agriculture

MDPI

Article

Evaluation of Crop to Crop Water Demand Forecasting: Tomatoes and Bell Peppers Grown in a Commercial Greenhouse

Dean C. J. Rice, Rupp Carriveau *, David S. -K. Ting and Mo'tamad H. Bata

Turbulence and Energy Laboratory, Ed Lumley Centre for Engineering Innovation, University of Windsor, Windsor, ON N9B 3P4, Canada; ricee@uwindsor.ca (D.C.J.R.); dting@uwindsor.ca (D.S.-K.T.); bata@uwindsor.ca (M.H.B.)

* Correspondence: rupp@uwindsor.ca; Tel.: +1-519-253-3000

Received: 24 October 2017; Accepted: 13 December 2017; Published: 20 December 2017

Abstract: Forecasting crop water demand is a critical part of any greenhouse's day-to-day operations. This study focuses on a region located in Essex County, Ontario Canada where water demand is dominated by commercial greenhouse operations (78% of capacity). Development of complex and elaborate forecasting methods such as artificial neural networks (ANN) can be costly to develop and implement, especially with the limited resources available to greenhouses. This study proposes simplified forecasting methods that would be used in conjunction with a more complex base model architecture. These simplified methods use one crop water usage as an indicator of another's, and is titled crop-to-crop forecasting (C2C). In this study, tomatoes and peppers were evaluated, and three C2C models were developed along with an ANN base model to provide a basis for evaluation. The models were created using a dataset containing hourly watering data along with climatic and temporal data for the period between June 2015 and August 2016. The three C2C architectures used were linear regression (LR), quotient method (QM), and feed-forward neural network (FFNN), compared with the (ANN) model, which is a feed-forward neural network with extra inputs (FFNN-EI). Each model was evaluated using the root mean squared error (RMSE) and the normalized root mean squared error (NRMSE). The results show that all C2C methods have higher RMSE and NRMSE than that of the base model, with an average RMSE increase of 12% for peppers and 29% for tomatoes.

Keywords: crop water demand; forecast; crop to crop modeling; ANN modeling; water utility demand prediction

1. Introduction

Development of water demand forecasting models can be time consuming and costly. For a water utility located in Essex County, Ontario, Canada, forecasting commercial greenhouse water demand has become a critical aspect of day-to-day operations. This is due to the fact that almost 80% of the utility water demand is attributed to commercial greenhouses. Current methods used in this region do not utilize current forecasting techniques [1], but instead rely on the greenhouse operators themselves to submit water requirements when a facility is being constructed. These water demands received from the greenhouse operators are fixed and are also estimated demands from when the greenhouse was built. Technology and growing practices are constantly changing, and water demand from when the greenhouse may have been constructed will vary greatly from current demand. This lack of proper forecasting technique poses a serious challenge on the water utility distribution system. In other words, without proper forecasting, the system has to be significantly oversized to avoid demand uncertainties. As shown in the United States Environmental Protection Agency study [2], the cost of distribution accounts for 80% of water utility energy expenses, which emphasizes how crucial proper forecasts can be. These forecasts not only affect day-to-day operations, but can also have an impact on future infrastructure projects, with incorrect forecasts leading to unnecessary network or plant upgrades. Another aspect that forecasting can impact is future development of the region. This is caused by inflated demand profiles assigned to consumers, which falsely indicates that the water utility is either at or near capacity, when in reality it is not. This could cause the utility to deny water permits for future developments, which will reduce income. All of these issues caused by incorrect forecasting can negatively impact the water utility's bottom line.

In this region, greenhouses grow a variety of crops, and with each crop comes a unique demand profile. Figure 1 shows the crop breakdown by area, with the water utility servicing over 720 hectares of greenhouse operations. It can be recognized that more than half of the total area (57%) is used for growing tomatoes. Examining the remaining crops shows that tomatoes, peppers, and cucumbers account for almost 90% of the water demand, with the remainder being used by other crops such as flowers, plugs, and various other specialty products. Current estimates for water usage are shown in Table 1. It can be observed that there is a large variation in water usage estimates between greenhouse operations, with the largest coming from flower growers. This variation in water requirements between crops makes it difficult to develop one forecasting model to reliably predict the industry demand. This research is conducted to evaluate the suitability of a simplified forecasting method that would accompany a more complex method involving external determinates of demand such as neural network models. This is done by evaluating crop watering data, in this case tomato and pepper crops, and looking for similarities in watering strategies that would allow the water demand of one crop (e.g., peppers) to be predicted using only another crop's usage data (e.g., tomatoes). This crop-to-crop (C2C) forecasting method would utilize data generated by a much more complex and specialized demand-forecasting model that would be developed and optimized for one crop, and in this case it should be developed for the tomato crop as it accounts for almost 60% of the land use. The simplified forecasting method would save on expenses related to developing and maintaining several more complex models and would also address the concerns of [1] regarding model simplicity and practicality. The two crops chosen for this study were tomatoes and peppers, which account for 69% of total greenhouse land use. These two crops were also chosen due to their similar range in current demand estimates as seen in Table 1, it follows then that they would potentially have the highest probability of similar demand patterns.

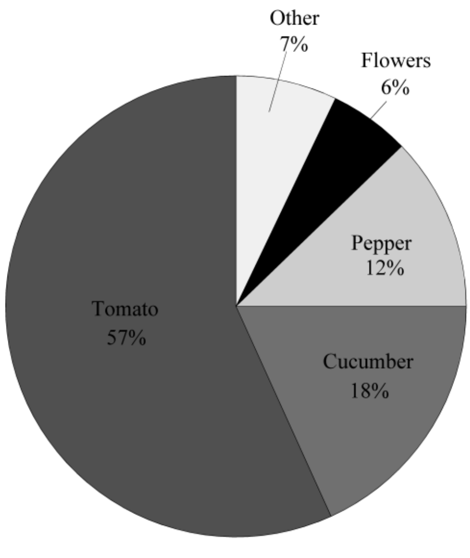

Figure 1. Crop breakdown by area.

Table 1. Current crop water demand estimates.

