

## Article

# Evaluation of Evaporation from Water Reservoirs in Local Conditions at Czech Republic

Eva Melišová <sup>1,2,\*</sup> , Adam Vizina <sup>1,2,†</sup> , Martin Hanel <sup>1,2,†</sup> , Petr Pavlík <sup>1,2</sup>  and Petra Šuhájková <sup>1,2</sup> 

<sup>1</sup> Department of Hydrology, T. G. Masaryk Water Research Institute, Podbabská 2582/30, 160 00 Prague, Czech Republic; adam.vizina@vuv.cz (A.V.); hanel@fzp.czu.cz (M.H.); petr.pavlik@vuv.cz (P.P.); petra.suhajkova@vuv.cz (P.Š.)

<sup>2</sup> Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 165 00 Prague, Czech Republic

\* Correspondence: eva.melissova@vuv.cz; Tel.: +420-220-197-263

† These authors contributed equally to this work.

**Abstract:** Evaporation is an important factor in the overall hydrological balance. It is usually derived as the difference between runoff, precipitation and the change in water storage in a catchment. The magnitude of actual evaporation is determined by the quantity of available water and heavily influenced by climatic and meteorological factors. Currently, there are statistical methods such as linear regression, random forest regression or machine learning methods to calculate evaporation. However, in order to derive these relationships, it is necessary to have observations of evaporation from evaporation stations. In the present study, the statistical methods of linear regression and random forest regression were used to calculate evaporation, with part of the models being designed manually and the other part using stepwise regression. Observed data from 24 evaporation stations and ERA5-Land climate reanalysis data were used to create the regression models. The proposed regression formulas were tested on 33 water reservoirs. The results show that manual regression is a more appropriate method for calculating evaporation than stepwise regression, with the caveat that it is more time consuming. The difference between linear and random forest regression is the variance of the data; random forest regression is better able to fit the observed data. On the other hand, the interpretation of the result for linear regression is simpler. The study introduced that the use of reanalyzed data, ERA5-Land products using the random forest regression method is suitable for the calculation of evaporation from water reservoirs in the conditions of the Czech Republic.

**Keywords:** evaporation; water reservoir; regression; observed data; ERA5-Land data; R language



**Citation:** Melišová, E.; Vizina, A.; Hanel, M.; Pavlík, P.; Šuhájková P. Evaluation of Evaporation from Water Reservoirs in Local Conditions at Czech Republic. *Hydrology* **2021**, *8*, 153. <https://doi.org/10.3390/hydrology8040153>

Academic Editors: Aristoteles Tegos and Nikolaos Malamos

Received: 31 August 2021

Accepted: 29 September 2021

Published: 12 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Water management, changes in natural water regime and sustainable landscape became an important topic of social discussions and policies not only in the Czech Republic, but also around the world [1]. It is clear that global and local climatic conditions are changing and will have an impact on the water management sector and therefore they should be given the highest attention. The evaporation in the Czech Republic also changes [2].

However, not only the climatic conditions change, but also the technology and knowledge that can be used in water management and specifically in hydrology. With the rapid development of remote sensing tools through recent decades an onset of easy-to-use high quality products supplied both professionals and public in water resources.

In recent years, there has been a significant development in the supply of information from remote sensing of the Earth utilizable in water management, not only for the professional public [3–5]. Another option is, for example, the use of globally available climate reanalyses or other available data sources. Despite the development of data availability and modelling tools, a question arises: How significant is the impact of the ongoing cli-

mate change on hydrological balance components and the consequent impact on water management [6]?

The hydrological balance is tied to rainfall-runoff processes, which are driven by climatic, geographical and geomorphological factors. The climatic factors include meteorological factors affecting the evaporation and evapotranspiration from the catchment, such as: precipitation, humidity, soil moisture, evaporation, air temperature, wind speed and direction and atmospheric pressure [7].

Recently, a number of studies pointed out that evapotranspiration significantly affects the hydrological balance. The key role of evapotranspiration in hydrological balance was the subject of many recent studies, e.g., [8–11]. And it is nowadays widely recognized, that on the most of the Earth's surface evaporation plays crucial role in the hydrological cycle.

The study [12] illustrates the impacts of climate change on the water cycle, which may impact from total evaporation, precipitation, atmospheric humidity and horizontal moisture transport at the global scale.

There are many methods to calculate evaporation, which can be calculated from free water, from the soil surface or from vegetation over a period of time. The evaluation of evaporation can be done by direct methods namely measurement or by indirect methods: empirical methods, remote sensing of the Earth on regional or global scales [13,14], the use of models that are classified as fully physically-based combination models, semi-physically based models or black-box models [15].

The total evaporation can be divided into actual, potential or reference evapotranspiration. The potential evaporation can be determined by empirical relationships or by measurement, the empirical relationships may differ in the input data or in the time step [8,16]. The calculation of the reference evapotranspiration is defined according to the FAO methodology, with the reference area being divided in [17].

The studies [18,19] evaluated evapotranspiration calculated on the base of empirical equations, which were divided into categories: mass-transfer, radiation based method and temperature-based method. The best equations from each category were then selected and compared based on the FAO and Penman–Monteith equations [20].

The estimation of reference evapotranspiration was used in the study [21], where the Penman–Monteith temperature-based equation achieved the best rating for the evaluation of reference evapotranspiration because it preserves the physical philosophy of the Penman–Monteith equation method. The method was applied at a global scale using the Köppen climate classification system with respect to the world dataset under different climate conditions. Calculation of reference evapotranspiration based on indirect methods can provide acceptable results when direct measurements of are not available [15].

Since most of the empirical formulas are based on geographical location, it is straightforward that the empirical calculation of evapotranspiration is not the same for different regions, due to the different climatic conditions [17]. National standards, legislation and expertise also takes place resulting that different methods are preferred in different countries, e.g., Netherlands—Makkink's method [22], Slovakia—Budyko's method [23], Bulgaria—Delibaltov–Hristov–Tsonev method [24].

The Penman–Monteith method is considered the sole standard for calculating reference evapotranspiration. The inputs to the equation are climatic data, solar radiation, air temperature, humidity and wind speed. It allows the calculation of evapotranspiration at different times of the year and in different regions, yet a precise measurement at a given location can easily replace the simplified Penman–Monteith equation [17].

Other methods of calculating evapotranspiration include the use of empirical relationships, e.g., the relationship between observed evaporation from evaporation stations and meteorological quantities, these relationships can be calculated either linearly or non-linearly [25,26] using machine learning algorithms [27,28] linear regression or random forest regression.

The assessment of long-term climate variables can be based on time series. The time series is a sequence of measurements recorded over time, that can be analysed using,

e.g., Least-Squares Spectral Analysis, Least-Squares Wavelet Analysis, Least-Squares Cross Wavelet Analysis [29].

Other methods for evaluation may include parametric and non-parametric trend tests, which are used in machine learning [30,31]. The parametric method (logistic regression, linear discriminant analysis and simple neural network) use a fixed number of parameters to build models, require fewer variables and the result may be affected by outliers. The non-parametric method (the Mann–Kendall, Spearman’s Rho and k-Nearest neighbors) use a flexible number of parameters, both variable and attribute can be used in the models, the result is not affected by outliers.

In this paper, we explore the relationships for the calculation of evaporation from water surface in the Czech Republic using reanalyzed climate data and the constructed linear models (LM) and random forest models (RFM) for the calculation of evaporation. Evaporation estimated from the derived models was compared with observed evaporation from evaporation stations. Finally, the derived relationships were applied to the selected water reservoirs.

Specifically, we aim to answer the following questions: Which statistical method for calculating evaporation achieves better linear regression or random forest regression? How many variables are important for determining the formula for calculating evaporation? How important is the geomorphological information (elevation and location) for calculating evaporation using linear and non-linear models? The main objective of the evaporation estimation from water surface was to derive a universal relationship for the whole territory of the Czech Republic.

