Hydrol. Earth Syst. Sci., 26, 4741–4756, 2022
https://doi.org/10.5194/hess-26-4741-2022

Hydrology and Earth System Sciences
Open Access
EGU

Pan evaporation is increased by submerged macrophytes

Brigitta Simon-Gáspár, Gábor Soós, and Angela Anda

Institute of Agronomy, Georgikon Campus, Hungarian University of Agriculture and Life Sciences, 8360 Keszthely, Hungary

Correspondence: Brigitta Simon-Gáspár (simon.gaspar.brigitta@uni-mate.hu)

Received: 18 November 2021 – Discussion started: 1 December 2021
Revised: 4 August 2022 – Accepted: 25 August 2022 – Published: 28 September 2022

Abstract. The topic of evaporation estimates is fundamental to land-surface hydrology. In this study, FAO-56 Penman–Monteith equation (FAO56–PM), multiple stepwise regression (MLR), and Kohonen self-organising map (K–SOM) techniques were used for the estimation of daily pan evaporation (E_p) in three treatments, where C was the standard class A pan with top water, S was a pan with sediment covered bottom, and SM was class A pan containing submerged macrophytes (*Myriophyllum spicatum*, *Potamogeton perfoliatus*, and *Najas marina*), at Keszthely, Hungary, in a six-season experiment, between 2015 and 2020. The modelling approach included six measured meteorological variables. Average E_p varied from 0.6 to 6.9 mm d^{-1} for C, 0.7 to 7.9 mm d^{-1} for S, and from 0.9 to 8.2 mm d^{-1} for SM during the growing seasons studied. Correlation analysis and K–SOM visual representation revealed that air temperature and global radiation had positive correlation, while relative humidity had a negative correlation with the E_p of C, S, and SM. The results showed that the MLR method provided close compliance ($R^2 = 0.58$–0.62) with the observed pan evaporation values, but the K–SOM method ($R^2 = 0.97$–0.98) yielded by far the closest match to observed evaporation estimates for all three pans.

To our best knowledge, no similar work has been published previously using the three modelling methods for seeded pan evaporation estimation.

The current study differs from previous evaporation estimates by using neural networks even with those pans containing sediments and submerged macrophytes. Their evaporation will be treated directly by K–SOM, in which the modelling is more than the simple E_p of a class A pan filled with clean tap water.

1 Introduction

Open water evaporation is one of the paramount elements of the hydrological cycle (Brutsaert, 1982). Evaporation losses from various surfaces appear to be increasing in recent decades (Mbangiwa et al., 2019). Due to climate change, it is also extremely important to determine evaporation as accurately as possible (Fournier et al., 2021), for which both direct and indirect methods are available. As a direct method, the evaporation pans (primarily the class A pan proposed by the World Meteorological Organization, WMO, are used extensively throughout the world to measure open water evaporation and to estimate reference evapotranspiration (Rahimikhoob, 2009; Fuentes et al., 2020). Measurements of pan evaporation may be spatially and temporally limited (Jensen et el., 1990; Rahimikhoob, 2009), like in case of maintenance problems which can affect the accuracy of evaporation measurements, e.g. most often turbidity of water, or watering of birds or other animals (Tabari et al., 2010).

To indirectly determine evaporation, several methods can be used: empirical equations are applied that estimate evaporation based on meteorological variables (air temperature, T_a; relative humidity, RH; global radiation, R_s), or transfer and water budget methods (Burman, 1976). The most widely used empirical formula is a FAO-56 Penman–Monteith equation (FAO56–PM) (Allen et al., 1998), which is the standard method for computation of daily reference evapotranspiration. However, measuring meteorological variables requires sophisticated instruments, which can often be challenging (Arunkumar and Jothiprakash, 2013; Sattari et al., 2020). The amount of required data and the difficulty of the estimation of the unknown meteorological elements may be additional problems (Sanikhani et al., 2015; Khatibi et al., 2020). Therefore, there is a need for alternative methods that are simple and effective, require fewer inputs, and are also

able to solve problems which are difficult to formalise (Sudheer et al., 2003; Kisi, 2015; Malik et al., 2020a).

A promising tool that can be used to estimate E_p and is a suitable alternative to the empirical models is different neural networks (Kim et al., 2015), thus neural networks are increasingly used in evaporation and evapotranspiration estimation (Kumar et al., 2002; Keskin and Terzi 2006; Rahimikhoob, 2009; Alsumaiei, 2020). The machine learning techniques can map high-dimensional data to a low-dimensional space and show some similar properties based on internal data relationships (Pearce et al., 2011; Zelazny et al., 2011). In recent years, machine learning techniques have been broadly employed in hydrological and environmental models, including to forecast evaporation (Wu et al., 2020). Numerous results in the literature indicate that machine learning algorithms such as artificial neural network (ANN), M5 model tree (M5T), support vector machines (SVMs), multivariate adaptive regression splines (MARS), gradient boosting with categorical features support (CatBoost), and random forest (RF) perform excellently in predicting pan evaporation as well (Dong et al., 2021).

Of the available methods, self-organising maps (SOMs) are able to handle noisy, irregular, and multivariate data well (Nakagawa, 2017). As a result, it has become one of the most popular neural network (NN) methods for data analysis (Nada et al., 2017). SOMs are used in many disciplines (Nakagawa 2017, 2020), such as agriculture (Li et al., 2019; Kumar et al., 2021a, b), ecology (Bedoya et al., 2009, Ristić et al., 2020), hydrology (Guntu et al., 2020; Rivas-Tabares, 2020; Lee and Kim, 2021), meteorology (Nada et al., 2017; Berkovic et al., 2021, Doan et al., 2021), and water management (Gu et al., 2019, Gholami et al., 2020; Lee et al., 2021). The unsupervised NNs, including Kohonen self-organising maps (K–SOMs), have several advantages (Kohonen, 1982, 2001). The essence of this method is to group the large-dimensional array of the input layer into a 2-dimensional array in the output layer, so that all variables of the input vectors can be found in each node of the output layer (Adeloye et al., 2011). Another advantage of K–SOM over traditional models is that it also has visualisation abilities (Hadjisolomou et al., 2018).

The study site, Lake Balaton, is the largest shallow freshwater lake in Central Europe with a surface area of 596 km^2 (Fig. 1). The three most dominant submerged macrophytes in Lake Balaton are *Potamogeton perfoliatus*, *Myriophyllum spicatum*, and *Najas marina*, therefore it was appropriate to include these three species in the observation. In Hungary, submerged macrophytes colonise in lakes in the summer season (from June to September). Evaporation of open water surfaces is usually measured by means of pans endowed with unrealistic properties. These pans are filled with clean tap water and the evaporated water is also replaced with tap water unlike in natural ecosystems. In nature, there may also be submerged macrophytes living in the open water. The presence of these plants is essential, and affects the chemical and physical water properties including its quality (Kimmel and Groeger, 1984; Zhang et al., 2017; Yan et al., 2019). Furthermore, the species that are rooted in the sediment can stabilise the sediment by inhibiting its resuspension (Madsen and Cedergreen, 2002; Vymazal, 2013).

Changes in the heat regime of a water body had been reported to result in alterations of macrophyte community composition (Barko et al., 1982; Poikane et al., 2015; Fritz et al., 2017; Kim and Nishihiro, 2020), which may affect the temporal appearance and spatial distribution of macrophytes in the future. As a result, due to global climate change, it is important to examine submerged macrophytes in all aspects, including their effect on evaporation.

The aim of the study was to investigate the effect of littoral sediment and macrophytes on lake evaporation, and not an introduction of a new method in pan evaporation estimation. The previous results in FAO-56 Penman–Monteith equation (Allen et al., 1998), Kohonen self-organising map techniques (Kohonen, 1982), and multiple stepwise regression are classic methods, highlighted widely by citations in the study. They are the tools in analysing the effect of sediment and macrophytes in pan (lake) evaporation estimation only. The novelty of the paper is in the way the evaporation estimation is carried out.

To our best knowledge, there are no studies attempting to project water bodies' evaporation using traditional A pan measurements, taking the macrophytes- and sediment-related factors into account under such climate conditions as our experimental site.

2 Materials and methods

2.1 Case study and data description

The climate of the region – see also Fig. 1 – is mild continental (Cfb) with warm, dry summers and fairly cold winters according to the Köppen–Geiger classification (Kottek, Grieser, Beck, Rudolf and Rubel, 2006). Months were included in the study (from June to September). Meteorological variables were recorded by a QLC-50 climate station (Vaisala, Helsinki, Finland) fitted with a CM-3 pyranometer (Kipp & Zonen Corp., Delft, the Netherlands) located at Keszthely agrometeorological research station (ARS) (latitude: 46°44′ N, longitude: 17°14′ E, elevation: 124 m a.s.l.) between 2015–2020. The ARS is placed on the area of the Hungarian University of Agriculture and Life Sciences. With the exception of wind speed, meteorological data of T_a, RH, R_s, daily maximum temperature (T_{max}), daily minimum temperature (T_{min}), and precipitation (P) were measured at 2 m above the ground surface. The height of wind speed (u) measurements was 10.5 m. The daily mean values of meteorological variables were calculated as average of 10 min observations of a 24 h period.

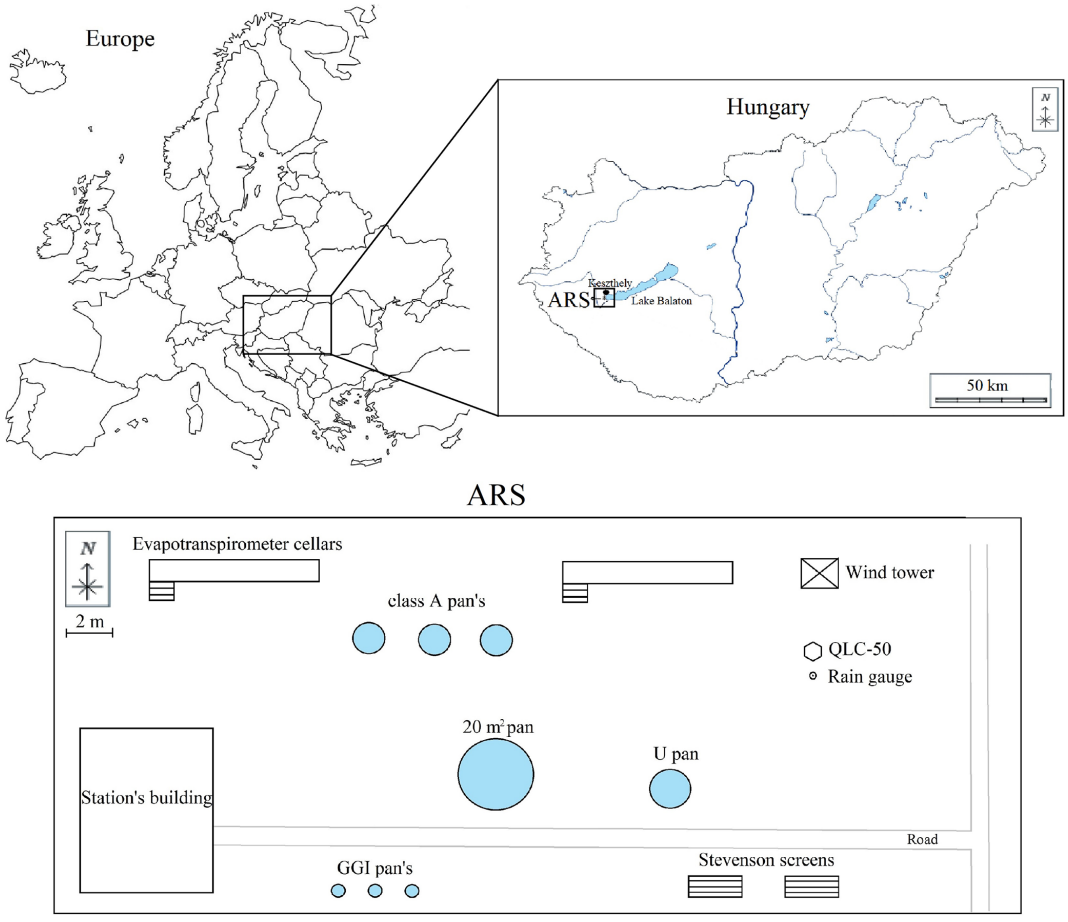

Figure 1. Location map of the study area with agrometeorological research station (ARS) at Keszthely, Hungary (from https://www.vectorstock.com, last access: 3 November 2021).

