Evaluating the Performance of Self-Organizing Maps to Estimate Well-Watered Canopy Temperature for Calculating Crop Water Stress Index in Indian Mustard (Brassica Juncea)

Navsal Kumar[1*], Vijay Shankar[2], Rabee Rustum[3], Adebayo J. Adeloye[4]

*[1] Ph.D. Research Scholar, Civil Engineering Department, National Institute of Technology Hamirpur, Himachal Pradesh – 177005, India; Email - navsal.happy@gmail.com (*Corresponding Author)*
[2] Associate Professor, Civil Engineering Department, National Institute of Technology Hamirpur, Himachal Pradesh – 177005, India; Email - vsdogra12@gmail.com
[3] Associate Professor, School of Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, Dubai - 294345, UAE; Email - r.rustum@hw.ac.uk
[4] Professor, School of Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, Edinburgh - EH14 4AS, UK; Email - a.j.adeloye@hw.ac.uk

Cite this article as:

Kumar, N., Shankar, V., Rustum, R., & Adeloye, A. J. (2021) Evaluating the Performance of Self-Organizing Maps to Estimate Well-Watered Canopy Temperature for Calculating Crop Water Stress Index in Indian Mustard (Brassica Juncea). *ASCE Journal of Irrigation and Drainage Engineering. 147*(2): 04020040. *https://doi.org/10.1061/(ASCE)IR.1943-4774.0001526*

Abstract

Crop Water Stress Index (CWSI) is a reliable indicator of water status in plants and has been utilized for stress monitoring, yield prediction, and irrigation scheduling. Despite this, however, its use is limited because its estimation requires the baseline temperatures under similar environmental conditions, which can be problematic. In this study, field crop experiments were performed to monitor the canopy temperature of Indian mustard (*Brassica Juncea*) from crop development through harvest under different irrigation treatment levels during 2017 and 2018 growing seasons. Kohonen Self-Organizing Map (KSOM), feed-forward neural network (FFNN) and multiple linear regression (MLR) models were developed for estimating the well-watered canopy temperature ($T_{c\text{-}ww}$) using air temperature and relative humidity as input predictor variables. Comparisons were performed between model estimated and measured $T_{c\text{-}ww}$ values. The findings indicate that the KSOM-modelled values presented a better agreement with the measured values in comparison to MLR and FFNN based estimates, with R^2 values of 0.978, 0.924 and 0.923 for KSOM, MLR and FFNN, respectively during model validation. The dry canopy temperature was estimated to be air temperature plus 2 °C. The CWSI computed using KSOM based estimates of $T_{c\text{-}ww}$ was compared with the CWSI obtained from measured values of $T_{c\text{-}ww}$. The results suggest a significant potential of KSOM for reliable estimation of the $T_{c\text{-}ww}$ for calculating the CWSI that can be automated for developing precision irrigation systems.

Keywords: Neural computing; multiple linear regression; Unsupervised learning; Model performance; Plant water status.

1. Introduction

Indian mustard (*Brassica Juncea*) accounts for nearly 90% of the rapeseed mustard cultivated area of India (MoEFCC 2016). It is a widely grown crop and the most prominent winter oilseed crop primarily producing vegetable oil along with vegetable, spice and fodder (Shekhawat et al. 2012). Although Indian mustard has a reputation of being tolerant to water stress (Wright et al. 1996; Kumar et al. 2020), irrigation schedule significantly affects its yield (Boomiraj et al. 2010; Mishra et al. 2019). Previous studies have indicated that frequent irrigation significantly increases stover yield but hampers the fruiting (Singh and Singh 2019). Moreover, the seed yield decreases significantly during drought or water-stressed conditions (Singh et at. 2018; Rana et al. 2019). This necessitates a thorough understanding of plant water status and associated degree of water stress, crop and water use efficiency. Monitoring tools capable of providing precise information regarding the water status of crops would, therefore, be useful for efficient irrigation scheduling and management (Adeyemi et al. 2018).

Infrared thermometry-based measurements of canopy temperature (T_c) have been acknowledged as a non-destructive and reliable plant water status indicator (Osroosh et al. 2015; Ihuoma and Madramotoo 2017; Romero-Trigueros et al. 2019). The utility of T_c for determining the water status in plants is based on the effect of relative transpirational cooling (Ehrler 1973; Hou et al. 2019). Apart from its dependence on plant water status, T_c is also governed by the prevailing environmental conditions including air temperature, wind speed, humidity and solar radiation (Poirier-Pocovi and Bailey 2020). Thus, T_c must be normalized before its application to account for the prevailing environmental dynamics (Gerhards et al. 2019). The most common approach to normalize the T_c is to use the crop water stress index (CWSI), initially proposed by Jackson et al. (1981).

CWSI is a simple tool that quantifies the crop water status for scheduling irrigation in crops (King and Shellie 2016). It has been used for monitoring water status in plants, detecting onset of moisture stress, predicting yield and scheduling irrigation in different crops (Yuan et al. 2004; Gontia and Tiwari 2008; Yildirim et al. 2012; Akkuzu et al. 2013; Gonzalez-Dugo et al. 2014; Bellvert et al. 2016; Kumar et al. 2020b; Anda et al. 2020). The limits of CWSI are 0 and 1, with 0 indicating the well-watered or non-water stressed condition and 1 representing the non-transpiring or severely water-stressed condition. CWSI is basically defined as (Jackson et al. 1981),

$$CWSI = \frac{[(T_c - T_a) - (T_{c-ww} - T_a)]}{[(T_{c-dry} - T_a) - (T_{c-ww} - T_a)]} \tag{1}$$

Where, T_c is the actual canopy temperature (°C); T_a is the air temperature (°C); $T_{c\text{-}ww}$ is the canopy temperature of a plant transpiring at full potential when the soil water is adequate (°C); and $T_{c\text{-}dry}$ is the canopy temperature of a non-transpiring plant due to stomatal closure when the soil becomes dry (°C). The terms ($T_{c\text{-}ww}$ - T_a) and ($T_{c\text{-}dry}$ - T_a) represent the lower and upper baseline temperatures, respectively.

There are two versions of CWSI in the literature, theoretical CWSI, and empirical CWSI. The difference in the versions is how the upper and lower baseline temperatures are calculated. The theoretical approach, initially given by Jackson et al. (1981) is based on the energy balance model. The baseline temperatures are calculated using Equation 2 and 3, respectively.

$$(T_{c-ww} - T_a) = \frac{r_a R_n}{\rho\, c_\rho} \frac{\gamma\left(1 + (r_{c-ww}/r_a)\right)}{\Delta + \gamma\left(1 + (r_{c-ww}/r_a)\right)} - \frac{(e_s - e_A)}{\Delta + \gamma\left(1 + (r_{c-ww}/r_a)\right)} \tag{2}$$

$$\left(T_{c-dry} - T_a\right) = \frac{r_a R_n}{\rho\, c_\rho} \frac{\gamma\left(1 + \left(r_{c-dry}/r_a\right)\right)}{\Delta + \gamma\left(1 + \left(r_{c-dry}/r_a\right)\right)} - \frac{(e_s - e_A)}{\Delta + \gamma\left(1 + \left(r_{c-dry}/r_a\right)\right)} \tag{3}$$

Where, γ = psychrometric constant (kPa °C^{-1}); R_n = net radiation (W m^{-2}); $r_{c\text{-dry}}$ = crop canopy resistance under dry conditions (sm^{-1}); $r_{c\text{-ww}}$ = crop canopy resistance under well-watered conditions (sm^{-1}); r_a = aerodynamic resistance (sm^{-1}); ρ = mean air density at constant pressure (Kg m^{-3}); c_ρ = heat capacity of air (J Kg^{-1} °C^{-1}); e_s = saturated vapour pressure (kPa); e_a = actual vapour pressure (kPa); and Δ = slope of saturated vapour pressure (kPa °C^{-1}).

The empirical approach was introduced by Idso et al. (1981) and considers the experimental observations of the baseline canopy temperatures. The lower baseline is generally obtained through a linear regression between ($T_c - T_a$) and vapor pressure deficit for potentially transpiring or well-watered crops, however, direct observations of $T_{c\text{-ww}}$ provide more accurate estimates of CWSI (Yuan et al. 2004). Previous studies have shown that the upper baseline which indicates a non-transpiring crop is well represented by air temperature plus a constant value (King and Shellie 2018; Adeyemi et al. 2018).

As seen above, the theoretical approach involves numerous complex meteorological data to compute the CWSI baselines. Although, the model has been found to precisely assess the crop water stress (Yuan et al. 2004; Heydari et al. 2019), its application in commercial crop production is limited due to requirement of complex input model parameters, particularly crop canopy resistance, aerodynamic resistance, and net radiation values (Al-Faraj et al. 2001). The empirical approach is simple to use and gives a reliable indication of crop water stress. It has, however, been shown that the $T_{c\text{-ww}}$ depends on the crop growth and the agro-climate in which it is grown (Kumar et al. 2019). Further, direct measurements of $T_{c\text{-ww}}$ and $T_{c\text{-dry}}$ under similar environmental conditions as the T_c are practically unfeasible due to experimental constraints, as both involve field soil water that is either undesirable ($T_{c\text{-dry}}$) or unattainable ($T_{c\text{-ww}}$) in practical conditions (Kumar et al. 2020a).

Artificial reference surfaces for estimating the baseline temperatures have been developed and successfully used under similar environmental conditions (Agam et al. 2013). These include the use of well-watered and water-stressed plots, leaves sprayed with water and covered with petroleum jelly and the application of wet and dry filter papers (Meron et al. 2010; Alchanatis et al. 2010). However, they require extensive maintenance and intensive data acquisition, which limits their use in precision irrigation systems (Maes and Steppe 2012). Numerical estimation of the baseline temperatures through physical models has also been found to give reliable results. Jones (1999) used the leaf energy balance model to develop the predictive equations for the baseline temperatures. The numerical estimation of the baseline temperature eliminates the need for an artificial reference surface, but it involves measurements of the equation parameters, routine observation of which is not feasible owing to the expensive instrumentation and lack of technical know-how (Park 2018). Hence, estimation of the baseline temperature through parsimonious predictive models using limited climatic data will enhance the utilization of CWSI as a tool for scheduling irrigation and monitoring crop stress (Osroosh et al. 2016; Egea et al. 2017).

The application of multiple linear regression (MLR) using climatic data including wind speed, vapor pressure deficit (VPD), air temperature, and solar radiation has been found to improve $T_{c\text{-}ww}$ prediction for a soybean crop, with the correlation coefficients ranging between 0.69-0.84 (Payero and Irmak 2006). The value of $T_{c\text{-}dry}$ has been observed to be equal to the air temperature plus a constant temperature, which varies with the crop type (O'Shaughnessy et al. 2011). King and Shellie (2016) reported on the application of artificial neural networks (ANN) in improving the $T_{c\text{-}ww}$ prediction using wind speed, air temperature, VPD, and solar radiation as input data. Although the ANN and MLR approaches have been successful in modeling complex, unknown relationships to predict physical variables, their predictions are sensitive to the availability and quality of input data used in model development. In other words, missing values or outliers in the input data can infuse large errors in their predictions (Adeloye et al. 2012). Indeed, ANN has been observed to give unrealistic results when such a noise is present in the input data (Rustum 2009).

