Moroccan Journal of Chemistry
ISSN: 2351-812X
http://revues.imist.ma/?journal=morjchem&page=login
salghi & al. / Mor. J. Chem. 8 N°1 (2020) 198-209

Calibration of GFS model by using real time weather data for water requirement forecasting of vegetables in Chtouka region

Athman Hafidi[1*], Ataa Abouatallah[2], Naziha AFFI[2], Rachid Salghi[1*]

[1]*Laboratory of Applied Chemistry and Environment, National School of Applied Science , Ibn Zohr University, PO Box 1136, 80000 Agadir, Morocco;*
[2] *Independent consultants, Agadir, Morocco;*

* *Corresponding author:*
r.salghi@uiz.ac.ma.

Received 25 Oct 2019,
Revised 21 Jan 2020,
Accepted 02 Feb 2020

Abstract

In the ultimate goal to forecast weather parameters in the agriculture field and precisely in predicting water requirements, a software was developed based on Global forecasting system. To assess its reliability, data of six weather stations were used in the Chtouka region to provide measured weather parameters and allow, thus, software calibration. Air temperature, air humidity, global radiation, evapotranspiration, precipitation and wind speed/direction are studied as weather parameters. Based on statistical comparisons using normalized root mean square error (NRSE), absolute percentage error (MAPE) and mean bias error (MBE), the software was revealed efficient to provide reliable temperature, global radiation and evapotranspiration forecasting but remains unable to predict relative humidity, wind speed/direction and precipitation in an accurate way.

Keywords: Weather Forecasting, Water Management, Irrigation, Software, Big Data

1. Introduction

World population is predicted to double in the next 50 years, so greater yields must be extracted from the current agricultural areas along with more marginal areas [1]. The major problem which farmers face is the lack of sufficient available water resources. In the near future, water scarcity will increase in many regions of the world due to climate change [2] and also to urbanization. Hence, the use of the limited water resources in a very efficient manner is becoming a necessity and an efficient irrigation water management is becoming crucial. The incorporation of different types of weather forecasts into irrigation scheduling can increase the profitability of irrigated agriculture and allows better hydrological resources management. Human beings have attempted to predict the weather informally for millennia and formally since the 19th century for a variety of public and private uses, and some authors believed that an economic value resides behind weather forecasts and tried to estimate it through different methods [3]. Weather forecasts provide useful knowledge that can prevent damages, life and property losses by predicting extreme events and increase the efficiency of operations [4]. In fact, Weather plays an important role in agricultural production. It has a profound influence on the growth, development and yields of a crop, incidence of pests and diseases, water needs and fertilizer requirements in terms of differences in nutrient mobilization due to water stresses and timeliness and effectiveness of prophylactic and cultural operations on crops. The aim of this work was to assess the reliability of the Global forecasting system to predict weather parameters in the agriculture field by using real time weather data. The study was conducted in Chtouka area located in south of Morocco. This region is the agricultural leader area in the kingdom, it produces the majority of the fruits and vegetables and ensures more than 90% of the Moroccan agricultural exportation when added to Souss region. Concerning social aspects, thanks to the importance of agricultural activity in the region, thousands of workers, engineers and technicians find employment on farms, packaging stations, transport ... etc. Nevertheless, that prosperity is threatened by two serious natural challenges: water resource shortage due to severe rainfall scarcity and unexpected “heat waves”. In fact, rainfall average over the last 20 years (1993-2012) is only 170 mm instead of 250 mm defined as region annual average. Yearly, the number of rainy days rarely exceeds 30. It is expected that, compared to nowadays, water availability decrease will reach 10-15% by 2020. Besides, the evapotranspiration (ETo) is 1800mm/year with more than 7mm/day in summer. Mean winter temperatures can drop to 5°C, while average summer temperatures can reach 36°C and the hardest challenge is the unexpected climate. In fact, because of weather randomly occurring phenomenon (for instance, heat waves), Souss Massa region undergoes, every year, hard damages on crops (flower abortion) and consequently, on farmer income. Trying to solve both of pre-discussed issues, we proposed to set weather monitoring stations on some defined locations within the Chtouka region, and use the collected data for weather forecasting software calibration.

2. Materials and methods

2.1. The global forecast system model (GFS)

It is a weather forecast model produced by the National Centers for Environmental Prediction. Several atmospheric and land-soil variables are available through this dataset (temperatures, winds, precipitation, soil moisture and atmospheric ozone concentration). The entire globe is covered by the GFS at a base horizontal resolution of 18 miles (28 kilometers) between grid points, which is used by the operational forecasters who predict weather out to 16 days in the future. Horizontal resolution drops to 44 miles (70 kilometers) between grid point for forecasts between one week and two weeks. The GFS model is a coupled model, composed of four separate models (an atmosphere model, an ocean model, a land/soil model, and a sea ice model), which work together to provide an accurate view of weather with decreased spatial resolution after 10 days [5]. In our case, we used the GFS to predict weather 48 hours in the future. Data are received every 3 hours. A web-based script was implemented to connect automatically to servers and

make extrapolations within space and time using real weather stations coordinates. That software is used to provide forecasted studied parameters. Table 1 shows the used bands, units and description of the corresponding parameters.