In this study, class A evaporation pans were used to determine daily evaporation (E_p). The class A pans were 1.21 m in diameter and 0.25 m in height located on an elevated ($\sim$ 0.15 m) wooden grid, with a water surface area of $\sim 1.15\ m^2$. The daily rate of E_p was calculated from the difference in water level for two consecutive days, considering any precipitation that may have fallen into the pans. The daily water loss was measured every morning at 07:00 LMT (local mean time).

In the ARS area, three class A pans were placed 5 m apart (Fig. 2). A class A pan was recommended by the WMO to be used as a standard treatment (control, C). Two class A pans were covered on the bottom with sediment to a thickness of 0.002 m (S). The used sediment was psammal/psammopelal (Ø $>$ 6 µm–2 mm, sand/sand with mud) with the following composition: quartz, calcite, aragonite, dolomite, muscovite, chlorite, feldspar, smectite, kaolinite, and pyrite (Anda et al., 2016). Submerged, freshwater aquatic macrophytes were planted in third class A pans with sediment-covered bottom (Anda et al., 2016, 2018). Macrophyte samples were gathered from Lake Balaton (Keszthely Bay) with similar water depth (0.6–0.8 m) each year. The amount of crop density was controlled monthly without variation in the green mass weight of crops between natural habitat and "seeded" class A pans. In the experimental area, three species of submerged, freshwater aquatic macrophytes: *Potamogeton perfoliatus*, *Myriophyllum spicatum*, and *Najas marina* were colonised. Due to the development of submersed macrophytes, class A pans were operational from June to September in the growing season 2015–2020.

In the last vegetation period, to detect vertical water temperature (T_w) profiles, four fastened thermistors of Delta Ohm HD-226-1 (accuracy: 0.3 °C) collected the temperature data at 0.05, 0.10, and 0.15 m depth from the pan bottom and on the water surface, at 10 min intervals. Hourly averaged T_w values were used in the analysis. To present diurnal variation in T_w and stratification, sample days were selected for clear-sky, calm, and cloudy weather conditions.

The weather of the studied growing seasons was specified by the monthly Thornthwaite index (TI) of the World Mete-

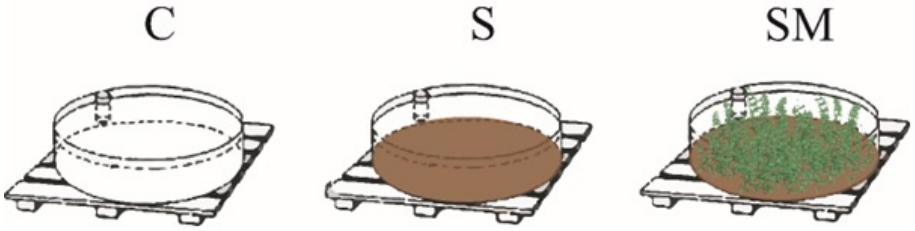

Figure 2. Class A pans with different treatments: C, S, and SM denote "empty", sediment-covered, and macrophyte-planted class A pans in the middle of the meteorological garden.

orological Organization (WMO) report (1975):

$$\mathrm{TI} = 1.65\left(\frac{P_\mathrm{m}}{T_\mathrm{am}} + 12.2\right)^{10/9}, \quad (1)$$

where P_m and T_am are the monthly sum of precipitation and the monthly mean air temperature, respectively.

In classifying the weather in each season's months, a 20 % deviation was assumed from climate norms (1981–2010), above and below the $\mathrm{TI}_\mathrm{norm}$ for both included meteorological variables (P_m and T_am), allowing the following weather classes to be distinguished:

- warm–dry month (h): $\mathrm{TI}_\mathrm{month} > \mathrm{TI}_\mathrm{norm} \times 0.8$;
- cooler–wet month (c): $\mathrm{TI}_\mathrm{month} > \mathrm{TI}_\mathrm{norm} \times 1.2$;
- month with normal weather (n): $\mathrm{TI}_\mathrm{norm} \times 0.8 \leq \mathrm{TI}_\mathrm{month} \leq \mathrm{TI}_\mathrm{norm} \times 1.2$.

By counting the highest number of months within each of these three groups, the season was considered to be either normal, cool, or warm.

2.2 Multiple stepwise regression (MLR)

The regression models are important tools for investigating relations between dependent and independent data (Razi and Athappilly, 2005), which is a method that has been used for a long time in the investigation of meteorological variables. Evaporation can be modelled by multiple linear regressions using different meteorological variables (e.g. T_a, RH, u) (Almedeij, 2012).

The MLR can be expressed by the following equation:

$$y = b_0 + b_1 x_1 + \ldots + b_k x_k + a, \quad (2)$$

where b_0, $b_1 \ldots$, and b_k are fitting constant, $x_1 \ldots$ and x_k represent the observed meteorological variables, and a is a random error term. The a is the remaining effects on estimated E_p (y) of variables not explicitly included in the model (Patle et al., 2020). The dependent variable, y, was E_p.

2.3 FAO-56 Penman–Monteith (FAO56–PM) method

The Penman–Monteith model is considered as the international standard for computing potential evapotranspiration and predicting crop water requirement. FAO56–PM may also be proper method to get pan evaporation with submerged macrophytes. Wang et al. (2021) reported that actual evaporation is important for hydrological research due to its direct impact on the hydrologic processes (water cycle, water resources management). The above authors concluded that to estimate pan evaporation, it is essential to find the proper formulation of the Penman–Monteith equation, a special case of the multiple stepwise regression methods. It may be especially true even in pans with seeded macrophytes. In accordance with composition of lake ecosystems, this is the method in evaporation estimation that implies living organism.

The reference evapotranspiration ET_0 was estimated by the WMO standardised FAO-56 Penman–Monteith method (Allen et al., 1998, 2005) at a daily step for short reference crops (clipped grass of 12 cm) as follows:

$$\mathrm{ET}_0 = \frac{0.408\Delta\,(R_\mathrm{n} - G) + \gamma\,\frac{900}{T_\mathrm{a}+273}u(e_\mathrm{s} - e_\mathrm{a})}{\Delta + \gamma(1 + 0.34u)}, \quad (3)$$

where R_n is net radiation (MJ m^{-2} d^{-1}), G is the soil heat flux density (MJ m^{-2} d^{-1}), T_a is the mean daily air temperature at 2 m height (°C), u is wind speed (m s^{-1}) at 2 m height, e_s is the saturation vapour pressure (kPa), e_a is the actual vapour pressure (kPa), Δ is the slope of the vapour pressure curve (kPa °C^{-1}), γ is a psychrometric constant (kPa °C^{-1}), and 0.408 is a conversion factor from MJ m^{-2} d^{-1} to equivalent evaporation in mm d^{-1}.

R_n was estimated from global radiation, mean daily temperature, the mean daily vapour pressure, the site latitude, and elevation after Allen et al. (2005). A fixed value of 0.23 was applied for the albedo. It was assumed that soil heat flux density was $G = 0$ on a daily basis. Detailed description of the process can be read in Soós and Anda (2014).

The Tetens equation (Monteith and Unsworth, 2008; Allen et al., 1998; Tetens, 1930) was used for calculating saturation vapour pressure (e_s) as follows:

$$e_\mathrm{s} = 0.6108 \times \exp(17.27T_\mathrm{a}/(T_\mathrm{a} + 237.3)), \quad (4)$$

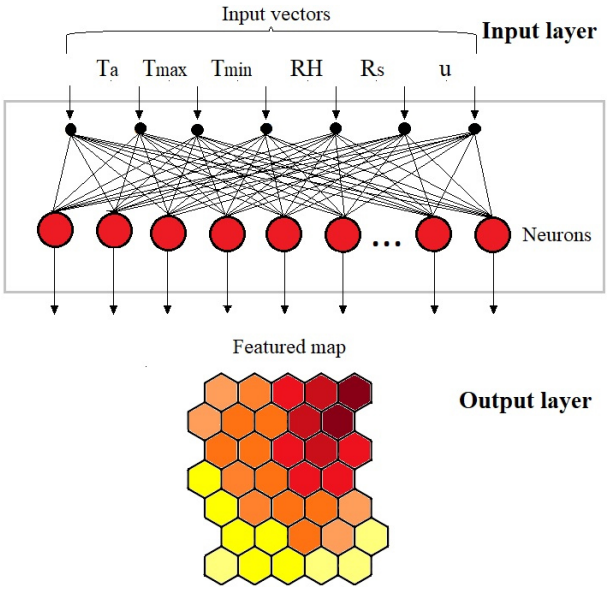

Figure 3. Illustration of the winning node and its neighbourhood in the Kohonen self-organising map (K–SOM).

where T_a is the air temperature in ° C. The actual vapour pressure, e_a, was calculated from the relative humidity (RH):

$$e_a = \left(\frac{RH}{100}\right) \times e_s. \quad (5)$$

2.4 Kohonen self-organisation map (K–SOM)

The K–SOM is a nonlinear mapping technique, which identifies groups of similarity in data sets without normal distribution assumption (Kohonen, 1982). SOM is a powerful and effective tool for complex data analyses such as data mining, estimation, and prediction. Using SOM, informative reference vectors are obtained via iterative updates under three main successive procedures: competition with nodes (1), selection of a winner node (2), and updating of the reference vector (3) (Yu et al., 2018). Every node has its vector adjusted according to sequential algorithm with the Gaussian neighbourhood function. The SOM consists of an input layer and an output layer (Park et al., 2006), where the output layer consists of so-called neurons, which are usually located in a hexagonal grid and are fully interconnected (Peeters et al., 2007). A schematic illustration of K–SOM is presented in Fig. 3. As similar input patterns could have different outputs, to determine the best output for a given input pattern is to use the mean output value as the clustered input patterns to the correspondent neuron, and then the closest (most similar) neuron would be directly used for the given input pattern (Chang et al., 2010; Kohonen, 1990).