On the contrary, unsupervised neural networks, known as Kohonen Self-Organizing Maps (KSOM) (Kohonen 1990; Kohonen et al. 1996) have no specific input or output arguments. KSOM clusters a large dimensional data into a small dimensional map, thus making any inherent correlations between the variables much more visible (Kothari and Islam 1999). The clustering enables effective replacement of the missing values or outliers by their corresponding features in the map, thereby causing no hindrance to the predictions of the model. Due to its versatility, the KSOM has been widely used in hydrological modeling including evapotranspiration modeling (Adeloye et al. 2011), global water flows assessment (Clark et al. 2015), water quality modeling (Rustum and Adeloye 2007; Ramachandran et al. 2019), streamflow forecasting (Mwale et al. 2014), rainfall-runoff modeling (Adeloye and Rustum 2012), soil moisture (Riese and Keller 2018), irrigation management (Ohana-Levi et al. 2019) and groundwater studies (Chen et al. 2018).

To the best of our knowledge, a KSOM has never been used to predict the baseline temperature ($T_{c\text{-}ww}$) for calculating the CWSI. Let alone the KSOM, even the application of ANN in this field has been reported only by King and Shellie (2016). Hence, the study aims to investigate the performance of KSOM to estimate the $T_{c\text{-}ww}$ for CWSI determination. The specific objectives are to:

1. Develop and validate a KSOM model to estimate the $T_{c\text{-}ww}$ and compare its values with experimentally derived $T_{c\text{-}ww}$.
2. Evaluate the performance of the KSOM model with multiple linear regression and feed-forward neural network models developed for estimating $T_{c\text{-}ww}$.

3. Apply the KSOM estimated $T_{c\text{-}ww}$ for predicting the CWSI in Indian mustard.

2. Materials and Methods

2.1 Agricultural plot and experimental details

The study was carried out during the 2017 and 2018 growing seasons at the agricultural experimental station of the National Institute of Technology, Hamirpur, India (altitude: 900 m asl; longitude: 76° 31' 33"; latitude: 31° 42' 40"). Field crop experiments were performed on Indian mustard (*Brassica Juncea*) from September to December. The climate of the study area is humid sub-tropical with seasonal mean values of relative humidity, air temperature, solar radiation and wind speed of 74.2 %, 19.10 °C, 0.16 kW m^{-2}, and 1.8 m s^{-1} respectively. The average seasonal rainfall is 65 mm. The soil in the experimental station had uniform sandy loam texture (silt = 24%, sand = 55% and clay = 21%) up to 1.6 m depth. The permanent wilting point (PWP) and field capacity (FC) of the soil obtained using pressure plate apparatus were 0.07 cm^3 cm^{-3} and 0.22 cm^3 cm^{-3} respectively. The available soil water (ASW), defined as the difference between FC and PWP, was estimated to be 0.15 cm^3 cm^{-3}. This is a relatively low ASW which should accelerate the drying up of the soil and hence make the determination of the $T_{c\text{-}dry}$ much more rapid. For soils with more water retention capacity, the drying process will be much slower especially during wet periods.

The experimental layout was designed using the randomized complete block design (RCBD). The field was divided into eight treatment plots (T1 to T8) with three replications (R1 to R3). Figure 1 shows the layout of the experimental plot. Irrigation in each trial was identical and provided for the application of eight levels of treatments, one for each of the 2m × 2m sized plots. The plots were separated from each other by embedding asbestos sheets 2m deep to prevent the horizontal flow of soil water.

Figure 1

The irrigation treatments were based on a specific level of soil water depletion (SWD) of the ASW in the crop root zone. Treatment T8 was not provided with any supplemental irrigation (except for pre-sowing and one for the crop survival) during the entire crop season. Treatment T1 was provided with frequent irrigations to maintain the water content near the FC. Treatments T8 and T1 were deliberately kept dry and well-watered, to allow the estimation of $T_{c\text{-}dry}$ and $T_{c\text{-}ww}$, respectively. The maximum level of SWD allowed in the treatments T2, T3, T4, T5, T6, and T7 was 10%, 20%, 30%, 40%, 50% and 60% of ASW, respectively. The soil water was monitored daily using a capacitance probe (*Sentek Sensor Technologies, SA, Australia*), which recorded the volumetric water content (VWC) every 0.1 m interval up to 1.6 m depth. The percentage SWD of ASW in the effective root zone was estimated using the relation $\mathrm{SWD} = (\mathrm{FC} - \mathrm{VWC})/\mathrm{ASW}$. Water was supplied to respective plots with the help of a water hose (surface irrigation) in calculated amounts (water meter installed at the inlet). A tipping bucket rain gauge was used for recording the rainfall.

The field was prepared using tilling and harrowing operations. At the beginning of the crop period, farmyard green manure was applied in all the plots. The crops were suitably fertilized during the growth stages with 100:40:40 Nitrogen-Phosphorus-Potassium (NPK) fertilizers. The crops were adequately spaced through the thinning process at 15-20 days after sowing (DAS). Treatment plots consisted of approximately 60 plants with five rows having twelve plants per row. Table 1 presents the relevant crop details. The crop growth period was divided into 4 stages viz. vegetative (initial stage), flowering (crop development stage), pod formation and seed development (mid-season stage) and maturity and harvest (late-season stage) as given in FAO-56 (Allen et al. 1998).

Table 1

2.2 Canopy temperature and weather monitoring

A multi-meter weather monitoring and data logging system (*METER Group Inc., Pullman, WA, USA*) installed near the field was utilized for recording relative humidity (RH) and air temperature (T_a). The climatic data were recorded at an interval of 10 minutes. The canopy temperature (T_c) was measured using a portable hand-held infrared thermometer (IRT) (MI-2H0, *Apogee Instruments Inc, North Logan, UT, USA*). The IRT operates within an atmospheric window of 8µm to 14µm with a response time less than 600 milliseconds and was accurate to ±0.3 °C. The T_c values were recorded between 12 PM and 2 PM under clear sky conditions. Each T_c observation was recorded from four directions (north, south, west and east) to avoid radiation effects. The recorded observations were averaged to determine the T_c of the treatment. The measurement of T_c began at 20 DAS when 70% of crop cover was achieved. The T_c measured from treatment T1 represented the $T_{c\text{-}ww}$ value. The value of $T_{c\text{-}dry}$ was based on T_c measurements made from T8 only when the crop was severely stressed and about to wilt. The collected data in 2017 was used for model development (training or calibration) while the data in 2018 was used for model validation. The statistical summary of the development and validation data sets is presented in Table 2.

Table 2

2.3 Kohonen Self-organizing maps

2.3.1 Basics of the Kohonen self-organizing maps

KSOM is a widely used neural network, which utilizes clustering for converting non-linear complex relationship between a high dimensional input data into a simple relationship on a low dimensional output display (Kohonen et al. 1996). The KSOM is also known as the Kohonen map or feature map. The units (nodes or neurons) of the map become tuned to input signal patterns based on unsupervised competitive learning. The clustering of the input data is performed in a way, such that similar patterns are represented by the same output unit, or by one of its neighboring units (Rustum 2009; Stefanovic and Kulasora 2011).

The KSOM consists of the high dimensional input layer and the low dimensional output layer. These layers are interconnected completely with each other as shown in Figure 2. The output layer contains 'M' neurons arranged in a 2-D grid. Each neuron consists of the same set of variables contained in the input vectors. The optimum value for M is determined using Equation 4 (Garcia and Gonzalez 2004),

$$M = 5\sqrt{N} \tag{4}$$

Where N is the total number of data samples. Once the value of M is obtained, the dimensions of the map, columns and rows are determined using Equation 5 (Garcia and Gonzalez 2004),

$$\frac{l_1}{l_2} = \sqrt{\frac{e_1}{e_2}} \tag{5}$$

Where l_1 and l_2 are the number of rows and columns of the map, respectively. e_1 and e_2 are the biggest and second-biggest eigenvalue of the training dataset, respectively.

Figure 2

2.3.2 Training the KSOM

Before the KSOM is trained, the high-dimensional input data is first normalized. A normalized input vector is then chosen randomly and presented to each of the neurons seeded with random values. The KSOM uses Euclidian distance (Equation 6) to identify the code vector most similar to the presented input vector.

$$D_i = \sqrt{\sum_{j=1}^{n} m_j\left(x_j - w_{ij}\right)^2} \tag{6}$$

Where, D_i is the Euclidian distance between input vector and code vector i; n is the dimension of the input vector; w_{ij} is the j^{th} element of code vector i; x_j is the j^{th} element of current input vector; m_j is mask, whose value is 0 when the given element x_j of the input vector is missing, otherwise it is 1. This becomes very useful while handling problems involving missing elements because all that needs to be done is to set the value of m_j for such elements as zero. The neuron for which D_i is minimum is chosen as the winning node or best matching unit (BMU) as shown in Figure 2. The code vectors of this BMU and its adjacent neurons are then adjusted to improve the agreement with the input data using Equation (7).

$$w_i(t+1) = w_i(t) + \alpha(t)h_{ci}(t)[x(t) - w_i(t)] \tag{7}$$

Where, w_i is the i^{th} code vector; t is the time; $\alpha(t)$ is the learning rate at t; and $h_{ci}(t)$ is the neighborhood function centered in the winner unit c at time t. In this way, each map unit develops internally the ability to identify input vectors like itself. This feature is referred to as self-organizing since the classification is achieved without providing any external output (Penn 2005). The process continues until an optimal number of iterations is reached or a specific error criterion is attained. The learning effectiveness of the

KSOM is affected by the neighborhood function and the learning rate and hence both must be chosen carefully as seen in Equations 8 and 9 respectively.

$$h_{ci}(t) = exp^{\left(-\|r_c - r_i\|^2/\left(2\sigma^2(t)\right)\right)} \tag{8}$$

$$\alpha(t) = \alpha_0 \left(\frac{0.005}{\alpha_0}\right)^{t/T} \tag{9}$$

where, T is the training length for convergence, usually taken as equal to $250/\sqrt{N}$ (Vesanto et al. 2000), α_0 is the initial learning rate, r_c is the position of node c on the KSOM grid, r_i is the position of node i on the grid, and $\sigma(t)$ is the neighborhood radius. Both $\alpha(t)$ and $\sigma(t)$ decreases monotonically with the increasing number of iterations.

The topographic and quantization errors are used to measure the quality of the trained KSOM. The errors are given by Equations 10 and 11 respectively.

$$t_e = \frac{1}{N}\sum_{i=1}^{N} u(X_i) \tag{10}$$

$$q_e = \frac{1}{N}\sum_{i=1}^{N} \|X_i - W_c\| \tag{11}$$

Where, t_e is the topographic error, q_e is the quantization error, X_i is the i^{th} input vector, W_c is the prototype vector of the winning node (BMU) for X_i; ||.|| represents the Euclidian distance (equation (6)), and u is a binary integer whose value is 1 if the first and second BMU are not adjacent units, otherwise zero.

