Journal of Hydrology 402 (2011) 144–154

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Technical Note

Guidelines on validation procedures for meteorological data from automatic weather stations

J. Estévez [a,*], P. Gavilán [b], J.V. Giráldez [c]

[a] Dpto. Ingeniería Rural, Área Proyectos de Ingeniería, Universidad de Córdoba, Ctra. Madrid, km 396, 14071 Córdoba, Spain
[b] IFAPA Centro Alameda del Obispo, Área de Producción Ecológica y Recursos Naturales, Av. Menéndez Pidal s/n, 14080 Córdoba, Spain
[c] Dpto. Agronomía and IAS-CSIC, Universidad de Córdoba, Ctra. Madrid, km 396, 14071 Córdoba, Spain

ARTICLE INFO

Article history:
Received 12 November 2010
Received in revised form 11 February 2011
Accepted 26 February 2011
Available online 9 March 2011
This manuscript was handled by K.P. Georgakakos, Editor-in-Chief, with the assistance of Rebecca Morss, Associate Editor

Keywords:
Quality control
Validation
Agrometeorology
Weather network

SUMMARY

Quality control is a major prerequisite for using meteorological information. High quality data sources are vital to scientists, engineers and decision makers alike. Validation of meteorological data ensures that the information needed has been properly generated and that it identifies incorrect values and detects problems that require immediate maintenance attention. In this work, several quality assurance procedures based on different criteria are proposed and applied to meteorological data from the Agroclimatic Information Network of Andalusia (Southern Spain) to assess their integrity and quality. The procedures include validations of record structure data, range/limits, time and internal consistency, persistence and spatial consistency tests. Quality assurance tests consist of procedures or rules against which data are tested, setting data flags to provide guidance to end users. The proposed system is capable of identifying several types of errors and is used as a tool that allows one to make decisions such as sensor replacement and to remove data prior to their application.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Meteorological observations and related environmental and geophysical measurements are necessary for a real-time preparation for weather analyses, forecasts and severe weather warnings, for the study of climate, for local weather-dependent operations, for hydrology and agricultural meteorology, and for research in meteorology and climatology (WMO, 2008). Some of the applications for meteorological data include: risk assessments, hydraulic structures design, crop water-use estimates, irrigation scheduling, input variables of climate change and hydrological models, active and passive renewable energy uses, etc. (Weiss and Robb, 1986; Meyer and Hubbard, 1992; Del Greco et al., 2005; Flores et al., 2005; Younes et al., 2005).

During the last two decades, the number of automated weather station networks has greatly increased throughout the world. This rapid development has been the consequence of the need to provide meteorological data in near-real time and the great evolution of automatic data acquisition systems (Miller and Barth, 2003). Irrigation Agroclimatic Information System (SIAR in Spanish language) is the most important automated meteorological network in Spain. It was installed in the period 1999–2000, covering the majority of irrigated areas of Spain. SIAR was established across all the Spanish regions for agronomic purposes (Pérez de los Cobos et al., 2003). The climate networks of many National Meteorological Services only provide temperature and precipitation data (Gázquez et al., 2003). Nevertheless, these data are not easily available with the frequency required (Meyer and Hubbard, 1992), in spite of World Meteorological Organization (WMO) resolutions requesting free and unrestricted exchange of climate information between National Organizations including the private sector (Cuadrat et al., 2002). Nowadays, the improvement and strengthening of meteorological observation networks to support data collection for operational applications and applied research in hydrological processes, agricultural meteorology or climate change models are required (WMO, 2006).

For purposes of scientific research and resource management, meteorological data from weather stations are recorded in large databases. All these efforts are enhanced by quality checking the data; however, no consistent methodologies are employed in many cases. Questionable results have been attributed to poor data quality as a consequence of non-existing or mixed quality control methods (Meek and Hatfield, 1994). There are three important reasons for applying quality control procedures to meteorological data: (i) to ensure that meteorological information is properly generated; (ii) to identify erroneous data involving inadequate decision making; and (iii) to detect and solve problems for a correct

* Corresponding author. Tel.: +34 957218343; fax: +34 957218550.
E-mail address: jestevez@uco.es (J. Estévez).

0022-1694/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2011.02.031

maintenance of the stations and periodic calibration of sensors (Doraiswamy et al., 2000).

Quality control of data is the best known component of quality management systems. It consists of the examination of data at weather stations and data centers with the aim of detecting errors (WMO, 1993). Different methods can be applied to ensure quality of meteorological data: periodic maintenance of stations and field sensor checks, validation of data using statistic procedures and, finally, sensor calibration. O'Brien and Keefer (1985) proposed a set of three computer-based rules, which were applied by Meek and Hatfield (1994) to validate meteorological data. These rules include computation of fixed or dynamic high/low bounds for each variable, use of fixed or dynamic rate of change limits for each variable and a continual no-observed-change in time limit. In the validation process, data of a doubtful quality must be detected and appropriately flagged. These quality control flags, which supplement but do not alter the data, are employed to describe which test the data failed. All archived meteorological data must be coupled with flags ("good", "suspect", "warning" or "failure") that indicate the level of confidence that network managers place upon their observations (Fiebrich and Crawford, 2001). Quality control flags are grouped into two main categories, "informative" and "severe", such as the California Irrigation Management Information System (CIMIS) (Snyder and Pruitt, 1992). In some cases, algorithms can be applied to correct erroneous data or to fill gaps, but both original and corrected data should always be archived in the databases (Reek et al., 1992). Finally, meteorological data considered as potentially erroneous should be verified and manually inspected by qualified personnel.

This paper presents some guidelines for applying quality control procedures to meteorological measurements according to several tests that use data from a single site (Meek and Hatfield, 1994; Shafer et al., 2000) and others that use data from multiple sites, comparing a station's data against neighboring stations (Reek et al., 1992; Hubbard, 2001). These validation methods have been applied to the Agroclimatic Information Network of Andalusia (RIAA), belonging to Spanish SIAR Network, using data from 1999 to 2006. The impact of error measurements in climate data on reference evapotranspiration (ET_0) has been recently studied for this network (Estévez et al., 2009). In general, valid meteorological data are required to make climate assessments and to make climate-related decisions. In semiarid regions, with a structural water deficit, the integrity and quality of these data are crucial to improving ET_0 estimates and precipitation, ensuring an adequate irrigation water management.

Table 1
Specifications of sensors installed on the RIAA automated weather stations.

Sensor	Variables	Accuracy	Range
Pt1000	Temperature	±0.2 °C (20 °C)	−39.2/60 °C
Humicap 180	Relative Humidity	±2% (0–90%) ±3% (90–100%)	0.8/100%
Young 05103	Wind Speed	±0.3 m s^{-1} (1–60 m s^{-1})	0/60 m s^{-1}
	Wind direction	±3°	0/360°
Skye SP1110	Solar radiation	±5%	350–1100 nm
ARG 100	Precipitation	0.2 mm/tip	

2. Material and methods

2.1. Data

Meteorological data used in the analysis were obtained from the RIAA. This network is currently composed of 99 automatic weather stations with the aim of providing ET_0 values and other meteorological data to improve irrigation water management (Gavilán et al., 2008). Daily data can be obtained from the Web at www.juntadeandalucia.es/agriculturaypesca/ifapa/ria. Automatic weather stations are controlled by a programmable CR10X datalogger (Campbell Scientific Inc., CSI, Logan, Utah) and are equipped with sensors for measuring temperature and relative

