



Heriot-Watt University
Research Gateway

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

Citation for published version:

Kumar, N, Shankar, V, Rustum, R & Adeloje, AJ 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*)', *Journal of Irrigation and Drainage Engineering*, vol. 147, no. 2, 0001526. [https://doi.org/10.1061/\(asce\)ir.1943-4774.0001526](https://doi.org/10.1061/(asce)ir.1943-4774.0001526)

Digital Object Identifier (DOI):

[10.1061/\(asce\)ir.1943-4774.0001526](https://doi.org/10.1061/(asce)ir.1943-4774.0001526)

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Peer reviewed version

Published In:

Journal of Irrigation and Drainage Engineering

Publisher Rights Statement:

© 2020, American Society of Civil Engineers. This is a post-peer-review, pre-copyedit version of an article published in Journal of Irrigation and Drainage Engineering. The final authenticated version is available online at: [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0001526](https://doi.org/10.1061/(ASCE)IR.1943-4774.0001526)

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

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 Shankar², Rabee Rustum³, Adebayo J. Adeloje⁴

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.adeloje@hw.ac.uk

Cite this article as:

Kumar, N., Shankar, V., Rustum, R., & Adeloje, 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](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-ww}) using air temperature and relative humidity as input predictor variables. Comparisons were performed between model estimated and measured T_{c-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-ww} was compared with the CWSI obtained from measured values of T_{c-ww} . The results suggest a significant potential of KSOM for reliable estimation of the T_{c-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 al. 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),

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

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

72 theoretical approach, initially given by Jackson et al. (1981) is based on the energy balance model. The
 73 baseline temperatures are calculated using Equation 2 and 3, respectively.

$$(T_{c-ww} - T_a) = \frac{r_a R_n}{\rho c_p} \frac{\gamma(1 + (r_{c-ww}/r_a))}{\Delta + \gamma(1 + (r_{c-ww}/r_a))} - \frac{(e_s - e_a)}{\Delta + \gamma(1 + (r_{c-ww}/r_a))} \quad (2)$$

$$(T_{c-dry} - T_a) = \frac{r_a R_n}{\rho c_p} \frac{\gamma(1 + (r_{c-dry}/r_a))}{\Delta + \gamma(1 + (r_{c-dry}/r_a))} - \frac{(e_s - e_a)}{\Delta + \gamma(1 + (r_{c-dry}/r_a))} \quad (3)$$

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

79 The empirical approach was introduced by Idso et al. (1981) and considers the experimental
 80 observations of the baseline canopy temperatures. The lower baseline is generally obtained through a
 81 linear regression between $(T_c - T_a)$ and vapor pressure deficit for potentially transpiring or well-watered
 82 crops, however, direct observations of T_{c-ww} provide more accurate estimates of CWSI (Yuan et al.
 83 2004). Previous studies have shown that the upper baseline which indicates a non-transpiring crop is
 84 well represented by air temperature plus a constant value (King and Shellie 2018; Adeyemi et al. 2018).

85 As seen above, the theoretical approach involves numerous complex meteorological data to
 86 compute the CWSI baselines. Although, the model has been found to precisely assess the crop water
 87 stress (Yuan et al. 2004; Heydari et al. 2019), its application in commercial crop production is limited
 88 due to requirement of complex input model parameters, particularly crop canopy resistance,
 89 aerodynamic resistance, and net radiation values (Al-Faraj et al. 2001). The empirical approach is
 90 simple to use and gives a reliable indication of crop water stress. It has, however, been shown that the
 91 T_{c-ww} depends on the crop growth and the agro-climate in which it is grown (Kumar et al. 2019). Further,
 92 direct measurements of T_{c-ww} and T_{c-dry} under similar environmental conditions as the T_c are practically
 93 unfeasible due to experimental constraints, as both involve field soil water that is either undesirable (T_{c-}
 94 $_{dry}$) or unattainable (T_{c-ww}) in practical conditions (Kumar et al. 2020a).

95 Artificial reference surfaces for estimating the baseline temperatures have been developed and
 96 successfully used under similar environmental conditions (Agam et al. 2013). These include the use of
 97 well-watered and water-stressed plots, leaves sprayed with water and covered with petroleum jelly and
 98 the application of wet and dry filter papers (Meron et al. 2010; Alchanatis et al. 2010). However, they
 99 require extensive maintenance and intensive data acquisition, which limits their use in precision
 100 irrigation systems (Maes and Steppe 2012). Numerical estimation of the baseline temperatures through
 101 physical models has also been found to give reliable results. Jones (1999) used the leaf energy balance
 102 model to develop the predictive equations for the baseline temperatures. The numerical estimation of

103 the baseline temperature eliminates the need for an artificial reference surface, but it involves
104 measurements of the equation parameters, routine observation of which is not feasible owing to the
105 expensive instrumentation and lack of technical know-how (Park 2018). Hence, estimation of the
106 baseline temperature through parsimonious predictive models using limited climatic data will enhance
107 the utilization of CWSI as a tool for scheduling irrigation and monitoring crop stress (Osroosh et al.
108 2016; Egea et al. 2017).

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

121 On the contrary, unsupervised neural networks, known as Kohonen Self-Organizing Maps
122 (KSOM) (Kohonen 1990; Kohonen et al. 1996) have no specific input or output arguments. KSOM
123 clusters a large dimensional data into a small dimensional map, thus making any inherent correlations
124 between the variables much more visible (Kothari and Islam 1999). The clustering enables effective
125 replacement of the missing values or outliers by their corresponding features in the map, thereby causing
126 no hindrance to the predictions of the model. Due to its versatility, the KSOM has been widely used in
127 hydrological modeling including evapotranspiration modeling (Adeloye et al. 2011), global water flows
128 assessment (Clark et al. 2015), water quality modeling (Rustum and Adeloye 2007; Ramachandran et
129 al. 2019), streamflow forecasting (Mwale et al. 2014), rainfall-runoff modeling (Adeloye and Rustum
130 2012), soil moisture (Riese and Keller 2018), irrigation management (Ohana-Levi et al. 2019) and
131 groundwater studies (Chen et al. 2018).

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

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

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

141 2. Materials and Methods

142 2.1 Agricultural plot and experimental details

143 The study was carried out during the 2017 and 2018 growing seasons at the agricultural experimental
144 station of the National Institute of Technology, Hamirpur, India (altitude: 900 m asl; longitude: 76° 31'
145 33"; latitude: 31° 42' 40"). Field crop experiments were performed on Indian mustard (*Brassica Juncea*)
146 from September to December. The climate of the study area is humid sub-tropical with seasonal mean
147 values of relative humidity, air temperature, solar radiation and wind speed of 74.2 %, 19.10 °C, 0.16
148 kW m⁻², and 1.8 m s⁻¹ respectively. The average seasonal rainfall is 65 mm. The soil in the experimental
149 station had uniform sandy loam texture (silt = 24%, sand = 55% and clay = 21%) up to 1.6 m depth.
150 The permanent wilting point (PWP) and field capacity (FC) of the soil obtained using pressure plate
151 apparatus were 0.07 cm³ cm⁻³ and 0.22 cm³ cm⁻³ respectively. The available soil water (ASW), defined
152 as the difference between FC and PWP, was estimated to be 0.15 cm³ cm⁻³. This is a relatively low
153 ASW which should accelerate the drying up of the soil and hence make the determination of the T_{c-dry}
154 much more rapid. For soils with more water retention capacity, the drying process will be much slower
155 especially during wet periods.