The importance of K–SOM in the field of environmental science lies in the fact that SOMs can be used for prediction and correlation analysis, mostly with visual representation (Barreto and Pérez-Uribe, 2007). An outstanding element of this is that K–SOM finds statistically significant dependencies among the variables in a multidimensional data sample. In the case where two variables are highly correlated, K–SOM produces two similar component planes (Barreto and Pérez-Uribe, 2007).

K–SOM as NN provides a method above the standard estimations of pan evaporation, which seems necessary to get evaporation of natural ecosystems including lakes. In other words, applying a method in which the pan evaporation is estimated from other, easily measurable meteorological parameters such as sun radiation, air temperature, and relative humidity, has primary importance. This approach has widely been used for pan evaporation projection among others by Kisi et al. (2016) and Lin et al. (2013). Kisi et al. (2016) compared the soft computing model K–SOM and multiple linear regression (MLR). The authors demonstrated the superiority of K–SOM over MLR even in the model performance.

2.5 Statistics and performance evaluation criteria

The Shapiro–Wilk test was used as a statistical test for normality, with a chosen alpha value of 0.05 ($p < 0.05$). Two-way analysis of variance (ANOVA) with Tukey's HSD test was performed to examine the impacts of treatments C, S, and SM on class A pan E_p. To study the impact of meteorological elements on E_p of C, S, and SM treatments, Pearson's correlation analysis was used. This, as well as the MLR, was carried out with SPSS Statistics software. In this study, the K–SOM algorithm was executed using MATLAB 2019b software. To train (years: 2015–2017) and test the models (years: 2018–2020), half of the data were used.

Performance of the proposed models is evaluated by computing statistical indices, such as root mean square error (RMSE), mean absolute error (MAE), scatter index (SI), and Nash–Sutcliffe efficiency (NSE) between observed and estimated values of E_p for the data sets considered. The RMSE range is zero to infinity ($0 < \text{RMSE} < \infty$); the lower the RMSE, the better the model's performance. The RMSE is proportional to the observed mean, as a result, SI (Shiri and Kişi, 2011) forms a good non-dimensional error measure. NSE (Nash and Sutcliffe, 1970; ASCE, 1993) compares the congruence between the observed and predicted data. A high value of NSE (NSE ≤ 1) indicates high efficiency of the model (Duan et al., 2016; Li and Liu, 2020).

These evaluation criteria calculate as the following equations:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(E_{\mathrm{p_{obs},}i} - E_{\mathrm{p_{est},}i})^2}{n}}, \quad (6)$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}\left|E_{\mathrm{p_{est},}i} - E_{\mathrm{p_{obs},}i}\right|}{n}, \quad (7)$$

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(E_{\mathrm{p_{obs},}i} - E_{\mathrm{p_{est},}i})^2}{\sum_{i=1}^{n}(E_{\mathrm{p_{est},}i} - E_{\mathrm{p_{est,m}}})^2}, \quad (8)$$

$$\mathrm{SI} = \sqrt{\frac{\sum_{i=1}^{N}\left[(E_{\mathrm{p_{est},}i} - E_{\mathrm{p_{est,m}}}) - (E_{\mathrm{p_{obs},}i} - E_{\mathrm{p_{obs,m}}})\right]}{\sum_{i=1}^{N}E_{\mathrm{p_{obs},}i}^2}}, \quad (9)$$

where $E_{\mathrm{p_{obs},}i}$, $E_{\mathrm{p_{est},}i}$ is observed and estimated pan evaporation values on the ith day; and $E_{\mathrm{p_{obs,m}}}$ and $E_{\mathrm{p_{est,m}}}$ is the mean value of $E_{\mathrm{p_{obs},}i}$ and $E_{\mathrm{p_{est},}i}$, respectively. The total number of testing patterns is denoted by n, and i represents the number of particular instances of the testing pattern.

3 Results

3.1 Meteorological variables and pan evaporation

The long-term (1971–2000) growing season's average T_a at Keszthely is 18.8 °C, the hottest month is July with a mean monthly T_a of 20.5 °C, while the coolest month is September (15.7 °C). In the study period, the seasonal mean T_a was 5.5 %–15.7 % higher than the 30-year average. Out of six seasons studied, three warm (2015, 2017, 2019) and three close to normal (2016, 2018, 2020) ones could be distinguished (Fig. 4). The seasonal mean T_a in warm seasons were 11.5 %–15.7 % higher than that of the climate norms.

The climate of Keszthely is characterised by highly variable and irregular P with a long-term seasonal total of 274.3 mm from June to September. Monthly seasonal mean precipitation sums varied from 78.5 mm (June) to 57.1 mm (September). Warm seasons (2015, 2017, 2019) were characteristically arid with 4.9 %–21.6 % less seasonal total P, respectively, compared to the 30-year average. In the other study seasons, there were 23.9 %–40.4 % more P (data not shown) than that of the climate norm.

Figure 4 displays the meteorological variables and observed daily E_p in different pan treatments determined in a box and whisker plot between growing season 2015–2020, indicating minimum, first quartile, median, third quartile, and maximum values. An increasing trend was observed in the T_min with an increment of 9.6 %, while the T_max exhibited an unchanged trend over the studied growing seasons. In the study location, there were hardly any differences in seasonal mean RH values (0.6 %–9.2 %) and daily R_s sums (21.3–24.3 W m^{-2}) between 2015 and 2020. The highest (1.6 m s^{-1}) and lowest (0.9 m s^{-1}) seasonal mean wind speeds were measured in 2016 and 2018, respectively.

Daily E_p rates were related to seasonal T_a variations and not to rainfall patterns. Higher daily mean water losses were registered during the warm–dry seasons (C: 3.5–3.8 mm d^{-1}, S: 4.2–4.3 mm d^{-1}, SM: 4.5–4.9 mm d^{-1}), while somewhat lower average E_p rates were measured in the three normal seasons (C: 3.0–3.5 mm d^{-1}, S: 3.4–4.0 mm d^{-1}, SM: 3.6–4.2 mm d^{-1}). As a result of pan seeding, differences in daily mean E_p rates were more pronounced in warm summers. In warm seasons, significant deviations of daily mean E_p between C and S ($p < 0.001$) and S and SM ($p < 0.001$) were observed. At the same time, significant differences in daily mean E_p between C and S ($p < 0.001$) and C and SM ($p < 0.001$) were registered in normal seasons. No significant impact of pan seeding in all the remaining treatments was detected ($p = 0.0693$–0.0896) (Fig. 4). A two-way ANOVA was conducted to explore the impact of the studied seasons and the treatment on E_p rates. There were significant main effects caused by the growing season ($F(5, 211) = 24.241$, $p = 0.001$) and the pan treatment ($F(2, 236) = 67.855$, $p = 0.001$) in the full dataset. The interaction between seasons and treatments was not significant ($F(10, 29) = 0.085$, $p = 0.503$). Tukey HSD post-hoc tests revealed significant differences among the three pan treatments ($p < 0.001$ for all pairwise comparisons) for the training, testing phase, and full dataset (Table 1).

The correlation of evaporation of different pan treatments with other meteorological variables is also given in Table 2. There was a statistically significant difference in evaporation rates of full datasets and in the case of training and testing datasets between the seeded and classic class A pan. The T_a, T_max, and R_s positively impacted the E_p, while RH had a negative correlation with E_p. In this study, u hardly affected the E_p rates irrespective to treatment. The descriptive statistics of both training and testing datasets showed that most of the meteorological variables and E_p were similar to the full data set.

On the basis of the daily variation of T_w in different depths, two time periods were distinguished (Fig. 5); daytime (07:00–18:00 LMT) and nighttime cooling (19:00–06:00 LMT). With clear-sky conditions, the surface T_w peaked at 14:00, irrespective to treatment. The magnitudes of surface T_w in daytime (between 07:00 and 14:00) increased from 21.6 to 37.5 °C in C, from 23.0 to 37.4 ° C in S, and from 19.8 to 38.0 ° C in SM. Then, with declining solar radiation, the T_w slightly decreased during the nighttime cooling to 21.2, 21.8, and 18.7 °C in C, S, and SM, respectively, until sunrise. In deeper water depth, a similar pattern of T_w with slightly smaller magnitudes was measured with a time lag of 1 to 2 h from the surface T_w. In the classic A pan, the T_w in deeper depth from the surface did not reduce as rapidly as T_w in seeded pans. On cloudy days, insignificant T_w differences less than 1 °C ($p = 0.059$–0.969) between the neighbouring layers were observed in every treatment.

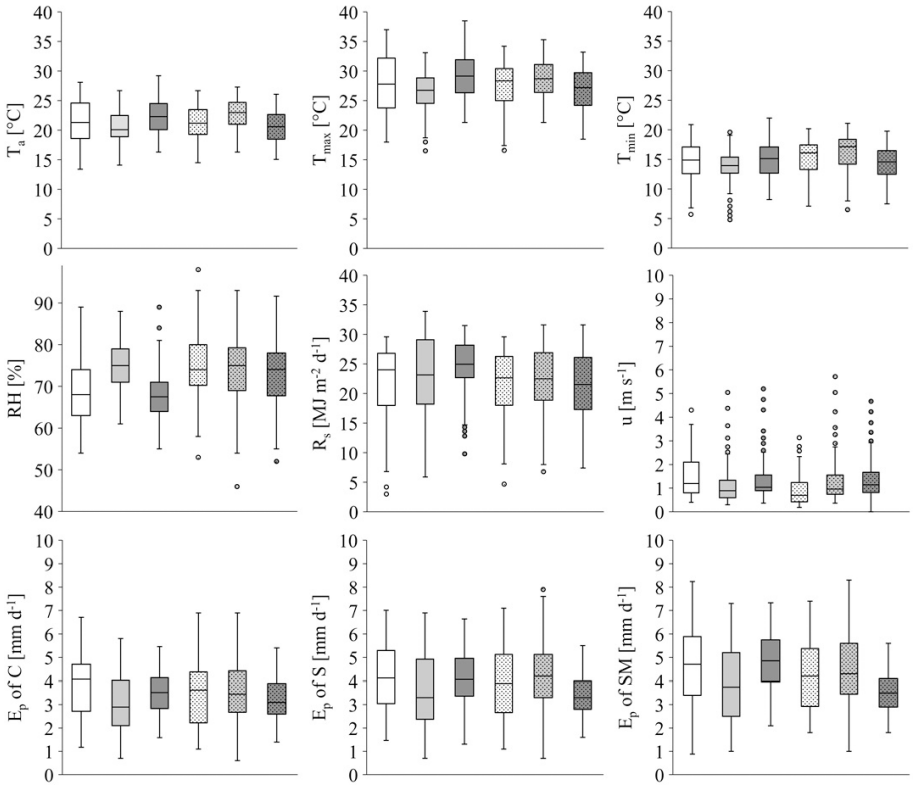

Figure 4. Box plot of meteorological parameters (T_a – daily mean temperature (°C), T_{max} – daily maximum temperature (°C), T_{min} – daily minimum temperature (°C), RH – relative humidity (%), R_s – global radiation (MJ m^{-2} d^{-1}), u – wind speed (m s^{-1})), and daily evaporation of different pan treatments (mm d^{-1}) (C – control, S – class A pan with sediment cover bottom, SM – class A pan with submerged macrophyte) in 2015–2020 growing seasons (June–September). The lower and upper ends of the box indicate the 25th and 75th percentiles of the variances, respectively, while the horizontal bar within the box indicates the median. The two horizontal bars indicate the range that covers 90 % of the variances. Outliers are indicated with circles.