Table 1: Predicted parameters

Band name	Description	Unit	Conversion
Temperature	2 m above ground	[K]	T(°C)=T-273.15
Relative Humidity	2 m above ground	[%]	
Wind component U	10 m above ground	m/s	
Wind component V	10 m above ground	m/s	
Solar radiation	Radiation at surface	W/m²	
Precipitation rate	Rain at surface	[kg/m^2]	

2.2. Studied region

The studied area is Chtouka region located in the southern part of Morocco. It is characterized by a semi-arid climate, dry and hot in summer and relatively cold in winter. Temperatures are moderate, with an annual average of 18.7°C, an annual minimum of 13.3°C and an annual, maximum of 24 °C. The daily maximum temperature can reach 40 °C in summer and as low as 5 °C in winter. As all south regions, Chtouka registered for several years water shortage and very irregular and very low underground and surface water resource renewal. Moreover, according to Seifennasr et al. [6], future climate prospection demonstrates critical evolution leading to an increase of air temperature by 2°C for the 2030-2049 time period, and by 4 to 5°C towards the end of the 21st century. Concerning precipitation, the found scenario shows a reduction of 10 to 30% for the period 2030-2049, up to 60% for 2080-2099. Those predictions concluded that water deficit will likely triple towards 2050 due to increase in water demand and decrease in aquifer recharge and dam storage. Six weather stations were set in Chtouka region within well-chosen locations in order to cover the entire studied region (Fig. 1): Ait Amira station (30.25230, -9.47389): It covers the ouestern part of Chtouka region. Ait Mimoun station (30.18640, -9.53990): It covers the southern area of Chtouka region. Tin Ait Brahim(30.15343, -9.46534): located in the center of the studied area. Tin Mansour (30.13013, -9.52414): is set in the center of the studied area. Biougra station (30.17290, -9.36372): It covers the northern area. Belfaa station (29.99700, -9.56478): It covers the eastern part of the studied region

Fig. 1: Geographical location of different weather stations collecting real data (A: Ait Amira, B: AitMimoun, C: Tin Ali Mansour, D: Tin AitBrahim, E: Biougra, F: Belfaa).

2.3. Measuring materiel and data logging

Studied weather data are temperature, relative humidity, wind speed/direction, rain and solar radiation. They were measured by different sensors every 15 min, than logged by a data logger and finally transmitted wirelessly to servers. The weather station is composed of: A data logger (A753-ADCON): fully integrated unit with modem and cell battery all together. Supporting analog and digital sensors. The Analog resolution is 16-bit and data memory is more than 5MB. Combisensor (HMP45-Vaissala) placed at 2m height above the ground. Measured relative humidity ranges from 0 to 100% with an accuracy of 2%. Air temperature parameter range is - 40°C to + 60°C offering an accuracy of ± 0,1°C. Solar radiation: is measured by a pyranometer (CMP3-Kipp & Zonen) for which the spectral range is 310 – 2800 nm and a sensitivity ranging between 5...20µV/Wm². Wind speed and direction: are measured by an aluminum wind set (Pro10/2 from ADCON) with a measuring range of 0.4 ... 75m/s - 0° - 360°, respectively. Precipitation: The mainly monitored precipitation parameter is rain which is measured using a rain gauge 0,2mm (from Lambrecht) with a 200 cm^2 orifice.

2.4. *Forecasted parameters*

2.4.1. Air temperature

Problems concerning the definition of mean daily air temperature have occupied the desks of scientists since the early 1800's [7]. The American Meteorological Society defines the mean temperature as: "Mean of the temperatures observed at 24 equidistant times in the course of a continuous 24-hour period (normally the mean solar day from midnight to mid-night according to time zone of the station) [8]. Nevertheless, about 100 years ago, Hartzell [9] used a planimeterto determine mean daily temperatures from thermograph charts. Maximum and minimum thermometers offered an irresistible convenience for observers [10]. By using them, it is possible to define an equation that allowed reading at a chosen time once per day. Due to this convenience, mean of minimum and maximum daily temperatures is used in about 43% of the world [11]. Considering foresaid arguments, mean daily temperature was calculated in a fifteen minutes basis and not as a daily average using the following formula:

$$T = \frac{Tmax+Tmin}{2}$$

2.4.2. Air relative humidity

Relative humidity is a term used to describe the amount of water vapor in a mixture of air and water vapor. It is defined as the partial pressure of water vapor in the air-water mixture, given as a percentage of the saturated vapor pressure under those conditions. The relative humidity of air thus changes not only with respect to the absolute humidity (moisture content) but also temperature and pressure, upon which the saturated vapor pressure depends. Relative humidity is often used instead of absolute humidity in situations where the rate of water evaporation is important, as it takes into account the variation in saturated vapor pressure [12]. The RH is the amount of moisture in the air (via moisture mass or vapor pressure) divided by the maximum amount of moisture that could exist in the air at a specific temperature (via max moisture mass or saturation vapor pressure). RH is expressed as a percentage and has no units since the units in both the numerator and denominator are the same. The percentage is found by multiplying the ratio by 100% [13].

$$\text{Relative humidity (\%)} = \frac{ea}{es} \times 100$$

Where *ea* is the air actual vapor pressure and *es* is the vapor pressure when the air is saturated. As far as all other parameters, they were measured as foresaid and then transferred to the logger before being added to the cloud.

Calculation and Statistical analysis

The choice of the relevant criteria allowing performance evaluation of the estimation methods is an important issue. Various statistical parameters can be used to measure the strength of the statistical relationship between the estimated values and the reference values [14]. To evaluate the used software, three statistic errors were considered:

The mean absolute percentage error (MAPE): is a measure of accuracy in a fitted time series value in statistics, specifically trending. It usually expresses accuracy as a percentage, and is defined by the following formula

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{I-Ip}{I}$$

Where, I is the actual value and Ip is the forecast value. The difference between I and Ip is divided by the actual value I. The absolute value of this calculation is summed for every fitted or forecast point in time and divided by the number of fitted points n. This makes it a percentage error that allows to compare the error of fitted time series that differ in level.