Crop	Water Demand Estimates (L/Plant/Day)
Tomato	2.7–4.5
Pepper	2.7–4.5
Cucumber	3.6–7.3
Flower	1.0–4.5

Current Literature

The content of this paper can be classified as crop water demand forecasting. There have been many studies focused on crop watering, of those few focus on greenhouse usage. The majority of crop irrigation studies focus on water-saving strategies and determining evapotranspiration rates [3–8] to optimize watering schemes, but do not address forecasting actual watering practices. These studies provide useful information for the greenhouse operators on how they can improve their watering schemes and reduce water usage, but do not address the concerns of the water utility. There is no shortage of water demand forecasting studies, but none focus on greenhouse water usage. Many studies focus on either residential water demand forecasting [9–11] or on entire service areas [12–14]. These studies are important and provide useful insights into modeling areas where water usage can be attributed largely to human water consumption. These studies do not address agricultural water demand including commercial scale greenhouse operations.

2. Materials and Methods

2.1. Data

The data used in this study was obtained from a local greenhouse operation. The data represents 2.43 ha of tomatoes and 3.64 ha of bell peppers and contains hourly water usage along with temporal and climactic data from 2 July 2015 through 8 August 2016 and contains 10,273 data points. Plant densities vary by crop type, and for this dataset the densities for tomatoes and peppers were approximately 28,800 and 33,500 plants per hectare respectively. The greenhouse was heated in the colder winter months and was naturally ventilated. The tomato crop was grown in a glass greenhouse and the pepper crop was grown in a polyethylene greenhouse. With current usage estimates listed in Table 1, examination of the dataset shows a maximum/average usage of 3.5/1.2 L and 4.3/1.7 L for pepper and tomato crops respectively. The dataset represents total water being delivered to the plants and does not differentiate alternate supply sources such as pond, well, or reuse water.

2.2. Preliminary Data Analysis

To determine the suitability of C2C forecasting, some preliminary data analysis must be performed. To compare the two crops, the data must be in the same units. As shown in the previous section, the dataset represented 3.64 ha of peppers and 2.43 ha of tomatoes, and simply using units of per hectare will not suffice, as both crops have different plant densities which would skew the results. To provide an unbiased assessment of water requirements, units of water consumption per plant was used. This allowed for a proper comparison of water requirements, and also provided a base demand that was scalable, and would be useful to the water utility as crop densities may vary from operation to operation. Figure 2 shows per plant tomato versus pepper water usage for the entire dataset. This figure shows a strong linear correlation (Pearson product-moment correlation coefficient of 0.95) between the water usage of both crops. There are outliers to this trend that can be seen along the x- and y-axis showing periods of zero watering of one crop and non-zero watering of the other, which may be attributed to non-plant watering events such as cleanouts at the end of the growth cycle. Overall this trend strengthened the use of C2C forecasting as it showed that as tomato watering increases, so does pepper watering.

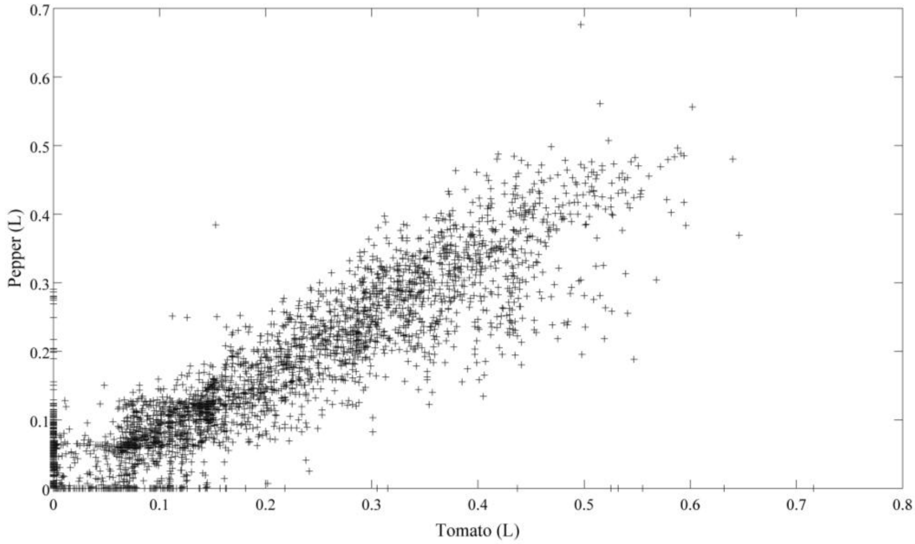

Figure 2. Per plant water usage.

Comparing water volume is not the only aspect that will impact the effectiveness of C2C forecasting. Time of use (time of day and time of year) could arguably have more impact on C2C forecasting than volume as C2C relies strictly on one crops time-series watering data to prepare forecasts. If the watering times are not similar C2C cannot be effective. To determine the similarities in daily watering times, Figure 3 was developed showing total water usage for each hour encompassing the entire dataset. Figure 3 shows that the vast majority of watering occurred between hours 9 and 19. This pattern held true for both crops, showing sharp increasing water usage between hours 9 and 10, with more incremental increase up until hour 14 where water usage begins to decline. It can be noticed that during hour 11, the water usage for tomatoes increased substantially to its maximum level and reduced during hour 12, after which it increased until hour 14 but did not reach the same usage level that was attained during hour 11. This deviation could have an impact on C2C forecasts, but since it only occurred during one hour with the remaining hours following the same pattern, the effects should be minimized.

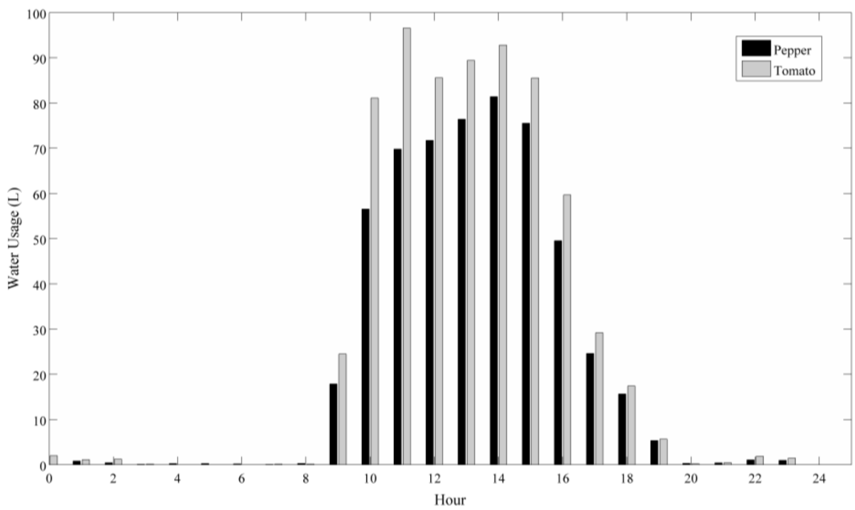

Figure 3. Total hourly water consumption by crop.