3.2 K–SOM features

Table 3 shows the usual parameter table for K–SOM. The following steps were required to present Fig. 5: inputs were normalised, the code book was generated, and the map size complied with the dimensions of the component planes. The neighbouring function of the pixels was Gaussian, the shapes of component planes were sheets, and the planes shapes were hexagonal. Two indicators are most often used to qualitatively evaluate the two main goals of the K–SOM algorithm: quantisation error (QE) and topographic error (TE) (Table 3). The QE shows how closely the map vectors match the data vectors, thereby quantifying map resolution (Kohonen, 1995). The TE, in turn, determines the extent to which the topology of the input data structure is preserved on the output map (Kiviluoto, 1996). QE and TE do not have a default value, but the smaller the QE and TE (if the values tend to be zero), the better the model is. In this study, the values of QE and TE were equal to 0.016 and 0.820, respectively, indicating that the K–SOM was appropriately trained in topology.

K–SOM can be interpreted using the output map and the individual component planes, so the relationships between each variable can be explored. The component planes help to visually illustrate areas in which the intensity of the relationship of the variables is high, low, or average, and thus helps to better understand the relationship between the E_p and meteorological variables. The component planes for each variable of the K–SOM model are shown in Fig. 6. Superimposed on K–SOM patterns of input meteorological variables, radiation, air temperatures including minimum and maximum, relative humidity, and wind speed could be captured revealing their co-variability with the pan evaporation.

In the map, the similar weight vectors have similar colours, based on the U matrix according to a naïve contraction model proposed by Himberg (2000) and Peeters et al. (2007). Among NN features, as the clustering, classification, prediction, and data mining in large datasets (Kohonen and Somervuo, 2002; Kalteh et al., 2008), only the prediction and data mining were applied in the study. As there was no group distinction (classification), the U matrix has not been presented here.

Table 1. The impact of sediment (S) and submerged aquatic macrophytes (SM) on evaporation rates (E_p) of class A pan (C) in the full data set (2015–2020), training (2015–2017), and testing (2018–2020) phase with 95 % confidence intervals.

Multiple comparisons						
					95 % Confidence interval	
(I) Treatment	(J) Treatment	Mean difference (I–J)	Std. error	Sig.	Lower bound	Upper bound
Full dataset (2015–2020)						
C	S	−0.490*	0.0733	**0.000**	−0.662	−0.318
	SM	−0.845*	0.0735	**0.000**	−1.017	−0.672
S	C	0.490*	0.0733	**0.000**	0.318	0.662
	SM	−0.355*	0.0733	**0.000**	−0.526	−0.183
SM	C	0.845*	0.0735	**0.000**	0.672	1.017
	S	0.355*	0.0733	**0.000**	0.183	0.526
Based on observed means. The error term is mean square (error) = 1.741						
Training data set (2015–2017)						
C	S	−0.712*	0.1066	**0.000**	−0.962	−0.462
	SM	−0.731*	0.1072	**0.000**	−0.982	−0.479
S	C	0.712*	0.1066	**0.000**	0.462	0.962
	SM	−0.019*	0.1124	**0.019**	−0.283	0.245
SM	C	0.731*	0.1072	**0.000**	0.479	0.982
	S	0.019*	0.1124	**0.019**	−0.245	0.283
Based on observed means. The error term is mean square (error) = 1.840						
Testing data set (2018–2020)						
C	S	−0.505*	0.0993	**0.000**	−0.738	−0.272
	SM	−0.716*	0.1001	**0.000**	−0.951	−0.481
S	C	0.505*	0.0993	**0.000**	0.272	0.738
	SM	−0.211*	0.0990	**0.045**	−0.443	0.022
SM	C	0.716*	0.1001	**0.000**	0.481	0.951
	S	0.211*	0.0990	**0.045**	−0.022	0.443

* The mean difference is significant at the 0.05 level. The bold mark indicates the significant difference.

A colour was assigned to a node in accordance with the relative value of the respective component in that node (Li et al., 2018). On the maps, the warm colours (red, orange) show positive correlation, and the cool colours (blue) show negative correlation between the study variables. The darker the colour on the map (both warm and cool colours), the stronger the correlation. Lighter colours indicate lower correlation. When one variable is red while the other one is blue on the same place of the heat map, the correlation between them will be negative. Thus, the correlation between the K–SOM modelled values of E_p, T_a, T_{min}, T_{max}, R_s, RH, and u becomes clearly visible. The colour gradient of E_p was similar to those for variables related to available energy (T_a, T_{min}, T_{max}, and R_s), indicating that these contribute most to the increase of E_p. The component planes also visually confirm the negative correlation between RH and E_p, with high values of the RH resulting in low values of the E_p.

3.3 FAO56–PM, MLR, and K–SOM models

Figure 6 depicts the time variation and X–Y scatter plots of the observed and estimated daily E_p values obtained by C, S, and SM during the testing period (2018–2020).

From Fig. 7, it can be observed that most of the estimated daily E_p values (for MLR and K–SOM) are close to the observed daily E_p values for all three pan treatments. The possible reason for low R^2 values in FAO56–PM might be the role of the variable that is the estimate in crop potential evapotranspiration and not evaporation in water bodies. The regression line is above the 1 : 1 line up to 4 mm, which means that the FAO56–PM and MLR models slightly overestimated the magnitude of the daily E_p values in different pan treatments.

Table 2. Statistics of meteorological variables (T_a – mean air temperature, T_{max} – maximum air temperature, T_{min} – minimum air temperature, RH – relative humidity, R_s – solar radiation, u – wind speed) and their correlation with evaporation (E_p) of C, S, and SM in the full time series (2015–2020), training (2015–2017), and testing phases (2018–2020). C, S, and SM are control class A pan, A pan with sediment cover–bottom, and A pan with planted freshwater submerged macrophyte, respectively.

Data set	Statistics	T_a (°C)	T_{max} (°C)	T_{min} (°C)	RH (%)	u (m s^{-1})	R_s (MJ m^{-2} d^{-1})	E_p of C (mm d^{-1})	E_p of S (mm d^{-1})	E_p of SM (mm d^{-1})
Full (2015–2020)	Average ± SD	21.1 ± 3.2	27.5 ± 4.0	14.8 ± 3.2	72.7 ± 8.0	1.3 ± 0.9	22.4 ± 6.0	3.4 ± 1.2	3.9 ± 1.4	4.3 ± 1.5
	Correlation with E_p of C	0.59**	0.53**	0.42**	−0.43**	0.01	0.50**	1.00	–	–
	Correlation with E_p of S	0.57**	0.51**	0.40**	−0.42**	0.03	0.53**	0.92**	1.00	–
	Correlation with E_p of SM	0.56**	0.50**	0.37**	−0.44**	0.01	0.52**	0.90**	0.93**	1.00
Training (2015–2017)	Average ± SD	20.9 ± 3.4	27.5 ± 4.4	14.4 ± 3.3	71.0 ± 7.5	1.4 ± 0.9	23.1 ± 6.1	3.4 ± 1.2	4.0 ± 1.4	4.4 ± 1.6
	Correlation with E_p of C	0.65**	0.59**	0.49**	−0.48**	0.05	0.51**	1.00	–	–
	Correlation with E_p of S	0.63**	0.58**	0.45**	−0.47**	0.00	0.56**	0.91**	1.00	–
	Correlation with E_p of SM	0.63**	0.57**	0.44**	−0.50**	0.04	0.54**	0.89**	0.93**	1.00
Testing (2018–2020)	Average ± SD	21.2 ± 2.9	27.4 ± 3.5	15.3 ± 3.0	74.2 ± 8.2	1.2 ± 0.9	21.8 ± 5.7	3.4 ± 1.2	3.9 ± 1.4	4.1 ± 1.4
	Correlation with E_p of C	0.53**	0.46**	0.35**	−0.41**	0.06	0.51**	1.00	–	–
	Correlation with E_p of S	0.51**	0.44**	0.35**	−0.39**	0.06	0.50**	0.92**	1.00	–
	Correlation with E_p of SM	0.49**	0.41**	0.33**	−0.38**	0.05	0.49**	0.92**	0.95**	1.00

** Correlation is significant at the 0.01 level (two-tailed).

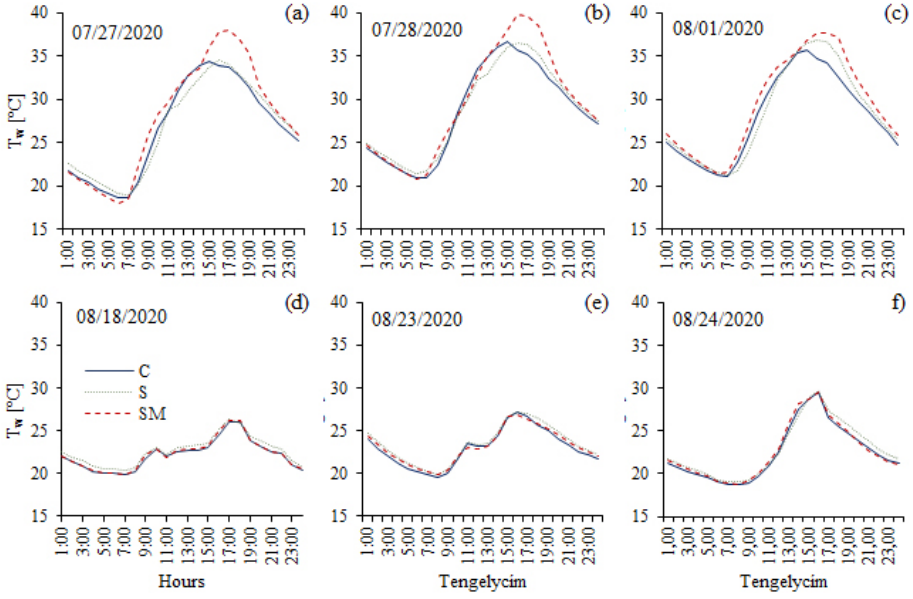

Figure 5. Water temperature of different pan treatments (C – class A pan/control; S – class A pan with sediment covered bottom; SM – class A pan with submerged macrophyte) in clear-sky and cloudy sample days. The layers represent the distance from the pan bottom. The lowest sensors' height was 5 cm.

