Modelling evapotranspiration of soilless cut roses ‘Red Naomi’ based on climatic and crop predictors

PATRICIA MALVA COSTA1, ISABEL PÔÇAS1,2,3,4, MÁRIO CUNHA1,3,4*

1Faculdade de Ciências da Universidade do Porto, Porto
2Linking Landscape, Environment, Agriculture and Food, Instituto Superior de Agronomia, Universidade de Lisboa, Lisboa, Portugal
3Geo-Space Sciences Research Centre, Universidade do Porto, Porto, Portugal
4Institute for Systems and Computer Engineering, Technology and Science INESC TEC, Campus da Faculdade de Engenharia da Universidade do Porto, Porto, Portugal

*Corresponding author: mccunha@fc.up.pt


Abstract: This study aimed to estimate the daily crop evapotranspiration (ETc) of soilless cut ‘Red Naomi’ roses, cultivated in a commercial glass greenhouse, using climatic and crop predictors. A multiple stepwise regression technique was applied for estimating ETc using the daily relative humidity, stem leaf area and number of leaves of the bended stems. The model explained 90% of the daily ETc variability ($R^2 = 0.90$, $n = 33$, $P < 0.0001$) measured by weighing lysimeters. The mean relative difference between the observed and the estimated daily ETc was 9.1%. The methodology revealed a high accuracy and precision in the estimation of daily ETc.

Keywords: Rosa hybrida L.; greenhouse crop; irrigation management; weighting lysimeter; multiple stepwise regression

Evapotranspiration (ET) represents the combination of evaporation from soil, substrate or plant surfaces and transpiration from the crop, thus integrating the interaction of atmospheric, plant and soil or substrate variables. In soilless crops, substrate is usually covered by a plastic film, so the evaporation component of the ET is often considered null.

ET can be directly obtained by several methods, e.g. lysimetry, of which the weighting lysimeter provides the most accurate data for short time periods (Allen et al. 2011). Alternatively, ET can be estimated through mathematical models based on meteorological data and crop predictors. During the last years, several models have been developed for ETc assessment of different crops in open-field conditions (Farahani et al. 2007), but little has been done for the soilless crop systems.

The FAO 56 methodology is a worldwide accepted modelling approach to determine ETc based on a reference evapotranspiration (ETo) and a crop coefficient (Kc), being the FAO Penman-Monteith equation the standard method for the computa-
tion of ETo from meteorological data (Allen et al. 1998). Although the FAO 56 methodology was developed for crops under open-field conditions it has been also applied to greenhouse crops such as tomato (Harel et al. 2014), melon, green pepper, green bean and watermelon (Orgaz et al. 2005). However, the FAO Penman-Monteith equation is demanding in terms of data input, making its application sometimes difficult in situations where weather data are limited, not available, or are unreliable. The Penman-Monteith ETc model, from which the FAO Penman – Monteith equation derived, is considered one of the primary models used in greenhouse horticulture. This model was first developed for open field conditions where climatic variables are more homogeneous (Morille et al. 2013).

The Stanghellini ETc model (Stanghellini 1987) was implemented to represent conditions in high technology controlled environment greenhouses. However this model requires the calibration of various hard-to-measure variables, e.g. aerodynamic and stomatal resistances (Villarreal-Guerrero et al. 2012). When comparing the Stanghellini and Penman–Monteith models to determine the ET of greenhouse bell pepper and tomato, (Bayer et al. 2013) concluded that even though the Stanghellini model showed the highest overall accuracy, there were no statistically significant differences between the ET predictions of both models.

In alternative, empirically-based (data-driven) ET models, i.e., not requiring a deep knowledge on biophysical mechanisms that produced the data, have been widely applied in the last years for ET greenhouse crops. Such techniques are less expensive, relatively easy to apply, and do not need a predefined structure of the model for estimating the ET. Simplified approaches to roses ETc determination, which included variables such as leaf area index, radiation intercepted by vegetation, and vapour pressure deficit (VPD) or energy input from heating system were implemented (Suay et al. 2003; Baas, Van Rijssel 2006; Mpusia 2006). Simple linear regression models of ETc against outside or inside solar radiation have also been proposed for practical management of irrigation in greenhouse crops (Bacci et al. 2011).