This paper is structured as follows: Section 2 introduces the area of interest and input data. The statistical method for evaluation evaporation with respect to goodness-of-fit (GOF) is evaluated in the R environment [32] and described Section 3. The results and discussion are in Section 4 along with a detailed evaluation of the goodness-of-fit (GOF) regression for evaporation stations and subsequently for water reservoirs. The paper is concluded in Section 5.

## 2. Study Area and Data

The study area is defined by the state border of the Czech Republic. Within the region (51°03' N to 48°33' N latitude and 12°05' E to E 18° 51' longitude) the long-term (1981–2010) mean annual precipitation totals at 709.5 mm, mean annual air temperature is 7.9 °C, mean runoff is 205.5 mm [33] and long-term runoff coefficient is thus 0.29 (29% of precipitation totals runs off).

Figure 1 describe long-term temperature, evaporation trend at evaporation station Hlasivo. The Hlasivo evaporation measuring station provides a consistent time series of 58 years, the evaporation values are measured by a 20 [m<sup>-2</sup>] benchmark evaporator. Other observed variables are: air temperature at 2 m [°C], water surface temperature in the evaporimeter [°C], relative humidity [%], global solar radiation [W·m<sup>-2</sup>] and wind speed at 2 m [m·s<sup>-1</sup>] [34].

Figure 2 shows the selected 24 evaporation stations and 33 water reservoirs. The evaporation stations were assigned to water reservoirs based on the Quitt classification and the elevation [35]. The elevation differences between the evaporation stations and water reservoirs do not exceed 100 m a.s.l. The Quitt classification divides the Czech Republic into three climatic regions (cold, moderately warm and warm regions), with an evaporation station in the same climatic region always assigned to a reservoir. The observed evaporation from the evaporation station was recorded between 1957 and 2019 (most evaporation station was recorded from 2005).

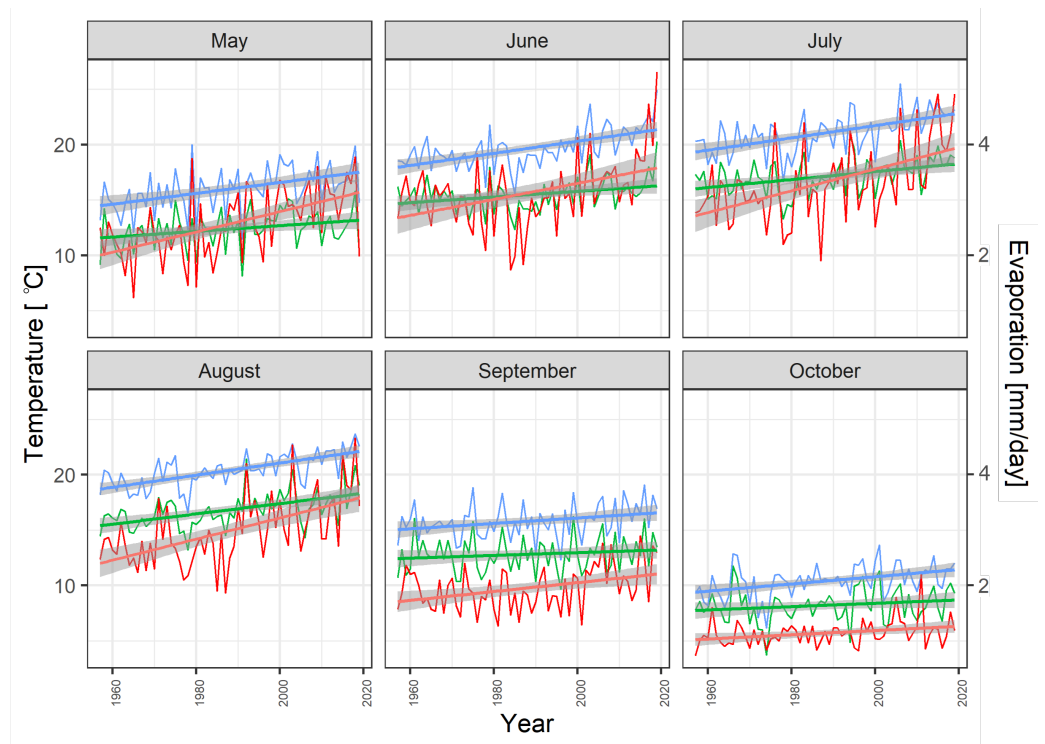


Figure 1. Mean monthly temperature (green lines), water temperature (blue line) and evaporation (red line) at Hlasivo evaporation station for period 1957–2019.

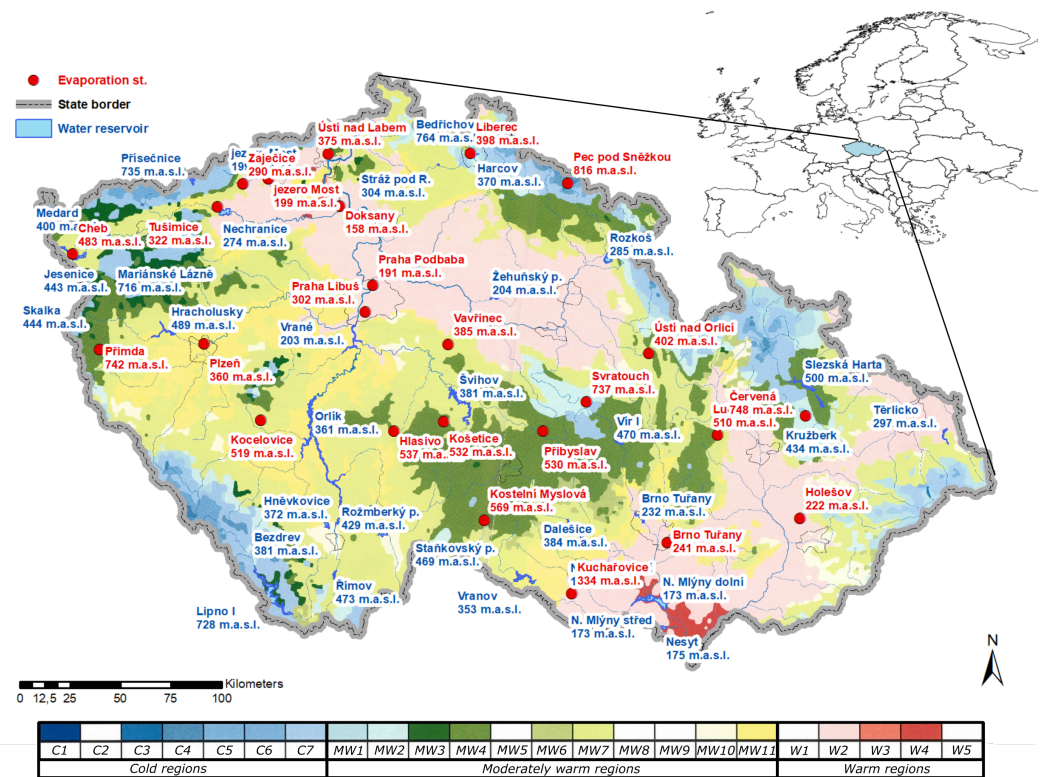


Figure 2. Study area: Czech Republic with climatic regions [35], blue color: water reservoirs with altitude of dam and red color: evaporation stations with altitude.

The data from the evaporimeter (EWM) were provided by the Czech Hydrometeorological Institute, Palivový kombinát Ústí, state-owned enterprise. The T. G. Masaryk Water

Research Institute, public research institution (TGM WRI, p.r.i.) provided data from the floating evaporator and data from the evaporation station Hlasivo.

Observed data from evaporation stations were aggregated into monthly step which were then used to evaluate evaporation from water reservoir surface, because the measured daily values are affected by random error [36]. The observed evaporation (may–october) is from 459 [mm·year<sup>-1</sup>] (Pec pod Sněžkou) to 760 [mm·year<sup>-1</sup>] (Holešov), mean evaporation (from evaporation stations) 627 [mm·year<sup>-1</sup>], minimum mean daily rate (1.38 [mm·year<sup>-1</sup>]) is in October and maximum mean daily rate is in July (4.53 [mm·year<sup>-1</sup>]), with maximum in June 2017 (5 stations exceeded 6.5 [mm·year<sup>-1</sup>]).