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

162 Figure 1

163 The irrigation treatments were based on a specific level of soil water depletion (SWD) of the
164 ASW in the crop root zone. Treatment T8 was not provided with any supplemental irrigation (except
165 for pre-sowing and one for the crop survival) during the entire crop season. Treatment T1 was provided
166 with frequent irrigations to maintain the water content near the FC. Treatments T8 and T1 were
167 deliberately kept dry and well-watered, to allow the estimation of T_{c-dry} and T_{c-ww} , respectively. The
168 maximum level of SWD allowed in the treatments T2, T3, T4, T5, T6, and T7 was 10%, 20%, 30%,
169 40%, 50% and 60% of ASW, respectively. The soil water was monitored daily using a capacitance
170 probe (*Sentek Sensor Technologies, SA, Australia*), which recorded the volumetric water content
171 (VWC) every 0.1 m interval up to 1.6 m depth. The percentage SWD of ASW in the effective root zone
172 was estimated using the relation $SWD = (FC - VWC)/ASW$. Water was supplied to respective plots
173 with the help of a water hose (surface irrigation) in calculated amounts (water meter installed at the
174 inlet). A tipping bucket rain gauge was used for recording the rainfall.

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

183 **Table 1**

184 **2.2 Canopy temperature and weather monitoring**

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

199 **Table 2**

200 **2.3 Kohonen Self-organizing maps**

201 **2.3.1 Basics of the Kohonen self-organizing maps**

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

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

$$213 \quad M = 5\sqrt{N} \quad (4)$$

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

$$216 \quad \frac{l_1}{l_2} = \sqrt{\frac{e_1}{e_2}} \quad (5)$$

217 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
 218 and second-biggest eigenvalue of the training dataset, respectively.

219 **Figure 2**

220 **2.3.2 Training the KSOM**

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

$$225 \quad D_i = \sqrt{\sum_{j=1}^n m_j (x_j - w_{ij})^2} \quad (6)$$

226 Where, D_i is the Euclidian distance between input vector and code vector i ; n is the dimension of the
 227 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,
 228 whose value is 0 when the given element x_j of the input vector is missing, otherwise it is 1. This becomes
 229 very useful while handling problems involving missing elements because all that needs to be done is to
 230 set the value of m_j for such elements as zero. The neuron for which D_i is minimum is chosen as the
 231 winning node or best matching unit (BMU) as shown in Figure 2. The code vectors of this BMU and
 232 its adjacent neurons are then adjusted to improve the agreement with the input data using Equation (7).

$$233 \quad w_i(t + 1) = w_i(t) + \alpha(t)h_{ci}(t)[x(t) - w_i(t)] \quad (7)$$

234 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
 235 function centered in the winner unit c at time t . In this way, each map unit develops internally the ability
 236 to identify input vectors like itself. This feature is referred to as self-organizing since the classification
 237 is achieved without providing any external output (Penn 2005). The process continues until an optimal
 238 number of iterations is reached or a specific error criterion is attained. The learning effectiveness of the

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

$$241 \quad h_{ci}(t) = \exp\left(-\|r_c - r_i\|^2 / (2\sigma^2(t))\right) \quad (8)$$

$$242 \quad \alpha(t) = \alpha_0 \left(\frac{0.005}{\alpha_0}\right)^{t/T} \quad (9)$$

243 where, T is the training length for convergence, usually taken as equal to $250/\sqrt{N}$ (Vesanto et al. 2000),
 244 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
 245 on the grid, and $\sigma(t)$ is the neighborhood radius. Both $\alpha(t)$ and $\sigma(t)$ decreases monotonically with the
 246 increasing number of iterations.

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

$$249 \quad t_e = \frac{1}{N} \sum_{i=1}^N u(X_i) \quad (10)$$

$$250 \quad q_e = \frac{1}{N} \sum_{i=1}^N \|X_i - W_c\| \quad (11)$$

251 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
 252 vector of the winning node (BMU) for X_i ; $\|\cdot\|$ represents the Euclidian distance (equation (6)), and u is
 253 a binary integer whose value is 1 if the first and second BMU are not adjacent units, otherwise zero.

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

260 **Figure 3**

261 **2.3.3 KSOM modeling**

262 KSOM modeling in the study was performed using the SOM toolbox for MATLAB (Vesanto et al.
 263 2000; Vatanen et al. 2015). The main objective of the study was to develop and evaluate a KSOM model
 264 for estimating the T_{c-ww} . For this purpose, the data of RH, T_a , and T_{c-ww} were used in the modeling. This
 265 was purposely done to evaluate the KSOM model using easily available limited climatic data. The
 266 dataset for model development considered 210 observations of each variable. Similarly, for model
 267 validation, a set of 225 data points were used. Table 2 provides the statistical summary of the
 268 development and validation data sets.

269 To minimize the potential bias of the autocorrelation in the predictive ability of the trained
270 maps, the input vectors of the training dataset were selected randomly and presented to the map in each
271 time step. The validation was crucial to establish the ability of the KSOM model to generalize. The T_{c-ww}
272 $_{ww}$ was omitted from the input vectors during the validation phase, indicating that the T_{c-ww} values were
273 missing. The BMU for each input vector of the validation phase was then determined to predict the
274 missing T_{c-ww} values as illustrated in Figure 3. After obtaining the T_{c-ww} values from the BMU's, they
275 were compared with their actual values for evaluating the performance of KSOM during validation.

276 **2.4 Multiple linear regression**

277 As noted earlier, two more modeling paradigms were considered for the prediction of the T_{c-ww} ,
278 namely multiple linear regression (MLR) and feed-forward neural network (FFNN). The description
279 of MLR is available in any standard statistical textbook and will hence not be repeated here. Details and
280 applications of MLR have been documented by Bottenberg and Ward (1963) and Aiken et al. (2012).
281 The MLR model was implemented using the Data Analysis toolbox in Microsoft Excel. Initially, a
282 regression equation was developed using the dataset of 2017. The regression equation consisted of T_{c-ww}
283 $_{ww}$ as the response variable and T_a and RH as the predictor variables. The equation was then applied to
284 the dataset of 2018 to estimate the T_{c-ww} . The estimated values were then compared with the actual
285 values for evaluating the performance of the MLR model.

286 **2.5 Feed Forward neural network**

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

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

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

314 **Figure 4**

315 **2.6 Statistical evaluation**

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

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

$$321 \quad MBE = \frac{1}{n} \sum_{i=1}^n (x_i - x'_i) \quad (12)$$

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

$$324 \quad MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x'_i| \quad (13)$$

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

$$327 \quad MSE = \frac{1}{n} \sum_{i=1}^n (x_i - x'_i)^2 \quad (14)$$

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

$$330 \quad PE = \left| \frac{\bar{x} - \bar{x}'}{\bar{x}} \right| \times 100 \quad (15)$$

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

$$333 \quad R^2 = \left[\frac{n \sum x x' - \sum x \sum x'}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum x'^2 - (\sum x')^2]}} \right]^2 \quad (16)$$

334 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-ww} during 2017 and 2018 cropping seasons are shown in Figure 5. There was no difference ($p>0.05$) between the measured T_{c-ww} across the three replications, hence, their mean is utilized for indicating the variation. However, a significant difference ($p\leq 0.05$) was observed between the measured T_{c-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 ($\alpha_0 = 0.5$) and neighborhood radius ($\sigma = \max(l_1, l_2)/4$) 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 (Kaltch 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-ww} , RH and T_a .

