum data in arizona. Transactions of the ASAE 29(2):522 – 524, doi:10.13031/2013.30184
Samani ZA (2000) Estimating solar radiation and evapotranspiration using minimum climatological data. Journal of Irrigation and Drainage Engineering 126(4):265–267, doi:10.1061/(ASCE)0733-9437(2000)126:4(265)
Sendanayake S, Miguntanna NP (2014) Estimating incident solar radiation in tropical islands with short term weather data. European Scientific Journal 10(3):401–412

Shahidian S, Serralheiro R, Serrano J, Teixeira J, Haie N, Santos F (2012) Chapter 4 hargreaves and other reduced-set methods for calculating evapotranspiration. In: Irmak A (ed) Evapotranspiration - Remote Sensing and Modeling, vol 33, InTech, pp 59–80, doi:10.5772/18059

Shahidian S, Serralheiro RP, Serrano J, Teixeira JL (2013) Parametric calibration of the hargreaves-samani equation for use at new locations. Hydrological Processes 27(4):605–616, doi:10.1002/hyp.9277

Subburayan S, Murugappan A, Mohan S (2011) Modified hargreaves equation for estimation of et0 in a hot and humid location in tamilnadu state, india. International Journal of Engineering Science and Technology 3(1):592–600

Trajkovic S (2005) Temperature-based approaches for estimating reference evapotranspiration. Journal of Irrigation and Drainage Engineering 131(4):316–323, doi:10.1061/(ASCE)0733-9437(2005)131:4(316)

Vanderlinden K, Giraldez JV, Meirvenne MV (2004) Assessing reference evapotranspiration by the hargreaves method in southern spain. Journal of Irrigation Drainage Engineering 130(3):184–191, doi: 10.1061/(ASCE)0733-9437(2004)130:3(18

Vicente-Serrano SM, Lanjeri S, López-Moreno JI (2007) Comparison of different procedures to map reference evapotranspiration using geographical information systems and regression-based techniques. International Journal of Climatology 28(8):1103–1118, doi:10.1002/joc.1460

Wood S (2006) Generalized Additive Models. An introduction with R. Chapman & Hall/CRC, Boca Raton

Xu CY, Singh VP (2000) Evaluation and generalization of radiation-based methods for calculating evaporation. Hydrological Processes 14(2):339–349, doi:10.1002/(SICI)1099-1085(20000215)14:2¡339::AID-HYP928¿3.0.CO;2-O

Xu CY, Singh VP (2001) Evaluation and generalization of temperature-based methods for calculating evaporation. Hydrological Processes 15(2):305–319, doi:10.1002/hyp.119

Xu CY, Singh VP (2002) Cross comparison of empirical equations for calculating potential evapotranspiration with data from switzerland. Water Resources Management 16(3):197–219, doi: 10.1023/A:1020282515975