Mean bias error (MBE)*:* is an indication of the average deviation of the predicted values from the corresponding measured data and can provide information on the long term performance of the models.

$$\text{MBE} = \frac{1}{N}\sum_{i=1}^{n}(Ipi - Ii)$$

Root mean square error (RMSE) and normalized RMSE: In statistics, RMSE measures residual errors which give a global idea of the difference between the observed and modeled values [15]. RMSE is a frequently used measure of the differences between values predicted by a model or an estimator and the values actually observed from the thing being modeled or estimated [16]. It provides information on the short term performance and is a measure of the variation of predicted values around the measured data. It indicates the scattering of data around the linear lines. Moreover, RMSE shows the efficiency of the developed network in predicting a future individual values. Large positive RMSE means big deviation in the predicted value form the real one while RMSE value equals to 0, the performance of the model shows excellence.

$$\text{RMSE} = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}(Ipi - Ii)^2\right)} \quad [17]$$

Where, I_{pi} is the predicted value, I_i is the measured one and n is the number of observations.

Because of heterogeneity within studied parameter scales and in order to get homogeneous and, hence, comparative results, we used the normalized root-mean-square deviation or error (NRMSD or NRMSE). Though there is no consistent means of normalization in the literature, common choices are the mean or the range (defined as the maximum value minus the minimum value) of the measured data. This value is commonly expressed as a percentage, where lower values indicate less residual variance [18].

N-RMSE = $\frac{\text{RSME}}{\text{MAX y} - \text{MIN y}}$x100, where y is the measured variable.

Regression analysis

Regression analysis can be defined as the establishment of relationships between a dependent variable (output) and independent variables (inputs). Either linear or nonlinear multiple regression analyses can be employed to construct plausible relationships describing the variation of a dataset including predetermined inputs and output. Multiple regression analysis is a simple and powerful tool for prediction and forecasting. Meaningful relationships between inputs and outputs can be established as far as it is assumed that the sample under consideration is representative of the population, and the variance of errors is homogenously distributed within the dataset. Both linear and nonlinear

models aim to minimize the sum of squared residuals, whereby smaller residuals would indicate better goodness-of-fit. Linear regression methods aim to fit a linear combination of functions to a predetermined dataset.

We assume that y_i (i =1, n) is the set of n reference values and e_i (i =1, n) is the set of the estimates and $\bar{y}$, $\bar{e}$ are mean of reference and estimates values respectively. The bias, linear correlation coefficient (r), Root Mean Squared Error (RMSE) can be calculated by using standard deviations of reference (σy) and estimate (σe) values, mean of reference and estimates values, estimated values and the reference values. The bias which is the difference between the mean estimate e and the mean reference value y. The statistical criteria formula of the linear correlation coefficient r is the following,

$$r = \frac{\sum_{i=1}^{n}(yi-y)(ei-e)}{n\ \sigma y\ \ \sigma e}$$

Where, r measures the proximity between estimate and reference. It is not sensitive to a bias.

Evapotranspiration (ETo)

As far as ETo computing, Jensen *et al.* [19] compared 20 reliable methods used for arid and humid locations. They found that the Penman-Monteith method as modified by Allen [20] was the most accurate for both of them [20]. That's why, we used the Penman-Monteith method based on air temperature, relative humidity, wind speed, and solar radiation data predicted and measured.

3.Results and discussion

Forecasted weather parameters values were collected using a software based on Global forecasting system and then compared with real values received from weather monitoring stations placed in different zones of Chtouka region in order to evaluate the model performance in terms of weather parameter prediction. The graph presenting the daily measured and forecasted air relative humidity shows that this climatic parameter is under-estimated when forecasted. In fact, real values are twice of forecasted ones since NRMSE and VAF are respectively, 38% and 44% showing that deviation between real and predicted air relative humidity is large (Fig. 2).

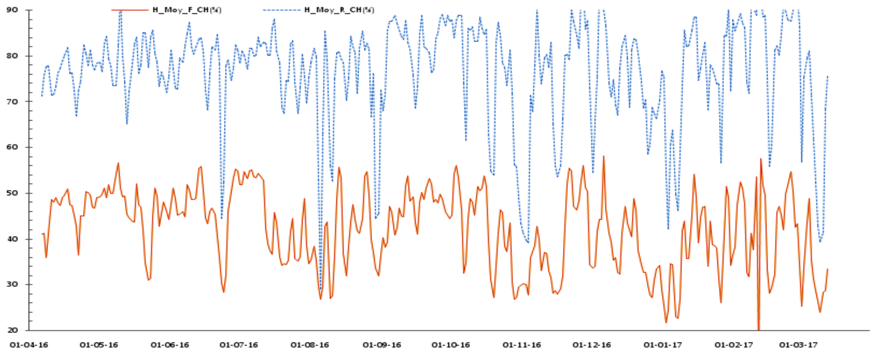

Fig. 2: Forecast and daily measured air relative humidity fluctuation.