Visual analysis of the component planes shows that the color (or grey) gradient of the plane for T_{c-ww} is parallel to the gradient of T_a , with high values of T_{c-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-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-ww} is lower than the T_a , confirming the accuracy of the model predictions.

Figure 6

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

380 **Figure 7**

381 **3.3 FFNN modeling results**

382 The feed-forward neural network (FFNN) model architecture was developed using several scenarios
383 based on trial-error and cross-validation. The best-performed model considered two input variables
384 (RH, T_a), one hidden layer (with 5 neurons) and an output variable (T_{c-ww}). Figure 8 shows the X-Y plot
385 of FFNN estimated and measured values of T_{c-ww} , which represent a good correlation with R^2 value of
386 0.90 and 0.92 during development and validation, respectively. The regression line slope during model
387 validation was significantly different ($p\leq 0.05$) from 1-1 line, which indicates a bias in FFNN
388 predictions. Table 3 shows the descriptive summary of the error statistics used in the study for
389 evaluating the performance of the FFNN model. The prediction results were similar to those reported
390 by King and Shellie (2016) who utilized FFNN modeling for estimating T_{c-ww} with four climatic
391 variables.

392 **Figure 8**

393 **3.4 MLR modeling results**

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

$$397 \quad T_{c-ww} = 1.296 + 3.948 \times RH + 0.744 \times T_a \quad (17)$$

398 X-Y plots of MLR estimated and measured values of T_{c-ww} shown in Figure 9 represent a good
399 correlation during development and validation with an R^2 value of 0.91 and 0.93, respectively. The
400 correlation between measured and MLR estimated values was similar to that of the FFNN model. The
401 regression line slope was significantly different ($p\leq 0.05$) from 1-1 line during model validation,

402 indicating a bias of the MLR model in estimating T_{c-ww} . Table 3 presents the error statistics for the
403 performance evaluation of the MLR model.

404 **Figure 9**

405 **Table 3**

406 **3.5 Comparison of KSOM, MLR and FFNN models**

407 Table 3 summarises the error statistics for performance evaluation of KSOM, FFNN and MLR models.
408 A comparison of the error statistics indicates that the performance of KSOM was much better than
409 FFNN and MLR in estimating the T_{c-ww} for Indian mustard. For example, the R^2 for the KSOM model
410 during validation was 0.98, whereas, for the FFNN and MLR models, it was 0.92 and 0.92 respectively.
411 Also, the FFNN and MLR model results were more biased than the KSOM model results during
412 validation, which is indicated by the bias error estimates. Table 3 shows that the errors corresponding
413 to MLR are similar to those of the FFNN model. A similar observation was reported by King and Shellie
414 (2016).

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

423 **Figure 10**

424 **3.6 Crop water stress index**

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

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

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

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

455 **Figure 11**

456 **Figure 12**

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

467 **4. Conclusion**

468 The current work presents a novel approach involving the application of Kohonen Self-Organising Map
469 (KSOM) in estimating the well-watered canopy temperature (T_{c-ww}) for computing the crop water stress

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

486 **Funding**

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

492 **Acknowledgments**

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

498 **Data Availability Statement**

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

501 **Software Availability Statement**

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

504 **References**

- 505 Adeloje, A. J., & Rustum, R. (2012). Self-organising map rainfall-runoff multivariate modelling for
506 runoff reconstruction in inadequately gauged basins. *Hydrology Research*, 43(5), 603-617.
- 507 Adeloje, A. J., Rustum, R., & Kariyama, I. D. (2011). Kohonen self-organizing map estimator for the
508 reference crop evapotranspiration. *Water Resources Research*, 47(8).
- 509 Adeloje, A. J., Rustum, R., & Kariyama, I. D. (2012). Neural computing modeling of the reference
510 crop evapotranspiration. *Environmental Modelling & Software*, 29(1), 61-73.
- 511 Adeyemi, O., Grove, I., Peets, S., Domun, Y., & Norton, T. (2018). Dynamic modelling of the baseline
512 temperatures for computation of the crop water stress index (CWSI) of a greenhouse cultivated lettuce
513 crop. *Computers and Electronics in Agriculture*, 153, 102-114.
- 514 Agam, N., Cohen, Y., Alchanatis, V., & Ben-Gal, A. (2013). How sensitive is the CWSI to changes in
515 solar radiation?. *International journal of remote sensing*, 34(17), 6109-6120.
- 516 Aiken, L. S., West, S. G., Pitts, S. C., Baraldi, A. N., & Wurpts, I. C. (2012). Multiple linear
517 regression. *Handbook of Psychology, Second Edition*, 2.
- 518 Akkuzu, E., Kaya, Ü., Çamoğlu, G., Mengü, G. P., & Aşık, Ş. (2013). Determination of crop water
519 stress index and irrigation timing on olive trees using a handheld infrared thermometer. *Journal of*
520 *irrigation and drainage engineering*, 139(9), 728-737.
- 521 Alchanatis, V., Cohen, Y., Cohen, S., Moller, M., Sprinstin, M., Meron, M., ... & Sela, E. (2010).
522 Evaluation of different approaches for estimating and mapping crop water status in cotton with thermal
523 imaging. *Precision Agriculture*, 11(1), 27-41.
- 524 Al-Faraj, A., Meyer, G. E., & Horst, G. L. (2001). A crop water stress index for tall fescue (*Festuca*
525 *arundinacea* Schreb.) irrigation decision-making—a fuzzy logic method. *Computers and electronics in*
526 *agriculture*, 32(2), 69-84.
- 527 Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for
528 computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome*, 300(9),
529 D05109.
- 530 Anda, A., Soós, G., Menyhárt, L., Kucserka, T., & Simon, B. (2020). Yield features of two soybean
531 varieties under different water supplies and field conditions. *Field Crops Research*, 245, 107673.
- 532 ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000). Artificial
533 neural networks in hydrology. II: Hydrologic applications. *Journal of Hydrologic Engineering*, 5(2),
534 124-137.
- 535 Bellvert, J., Marsal, J., Girona, J., Gonzalez-Dugo, V., Fereres, E., Ustin, S. L., & Zarco-Tejada, P. J.
536 (2016). Airborne thermal imagery to detect the seasonal evolution of crop water status in peach,
537 nectarine and Saturn peach orchards. *Remote Sensing*, 8(1), 39.
- 538 Boomiraj, K., Chakrabarti, B., Aggarwal, P. K., Choudhary, R., & Chander, S. (2010). Assessing the
539 vulnerability of Indian mustard to climate change. *Agriculture, ecosystems & environment*, 138(3-4),
540 265-273.
- 541 Bottenberg, R. A., & Ward, J. H. (1963). *Applied multiple linear regression* (Vol. 63, No. 6). 6570th
542 Personnel Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Lackland
543 Air Force Base.

544 Bowden, G. J., Dandy, G. C., & Maier, H. R. (2005). Input determination for neural network models in
545 water resources applications. Part 1—background and methodology. *Journal of Hydrology*, 301(1-4),
546 75-92.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

590 Kalteh, A. M., Hjorth, P., & Berndtsson, R. (2008). Review of the self-organizing map (SOM) approach
591 in water resources: Analysis, modelling and application. *Environmental Modelling & Software*, 23(7),
592 835-845.