To examine both crops seasonal water usage trends Figure 4 was developed. This figure shows the monthly water usage for both crops for the entire dataset. Examination of Figure 4 shows a distinct pattern in water usage, starting with low use at the beginning of the year with a steady increase until usage peaked at month 7 (July), after which there was a sharp decline in usage for the remainder of the year. This pattern was consistent for both crops, with tomatoes having a larger magnitude. The most noticeable feature in this yearly pattern is the sharp increase in water usage between months 5 and 6 (May and June), and the sharp decrease in usage between months 7 and 8 (July and August), which was consistent for both crops. These increases in crop watering during these summer months (June to August) could be attributed to the higher levels of solar radiation and elevated outdoor temperatures which have been shown to have a major impact on pepper water consumption [15] along with the United Nations study [16] which show that solar radiation has the largest effect on evapotranspiration rates of various crops.

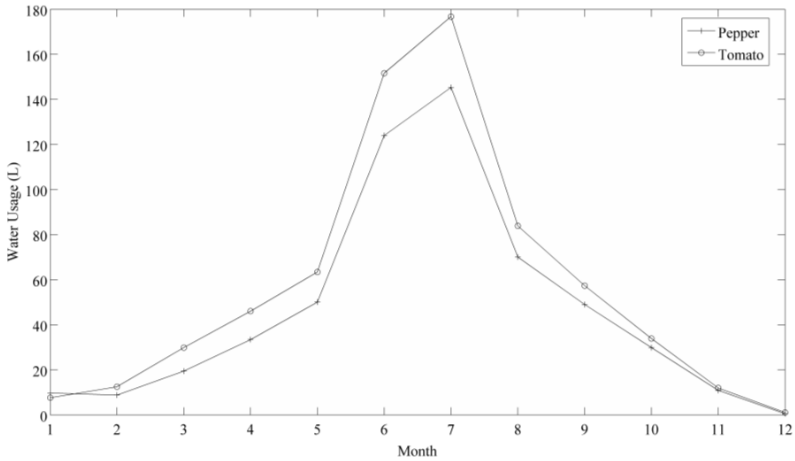

Figure 4. Monthly crop water usage.

The water loss from the land to the atmosphere due to Evapotranspiration (ET) can be divided into two losses, evaporation and transpiration. Evaporation is where water in liquid state converted to vapor from surfaces such as pavement, soil, etc. Transpiration is basically evaporation of water from plant leaves and it is dominant in greenhouses. However, the effect of evaporation cannot be ignored, "Evaporation and transpiration occur simultaneously and both depend on solar radiation, air temperature, relative humidity (i.e., vapor pressure deficit) and wind speed" [17].

Evapotranspiration is also a function of other factors, such as soil type, crop type, plant size, leaf texture, etc. For example, when the crop is small, water is predominately lost by soil evaporation, but once the crop is well developed and completely covers the soil, transpiration becomes the main process [18].

Evapotranspiration (ET) rates can be estimated using three approaches, ET measurement, ET estimated from pan evaporation, and ET computed from meteorological data. ET measurement uses methods that are often expensive and demanding in terms of accuracy of measurements and can only be fully exploited by well-trained research personnel. ET estimation from pan evaporation can provide an index of the process generally but does not account for some significant losses [18]. The third approach is collecting meteorological data to estimate the reference evapotranspiration (ETo), which can be achieved empirically using various methods, some of which is developed from the well-known Penman equation [19]. Most of the later equations used to estimate ETo were derived from the Penman equation; the developed equations used the reference evapotranspiration term (ETo) to estimate ET of a hypothetical reference crop using climatic data "independent of the crop type, crop development and management practices" [17].

After estimating ETo, ET can be calculated by multiplying ETo by a crop specific coefficient (kc). For the evapotranspiration inside the greenhouse [20] observed that the reference Evapotranspiration (ETo) inside greenhouses was always lower, with a range of 45–77% of that verified outside [21].

Although it is important to estimate ET for the plant water needs, the greenhouse operator designs water schedules are based on past crop yields, and empirical estimations are based on climate, and in some cases intuition. Because of this and the fact that leaf area can differ significantly from year to year makes it nearly impossible for the water utility to predict accurately. As a result, for the proposed model in this study, water usage per plant was used because regardless of each individual plant's needs, all plants received the same amount of water at the same times.

Another aspect to consider when comparing crop water usage is fruit mass, or yield. The assumption could be made that the greater the mass of fruit from a plant, the greater the water requirement. For this dataset, each tomato plant yielded on average 24 kg compared to 9 kg for a pepper plant. This large difference in fruit mass points to a substantial increase in water consumption of a tomato plant compared to a pepper plant, but examination of total water usage per plant for the entire dataset showed different results. The total water used per tomato plant was 675.9 L and 550.9 L for a pepper plant, which showed an increase in usage for tomatoes, but nowhere near the expected values, considering the difference in yield. This shows that water usage and yield are not linearly related and expected yield cannot be used as a direct indicator of water usage.

Sensitivity Analysis

In order to determine if one dataset can be predicted using the other, the factors influencing water demand of both crop types must be compared. The extended Fourier Amplitude Sensitivity Test (eFAST) was used by [15] to determine the driving factors in bell pepper water usage. This study carried out the same procedure which was developed by [22]. Using the neural network toolbox in MATLAB (MathWorks, Inc., Natick, MA, USA), a two-layer feed-forward neural network (FFNN) with eight hidden layers was developed to simulate the watering system. This model was trained using Levenberg-Marquardt back propagation, and was trained, tested, and validated using 75%, 10%, and 15% of the data respectively. The model performance was evaluated using the root-mean squared error (RMSE), which was 0.042 L.