Nevertheless, both of two variables are linearly related since R^2 is about 45% (Fig. 3). Comparing that situation with the same works (same parameter, same studied period) that was performed within the Souss Upstream and Souss center regions, we notice that the correlation between measured and forecasted air relative humidity is higher in Souss Upstream (88%), lower in Souss Center (70%) [21], much lower in Souss downstream (35%) [22] and Chtouka region (45%). Thereby, it seems that this forecast software efficiency to predict air relative humidity in Souss Massa region is

affected by coastal proximity. Temperature parameter is frequently used as an explanatory variable for forecasting other variables and processes, such as stream flow, electricity loading, river water, surface energy fluxes or solar radiation [23]. Hence, accurate prediction of temperature can provide higher forecasting accuracy of other variables [24]. Fig.4 shows forecast and measured daily averaged air temperature evolution during the year 2016-2017. Seasonal and daily temperature variation tendency is about similar since they are highly correlated (R2 = 85%) (Fig.5). NRMSE, VAF ratio and MAPE are respectively, 28%, 77% and 19% proving a high performance according to temperature parameter forecasting. Comparing software weather parameter forecasting reliability, it appears that air temperature is the most perfectly predictable parameter.

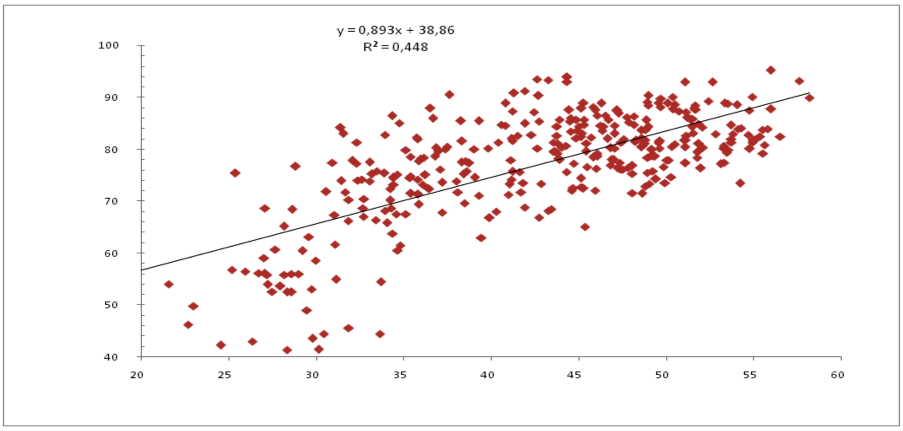

Fig. 3: Linear regression relating daily mean forecast and measured air relative humidity.

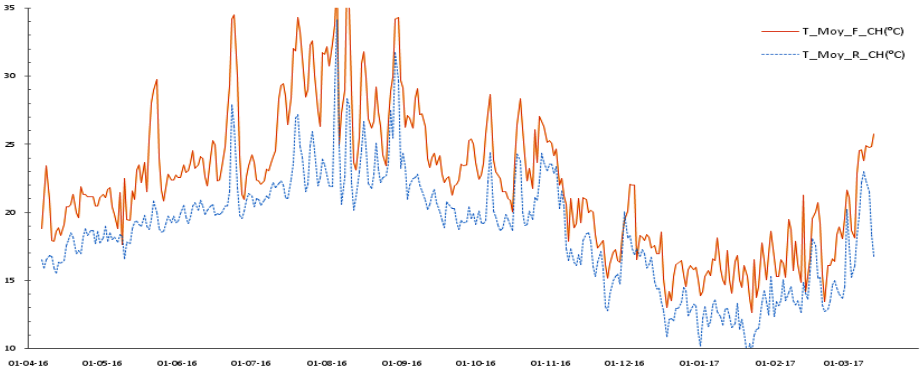

Fig. 4: Forecast and measured daily mean air temperature fluctuation.

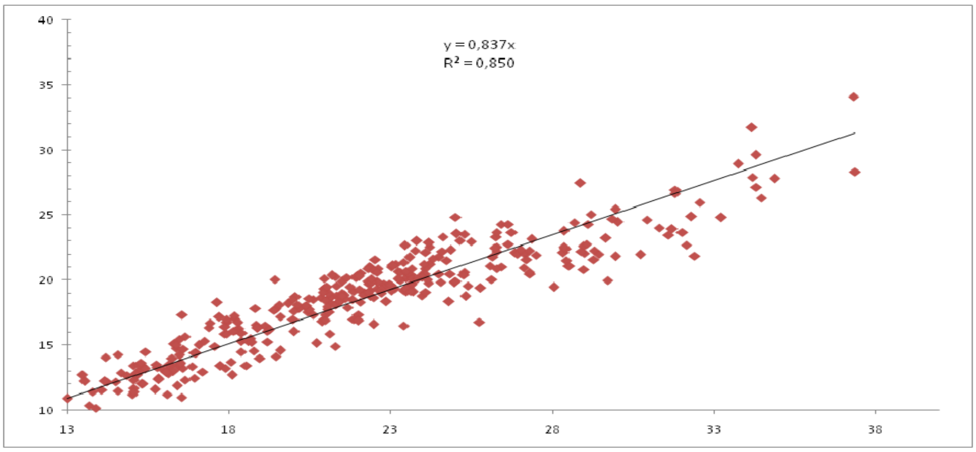

Fig. 5: Linear regression relating daily forecast and measured temperature.

Fig. 6 illustrates global radiation variation during the studied period. As far as trend, forecasted and measured values follow same trend showing similar seasonal variation. That fact is corroborated by the high linear regression between forecasted and measured global radiation proved by its coefficient that is about 70% (Fig. 7). Previous similar study was carried out in Souss upstream and center revealed that such relationship is likewise good and even better (R^2> 80%). Nevertheless, it is obvious, according to statistical parameters, that values of global radiation are badly forecasted since NRMSE, VAF and MAPE are, respectively, 29, 66% and 30%. The large variance between forecast and measured global radiation values in one hand and the high linear correlation in the other hand could suggest the existence of an increase ratio to bring back forecast global radiation to measured one.