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

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

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

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

602 Kothari, R., & Islam, S. (1999). Spatial characterization of remotely sensed soil moisture data using
603 self organizing feature maps. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2), 1162-
604 1165.

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

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

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

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

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

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

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

- 624 Ohana-Levi, N., Bahat, I., Peeters, A., Shtein, A., Netzer, Y., Cohen, Y., & Ben-Gal, A. (2019). A
625 weighted multivariate spatial clustering model to determine irrigation management zones. *Computers
626 and Electronics in Agriculture, 162*, 719-731.
- 627 O'shaughnessy, S. A., Evett, S. R., Colaizzi, P. D., & Howell, T. A. (2011). Using radiation
628 thermography and thermometry to evaluate crop water stress in soybean and cotton. *Agricultural Water
629 Management, 98*(10), 1523-1535.
- 630 Osroosh, Y., Peters, R. T., Campbell, C. S., & Zhang, Q. (2015). Automatic irrigation scheduling of
631 apple trees using theoretical crop water stress index with an innovative dynamic threshold. *Computers
632 and Electronics in Agriculture, 118*, 193-203.
- 633 Osroosh, Y., Peters, R. T., Campbell, C. S., & Zhang, Q. (2016). Comparison of irrigation automation
634 algorithms for drip-irrigated apple trees. *Computers and Electronics in Agriculture, 128*, 87-99.
- 635 Park, S., 2018. *Estimating plant water stress and evapotranspiration using very-high-resolution (VHR)
636 UAV imagery* (Doctoral dissertation).
- 637 Payero, J. O., & Irmak, S. (2006). Variable upper and lower crop water stress index baselines for corn
638 and soybean. *Irrigation Science, 25*(1), 21-32.
- 639 Penn, B. S. (2005). Using self-organizing maps to visualize high-dimensional data. *Computers &
640 Geosciences, 31*(5), 531-544.
- 641 Poirier-Pocovi, M. and Bailey, B.N., 2020. Sensitivity analysis of four crop water stress indices to
642 ambient environmental conditions and stomatal conductance. *Scientia Horticulturae, 259*, p.108825.
- 643 Ramachandran, A., Rustum, R., & Adeloye, A. J. (2019). Anaerobic digestion process modeling using
644 Kohonen self-organising maps. *Heliyon, 5*(4), e01511.
- 645 Rana, K., Parihar, M., Singh, J.P. and Singh, R.K., 2019. Effect of sulfur fertilization, varieties and
646 irrigation scheduling on growth, yield, and heat utilization efficiency of indian mustard (Brassica Juncea
647 L.). *Communications in Soil Science and Plant Analysis*, pp.1-11.
- 648 Riese, F. M., & Keller, S. (2018, July). Introducing a framework of self-organizing maps for regression
649 of soil moisture with hyperspectral data. In *IGARSS 2018-2018 IEEE International Geoscience and
650 Remote Sensing Symposium* (pp. 6151-6154). IEEE.
- 651 Romero-Trigueros, C., Bayona Gambín, J.M., Nortes Tortosa, P.A., Alarcón Cabañero, J.J. and Nicolás
652 Nicolás, E., 2019. Determination of Crop Water Stress Index by Infrared Thermometry in Grapefruit
653 Trees Irrigated with Saline Reclaimed Water Combined with Deficit Irrigation. *Remote Sensing, 11*(7),
654 p.757.
- 655 Rustum, R. (2009). *Modelling activated sludge wastewater treatment plants using artificial intelligence
656 techniques (fuzzy logic and neural networks)* (Doctoral dissertation, Heriot-Watt University).
- 657 Rustum, R., & Adeloye, A. J. (2007). Replacing outliers and missing values from activated sludge data
658 using Kohonen self-organizing map. *Journal of Environmental Engineering, 133*(9), 909-916.
- 659 Shekhawat, K., Rathore, S. S., Premi, O. P., Kandpal, B. K., & Chauhan, J. S. (2012). Advances in
660 agronomic management of Indian mustard (Brassica Juncea (L.) Czernj. Cosson): an
661 overview. *International Journal of Agronomy, 2012*. doi:10.1155/2012/408284
- 662 Singh, P.K. and Singh, A.K., 2019. Effect of sowing dates and irrigation schedules on performance of
663 Indian mustard (Brassica juncea L.) and water use efficiency. *Journal of Soil and Water Conservation,
664 18*(2), pp.164-167.

- 665 Singh, V.V., Garg, P., Meena, H.S., & Meena, M.L. (2018). Drought Stress Response of Indian Mustard
666 (*Brassica juncea* L.) Genotypes. *Int. J. Curr. Microbiol. App. Sci.* 7(3), 2519-2526.
- 667 Stefanovic, P., & Kurasova, O. (2011). Visual analysis of self-organizing maps. *Nonlinear Analysis:*
668 *Modelling and Control*, 16(4), 488-504.
- 669 Vatanen, T., Osmala, M., Raiko, T., Lagus, K., Sysi-Aho, M., Orešič, M., ... & Lähdesmäki, H. (2015).
670 Self-organization and missing values in SOM and GTM. *Neurocomputing*, 147, 60-70.
- 671 Vesanto, J., Himberg, J., Alhoniemi, E., & Parhankangas, J. (2000). SOM toolbox for Matlab
672 5. *Helsinki University of Technology, Finland*, 109.
- 673 Wright, P. R., Morgan, J. M., & Jessop, R. S. (1996). Comparative adaptation of canola (*Brassica napus*)
674 and Indian mustard (*B. juncea*) to soil water deficits: plant water relations and growth. *Field Crops*
675 *Research*, 49(1), 51-64.
- 676 Yildirim, M., Demirel, K., & Bahar, E. (2012). Effect of restricted water supply and stress development
677 on growth of bell pepper (*Capsicum Annuum* L.) under drought conditions. *Journal of Agro Crop*
678 *Science*, 3(1), 1-9.
- 679 Yuan, G., Luo, Y., Sun, X., & Tang, D. (2004). Evaluation of a crop water stress index for detecting
680 water stress in winter wheat in the North China Plain. *Agricultural Water Management*, 64(1), 29-40.

