Limnol. Rev. (2020) 20, 1: 33-39
DOI 10.2478/limre-2020-0004

sciendo

The use of numerical methods in the design of a formula that returns the surface of the leaves of *Phragmites australis (Cav.) Trin. Ex Steud.*

Katarzyna Krzyżanowska[1,*], Paweł Krzyżanowski[2]

[1] Department of Limnology, University of Gdańsk, Bażyńskiego 4, 80-309 Gdańsk, Poland,
e-mail: katarzyna.krzyzanowska@ug.edu.pl (*corresponding author)
[2] Vector Technologies, Krzemowa 6, 81-577 Gdynia, Poland

Abstract: The paper presents the results of calculations and a verification of numerical models developed for estimating the surface of leaves of the common reed (*Phragmites australis (Cav.) Trin. Ex Steud.*). The research sample consisted of 137 leaves collected from the rush zone of Lake Raduńskie Górne in 2018. The total area of leaves obtained for testing was 1932.3 cm^2. To derive a formula that returns the surface of common reed foliage regression models were used – MLR (Multiple Linear Regression) and SLR (Stepwise Linear Regression). It has been shown that the measurement of basic leaf dimensions (i.e. length – L, mid-width – WM and maximum width – WX) makes it possible to define an empirical formula which, with an average accuracy of 99.9%, allows the real surface of leaves to be estimated. The modelling results were compared with formulas currently used in practice, and the measurement errors were determined using these formulas. It has been shown that the formulas used to date are subject to RMSE to the value of 1.19-2.52. The application of the developed formula (A = 0.4486 – 0.046 L + 7.9267 WM – 5.8121 WX + 0.5853 L · WX) will significantly reduce errors in leaf surface estimation (RMSE = 0.86) and thus the amount of reed transpiration and evapotranspiration, especially in the case of handling small samples (number of leaves and measurements).

Key words: phragmites australis, common reed, numerical methods

Introduction

Common reed (*Phragmites australis (Cav.) Trin. Ex Steud.*) is a grass plant (Srivastava et al. 2014), which according to the Raunkićr classification of plant life-forms belongs to cryptophytes, more precisely – helophytes, i.e. growing in mud (Raunkićr 1907). Like most grass plants, the reed is a hardy species, well adapted to a wide spectrum of environmental conditions (Ostendorp 1991). The geographical distribution of Phragmites spp. extends from cold temperate regions to the wetlands of hot and humid tropical zones (Lessmann et al. 2001), making it one of the few cosmopolitan species (Den Hartog et al. 1989; Clevering and Lissner 1999). The best environmental conditions for common reed are in transition zones (ecotones) of rivers, wetlands and lakes (Srivastava et al. 2014). Nearly 99% of lakes on earth have an extensive littoral zone (Wetzel 2001), which is a potential environment for the development of wetland vegetation (Thomaz and Cunha 2010). On the other hand, the common reed is a highly invasive species, forming thick patches, which often control the growth and development of native vegetation in coastal zones (Minchinton and Bertness 2003).

Because the expansion of common reed leads to a reduction of floristic diversity (Bresciani et al. 2011), in 2014-2015 studies were carried out on the transpiration of rushes and their impact on the water balance of lakes (Grabowska et al. 2016). These studies showed that the average daily volume of common reed transpiration during the study period was 3.9 mm, and the daily maxima reached 12.1-12.5 mm. The average monthly transpiration sums ranged from 42.8 (October) to 208.5 mm (August). During the growing season,

water losses in the transpiration process reached 872-971 mm. These values indicate that transpiration of common reed during the growing season is 60% higher than evaporation from an open water surface. In addition, the total water losses of common reed patches in the evapotranspiration process are twice as high as the water losses resulting from the evaporation of the water surface without rush vegetation (Grabowska et al. 2016). These results are confirmed by the conclusions of other authors who have studied the evapotranspiration of aquatic vegetation, including Otis (1914), Kuznecov (1949: after Uryvaev 1953), Kiendal (1953), Bernatowicz et al. (1976), Herbst and Kappen (1999), Xu and Ma (2011), Rösler et al. (2012).

Currently, there are many methods for assessing the size of evapotranspiration (ET), and direct methods and model approaches are considered the most representative (Xu and Ma 2011; Rana and Katerji 2000). Various models are commonly used to estimate ET, including Turc (1961), Priestley-Taylor (1972), Grabarczyk (1976), Linacre (1977), Hargreaves-Samani (1985) and Penman-Monteith (Allen et al. 1998). As a result of differences in assumptions, data requirements and different climatic conditions in which these models were developed, the results of evapotranspiration obtained using them may be incorrectly estimated. None of the currently functioning models includes the foliage surface as a variable determining the size of the evapotranspiration stream, even though the intensity of this process also depends on the number and size of leaves (Kacperska 2009).