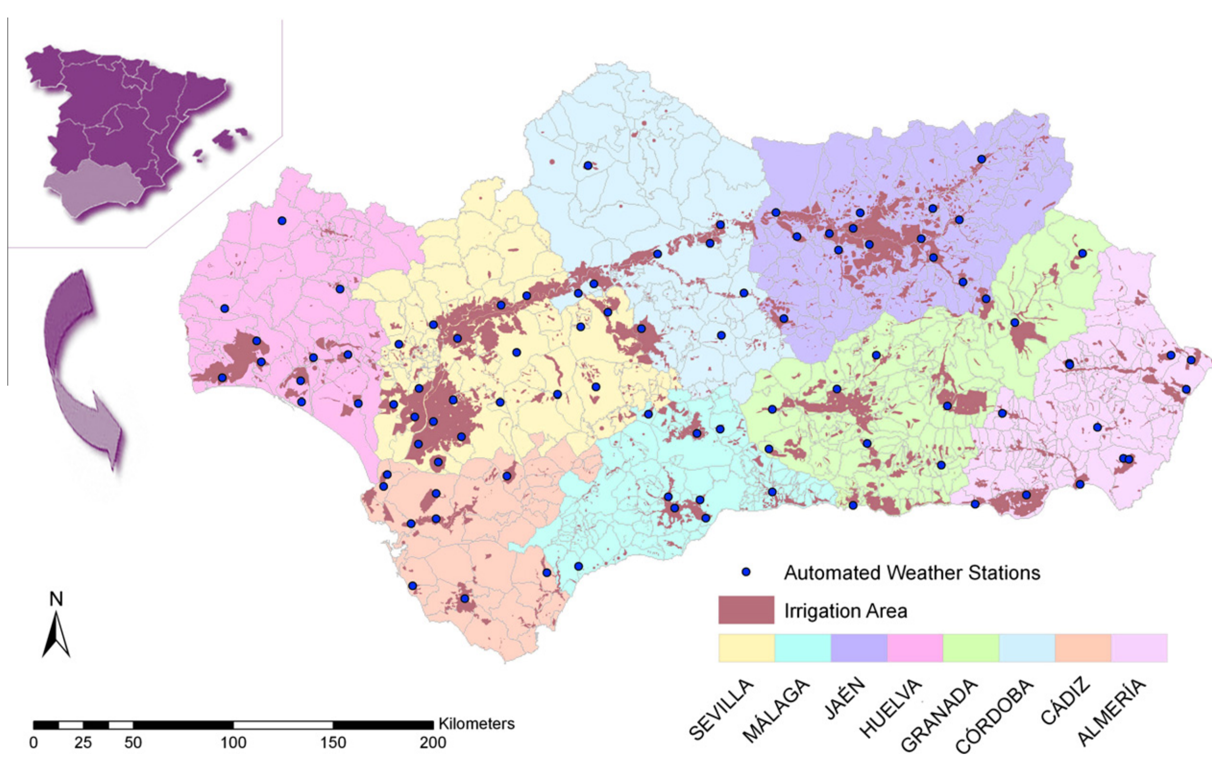

Fig. 1. Spatial distribution of automated weather stations from Agroclimatic Information Network of Andalusia.

Table 2
Location and elevation of automated weather stations validated in this study (Agroclimatic Information Network of Andalusia).

Stations	Latitude (°)	Longitude (°)	Elevation (m)
La Mojonera (AL)	36.788	−2.703	142
Almería (AL)	36.836	−2.401	22
Tabernas (AL)	37.092	−2.301	435
Fiñana (AL)	37.157	−2.837	971
V. Fátima-Cuevas (AL)	37.390	−1.769	185
Huércal-Overa (AL)	37.413	−1.883	317
Cuevas Almanz. (AL)	37.257	−1.799	20
Adra (AL)	36.747	−2.991	42
Níjar (AL)	36.951	−2.156	182
Tíjola (AL)	37.379	−2.458	796
Basurta-Jerez (CA)	36.758	−6.016	60
Jerez Frontera (CA)	36.643	−6.012	32
Villamartín (CA)	36.845	−5.621	171
Conil Frontera (CA)	36.337	−6.130	26
Vejer Frontera (CA)	36.286	−5.838	24
Jimena Frontera (CA)	36.413	−5.384	53
Puerto Sta. María (CA)	36.617	−6.151	20
Bélmez (CO)	38.255	−5.208	523
Adamuz (CO)	37.998	−4.443	90
Palma del Río (CO)	37.676	−5.246	134
Hornachuelos (CO)	37.720	−5.158	157
El Carpio (CO)	37.915	−4.502	165
Córdoba (CO)	37.861	−4.800	117
Santaella (CO)	37.523	−4.884	207
Baena (CO)	37.692	−4.304	334
Baza (GR)	37.565	−2.766	814
Puebla D.Fadriq. (GR)	37.877	−2.380	1110
Loja (GR)	37.170	−4.136	487
Pinos Puente (GR)	37.262	−3.771	594
Iznalloz (GR)	37.417	−3.550	935
Jerez Marques. (GR)	37.191	−3.148	1212
Cádiar (GR)	36.924	−3.182	950
Zafarraya (GR)	36.991	−4.152	905
Almuñécar (GR)	36.740	−3.676	49
Padul (GR)	37.020	−3.598	781
Tojalillo-Gibraleón (HU)	37.319	−7.026	52
Lepe (HU)	37.240	−7.243	74
Gibraleón (HU)	37.413	−7.058	169
Moguer (HU)	37.147	−6.791	87
Niebla (HU)	37.348	−6.733	52
Aroche (HU)	37.959	−6.943	299
Puebla Guzmán (HU)	37.553	−7.246	288
El Campillo (HU)	37.662	−6.598	406
Palma Condado (HU)	37.368	−6.540	192
Almonte	37.151	−6.470	18
Moguer-Cebollar (HU)	37.241	−6.800	63
Huesa (JA)	37.748	−3.060	793
Pozo Alcón (JA)	37.673	−2.928	893
S.José Propios (JA)	37.858	−3.229	509
Sabiote (JA)	38.080	−3.234	822
Torreblascopedro (JA)	37.989	−3.688	291
Alcaudete (JA)	37.578	−4.077	645
Mancha Real (JA)	37.917	−3.595	436
Úbeda (JA)	37.943	−3.299	358
Linares (JA)	38.060	−3.648	443
Marmolejo (JA)	38.057	−4.129	208
Chiclana Segura (JA)	38.303	−2.954	510
Higuera Arjona (JA)	37.950	−4.006	267
Santo Tomé (JA)	38.030	−3.081	571
Jaén (JA)	37.891	−3.770	299
Málaga (MA)	36.757	−4.536	68
Vélez-Málaga (MA)	36.797	−4.130	49
Antequera (MA)	37.056	−4.558	457
Estepona (MA)	36.446	−5.208	199
Archidona (MA)	37.078	−4.428	516
Sierra Yeguas (MA)	37.139	−4.834	464
Churriana (MA)	36.675	−4.501	32
Pizarra (MA)	36.767	−4.713	84
Cártama (MA)	36.718	−4.676	95
Palacios-Villafran.(SE)	37.180	−5.937	21
Cabezas S. Juan (SE)	37.016	−5.883	25
Lebrija I (SE)	36.977	−6.125	25
Lebrija 2 (SE)	36.900	−6.009	40
Aznalcázar (SE)	37.152	−6.271	4
Puebla del Río (SE)	37.227	−6.132	25
Puebla del Río II (SE)	37.081	−6.045	41
Écija (SE)	37.594	−5.075	125
La Luisiana (SE)	37.526	−5.226	188
Osuna (SE)	37.256	−5.133	214
La Rinconada (SE)	37.458	−5.923	37
Sanlúcar la Mayor (SE)	37.422	−6.253	88
Villan.Río-Minas (SE)	37.614	−5.682	38
Lora del Río (SE)	37.660	−5.538	68
Los Molares (SE)	37.177	−5.671	90
Guillena (SE)	37.515	−6.062	191
Puebla Cazalla (SE)	37.219	−5.349	229
Carmona-Tomejil (SE)	37.401	−5.586	79

AL = Almería; CA = Cádiz; CO = Córdoba; HU = Huelva; GR = Granada; JA = Jaén; MA = Málaga; SE = Sevilla.

humidity of the air (HMP45C probe, Vaisala, Helsinky, Finland), solar radiation (pyranometer SP1110, Skye, Llandrindod Wells, United Kingdom), wind speed and direction (wind monitor RM Young 05103, Traverse City, Mich.), and rainfall (tipping bucket rain gauge ARG 100). The measurement heights and more details have been outlined in Gavilán et al. (2006) and are not repeated here, although specifications of the sensors are reported in Table 1. Daily and semihourly values are stored for each meteorological variable. Meteorological data from 1999 to 2006 were used in this study. Variables validated were: daily and semihourly solar radiation (Rs and Rs_{sh}, respectively), daily mean, maximum and minimum air temperatures (T_m, T_x and T_n, respectively), semihourly air temperature (T_{sh}), daily mean, maximum and minimum relative humidity of the air (RH_m, RH_x and RH_n, respectively), semihourly relative humidity (RH_{sh}), daily mean and maximum wind speed (U_m and U_x, respectively), semihourly wind speed (U_{sh}), daily amount of precipitation (P) and semihourly amount of precipitation amount (P_{sh}). The geographical distribution of the stations is shown in Fig. 1. The stations are evenly distributed throughout the irrigated areas. Site elevations range from 4 to 1212 m above mean sea level, longitude, from 1769 to 7246W and latitude, from 36,286 to 38,255N (Table 2).

2.2. Validation procedures

Validation methods applied to identify erroneous data from sensor measurements are described in this section. Initially, they are necessary to verify that all possible data have been collected and that the record structure is correct, complete and without any gaps in the data files. Meteorological data not fulfilling the above requirements are flagged as erroneous.

However, the automated analysis is not sometimes capable of recognizing subtle problems in meteorological observations, such as instrument drift. Furthermore, some localized phenomena (e.g. thunderstorms, heatbursts) may be erroneously flagged by these automatic quality control algorithms (Shafer et al., 2000). Visual analysis is the last step in the validation process and it must be carried out by well-trained technicians or meteorologists, helping to identify problems in the climate dataset. Usual methods of visual inspection of meteorological observations are real-data monitoring, mapping extreme values or cumulative totals during the period of interest. Also, time-series graphs depicting data from the site and parameter in question can be created or a long-term analysis such as ‘‘double mass analysis’’ technique can be conducted (Allen et al., 1998) Obviously, depending on the availability of long-term series and qualified personal, and on the infrastructure of the network itself, some methods could be appropriately applied and some not.

Table 3
Quality control procedures applied to RIAA network for daily and semihourly data.