The relationships for calculating evaporation from the water surface were developed using linear and nonlinear regression. Measured evaporation from evaporation stations serves as the dependent variable. ERA5-Land climate reanalysis data were used for the non-dependent variables from 1981 to 2019.

#### *Climate Reanalysis*

The purpose of the reanalysis is to provide an estimate of quantities describing atmospheric, climatic and hydrological processes and behavior of oceans with global coverage and relatively high spatiotemporal resolution.

The reanalyses are outputs of various models, usually including a hydrological, atmospheric and ocean model and a model of the Earth's surface. The advantage is the provision of multidimensional spatially complete and coherent information about the global circulation and hydroclimatic quantities. Climate reanalyses are generated in a similar manner as in numerical weather forecasts, where the prediction models based on the development of the climate system from the initial state are used to predict the future state of the atmosphere. The initial state of the climate is a key input into the forecast determining the future development of the model simulation. Data assimilation is used to estimate the initial state that best matches the available data, while taking into account model errors. The climate reanalysis is performed as the only version of data assimilation that includes the use of the prediction model [37].

The reanalysis uses a combination of modeled data and observed data with emphasis on the laws of physics. The data are stored in the ECMWF archive and copied to the COPERNICUS Climate Data Store archive, from where they are freely downloadable using the CDS catalog or the CDS API application in the GRIB or NetCDF format.

The data was downloaded in NetCDF, which is a common format in drought or flood forecasting [38]. The spatial resolution is 0.1° × 0.1°, which represents approximately a grid of 9 km × 9 km.

The data set consisting of 2 m temperature [K], skin temperature [K], 2 m dew-point temperature [K], 10 m v-component of wind [m·s<sup>-1</sup>], surface pressure [Pa], surface net solar radiation [J·m<sup>-2</sup>] was selected to calculate evaporation from water reservoir. Temperature units [K] were converted to [°C] and energy units from [J·m<sup>-2</sup>] were converted to [W·m<sup>-2</sup>], values divided by the accumulation time expressed in seconds. Relative humidity [%] was calculated using the August–Roche–Magnus approximation [39], where the input data were dew point and temperature.

In the final dataset preparation, evaporation data from evaporation stations and geomorphological variables (elevation, latitude and longitude) were added to the reanalyzed data.

### **3. Methods**

Statistical methods of linear and non-linear regression (random forest regression) were used to evaluate evaporation from the water reservoir. In this case, the main objective of the regression is to determine the best fit between the observed values from the evaporation stations and the variables from the ERA5-Land project. The resulting linear and non-linear models were evaluated based on cross validation and goodness-of-fit (GOF): mean absolute error (MAE), root mean squared error (RMSE), coefficient of determination (R<sup>2</sup>)



and relative error (RERR). This section introduced building linear and non-linear models and their evaluating.

### 3.1. Linear Regression

Linear regression attempts to explain the values of a dependent variable through other quantities. In our case, an attempt was made to explain the dependent variable (evaporation value or evaporation rate from evaporimeter stations and evaporimeters EWM) using other variables (air temperature, surface temperature, wind, surface net solar radiation, dew point, pressure, latitude and altitude, evaporation type distribution) using 18 linear models created by sequential testing manually (8 models) and on the basis of stepwise regression (10 models).

The first set of models (built manually) was evaluated based on the Akaike Information Criterion (AIC) [40] value and the QQ plot was used for visual diagnostics [41]. The value of AIC is the sum of two terms, the first is proportional to the logarithm of the residual sum of squares, the second term is proportional to the complexity of the model (number of its members). When building the LM models, it can often happen that more independent variables reduce the sum of residues (improves the fit of the model with the observed data), however, this can result in an overfitted LM. The part of the AIC that penalizes the complexity of the model should prevent overfitting. When verifying the assumptions of the model (normality of residues), the QQ plot of residues can help. In the QQ plot of residues, two quantiles are plotted against each other—the theoretical quantile from distribution and the quantile with the actual residues of the model.

The second part of the linear models was developed using stepwise regression. R-packages *caret*, *leaps*, *MASS* [42] were used for this regression. The R-package *caret* uses the principle of machine learning and the R-package *leaps* are used to calculate the stepwise regression. The R-package *caret* has a function *train()*, which allows the implementation of a sequential selection of predictors, where the linear regression selection is selected:

- *leapBackward*,
- *leapForward*,
- *leapSeq*.

In this work, a method with backward selection was selected. The hyperparameter *nvmax* corresponds to the maximum number of predictors that are included in the model. In this work, 11 predictors were used. Furthermore, it is also possible to set the parameters of the validation method, in this work it was cross validation with 500 iterations.

### 3.2. Random Forest Regression

Random forest (RF) is a combined learning method for classification and regression that creates multiple decision trees during learning and then outputs the modus (most frequent value) of the classes returned by each tree to form a regression forest. The resulting regression function is defined as a weighted average of the regression functions of multiple trees. Regression forests belong to the so-called committee or ensemble methods, the main idea of which is to combine several separate models into a single ensemble. Thus, it uses the so-called collective decision [26,43]. A random forest consists of a set of trees  $T_1, \dots, T_N$  whose classification or regression functions can be expressed as follows:

$$h(\mathbf{X}, \mathbf{O}_1), \dots, h(\mathbf{X}, \mathbf{O}_N), \quad (1)$$

where  $h$  is a function,  $\mathbf{X}$  is a predictor and  $\mathbf{O}_1, \dots, \mathbf{O}_N$  are independent equally distributed random vectors. For the Random forests method, binary trees of type CART [44] are used. Similar to the creation of individual trees or other calibrations, a split into test and training sets is used. The R-package *randomforest* [27] was used in this work.

Random forest is an approach to build predictive models for both classification and regression tasks. It is a way to combine poorer performing baseline models to obtain better predictive models. Due to their simple nature, low assumptions and high performance,

RF models have been widely used in machine learning. The term “forest” refers to a set of decision trees that are themselves “weak” classifiers. A regression forest does not have the same predictive power as a stand-alone regression tree. If a single tree splits into a single criterion, it is very sensitive to changes. RF models classify variables based on their importance to achieve the best RF model [45].

### 3.3. Evaluation of Regression

Cross validation is used to improve the quality of regression models [46]. Depending on the method chosen, cross-validation is divided into k-fold cross validation, k-fold cross validation and leave-one-out. In our experiment, the method selected was leave-one-out validation. The dataset was split into training and test data, with one subset of data removed for the training data. The dataset consisted of the selected stations and in the training data the subset consisted of one sampled station, for a total of 24 stations, resulting in 24 iterations. Goodness-of-fit (GOF) criteria were used for further evaluation.

### 3.4. Evaluation of Regression by Goodness-of-Fit (GOF)

The linear regression and random forest regression set were evaluated based on their GOF ( $R^2$  [47], RMSE [48], MAE [48] and RERR [49]). This means that we would like to identify the best model which is the most suitable for the calculation of evaporation in the Czech Republic.

(i) The  $R^2$  is given by:

$$R^2 = 1 - \frac{RSS}{TSS}, \quad (2)$$

where  $RSS$  is the residual sum of squares and  $TSS$  the total sum of squares from predicted evaporation values  $Ep$  and of tested data of cross validation  $Et$ .

- $R^2$  indicates a measure of the quality of the regression model and explains the proportion of variability in the dependent variable of the model  $R^2$ , it may attain maximum value of 1, which means perfect prediction of the dependent variable. Conversely, value of 0 means that the model provides no information for understanding the dependent variable and is useless.

(ii) RMSE is given by:

$$RMSE = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n (Ep_i - Et_i)^2}, \quad (3)$$

where  $Ep_i$  is predicted evaporation values  $i$ -th case,  $Et_i$  tested data from cross validation and  $N$  is the total number of simulated values.

- It was used as the standard statistical metric providing a relatively high weight to large errors.