The size of the transpiration flux can also be represented by calculating the transpiration intensity. In this case, it is necessary to determine the surface of the leaves (g dm^{-2} h^{-1}) (Czerwiński 1980). There are foliage scanning devices on the market. They can be successfully used for plants with relatively small transpiring organs. However, the mature leaves of common reed are too long for the scanner to accurately digitalise them. Moreover, the use of the device in the field is difficult or impossible due to changing weather conditions, to which the scanners are not always adapted. In the previously published studies, the authors (Królikowska 1971; Ho 1979; Herbst and Kappen 1999) refer to the surface of common reed foliage by providing empirical formulas by which their sizes were measured. So far, no papers have been published indicating the correctness of estimating the area of foliage using these formulas and describing the error accompanying the calculations.

To improve the measurement of leaf surface, experimental research was undertaken in 2014 at the Limnological Station of the University of Gdańsk in Borucino. In the middle of each month of the summer season (May-Oct) 2014, measurements were made of all the characteristic lengths of leaves contained in an evaporimeter. The result of the research was the implementation of a mathematical approximating function returning the surface of common reed foliage. The calculation of the dimensions of the foliage to the surface, however, was not described by a formula that would allow its widespread and straightforward use.

The main purpose of this study is, therefore: (*i*) to show that using measurements of morphological features (length (L), mid-width (WM) and maximum width (WX)) of common reed foliage together with numerical methods, it is possible to determine a formula that returns the surface of the leaves, and the obtained result will be burdened with the smallest possible error, (*ii*) to compare and verify the correctness of the estimation of the foliage area calculated by formulas published in the literature of the subject and earlier studies.

Study area

Field tests included a sampling of foliage (137 leaves) from a site located in the shore zone of Lake Raduńskie Górne, located in the eastern part of the Kashubian Lake District 54°15' N, 17°59' E) (Fig. 1). A sample was taken from an area located in the northern part of the lake.

According to Köppen, the sampling site is located in an area of humid continental climate with mild summer (Dfb) (Köppen 1936). In the light of Poland's climate regionalisation based on the occurrence of certain types of weather, the research area is located within the East Pomeranian region (Woś 1999).

Sampling was conducted on 14 and 15 July 2018, when common reed leaves were well developed. Two sampling dates were decided on due to the limited time necessary to digitalise the leaves. Over time, the collected leaves withered (curled), and their accurate scanning was difficult.