Validation procedures	Global radiation daily (MJ m^2 d^{-1}) semihourly (W m^{-2})	Relative humidity (%)	Air temperature (°C)	Wind speed (m s^{-1})	Precipitation (mm)
Range test	$-1 < Rs_{sh} < 1500$ (Shafer et al., 2000); $0.03R_a \leqslant Rs, Rs_{sh} \leqslant R_a$; $Rs, Rs_{sh} < 1.1R_{so}$ (Allen, 1996; Geiger et al., 2002; Moradi, 2007)	$0.8 < RH < 103$ (Table 1) (Shafer et al., 2000)	$-30 < T < 50$ (Shafer et al., 2000); $T_{LOW} < T < T_{HIGH}$ (AEMET, 2008)	$0 < U, U_{sh} < 60.3$ (Table 1); $0 < U_x < 100$ (Shafer et al., 2000)	$0 \leqslant P_{sh} \leqslant 120$ (Zahumensky, 2004); $0 \leqslant P < 508$ (Shafer et al., 2000); $0 \leqslant P, P_{sh} \leqslant P_{MAX}$ (AEMET, 2008)
Step test	$0 \leqslant \|Rs_{sh} - Rs_{sh\text{-}2}\| \leqslant 555$ (Meek and Hatfield, 1994)	$\|RH_{sh} - RH_{sh\text{-}1}\| < 45$ (Zahumensky, 2004)	$\|T_{sh} - T_{sh\text{-}2}\| < 4$; $\|T_{sh} - T_{sh\text{-}4}\| < 7$; $\|T_{sh} - T_{sh\text{-}6}\| < 9$; | $T_{sh} - T_{sh\text{-}12}\| < 15$; | $T_{sh} - T_{sh\text{-}24}\| < 25$; (WMO, 1993)	$\|U(d) - U(d-1)\| < 10$; $\|U_{sh}\text{-}U_{sh\text{-}2}\| < 10$; (Meek and Hatfield, 1994)	
Internal consistency test		$RH_x > RH_m > RH_n$ (Reek et al., 1992; Feng et al., 2004) $RH_x > \max(RH_{sh})$; $RH_n < \min(RH_{sh})$ (Vejen et al., 2002)	$T_x > T_m > T_n$; $T_x(d) > T_n(d-1)$; $T_n(d) \leqslant T_x(d-1)$ (Reek et al., 1992; Feng et al., 2004) $T_x > \max(T_{sh})$ $T_n < \min(T_{sh})$ (Vejen et al., 2002)	Speed = 0 and direction = 0; speed ≠ 0 and direction ≠ 0 (DeGaetano, 1997; Zahumensaky, 2004) $U_x(d) > U(d)$; $U_x(d) > \max(U_{sh})$ (Vejen et al., 2002)	Diurnal precipitation events are true if: $K_T^* < 0.5$ *and* $RH^* > 80\%$ (Estévez, 2008) $P_{sh(0-3h)} \leqslant P_{sh(0-6h)}$; $P_{sh(0-12h)} \leqslant P_{sh(0-24h)}$ (Vejen et al., 2002)
Persistence test	$Rs(d) \neq Rs(d-1) \neq Rs(d-2)$; $Rs_{sh} \neq Rs_{sh\text{-}2} \neq Rs_{sh\text{-}4} \neq Rs_{sh\text{-}6}$ (Meek and Hatfield, 1994)	$RH(d) \neq RH(d-1) \neq RH(d-2)$ (Meek and Hatfiled, 1994) $\sigma RH_{sh} > 1$ (Zahumensky, 2004)	$T(d) \neq T(d-1) \neq T(d-2)$; $T_{sh} \neq T_{sh\text{-}2} \neq T_{sh\text{-}4} \neq T_{sh\text{-}6}$ (Meek and Hatfield, 1994)	$U(d) \neq U(d-1) \neq U(d-2)$; $U_x(d) \neq U_x(d-1) \neq U_x(d-2)$; $U_{sh} \neq U_{sh\text{-}2} \neq U_{sh\text{-}4} \neq U_{sh\text{-}6}$ (Meek and Hatfield, 1994)	
Spatial consistency test			$T'\text{-}fs' \leqslant T^* \leqslant T' + fs'$ where $T^* = T_x, T_n, T_m$; $T' = T'_x, T'_n, T'_m$ (reference series); s' = weighting root-mean-square error; f = factor		

Rs_{sh}: semihourly global radiation (W m^{-2}); Rs: daily global radiation (MJ m^2 d^{-1}); R_{so} : daily or semihourly global radiation under clear sky conditions (MJ m^2 d^{-1} or W m^{-2}); R_a: daily or semihourly extraterrestrial radiation (MJ m^2 d^{-1} or W m^{-2}); RH: daily maximum, minimum or mean relative humidity (%); RH_{sh}: semihourly relative humidity (%); T: daily or semihourly temperature (°C); T_{LOW}: minimum temperature value measured in long-term observatory closer to each station (°C); T_{HIGH}: maximum temperature value measured in long-term observatory closer to each station (°C); T_m: daily mean temperature (°C); T_x: daily maximum temperature (°C); T_n: daily minimum temperature (°C); T_{sh}: semihourly temperature (°C); U: daily mean wind speed (m s^{-1}); U_x: daily maximum wind speed (m s^{-1}); U_{sh}: semihourly wind speed (m s^{-1}); P: daily precipitation (mm); P_{sh}: semihourly precipitation (mm); $P_{sh\ (0\text{-}Xh)}$: amount of semihourly precipitation from 0 to Xh (mm); P_{MAX}: daily or semihourly extreme value measured in long-term observatory closer to each station; K_T*: mean "clearness" index during diurnal precipitation event; RH*: mean relative humidity during diurnal precipitation event; d: day d; d-1: day before day d; d-2: day before day d-1; sh: semihour sh; sh-1: semihour before 30 minutes; sh-2: semihour before one hour; sh-4: semihour before two hours; sh-6: semihour before three hours, etc.
σ_{RHsh}: standard deviation of semihourly relative humidity data.

Procedures recommended for assessing quality of net radiation, soil heat flux and vapor pressure (Allen et al., 1998), sunshine duration (Allen et al., 1994), barometric pressure (Meek and Hatfield, 1994) or pan evaporation (Feng et al., 2004) are not described in the next sections. The quality control methods applied in this work are: range/limit (fixed or dynamic) test, step test, internal consistency test, persistence test and spatial consistency test. Meteorological data which do not pass the fixed range test should be flagged as erroneous. Data rejected by the other tests should be flagged as suspect, and they should be validated by manual inspection.

2.2.1. Range test

The range test is based upon a combination of performance specifications for each sensor and physical/climate extremes for each location and variable. Limits of different meteorological variables may depend on climate conditions at the station's site and on a season. Any observation that occurs outside the maximum or minimum allowable value is properly flagged. Measured data should be inside this upper and lower threshold in order to be considered valid (Reek et al., 1992; Meek and Hatfield, 1994; Shafer et al., 2000; Feng et al., 2004). There are two types of ranges: fixed (physics and instrumental) and dynamic ranges.

The fixed range test compares the value of a meteorological variable with established extreme values. Any observation that occurs outside the acceptable range is flagged as erroneous and is not validated by the next tests. The fixed limits proposed for different variables are based on Table 1 and on ranges proposed by some authors (Table 3).

Dynamic bounds for each meteorological variable are based on extreme values measured for each location or on the theoretical possible extreme for each site and time period, as Allen (1996) proposed for solar radiation. Variable high/low extreme values recorded for temperature and precipitation used in this test were reported by the Spanish Meteorological Agency (AEMET, 2008) and they were obtained from long-term meteorological observatories closer to each station. Data rejected by the dynamic range test are flagged as suspect and they should be verified by manual inspection. If a suspected measured datum (potential outlier) is correctly verified, it is stored as a new extreme value for the corresponding location.

Daily or semihourly solar radiation data (Rs and Rs_{sh}, respectively) can be validated by applying the dynamic test described by Moradi (2007), where measured data are compared with extraterrestrial solar radiation values (R_a) for each location, according to:

$$0.03R_a \leqslant Rs(\text{or } Rs_{sh}) \leqslant R_a \quad (1)$$

The low dynamic bound is the same as that proposed by Geiger et al. (2002) for the atmospheric transmittance coefficient, defined as the ratio of measured solar radiation to extraterrestrial radiation. A more restrictive test, and a very useful one as a quality control procedure, is based on the comparison of measured solar radiation with solar radiation expected under clear sky conditions (R_{so}), This procedure was suggested by Allen (1996), plotting both curves at daily, hourly or shorter periods. Geiger et al. (2002) proposed the same high bound incremented in 10%, according to:

$$Rs(\text{or } Rs_{sh}) < 1.10R_{so} \quad (2)$$

R_{so} can be computed as

$$R_{so} = K_T R_a \quad (3)$$

where K_T is a "clearness" index, which may range from 0.7 to 0.8, depending on atmospheric clarity (dust, pollution, humidity, etc.) and elevation and sun angle. Both R_a and K_T can be estimated with the method proposed by Allen (1996). A simple prediction for K_T, where only site elevation is considered, was developed by Allen et al. (1994) by regression of the integrated Beer's radiation extinction function (Monteith and Unsworth, 1990) with an elevation between sea level and 3000 m

$$K_T = 0.75 + 2 \cdot 10^{-5} z \quad (4)$$

where z is the station elevation above mean sea level (m). Improved estimates of R_{so} can be used and have been outlined in Allen et al. (1998) and are not repeated here.