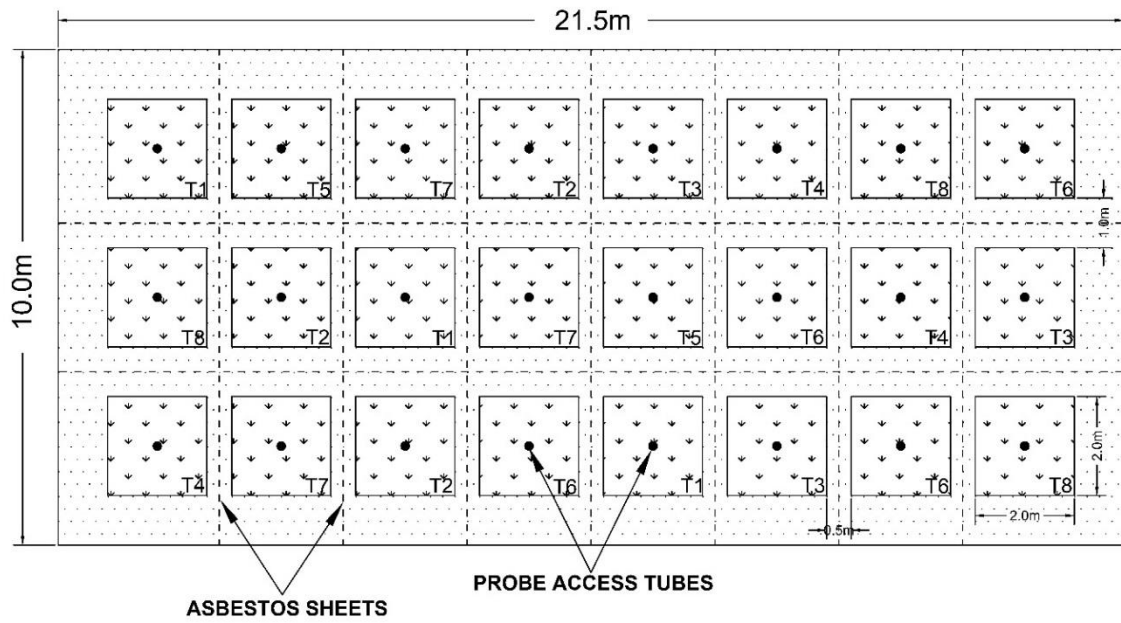


Figure 1. The layout of the experimental plot (T1- Well-watered plot; T2-10% SWD; T3-20% SWD; T4-30% SWD; T5-40% SWD; T6-50% SWD; T7-60% SWD; and T8-Maximum stressed plot.)

681

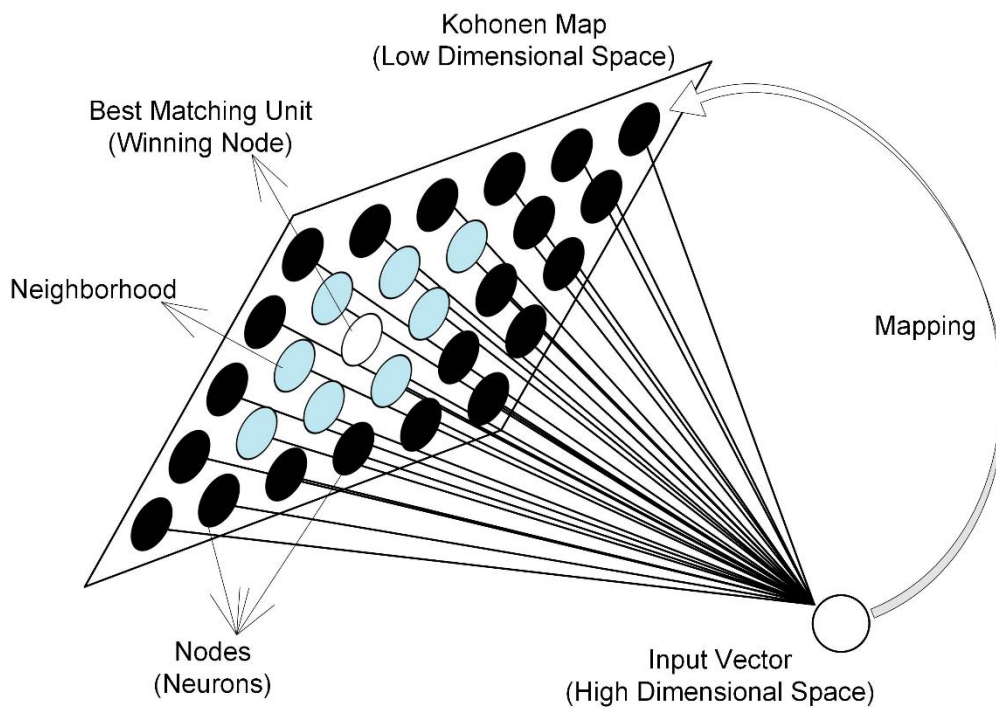


Figure 2. Representation of the winning node and its neighbors in a KSOM

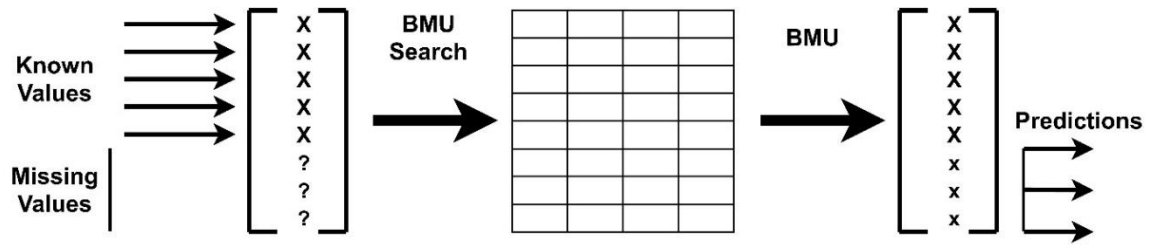


Figure 3. Prediction of the missing component of the input vector using the Kohonen Self Organizing Map.

683

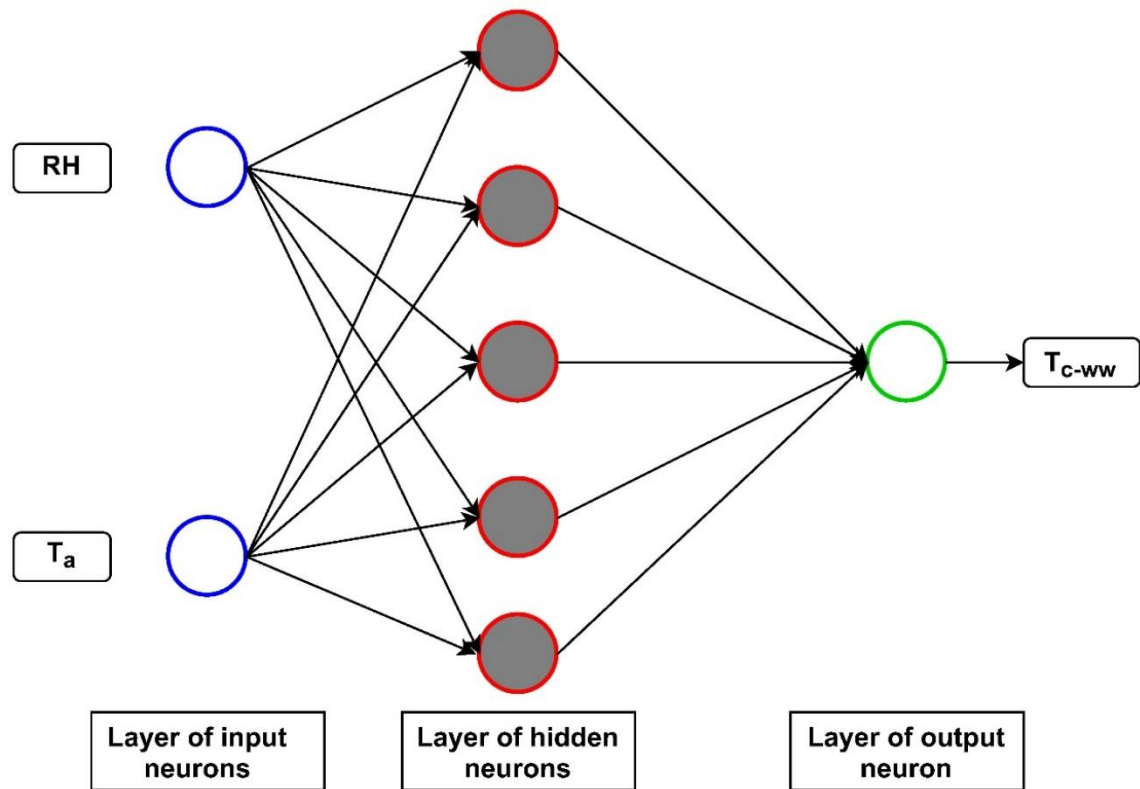


Figure 4. Schematic representation of the feed-forward neural network modeling architecture with two inputs and one hidden layer. RH – Relative humidity, T_a – Air temperature and T_{c-ww} – Well-watered canopy temperature.