(iii) MAE is given by:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Ep_i - Et_i|, \quad (4)$$

The mean absolute error (MAE) is calculated as the average of the absolute differences between the predicted evaporation values  $Ep_i$  and tested data from cross validation  $Et_i$ .

- MAE is used to measure how close the predictions or forecasts are to the final results. ‘Absolute’ means that negative values are converted to positive values. The error is less sensitive to occasional very large errors because it does not amplify calculation errors.

(iv) RERR is given by:

$$\delta = \frac{|Ep - Et|}{Et}. \quad (5)$$

is the ratio of the absolute error between  $E_p$ -predicted evaporation values and  $E_t$ -tested data to the true of the value  $E_p$ -predicted evaporation values.

- It is a dimensionless quantity and can be given in percentages, it may attain both positive and negative values. Relative error can be used to compare quantities with different dimensions.

### 3.5. Final Evaluation of Regression Models

The last step of the evaluation was to create a scoring matrix and consecutively remove the models from the end (order of removal was from the worst models to the best). In order for the removal to occur, the individual models had to be ranked (from best to worst) or standardized using a GOF. Based on this procedure, the final evaluation was performed.

## 4. Results and Discussion

In this section, a detailed evaluation of linear and random forest regression with respect to GOF ( $R^2$ , RMSE, MAE and RERR) is presented. After evaluating all GOFs, RMSE was selected. Then, the best evaporation formulas are selected from the group of linear models (LM) and random forest models (RFM). Selected models were used to calculate evaporation from the water reservoirs.

### 4.1. Evaluation of Regression Models

Regression models LM and RFM were evaluated by cross-validation. The cross-validation procedure was as follows:

- In the training data, one station out of 24 stations was selected and validation of the inferred patterns from 23 stations was performed for this station.
- Validation was carried out successively for all stations and models.
- For validations, the goodness-of-fit  $R^2$ , RMSE, MAE and RERR were calculated.
- Based on the RMSE, the function of R [32] `rank()` was used, which lists the order of individual values corresponding in an ascending order to the sorted vector. After creating a unique identifier, a matrix was created where the models were on the  $x$ -axis and on the stations on the  $y$ -axis were. Based on this matrix, the best models were selected.

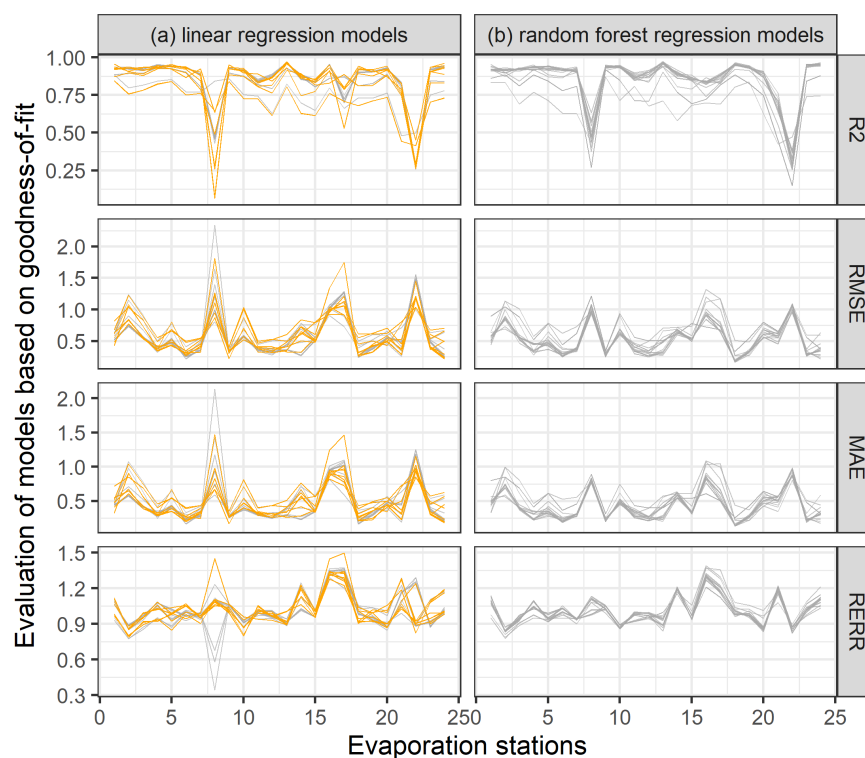
The models were evaluated and compared using GOF (see Figure 3). The results show that RF models can fit the data better than LM models. RF models are more consistent than LM models for all criterion functions. It can also be seen from the graph and results that for some stations the models do not achieve a good fit.

Outliers (the worst 10% GOF values) are present in all LM models, which also happens in RF models, but on a smaller scale. The outliers corresponded to 70% of the maximum value, thus setting the limit value for selected GOF. Table 1 shows evaporative stations that have exceeded the limit values for the selected GOF.

**Table 1.** Evaluation of evaporation stations based on GOF.

Goodness-of-Fit	Limit Values	Station above the Limit Value
$R^2$	<0.2	Hlasivo, reservoir Most
RMSE	<1.5	Hlasivo, reservoir Most, Praha Podbaba
MAE	<1	Hlasivo, reservoir Most, Praha Podbaba, Praha Libuš, Dukovany
RERR	<1.3	Hlasivo, reservoir Most, Praha Podbaba, Praha Libuš





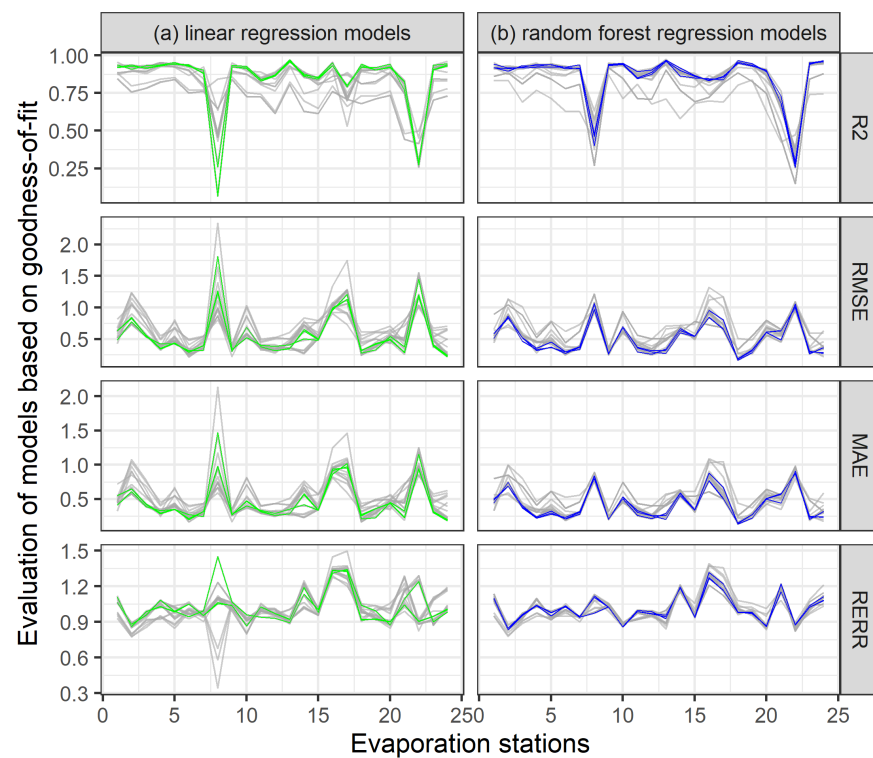
**Figure 3.** Model evaluation using GOF ( $R^2$ , RMSE, MAE, RERR). The lines in the plot represent the LM and RFM models. Part (a) linear regression models (18 models) is divided into two parts: orange line: models created manually (8 models), grey part: stepwise regression was used (10 models). Part (b) random forest regression models (15 models).

By the method of sorting, using the R function `rank()`, 3 linear and 3 random forest models (RFM) were selected. The selected regression models with average values of GOF are presented in Table 2.