However, above 4 mm daily E_p, the FAO56–PM and MLR models already underestimated the observed E_p values. The daily E_p values of C, S, and SM of the K–SOM model follow the 1 : 1 line most accurately. For all three models, R^2 values were highest for SM treatment (FAO56–PM: 0.1393, MLR: 0.6242, K–SOM: 9864). In the case of K–SOM, it can also be observed that low E_p values are overestimated, while higher E_p values are underestimated, although the estimated "middle" E_p values (which occur most frequently in a growing season) were close to the observed E_p values regardless of pan treatment. A greater degree of underestimation is observed for SM treatment for K–SOM.

In this study, we developed E_p models based on three different approaches (FAO56–PM, MLR, and K–SOM) with daily meteorological variables, and tested the performance of the models by four commonly used statistical indicators (MAE (Ideal = 0, $(0, +\infty)$), RMSE (Ideal = 0, $(0, +\infty)$), NSE (Ideal = 1, $(-\infty, 1)$), and SI (Ideal = 0, $(0, +\infty)$)). Figure 8 shows the overall performance of the three predicted methods at the three pan treatments.

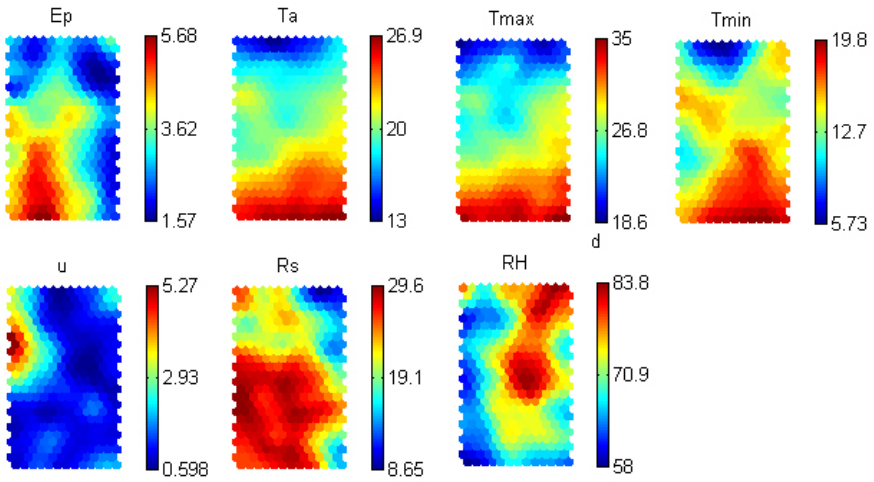

Figure 6. Kohonen self-organising map (K–SOM) visualisation of pan evaporation and meteorological variables assessment (T_a – daily mean temperature (°C), T_{max} – daily maximum temperature (°C), T_{min} – daily minimum temperature (°C), RH – relative humidity (%), R_s – global radiation (MJ m^{-2} d^{-1}), u – wind speed (m s^{-1}), and E_p – daily evaporation (mm d^{-1})). The bars indicate the intensity of the variables: the red colour is high importance and the blue colour is low importance.

Table 3. Characteristics of trained Kohonen self-organising map (K–SOM) model.

Characteristics	Values
Normalisation method	$x' = (x - \overline{x})/\sigma_x$ variance
Codebook	312×3
Map size	24×13
Neighbourhood function	Gaussian
Shape	Sheet
Lattice	Hexagonal
Final topographic error (TE)	0.820
Final quantisation error (QE)	0.016

The K–SOM models (RMSE = 0.222–0.253; NSE = 0.761–0.951; SI = 0.065–0.074) performed the best in the testing period, their RMSE and MAE were lower, and their NSE was higher than those of FAO56–PM and MLR models regardless of pan treatment (C: 0.951; S: 0.906; SM: 0.761). Additionally, the MAE value for treatments C and S was the lowest in the K–SOM models (MAE = 0.164 and MAE = 0.338, respectively); in contrast, the FAO56-PM had the best MAE value for SM treatment (MAE = 0.601).

Overall, the MLR (RMSE = 0.834; MAE = 0.660; S = 0.217) was slightly superior to FAO56–PM (RMSE = 0.877; MAE = 0.675; SI = 0.220) in the S, and there was only a small difference in the value of NSE between the two models (MLR: 0.572; FAO56-PM: 0.580). In the C treatment, RMSE (0.796) and SI (0.200) were lower for FAO56–PM, while MAE (0.648) and NSE (0.531) values were more favourable for the MLR model. Nevertheless, both the K–SOM model and MLR model were better than the FAO56–PM model during the testing period for "non-empty" treatments (S and SM).

4 Discussion

To date, there is little information about the impact of submerged aquatic macrophytes on E_p rate. According to a previous study in India (Kota, Rajasthan), water hyacinth evapotranspirated 26 % more water than free water surface in a 9-month experiment (Brezny et al., 1973). In the same place as this study, Anda et al. (2016, 2018) have shown that the presence of sediment increases the evaporation of the class A pans by an average of 12.7 %, and the submerged aquatic macrophytes by an average of 21.3 %, between 2014 and 2016. Jiménez-Rodríguez et al. (2019) reported that the observed E_p were higher for aquatic plants than the open water cover in Palo Verde National Park, Costa Rica, between December 2012 and January 2013 (45 d). Concerning the relationship between pan treatments and meteorological variables, it can be concluded that positive correlation was observed with most meteorological variables, while a negative correlation was observed with RH. This result was supported by other studies in the literature (Sheffield et al., 2017). In this study, u hardly affected the E_p rates of each treatment. This does not confirm the conclusions made by earlier studies (McVicar et al., 2012). This may be due to the fact that Keszthely is sheltered by surrounding mountains causing lower wind speeds (Anda et al., 2016).

Daily mean T_w increases were 5.4 and 4.5 °C in S and SM, respectively, compared to C during clear-sky conditions. Despite the less intense stratification on overcast days, T_w of seeded pans was 5.4 °C higher than that of daily mean T_w of C.

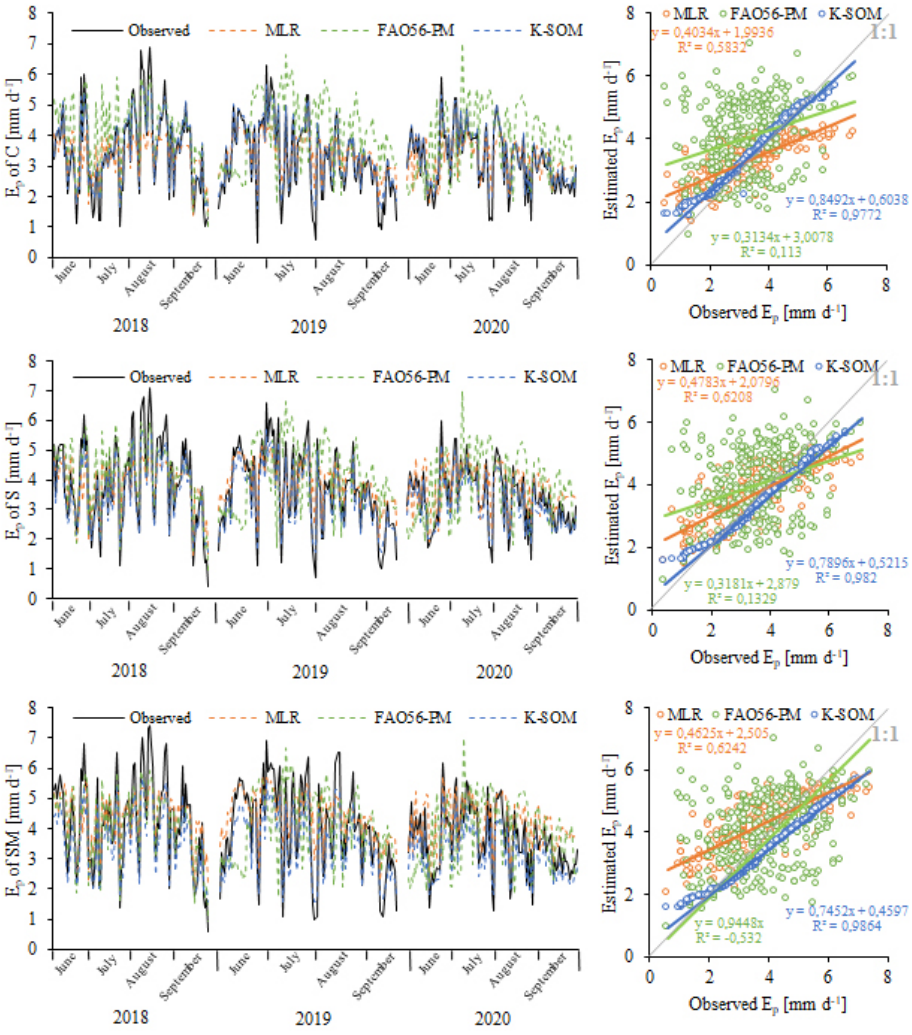

Figure 7. Time series and X–Y scatter plot of observed and predicted daily pan evaporation (E_p) in different pan treatment (C – control, S – pan with sediment cover bottom, SM – pan with submerged macrophytes) by daily multiple stepwise regression (MLR), FAO-56 Penman–Monteith (FAO56–PM), and Kohonen self-organisation map (K–SOM) models during testing period (2018–2020 growing seasons). All probability levels were equal to $p < 0.001$.

In accordance with shallow lake stratification results of Jacobs et al. (1998), increased stratification was evident in daytime, but the number of layers strongly depended on macrophyte presence. More moderate T_w layer differences were also present at night. The stratification was the most intense with three significantly different layers ($p < 0.001$) in seeded pans, during clear-sky daytime. At the same time, the number of layers with varied T_w was only two ($p < 0.001 - p = 0.012$) in classic A and sediment covered pans. Results in the study were confirmed by Andersen et al. (2017) concluding that shallow lakes colonised by submerged macrophytes strongly stratify the water body, mainly during the daytime. The reason for this stratification is the dissipating turbulent kinetic energy and absorbing heat (Vilas et al., 2018). The plants may act as a barrier to seeded pans water mixing, attenuating underwater light, thereby enhancing the thermal stratification inside the pan's water column.

The strength of stratification, the daily mean T_w differences between the surface and bottom water were 2.5 ($p = 0.005$), 3.0 ($p < 0.001$), and 6.5 °C ($p < 0.001$) in C, S, and SM, respectively, on cloudless days. At night-time cooling, variation in T_w between different layers was less pronounced, remaining below 1 °C ($p < 0.001$–$p = 0.005$).