The practical applications of the KSOM include data reduction for model identification, prediction, non-linear interpolation, generalization and compression of information (Kohonen 1996). In the present study, the KSOM is applied for prediction purpose as illustrated in Figure 3. Firstly, the available data is used to train a model. Once the model is trained, the depleted vector in which the predictand variable is either deliberately removed or missing is shown to the KSOM to find its BMU. The values of the missing variables are then obtained as their corresponding values in the BMU.

Figure 3

2.3.3 KSOM modeling

KSOM modeling in the study was performed using the SOM toolbox for MATLAB (Vesanto et al. 2000; Vatanen et al. 2015). The main objective of the study was to develop and evaluate a KSOM model for estimating the $T_{c\text{-}ww}$. For this purpose, the data of RH, T_a, and $T_{c\text{-}ww}$ were used in the modeling. This was purposely done to evaluate the KSOM model using easily available limited climatic data. The dataset for model development considered 210 observations of each variable. Similarly, for model validation, a set of 225 data points were used. Table 2 provides the statistical summary of the development and validation data sets.

To minimize the potential bias of the autocorrelation in the predictive ability of the trained maps, the input vectors of the training dataset were selected randomly and presented to the map in each time step. The validation was crucial to establish the ability of the KSOM model to generalize. The T_{c-ww} was omitted from the input vectors during the validation phase, indicating that the T_{c-ww} values were missing. The BMU for each input vector of the validation phase was then determined to predict the missing T_{c-ww} values as illustrated in Figure 3. After obtaining the T_{c-ww} values from the BMU's, they were compared with their actual values for evaluating the performance of KSOM during validation.

2.4 Multiple linear regression

As noted earlier, two more modeling paradigms were considered for the prediction of the T_{c-ww}, namely multiple linear regression (MLR) and feed-forward neural network (FFNN). The description of MLR is available in any standard statistical textbook and will hence not be repeated here. Details and applications of MLR have been documented by Bottenberg and Ward (1963) and Aiken et al. (2012). The MLR model was implemented using the Data Analysis toolbox in Microsoft Excel. Initially, a regression equation was developed using the dataset of 2017. The regression equation consisted of T_{c-ww} as the response variable and T_a and RH as the predictor variables. The equation was then applied to the dataset of 2018 to estimate the T_{c-ww}. The estimated values were then compared with the actual values for evaluating the performance of the MLR model.

2.5 Feed Forward neural network

ANN is successfully used for modeling unknown, complex relationships to predict physical conditions (or variables). The ANN has wide applications in water resources sector including evapotranspiration modeling, reservoir operations management, rainfall-runoff modeling, streamflow prediction, and many more (ASCE 2000). The FFNN is the most commonly used ANN algorithm in which, several forward and backward passes are made through a network until a specified target error or a maximum number of epochs is reached (Jain and Kumar 2007). Normally, the network is trained using an input-output pair to estimate the synaptic weights (Bowden et al. 2005). A network architecture essentially consists of an input layer, a hidden layer, and an output layer. The network architecture along with the synaptic weights together constitutes the model and is stored. When new inputs are presented to the model, it uses the training experience to predict the output.

The neural network toolbox of MATLAB was used to develop and validate the FFNN model. The development dataset (2017) was randomly partitioned into datasets for training (70%), validation while training (15%) and testing (15%). While the random nature of partitioning data might suggest the need for repeat trials, the data record used for the analysis is unlikely to produce a radically different outcome from the single randomization, thus making repetitions unnecessary. The input data were preprocessed, and the variables were normalized to a range of −1 to +1 before presenting them to the network.

A multilayer perceptron FFNN architecture was used to estimate the $T_{c\text{-}ww}$. The neurons in the hidden layer used a hyperbolic tangent activation function and the neuron in the output layer used a logistic activation function. The network architectures were evaluated with up to ten neurons in the hidden layer. The trial-error method based on minimizing the error and maximizing the correlation within the training dataset while utilizing minimum number of hidden neurons to avoid over-fitting the model was used to select the best network architecture. The FFNN architecture consisting of a hidden layer (5 neurons) and an output neuron was selected in the study (Figure 4). The Levenberg-Marquardt algorithm was applied for training the network using the training dataset due to its faster convergence and small residuals (errors) than other algorithms tested. The performance of the developed FFNN model was then validated with the dataset of 2018.

Figure 4

2.6 Statistical evaluation

The performance of the models KSOM, MLR, and FFNN were evaluated using qualitative (graphical regressions) and quantitative (error statistics) comparisons. The regression line significance was evaluated using the analysis of variance (ANOVA) test statistics. Following error statistics were used in the study:

1. The mean bias error (MBE) measures the average bias in the model predicted values.

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(x_i - x'_i) \tag{12}$$

2. The mean absolute error (MAE) measures the average of the absolute errors of the model predicted values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - x'_i| \tag{13}$$

3. The mean square error (MSE) measures the average of the square of the errors of the model predictions.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - x'_i)^2 \tag{14}$$

4. The percent error (PE) expresses the difference between a predicted and actual value, divided by the actual value.

$$PE = \left|\frac{\bar{x} - \overline{x'}}{\bar{x}}\right| \times 100 \tag{15}$$

5. The correlation coefficient (R^2) assesses the effectiveness of the model in predicting actual values.

$$R^2 = \left[\frac{n\sum x \times x' - \sum x \sum x'}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum x'^2 - (\sum x')^2]}}\right]^2 \tag{16}$$

where, *x'* is the model predicted value; *x* is the actual value; *n* is the number of samples.

3. Results and Discussion

3.1 Measured well-watered canopy temperature

The time series plot of measured $T_{c\text{-}ww}$ during 2017 and 2018 cropping seasons are shown in Figure 5. There was no difference ($p>0.05$) between the measured $T_{c\text{-}ww}$ across the three replications, hence, their mean is utilized for indicating the variation. However, a significant difference ($p\leq0.05$) was observed between the measured $T_{c\text{-}ww}$ across the 2017 and 2018 cropping seasons. This is not surprising, given the variability in the environmental factors during both seasons as shown in Table 2.

Figure 5

3.2 KSOM modeling results

The KSOM model development and validation was based on the dataset from 2017 and 2018 growing seasons respectively. Initially, default values of learning rate (α_0 = 0.5) and neighborhood radius (σ_0 = max(l_1,l_2)/4) were used to train the model in the SOM toolbox, where l_1 and l_2 are the dimensions of map computed using Equation (5). The toolbox uses Equation (4) to compute the size (number of units or neurons) of the map, however, the final units on map (M) are adjusted such that it equals the product of l_1 and l_2. The KSOM model has the map size of M = 72 units having dimensions 12×6. The topographic and quantization errors in the map are 0.427 and 0.109, respectively.

A significant feature of the KSOM is the development of the component planes which enables visualization of the correlation between the variables. The component planes for each variable in the KSOM are shown in Figure 6. Each plane is a sliced version of the KSOM and contains a single vector variable which represents its value in each map unit (Kalteh et al. 2008). The component planes are filled using colored or grey shades to reflect the feature values of each KSOM unit in the 2-D lattice, in such a way that, the darker the color, the lower the relative value of the component of the corresponding variable. In this way, the component planes visually indicate the regions in which a variable is high, low or average. This facilitates visual interpretation of the correlation between KSOM modeled values of $T_{c\text{-}ww}$, RH and T_a.

Visual analysis of the component planes shows that the color (or grey) gradient of the plane for $T_{c\text{-}ww}$ is parallel to the gradient of T_a, with high values of $T_{c\text{-}ww}$ being correlated with the high values of T_a and vice-versa. The component plane also confirms a negative correlation of RH with $T_{c\text{-}ww}$ and T_a, with low values of the former associated with the high values of the latter. A lower value of RH corresponds to a higher water deficit, resulting in an increase in the transpiration from crops (under potential soil water conditions), thereby causing relative transpirational cooling of the leaf surface. Now, by looking at the right bottom of the component plane of each variable, it can be seen that, at low values of RH, the $T_{c\text{-}ww}$ is lower than the T_a, confirming the accuracy of the model predictions.

Figure 6

Table 3 summarizes the error statistics for evaluating the performance of the KSOM model during development and validation. The correlation between measured and estimated values of T_{c-ww} was high, with R^2 equal to 0.981 and 0.978 during development and validation respectively, which indicate an excellent performance of the KSOM model in estimating the T_{c-ww} for Indian mustard. The KSOM utilized only two variables (T_a and RH) and still presented exemplary results. Linear regression between KSOM estimated and measured values of T_{c-ww} demonstrate a uniform scatter around the 1-1 line as shown by the X-Y plots (Figure 7). The regression line slope was not different ($p>0.05$) from 1-1 line during development and validation, indicating negligible bias in the model predictions. This is further substantiated by the low bias error values given in Table 3. The results in Figure 7 also indicate that the residuals of the prediction are random and normally distributed, hence a formal analysis of the residuals is not performed.

Figure 7

3.3 FFNN modeling results

The feed-forward neural network (FFNN) model architecture was developed using several scenarios based on trial-error and cross-validation. The best-performed model considered two input variables (RH, T_a), one hidden layer (with 5 neurons) and an output variable (T_{c-ww}). Figure 8 shows the X-Y plot of FFNN estimated and measured values of T_{c-ww}, which represent a good correlation with R^2 value of 0.90 and 0.92 during development and validation, respectively. The regression line slope during model validation was significantly different ($p\leq0.05$) from 1-1 line, which indicates a bias in FFNN predictions. Table 3 shows the descriptive summary of the error statistics used in the study for evaluating the performance of the FFNN model. The prediction results were similar to those reported by King and Shellie (2016) who utilized FFNN modeling for estimating T_{c-ww} with four climatic variables.

Figure 8

3.4 MLR modeling results

Estimation of T_{c-ww} using multiple linear regression (MLR) with the same input data (T_a, RH) provided the results similar to FFNN, during both development and validation (Table 3). The MLR equation in terms of T_a and RH is found to be as in Equation 17.

$$T_{c-ww} = 1.296 + 3.948 \times RH + 0.744 \times T_a \tag{17}$$

X-Y plots of MLR estimated and measured values of T_{c-ww} shown in Figure 9 represent a good correlation during development and validation with an R^2 value of 0.91 and 0.93, respectively. The correlation between measured and MLR estimated values was similar to that of the FFNN model. The regression line slope was significantly different ($p\leq0.05$) from 1-1 line during model validation, indicating a bias of the MLR model in estimating $T_{c\text{-}ww}$. Table 3 presents the error statistics for the performance evaluation of the MLR model.

Figure 9

Table 3

3.5 Comparison of KSOM, MLR and FFNN models

Table 3 summarises the error statistics for performance evaluation of KSOM, FFNN and MLR models. A comparison of the error statistics indicates that the performance of KSOM was much better than FFNN and MLR in estimating the $T_{c\text{-}ww}$ for Indian mustard. For example, the R^2 for the KSOM model during validation was 0.98, whereas, for the FFNN and MLR models, it was 0.92 and 0.92 respectively. Also, the FFNN and MLR model results were more biased than the KSOM model results during validation, which is indicated by the bias error estimates. Table 3 shows that the errors corresponding to MLR are similar to those of the FFNN model. A similar observation was reported by King and Shellie (2016).