As far as evapotranspiration is concerned, the fig. 8 shows its variation during the study period. The first finding is that winter period registered closer forecast and measured values in contrast to summer period. Besides, as it was concluded for fore discussed parameters, both forecast and measured ETo have almost similar trends along the observation period which is explained by the linear regression relationship quite high (52%) (Fig.9).

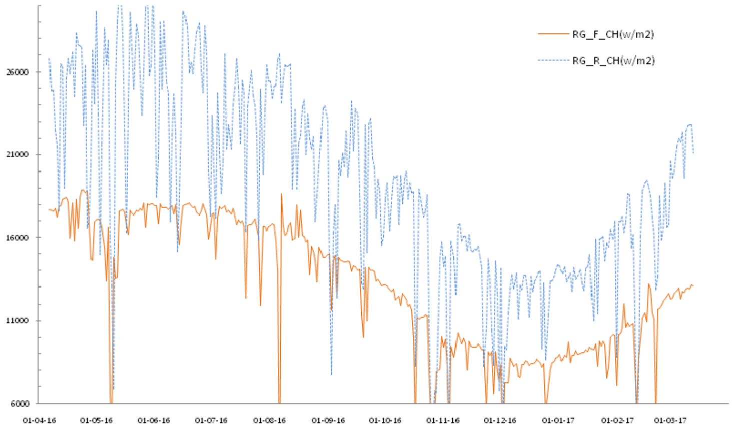

Fig. 6: Forecast and measured daily global radiation fluctuation.

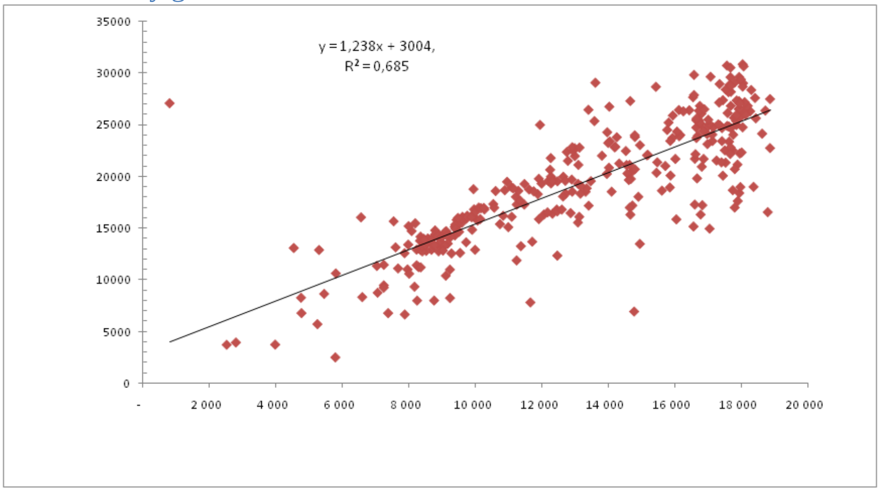

Fig. 7: Linear regression relating daily forecast and measured global radiation.

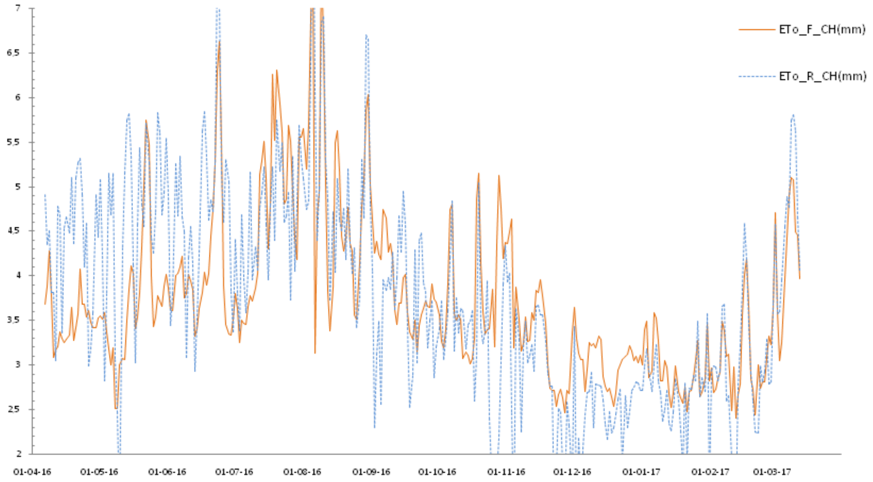

Fig. 8: Forecast and measured daily evapotranspiration.