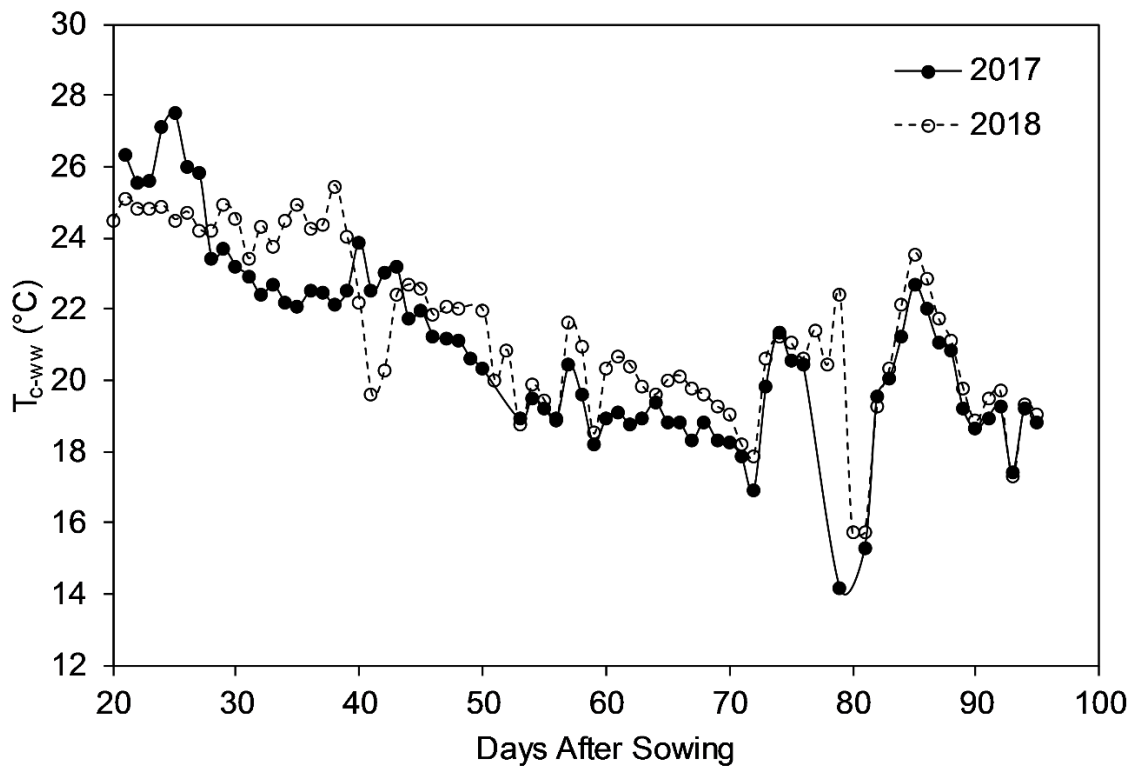


Figure 5. Time series plot of well-watered canopy temperature of Indian mustard during the crop period

685

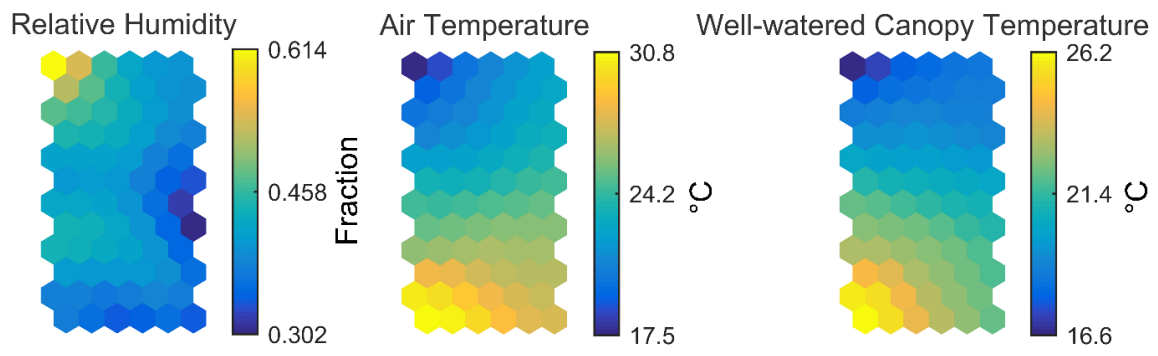


Figure 6. KSOM component planes.

686

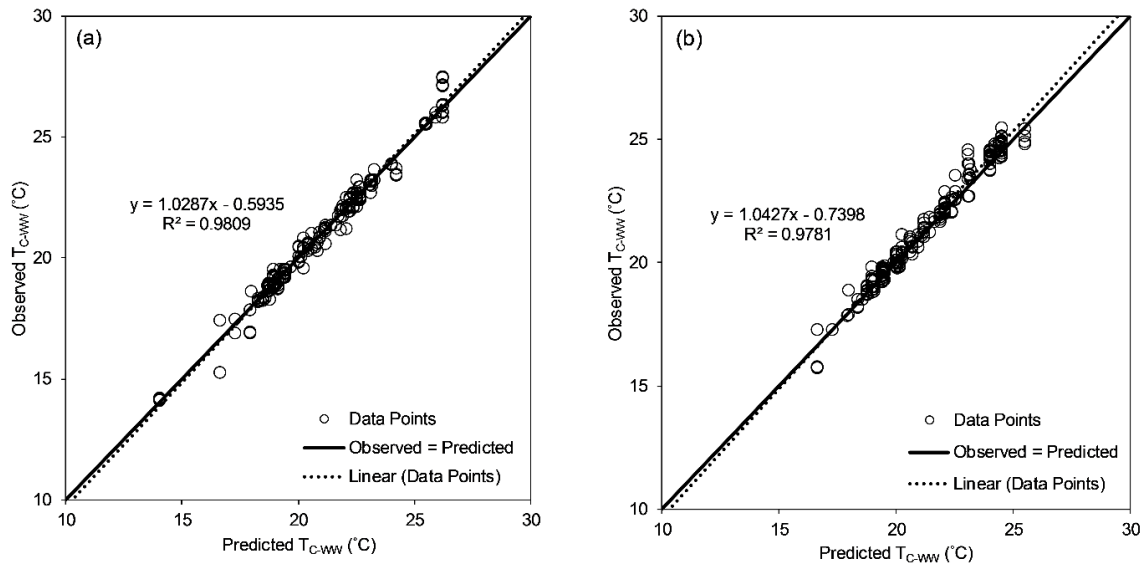


Figure 7. X-Y scatter plot of KSOM predicted and measured values of T_{c-ww} during (a) model development, and (b) model validation