Figure 10 shows the time series plots of the measured and model estimated values of $T_{c\text{-}ww}$ during development and validation, which further strengthens the efficacy of the KSOM model. In Figure 10, it can be seen, that the KSOM estimated $T_{c\text{-}ww}$ values are close to the measured values during the crop period, whereas those estimated using MLR and FFNN, although provided good results for the most part of the crop period, performed relatively poor during the mid and late growth seasons. Also, the performance of KSOM was better than MLR and FFNN during the most important validation phase. From this discussion, it can be inferred that KSOM can adequately model the $T_{c\text{-}ww}$, and its performance is better than FFNN and MLR models.

Figure 10

3.6 Crop water stress index

A further objective was to compute the crop water stress index (CWSI) of Indian mustard under different levels (T1 – T8) of soil water depletion (SWD). This objective was kept particularly to evaluate the performance of the KSOM estimated $T_{c\text{-}ww}$ in calculating the CWSI. Figures 11 and 12 show the time series plot of empirical CWSI for Indian mustard during 2017 and 2018 respectively. The empirical CWSI was computed using Equation (1) based on measured $T_{c\text{-}ww}$ ($CWSI_{measured}$) and KSOM estimated $T_{c\text{-}ww}$ ($CWSI_{KSOM}$). As previously indicated, the value of $T_{c\text{-}dry}$ was based on T_c measurements made from treatment T8 under maximum water-stressed conditions (CWSI ~ 0.8-1.0). For example, as seen in Figure 11 (T8), the T_c values during (35-40 DAS), (72-78 DAS) and (90-95 DAS) were utilized for computing the value of $T_{c\text{-}dry}$. The mean of these observations was ~T_a + 2 °C. Similar observations were obtained during 2018 cropping season. Hence, the value of $T_{c\text{-}dry}$ for the present study was considered equal to T_a + 2 °C.

ANOVA results indicated a significant difference ($p \leq 0.05$) between the empirical CWSI obtained for treatments T1-T8. This is not surprising, since irrigation was supplied at a specific level of SWD in each treatment, and the resulting CWSI was likely to be different. In Figures 11 and 12, it is observed that the CWSI reaches a certain level and then drops due to irrigation or rainfall (wetting event). This can be seen as an inverse scenario of soil water, which decreases with time, reaching a minimum, and then rises due to a wetting event.

It is evident from Figures 11 and 12, that the $CWSI_{KSOM}$ closely matched with the $CWSI_{measured}$ during both model development and validation. A closer observation reveals that $CWSI_{KSOM}$ estimates presented a better agreement with $CWSI_{measured}$ for treatments T5-T8, as compared to treatments T1-T4. This could be because, CWSI computations are more sensitive to $T_{c\text{-}ww}$ values at lower SWD levels, and even a minute error in the estimation of $T_{c\text{-}ww}$ could result in much more enhanced error in CWSI. This observation regarding the sensitivity of CWSI to different SWD levels is consistent with the findings of Colaizzi et al. (2003 a, b). At higher SWD levels, the results were exemplary which indicates the potential of $CWSI_{KSOM}$ under water-stressed scenarios. Hence, KSOM provides a reliable alternative to other algorithms with complex computations and extensive data requirements.

A critical observation regarding the maximum value of CWSI in each treatment can be made since irrigation scheduling through the CWSI approach is based on its value. Further evaluation of these results and comparison thereof with the SWD, water use efficiency and yield, will provide an insight into the scheduling criterion for Indian mustard adopting a simple KSOM based approach.

Figure 11

Figure 12

However, like any other modeling technique, the KSOM model developed in this study has some limitations, and these should be kept in mind while applying the model. As common with most data-driven approaches, the model performance is limited by the number of data points used in model development. In the present study, data from a single crop period with three replications have been used. Though the model performance was good, a relatively larger data set can increase performance, since more patterns can be extracted from them. Therefore, studies on more crops during other seasons need to be conducted to induce a generalization in the KSOM model. Another limitation is that the model is developed based on experiments performed in a single agro-climate. However, the analysis used in the study can be easily extended over more data, hence this should not be seen as a major problem.

4. Conclusion

The current work presents a novel approach involving the application of Kohonen Self-Organising Map (KSOM) in estimating the well-watered canopy temperature ($T_{c\text{-}ww}$) for computing the crop water stress index (CWSI). Field crop experiments on Indian mustard were performed in a humid sub-tropical agro-climate, during the 2017 and 2018 cropping seasons. Field measurements of $T_{c\text{-}ww}$ were obtained from a well-watered irrigated treatment. The performance of the KSOM, MLR and FFNN models was evaluated with the observed values of $T_{c\text{-}ww}$. The results based on the error statistics and graphical comparisons indicated that the KSOM model outperformed the MLR and FFNN models in estimating the $T_{c\text{-}ww}$. The KSOM estimated $T_{c\text{-}ww}$ was further used for computing the empirical CWSI in various treatments irrigated at different levels of soil water depletion. Visual observation in different treatments indicated that KSOM based empirical CWSI was closely related to the field-based empirical CWSI. The predictions of the KSOM model were reliable during development and validation. A unique feature of KSOM is that its predictive ability is unencumbered even if some of its input variables are missing, which is not the case with either FFNN or MLR modeling approaches. The CWSI based on KSOM estimated $T_{c\text{-}ww}$ provides a simple alternative to other complex algorithms for monitoring crop stresses and irrigation scheduling applications. The KSOM model developed in the study is expected to work well in similar agro-climates. Further research should concentrate on the application of KSOM modeling in estimating the $T_{c\text{-}ww}$ and subsequently calculating the CWSI for different crops, across different agro-climates.

Funding

The work received external funding from UK-NERC (Award No. NE/N016394/1) and Indian-MoES (Award No. MoES/NERC/IA-SWR/P3/10/2016-PC-II) through a scientific research collaborative project "Sustaining Himalayan Water Resources in a changing climate (SusHi-Wat)". The study is an outcome of visiting fellowship (ODF/2018/000374) awarded by SERB, DST (Govt. of India) to Navsal Kumar for researching at Heriot-Watt University, Edinburgh (UK).

Acknowledgments

The authors are grateful to the Institute for Infrastructure and Environment, Heriot-Watt University (UK) and Department of Civil Engineering, NIT Hamirpur (India) for providing necessary technical guidance, experimental facilities, and support for the study. The authors are thankful to the three anonymous reviewers whose critical suggestions and feedback assisted the authors in improving the quality of the manuscript.

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request (field experimental data).

Software Availability Statement

The SOM Toolbox (version 2.1) for MATLAB used in the present study is freely available to download from GITHUB (https://github.com/ilarinieminen/SOM-Toolbox).

References

Adeloye, A. J., & Rustum, R. (2012). Self-organising map rainfall-runoff multivariate modelling for runoff reconstruction in inadequately gauged basins. *Hydrology Research*, *43*(5), 603-617.

Adeloye, A. J., Rustum, R., & Kariyama, I. D. (2011). Kohonen self-organizing map estimator for the reference crop evapotranspiration. *Water Resources Research*, *47*(8).

Adeloye, A. J., Rustum, R., & Kariyama, I. D. (2012). Neural computing modeling of the reference crop evapotranspiration. *Environmental Modelling & Software*, *29*(1), 61-73.

Adeyemi, O., Grove, I., Peets, S., Domun, Y., & Norton, T. (2018). Dynamic modelling of the baseline temperatures for computation of the crop water stress index (CWSI) of a greenhouse cultivated lettuce crop. *Computers and Electronics in Agriculture*, *153*, 102-114.

Agam, N., Cohen, Y., Alchanatis, V., & Ben-Gal, A. (2013). How sensitive is the CWSI to changes in solar radiation?. *International journal of remote sensing*, *34*(17), 6109-6120.

Aiken, L. S., West, S. G., Pitts, S. C., Baraldi, A. N., & Wurpts, I. C. (2012). Multiple linear regression. *Handbook of Psychology, Second Edition*, *2*.

Akkuzu, E., Kaya, Ü., Çamoğlu, G., Mengü, G. P., & Aşik, Ş. (2013). Determination of crop water stress index and irrigation timing on olive trees using a handheld infrared thermometer. *Journal of irrigation and drainage engineering*, *139*(9), 728-737.

Alchanatis, V., Cohen, Y., Cohen, S., Moller, M., Sprinstin, M., Meron, M., ... & Sela, E. (2010). Evaluation of different approaches for estimating and mapping crop water status in cotton with thermal imaging. *Precision Agriculture*, *11*(1), 27-41.

Al-Faraj, A., Meyer, G. E., & Horst, G. L. (2001). A crop water stress index for tall fescue (Festuca arundinacea Schreb.) irrigation decision-making—a fuzzy logic method. *Computers and electronics in agriculture*, *32*(2), 69-84.

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome*, *300*(9), D05109.

Anda, A., Soós, G., Menyhárt, L., Kucserka, T., & Simon, B. (2020). Yield features of two soybean varieties under different water supplies and field conditions. *Field Crops Research*, *245*, 107673.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000). Artificial neural networks in hydrology. II: Hydrologic applications. *Journal of Hydrologic Engineering*, *5*(2), 124-137.

Bellvert, J., Marsal, J., Girona, J., Gonzalez-Dugo, V., Fereres, E., Ustin, S. L., & Zarco-Tejada, P. J. (2016). Airborne thermal imagery to detect the seasonal evolution of crop water status in peach, nectarine and Saturn peach orchards. *Remote Sensing*, *8*(1), 39.

Boomiraj, K., Chakrabarti, B., Aggarwal, P. K., Choudhary, R., & Chander, S. (2010). Assessing the vulnerability of Indian mustard to climate change. *Agriculture, ecosystems & environment*, *138*(3-4), 265-273.

Bottenberg, R. A., & Ward, J. H. (1963). *Applied multiple linear regression* (Vol. 63, No. 6). 6570th Personnel Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Lackland Air Force Base.

Bowden, G. J., Dandy, G. C., & Maier, H. R. (2005). Input determination for neural network models in water resources applications. Part 1—background and methodology. *Journal of Hydrology*, *301*(1-4), 75-92.

Chen, I. T., Chang, L. C., & Chang, F. J. (2018). Exploring the spatio-temporal interrelation between groundwater and surface water by using the self-organizing maps. *Journal of Hydrology*, *556*, 131-142.

Clark, S., Sarlin, P., Sharma, A., & Sisson, S. A. (2015). Increasing dependence on foreign water resources? An assessment of trends in global virtual water flows using a self-organizing time map. *Ecological Informatics*, *26*, 192-202.

Colaizzi, P. D., Barnes, E. M., Clarke, T. R., Choi, C. Y., & Waller, P. M. (2003a). Estimating soil moisture under low frequency surface irrigation using crop water stress index. *Journal of irrigation and drainage engineering*, *129*(1), 27-35.