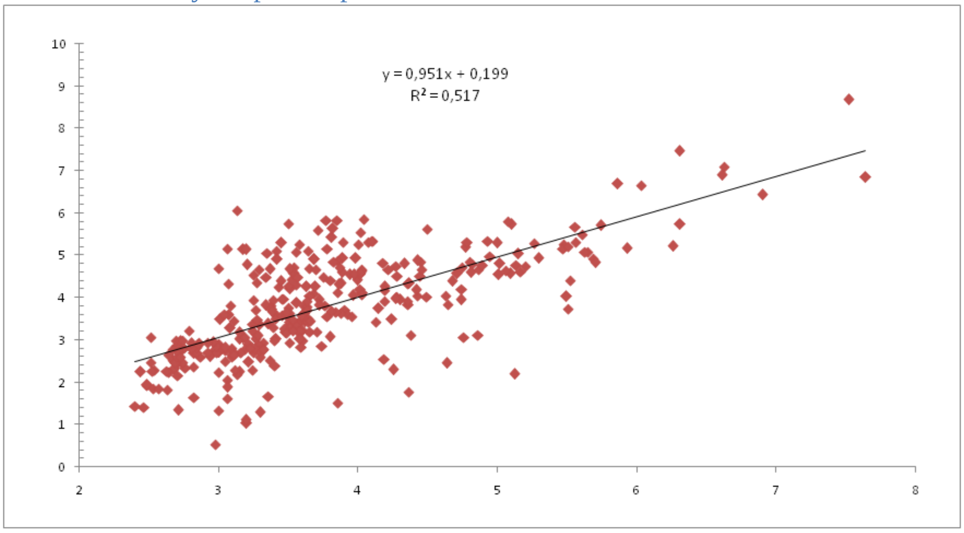

Fig. 9: Linear regression relating daily forecast evapotranspiration.

Regarding the two variable proximity assessing forecasting accuracy, the ETo parameter is the most exactly predicted. In fact, NRMSE is about 18 which is the least one within all the studied parameters. In addition, MBE is about 2, VAF is 52% and MAPE is 7% leading to conclude that this parameter is well forecasted. The used formula to calculate ETo (Penman Monteith) is a non-linear model. That's why it can be reliably forecasted despite that some other fore mentioned used parameters are not. Same conclusion was found by Ben Daoud *et al.* [25] who noticed that forecast of evapotranspiration is relevant. According to the Fig. 10, it can be noticed that precipitation parameter forecasting leads to higher values comparing to measured ones. Such observation is corroborated by statistic. In fact, NRMSE, MBE, VAF and MAPE are, respectively, 28, 60%, 67% and 50% showing very high difference between forecast and measured precipitation values. Moreover, MBE value is too large to prove the overestimation of precipitation forecast values, fact corroborated by the Fig. 11. In the same way, the linear regression between forecast and measured precipitation is too low (R^2=20%) which is in contrast with Ben Daoud *et al.* [25] who confirmed that precipitation is one of most predictable weather parameter. Linear regressions presented by Figs.13and 15 forecast wind speed and direction seem to be weakly associated to their measured homologous since R^2 are about 30%. Same finding proved through high calculated statistical parameters since wind speed NRMSE, MBE and MAPE are, respectively, 28, 41, 25% in addition to wind direction for which NRMSE, MBE and MAPE are 85, 9511 and 36%. Those very large

values indicate that wind direction is more unpredictable than wind speed and even than all other fore-discussed parameters (fig. 12 and 14).

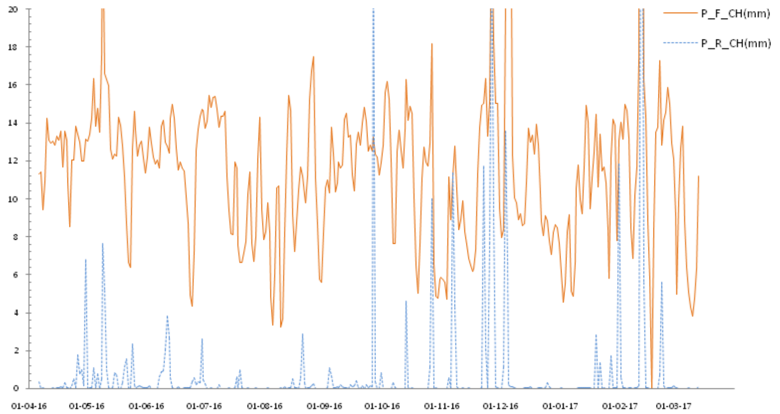

Fig. 6: Forecast and measured daily mean precipitation.

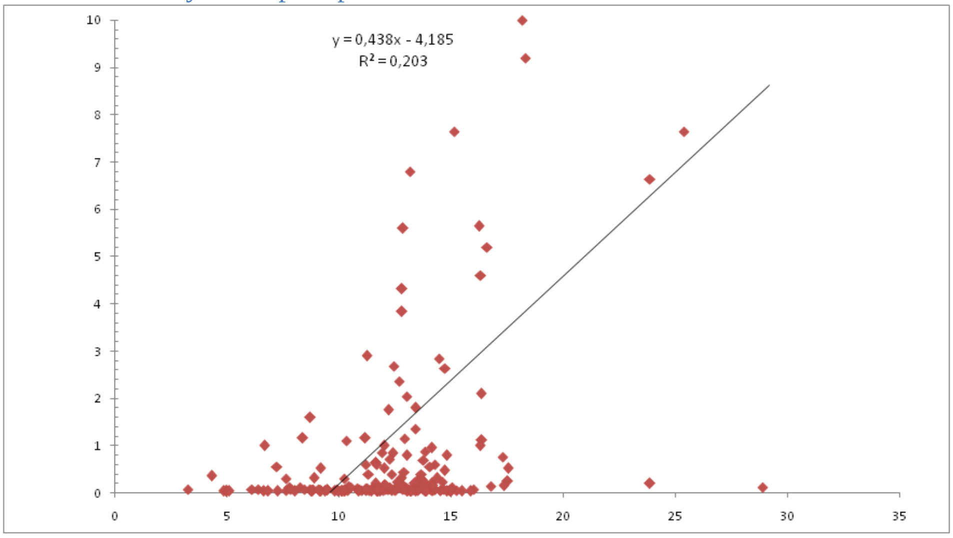

Fig. 7: Linear regression relating daily mean forecast and measured precipitation.