687

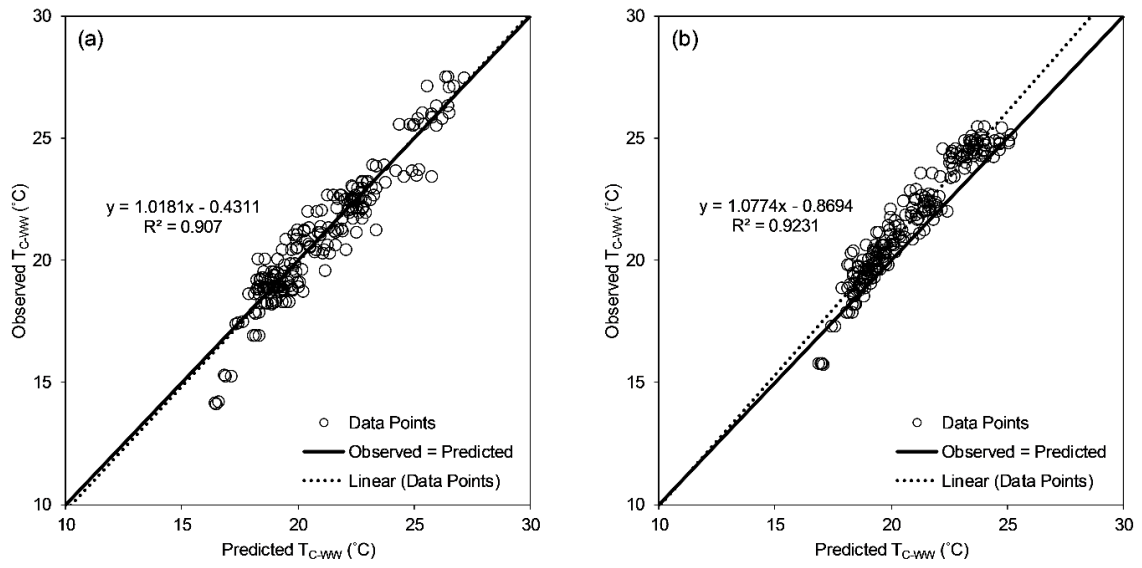


Figure 8. X-Y scatter plot of FFNN predicted and measured values of T_{c-ww} during (a) model development, and (b) model validation

688

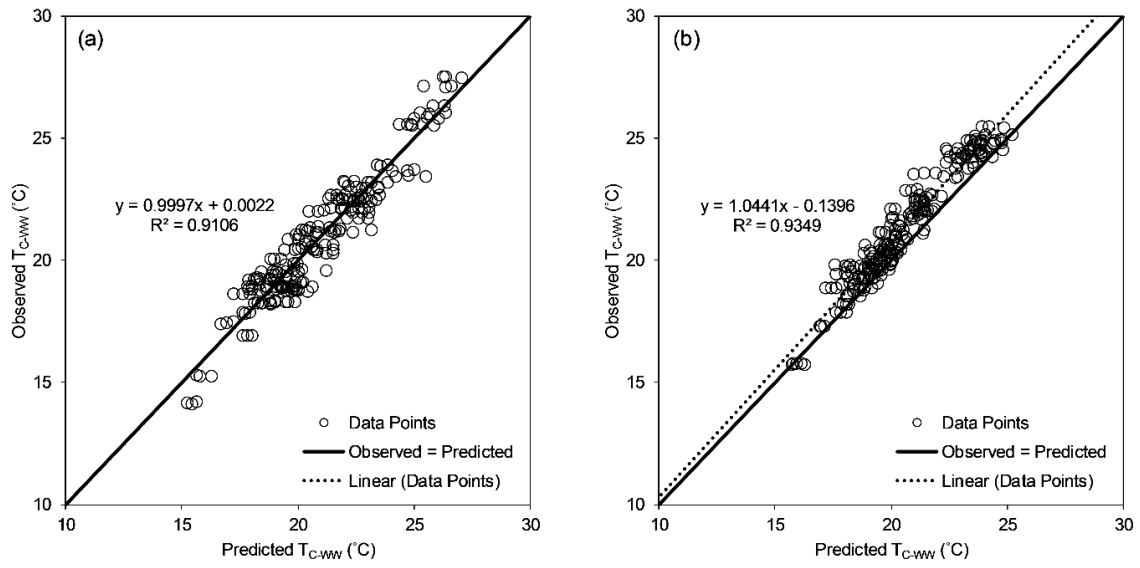


Figure 9. X-Y scatter plot of MLR predicted and measured values of T_{c-ww} during (a) model development, and (b) model validation

689

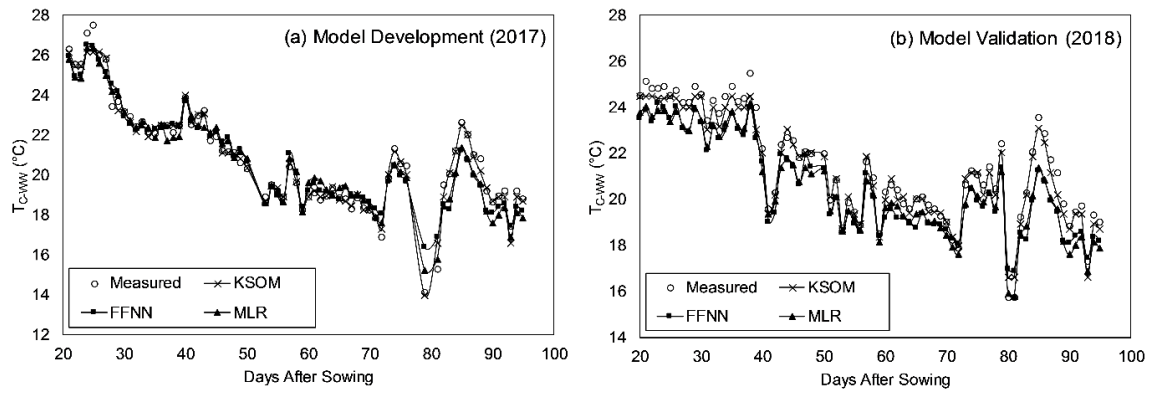


Figure 10. Time series plot of measured and predicted well-watered canopy temperature of Indian mustard during the crop period for the growing season (a) 2017 and (b) 2018