Nevertheless, there is still no standard and accurate method available for greenhouse’s ETc estimation (Gavilan et al. 2015). In most practical situations, the existing methodologies for direct estimation of ETc are not feasible, e.g., due to the lack of easy-to-use equipment, the high costs or other resources such as specialized technicians. Due to these limitations, ETc modeling continues to stand out as an alternative to its direct measurement.

In this context, the main goal of this study was to establish a dynamic data-driven predictive model for estimating ETc in soiless cut roses ‘Red Naomi’ for daily management of irrigation. The specific goals include (i) to develop a predictive model for ETc based on climatic and crop predictors; and (ii) contribute to efficient greenhouse irrigation automation.

**MATERIAL AND METHODS**

**Study area.** This study was carried out in a commercial glass greenhouse, “Venlo” type, in the company Floralves, located in Vila do Conde (41°19’40.8’’N 8°42’17.4’’W), in the North of Portugal. The greenhouse has a North-South orientation and an area of approximately 1 ha.

The greenhouse is exclusively occupied by the production of cut roses (*Rosa hybrida* L.) grown in a substrate cultivation system composed by coir. The cultivar ‘Red Naomi,’ which occupied about 30% of the greenhouse area, was used for the case study. The plants used in this study were transplanted in 2011 with a density of eight plants per m².

The crop was irrigated by a closed drip system with a 2 l/h discharge. A standard nutrient solution for cut roses was applied. Irrigation frequency was based on solar radiation with the sensor located outside the greenhouse. A target value of 40% for leaching fraction was established in order to maintain optimal conditions of water supply to the plants. The irrigation water was reused after disinfection with ultraviolet light. The target limit values for electrical conductivity and pH were 1.5 dS/m and 5.4, respectively.

The crop was managed by the producer following standard cultural practices, using the stem-bending system, where non-productive and lesser quality stems are bent into the canopy or aisle. The greenhouse was whitewashed in July by spraying with a thinner layer of lime.

**Data collection.** Evapotranspiration. The data used for developing the daily ETc model were collected in 22-days between mid-July and October of 2014, covering the full crop cycle. Crop management practices by the producer interfered with continuous data collection. The experimental de-
The accuracy of irrigation and drainage indirect measurements was tested (Table 1). The key points determined graphically (M2 and M3, Fig. 1) based on the records of the weighing lysimeter were consistent with observed data. The comparison between the irrigation amounts observed and determined by the lysimeter showed a mean absolute error (MAE) of 3.93 ml, corresponding to a mean relative error (MRE) of 13.53%, while for the drainage the MAE was 11.61 ml representing a MRE of 13.53%.

**Climate data.** Air temperature and relative humidity were obtained from two sensors located inside the greenhouse and recorded every minute in an automatic recording software. The daily vapour pressure deficit (VPDd) was calculated according to (Allen et al. 1998).

**Leaf area.** Crop variables such as leaf area (LA) and number of leaves of the bent stems (NL) of the sampling unit plants were also considered as input predictors in the daily ETc estimation model. The LA was obtained as described in (Costa et al. 2016). During the experiment, the LA of erect stems (SLA) of each sampling unit was measured once a week. A monthly count of the number of expanded leaves in the bent stems was made. The number of stems of the sampling units was checked two to four days a week between SLA measurements. The number of bent stems of the sampling units was also checked, between monthly measurements, and the number of expanded leaves corrected if needed.

**Irrigation and drainage.** The crop irrigation was managed by means of a radiation sensor. The daily

---

**Table 1. Statistics between direct (observed) and indirect irrigation and drainage measurements**

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Units</th>
<th>Irrigation observed</th>
<th>Irrigation weighting lysimeter</th>
<th>Drainage observed</th>
<th>Drainage weighting lysimeter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>ml</td>
<td>265.13</td>
<td>232.75</td>
<td>107.10</td>
<td>105.40</td>
</tr>
<tr>
<td>CV</td>
<td>%</td>
<td>3.93</td>
<td>8.22</td>
<td>41.66</td>
<td>41.60</td>
</tr>
<tr>
<td>MAE</td>
<td>ml</td>
<td>33.90</td>
<td>11.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRE</td>
<td>%</td>
<td>12.77</td>
<td>13.53</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CV – coefficient of variation; MAE – mean absolute error; MRE – mean relative error
leaching fraction of 12 substrate bags (three per sector), including the four sampling units of this study, was monitored three to four times a week. The daily leaching fraction was calculated through the mean of the leaching fractions of the monitored substrate bags collected during a 24 hours period. This daily leaching fraction was used to regulate the irrigation in each sector. As the irrigation system was set for automatically initiate based on a threshold of accumulated radiation, a predefined value of daily leaching fraction (40%) was considered for determining the appropriate threshold. However, values of daily leaching fraction between 35 and 45% were also accepted. The collection of the leaching fraction was made with drainage lysimeters.