To perform the eFAST procedure, SimLab sensitivity analysis software (InstallShield Software Corporation, Schaumburg, IL, USA) was used. SimLab generates a unique set of inputs based on the range of data that exists in the training dataset. The sample size used was 1480 based on the recommendation of [23]. These inputs are then fed through the FFNN and the results are recorded. These results are then introduced back into SimLab where it carries out the eFAST procedure and produces a set of sensitivities for each input factor. Since SimLab uses eFAST there are two sensitivities for each input, the first (S_i) and total indices (S_{Ti}). The first indices report the overall sensitivity of each input while the total indices accounts for interaction between the input factors. The results of the eFAST are displayed in Table 2. It can be seen that the results of the sensitivity analysis for the tomato crop are similar to the findings of [15] show that the top four most influential input factors for bell peppers were: Time, Solar Radiation, Outdoor Temperature, and Cumulative Solar Radiation. This shows that both pepper and tomato watering schemes are based on the same set of indicators and that combined with the findings of the previous sections show that it is possible to predict either crops water use if the other crops water use is known.

Table 2. First and Total Indices for all input factors in order of importance.

Input Factor	First Order Indices (S_i)		Total Indices (S_{Ti})	
	Value	**Rank**	**Value**	**Rank**
Time	0.396	1	0.622	1
Solar Radiation	0.249	2	0.368	2
Outdoor Temperature	0.038	3	0.174	3
Cumulative Solar Radiation	0.016	4	0.106	4
Greenhouse Relative Humidity	0.012	5	0.097	5
Month	0.009	6	0.096	6
Wind	0.006	7	0.065	7
Greenhouse Temperature	0.003	8	0.060	8
Σ	0.729		1.588	

2.3. Forecasting Methodology

The focus of this study is to determine if there is a way to predict greenhouse water usage without the use of complex and costly modeling techniques. C2C methods save time and cost in predicting watering patterns for one crop when knowing the watering patterns of other crops. This was achieved in this study by showing that instead of using individual studies to model crops individually, a well modeled crop can be used to predict other crops that show similarities. The methods used to achieve this goal are linear regression (LR) and a method titled the quotient method (QM), which will be discussed later in the section. The rationale behind choosing these two methods is that they will each produce either a simple equation or multiplier that can be directly used to forecast. These methods do not require the use of further modeling software to forecast after the initial values are determined. This is stated knowing that all models should be re-evaluated over time to ensure their accuracy and relevance. Since two crops are being used to evaluate these methods, two models will be created for each. One model will forecast tomato water usage using pepper data, and the other will forecast pepper usage using tomato data. Since the two model architectures being examined are on the elementary side, there must be some measure or comparison to a more complex and more accurate model. Including artificial neural network models (ANN) will provide a baseline for which the errors of the C2C models can be compared. The use of ANN models undermines the practicality aspect of this study, so it is being used as a comparative value to indicate the potential "best-case" performance. Each model will be developed with a hold back data set of one day and have a forecast horizon of one day, with forecasts occurring hourly.

To provide a measure of model performance for these methods, the root-mean-squared error (RMSE) and the normalized root-mean-squared error (NRMSE) will be used, and are represented by Equations (1) and (2), where n represents the number of data points, $\hat{Y}_i$ represents the forecasted water usage, and Y_i represents the actual data. The reasoning for using RMSE is due to the unique dataset, which contains numerous zero water usage data points, which can cause serious error when using other error measures such as mean absolute percent error (MAPE) which uses division by the observed value, which in this case would be zero, RMSE was also chosen based on the recommendations of Donkor et al. (2014) [1]. The inclusion of NRMSE is to reconcile the error into terms that can be compared when evaluating data sets with different ranges and means.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i^{\,2} - Y_i^2\right)} \tag{1}$$

$$\text{NRMSE} = \frac{\text{RMSE}}{\overline{Y_i}} \tag{2}$$

2.3.1. Linear Regression

Because of the seemingly linear relationship shown in Figure 2 between both crops' watering schemes, the use of a linear regression model was appropriate. In a typical water demand forecasting situation, the use of a single linear regression model is unwarranted due to its inability to handle multiple indicators of water demand usually present in a complex highly non-linear relationship between consumers and various socio-economic and climactic factors [24]. In this case the model involved two variables, which were treated as both dependent and independent, and were shown to have a linear relationship. By determining a best-fit line for the data using the least squares method to develop the coefficients (β_0, β) for Equation (3), the linear regression model will provide demand forecasts solely based on the linear relationship between the two crops, where W_t and W_p represent the hourly tomato and pepper water usage. Both W_t and W_p will be switched in order to develop an equation for both crops.

$$W_t = \beta_0 \pm \beta W_p \tag{3}$$

2.3.2. Quotient Method

The quotient method is the least complex method used in this study. Examining the total water usage for each crop type for the entire dataset and then dividing the values will produce a quotient (q), or how many times more water is used by tomatoes compared to peppers (or vice versa), as seen in Equation (4) where W_p and W_t are the hourly water usage values for peppers and tomatoes per plant, respectively, and q_p representing the quotient for the pepper crop. Equation (4) can also be modified by switching tomato and pepper values to give the quotient for tomatoes (q_t). The quotients can then be used in Equation (5) to provide forecasts for each crop. This method may seem elementary, but the previous sections show that there are many similarities in the watering habits for both crop types, which may make this forecasting option viable due to the simplicity and the costs saved, over a more traditional method.

$$q_p = \frac{\Sigma W_p}{\Sigma W_t} \tag{4}$$

$$W_p = W_t \times q_p \tag{5}$$

2.3.3. Feed-Forward Neural Network

The inclusion of a feed-forward neural network (FFNN) model in this study was for comparison, as FFNN models have been shown to outperform traditional methods in water demand forecasting [12,25–27]. The use of the FFNN will provide a benchmark with which both previous methods can be compared. Two FFNN models were developed, one a C2C model, and the other using temporal and climactic data as indicators of demand as recommended by [15]. The second model was titled FFNN_EI for external inputs (EI). The FFNN, generated using the Neural Network toolbox in MATLAB, contained one sigmoid hidden layer and one linear output layer and was trained using Levenberg-Marquardt back propagation. The data was divided with 65%, 20%, and 15% of the data used for training, testing, and validation. The network contains one input and five hidden layers. This network was chosen based on the guidelines of [28,29]. Due to the nature of FFNN models, there is no universal development method that will provide optimal results. To address this issue, a FFNN was developed and retrained 10 times, with the iteration with the lowest error being chosen for final forecasting.