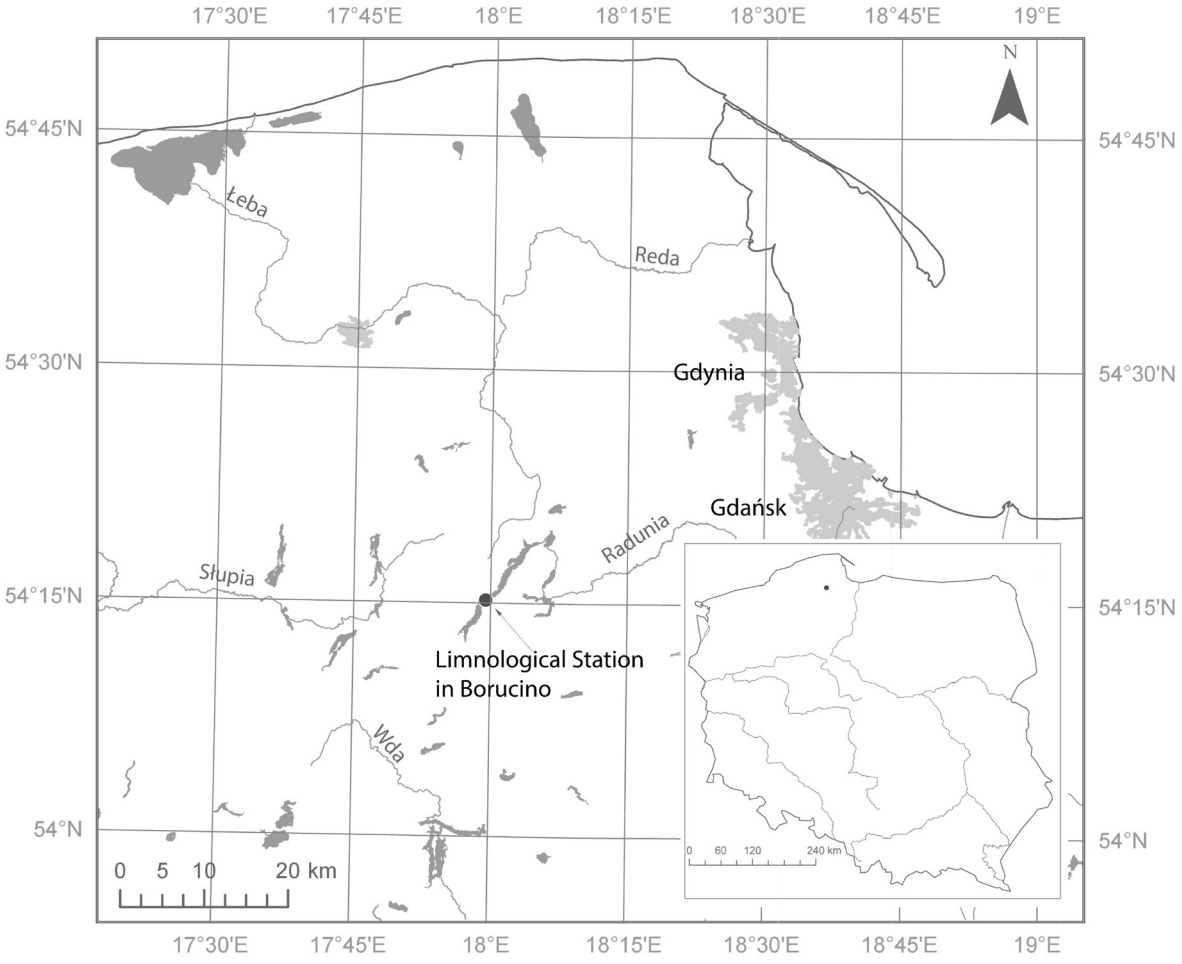

Fig. 1. Location of the sampling site

Methods

All leaves collected in the sample were scanned at 300 dpi. The exact surface (dependent variable – explained), as well as the length, maximum and mid-width (independent variables – explanatory variables), were determined in ImageJ. The measured morphological features of the leaves were subjected to a correlation analysis, which is a method of selecting variables. Using this method allowed the relationships between individual foliage features to be analysed.

Next, a series of mathematical analyses were performed in the MATLAB programming environment. Numerical calculations were based on the quantitative approach to the relationship between many explanatory variables and the explained variable. Based on the explanatory variables or their combination, the designated functions estimated the surface of the foliage. The procedure for selecting the reed leaf surface function was divided into two stages corresponding to the methods used to determine the formulas. First, the Multiple Linear Regression (MLR) method was used, including calculations using the Least Absolute Residual (LAR) function. Then the Stepwise Linear Regression (SLR) method was applied.

The validation of the formulas determined by the MLR method was based on the division of data into two datasets – test and training (containing 69 and 68 leaves, respectively). The goodness of matching the function to the "learning" data was analysed based on the Sum of Squared Errors (SSE), correlation coefficient (R^2), adjusted R^2 (R^2 adj.) and Root Mean Square Error (RMSE). The validation of the formulas with respect to the test dataset was described using validation SSE (SSE_v) and validation RMSE ($RMSE_v$). The following formulas were used in

the calculations:

$$SE = \sum_{i=1}^{N}(x_i - \hat{x}_i)^2 \quad [1]$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i - x_i)^2} \quad [2]$$

$$R^2 = \frac{\sum_{i=1}^{N}(\hat{x}_i - x_i)^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad [3]$$

where: N – number of observations, $\hat{x}_i$ – value of the theoretical explained (modelled) variable, $\bar{x}_i$ – real value of the variable at the moment t, x - arithmetic mean of empirical values of the explained variable.

The goodness of matching the SLR function was determined by means of 5-fold cross-validation, where the sample was divided into five subsets, each of which was then used as a test dataset, while the others together as a training dataset. All five cross-validation results were analysed for compliance with training and test datasets. Therefore, five independent, well-correlated datasets with given functions were determined. In the last stage, all data were combined, and the SLR function describing the leaf surface was calculated on the original sample.

Results

The matrix graph and correlation coefficients for common reed leaves taken from the coastal zone are presented in Figure 2. The assumption of the regression models that were used is to assume the number of observations n greater or equal to the number of estimated parameters, i.e. $n \geq k + 1$ (Stanisz 2007). Without this condition, it would be impossible to solve the system of equations that guarantee the existence of parameters. In order to draw substantive conclusions, it is essential for the number n to be many times greater than the number of estimated parameters – the greater the number of observations, the greater the precision of estimating the parameters of the function. Roscoe (Roscoe 1975: after Sekaran 2003) proposes rules that suggest taking a sample from 30-500, which is sufficient for most studies, and if multiple regression is used, accepting a sample several times larger (10 or more preferred) than the number of variables in the research.