**Table 2.** Average values of goodness-of-fit of selected models.

ID	$R^2$	RMSE	MAE	RERR
LM1	0.85	0.58	0.47	1.04
LM7	0.84	0.56	0.45	1.01
LM8	0.84	0.56	0.46	1.02
RFM4	0.86	0.51	0.42	1.02
RFM5	0.86	0.51	0.42	1.02
RFM15	0.86	0.51	0.42	1.01

The top 3 linear models according to all criterion functions are LM1, LM7 and LM8 and the top 3 RFM are RFM4, RFM5 and RFM15. The selected models are shown in Figure 4, green line represents linear models and blue line represents random forest models. The average value of RMSE for the selected linear models is 0.57, the minimum value is 0.22. The selected RFM had an average RMSE value of 0.51 and a minimum value of 0.18. The models that were designed based on stepwise regression achieved worse results than the models that were built manually based on data analysis. Models designed using manual regression achieved better results; however, some models designed using stepwise regression achieved good results in some cases, with less demanding inputs. The linear models were further supplemented with LM12, which also showed good results and the derived equation is more useful for practice due to its simplicity. All regression models are presented in the Sect. Appendix A in Table A1.



**Figure 4.** Evaluation of models using GOF, where the best models are selected. Part (a) linear regression models, green lines are LM1, LM7 and LM8. Part (b) random forest regression models, blue lines are RFM4, RFM5 and RFM15.

Selected regression formulas based on linear models:

$$LM1 = 45.84 + (0.173T \cdot (-0.004R))^{0.0008} - 0.183D - 0.0002P - 0.0002asl - 0.475Y + 0.063X, \quad (6)$$

$$LM7 = 16.97 + 0.082W + (0.235T \cdot (-0.263D))^{0.007} + 0.008R - 0.0003asl - 0.368Y + 0.063X, \quad (7)$$

$$LM8 = 17.33 + 0.055X - 0.367Y - 0.0003asl + (0.2134T \cdot (-0.277D))^{0.009} + 0.008R, \quad (8)$$

$$LM12 = 19.82 + 0.302ST + 0.006R - 0.170D - 0.419Y. \quad (9)$$

Selected variables for best random forest regression models:

$$RFM4 : W, (T * D), R, asl, Y, X, \quad (10)$$

$$RFM5 : X, Y, asl, (T * D), R, \quad (11)$$

$$RFM15 : W, T, ST, R, D, P, H, asl, Y, X. \quad (12)$$

where:

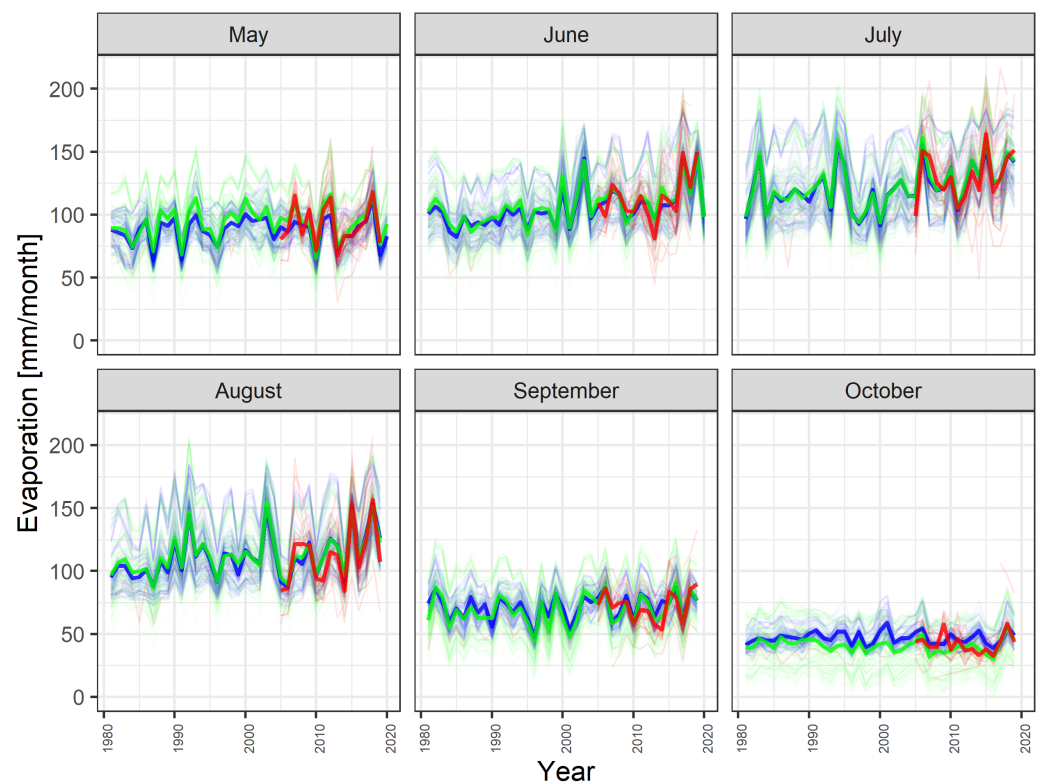
- LM1, LM7, LM8, LM12, RFM4, RFM5, RFM15 are formula identifiers for evaporation,
- T ... temperature (2 m) [°C],
- ST ... surface temperature [°C],
- P ... surface pressure [Pa],
- W ... wind speed [ $\text{m} \cdot \text{s}^{-1}$ ],
- R ... surface net solar radiation [ $\text{W} \cdot \text{m}^{-2}$ ],
- D ... dew point [°C],
- H ... relative humidity [%],
- asl ... elevation above sea level [m],
- X ... longitude,
- Y ... latitude.

#### 4.2. Model Application to Water Reservoirs

For testing, the best LM models (LM1, LM7, LM8 and LM12), RF models (RFM4, RFM5, RFM15) already described above were applied to selected reservoirs in the Czech Republic for the period May–October. The selection of the May–October period is because the evaporation from the observed data in the winter months is not measured due to

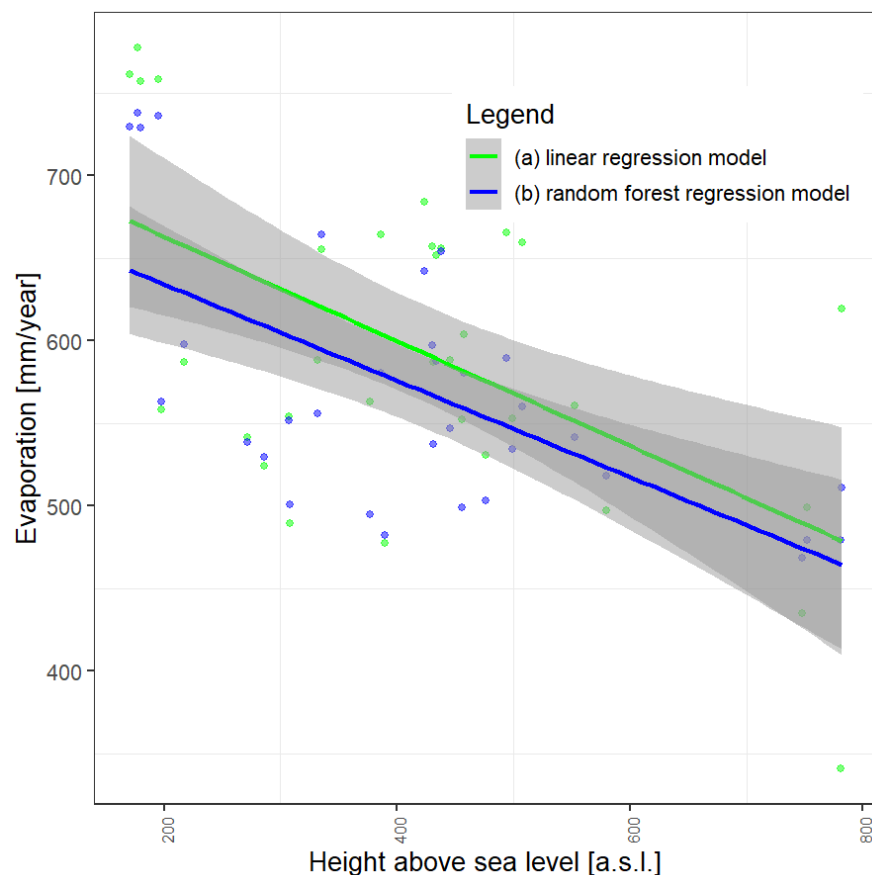
freezing. The calculated evaporation values for the water reservoirs are introduced in the Sect. Appendix A in Table A2.