Colaizzi, P. D., Barnes, E. M., Clarke, T. R., Choi, C. Y., Waller, P. M., Haberland, J., & Kostrzewski, M. (2003b). Water stress detection under high frequency sprinkler irrigation with water deficit index. *Journal of Irrigation and Drainage Engineering*, *129*(1), 36-43.

Egea, G., Padilla-Díaz, C.M., Martinez-Guanter, J., Fernández, J.E. and Pérez-Ruiz, M., 2017. Assessing a crop water stress index derived from aerial thermal imaging and infrared thermometry in super-high density olive orchards. *Agricultural water management*, *187*, pp.210-221.

Ehrler, W. L. (1973). Cotton leaf temperatures as related to soil water depletion and meteorological factors 1. *Agronomy Journal*, *65*(3), 404-409.

García, H. L., & González, I. M. (2004). Self-organizing map and clustering for wastewater treatment monitoring. *Engineering Applications of Artificial Intelligence*, *17*(3), 215-225.

Gerhards, M., Schlerf, M., Mallick, K. and Udelhoven, T., 2019. Challenges and Future Perspectives of Multi-/Hyperspectral Thermal Infrared Remote Sensing for Crop Water-Stress Detection: A Review. *Remote Sensing*, *11*(10), p.1240.

Gontia, N. K., & Tiwari, K. N. (2008). Development of crop water stress index of wheat crop for scheduling irrigation using infrared thermometry. *Agricultural water management*, *95*(10), 1144-1152.

González-Dugo, V., Zarco-Tejada, P. J., & Fereres, E. (2014). Applicability and limitations of using the crop water stress index as an indicator of water deficits in citrus orchards. *Agricultural and forest meteorology*, *198*, 94-104.

Heydari, A., Bijanzadeh, E., Naderi, R. and Emam, Y., 2019. Quantifying water stress in canola (Brassica napus L.) using crop water stress index. *Iran Agricultural Research*, *38*(1), pp.1-8.

Hou, M., Tian, F., Zhang, L., Li, S., Du, T., Huang, M. and Yuan, Y., 2019. Estimating crop transpiration of soybean under different irrigation treatments using thermal infrared remote sensing imagery. *Agronomy*, *9*(1), p.8.

Idso, S. B., Jackson, R. D., Pinter Jr, P. J., Reginato, R. J., & Hatfield, J. L. (1981). Normalizing the stress-degree-day parameter for environmental variability. *Agricultural Meteorology*, *24*, 45-55.

Ihuoma, S. O., & Madramootoo, C. A. (2017). Recent advances in crop water stress detection. *Computers and Electronics in Agriculture*, *141*, 267-275.

Indian Ministry of Environment, Forest and Climate Change (MoEFCC) / Directorate of Rapeseed Mustard Research. (2016). Biology of *Brassica juncea* (Indian Mustard). New Delhi.

Jackson, R. D., Idso, S. B., Reginato, R. J., & Pinter Jr, P. J. (1981). Canopy temperature as a crop water stress indicator. *Water resources research*, *17*(4), 1133-1138.

Jain, A., & Kumar, A. M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, *7*(2), 585-592.

Jones, H. G. (1999). Use of infrared thermometry for estimation of stomatal conductance as a possible aid to irrigation scheduling. *Agricultural and forest meteorology*, *95*(3), 139-149.

Kalteh, A. M., Hjorth, P., & Berndtsson, R. (2008). Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environmental Modelling & Software*, *23*(7), 835-845.

King, B. A., & Shellie, K. C. (2016). Evaluation of neural network modeling to predict non-water-stressed leaf temperature in wine grape for calculation of crop water stress index. *Agricultural water management*, *167*, 38-52.

King, B. A., & Shellie, K. C. (2018). Wine grape cultivar influence on the performance of models that predict the lower threshold canopy temperature of a water stress index. *Computers and Electronics in Agriculture, 145,* 122-129.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE, 78*(9), 1464-1480.

Kohonen, T., Oja, E., Simula, O., Visa, A., & Kangas, J. (1996). Engineering applications of the self-organizing map. *Proceedings of the IEEE, 84*(10), 1358-1384.

Kothari, R., & Islam, S. (1999). Spatial characterization of remotely sensed soil moisture data using self organizing feature maps. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(2), 1162-1165.

Kumar, N., Adeloye, A. J., Shankar, V., & Rustum, R. (2020a). Neural computing modelling of the crop water stress index. *Agricultural Water Management*, *239*, 106259.

Kumar, N., Poddar, A., & Shankar, V. (2019, August). Optimizing irrigation through environmental canopy sensing–A proposed automated approach. In *AIP Conference Proceedings* (Vol. 2134, No. 1, p. 060003). AIP Publishing LLC.

Kumar, N., Poddar, A., Shankar, V., Ojha, C. S. P., & Adeloye, A. J. (2020b). Crop water stress index for scheduling irrigation of Indian mustard (Brassica juncea) based on water use efficiency considerations. *Journal of Agronomy and Crop Science, 206*(1), 148-159.

Maes, W. H., & Steppe, K. (2012). Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review. *Journal of Experimental Botany*, *63*(13), 4671-4712.

Meron, M., Tsipris, J., Orlov, V., Alchanatis, V., & Cohen, Y. (2010). Crop water stress mapping for site-specific irrigation by thermal imagery and artificial reference surfaces. *Precision agriculture*, *11*(2), 148-162.

Mishra, J., Singh, R.K., DeshrajYadaw, S.S. and Mishra, A.P., 2019. Quality of Indian mustard [Brassica juncea (L.) Czernj and Cosson] as influenced by tillage and irrigation frequency. *Journal of Pharmacognosy and Phytochemistry*, *8*(1), pp.2280-2283.

Mwale, F. D., Adeloye, A. J., & Rustum, R. (2012). Infilling of missing rainfall and streamflow data in the Shire River basin, Malawi–A self-organizing map approach. *Physics and Chemistry of the Earth, Parts A/B/C, 50,* 34-43.

Ohana-Levi, N., Bahat, I., Peeters, A., Shtein, A., Netzer, Y., Cohen, Y., & Ben-Gal, A. (2019). A weighted multivariate spatial clustering model to determine irrigation management zones. *Computers and Electronics in Agriculture*, *162*, 719-731.

O'shaughnessy, S. A., Evett, S. R., Colaizzi, P. D., & Howell, T. A. (2011). Using radiation thermography and thermometry to evaluate crop water stress in soybean and cotton. *Agricultural Water Management*, *98*(10), 1523-1535.

Osroosh, Y., Peters, R. T., Campbell, C. S., & Zhang, Q. (2015). Automatic irrigation scheduling of apple trees using theoretical crop water stress index with an innovative dynamic threshold. *Computers and Electronics in Agriculture*, *118*, 193-203.

Osroosh, Y., Peters, R. T., Campbell, C. S., & Zhang, Q. (2016). Comparison of irrigation automation algorithms for drip-irrigated apple trees. *Computers and Electronics in Agriculture*, *128,* 87-99.

Park, S., 2018. *Estimating plant water stress and evapotranspiration using very-high-resolution (VHR) UAV imagery* (Doctoral dissertation).

Payero, J. O., & Irmak, S. (2006). Variable upper and lower crop water stress index baselines for corn and soybean. *Irrigation Science*, *25*(1), 21-32.

Penn, B. S. (2005). Using self-organizing maps to visualize high-dimensional data. *Computers & Geosciences*, *31*(5), 531-544.

Poirier-Pocovi, M. and Bailey, B.N., 2020. Sensitivity analysis of four crop water stress indices to ambient environmental conditions and stomatal conductance. *Scientia Horticulturae*, *259*, p.108825.

Ramachandran, A., Rustum, R., & Adeloye, A. J. (2019). Anaerobic digestion process modeling using Kohonen self-organising maps. *Heliyon*, *5*(4), e01511.

Rana, K., Parihar, M., Singh, J.P. and Singh, R.K., 2019. Effect of sulfur fertilization, varieties and irrigation scheduling on growth, yield, and heat utilization efficiency of indian mustard (Brassica Juncea L.). *Communications in Soil Science and Plant Analysis*, pp.1-11.

Riese, F. M., & Keller, S. (2018, July). Introducing a framework of self-organizing maps for regression of soil moisture with hyperspectral data. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 6151-6154). IEEE.

Romero-Trigueros, C., Bayona Gambín, J.M., Nortes Tortosa, P.A., Alarcón Cabañero, J.J. and Nicolás Nicolás, E., 2019. Determination of Crop Water Stress Index by Infrared Thermometry in Grapefruit Trees Irrigated with Saline Reclaimed Water Combined with Deficit Irrigation. *Remote Sensing*, *11*(7), p.757.

Rustum, R. (2009). *Modelling activated sludge wastewater treatment plants using artificial intelligence techniques (fuzzy logic and neural networks)* (Doctoral dissertation, Heriot-Watt University).

Rustum, R., & Adeloye, A. J. (2007). Replacing outliers and missing values from activated sludge data using Kohonen self-organizing map. *Journal of Environmental Engineering*, *133*(9), 909-916.

Shekhawat, K., Rathore, S. S., Premi, O. P., Kandpal, B. K., & Chauhan, J. S. (2012). Advances in agronomic management of Indian mustard (Brassica Juncea (L.) Czernj. Cosson): an overview. *International Journal of Agronomy*, *2012*. doi:10.1155/2012/408284

Singh, P.K. and Singh, A.K., 2019. Effect of sowing dates and irrigation schedules on performance of Indian mustard (Brassica juncea L.) and water use efficiency. *Journal of Soil and Water Conservation*, *18*(2), pp.164-167.

Singh, V.V., Garg, P., Meena, H.S., & Meena, M.L. (2018). Drought Stress Response of Indian Mustard (Brassica juncea L.) Genotypes. *Int. J. Curr. Microbiol. App. Sci. 7*(3), 2519-2526.

Stefanovic, P., & Kurasova, O. (2011). Visual analysis of self-organizing maps. *Nonlinear Analysis: Modelling and Control*, *16*(4), 488-504.

Vatanen, T., Osmala, M., Raiko, T., Lagus, K., Sysi-Aho, M., Orešič, M., ... & Lähdesmäki, H. (2015). Self-organization and missing values in SOM and GTM. *Neurocomputing*, *147*, 60-70.

Vesanto, J., Himberg, J., Alhoniemi, E., & Parhankangas, J. (2000). SOM toolbox for Matlab 5. *Helsinki University of Technology, Finland*, *109*.

Wright, P. R., Morgan, J. M., & Jessop, R. S. (1996). Comparative adaptation of canola (Brassica napus) and Indian mustard (B. juncea) to soil water deficits: plant water relations and growth. *Field Crops Research*, *49*(1), 51-64.

Yildirim, M., Demirel, K., & Bahar, E. (2012). Effect of restricted water supply and stress development on growth of bell pepper (Capsicum Annuum L.) under drought conditions. *Journal of Agro Crop Science*, *3*(1), 1-9.