In addition to stratification, the macrophytes have strengthened the daily variation of T_w in different depth. A 0.3 °C increase in daily mean surface T_w of seeded pans related to C was obtained during daytime, with a variation (T_{max}–T_{min}) of 18.4 and 19.3 °C in C and SM, respectively. On the bottom, an opposite trend in daytime mean T_w was detected; the seeded pans T_w in 0.05 m depth was 3.1 °C ($p = 0.040$) cooler than that of the T_w of C. Probably

the macrophyte presence resulted in insufficient downward heat transport, maintaining the more stratified water body of seeded pans. Herb and Stefan (2004) also found reduced turbulent mixing in shallow Otter Lake, Minnesota, with rooted macrophytes. The authors observed that T_w fluctuations at 20 cm depth were 3 °C in open water and 4.5 °C in lake water with macrophyte cover. Evapotranspiration functions of SM fitted to surface T_w evolution; the higher the surface T_w, the more intense the E_p rate was measured in SM related to E_p of classic A pan.

Many researchers have conducted research with neural networks aimed at the estimation of E_p as a function of meteorological variables (Keskin and Terzi, 2006). Several of these researchers found better results in E_p estimation with neural network approach than those obtained from the Priestley–Taylor and the Penman methods (Rahimikhoob, 2009; Malik et al., 2020b). Consistent with other studies, this study demonstrated that modelling of E_p is possible through the use of K–SOM technique in addition to the FAO56–PM and MLR methods. The comparison results indicated that, in general, the K–SOM model was superior to the FAO56–PM and MLR methods. Chang et al. (2010) used different methods to estimate pan evaporation, including also the K–SOM and the FAO56–PM. According to the results of Chang et al. (2010), K–SOM was the best of the studied methods, and it was found that the Penman–Monteith method is also likely to underestimate evaporation. Malik et al. (2017) used four heuristic approaches and two climate-based models to approximate monthly pan evaporation, where the K–SOM model performed better than the climate-based models. The regression line in scatter plots has R^2 as 0.937 for K–SOM model at Pantnagar and Ranichauri (India). In the study of Malik et al. (2017), RMSE values were 0.685 and 1.126 for K–SOM, when 50 % of the total available data was used in the testing of models in two stations.

5 Conclusions

The E_p of a class A pan with submerged aquatic macrophytes and with a sediment-covered bottom was observed at Keszthely, over six consecutive (2015–2020) growing seasons. In this study, it was attempted to model E_p by employing models consisting of FAO56–PM, MLR, and K–SOM, using daily pan evaporation values in different class A pan treatments (C, S, SM). The E_p rate of SM and S was always significantly higher than that of the "empty" class A pan each growing season. The presence of submerged macrophyte resulted in a higher E_p than in the sediment-covered class A pan.

Macrophyte-induced thermal stratification in water bodies (lakes/evaporation pans) emerges only in the vegetation period, during macrophyte development. One less layer in classic A pan compared to macrophyte seeded pans was probably due to modified T_w stratification causing changed water column stability. Wider T_w values-induced dynamics presented in the macrophyte seeded pans demonstrated the possibility of developing a more heterogenous environment for aquatic ecosystems. Macrophyte-induced modified thermal stratification with higher surface T_w could explain the increased E_p in seeded pans. Modified E_p of seeded pans made those values closer to the E_p of natural lakes with submerged macrophytes. While the T_w stratification trend in SM was similar to that of natural shallow lakes, it may also provide a new consideration for routine hydrometeorological management. T_w distribution in macrophyte-covered lakes impacts other physical properties such as nutrient cycling, dissolved oxygen, etc. When treating E_p from a pan to that from a vegetated surface including lakes or other aquatic habitats, to improve evaporation estimation, multidimensional approximation is necessary, offering simple methods for end-users including hydrologists, meteorologists, or any other specialists.

Daily E_p rates for all pan treatments were related to seasonal T_a variations. Correlation analysis revealed that T_a, T_{max}, T_{min}, and R_s had a positive correlation with pan evaporation, whereas RH had a negative correlation (-0.42 to -0.44) with E_p of C, S, and SM in the full dataset. Among all, the R (correlation coefficient) of T_a (ranging from 0.56–0.59) had a stronger positive correlation followed by R of T_{max} (ranging from 0.50–0.53) and R of R_s (ranging from 0.50–0.53). The relationship with u was low for the E_p of the three treatments, which can be explained by the low u of Keszthely in the growing seasons. Using the visualisation capability of the K–SOM, it was clearly confirmed that the E_p was more closely correlated with the variables related to available energy than the RH.

The performance accuracy of the different applied models was evaluated with RMSE, MAE, NSE, and SI statistics. Results showed that the K–SOM model has accuracy in prediction precision over the FAO56–PM and MLR models. Comparing the FAO56–PM and MLR models, MLR performed better in this study in S and SM treatments.

Since the E_p of one sample place was included in the study, the "generic" impact of submerged macrophytes on E_p was not fully discussed; maybe for different reasons, our results in other sites became variable. More surveys are needed to reveal the applicability of planted standard A pan E_p for different geographical and climatic conditions.

A possible application value of the study is in validating the presence of littoral sediments and macrophytes in evaporation estimation; the amount of lost water by wetlands that can easily be accounted in the prediction of their performance. Results from the study may also contribute to the protection of aquatic plants and to environmental management of wetlands also in other regions of the world. Management strategies aiming to estimate accurate water budget terms including evaporation can be a realistic aim for preventing further inaccurate water loss projections.

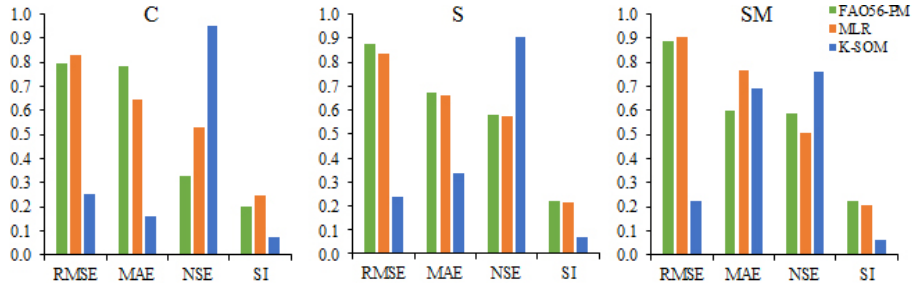

Figure 8. Error statistics (root mean square error – RMSE, mean absolute error – MAE, scatter index – SI and Nash–Sutcliffe efficiency – NSE) for the multiple stepwise regression (MLR), FAO-56 Penman–Monteith reference crop evapotranspiration (FAO56–PM), and Kohonen self-organisation map (K–SOM) models during the testing period for different pan treatments (C is standard class A pan with clean water, S is class A pan with sediment cover bottom, and SM is class A pan containing submerged macrophyte).

Data availability. The meteorological datasets are downloaded from https://odp.met.hu/climate/station_data_series/daily/ (last access: 12 February 2021).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/hess-26-4741-2022-supplement.

Author contributions. BSG and AA designed the experiments. BSG and SG carried them out. BSG produced all figures and tables and formatted the article. BSG and AA prepared the article. BSG, SG, and AA reviewed, revised, and supervised the progress of the paper.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Financial support. This research has been supported by the Ministry for Innovation and Technology from the National Research, Development and Innovation Fund (grant no. PD 138660).

Review statement. This paper was edited by Daniel Green and reviewed by Meine van Noordwijk and four anonymous referees.

References

Adeloye, A. J., Rustum, R., and Kariyama, I. D.: Kohonen self-organizing map estimator for the reference crop evapotranspiration, Water Resour. Res., 47, W08523, https://doi.org/10.1029/2011WR010690, 2005.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements, Irrigation and Drainage Paper 56, Food and Agriculture Organization of the United Nations: Rome, Italy, http://www.fao.org/3/x0490e/x0490e00.htm (last access: 14 March 2021), 1998.

Allen, R. G., Walter, I. A., Elliott, R., Howell, T., Itenfisu, D., Jensen, M., and Synder, R. L.: The ASCE Standardized Reference Evapotranspiration Equation, Final Report (ASCE–EWRI), Task Committee on Standardization of Reference Evapotranspiration, Environmental and Water Resources Institute of the American Society of Civil Engineers: Reston, VA, USA, https://doi.org/10.1061/9780784408056, 2005.

Almedeij, J.: Modeling pan evaporation for Kuwait by multiple linear regression, Sci. World J., 2012, 574742, https://doi.org/10.1100/2012/574742, 2012.

Alsumaiei, A. A.: Utility of Artificial Neural Networks in Modeling Pan Evaporation in Hyper-Arid Climates, Water, 12, 1508, https://doi.org/10.3390/w12051508, 2020.

An, N., Wang, K., Zhou, C., and Pinker, R. T.: Observed variability of cloud frequency and cloud-based height within 3600 m above the surface over the contiguous United States, J. Climate, 30, 3725–3742, https://doi.org/10.1175/JCLI-D-16-0559.1, 2017.

Anda, A., Simon, B., Soós, G., Menyhárt, L., Teixeira da Silva, J. A., and Kucserka, T.: Extending Class A pan evaporation for a shallow lake to simulate the impact of littoral sediment and submerged macrophytes: a case study for Keszthely Bay (Lake Balaton, Hungary), Agr. Forest Meteorol., 250, 277–289, https://doi.org/10.1016/j.agrformet.2018.01.001, 2018

Anda, A., Simon, B., Soos, G., Teixeira da Silva, J. A., and Kucserka, T.: Effect of submerged, freshwater aquatic macrohytes and littoral sediments on pan evaporation in the Lake Balaton region, Hungary, J. Hydrol., 542, 615–626, https://doi.org/10.1016/j.jhydrol.2016.09.034, 2016.

Anda, A., Soos, G., Teixeira da Silva, J. A., and Kozma-Bognár, V.: Regional evapotranspiration from a wetland in Central Europe, in a 16-year period without human intervention, Agric. Forest Meteor., 205, 60–72, https://doi.org/10.1016/j.agrformet.2015.02.010, 2015.

Andersen, M. R., Sand-Jensen, K., Iestyn Woolway, R., and Jones, I. D.: Profound daily vertical stratification and mixing in a small, shallow, wind-exposed lake with submerged macrophytes, Aquat. Sci., 79, 395–406, https://doi.org/10.1007/s00027-016-0505-0, 2017.

Arunkumar, R. and Jothiprakash, V.: Reservoir evaporation prediction using data driven techniques, J. Hydrol. Eng., 18, 40–49, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000597, 2013.

Barko, J. W., Hardin, D. G., and Matthews, M. S.: Growth and morphology of submersed freshwater macrophytes in relation to light and temperature, Can. J. Botany, 60, 877–887, https://doi.org/10.1139/b82-113, 1982

Barreto, S. M. A. and Pérez-Uribe, A.: Improving the Correlation Hunting in a Large Quantity of SOM Component Planes, in: Artificial Neural Networks – ICANN 2007, edited by: de Sá, J. M., Alexandre, L. A., Duch, W., and Mandic, D., Lect. Notes Comput. Sc., 4669, Springer, Berlin, Heidelberg, 379–388, https://doi.org/10.1007/978-3-540-74695-9_39, 2007.