The difference and seasonality in evaporation between the water reservoirs is described in Figure 5 where green lines represent linear models (LM), blue lines random forest models (RFM) and red lines introduce observed data. The average across all data is represented by the bold lines. The mean value from LM models and RFM models over the period (1981–2020) for reservoirs for May–October is 546.54 [ $\text{mm}\cdot\text{year}^{-1}$ ] and for RFM is 546.02 [ $\text{mm}\cdot\text{year}^{-1}$ ]. The mean value of the evaporation stations (2005–2019) is 497.26 [ $\text{mm}\cdot\text{year}^{-1}$ ]. The highest increase in evaporation is observed in the month of July, however, in the summer months (June–August) a significant increase in evaporation can be observed for all models.



**Figure 5.** Monthly evaporation for the water reservoirs throughout 1981–2019 (green lines represent linear regression, blue lines represent random forest regression and red lines observed data) and from evaporation stations throughout 2005–2019. Bold lines represent means for water reservoirs and evaporation’s stations.

Top models LM1 and RFM12 are compared with elevation for the whole water reservoir. The following Figure 6 shows the relationship between elevation and evaporation, where the green line represents linear regression model and blue line represents random forest model. The elevation of water reservoirs is 170.54–781.91 m a.s.l. The evaporation decreases with the elevation above sea level. Both models are influenced by local conditions because both models have input geographic coordinates and elevation.



**Figure 6.** The lines (a) and (b) represent the relationship between yearly total evaporation and elevation based on the derived formulas. The points represent yearly evaporation and altitude for both models.

The results of the study will be implemented to the hydrological model Bilan [50,51] and for assessing climate change studies in the Czech Republic [52].

## 5. Concluding Remarks

The main objective of the estimation of evaporation from the water reservoirs was to derive a universal relationship for the whole territory of the Czech Republic.

The estimation of evaporation from water reservoirs is complicated because a large number of water reservoirs do not have observed evaporation data. In this work, Quitt's climate classification was used to assign an evaporimeter station that is not near a reservoir to a given reservoir based on climate region and elevation. Within the Czech Republic, the evaporation value from water reservoirs is determined on the basis of a handling order, which is established according to a Czech technical standard which is based on old climatic data and does not deal with climate change. For this reason, the determination of the evaporation from water reservoirs is based on estimation using statistical methods rather than exact measurement.

The ERA5-Land climate reanalysis data were used for derivation and were chosen for their comprehensiveness, availability, high spatial resolution, long time series and advantageous management. Relative humidity was included into the results based on the calculated August–Roche–Magnus approximation. The climate reanalysis data were exported for stations and water reservoirs.

The derivation of the relationship for evaporation was based on the multiple linear regression method, where the values of the dependent variable (evaporation) were sought, based on two or more variables (predictors: air temperature, surface temperature, wind speed, surface net solar radiation, dew point, surface pressure, dew point, altitude, latitude,

longitude and calculated humidity). The construction of the models was done (i) manually, where the evaluation was done using the AIC parameter and the quantile–quantile (plot–QQ) was used for visual diagnostics, this method was time consuming, (ii) using stepwise regression, where the predictors are entered sequentially and models from one to  $X$ -selected variables were generated, this method is not time consuming. Random forest regression was used to account for non-linear relationships. Linear and random forest regression models were cross-validated and evaluated using criterion functions ( $R^2$ , RMSE, MAE and RERR). Finally, 3(+1) LM models and 3 RF models were selected. The models contained a large number of independent variables (6–7), possibly leading to model overfitting and therefore another model was selected which performed best for the RMSE criterion function and is based only on 4 independent variables and is therefore more user friendly.

It turned out that geomorphological information (elevation, location) appeared more in the manually derived models as opposed to models constructed using the stepwise regression method. When comparing linear models (LM) and random forest models (RFM), LM was found to have much more variability in the outcome compared to the RFM. The advantage of RFM is their adaptability, but the subsequent interpretation of the results can be a problem. This has been shown in the design of LM and RFM as well as when applying the proposed models to water reservoirs.

Evaporation values for the period 1981–2019 were calculated for the selected water reservoirs and selected formulas based on ERA5-Land climate reanalysis data.

For the evaluation of evaporation, models from LM and RFM models were used. Among the best models that were evaluated by linear regression, models LM1 from the manual linear regression group and LM12 from the stepwise regression group were used. Model LM1 was selected as the best model among the six predictors. The LM1 model can be replaced by an alternative model LM12 with which also performed satisfactorily with four predictors.

**Author Contributions:** Conceptualization, A.V., E.M. and M.H.; methodology, A.V., E.M. and M.H.; validation, A.V., E.M., M.H., P.P. and P.Š.; formal analysis, A.V., E.M., M.H. and P.P.; investigation, A.V.; resources, A.V.; data curation, A.V., E.M. and M.H.; writing—original draft preparation, A.V., P.P. and P.Š.; writing—review and editing, A.V., E.M., M.H. and P.P.; visualization, A.V., E.M., M.H. and P.P.; supervision, A.V. and M.H.; project administration, P.Š.; funding acquisition, A.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by T. G. Masaryk Water Research Institute, public research institution (VÚV TGM, v.v.i.) grant number 3600.52.26.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data underlying the analyses in this paper are available from the lead authors E.M., A.V., M.H. and P.P.

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

## Abbreviations

The following abbreviations are used in this manuscript:

LM	Linear Model
RFM	Random forest Model
GOF	goodness-of-fit
EWM	evaporimeter
ERA5-Land	climate reanalysis product
ECMWF	European Centre for Medium-Range Weather Forecasts
COPERNICUS	European Union’s Earth Observation Programme
CDS	Climate Data Store



API	Application Program Interfaces
GRIB	General Regularly-distributed Information in Binary form
NetCDF	Network Common Data Form
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
R <sup>2</sup>	Coefficient of Determination
RERR	Relative Error
AIC	Akaike Information Criterion
QQ	Quantile-Quantile Plot

## Appendix A

**Table A1.** Review of linear models and random forest models, where evaporation E [mm·month<sup>-1</sup>], temperature T (2 m) [°C], surface temperature ST [°C], wind speed W [m·s<sup>-1</sup>], surface net solar radiation R [W·m<sup>-2</sup>], dew point D [°C], relative humidity H [%], surface pressure P [Pa], elevation above sea level asl [m], longitude X, latitude Y.