Fig. 2 shows examples of using Eqs. (2)–(4) for validating Rs data using estimated R_{so} envelopes in two different locations for daily time steps. Fig. 2a and b show daily Rs data measured at "Las Cabezas de San Juan" (Sevilla) and Córdoba, respectively, both from 2001 to 2006. Rs measurements on clear-sky days in Córdoba followed the calculated clear-sky envelopes estimated using equations 3 and 4 during all seasons (Fig. 2b). Only two solar radiation data were below the dynamic bound, maybe as a consequence of pyranometer conditions not being clean. Comparison of R_{so} curve with measured solar radiation from "Las Cabezas de San Juan" (Sevilla) station indicates that the pyranometer was measuring about 20% high on clear-sky days from spring 2002 to summer 2006. Outliers detected increased due to an incorrect calibration of the pyranometer. Generated alarms allowed the replacement of the sensor with a new one. Measurement data decreased and remained very close to the computed clear-sky envelope after replacement of the pyranometer (Fig. 2a).

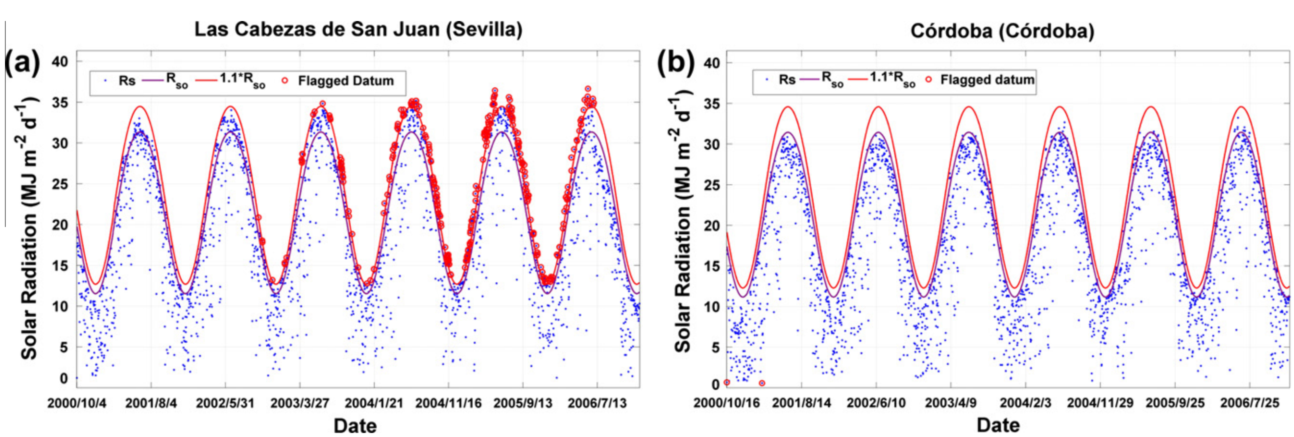

Fig. 2. Application of dynamic range test to daily solar radiation data for "Las Cabezas de San Juan (Sevilla)" (a) and for "Córdoba (Córdoba)" stations (b). Rs: Measured daily solar radiation; R_{so}: solar radiation expected under clear sky conditions.

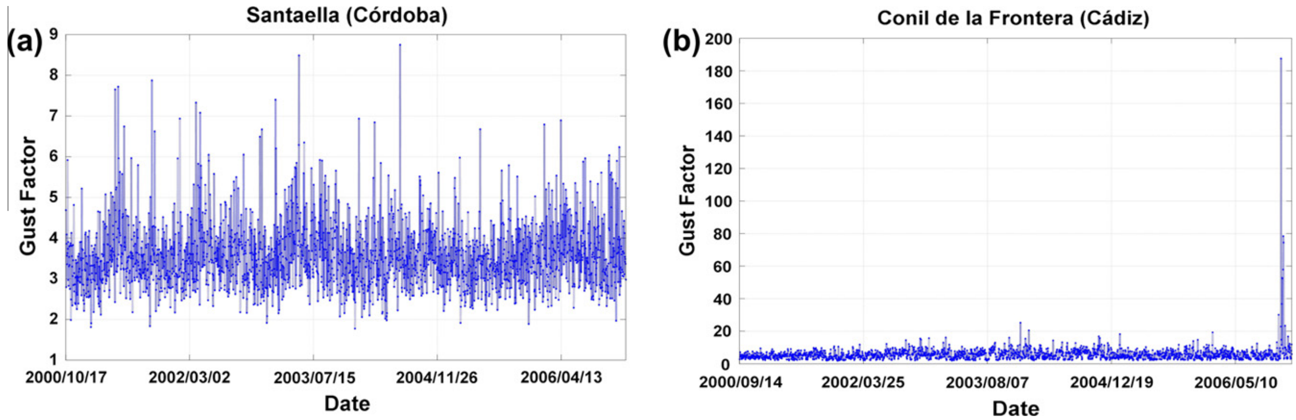

Fig. 3. Daily evolution of gust factor (ratio of daily maximum wind speed to daily mean wind speed) for "Santaella (Córdoba)" (a) and "Conil de la Frontera (Cádiz)" (b) stations.

2.2.2. Step test

Quality control procedures based on time consistency (step test) compare the difference between successive measurements. If the difference exceeds an allowed value, a different one for each parameter, both observations are flagged as suspect (Shafer et al., 2000). Therefore, this test checks the excessive rate of change of two consecutive values.

For validating semihourly global radiation data, a modified step test proposed by Meek and Hatfield (1994) can be applied as

$$0 \leqslant |Rs_{\mathrm{sh}} - Rs_{\mathrm{sh}-2}| \leqslant 555Wm^{-2} \quad (5)$$

where Rs_{sh} is the solar radiation measured at hour h, and $Rs_{\mathrm{sh\text{-}2}}$ the solar radiation measured one hour before (Table 3). This test was used for validating hourly temperature and relative humidity data in some USA networks (Graybeal et al., 2004a) and it has proven to be useful for detecting erroneous readings due to loose wires or datalogger problems (Shafer et al., 2000). Recent works used this test with data recorded at a higher sample rate (Schroeder et al., 2005). An adaptation of this test was proposed by Lanzante (1996) for temperature, requiring ten years of data. DeGaetano (1997) proposed some specific algorithms to validate wind speed and direction data based on time consistency, requiring instantaneous values or 1-min, 2-min average. Verifications of temporal consistency of meteorological data are based on information redundancy of successive observations at a given weather station. In this sense, the World Meteorological Organization (WMO, 1993) proposed some recommendations of tolerance for temperature (Table 3), dew point temperature and atmospheric pressure. Zahumensky (2004) recommended a possible limit of a maximum variability of 10-min relative humidity values. A modified step test for RIAA network is proposed for semihourly RH data (Table 3).

The procedures based on time consistency that were applied, according to the available meteorological data from RIAA, are summarized in Table 3.

2.2.3. Internal consistency test

This test is based on the verification of physics or climatologic consistency of each observed parameter or on the relation between two measured variables (Grüter et al., 2001). Several meteorological variables measured at the same location and time should be meteorologically consistent with each other. Otherwise, both observations would be flagged as suspect. For example, an average value will always be lower than the maximum instantaneous value, or the amount of precipitation occurring in 6 h will always be lower or equal to the amount of precipitation in 24 h (Vejen et al., 2002). Therefore, similar algorithms to Equation 6 for several variables are summarized in Table 3. Semihourly data in RIAA network is the average of three instantaneous values and extreme (maximum or minimum) daily values are the lowest or highest of the instantaneous data of a whole day. In Equation 6, U_x is the daily maximum instantaneous wind speed value and $\max(U_{\mathrm{sh}})$ is the semihourly maximum value from 48 wind speed records from a specific day.

$$U_x > \max(U_{\mathrm{sh}}) \quad (6)$$

Also, it is usual to use checks such as $T_x(\mathrm{d}) > T_m(\mathrm{d}) > T_n(\mathrm{d})$ or $T_x(\mathrm{d}) > T_n(\mathrm{d}-1)$ for a given day d (Reek et al., 1992; Feng et al., 2004). Another check is to verify that wet bulb temperature does not exceed dry bulb temperature (Graybeal et al., 2004a), and that it does not exceed dew point temperature (Graybeal et al., 2004b). For manual stations, a limit for the daily range temperature (T_x–T_n) can be used for temperature validation data (Reek et al., 1992; Feng et al., 2004) or its relation to the dewpoint depression to check both parameters (Graybeal et al., 2004b). To validate wind speed and direction data, the following conditions will be true: wind speed = 0 and wind direction = 0; wind speed ≠ 0 and wind direction ≠ 0 (DeGaetano, 1997; Zahumensky, 2004). Relationships between peak-gust, fastest-mile and 5-min wind speeds can be used to validate daily wind speed data (Graybeal, 2005). The problem is that these kinds of records are not measured in the majority of automated weather stations. Maximum wind speed data, if available, can assist in the assessment of daily wind speed data. Gust factor (ratio of maximum wind speed to mean daily wind speed) is a useful index for checking anemometers measurements (ASCE-EWRI, 2005). This factor often increases as contamination increases the friction in the bearings. As an example, this factor is represented in Fig. 3 for two different locations for several years. At the "Santaella" (Córdoba) station (Fig. 3a) the gust factor values showed good behavior, but the data from "Conil de la Frontera" (Cádiz) station showed some anomalies at the end of 2006 (Fig. 3b). This short period of excessively large values indicate that the anemometer may be malfunctioning.