The water distribution uniformity of the greenhouse sectors (sectors average of 96%) in study was evaluated accordingly to NSW (2009) method. The water intake was determined indirectly by collecting and measuring the volume of a “reference dripper”, with a graduated cylinder, from one substrate bag contiguous to the monitored substrate bags.

**Development of the model for daily ETc estimation.** The dependent variable ETc (range 0.4–4.6 mm/day; average ($\bar{x}$) = 3.0; coefficient of variation (CV) = 34.2%) was regressed against the potential predictors from climatic and crop variables. The climatic potential predictors tested were: Td – daily average air temperature (range 18.1–25.4°C; $\bar{x}$ = 21.7; CV = 8.7%); T1 – average air temperature between midnight and 6 a.m. (range 12.7–21.9 °C; $\bar{x}$ = 18.1; CV = 12.8%); T2 – average air temperature between 6 a.m. and the beginning of the first watering of the day (range 16.6–24.7°C; $\bar{x}$ = 20.7; CV = 9.7%); T10 – average air temperature between the end of the drainage of the last watering and 9 p.m. (range 18.6–26.7°C; $\bar{x}$ = 22.5; CV = 8.8%); T11 – average air temperature between 9 p.m. and midnight (range 14.8–22.5°C; $\bar{x}$ = 18.8; CV = 11.6%); RHd – daily average relative humidity (range 80–100%; $\bar{x}$ = 86.1%; CV = 5.6%); RH2 – average relative humidity between 6 a.m. and the beginning of the first watering of the day (range 84.4–100%; $\bar{x}$ = 89.6%; CV = 3.8%); RH10 – average relative humidity between the end of the drainage of the last watering and 9 p.m. (range 72.5–100%; $\bar{x}$ = 83.9%; CV = 9.1%) and VPDd (range 0.1–1.6 KPa; $\bar{x}$ = 0.9 KPa; CV = 32.7%). Potential predictors also included the erect stem leaf area (SLA) – (range 588.3–4345.2 cm²; $\bar{x}$ = 2,198.8 cm²; CV = 46.7%) and the number of leaves of the bended stems (NL) – (range 46–168; $\bar{x}$ = 107.2; CV = 42%) on the sampling unit. Combinations such as the logarithm of SLA, LN, T2, T11 and RHd, square root of SLA, LN and square of Td, T2, T10, T11 and RHd were also tested in model development.

For the model development an initial number of 42 observations were used, from which 21% (9 observations) were used for external validation. Multiple stepwise regression was applied to identify robust predictors. A critical level of 5% for the $P$-value of the Student’s $t$-test was set to include a variable in the model. At each variable inclusion step, the coefficient of determination ($R^2$) value and the change in $R^2$ value were also calculated. Assumptions of normality, homoscedasticity, and the existence of multicollinearity among the independent variables were verified. The variance inflation factor (VIF) and tolerance (T) were calculated to assess the collinearity between the model variables. The variables with VIF $>$ 10 and $T < 0.1$ were excluded from the model (MONTGOMERY, PECK 1992). Inferences about the regression parameters were checked by the 95% confidence intervals and the Student’s $t$-tests ($P < 0.05$) for the null hypothesis of the regression parameters being equal to zero. The $F$-test ($P < 0.005$) was calculated to test the significance of the independent variables as a group for predicting the ETc. The regression mean prediction interval for the 95% probability level was calculated according to MONTGOMERY, PECK (1992) and was presented graphically.

**Model validation and prediction accuracy.** Nine observations not used in the model parameter estimation were considered for the external validation, aiming to evaluate the prediction reliability. An additional validation was applied over the full set of data ($n = 33$) using the “leave-one-out” (LOO) cross-validation method (CUNHA et al. 2010).