ID	Manual Linear Regression
<b>LM1</b>	<b><math>E \sim (T * R) + D + P + asl + Y</math></b>
LM2	$E \sim T + (ST^2) + R + Y$
LM3	$E \sim H + W + T + ST + asl$
LM4	$E \sim W + T + asl + (X * Y)$
LM5	$E \sim W + T + (X * Y) + asl$
LM6	$E \sim W + T + R + (X * Y) + asl$
<b>LM7</b>	<b><math>E \sim W + (T * D) + R + asl + Y + X</math></b>
<b>LM8</b>	<b><math>E \sim X + Y + asl + (T * D) + R</math></b>
ID	Stepwise regression
LM9	$E \sim ST$
LM10	$E \sim ST + R$
LM11	$E \sim ST + R + Y$
<b>LM12</b>	<b><math>E \sim ST + R + D + Y</math></b>
LM13	$E \sim ST + R + D + asl + Y$
LM14	$E \sim ST + R + D + P + asl + Y$
LM15	$E \sim ST + R + D + P + asl + Y + X$
LM16	$E \sim W + ST + R + D + P + asl + Y + X$
LM17	$E \sim W + T + ST + R + D + P + asl + Y + X$
LM18	$E \sim W + T + ST + R + D + P + H + asl + Y + X$
ID	Random forest regression
RFM1	$E \sim (T * R) + D + P + asl + Y$
RFM2	$E \sim H + W + T + ST + asl$
RFM3	$E \sim W + T + asl + (X * Y)$
<b>RFM4</b>	<b><math>E \sim W + (T * D) + R + asl + Y + X</math></b>
<b>RFM5</b>	<b><math>E \sim X + Y + asl + (T * D) + R</math></b>
RFM6	$E \sim ST$
RFM7	$E \sim ST + R$
RFM8	$E \sim ST + R + Y$
RFM9	$E \sim ST + R + D + Y$
RFM10	$E \sim ST + R + D + asl + Y$
RFM11	$E \sim ST + R + D + P + asl + Y$
RFM12	$E \sim ST + R + D + P + asl + Y + X$
RFM13	$E \sim W + ST + R + D + P + asl + Y + X$
RFM14	$E \sim W + T + ST + R + D + P + asl + Y + X$
<b>RFM15</b>	<b><math>E \sim W + T + ST + R + D + P + H + asl + Y + X</math></b>

**Table A2.** Average year evaporation [mm·year<sup>-1</sup>] values for selected reservoirs by model.

	Water Reservoir	LM 1	LM 7	LM 8	LM 12	RF 4	RF 5	RF 15
1	Mariánské Lázně	456.57	406.37	407.53	460.76	448.25	441.19	440.32
2	Medard	517.42	419.89	420.58	445.50	464.99	461.08	452.22
3	Nesyt	734.47	738.63	738.35	720.95	693.63	694.18	693.05
4	Rožmberk	621.84	577.57	579.15	581.44	581.58	576.09	555.99
5	Staňkovský pond	629.94	570.13	570.85	586.79	566.87	562.56	550.49
6	Bezdrv	629.52	568.50	568.80	564.59	571.94	567.83	564.59
7	jezero Most	507.42	470.84	468.07	449.22	540.34	546.20	510.86
8	Bedřichov	302.96	379.47	381.93	398.94	442.42	435.92	442.27
9	Brno	612.95	625.37	626.29	606.97	627.63	625.78	609.18
10	Dalešice	617.56	606.14	606.28	607.93	608.56	607.86	597.58
11	Harcov	437.17	405.10	408.10	404.30	452.33	447.10	445.74
12	Hněvkovice	616.43	565.34	566.31	570.69	569.44	565.61	546.01
13	Nové Mlýny dolní	715.26	723.92	724.69	703.13	690.91	692.57	689.31
14	Nové Mlýny horní	716.69	718.96	719.29	696.68	684.96	685.64	682.96
15	Nové Mlýny střed	719.17	718.65	719.25	693.89	685.19	685.79	682.40
16	Orlík	566.86	536.74	536.87	551.54	557.05	557.89	539.68
17	Přísečnice	392.88	364.87	363.56	404.01	448.05	445.16	429.48
18	Rozkoš	483.40	480.53	481.91	453.28	514.23	513.06	489.23
19	Skalka	485.34	421.82	425.00	460.64	472.03	467.27	459.67
20	Slezská Harta	455.25	487.71	483.71	463.05	507.73	511.28	480.48
21	Stráž pod Ralskem	447.56	439.12	441.68	430.19	464.72	459.26	462.60
22	Těrlicko	510.93	547.67	546.09	492.15	538.75	537.80	507.56
23	Vranov	646.21	617.51	616.28	621.22	602.25	599.90	590.68
24	Vrané	547.91	543.23	540.54	543.30	512.07	509.25	520.30
25	Vír I	521.65	513.87	512.16	510.30	528.49	531.97	504.17
26	Hracholusky	543.98	491.44	490.70	513.37	500.41	497.08	491.13
27	Jesenice	507.08	421.91	424.18	456.80	468.90	464.60	456.41
28	Kružberk	508.67	503.72	499.90	480.90	512.99	515.47	489.92
29	Lipno I	586.06	508.19	510.88	541.07	516.80	512.38	476.28
30	Nechranice	493.24	479.92	476.84	465.56	502.13	500.34	490.44
31	Římov	625.90	558.38	559.47	564.04	560.89	557.42	524.97
32	Švihov	546.86	527.77	526.51	533.85	513.68	512.63	506.66
33	Žehuňský pond	536.57	564.87	563.07	548.24	543.89	539.64	551.57

## References

- Beran, A.; Hanel, M.; Nesládková, M.; Vizina, A. Increasing water resources availability under climate change. *Procedia Eng.* **2016**, *162*, 448–454. [\[CrossRef\]](#)
- Možný, M.; Trnka, M.; Vlach, V.; Vizina, A.; Potopová, V.; Zahradníček, P.; Štěpánek, P.; Hájková, L.; Staponites, L.; Žalud, Z. Past (1971–2018) and future (2021–2100) pan evaporation rates in the Czech Republic. *J. Hydrol.* **2020**, *590*, 125390. [\[CrossRef\]](#)
- Tang, Q.; Gao, H.; Lu, H.; Lettenmaier, D.P. Remote sensing: Hydrology. *Prog. Phys. Geogr.* **2009**, *33*, 490–509. [\[CrossRef\]](#)
- Schultz, G.A.; Engman, E.T. *Remote Sensing in Hydrology and Water Management*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
- Calera, A.; Campos, I.; Osann, A.; D'Urso, G.; Menenti, M. Remote sensing for crop water management: From ET modelling to services for the end users. *Sensors* **2017**, *17*, 1104. [\[CrossRef\]](#)
- Beran, A.; Hanel, M.; Nesládková, M.; Vizina, A.; Kožíň, R.; Vyskoč, P. Climate change impacts on water balance in Western Bohemia and options for adaptation. *Water Supply* **2019**, *19*, 323–335. [\[CrossRef\]](#)
- Chahine, M.T. The hydrological cycle and its influence on climate. *Nature* **1992**, *359*, 373–380. [\[CrossRef\]](#)
- Zhao, L.; Xia, J.; Xu, C.; Wang, Z.; Sobkowiak, L.; Long, C. Evapotranspiration estimation methods in hydrological models. *J. Geogr. Sci.* **2013**, *23*, 359–369. [\[CrossRef\]](#)
- Beran, A.; Kašpárek, L.; Vizina, A.; Šuhájková, P. Water loss by evaporation from free water surface. *Vodohospod. Tech.-Ekon. Inf.* **2019**, *61*, 12–18.
- Šuhájková, P. Evaporation from VÚV TGM evaporation stations. *Vodohospod. Tech.-Ekon. Inf.* **2020**, *5*, 16–27.
- Markonis, Y.; Kumar, R.; Hanel, M.; Rakovec, O.; Máca, P.; AghaKouchak, A. The rise of compound warm-season droughts in Europe. *Sci. Adv.* **2021**, *7*, eabb9668. [\[CrossRef\]](#)
- Rodell, M.; Beaudoin, H.K.; L'ecuyer, T.; Olson, W.S.; Famiglietti, J.S.; Houser, P.R.; Adler, R.; Bosilovich, M.G.; Clayson, C.A.; Chambers, D.; et al. The observed state of the water cycle in the early twenty-first century. *J. Clim.* **2015**, *28*, 8289–8318. [\[CrossRef\]](#)