Yuan, G., Luo, Y., Sun, X., & Tang, D. (2004). Evaluation of a crop water stress index for detecting water stress in winter wheat in the North China Plain. *Agricultural Water Management*, *64*(1), 29-40.

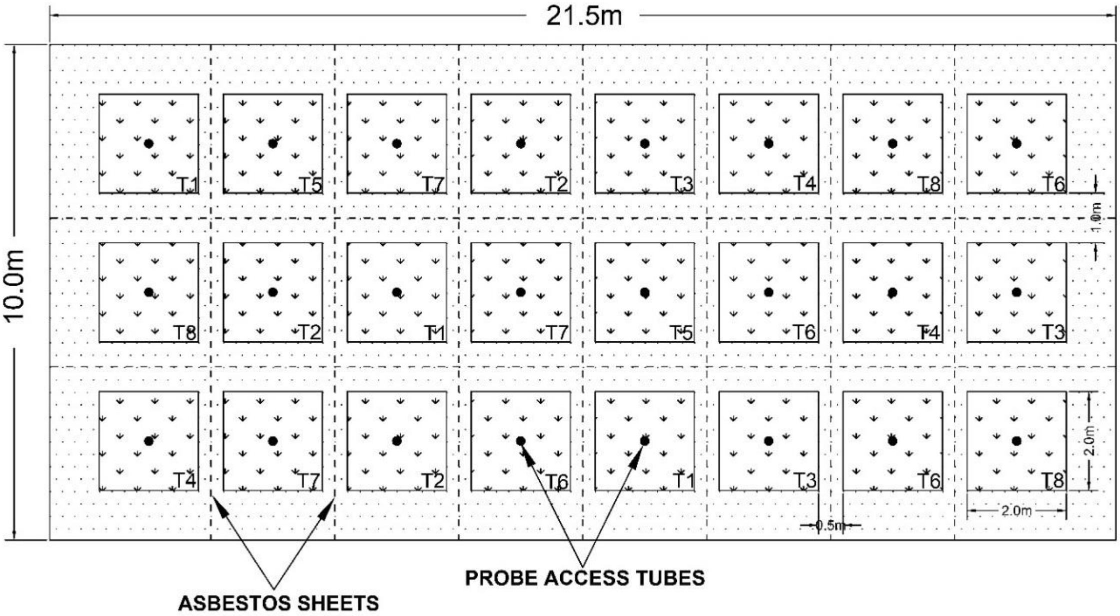

Figure 1. The layout of the experimental plot (T1- Well-watered plot; T2-10% SWD; T3-20% SWD; T4-30% SWD; T5-40% SWD; T6-50% SWD; T7-60% SWD; and T8-Maximum stressed plot.)

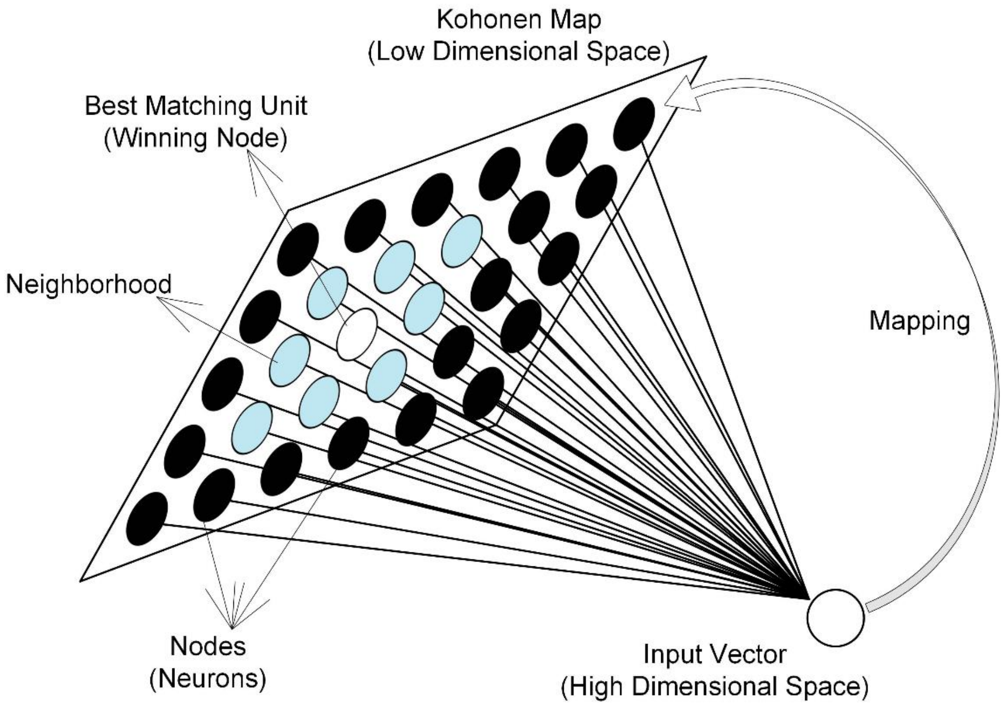

Figure 2. Representation of the winning node and its neighbors in a KSOM

682

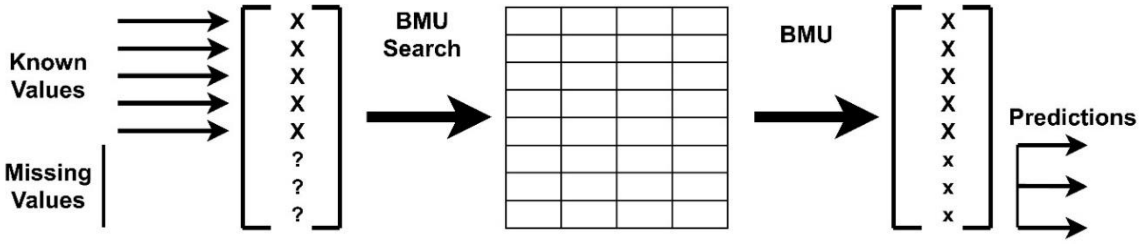

Figure 3. Prediction of the missing component of the input vector using the Kohonen Self Organizing Map.

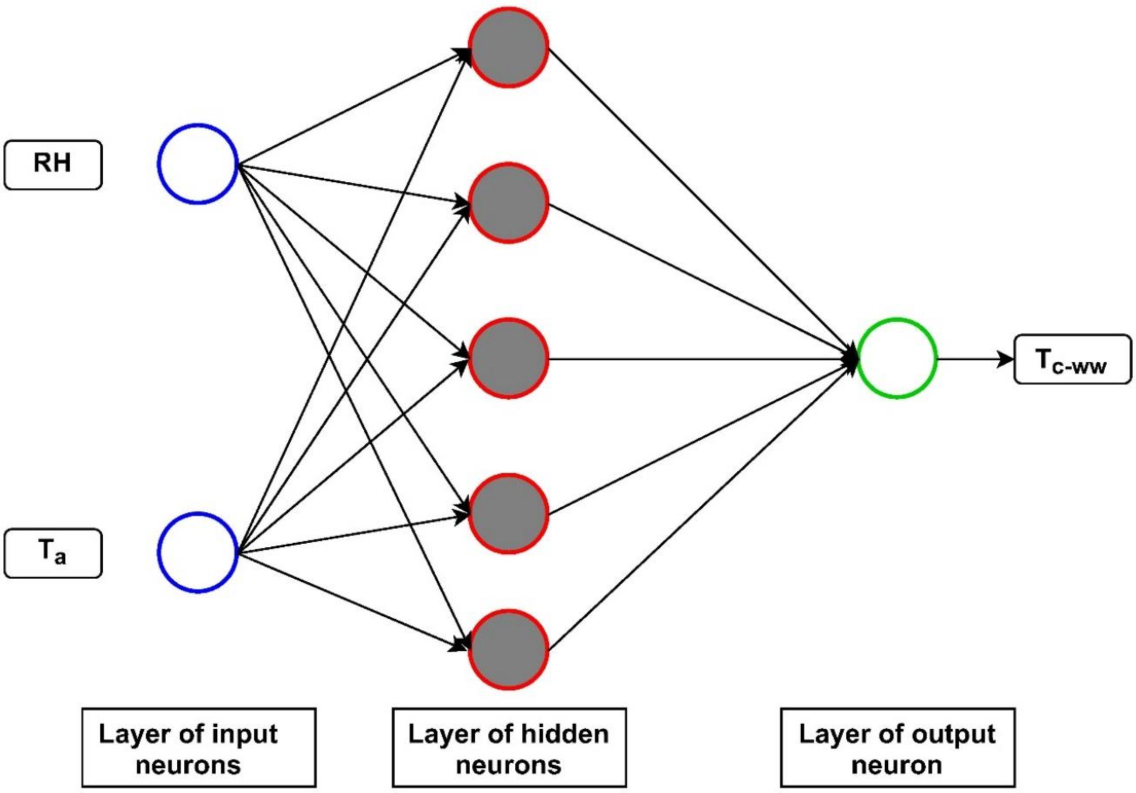

Figure 4. Schematic representation of the feed-forward neural network modeling architecture with two inputs and one hidden layer. RH – Relative humidity, T_a – Air temperature and $T_{c\text{-}ww}$ – Well-watered canopy temperature.

684

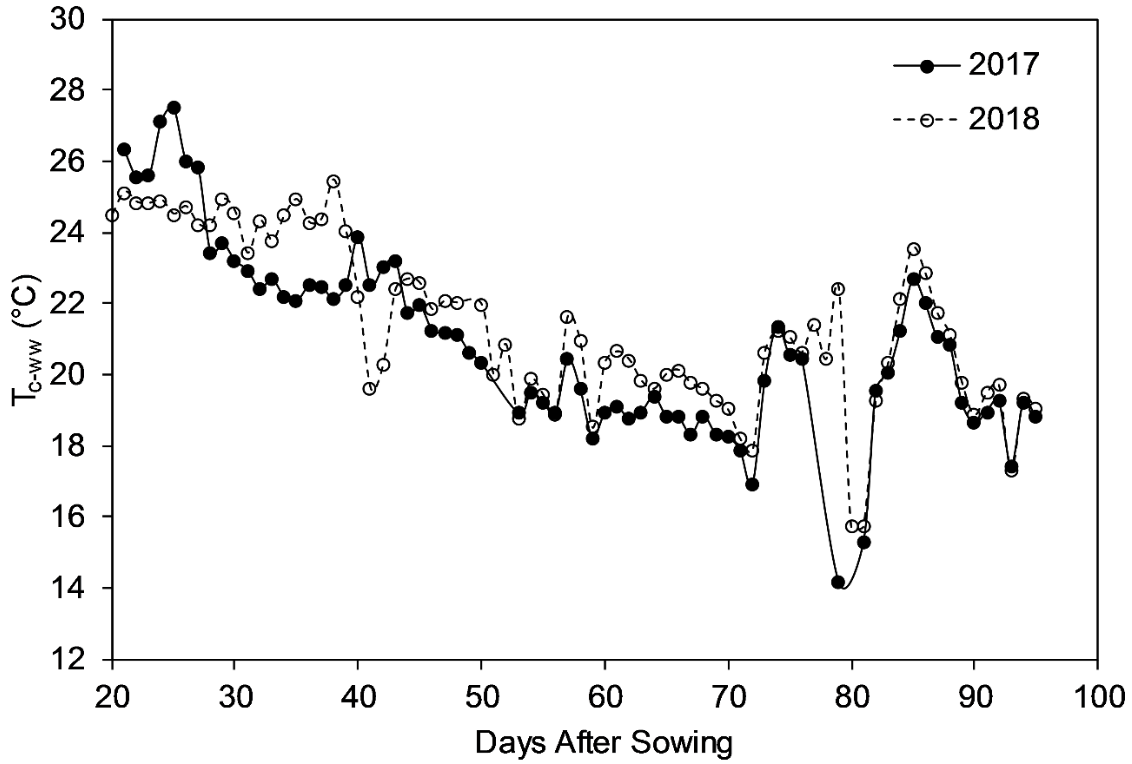

Figure 5. Time series plot of well-watered canopy temperature of Indian mustard during the crop period