For precipitation, Estévez (2008) reported -using SIAR network data in Southern Spain- that in semiarid regions like Andalusia the majority of diurnal rainfall events occurs with high relative humidity values (higher than 75–80%) and low "clearness" index values (K_T). This index was introduced by Allen (1996) and Estévez (2008) proposed that a diurnal event would be true if, on average, $K_T < 0.5$ and $RH > 80\%$ during the event time period. This test is very useful in agrometeorological networks, where the weather stations are located covering irrigation areas, because rain gauges can receive water from sprinkler irrigation systems or can register pulses

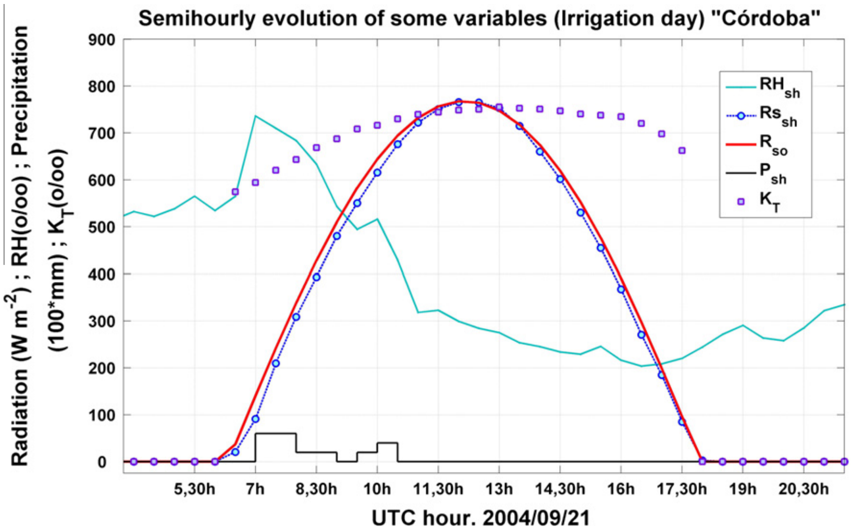

Fig. 4. Semihourly evolution of relative humidity (RH_{sh}), global radiation (Rs_{sh}), expected global radiation under clear sky conditions (R_{so}), precipitation (P_{sh}) and "clearness" index (K_T) at "Córdoba" station during an irrigation day.

because of some farm work machinery being near the station. Fig. 4 shows the evolution of several meteorological parameters on a typical irrigation day, when precipitation records occurred between 7 h and 10.30 h, under clear-sky conditions ($K_T > 0.5$) and when relative humidity values were lower than 60%. This false rainfall event was due to an irrigation event.

2.2.4. Persistence test or time series consistency test

This test checks the variability of the measurements. When a sensor fails it will often report a constant value, thus the standard deviation will become low. When the sensor is out for an entire reporting period, its standard deviation will be zero. In Shafer et al. (2000), the persistence test checks an entire day's data from one station, one parameter at a time. The mean and standard deviation for that parameter are computed. If the standard deviation is below an acceptable minimum, the corresponding data values are flagged as suspect. Because the persistence test uses aggregate statistics, it cannot discern which observations within the time period are responsible for this alert. So, all data for that period are flagged as suspect or warning. A "delta test" is also used to check the largest difference between any pair of observations within a selected time range, usually 24 h (Shafer et al., 2000; Fiebrich and Crawford, 2001; Schroeder et al., 2005). If the difference is less than a minimum acceptable change, all data for the corresponding day are flagged. A test called "flatliner check" has been used for temperature, which analyzes data values of five consecutive days (Reek et al., 1992; Kunkel et al., 1998). For the rest of meteorological measurements, excluding precipitation, Meek and Hatfield (1994) proposed a continual no-observed-change in time limit and it was applied in this work for different variables (Table 3). A specific test for relative humidity is proposed by Zahumensky (2004), checking on a minimum required variability of instantaneous values during a certain time period. A possible limit for 30 min period is proposed as

$$\sigma_{RH_{sh}} > 0.01 \quad (7)$$

where σ_{RHsh} is the daily standard deviation of semihourly data of relative humidity.

2.2.5. Spatial consistency test

These tests are designed to detect gross errors in individual observations. They are based on the comparison of data from one station with corresponding values of surrounding stations (Wade, 1987; Gandin, 1988; Eischeid et al., 1995; Hubbard, 2001). More subtle errors are identified through manual inspection. In the test proposed by Shafer et al. (2000), observations are weighted according to their distance from the station being evaluated. The difference between the observed and estimated value is compared to the standard deviation, flagging as being suspect any observation whose difference exceeds twice the standard deviation. To apply this test properly, at least six weather stations are necessary. This procedure is based on the Barnes algorithm (Barnes, 1964), also used for meteorological variables measured as a higher sample rate (Schroeder et al., 2005). The majority of algorithms employed to estimate values at a station from data of neighboring stations are based on the inverse distance weighting technique. This technique weights the values at surrounding stations according to the inverse of the distance separating the locations (Isaaks and Srivastava, 1989). The problem is that the nearest station does not always register the closest value. It depends on terrain topography and the particular variability of each meteorological parameter (Wade, 1987; Guttman et al., 1988). Recently, Hubbard et al. (2005) developed a new test called the spatial regression test. It is based on weighting data of each station according to the root-mean-square error, improving the results of those based on distance for variables like maximum and minimum temperature and precipitation (Hubbard and You, 2005). This test checks that the variable falls inside the confidence interval formed from estimates based on N "best fit" neighboring stations during a time period of length n, which was taken as 24 days for this study, and N was set to 8. The algorithm applied to temperature data is shown in Table 3. Feng et al. (2004) applied a similar spatial consistency test for several climate variables.

3. Quality control

The results of running the procedures on the dataset for each climate variable are reported below and they are summarized in Table 4. A previous analysis of meteorological database for detecting gaps was necessary before applying quality control tests. In this sense, no trend is observed, 2005 and 2006 being the years with fewer gaps.

3.1. Solar radiation

"Puebla de Cazalla" (Sevilla) was the only one weather station where the fixed range test detected a daily value outside the range. Therefore, only this solar radiation datum was rejected by this test,

Table 4
Percentage of flagged meteorological data by quality control tests (Max = maximum; Avg = average; St. Dev.=standard deviation of total network data). The percentages on each cell correspond to Max/Avg (St. Dev).

Variables		Tests					
		Range (fixed limits)	Range (dynamic limits)	Step	Internal consistency	Persistence test	Spatial consistency
Solar	Rs	0.049/0.001(0.005)	11.80/0.34(1.32)			0.82/0.01(0.09)	
Radiation	Rs_{sh}	0.016/0.000(0.002)	7.04/0.80(1.10)	0.24/0.10(0.04)		0/0(0)	
	RH_x	1.63/0.07(0.20)			0.649/0.044(0.080)	16.019/0.252(1.717)	
	RH_n	0/0(0)			0.240/0.007(0.030)	0.147/0.006(0.029)	
Relative humidity	RH_m	0/0(0)			0.239/0.004(0.029)	0.234/0.008(0.034)	
	RH_{sh}	0/0(0)		0.021/0.002(0.004)		1.535/0.094(0.194)	
Air	T_x	0/0(0)	0/0(0)		0.091/0.012(0.023)		
Temperature	T_n	0/0(0)	0/0(0)		0.175/0.005(0.022)	0/0(0)	24.41/0.83(2.75)
	T_m	0/0(0)	0/0(0)		0/0(0)	0/0(0)	10.62/0.60(1.34)
	T_{sh}	0/0(0)	0/0(0)	0.77/0.11(0.14)		0/0(0)	28.3/0.90(3.3)
Precipitation	P	0.105/0.001(0.011)	0/0(0)			0/0(0)	
	P_{sh}	0.003/0.000(0.000)	0/0(0)		0.372/0.015(0.046)		
Wind	U_x	0/0(0)			2.70/0.32(0.42)	3.78/0.05(0.43)	
Speed	U	0/0(0)		0.09/0.00(0.00)	0.65/0.01(0.07)	2.31/0.04(0.33)	
	U_{sh}	0/0(0)		0/0(0)	0.93/0.06(0.15)	3.07/0.12(0.45)	