3. Results and Discussion

3.1. C2C Model Development Using Real Data

The results of the LR model are shown in Equations (6) and (7). These equations show that there is an increasing trend in water usage for both crops, which follows the findings of the previous sections.

$$W_t = 4.52 \times 10^{-3} + 1.14 \times W_p \tag{6}$$

$$W_p = 1.28 \times 10^{-6} + 0.82 \times W_t \tag{7}$$

The results of Equation (4), the QM, were 1.227 and 0.815 for W_t and W_p, respectively, and the results of Equations (6) and (7), the LR, were 0.0693 and 0.0568 for W_t and W_p, respectively. This means that excluding the holdback data of one day, q_p was equal to 1.50 and 1.22 for QM and LR, respectively. Tomatoes used on average roughly 1.35 times more water per plant than peppers. Table 3 displays the error results of each forecasting method for day-ahead forecasts. For both LR and QM methods the NRMSE was almost identical, at 44% and 45%, respectively. The performance of the FFNN was only marginally better, with a NRMSE of 38% when forecasting tomatoes and 41% when forecasting peppers. The FFNN_EI model forecasting tomatoes showed the best performance with a NRMSE of 29%, which was approximately 16% less than the LR and QM models, and approximately 11% less than the FFNN model. If the models were judged purely on error measures, the FFNN_EI model would be the best ideal choice. But the use of a FFNN model comes with added complexity and likely added costs associated with implementation and maintenance. The increase in performance, measured by the decrease in error between the FFNN_EI and the C2C methods was at minimum 9% and maximum 16% of the average water usage. This difference in performance would need to be evaluated by the water utility to determine if the cost increase is worth the increase in performance. One notable result was that there were very minor differences when forecasting tomatoes or peppers for all model architectures, which shows that there was no advantage to forecasting either crop. This was useful as the water utility could determine which crop dominated the demand mix, and develop a dedicated base model without concern that crop type would affect the performance of the C2C forecast. The assumption of a more complex model involving external temporal and climactic data was validated, as the FFNN_EI model showed the best performance in hourly day-ahead forecasts. To summarize, with the limitations of using C2C method for two crops that share high similarities in time and volume of watering, implementing such a simple and useful tool can benefit the water utility where costly and high-performance methods are not an option.

Table 3. Results in (%) of day-ahead forecasting methods using actual data.

Model	LR		QM		FFNN		FFNN_EI	
	RMSE	NRMSE	RMSE	NRMSE	RMSE	NRMSE	RMSE	NRMSE
Tomato with Pepper	2.9	44	3.0	45	2.5	38	1.9	29
Pepper with Tomato	2.4	45	2.4	45	2.2	41	1.6	30

LR: Linear Regression; QM: Quotient Method; FFNN: Feed-Forward Neural Network; FFNN_EI: Feed-Forward Neural Network_External Inputs; RMSE: Root Mean Square Error; NRMSE: Normalized Root Mean Square Error.

The results of day-ahead forecasts for each model are displayed in Figures 5 and 6. The daily pattern can be observed, with all watering taking place between hour 8 and hour 20. All forecast methods were able to determine these start and stop times, with no spike in usage occurring before or after the actual start and stop hours. One issue that was noticed was the inability of the LR model to forecast zero usage points. As shown in Equations (6) and (7), the LR model produced a minimum of 4.52×10^{-3} L when forecasting tomato usage and 1.28×10^{-6} L when forecasting pepper usage. The LR model was not the only model to have issues forecasting zero usage data; the FFNN and FFNN_EI models also did not forecast zero, but for each instance, forecasted values ranged from 3.81×10^{-3} L to 1.01×10^{-4} L. This issue could cause difficulty when the forecast is extrapolated over the millions of plants in the region. The QM however, was able to forecast the zero usage periods as they occurred during the same times for both crops. Examining both figures shows how similar the two crops watering schemes are. The differences in watering times become apparent in the forecasts during hour 9 and 11, which show peaks occurring at opposite times for each crop. This is where the majority of the difference in forecasts occurs. The forecast for tomatoes followed the same pattern as the actual usage data for peppers, except at a higher magnitude, and the same applied to the pepper forecasts, as they resembled actual tomato usage. This applies to all C2C models, but not the FFNN_EI model. One notable feature in Figure 6 is how close each of the forecasting methods is compared to those in Figure 5. Examining the results of the FFNN_EI model showed better adherence to the actual usage data, particularly during the peaks in usage occurring at hour 11 for tomatoes and hour 10 for peppers, where the other models struggle.

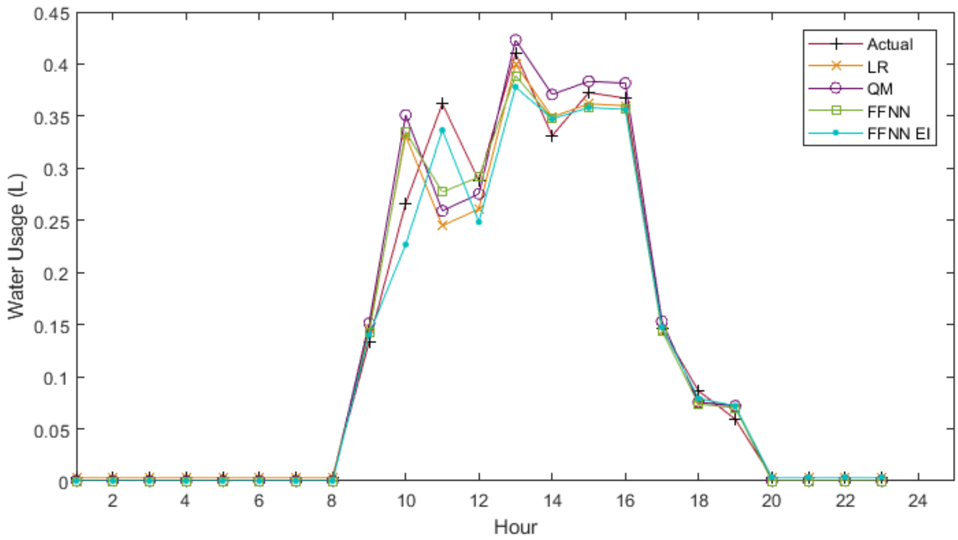

Figure 5. Day-ahead forecasts for tomatoes.