690

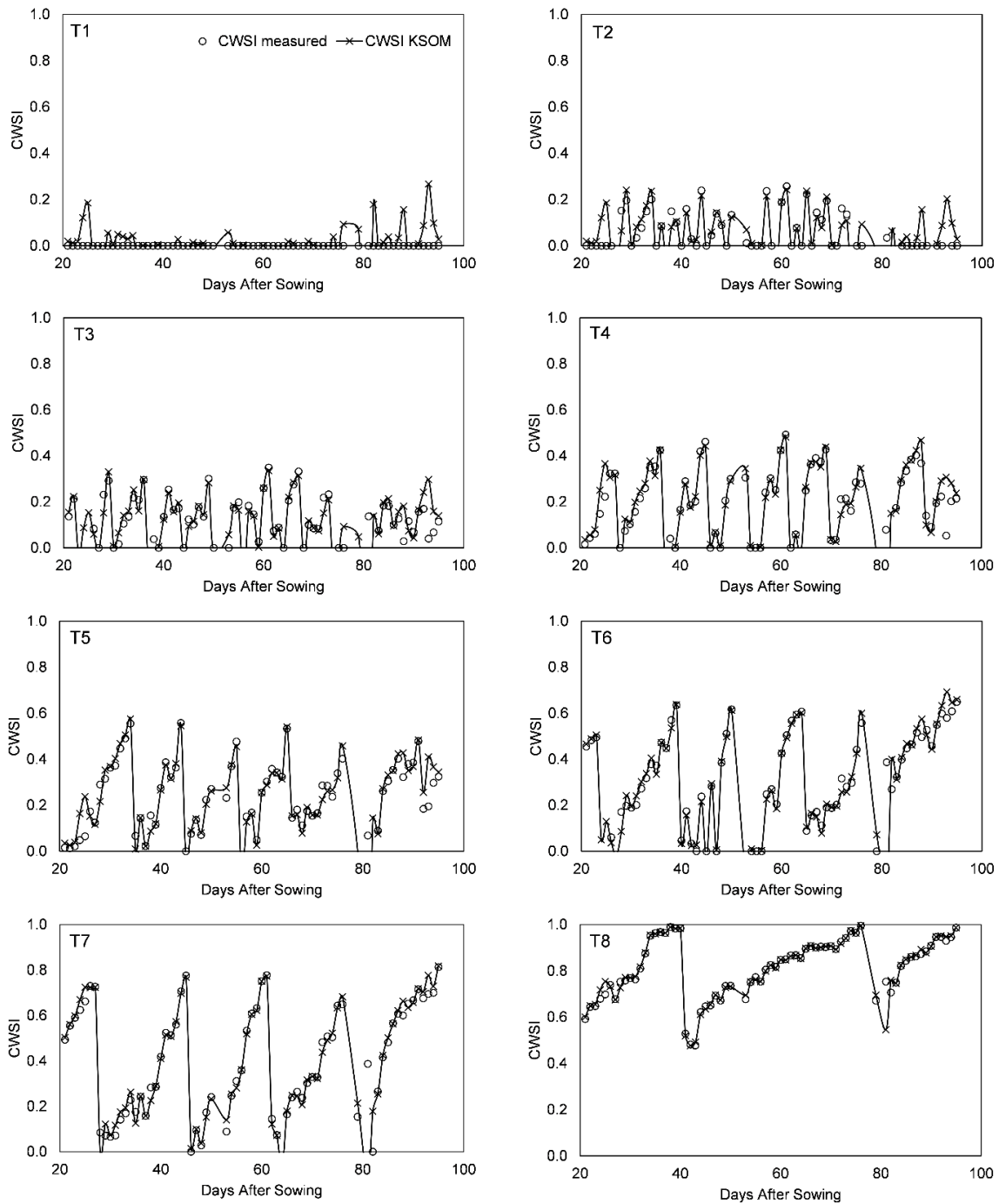


Figure 11. Comparison between observed CWSI (based on measured values of T_{c-ww}) and predicted CWSI (based on KSOM estimated values of T_{c-ww}) for different irrigation treatments during model development (2017)

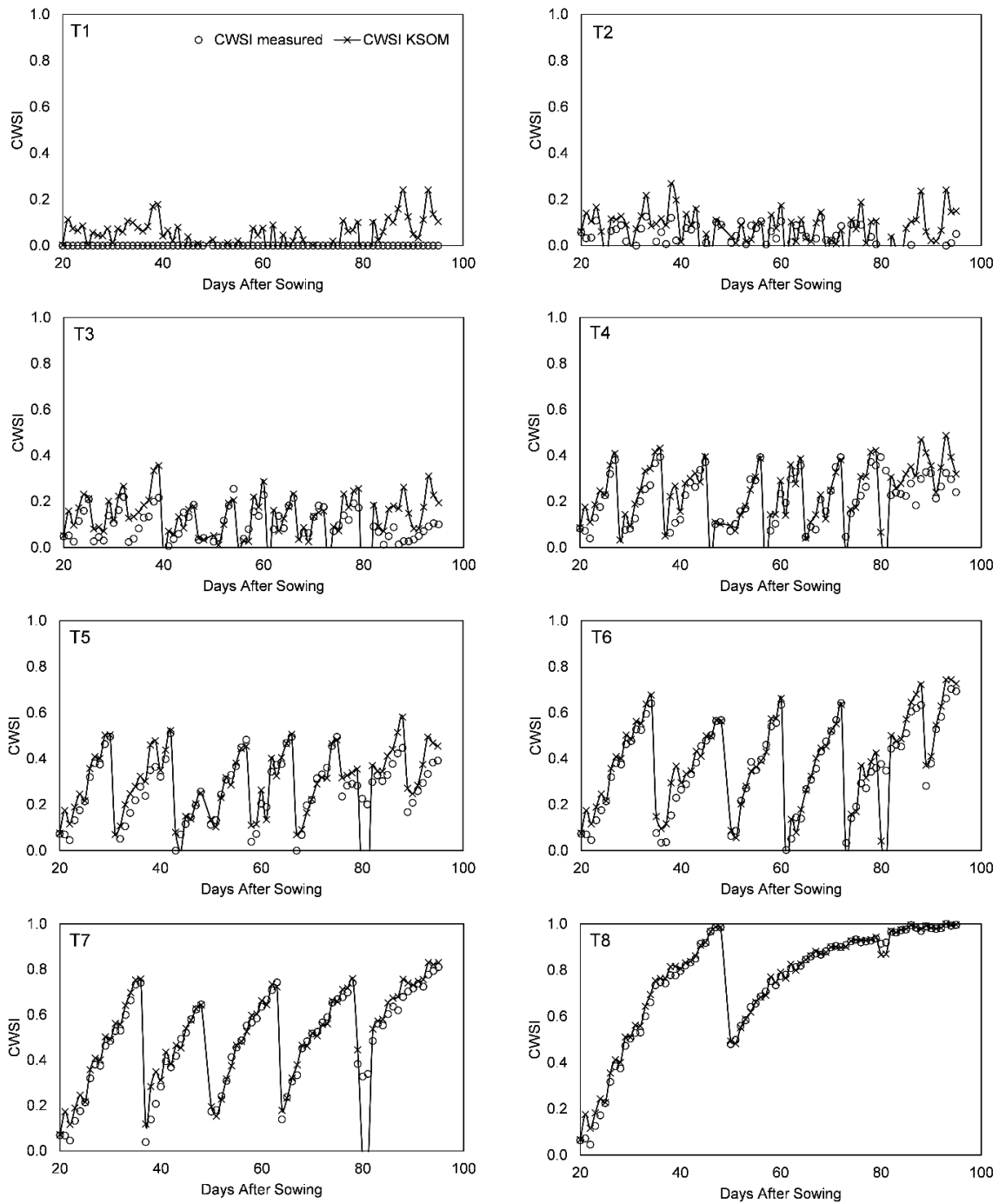


Figure 12. Comparison between observed CWSI (based on measured values of T_{c-wv}) and predicted CWSI (based on KSOM estimated values of T_{c-wv}) for different irrigation treatments during model validation (2018)

693

694 **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 (<i>Brassica Juncea</i>)	P.T. 303	95	20	25	30	20	40 × 15	22 nd September 2017	25 th December 2017
		95	20	25	30	20	40 × 15	25 th September 2018	28 th December 2018

695
696

* 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).

697 **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-ww}	Development	27.51	14.13	20.91	1.64
			Validation	25.49	15.73	21.41	2.35

698

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									