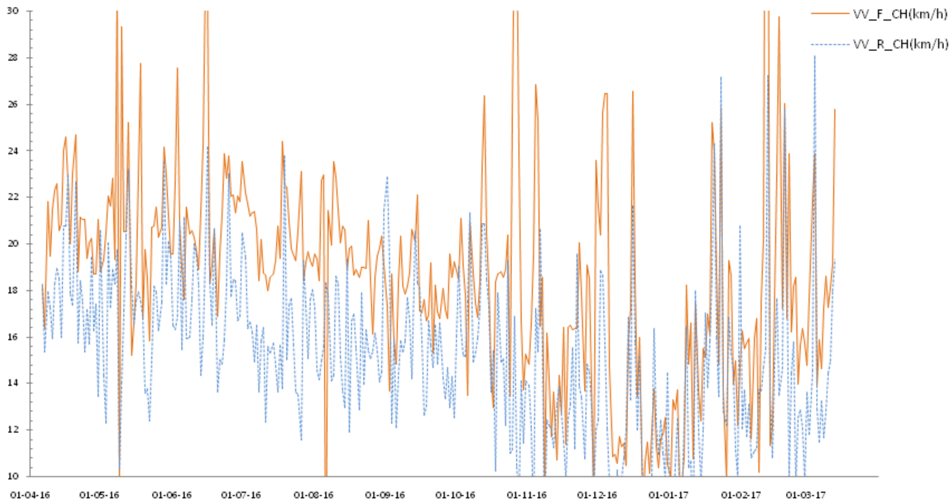

Fig. 8: Forecast and measured wind speed fluctuation.

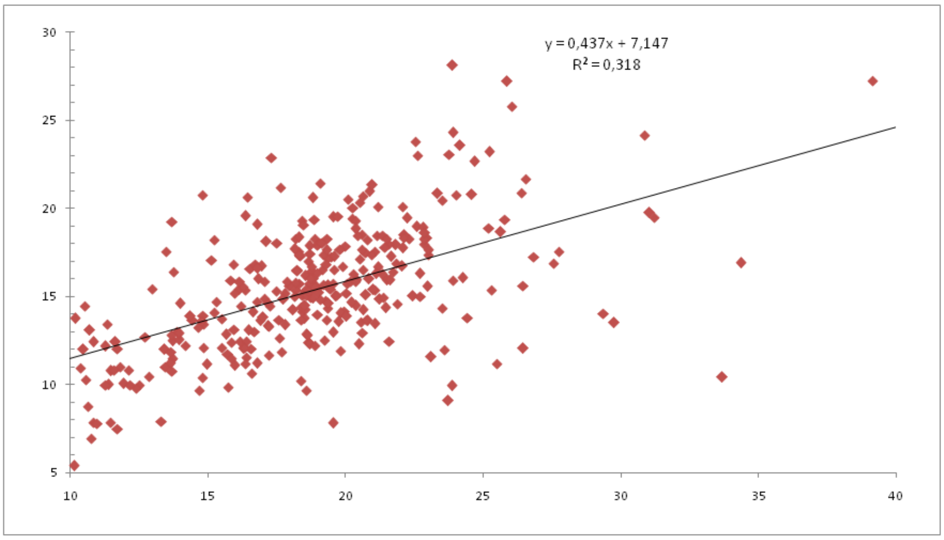

Fig. 9: Linear regression relating forecast and measured wind speed.

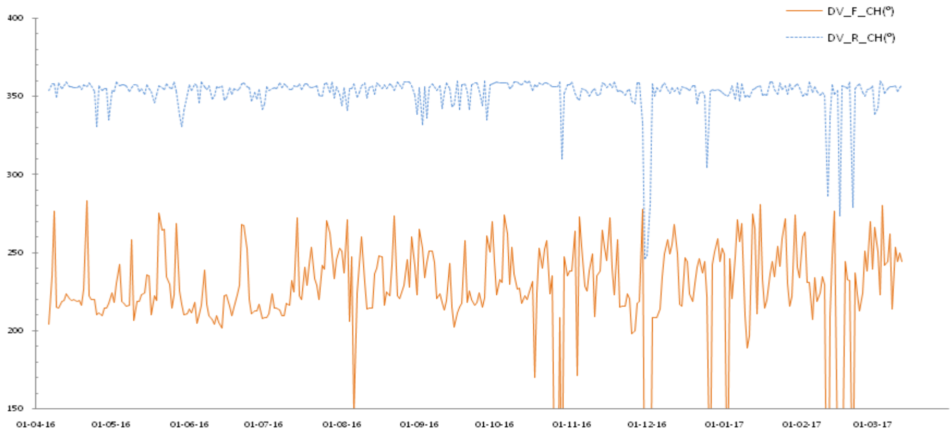

Fig. 10: Forecast and measured wind direction fluctuation.