representing approximately 0.05% of the fraction errors caught by this procedure in that station, and 0.001% on average. The rest of the daily solar radiation data (89 stations) were not flagged. Similar results were obtained after applying this test to semihourly data. Only 14 values were flagged for "Tíjola" (Almeria) station, corresponding to 0.016% of solar radiation data at this station. The average percentage of flagged daily solar radiation data after application of dynamic range test was 0.34%, with a maximum of 11.80% for "Las Cabezas de San Juan" (Sevilla) station. 30 of the 90 stations do not generate any alert using this test. The majority of values out of range were detected in 2005, due to the non calibration of pyranometers as a consequence of silicon photocell aging. The average percent for semihourly solar radiation data flagged by dynamic range test was 0.80%, with a maximum of 7.04% for the same station as in daily validation. All stations had flagged semihourly values, with a standard deviation of 1.10%. The majority of flagged data, as in the daily validation process, was detected in 2005. These results indicate that the percentage of semihourly flagged data detected by this test was more than twice those found in the daily dataset. The increase in measured semihourly solar radiation data to above the R_{so} envelope is caused by the reflection of direct beam radiation from nearby clouds (Allen, 1996; ASCE-EWRI, 2005). This phenomenon and the model mismatch to estimate this envelope (especially at sunrise and sunset) caused that semihourly flagged data fraction to be higher than the daily fraction. This dynamic test has been frequently used to verify the integrity of solar radiation data before use them to estimate other variables. Jensen et al. (1997) and Allen (2000) remarked on the usefulness of this test to check pyranometers measurements before estimating ET_0. Other authors reported some limitations caused by uncertainty atmosphere turbidity (Irmak et al., 2003) or the degree of cloudiness during summer months (Yoder et al., 2005), especially in humid climates. Nevertheless, these authors point out draft its great utility as a previous step to estimate net radiation. In quantitative terms, there are no similar studies that apply this kind of test to a wide weather station network. Geiger et al. (2002) applied a similar procedure for daily data to only one station, resulting in 0.11% of flagged data, a similar percentage to many stations analyzed in this work. Meek and Hatfield (1994) applied a test based on a different envelope solar radiation for clear-sky conditions (Howell et al., 1984), obtaining 7% of flagged data for hourly data and no daily flagged data for two stations.

A step test applied to semihourly data was able to detect 0.1% of anomalous values, with a maximum of 0.24% in "Jerez del Marquesado" (Granada) station. On average, 0.01% of daily solar radiation data were flagged by a persistence test. Only four stations had flagged data, "Almonte" (Huelva) being the weather station with the highest percentage of flagged data (0.82%). Finally, no semihourly data were flagged by the persistence test (Table 4). These percentages are similar to those obtained by Meek and Hatfield (1994) after analyzing two stations. In their work, the persistence test for hourly data did not detect flagged data either, and fewer records were flagged by step test, like the present study. Similarly, all daily data correctly passed the persistence test, like 86 of the 90 stations analyzed in this paper.

3.2. Relative humidity

The fixed range test flagged only RH_x data, no errors for RH_n, RH_m and RH_{sh} data were detected (Table 4). These flagged data fell between 100% and 103%, indicating an incoherent meteorological record, but a correct functioning of the HUMICAP probe sensor (Table 1). These values were corrected to the maximum physic value of relative humidity (100%). As expected in a semiarid region, only a few RH data were flagged by this test. The average percentage of RH_x data flagged was 0.07%, "Lepe" (Huelva) being the location with the maximum percentage (1.63%). No outliers were found in 69 weather stations.

The greatest percentage of flagged data by internal consistency tests was detected for RH_x (0.044% on average), "Padul" (Granada) being the station with the highest percentage (0.649%). There were 43 weather stations without RH_x flagged data after applying internal consistency tests. The average percentage of RH_n flagged data was 0.007%, with a maximum (0.240%) found in "Guillena" (Sevilla) station, and there were 83 locations without flagged data. The percentages of RH_m flagged data were lower, with only 2 stations with alerts. These results are in accordance with Feng et al. (2004) for RH_m, in which the total of flagged data for the network studied by this kind of procedure was 0.0039%. Throughout the time period analyzed, no trend was observed for percentages of flagged data. The highest percentage of data flagged by persistence test was obtained for RH_x, with an average equal to 0.252% corresponding to all stations and the highest value detected in "Isla Mayor" (Sevilla) station (16.019%). All RH_x flagged data ranged from 90% and 100%. In this range, the sensor response is low due to hysteresis phenomena. In this case, there is a slight increasing trend of flagged data percentage through the years. Results obtained for RH_n and RH_m were similar and only data from four weather stations were flagged. In addition, this persistence test flagged 0.09% RH_{sh} data, with the highest percentage detected in "Puebla de Guzmán" (Huelva) station (1.53%). No errors were detected in 32 stations of the network. These results, on average, are similar to those reported by Graybeal et al. (2004a) for hourly data (0.1%). Finally, results obtained after applying the step test were the lowest percentages, 0.002% on average and 0.021% of flagged data as the maximum in "Higuera Arjona" (Jaén) station. No records were flagged by this test in 55 weather stations, the majority of errors being associated with rain events, and, consequently, to high values of relative humidity, of over 90%.

3.3. Wind speed

No data were flagged applying the range test defined by fixed limits. Nevertheless, internal consistency routines together with step and persistence tests were able to detect anomalous data (Table 4). For maximum wind speed (U_x) values, the average percentage of flagged data by an internal consistency test was 0.32%, with the highest percentage (2.79%) detected in "Almuñécar" (Granada) location and there were 11 stations with no errors found. The average percentage of flagged data for daily wind speed was 0.01%, with the highest percentage (0.65%) detected in "Padul" (Granada) station. Results obtained for semihourly data were similar to percentages of detected errors described by DeGaetano (1997) for hourly data. In general, the percentages of daily flagged data for wind speed are higher than those reported by Feng et al. (2004) and Schroeder et al. (2005). On average, all wind speed data passed the step test satisfactorily, only some daily data from "Jimena de la Frontera" (Cádiz) station were flagged (0.09%). These results obtained for the step test are similar to those reported by Meek and Hatfield (1994). Only 0.12% of semihourly wind speed data were flagged by persistence test, with no errors found in 40 stations. Percentages of flagged data for maximum and mean daily wind speed were, on average, 0.05% and 0.04% respectively, with the highest percentages being detected in "Almuñécar" (Granada) station. Visual inspection applied to wind speed data from this station showed a constant value for a time period corresponding to maintenance routines of technicians. Meek and Hatfield (1994) reported no errors detected by this test for daily data and 0.3% of flagged data for hourly values. However, Schroeder et al. (2005) described similar results to the present study (0.07%).

3.4. Air temperature

The application of the range test to air temperature data did not flag any values, as can be seen in Table 4. However, results from the internal consistency test showed, on average, 0.012% of anomalous daily maximum data and 0.091% as the highest percentage of flagged data in "La Puebla del Rio II" (Sevilla) station, no errors being identified in 69 stations. These results are slightly higher than those reported by Feng et al. (2004) in 726 weather stations and very similar to those obtained by Reek et al. (1992) in 138 stations (0.019%). The average percentage of flagged minimum air temperature data was 0.005%, with the maximum percentage obtained in "Adra" (Almería) station (0.175%) and no errors detected in 85 stations. For semihourly temperature data, 0.11% of them, on average, were flagged by step test, with the maximum percentage detected in "Úbeda" (Jaén) station. Similar results for this test were reported by different authors (Eskridge et al., 1995; Kunkel et al., 1998 and Graybeal et al., 2004a), the percentages described by Meek and Hatfield (1994) being nearly equal in two stations (0.1%). This last work reported the same results for the persistence test as those obtained in the present study (no erroneous data detected).

The spatial consistency procedure conducted for air temperature data (T_x, T_m and T_n) in this study was the spatial regression weighted estimate (Hubbard et al., 2005), with characteristics that make it possible to build statistical confidence intervals for testing data at the target station. Results from the application of this test at two stations are shown in Fig. 5a and b. The spatial analysis at "Los Molares" (Sevilla) station indicated that approximately 1% of the data would be flagged for maximum, minimum and average temperature if f values ranged from 1.7 to 2.5. However, at "Belmez" (Córdoba) station, f should be higher than 3.5 to detect the same percentage of potential errors. This different behavior is due to "Los Molares" (Sevilla) station being surrounded by nearby stations with similar characteristics, all of them located in the lower region of Guadalquivir valley. The reverse occurs at "Belmez" (Córdoba) station, mainly for being an isolated station, the only one located in the Northern Province of Córdoba. A more detailed explanation of the methodology and parameters employed in the spatial consistency test are reported by Estévez (2008) and its application to other climate data (Estévez and Gavilán, 2006). Nevertheless, percentages of flagged data (at $f = 3$) by this test are reported in Table 4. These results, on average, are similar to those reported by Hubbard et al. (2005) for maximum and minimum temperature under a variety of climate conditions.