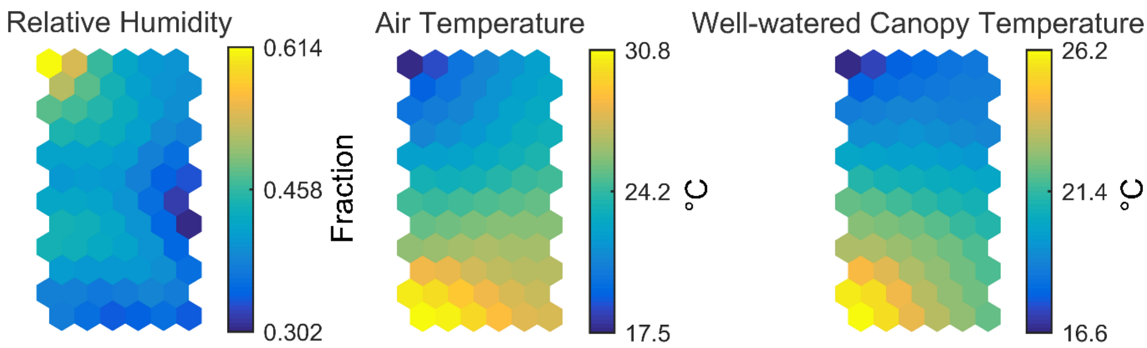

Figure 6. KSOM component planes.

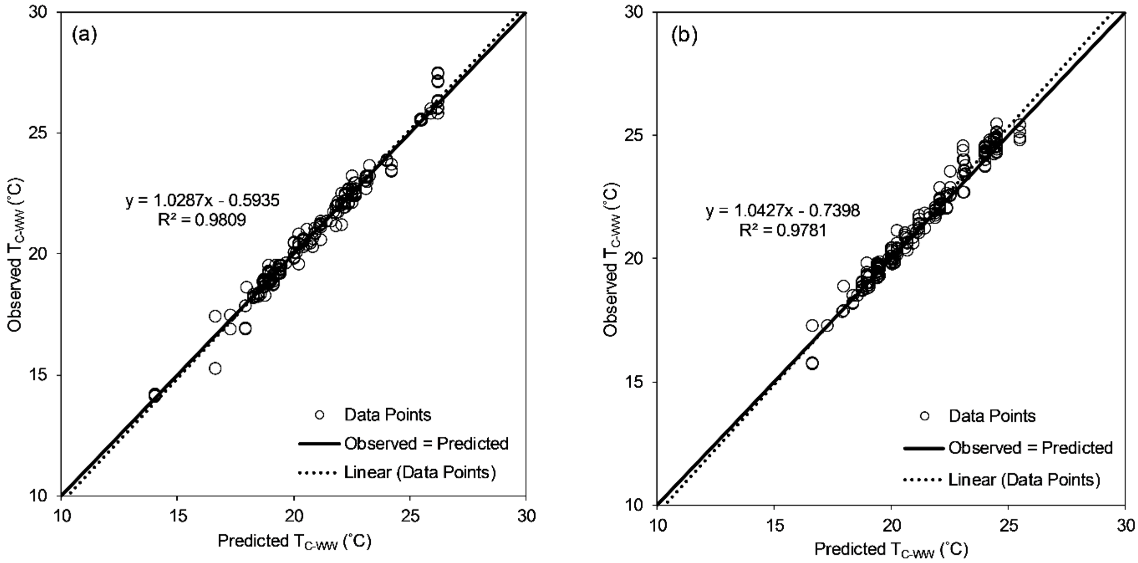

Figure 7. X-Y scatter plot of KSOM predicted and measured values of $T_{c\text{-}ww}$ during (a) model development, and (b) model validation

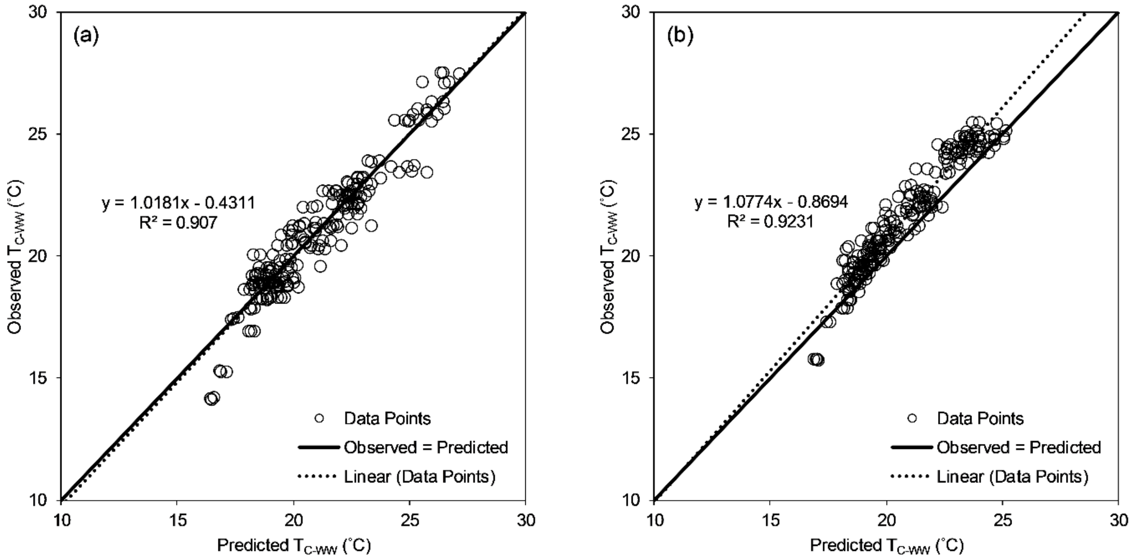

Figure 8. X-Y scatter plot of FFNN predicted and measured values of $T_{c\text{-}ww}$ during (a) model development, and (b) model validation

688

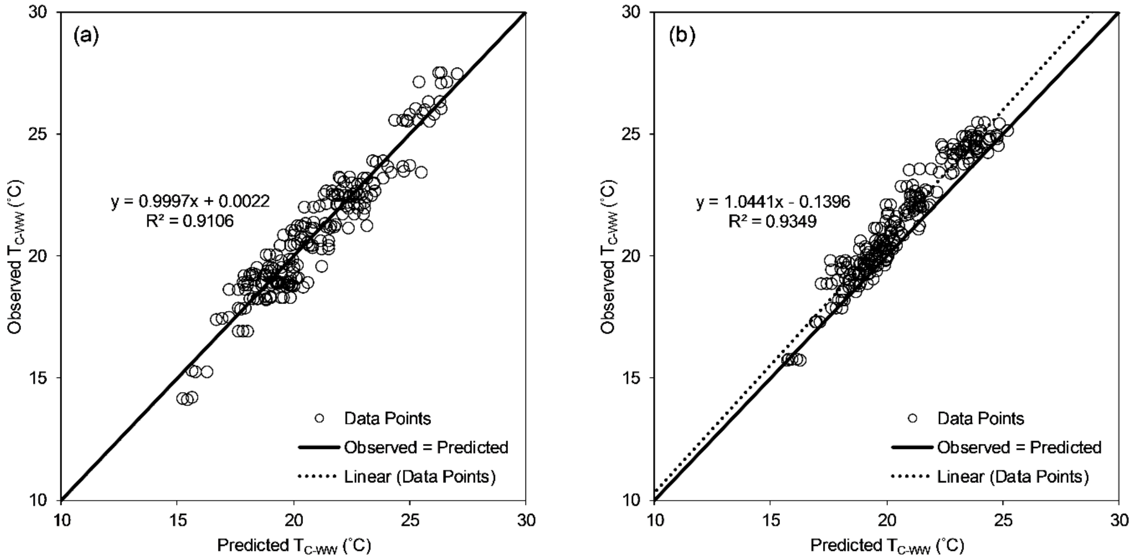

Figure 9. X-Y scatter plot of MLR predicted and measured values of $T_{c\text{-}ww}$ during (a) model development, and (b) model validation

689

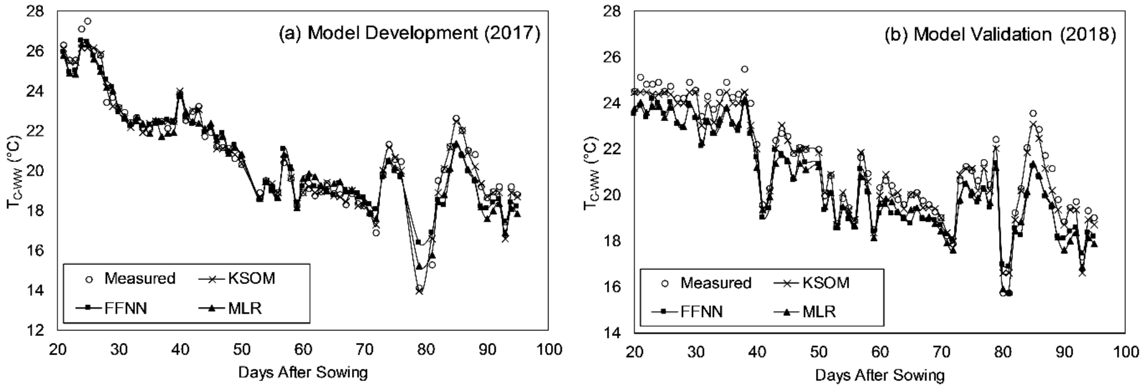

Figure 10. Time series plot of measured and predicted well-watered canopy temperature of Indian mustard during the crop period for the growing season (a) 2017 and (b) 2018