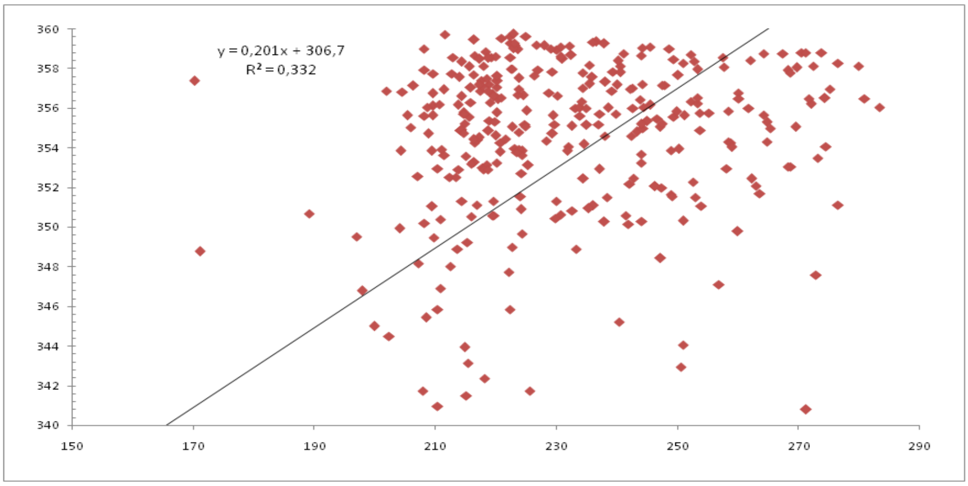

Fig. 11: Linear regression relating forecast and measured wind direction.

4.Conclusion

In this paper, the reliability of a forecasting software based on Global forecasting system was tested. The assessment of this model reliability to predict weather parameters in Chtouka area shows that this software is efficient when predicting temperature and referential evapotranspiration with a correlation of 85% and 52% respectively. However, it

is moderately efficient when predicting air relative humidity and global radiation which needs more efforts in calibration. Whereas, this model is unreliable to predict wind speed/direction and precipitation parameters. These obtained results are currently used to calibrate the Global forecasting system software in order to develop a new model called Yobeen which will be more precise in predicting weather parameters in the south region of Morocco. This latter will be described in details in our future work.

References

[1] M.M. Chaves, B. Davies, *Functional Plant Biology*, 37 (2010) iii-vi.

[2] P. Martin, C. Osvaldo, P. Jean, V. Paul, H. Clair, Printed in USA at the University Press, New York (2007).

[3] X. Cai, M. Hejazi and D. Wang, *Journal of Water Resources Planning and Management*, 137 (5) (2011) 391-403

[4] L. Al-Matarneh, A., Sheta, S., Bani-Ahmad, J., Alshaer, I., Al-oqily, *International Journal of Multimedia and Ubiquitous Engineering*, 9 (12) (2014) 343–366.

[5] M. O. Mughal, M. Lynch, F. Yu, B. McGann, F. Jeanneret, J. Sutton, *Environmental Modelling& Software*, 90 (2017) 107-125.

[6] M. Seifennasr, Z. Rashyd, A. Hirich, G.B. Caroletti, Z.A. Lhoussaine El Morjani, H. Beraaouz, R., McDonnell, R. Choukr-Allah, *The Science of the total environment*, 573 (2016) 862-875.

[7] O. Bonacci, I. Željković, R. ŠakićTrogrlić, J. Milkovi, *Theoretical and Applied Climatology*, 114 (2013) 271–279.

[8] TS. Glickman (ed), Glossary of meteorology, Second Edition, published by American Meteo-rological Society, Boston 2000

[9] FZ. Hartzell, *Monthly Weather Review*, 47(11) (1919) 799–801

[10] G. Conner, S. Foster, Paper 4.3, 17th Conference on Applied Climatology, 88th Annual Meeting American Meteorological Society, 12 August 2008, Whistler British Columbia, Canada, 2008

[11] NOAA (National Oceanic and Atmospheric Administration U. S.), *World Weather Records* Vols 1–6 (1997)

[12] S. AlSadi, T. Khatib, *Journal of Asian Scientific Research*, 2 (2) (2012) 80-85.

[13] L.N. Gunawardhana, R. Al-Rawas, and S. Kazama, *Meteorological Applications*; 24 (2017) 551–559

[14] M., Sahin, B.Y, Yildiz, O., Senkal, V, Pestemalci, *Gazi University Journal of Science*, 25(2) (2012) 327-335

[15] MW. Gardner, SR. Dorling, *Atmospheric Environment;* 34 (2000) 21–34.

[16] M.G., Kendall, A., Stuart, *Charles Griffin & Company Press*, London, (1963)

[17] C. Gokceoglu, *Engineering Geology*; 66, (2002), 39–51

[18] R. Pontius, O. Thontteh, H. Chen, *Environmental Ecological Statistics*, 15 (2) (2008) 111–142

[19] M.E. Jensen, R.D. Burman and R.G. Allen, *American Society of Civil Engineers and Rep. On Eng. Prac*. No. 70, New York, NY. (1990).

[20] R.G. Allen, *Journal of Irrigation and Drainage Engineering*, 112(4) (1986) 348-368

[21] A. Hafidi, R. Salghi, A. Abouatallah, M. Ayoub, F. Ayoub, N. AFFI, *International Journal of Engineering Research and Applications* (IJERA), 7 (11) (2017) 01–11.

[22] A. Hafidi, A. Abouatallah, N. AFFI, A, Chaouiki, R. Salghi, *Applied Journal of Environmental Engineering Science*, 5 (1) (2019) 42-52.

[23] S. Shamshirband, K., Mohammadi, H.L. Chen, G.N. Samy, D. Petković, C., Ma, *Journal of Atmospheric and Solar-Terrestrial Physics*, 134 (Suppl. C) (2015) 109–117.

[24] K. Mohammadi, S. Shamshirband, A.S. Danesh, M.S. Abdullah, M. Zamani, *Theoretical and Applied Climatology*, 125 (1–2) (2016) 101–112

[25] A. Ben Daoud, E. Sauquet, M. Lang, G. Bontron and C. Obled, *Advances in Geosciences*, 29 (2011) 103-107.