The list of morphological characteristics of the foliage taken in the sample is summarised in Table 1.

The matrix graph and correlation coefficients (Fig. 2) for common reed leaves taken from the

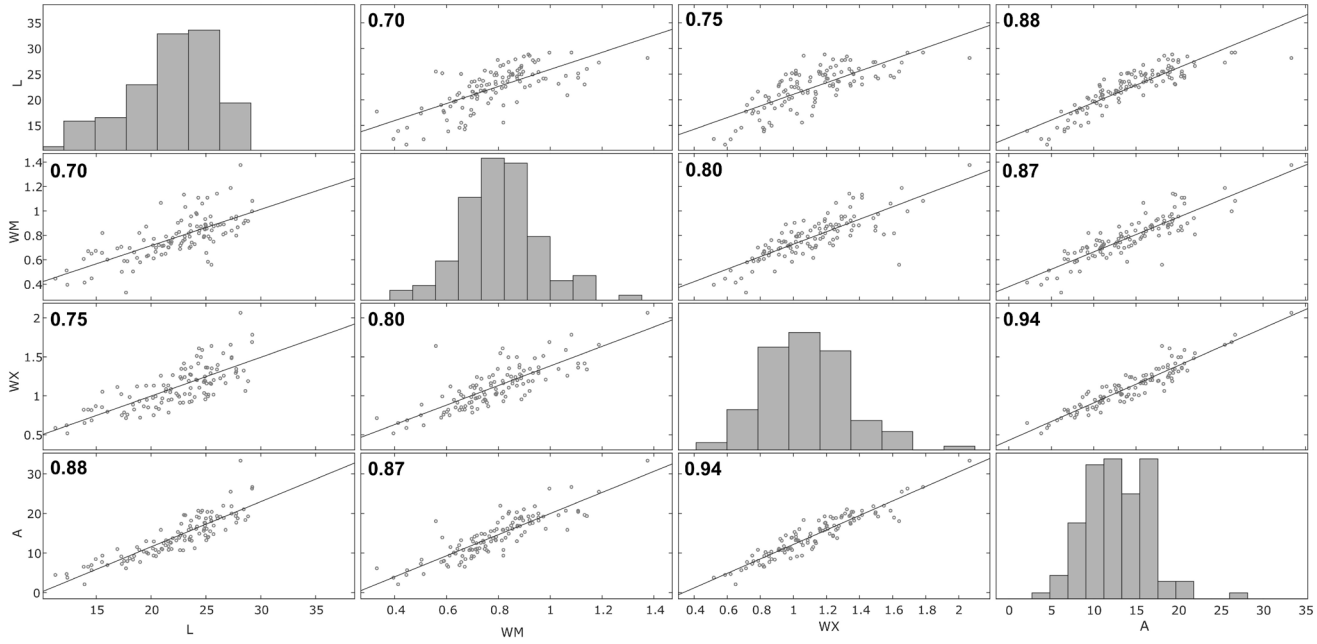

Fig. 2. Matrix charts and correlation coefficients of individual morphometric features of common reed leaves taken from the shore zone of Lake Raduńskie Górne

edge zone indicate that all linear leaf dimensions are highly positively correlated with its surface (R2 ∈ <0.70, 0.94>).

First, the MLR method was used to determine the formula. The results showed that the best matches are obtained for polynomial functions. The derived polynomial functions were then validated. At the validation stage, the LAR function was rejected, which despite the best fit to the training dataset (R2 = 0.989, RMSE = 0.5677), does not work for prediction. RMSE for the training dataset is much lower than for the "test" set. The difference at the level of 0.726 (more than twice the value of RMSE) is a typical example of excessive model matching (the model is best suited to the input data (training dataset), but thus it is not possible to use it for another dataset).