The model adequacy was assessed by the percentage of variance explained by the model, expressed by the R-square ($R^2$). We used $R^2$-LOO to show the proportion of variance explained by cross validated predictions. Also, several goodness-of-fit indicators were used for calibration and validation data-sets: i) residual indices: the root mean square error (RMSE, mm/day), the mean relative error (MRE, %) and the mean absolute error (MAE, mm/day), and ii) association measures based on the linear regression through the origin between pairs of observed and modeled ETc. We performed all analyses using IBM SPSS statistical software (Version 23.0., IBM Corp., USA).
RESULTS

The analysis of the coefficients of variation (CV) for the potential predictors used for modelling ranged between 3.8 % and 46.7%. The dependent variable also showed a high CV value (34%).

Three predictors were selected by the stepwise multiple regression model for the estimation of daily ETc: Ln_RHd, SLA and NL. Predictors and their corresponding regression coefficients for modelling the ETc are presented in Table 2. The intercept term and the regression coefficients of predictors are significantly different from zero on the basis of $t$-test at 5% level (Table 2). Also, the 95% confidence interval of the regression parameters does not contain zero, which imply that they are statistically significant. The $R^2$ stepwise represents the $R^2$ of the regression when each model predictor was added to the model. The model is statistically significant on the basis of the $F$-test ($P < 0.000$). VIF was lower than 10 and $T$ was higher than 0.1, indicating the inexistence of collinearity between variables.

The MAE, MRE and RMSE obtained for the model validation procedures, both with external data and the cross-validation, showed good results (Table 2). For the external validation, values of 0.21 mm/day and 0.38 mm/day were obtained for RSME and MAE, respectively. Lower values were obtained in the cross-validation, with RSME of 0.13 mm/day. and MAE of 0.28 mm/day. The MRE value of the cross-validation was 10.42%, lower than the value obtained by the external validation (MRE = 14.96 %).

When the observed daily ETc was plotted against the modelled daily ETc (Fig. 2), the regression slope was very close to one (0.99) and the coefficient of determination was 89%. Moreover, the MRE between the observed and estimated daily ETc was around

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Adjustment and diagnostic tests</th>
<th>Residual Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$ $\pm$ SE $t$-Student*</td>
<td>$R^2$ SE T VIF test</td>
</tr>
<tr>
<td>Constant</td>
<td>57.42 $\pm$ 5.14</td>
<td>0.74 0.53 0.86 1.17 MAE</td>
</tr>
<tr>
<td>Ln_RHd</td>
<td>$-$12.66 $\pm$ 1.13 $&lt; 0.000$</td>
<td>0.85 0.41 0.93 1.10 MRE</td>
</tr>
<tr>
<td>SLA</td>
<td>0.03 $\pm$ 0.06 $&lt; 0.000$</td>
<td>0.90 0.34 0.88 1.14 RMSE</td>
</tr>
<tr>
<td>NL</td>
<td>0.01 $\pm$ 0.01 $&lt; 0.001$</td>
<td></td>
</tr>
</tbody>
</table>

*Selected variables using the stepwise regression method ($P < 0.05$) and the value of tolerance value ($T$) and variance inflation factor (VIF) were logarithm of daily average relative humidity (Ln_RHd), erect stem leaf area (SLA), and number of leaves of the bended stems (NL), the other variables included in the model were not considered statistically significant by the stepwise regression method; *probability associated with the Student’s $t$-test.

Fig. 2. Regression through the origin between observed and modelled (estimated and LOO – “leave-one-out” cross-validation method)) daily crop evapotranspiration. The regression 1 : 1 is represented by the dashed line.
9.1%, corresponding to a MAE of 0.25 mm/day and RMSE of 0.10 mm/day. Similar results were obtained when comparing the cross validation ETc values with the observed daily ETc values (regression slope close to one and coefficient of determination of 87%).

Daily observed and modelled ETc values (both estimated and using the LOO cross validation) were compared, showing a good similarity between the sets of data (Fig. 3). The model was able to accommodate a wide range of ETc values (CV = 34.2%), which included the minimum value of 0.4 mm/day and a maximum of 4.61 mm/day. The data set for the model development covered the full crop cycle, thus contributing to the model robustness.

Additionally, in 79% (LOO cross validation) and 82% (estimation) of the cases, the errors between observed and modelled values were lower than 15% (Fig. 4), all consistently inside the prediction interval (Fig. 3).