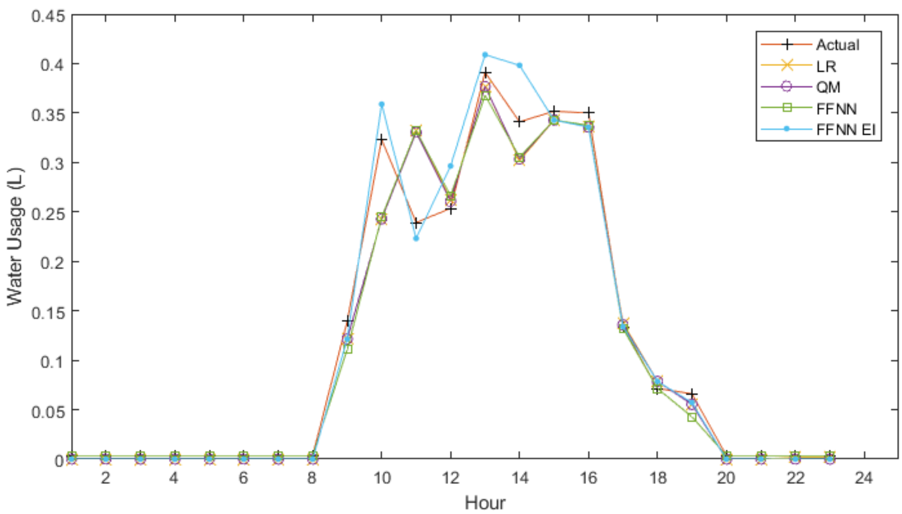

Figure 6. Day-ahead forecasts for peppers.

3.2. C2C Model Development Using Base Model Data

Table 4 shows the results of C2C model development using the FFNN_EI base model as input criteria. Since modeled data was used as input data for the C2C models, a higher error was expected. This is shown in Table 4, with all C2C models producing a higher error when developed using base model data. C2C forecasts for tomatoes saw an average increase in RMSE of 29%, with the smallest increase occurring in the LR model (25%). Forecasts for peppers showed a much smaller increase in RMSE; with an average increase of 12%, with the smallest increase occurring with the LR model (11%). This disparity in performance is likely caused by the amplification of larger errors in the base model forecast. Examination of Table 3 shows that the base FFNN_EI model has similar RMSE and NRMSE (29% and 30%) for both crops, but what the error measure does not account for is the occurrences of large errors. It can be seen in Figure 6 that the FFNN_EI model had generally less errors, but the errors that occurred were larger than those in Figure 5. These large error occurrences appear to have been amplified when the FFNN_EI results were used to generate C2C forecasts. This shows the impact that base model development can have on future forecasts developed using the base model.

Table 4. Results in (%) of day-ahead forecasting methods using base model.

Model	LR		QM		FFNN	
	RMSE	NRMSE	RMSE	NRMSE	RMSE	NRMSE
Tomato with Pepper	3.9	59	4.3	64	3.8	51
Pepper with Tomato	2.7	51	2.8	52	2.5	47

LR: Linear Regression; QM: Quotient Method; FFNN: Feed-Forward Neural Network; FFNN_EI: Feed-Forward Neural Network_External Inputs; RMSE: Root Mean Square Error; NRMSE: Normalized Root Mean Square Error.

As shown by the results of the previous section, examination of only error measures was not sufficient to make general claims about model performance. Figure 7 shows the day-ahead forecast for tomatoes using the base model (FFNN_EI) forecast data. It can be observed that the LR and QM models followed the same pattern that exists in the base model for peppers (Figure 6). The FFNN model however, did not follow the same pattern over the entire forecast. The FFNN model appeared to smooth out the forecast between the hours of 12 and 17, where some of the larger base model error existed. The same held true for the forecast model for peppers, with the exception of the smoothing of the FFNN model, as shown in Figure 8. When examining both figures it can be seen that the forecast for peppers (Figure 8) showed better adherence to the actual demand pattern, than that of the tomato forecast. Both forecasts suffered from the same issues that arose in the previous section, where peppers and tomatoes had slightly different daily demand patterns, which was shown by the lack of forecast accuracy between the hours 9 and 12. Overall, forecasts for both tomatoes and peppers suffered from the same issues as the forecasts developed with real watering data. The main difference was that the errors that existed in the base model were amplified, stressing the importance of proper base model development.

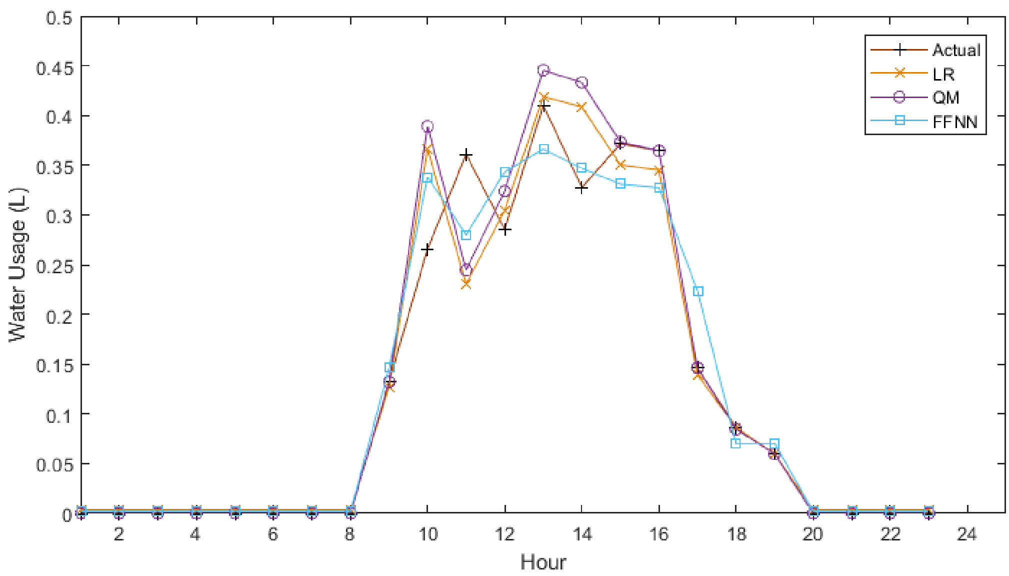

Figure 7. Day-ahead forecast for tomatoes using base model data.