690

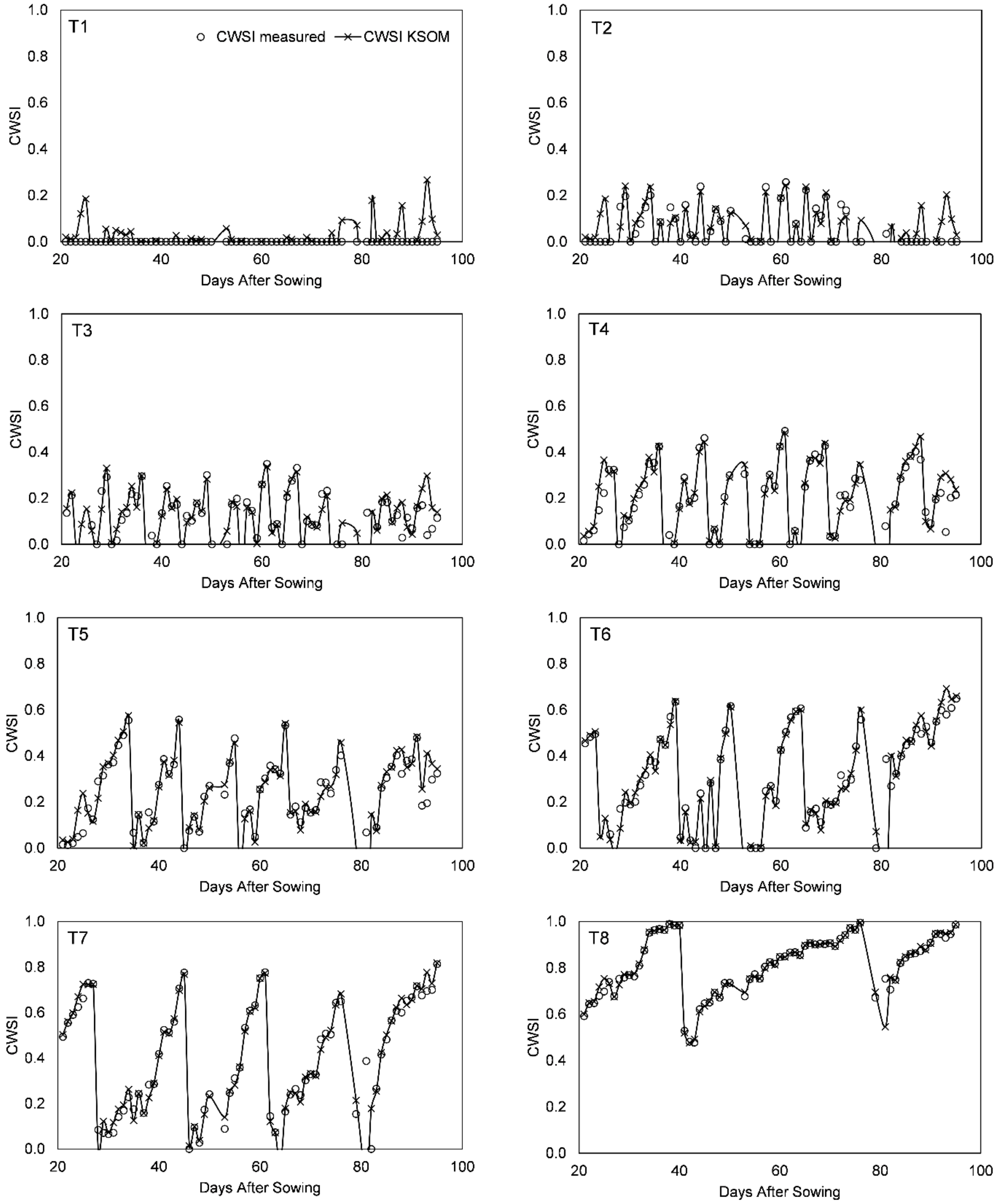

Figure 11. Comparison between observed CWSI (based on measured values of $T_{c\text{-}ww}$) and predicted CWSI (based on KSOM estimated values of $T_{c\text{-}ww}$) for different irrigation treatments during model development (2017)

691

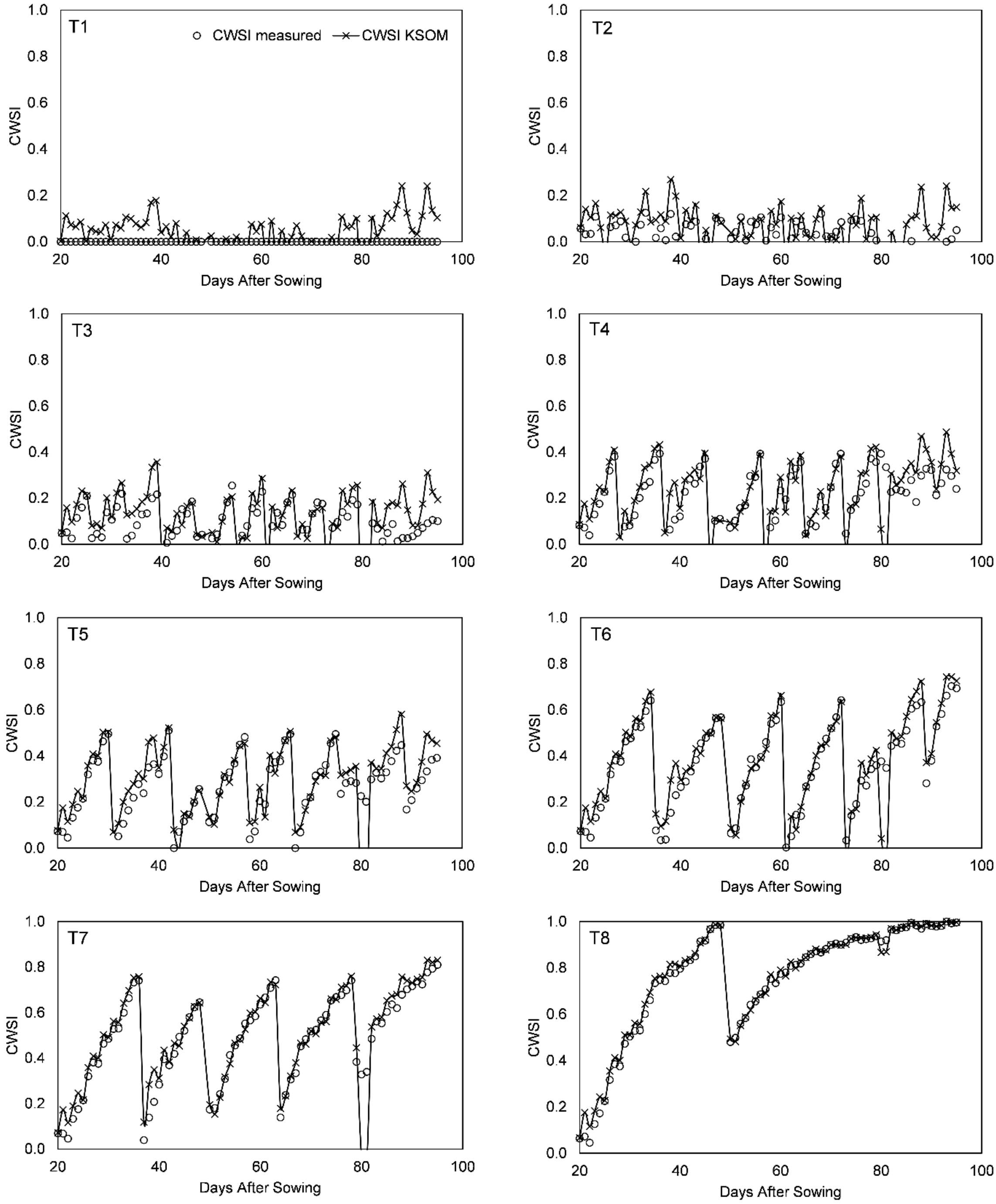

Figure 12. Comparison between observed CWSI (based on measured values of $T_{c\text{-}ww}$) and predicted CWSI (based on KSOM estimated values of $T_{c\text{-}ww}$) for different irrigation treatments during model validation (2018)

692

Table 1 Details of crop variety sown, growth stages, crop duration and spacing

Crop	Variety sown	Crop duration (Days)	Growth stages (Days)*				Spacing (cm)	Date of Sowing	Date of Harvesting
			I	II	III	IV			
Indian mustard (*Brassica Juncea*)	P.T. 303	95	20	25	30	20	40×15	22nd September 2017	25th December 2017
		95	20	25	30	20	40×15	25th September 2018	28th December 2018

* *I - vegetative (initial stage), II - flowering (crop development stage), III - pod formation and seed development (mid-season stage), IV - maturity and harvest (late-season stage).*

Table 2 Statistical summary of data used for model development and validation

Variable	Units	Symbol	Dataset	Maximum	Minimum	Mean	SD
Relative Humidity	Fraction	RH	Development	0.88	0.21	0.41	0.10
			Validation	0.65	0.22	0.42	0.09
Air Temperature	°C	T_a	Development	32.4	14.1	24.18	3.64
			Validation	30.3	16.1	23.76	3.16
Well-watered Canopy Temperature	°C	$T_{c\text{-}ww}$	Development	27.51	14.13	20.91	1.64
			Validation	25.49	15.73	21.41	2.35

Table 3 Descriptive summary of error statistics for modelling well-watered canopy temperature (T_{c-ww}) in Indian mustard

Modelling Phase	Statistics	Mean			Maximum			Minimum			Standard deviation		
		KSOM	MLR	FFNN	KSOM	MLR	FFNN	KSOM	MLR	FFNN	KSOM	MLR	FFNN
Model Development (2017)	Observed (°C)	20.910	20.910	20.910	27.51	27.51	27.51	14.13	14.13	14.13	2.646	2.646	2.646
	Predicted (°C)	20.904	20.915	20.962	26.176	27.051	27.135	14.01	15.225	16.411	2.526	2.526	2.526
	Bias error (°C)	0.006	-0.004	-0.051	1.333	1.739	1.792	-1.354	-2.068	-2.365	0.373	0.791	0.808
	Absolute error (°C)	0.240	0.660	0.638	1.354	2.068	2.365	0.002	0.000	0.000	0.285	0.433	0.496
	Square error (°C)	0.138	0.623	0.652	1.835	4.276	5.596	0.000	0.000	0.000	0.335	0.711	0.992
	Percent error (%)	1.163	3.232	3.177	8.878	10.039	16.741	0.009	0.002	0.000	1.458	2.204	2.802
	R^2	0.981	0.910	0.907									
Model Validation (2018)	Observed (°C)	21.409	21.409	21.409	25.49	25.49	25.49	15.73	15.73	15.73	2.350	2.350	2.350
	Predicted (°C)	21.241	20.638	20.641	25.487	25.212	25.156	16.614	15.71	16.859	2.228	2.176	2.095
	Bias error (°C)	0.167	0.770	0.768	1.512	2.544	2.660	-0.884	-0.539	-1.348	0.360	0.607	0.703
	Absolute error (°C)	0.301	0.807	0.869	1.512	2.544	2.660	0.000	0.008	0.003	0.258	0.556	0.571
	Square error (°C)	0.157	0.961	1.082	2.287	6.470	7.078	0.000	0.000	0.000	0.280	1.151	1.243
	Percent error (%)	1.413	3.706	4.030	6.155	11.337	11.311	0.002	0.045	0.020	1.248	2.490	2.573
	R^2	0.978	0.924	0.923									