**DISCUSSION**

When analysing the CV of the potential predictors used for the daily ETc modelling, we concluded that, with the exception of the variables Td, T2, T10, RHd, RH2 and RH10, there was a marked variability in the selected descriptors, which allowed the formulation of ETc estimates over a wide validation interval and contributed to the model robustness.

In the daily ETc model (Table 2), the regression coefficients for crop predictors related with the LA are positive while for the climatic predictors is negative, thus consistent with what would be expected for their impact on ETc (Allen et al. 1998). Hence, an increase in the LA and a decrease in the RHd results in an increase in the ETc. The model presented a satisfactory fit when using solely the RHd, with $R^2$ values of 0.74, which increased to 0.90 when the variables related with crop predictors were added to the model. Thus, 90% of the daily variability of ETc in different crop stages can be explained by the three predictors. The selection of these predictors highlights the importance of crop variables such as LA (Stanghellini 1987; Baille et al. 1994; Suay et al. 2003). Cut roses, unlike most crops, are continuously harvested, and thereby exhibit a large time-variability of their transpiration area, making LA an important parameter when modelling roses ET (Raviv, Blom 2001). The RHd

Fig. 3. Overall comparison between observed and modelled (estimated and predicted – LOO) daily ETc for 33 observations. The predicted was obtained from the leave-one-out cross validation procedure (LOO). External thicker lines are the prediction interval ($\alpha = 95\%$).

Fig. 4. Frequency levels of errors between each pair of observed and modelled (estimated or predicted) daily crop evapotranspiration.
is also a key parameter for the ET demand (Mpusia 2006). Although the VPD (inversely proportional to air humidity) is considered one of the main factors affecting greenhouse crop transpiration, along with solar radiation and stomatal resistances (Raviv, Blom 2001; Katsoulas, Kittas 2011), it was not selected as a predictor for our model. Similarly, the VPD did not improve regression models for estimating crop transpiration in a greenhouse study with roses (Baas, Van Rijssel 2006). The model developed in our study presented a $R^2$ very similar to the values obtained by (Suay et al. 2003) and (Baas, Van Rijssel 2006) although a lower number of observations were used in our study when compared to (Suay et al. 2003).

When the observed daily ETc values were plotted against the modelled daily ETc, the regression slope was close to one (0.99) and the coefficient of determination was 89%, showing the model’s ability to estimate daily ETc with high accuracy and precision (Fig. 2).

The model good predictive performance is also supported by the results obtained when comparing the LOO cross validation ETc values with the observed daily ETc values (Fig. 2; regression slope close to one and the coefficient of determination of 87%). The similarity between observed and modelled ETc data (Fig. 3) also indicates good performance of the developed model.

When compared to other simplified models used to determine the transpiration of roses (Suay et al. 2003; Baas, Van Rijssel 2006; Mpusia 2006) the developed model presents the advantage of using less predictors, which can be more easily obtained, without a large cost-investment. In the simplified roses transpiration models used by Suay et al. (2003), Mpusia (2006) and Baas and Van Rijssel (2006), radiation and VPD inputs or energy input from heating system are used. However, in an operational context, most greenhouses in the Mediterranean area have minimal climate control equipment (Montero et al. 2011), making data collection of such variables a difficult task. Our model presents the advantage of using more user-friendly inputs, such as dRH, a common variable controlled by producers. The main advantage and novelty of our model is its applicability to different stages of the crop development, by using the crop parameter SLA. The non-destructive method for determining the SLA for the cultivar ‘Red Naomi’ developed by Costa et al. (2016) allows the model to accommodate the variation of the LA throughout the entire crop cycle instead of being limited to a specific stage of crop development. This advantage is particularly relevant in roses, where harvesting is continuous. However, the determination of crop predictors (SLA and NL) selected by the model is time consuming, which can cause some constrains.

The developed model, like the ones above mentioned, has the disadvantage of using predictors that can only be determinate at the end of the day (in our study, is RHd). However, relative humidity inside the greenhouse can be predicted, primarily from external relative humidity, ventilation rate and ET (Litago et al. 2005). In either case, the RHd can be adjusted throughout the day, based on data from the greenhouse sensor. Future studies should be undertaken to test the robustness of the model in different greenhouse conditions, other cut rose cultivars, and types of greenhouses.

Acknowledgements

The authors thank Sara Gonçalves for her help. We also thank Floralves Company and Sérgio Alves.

References


Received for publication August 22, 2017
Accepted after corrections July 27, 2018