13. Glenn, E.P.; Huete, A.R.; Nagler, P.L.; Hirschboeck, K.K.; Brown, P. Integrating remote sensing and ground methods to estimate evapotranspiration. *Crit. Rev. Plant Sci.* **2007**, *26*, 139–168. [[CrossRef](#)]
14. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; others. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [[CrossRef](#)]
15. Srivastava, A.; Sahoo, B.; Raghuvanshi, N.S.; Singh, R. Evaluation of Variable-Infiltration Capacity Model and MODIS-Terra Satellite-Derived Grid-Scale Evapotranspiration Estimates in a River Basin with Tropical Monsoon-Type Climatology. *J. Irrig. Drain. Eng.* **2017**, *143*, 04017028. [[CrossRef](#)]
16. Singh, V.P.; Xu, C.Y. Evaluation and Generalization of 13 Equations for Determining Free Water Evaporation. *Hydrol. Process.* **1997**, *11*, 311–323. [[CrossRef](#)]
17. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration—Guidelines for Computing Crop Water Requirements-FAO Irrigation and Drainage Paper 56*; FAO: Rome, Italy, 1998; Volume 300, p. D05109.
18. Xu, C.Y.; Singh, V.P. Evaluation and generalization of radiation-based methods for calculating evaporation. *Hydrol. Process.* **2000**, *14*, 339–349. [[CrossRef](#)]
19. Xu, C.Y.; Singh, V.P. Evaluation and generalization of temperature-based methods for calculating evaporation. *Hydrol. Process.* **2001**, *15*, 305–319. [[CrossRef](#)]
20. Xu, C.Y.; Singh, V.P. Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. *Water Resour. Manag.* **2002**, *16*, 197–219. [[CrossRef](#)]
21. Almorox, J.; Senatore, A.; Quej, V.H.; Mendicino, G. Worldwide assessment of the Penman–Monteith temperature approach for the estimation of monthly reference evapotranspiration. *Theor. Appl. Climatol.* **2018**, *131*, 693–703. [[CrossRef](#)]
22. Jacobs, A.F.G.; De Bruin, H.A.R. Makkink’s equation for evapotranspiration applied to unstressed maize. *Hydrol. Process.* **1998**, *12*, 1063–1066. [[CrossRef](#)]
23. Budyko, M.I. *Climate and Life*; Academic Press: Cambridge, MA, USA, 1974.
24. Moteva, M.; Kazandjiev, V.; Zhivkov, Z.; Kireva, R.; Mladenova, B.; Matev, A.; Kalaydzhieva, R. Estimation of Crop Evapotranspiration in Bulgaria Climate Conditions. *Sci. Pap.-Ser. A Agron.* **2014**, *57*, 255–263.
25. Beran, A.; Vizina, A. Derivation of regression equations for calculation of evaporation from a free water surface and identification of trends in measured variables in Hlasivo station. *Vodohospod. Tech.-Ekon. Inf.* **2013**, *4*, 4–8.
26. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
27. Ishwaran, H.; Kogalur, U.B. *Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC)*; R Package Version 2.10.1; 2021. Available online: <https://rdrr.io/cran/randomForestSRC/man/rfsrc.html> (accessed on 25 August 2021).
28. Kumar, M.; Bandyopadhyay, A.; Raghuvanshi, N.S.; Singh, R. Comparative study of conventional and artificial neural network-based ETo estimation models. *Irrig. Sci.* **2008**, *26*, 531–545. [[CrossRef](#)]
29. Ghaderpour, E.; Vujadinovic, T.; Hassan, Q.K. Application of the Least-Squares Wavelet software in hydrology: Athabasca River Basin. *J. Hydrol. Reg. Stud.* **2021**, *36*, 100847. [[CrossRef](#)]
30. Shadmani, M.; Marofi, S.; Roknian, M. Trend Analysis in Reference Evapotranspiration Using Mann-Kendall and Spearman’s Rho Tests in Arid Regions of Iran. *Water Resour. Manag. Int. J. Publ. Eur. Water Resour. Assoc. (EWRA)* **2012**, *26*, 211–224. [[CrossRef](#)]
31. Chen, Y.; Guan, Y.S.G.; Zhang, D. Investigating trends in streamflow and precipitation in Huangfuchuan Basin with wavelet analysis and the Mann-Kendall test. *Water* **2016**, *8*, 77. [[CrossRef](#)]
32. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.
33. Tolasz, R.; Míková, T.; Valeriánová, A.; Voženílek, V. Atlas Podnebí Česka. Palacký University Olomouc, Czech Hydrometeorological Institute. 2007. Available online: <https://www.geoinformatics.upol.cz/publikace/atlas-podnebi-ceska> (accessed on 25 August 2021).
34. Šuhájková, P.; Kožín, R.; Beran, A.; Melišová, E.; Vizina, A.; Hanel, M. Update of empirical relationships for calculation of free water surface evaporation based on observation at Hlasivo station. *Vodohospod. Tech.-Ekon. Inf.* **2019**, *61*, 4–11.
35. Vondráková, A.; Vávra, A.; Voženílek, V. Climatic regions of the Czech Republic. *J. Maps* **2013**, *9*, 425–430. [[CrossRef](#)]
36. Mrkvičková, M. Evaluation of the measurements at the Hlasivo evaporation station. *Vodohospod. Tech.-Ekon. Inf.* **2007**, *49*, 9–11.
37. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.A.; Balsamo, G.; Bauer, P.; et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 553–597. [[CrossRef](#)]
38. Muñoz-Sabater, J. ERA5-Land monthly averaged data from 1981 to present. *Copernic. Clim. Chang. Serv. (C3s) Clim. Data Store (CDS)* **2019**, *10*, 24381.
39. Alduchov, O.A.; Eskridge, R.E. Improved Magnus form approximation of saturation vapor pressure. *J. Appl. Meteorol.* **1996**, *35*, 601–609. [[CrossRef](#)]
40. Akaike, H. Factor analysis and AIC. In *Selected Papers of Hirotugu Akaike*; Springer: Berlin/Heidelberg, Germany, 1987; pp. 371–386.
41. Hadley, W. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016.
42. Bruce, P.; Bruce, A. Practical Statistics for Data Scientists: 50 Essential Concepts. 2017. Available online: <http://www.sthda.com/english/articles/37-model-selection-essentials-in-r/154-stepwise-regression-essentials-in-r/> (accessed on 11 March 2018).
43. Aleh, M. Regresní Stromy. Master’s Thesis, Univerzita Karlova v Praze, Staré Město, Czech, 2013.
44. Antipov, E.A.; Pokryshevskaya, E.B. Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. *Expert Syst. Appl.* **2012**, *39*, 1772–1778. [[CrossRef](#)]

45. Nhu, V.H.; Shahabi, H.; Nohani, E.; Shirzadi, A.; Al-Ansari, N.; Bahrami, S.; Miraki, S.; Geertsema, M.; Nguyen, H. Daily Water Level Prediction of Zrebar Lake (Iran): A Comparison between M5P, Random Forest, Random Tree and Reduced Error Pruning Trees Algorithms. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 479. [[CrossRef](#)]
46. Westerhuis, J.A.; Hoefsloot, H.C.; Smit, S.; Vis, D.J.; Smilde, A.K.; van Velzen, E.J.; van Duijnhoven, J.P.; van Dorsten, F.A. Assessment of PLS-DA cross validation. *Metabolomics* **2008**, *4*, 81–89. [[CrossRef](#)]
47. Nagelkerke, N.J. A note on a general definition of the coefficient of determination. *Biometrika* **1991**, *78*, 691–692. [[CrossRef](#)]
48. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [[CrossRef](#)]
49. Park, H.; Stefanski, L. Relative-error prediction. *Stat. Probab. Lett.* **1998**, *40*, 227–236. [[CrossRef](#)]
50. Vizina, A.; Horáček, S.; Hanel, M. Recent Developments of the Bilan Model. *Vodohospod. Tech.-Ekon. Inf.* **2015**, *57*, 7–10.
51. Melišová, E.; Vizina, A.; Staponites, L.R.; Hanel, M. The Role of Hydrological Signatures in Calibration of Conceptual Hydrological Model. *Water* **2020**, *12*, 3401. [[CrossRef](#)]
52. Vyskoč, P.; Beran, A.; Peláková, M.; Kožín, R.; Vizina, A. Preservation of drinking water demand from water reservoirs in climate change conditions. *Vodohospod. Tech.-Ekon. Inf.* **2021**, *63*, 4–18.