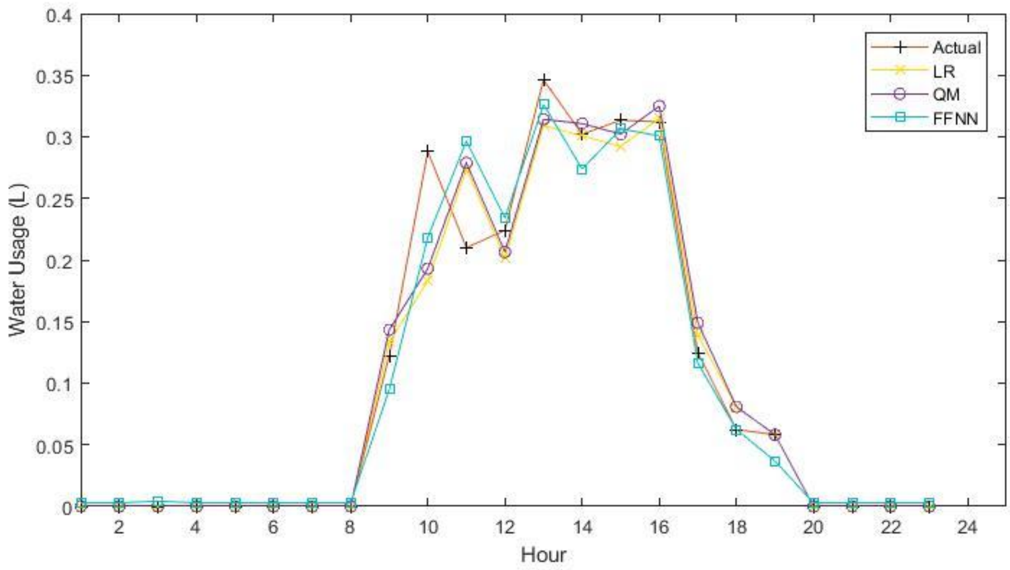

Figure 8. Day-ahead forecast for peppers using base model data.

4. Conclusions

The main goal of this study was to determine the effectiveness of using simplified C2C forecasting methods to determine greenhouse crop watering needs. The results show that both LR and QM models have similar performance having NRMSE's of 44% to 45%, while both FFNN model and FFNN-EI model had better performance. One issue with using C2C forecasting is the apparent ghosting of the independent crop on the forecast. This is evident during hour 10 for peppers and 11 for tomatoes, where only the FFNN_EI model was able to forecast the correct time of this peak for both crops. This shows a weakness in the C2C method, but it still provides better forecasts than the current fixed demand model. This study also shows that the FFNN models perform better than the more traditional LR method and the simplified QM. The use of a FFNN model in C2C forecasting is shown to be ineffective as the FFNN model using external inputs (FFNN_EI) showed better performance, NRMSE of 38% compared to 29%, while maintaining the same level of complexity. As expected, the proposed base model FFNN_EI showed the best performance with a NRMSE of 29% and 30% for tomatoes and peppers, respectively.

C2C forecasting using base model data (FFNN_EI) showed larger RMSE and NRMSE for both forecasts, with the tomato forecast producing and average increase in RMSE of 25% and peppers 12%. This disparity in error can be attributed to the magnitude of error present in the base model. Although both base models had similar NRMSE (29% and 30%), the model for peppers had a number of occurrences of large errors, while the tomato model was more consistent. These large error occurrences appear to be amplified when the base model is used to generate future forecasts, as the increase in error between crop forecasts is 13%. This error increase shows that base model development is crucial for the effectiveness of C2C forecasting. As for performance, the FFNN model showed the lowest NRMSE for both crops, with the lowest occurring with the pepper crop (47%). The increase in RMSE between the FFNN model and the LR and QM models ranges from 8% to 12% for peppers and 3% and 13% for tomatoes, respectively.

Since greenhouses grow a variety of crops other than tomatoes and peppers, it would be advisable to perform a similar analysis on other crops to determine similarities in watering habits and evaluate the effectiveness of C2C forecasting on a broader scale. The performance of all C2C model architectures remained relatively consistent for both tomato and pepper crops (NRMSE of 44% to 45%) when using actual watering data as inputs. This shows that there is no benefit to using a particular crop for the base model for which C2C forecasts are generated and allows the water utility to develop a base model for the most prevalent crop. By developing a base model for the largest crop in the region the water utility will reduce forecast error since the base model has better performance than C2C methods.

The conclusion of this study shows that C2C forecasting can be achieved, although with higher error than more complex model architectures involving external inputs. These performance increases should be evaluated by the water utility to determine if the costs associated with implementing a FFNN model justify the increase in performance. The use of C2C forecasts will provide an improvement to the current fixed demand methods used in the study region, but will require the development and implementation of a base model in order to function.

Acknowledgments: The Authors would like to thank Gene Ingratta of Allegro Acres Inc. (Windsor-ON, Canada), Rodney Bouchard of Union Water Supply System, the Ontario Clean Water Agency, and the Ontario Greenhouse Vegetable Growers for their continued support and guidance.