Bedoya, D., Novotny, V., and Manolakos, E. S.: Instream and off stream environmental conditions and stream biotic integrity importance of scale and site similarities for learning and prediction, Ecol. Model., 220, 2393–2406, https://doi.org/10.1016/j.ecolmodel.2009.06.017, 2009.

Berkovic, S., Mendelsohn, O. Y., Ilotoviz, E., and Raveh-Rubin, S.: Self-organizing map classification of the boundary layer profile: A refinement of Eastern Mediterranean wintersynoptic regimes, Int. J. Climatol., 41, 3317–3338, https://doi.org/10.1002/joc.7021, 2021.

Brezny, O., Mehta, I., and Sharmas, R. K.: Studies of evapotranspiration of some aquatic weeds, Weed Sci., 21, 197–204, https://doi.org/10.1017/S0043174500032112, 1973.

Brutsaert, W. H.: Evaporation into the Atmosphere, Springer, 12–36, https://doi.org/10.1007/978-94-017-1497-6, 1982.

Burman, R. D.: Intercontinental comparison of evaporation estimates, J. Irr. Drain. Div.-ASCE, 102, 109–118, https://doi.org/10.1061/JRCEA4.0001076, 1976.

Chang, F. J., Chang, L. C., Kao, H. S., and Wu, G. R.: Assessing the effort of meteorological variables for evaporation estimation by self-organizing map neural network, J. Hydrol., 384, 118–129, https://doi.org/10.1016/j.jhydrol.2010.01.016, 2010.

Doan, Q.-V., Kusaka, H., Sato, T., and Chen, F.: S-SOM v1.0: a structural self-organizing map algorithm for weather typing, Geosci. Model Dev., 14, 2097–2111, https://doi.org/10.5194/gmd-14-2097-2021, 2021.

Dong, L., Zeng, W., Wu, L., Lei, G., Chen, H., Srivastava, A. K., and Gaiser, T.: Estimating the pan evaporation in Northwest China by coupling CatBoost with Bat algorithm, Water, 13, 256, https://doi.org/10.3390/w13030256, 2021.

Duan, W. Y., Han, Y., Huang, L. M., Zhao, B. B., and Wang, M. H.: A hybrid EMD-SVR model for the short-term prediction of significant wave height, Ocean Eng., 124, 54–73, https://doi.org/10.1016/j.oceaneng.2016.05.049, 2016.

Fournier, J., Thiboult, A., Nadeau, D. F., Vercauteren, N., Anctil, F., Parent, A.-C., Strachan, I. B., and Tremblay A.: Evaporation from boreal reservoirs: Acomparison between eddy covariance observations andestimates relying on limited data, Hydrol. Process., 35, e14335, https://doi.org/10.1002/hyp.14335, 2021.

Fritz, C., Schneider, T., and Geist, J.: Seasonal Variation in Spectral Response of Submerged Aquatic Macrophytes: A Case Study at Lake Starnberg (Germany), Water, 9, 527, https://doi.org/10.1016/10.3390/w9070527, 2017.

Fuentes, van Ogtrop, F., and Vervoot, R. W.: Long term surface water trends and relationship with open water evaporation losses in the Namoi catchment, Australia, J. Hydrol., 584, 124714, https://doi.org/10.1016/j.jhydrol.2020.124714, 2020.

Gholami, V., Sahour, H., and Hadian, M. A.: Mapping soil erosion rates using self-organizing map (SOM) and geographic information system (GIS) on hillslopes, Earth Sci. Inform., 13, 1175–1185, https://doi.org/10.1007/s12145-020-00499-w, 2020.

Gu, Q., Hu, H., Ma., L., Sheng, L., Yang, S., Zhang, X., Zhang, M., Zheng, K., and Chen, L.: Characterizing the spatial variations of the relationship between land use and surface water quality using self-organizing map approach, Ecol. Indicat., 102, 633–643, https://doi.org/10.1016/j.ecolind.2019.03.017, 2019.

Guntu, R. K., Maheswaran, R., Agarwal, A., and Singh, V. P.: Accounting for temporal variability for improved precipitation regionalization based on self-organizing map coupled with information theory, J. Hydrol., 590, 125236, https://doi.org/10.1016/j.jhydrol.2020.125236, 2020.

Hadjisolomou, E., Stefanidis, K., Papatheodorou, G., and Papastergiadou, E.: Assessment of the eutrophication-related environmental parameters in two mediterranean lakes by integrating statistical techniques and self-organizing maps, Int. J. Env. Res. Pub. He., 15, 547, https://doi.org/10.3390/ijerph15030547, 2018.

Himberg, J.: A SOM Based Cluster Visualization and Its Application for False Coloring, Proceedings of International Joint Conference on Neural Networks (IJCNN2000), 3, 587–592, https://doi.org/10.1109/IJCNN.2000.861379, 2000.

Jacobs, A. F. G., Heusinkveld, B. G., and Nieveen, J. P.: Temperature Behavior of a Natural Shallow Water Body during a Summer Period, Theor. Appl. Climatol., 59, 121–127, https://doi.org/10.1007/s007040050017, 1998.

Jensen, M. E., Burman, R. D., and Allen, R. G.: Evapotranspiration and irrigation water requirements, American Society of Civil Engineers 70, New York, 332 pp., https://doi.org/10.1061/9780784414057, 1990.

Jiménez-Rodríguez, C. D., Esquivel-Vargas, C., Coenders-Gerrits, M., and Sasa-Marín, M.: Quantification of the Evaporation Rates from Six Types of Wetland Cover in Palo Verde National Park, Costa Rica, Water, 11, 674, https://doi.org/10.3390/w11040674, 2019.

Kalteh, A. M., Hjorth, P., and Berndtsson, R.: Review of the self-organizing map (SOM) approach in water resources: analysis, modelling and application, Environ. Model. Softw., 23, 835–845, https://doi.org/10.1016/j.envsoft.2007.10.001, 2008.

Keskin, M. E. and Terzi, O.: Artificial neural network models of daily pan evaporation, J. Irrig. Drain. Eng., 11, 65–70, https://doi.org/10.1061/(ASCE)1084-0699(2006)11:1(65), 2006.

Khatibi, R., Ghorbani, M. A., Naghshara, S., Aydin, H., and Karimi, V.: Introducing a framework for "inclusive multiple modelling" with critical views on modelling practices-Applications to modelling water levels of Caspian Sea and Lakes Urmia and Van, J. Hydrol., 587, 124923, https://doi.org/10.1016/j.jhydrol.2020.124923, 2020.

Kim, J. Y. and Nishihiro, J.: Responses of lake macrophyte species and functional traits to climate and land use changes, Sci. Total Environ., 736, 139628, https://doi.org/10.1016/j.scitotenv.2020.139628, 2020.

Kim, S., Shiri, J., Singh, V. P., Kisi, O., and Landeras, G.: Predicting daily pan evaporation by soft computing models

with limited climatic data, Hydrolog. Sci. J., 60, 1120–1136, https://doi.org/10.1080/02626667.2014.945937, 2015.

Kimmel, B. L. and Groeger, A. W.: Factors controlling primary production in lakes and reservoirs: a perspective, Lake Reserv. Manage., 1, 277–281, https://doi.org/10.1080/07438148409354524, 1984.

Kisi, O., Genc, O., Dinc, S., and Zounemat-Kermani, M.: Daily pan evaporation modeling using chi-squared automatic interactiondetector, neural networks, classification and regression tree, Comput. Electron. Agr., 122, 112–117, https://doi.org/10.1016/j.compag.2016.01.026, 2016.

Kisi, O.: Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree, J. Hydrol., 528, 312–320, https://doi.org/10.1016/j.jhydrol.2015.06.052, 2015.

Kiviluoto, K.: Topology preservation in self-organizing maps, Proceedings of International Conference on Neural Networks, 294–299, https://doi.org/10.1109/ICNN.1996.548907, 1996.

Kohonen, T.: Self-organizing formation of topologically correct feature maps, Biol. Cybern., 43, 59–69, https://doi.org/10.1007/BF00337288, 1982.

Kohonen, T.: The self-organizing map. Proceedings of the IEEE, 78, 1464–1480, https://doi.org/10.1109/5.58325, 1990.

Kohonen, T.: Self-Organizing Maps, 3rd edition, Berlin, Heildelberg: Springer-Verlag, p. 501, 2001.

Kohonen, T. and Somervuo, P.: How to make large self-organizing maps for nonvectorial data. Neural Netw., 15, 945–952, https://doi.org/10.1016/S0893-6080(02)00069-2, 2002.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: WorldMap of the Köppen-Geiger climate classification updated, Meteorol. Z., 15, 259–263, https://doi.org/10.1127/0941-2948/2006/0130, 2006.

Kumar, M., Raghuwanshi, N. S., Singh, R., Wallender, W. W., and Pruitt, W. O.: Estimating evapotranspiration using artificial neural network, J. Irrig. Drai. Eng., 128, 224–233, https://doi.org/10.1061/(ASCE)0733-9437(2002)128:4(224), 2002.

Kumar, N., Rustum, R., Shankar, V., and Adeloye, A. J.: Self-organizing map estimator for the crop water stress index, Comput. Electron. Agr., 187, 106232, https://doi.org/10.1016/j.compag.2021.106232, 2021a.

Kumar, N., Shankar, V., Rustum, R., and Adeloye, A. J.: Evaluating the Performance of Self-Organizing Maps to Estimate Well-Watered Canopy Temperature for Calculating Crop Water Stress Index in Indian Mustard (Brassica juncea), J. Irrig. Drain. Eng., 147, 04020040, https://doi.org/10.1061/(ASCE)IR.1943-4774.0001526, 2021b.

Lee, E. and Kim, S.: Characterization of soil moisture response patterns and hillslope hydrological processes through a self-organizing map, Hydrol. Earth Syst. Sci., 25, 5733–5748, https://doi.org/10.5194/hess-25-5733-2021, 2021.

Lee, C.-M., Choi, H., Kim, Y., Kim, M, Kim, H., and Hamm, S.-Y.: Characterizing land use effect on shallow groundwater contamination by using self-organizing map and buffer zone, Sci. Total Environ., 800, 149632, https://doi.org/10.1016/j.scitotenv.2021.149632, 2021.

Li, M. and Liu, K.: Probabilistic prediction of significant wave height using dynamic bayesian network and information flow, Water, 12, 2075, https://doi.org/10.3390/w12082075, 2020.

Li, Y., Wright, A., Liu, H., Wang, J., Wang, G., Wu, Y., and Dai, L.: Land use pattern, irrigation, and fertilization effects of rice-wheat rotation on water quality of ponds by using self-organizing map in agricultural watersheds, Agr. Ecosyst. Environ., 272, 155–164, https://doi.org/10.1016/j.agee.2018.11.021, 2019.