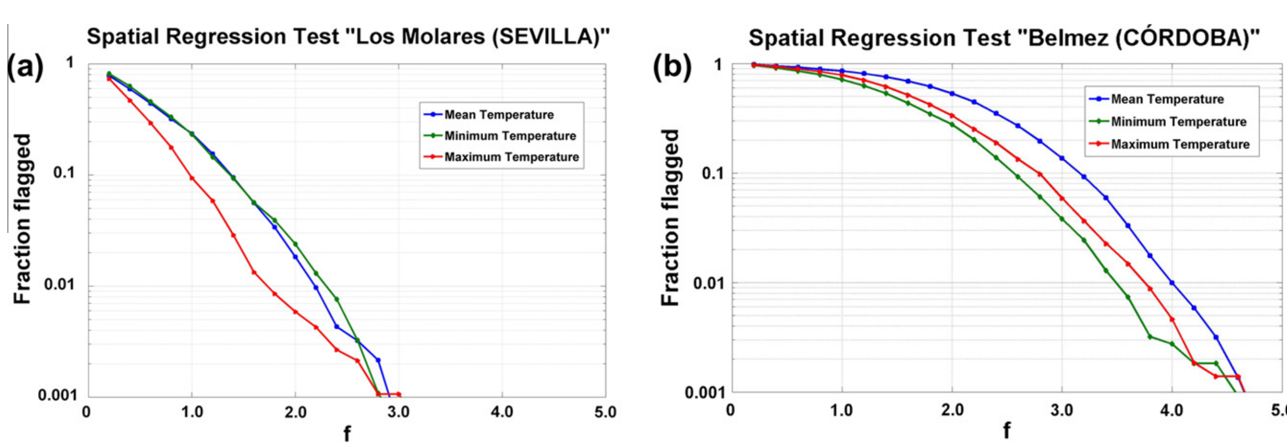

Fig. 5. Results from the spatial consistency test showing the fraction of all data flagged for maximum, minimum and mean temperature for the values of f shown at "Los Molares (Sevilla)" (a) and "Belmez (Córdoba)" (b) stations.

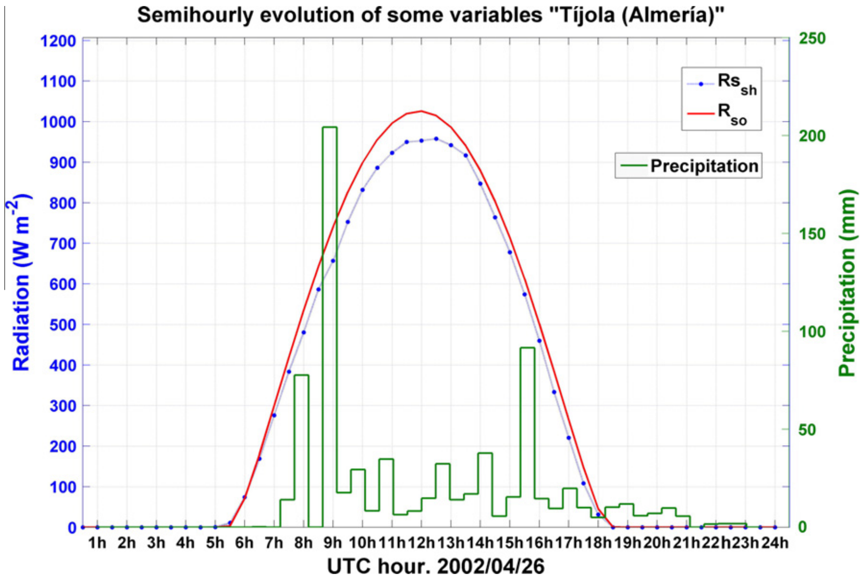

Fig. 6. Semihourly evolution of global radiation (Rs_{sh}), expected global radiation under clear sky conditions (R_{so}) and Precipitation at "Tíjola (Almeria)" station on 2002/04/26.

3.5. Precipitation

The range test applied to precipitation data for daily and semihourly values only identified erroneous values in "Tíjola" (Almería) station, for both types of records (Table 4). Only three values from more than seven million records were flagged for semihourly data. For the daily time step, only one precipitation datum was identified as being erroneous, corresponding to the same date of semihourly data flagged. Fig. 6 shows the evolution of some variables at this station for the date marked after range test application. Visual inspection and information from routine maintenance schedule were very useful for checking some alerts generated by the quality control system. In this case, the outliers detected by the range test were due to operational procedures to check whether the rain gauge was working properly during the maintenance works. Similar results were reported by Reek et al. (1992) and some works with no data flagged by this test were reported by Meek and Hatfield (1994) and Feng et al. (2004).

Results obtained from internal consistency test showed a maximum percentage of errors (0.372%) at the Córdoba station, where the sprinkler irrigations frequently causing these "false positives" is a well known fact, due to irrigation water being received by the rain gauges. In addition, 0.015% of data were flagged on average by this test, with no erroneous values identified in 36 weather stations. Vibration of farm machinery near the station can also record accidental pulses in the rain gauge. These input values are stored as precipitation data, although they do not correspond to rain, so it is necessary to correct them in order to prevent their inclusion in the soil water balance.

4. Summary and conclusions

A growing number of climate change and variability studies, as well as applied research towards improving engineering design climatographies, require high-quality meteorological data. It is known that long-term and extreme-value climate data sets are necessary for accurate and reliable estimates and assessments, being the validation of these data necessary for present and future uses. Several quality assurance procedures have been introduced for assessing integrity of meteorological data from automated weather station networks for hydrological, agricultural and other applications. Range, step, internal consistency, persistence and spatial consistency tests have been applied to the 1999–2006 dataset in the Andalusia region (southern Spain). The results of applying different validation tests reveal the existence of erroneous and potentially inaccurate data in the two databases validated (daily and semihourly). The temporal analysis of flagged data identified by these procedures and the use of dynamically generated control charts have allowed the making of such important decisions as the replacement of sensors or debugging of flagged records before being made available to users. Moreover, these guidelines have been presented for assessing quality of climate data, in science, as a prerequisite to their use for the estimation of environmental parameters or their introduction as baseline data in climate change or hydrological models. High quality data sources are vital to scientists, engineers, and decision makers alike. All the recommended guidelines are simple and straightforward to use and serve as preliminary filters with which to scrutinize weather measurements. The purpose of using quality assurance flags is to provide guidance to those who use data from the Agroclimatic Information Network of Andalusia. Because no data are altered, the user can make the final judgment as to their quality. Finally, constant improvement of the quality assurance procedures is necessary according to the latest available technology and techniques in data analysis, this validation data being only part of the quality management system, also composed of sensor calibration, field maintenance routines and manual inspection.

Acknowledgments

This work was supported by the Spanish National Institute of Agricultural Research (INIA) under Project RTA04-063. The authors are grateful to Mr. Juan de Haro for his valuable work in the Agroclimatic Information Network of Andalusia.

References

AEMET, 2008. Resumen de extremos climatológicos en España. <http://www.aemet.es/documentos/es/divulgacion/resumen_efemerides/Resumen_extremos.pdf> (accessed 03.08.10, in Spanish).

Allen, R.G., 1996. Assessing integrity of weather data for reference evapotranspiration estimation. Journal of Irrigation and Drainage Engineering ASCE 122, 97–106.