Author Contributions: D.C.J.R conceived the concept for the manuscript and performed the majority of initial draft development. This proceeded under the guidance and supervision of R.C. and D.S.-K.T. R.C. provided expertise in greenhouse and water utility operation—directing data gathering. D.S.-K.T. provided expertise in data illustration, interpretation, and model validation practices. M.H.B developed subsequent and ultimately final drafts of the manuscript that brought a more complete consideration of greenhouse growing conditions. The interactions of the listed authors represented a true collaborative effort in this publication.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Donkor, E.A.; Mazzuchi, T.; Soyer, R.; Roberson, A. Urban Water Demand Forecasting: Review of Methods and Models. *J. Water Resour. Plan. Manag.* **2014**, *140*, 146–159. [CrossRef]
2. United States Environmental Protection Agency. Strategies for Saving Energy at Public Water Systems. 2013. pp. 1–16. Available online: http://water.epa.gov/type/drink/pws/smallsystems/upload/epa816f13004.pdf (accessed on 15 January 2015).
3. Bernier, M.-H.; Madramootoo, C.A.; Mehdi, B.B.; Gollamudi, A. Assessing On-Farm Irrigation Water Use Efficiency in Southern Ontario. *Can. Water Resour. J.* **2010**, *35*, 115–130. [CrossRef]
4. Daghighi, A.; Nahvi, A.; Kim, U. Optimal Cultivation Pattern to Increase Revenue and Reduce Water Use: Application of Linear Programming to Arjan Plain in Fars Province. *Agriculture* **2017**, *7*, 73. [CrossRef]
5. Capraro, F.; Patino, D.; Tosetti, S.; Schugurensky, C. Neural network-based irrigation control for precision agriculture. In Proceedings of the 2008 IEEE International Conference on Networking, Sensing and Control, ICNSC, Sanya, China, 6–8 April 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 357–362.
6. Orgaz, F.; Fernández, M.D.; Bonachela, S.; Gallardo, M.; Fereres, E. Evapotranspiration of horticultural crops in an unheated plastic greenhouse. *Agric. Water Manag.* **2005**, *72*, 81–96. [CrossRef]
7. Ma, L.; He, C.; Wang, Z. The research for the greenhouse water evaporation based on the environmental factors. *Adv. J. Food Sci. Technol.* **2013**, *5*, 1049–1054. [CrossRef]
8. Thompson, R.B.; Gallardo, M.; Valdez, L.C.; Fernández, M.D. Using plant water status to define threshold values for irrigation management of vegetable crops using soil moisture sensors. *Agric. Water Manag.* **2007**, *88*, 147–158. [CrossRef]
9. Polebitski, A.S.; Palmer, R.N. Seasonal Residential Water Demand Forecasting for Census Tracts. *J. Water Resour. Plan. Manag.* **2010**, *136*, 27–36. [CrossRef]
10. Dong, Y.; Zhou, W. Urban residential water demand forecasting in Xi'an based on RBF model. In Proceedings of the 2009 International Conference on Energy and Environment Technology, Guilin, China, 16–18 October 2009; pp. 901–904.
11. Bennett, C.; Stewart, R.A.; Beal, C.D. ANN-based residential water end-use demand forecasting model. *Expert Syst. Appl.* **2013**, *40*, 1014–1023. [CrossRef]
12. Ghiassi, M.; Zimbra, D.K.; Saidane, H. Urban Water Demand Forecasting with a Dynamic Artificial Neural Network Model. *J. Water Resour. Plan. Manag.* **2008**, *134*, 138–146. [CrossRef]
13. Burney, N.; Mukhopadhyay, A.; Al-Mussallam, N.; Akber, A.; Al-Awadi, E. Forecasting of Freshwater demand in Kuwait. *Arab. J. Sci. Eng.* **2001**, *26*, 99–113.
14. Bougadis, J.; Adamowski, K.; Diduch, R. Short-term municipal water demand forecasting. *Hydrol. Process.* **2005**, *19*, 137–148. [CrossRef]

15. Rice, D.C.J.; Carriveau, R.; Ting, D.S.-K. Commercial greenhouse water demand sensitivity analysis: Single crop case study. *Water Sci. Technol. Water Supply* **2016**, *16*, 1185–1197. [CrossRef]
16. Doorenbos, J.; Pruitt, W.O. *Guidelines for Predicting Crop Water Requirements*; Food and Agriculture Organization of the United Nations: Rome, Italy, 1977; p. 154.
17. Zotarelli, L.; Dukes, M.D.; Romero, C.C.; Migliaccio, K.W.; Morgan, K.T. *Step by Step Calculation of the Penman-Monteith Evapotranspiration (FAO-56 Method)*; Document AE459; Agricultural and Biological Engineering Department, UF/IFAS Extension: Gainesville, FL, USA, 2010.
18. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *"Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements." Irrigation and Drainage Paper No. 56*; Food and Agriculture Organization of the United Nations: Rome, Italy, 1998.
19. Penman, H.L. Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **1948**, *193*, 120–145. [CrossRef] [PubMed]
20. Farias, J.R.B.; Bergamaschi, H.; Martins, S.R. Evapotranspiration inside plastic greenhouses. *Braz. J. Agrometeorol.* **1994**, *2*, 17–22.
21. Fernandes, C.; Corá, J.E.; de Araújo, J.A.C. Reference evapotranspiration estimation inside greenhouses. *Sci. Agricola* **2003**, *60*, 591–594. [CrossRef]
22. Saltelli, A.; Chan, K.; Scott, E.M. *Sensitivity Analysis*; John Wiley & Sons, Ltd.: Chichester, UK, 2008.
23. Saltelli, A.; Tarantola, S.; Chan, P.-S. A Quantitative Model-Independant Method for Global Sensitivity of Model Output Analysis. *Technometrics* **1999**, *41*, 39–56. [CrossRef]
24. Herrera, M.; Torgo, L.; Izquierdo, J.; Pérez-García, R. Predictive models for forecasting hourly urban water demand. *J. Hydrol.* **2010**, *387*, 141–150. [CrossRef]
25. Jain, A.; Ormsbee, L.E. Short-term Water demand forecast modeling techniques—Conventional Methods versus AI. *J. Am. Water Works Assoc.* **2002**, *94*, 64–72.
26. Adamowski, J.; Karapataki, C. Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: Evaluation of different ANN learning algorithms. *J. Hydrol. Eng.* **2010**, *15*, 729–743. [CrossRef]
27. Mitrea, C.A.; Lee, C.K.M.; Wu, Z. A Comparison between Neural Networks and Traditional Forecasting Methods: A Case Study. *Int. J. Eng. Bus. Manag.* **2009**, *1*, 19–24. [CrossRef]
28. Hunter, D.; Yu, H.; Pukish, M.S., III; Kolbusz, J.; Wilamowski, B.M. Selection of Proper Neural Network Sizes and Architectures—A Comparative Study. *IEEE Trans. Ind. Inform.* **2012**, *8*, 228–240. [CrossRef]
29. Jinchuan, K.; Xinzhe, L. Empirical analysis of optimal hidden neurons in neural network modeling for stock prediction. In Proceedings of the 2008 Pacific-Asia Workshop on Computational Intelligence and Industrial Application, PACIIA 2008, Wuhan, China, 19–20 December 2008; Institute of Electrical and Electronics Engineers Computer Society: Washington, DC, USA, 2008; pp. 828–832.