Lin, G. F., Lin, H. Y., and Wu, M. C.: Development of a support-vector-machine-based model for daily pan evaporation estimation. Hydrol. Process., 27, 3115–3127, https://doi.org/10.1002/hyp.9428, 2013.

Madsen, T. V. and Cedergreen, N.: Sources of nutrients to rooted submerged macrophytes growing in a nutrient-rich stream, Freshwater Biol., 47, 283–291, https://doi.org/10.1046/j.1365-2427.2002.00802.x, 2002.

Malik, A., Kumar, A., and Kisi, O.: Monthly pan-evaporation estimation in Indian central Himalayas using different heuristic approaches and climate based models, Comput. Electron. Agr., 143, 302–313, https://doi.org/10.1016/j.compag.2017.11.008, 2017.

Malik, A., Rai, P., Heddam, S., Kisi, O., Sharafati, A., Salih, S. Q., Al-Ansari, N., and Yaseen, Z. M.: Pan evaporation estimation in Uttarakhand and Uttar Pradesh States, India: validity of an integrative data intelligence model, Atmosphere, 11, 553, https://doi.org/10.3390/atmos11060553, 2020a.

Malik, A., Kumar, A., Kim, S., Kashani, M. H., Karimi, V., Sharafati, A., Ghorbani, M. A., Al-Ansari, N., Salih, S. Q., Yaseen, Z. M., and Chau, K.-W.: Modeling monthly pan evaporation process over the Indian central Himalayas: application of multiple learning artificial intelligence model, Eng. Appl. Comp. Fluid., 14, 323–338, https://doi.org/10.1080/19942060.2020.1715845, 2020b.

Mbangiwa, N. C., Savage, M. J., and Mabhaudhi, T.: Modelling and measurement of water productivity and total evaporation in a dryland soybean crop, Agric. For. Meteorol., 266, 65–72, https://doi.org/10.1016/j.agrformet.2018.12.005, 2019.

McVicar ,T. R., Roderick, M. L., Donohue, R. J., Li, L. T., van Niel, T. G., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N. M., Mescherskaya, A. V., Kruger, A. C., Rehman, S., and Dinpashohl, Y.: Global review and synthesis of trends in observed terrestrial near-surface wind speeds: implications for evaporation, J. Hydrol., 416, 182–205, https://doi.org/10.1016/j.jhydrol.2011.10.024, 2012.

Monteith, J. L. and Unsworth, M. H.: Principles of Environmental Physics, Third Ed. AP, Amsterdam, ISBN: 9780080924793, 2008.

Nada, T., Sahoo, B., and Chatterjee, C.: Enhancing the applicability of Kohonen Self-Organizing Map (KSOM) estimator for gap-filling in hydrometeorological timeseries data, J. Hydrol., 549, 133–147, https://doi.org/10.1016/j.jhydrol.2017.03.072, 2017.

Nakagawa, K., Amano, H., Kawamura, A., and Berndtsson, R.: Classification of groundwater chemistry in Shimabara, using self-organizing maps, Hydrol. Res., 48, 840–850, https://doi.org/10.2166/nh.2016.072, 2017.

Nakagawa, K., Yu, Z.-Q., Berndtsson, R., and Hosono, R.: Temporal characteristics of groundwater chemistry affected by the 2016 Kumamoto earthquake using self-organizing maps, J. Hydrol., 582, 124519, https://doi.org/10.1016/j.jhydrol.2019.124519, 2020.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – a discussion of principles, J. Hydrol., 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Park, Y. S., Lek, S., Scardi, M., Verdonschot, P., and Jorgensen, S. E.: Patterning exergy of benthic macroinvertebrate communities using self-organizing maps, Ecol. Model., 195, 105–113, https://doi.org/10.1016/j.ecolmodel.2005.11.027, 2006.

Patle, G. T., Chettri, M., and Jhajharia, D.: Monthly pan evaporation modelling using multiple linear regression and artificial neural network techniques, Water Supply, 20, 800–808, https://doi.org/10.2166/ws.2019.189, 2020.

Pearce, A. R., Rizzo, D. M., and Mouser, P. J.: Subsurface characterization of groundwater contaminated by landfill leachate using microbial community profile data and a nonparametric decision-making process, Water Resour. Res., 47, W06511, https://doi.org/10.1029/2010WR009992, 2011.

Peeters, L., Bação, F., Lobo, V., and Dassargues, A.: Exploratory data analysis and clustering of multivariate spatial hydrogeological data by means of GEO3DSOM, a variant of Kohonen's Self-Organizing Map, Hydrol. Earth Syst. Sci., 11, 1309–1321, https://doi.org/10.5194/hess-11-1309-2007, 2007.

Poikane, S., Birk, S., Böhmer, J., Carvalho, L., de Hoyos, C., Gassner, H., Hellsten, S., Kelly, M., Lyche Solheim, A., Olin, M., Pall, K., Phillips, G., Portielje, P., Ritterbusch, B., Sandin, L., Schartau, A. K., Solimini, A. G., van den Berg, M., Wolfram, G., and van de Bund, W.: A hitchhiker's guide to european lake ecological assessment and intercalibration, Ecol. Indic., 52, 533–544, https://doi.org/10.1016/j.ecolind.2015.01.005, 2015.

Rahimikhoob, A.: Estimating daily pan evaporation using artificial neural network in a semi-arid environment, Theor. Appl. Climatol., 98, 101–105, https://doi.org/10.1007/s00704-008-0096-3, 2009.

Razi, M. A. and Athappilly, K.: A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models, Expert Syst. Appl., 29, 65–74, https://doi.org/10.1016/j.eswa.2005.01.006, 2005.

Ristić, S., Stamenković, S., Piperac, M. S., Šajn, R., Kosanić, M., and Ranković, B.: Searching for lichen indicator species: the application of self-organizing maps in air quality assessment – a case study from Balkan area (Serbia), Environ. Monit. Assess, 192, 693, https://doi.org/10.1007/s10661-020-08633-3, 2020.

Rivas-Tabares, D., de Miguel, Á., Willaarts, B., and Tarquis, A. M.: Self-organizing map of soil properties in the context of hydrological modeling, Appl. Math. Model., 88, 175–189, https://doi.org/10.1016/j.apm.2020.06.044, 2020.

Sanikhani, H., Kisi, O., Kiafar, H., and Ghavidel, S. Z. Z.: Comparison of different data-driven approaches for modeling lake level fluctuations: the case of Manyas and Tuz Lakes (Turkey), Water Resour. Manag., 29, 1557–1574, https://doi.org/10.1007/s11269-014-0894-6, 2015.

Sattari, M. T., Apaydin, H., and Shamshirband, S.: Performance Evaluation of Deep Learning-Based Gated Recurrent Units (GRUs) and Tree-Based Models for Estimating ETo by Using Limited Meteorological Variables, Mathematics, 8, 972, https://doi.org/10.3390/math8060972, 2020.

Sheffield, J., Goteti, G., and Wood, E. F.:. Development of a 50-Year high-resolution global dataset of meteorological forcings for land surface modelling, J. Climate, 19, 3088–3111, https://doi.org/10.1175/JCLI3790.1, 2006.

Shiri, J. and Kisi, O.: Application of artificial intelligence to estimate daily pan evaporation using available and estimated climatic data in the Khozestan Province (South-Western Iran), J. Irrig. Drain. Eng., 137, 412–425, https://doi.org/10.1061/(ASCE)IR.1943-4774.0000315, 2011.

Soós, G. and Anda, A.: A methodological study on local application of the FAO-56 Penman-Monteith reference evapotranspiration equation, Georgikon Agricult., 18, 71–85, 2014.

Sudheer, K. P., Gosain, A. K., and Ramasastri, K. S.: Estimating actual evapotranspiration from limited climatic data using neural computing technique, J. Irrig. Drain. Eng., 129, 214–218, https://doi.org/10.1061/(ASCE)0733-9437(2003)129:3(214), 2003.

Tabari, H., Marofi, S., and Sabziparvar, A. A.: Estimation of daily pan evaporation using artificial neural network and multivariate non-linear regression, Irrigation Sci., 28, 399–406, https://doi.org/10.1007/s00271-009-0201-0, 2010.

Tetens, O.: Über einige meteorologische Begriffe, Z. Geophys., 6, 297–309, 1930.

Vilas, M. P., Marti, C. L., Oldham, C. E., and Hipsey, M. R.: Macrophyte-induced thermal stratification in a shallow urban lake promotes conditions suitable for nitrogen-fixing cyanobacteria, Hidrobiologica, 806, 411–426, https://doi.org/10.1007/s10750-017-3376-z, 2018.

Vymazal, J.: Emergent plants used in free water surface constructed wetlands: a re-view, Ecol. Eng., 61, 582–592, https://doi.org/10.1016/J.ECOLENG.2013.06.023, 2013.

Wang, L., Hang, S., and Tian, F.: Comparison of formulating apparent potential evaporation with pan measurements and Penman methods. J. Hydrol., 592, 1258162021, https://doi.org/10.1016/j.jhydrol.2020.125816, 2021.

William, R. H. and Heinz, G. S.: Temperature Stratification nand Mixing Dynamics in a Shallow Lake With Submersed Macrophytes, Lake Reserv. Manage., 20, 296–308, https://doi.org/10.1080/07438140409354159, 2004.

WMO Report: Drought and Agriculture, WMO Techn. Note No. 138, 1975.

Wu, L., Huang, G., Fan, J., Ma, X., Zhou, H., and Zeng, W.: Hybrid extreme learning machine with metaheuristic algorithms for monthly pan evaporation prediction, Comput. Electron. Agr., 168, 105115, https://doi.org/10.1016/j.compag.2019.105115, 2020.

Yan, K., Yuan, Z., Goldberg, S., Gao, W., Ostermann, A., Xu, J., Zhang, F., and Elser, J.: Phosphorus mitigation remains critical in water protection: A review andmeta-analysis from one of China's most eutrophicated lakes, Sci. Total Environ., 689, 1336–1347, https://doi.org/10.1016/j.scitotenv.2019.06.302, 2019.

Yu, Z. Q., Amano, H., Nakagawa, K., and Berndtsson, R.: Hydrogeochemical evolution of groundwater in a Quaternary sediment and Cretaceous sandstone unconfined aquifer in Northwestern China, Environ. Earth Sci., 77, 629, https://doi.org/10.1007/s12665-018-7816-5, 2018.

Zelazny, M., Astel, A., Wolanin, A., and Malek, S.: Spatiotemporal dynamics of spring and stream water chemistry in a high-mountain area, Environ. Pollut., 159, 1048–1057, https://doi.org/10.1016/j.envpol.2010.11.021, 2011.

Zhang, Y., Jeppesen, E., Liu, X., Qin, B., Shi, K., Zhou, Y., Thomaz, S. M., and Deng, J.: Global loss of aquatic vegetation in lakes, Earth-Sci. Rev., 173, 259–265, https://doi.org/10.1016/j.earscirev.2017.08.013, 2017.