Allen, R.G., 2000. Using the FAO-56 dual crop coefficient method over an irrigated regions as part of an evapotranspiration intercomparison study. Journal of Hydrology 229, 27–41.
Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration: Guidelines for computing Crop Water Requirements. FAO Irrigation and Drainage Paper no. 56, Rome.
Allen, R.G., Smith, M., Pereira, L.S., Perrier, A., 1994. An update for the calculation of reference evapotranspiration. ICID Bulletin 43, 35–92.
ASCE-EWRI, 2005. The ASCE Standardized Reference Evapotranspiration Equation. Environmental and Water Resources Institute of the ASCE Standardization of Reference Evapotranspiration Evapotranspiration Task Committee. American Society of Civil Engineers, Reston, Virginia, p. 216.
Barnes, S.L., 1964. A technique for maximizing details in numerical weather map analysis. Journal of Applied Meteorology 3, 396–409.
Cuadrat, J.M., Vicente, S.M., Saz, M.A., 2002. Fuentes de información en climatología: Incertidumbres de las series de datos climáticos en España. VII Reunión Nacional de Climatología, 27–29 June 2002, Albarracín (Zaragoza), Spain.
DeGaetano, A.T., 1997. A quality control procedure for hourly wind data. Journal of Atmospheric and Oceanic Technology 14, 308–317.
Del Greco, S., Urzen, M., Ansari, S., 2005. Automated Validation Precipitation Data Using GIS and Spatial Interpolation Schemes. General Assembly European Geosciences Union, 24–29 April 2005. Vienna, Austria.
Doraiswamy, P.C., Pasteris, P.A., Jones, K.C., Motha, R.P., Nejedlik, P., 2000. Techniques for methods of collection, database management and distribution of agrometeorological data. Agricultural and Forest Meteorology 103, 83–97.
Eischeid, J.K., Baker, T.K., Díaz, H.F., 1995. The quality control of long-term climatological data using objective data analysis. Journal of Applied Meteorology 34, 2787–2795.
Eskridge, R.E., Alduchov, O.A., Chernykh, I.V., Panmao, Z., Polansky, A.C., Doty, S.R., 1995. A comprehensive aerological reference data set (CARDS): rough and systematic errors. Bulletin of the American Meteorological Society 76, 1759–1775.
Estévez, J., 2008. Diseño de un sistema integrado para el control de calidad de datos de estaciones meteorológicas automáticas. PhD dissertation, Department of Agronomy, University of Cordoba, Cordoba, Spain.
Estévez, J., Gavilán, P., 2006. Quality control procedures in the Agroclimatic Stations Network of Andalusia. In: Proceedings of the 4th International Conference on Experiences with Automated Weather Stations, 24–26 May 2006. Lisbon, Portugal.
Estévez, J., Gavilán, P., Berengena, J., 2009. Sensitivity analysis of a Penman-Monteith type equation to estimate reference evapotranspiration in southern Spain. Hydrological Processes 23, 3342–3353.
Feng, S., Hu, Q., Qian, Q., 2004. Quality control of daily meteorological data in China, 1951–2000: a new dataset. International Journal of Climatology 24, 853–870.
Fiebrich, C.A., Crawford, K.C., 2001. The impact of unique meteorological phenomena detected by the Oklahoma Mesonet and ARS Micronet on automated quality control. Bulletin of the American Meteorological Society 82, 2173–2187.
Flores, P., Tapia, A., Tapia, G., 2005. Application of a control algorithm for wind speed prediction and active power generation. Renewable Energy 30, 523–536.
Gandin, L.S., 1988. Complex quality control of meteorological data. Monthly Weather Review 116, 1137–1156.
Gavilán, P., Lorite, I.J., Tornero, S., Berengena, J., 2006. Regional calibration of Hargreaves equation for estimating reference ET_0 in a semiarid environment. Agricultural Water Management 81, 257–281.
Gavilán, P., Estévez, J., Berengena, J., 2008. Comparison of standardized reference evapotranspiration equations in Southern Spain. Journal of Irrigation and Drainage Engineering ASCE 134, 1–12.
Gázquez, A., Perpinyà, M., Peña, J., 2003. The XAC –Agroclimatology Network of Catalonia (La Xarxa Agrometeorològica de Catalunya). In: 3rd International Conference on Experiences with Automatic Weather Stations, 19–21 February 2003, Torremolinos (Málaga), Spain.
Geiger, M.L., Diabate, L.M., Wlad, L., 2002. A web service for controlling the quality of measurements of global solar irradiation. Solar Energy 73, 475–480.
Graybeal, D.Y., 2005. Relationships among daily mean and maximum wind speeds with application to data quality assurance. International Journal of Climatology 26, 29–43.
Graybeal, D.Y., DeGaetano, A.T., Eggleston, K.L., 2004a. Improved quality assurance for historical hourly temperature and humidity: development and application to environmental analysis. Journal of Applied Meteorology 43, 1722–1735.
Graybeal, D.Y., DeGaetano, A.T., Eggleston, K.L., 2004b. Complex quality assurance of historical hourly surface airways meteorological data. Journal of Atmospheric and Oceanic Technology 21, 1156–1169.
Grüter, E., Häberli, C., Küng, U., Mumenthaler, P., Mettler, J., Bassi, M., Konzelmann, T., Dösseger, R., 2001. The next generation of quality control tools for meteorological data at MeteoSwiss DACH-MT 2001.
Guttman, N.B., Karl, C., Reek, T., Shuler, V., 1988. Measuring the performance of data validators. Bulletin of American Meteorological Society 69, 1448–1452.
Howell, T.A., Meek, D.W., Phene, C.J., Davis, K.R., McCormick, R.L., 1984. Automated weather data collection for research on irrigation scheduling. Transactions of ASAE 27 (2), 386–391. 396.
Hubbard, K.G., 2001. Multiple Station Quality Control Procedures. Automated Weather Stations for Applications in Agriculture and Water Resources Management. World Meteorological Organization Tech. Doc. AGM-3 WMO/TD No. 1074, pp. 133–136.
Hubbard, K.G., Goddard, S., Sorensen, W.D., Wells, N., Osugi, T.T., 2005. Performance of quality assurance procedures for an applied climate information system. Journal of Atmospheric and Oceanic Technology 22, 105–112.
Hubbard, K.G., You, J., 2005. Sensitivity analysis of quality assurance using the spatial regression approach. A case study of the maximum/minimum air temperature. Journal of Atmospheric and Oceanic Technology 22, 1520–1530.
Irmak, S., Irmak, A., Allen, R.G., Jones, J.W., 2003. Solar and Net Radiation-Based equations to estimate reference evapotranspiration in humid climates. Journal of Irrigation and Drainage Engineering ASCE 129, 336–347.
Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford Univ. Press, Oxford.
Jensen, D.T., Hargreaves, G.H., Temesgen, B., Allen, R.G., 1997. Computation of ET_0 under nonideal conditions. Journal of Irrigation and Drainage Engineering ASCE 123, 394–400.
Kunkel, K.E., Andsager, K., Conner, G., Decker, W.L., Hillacker, H.J., Kmox, P.N., Nurnberger, F.V., Rogers, J.C., Scheeringa, K., Wendland, W.M., Zandlo, J., Angel, J.R., 1998. An expanded digital daily database for climatic resources applications in the midwestern United States. Bulletin of American Meteorological Society 79, 1357–1366.
Lanzante, J.R., 1996. Resistant, robust and nonparametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. International Journal of Climatology 16, 1197–1226.
Meek, D.W., Hatfield, J.L., 1994. Data quality checking for single station meteorological databases. Agricultural and Forest Meteorology 36, 85–109.
Meyer, S.J., Hubbard, K.G., 1992. Nonfederal automated weather stations and networks in the United States and Canada: a preliminary survey. Bulletin of the American Meteorological Society 73, 449–457.
Miller, P.A., Barth, M.F., 2003. Ingest, integration, quality control and distribution of observations from State Transportation Departments using MADIS. In: 19th International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology (IIPS) February 9–13, 2003, Long Beach, California
Monteith, J.L., Unsworth, M.H., 1990. Principles of Environmental Physics, second ed. Edward Arnold, London, UK.
Moradi, I., 2007. Quality control of global solar radiation using sunshine duration hours. Energy 34, 1–6.
O'Brien, K.J., Keefer, T.N., 1985. Real-time data verification. Computer applications in water resources. In: Proc. ASCE Spec. Conf., Buffalo, NY, ASCE, New York, NY, pp. 764–770.
Pérez de los Cobos, P., Carazo, J.I., Padilla, F., 2003. Agroclimatic information system for irrigation areas. In: 3rd International Conference on Experiences with Automatic Weather Stations. 19, 20 y 21 Febrero 2003. Torremolinos (Málaga).
Reek, T., Doty, S.R., Owen, T.W., 1992. A deterministic approach to the validation of historical daily temperature and precipitation data from the Cooperative Network. Bulletin of the American Meteorological Society 73, 753–762.
Schroeder, J.L., Burgett, W.S., Haynie, K.B., Sonmez, I., 2005. The West Texas Mesonet: a technical overview. Journal of Atmospheric and Oceanic Technology 22, 211–222.
Shafer, M.A., Fiebrich, C.A., Arndt, D.S., Fredrickson, S.E., Hughes, T.W., 2000. Quality assurance procedures in the Oklahoma Mesonet. Journal of Atmospheric and Oceanic Technology 17, 474–494.
Snyder, R.L., Pruitt, W.O., 1992. Evapotranspiration data management in California. In: Irrigation and Drainage Session Proceedings Water Forum 1992, ASCE. August 2–6, 1992, Baltimore, MD., pp.128–133.
Vejen, F., Jacobson, C., Fredriksson, U., Moe, M., Andresen, L., Hellsten, E., Rissanen, P., Palsdottir, T., Arason, T., 2002. Quality control of Meteorological Observations. Report 8/2002 KLIMA. Norweigan Meteorological Institute, Oslo, Norway.
Wade, C.G., 1987. A quality control program for surface mesometeorological data. Journal of Atmospheric and Oceanic Technology 4, 435–453.
Weiss, A., Robb, J.G., 1986. Results and interpretations from a survey on agriculturally related weather information. Bulletin of the American Meteorological Society 67, 10–15.
World Meteorological Organization, 1993. Guide on the Global Data-Processing System. WMO-No. 305, Geneva, Switzerland.
World Meteorological Organization, 2006. Commission for Agricultural Meteorology (CAgM): The First Fifty Years. WMO-No. 999, Geneva, Switzerland.
World Meteorological Organization, 2008. Guide to Meteorological Instruments and Methods of Observations. WMO-No.8, Geneva, Switzerland.
Yoder, R.E., Odhiambo, L.O., Wright, W.C., 2005. Effects of vapor-pressure déficit and net-irradiance calculation methods on accuracy of standardized Penman-Monteith equation in humid climate. Journal of Irrigation and Drainage Engineering ASCE 13, 228–237.
Younes, S., Claywell, R., Muneer, T., 2005. Quality control of solar radiation data: present status and proposed new approaches. Energy 30, 1533–1549.
Zahumensky, I., 2004. Guidelines on Quality Control Procedures for Data from Automatic Weather Stations. WMO-No. 955. Geneva, Switzerland.