Two functions determined by the MLR method were selected for further analysis – $L_WX_{reduced}$ [4] and $L_WX_{reduced2}$ [5]. Both were determined based on the maximum length and width. However, they differed in the number of coefficients appearing in the function equation. This is the case of the function $L_WX_{reduced2}$ which boils down to its simplification. Both selected functions have high corrected correlation coefficients R2, and the difference from the function L_WX (with more coefficients) is small (0.07 for function A and 1.17% for function B). In addition, the differences between the RMSE values for the training and test formula are the smallest among all for the formulas (0.1526 for function A and 0.1412 for function B). The models determined by the MLR method took the form of equations 4 and 5:

$$L_WX_{reduced} = -0.9132 + 0.1882\ L - 0.8298\ WX + 0.3721\ L \cdot WX + 1.809\ WX^2 \quad [4]$$

$$L_WX_{reduced2} = -11.33 + 0.4894\ L + 13.14\ WX \quad [5]$$

In the last step, the function was also calculated by the SLR method. The equation took the following form:

$$L_WX_WM = 0.4486 - 0.046\ L + 7.9267\ WM - 5.8121\ WX + 0.5853\ L \cdot WX \quad [6]$$

The correlation coefficient was $R^2 = 0.999$, and RMSE = 0.86.

The equation was determined on the basis of stepwise regression. At first, the initial model containing the explanatory variables was adjusted. In the next iteration, the variable with the lowest *p*-value (test probability) meeting the significance criterion was selected for the model. After adding the new variable in the next step, the significance level of all potential explanatory variables already in the model was validated. After validation, the regression algorithm again selected new variables. In this way, after several iterations, the model contained the minimum number of explanatory variables meeting the required significance criterion.

Discussion

To determine the effectiveness of prediction with the functions set out in this study in relation to the formulas used in literature (A_K – Królikowska 1971; equation 7; A_{HK} – Herbst and Kappen 1998, eqn 8; A_H – Ho 1979, eqn 9):

$$A_K = 0.57 \cdot L \cdot WX, \quad [7]$$

$$A_{HK} = 0.51 \cdot L \cdot WM + 5.7, \quad [8]$$

$$A_H = 0.55 \cdot L \cdot WX, \quad [9]$$

and the numerical function (approximation) prepared as part of previous, preliminary research (function in the form of code implemented into the MATLAB programming environment), the area of all leaves taken from the lakeshore zone by the above-mentioned methods was calculated (Table 2).

The best data compliance (99.99%) was obtained using the L_WX_WM (SLR) function (Table 2). This function also shows the smallest RMSE with a standard deviation (RMSE = 0.86, std.dev = 2.33) (Fig. 3), which means that the surface calculated by the formula is the closest to the real surface of all the functions presented. The weakest match is obtained by the approximation method, which overstates the results by an average of nearly 10%. It has been shown that formulas functioning in the literature of the subject generally map the surface of the foliage. The formula of Ho (1979) returns the results closest to real (99.2% compliance, RMSE = 1.19, std.dev. RMSE = 3.21), the formula by Królikowska (1971) overestimates the results by nearly 3%, while the Herbst and Kappen function (1998) has the weakest prediction, which exaggerates the results by 5.3% and on average the surface differs from the actual by 2.52 cm^2. The results for the two recommended functions L_WX_WM and $L_WX_{reduced2}$ available in the literature by Ho (1979) reflect the surface of the

Table 1. Average total length (L), mid-width (WM), maximum width (WX), average leaf length at the maximum width and its percentage relative to total length (L (WX)), and average real area (A) of common reed leaves taken in the sample

L	WM	WX	L(WX)	L(WX)	A
[cm]	[cm]	[cm]	[cm]	[%]	[cm^2]
22.40	0.78	1.10	6.24	27.9	14.10

Table 2. List of polynomial functions returning the leaf surface with their validation. Abbreviations used in function names: L – leaf length, WX – maximum width, WM – leaf mid-width, LAR – Least Absolute Residuals, match type (first and second-degree polynomial) successively describe the variables given in the match name, SSE – Sum Square Errors, R^2 – determination coefficient, RMSE – Root Mean Squared Error

Match name	Match type	SSE	R^2	R^2 adj.		No. of regression coefficients	SSE_v	$RMSE_v$
$L_WX_{reduced}$	1st and 2nd degree polynomial	77.559	0.959	0.956	1.101	5	106.820	1.253
$L_WX_{reducedLAR}$	1st and 2nd degree polynomial	20.629	0.989	0.988	0.568	5	113.798	1.294
L_WX	1st and 2nd degree polynomial	75.123	0.960	0.957	1.092	6	117.877	1.317
$L_WX_{reduced2}$	1st and 2nd degree polynomial	102.067	0.946	0.944	1.247	3	130.410	1.385
L_WM	1st and 2nd degree polynomial	97.380	0.948	0.944	1.243	6	271.716	1.999
WM_WX	1st and 2nd degree polynomial	145.391	0.922	0.916	1.519	6	297.869	2.093

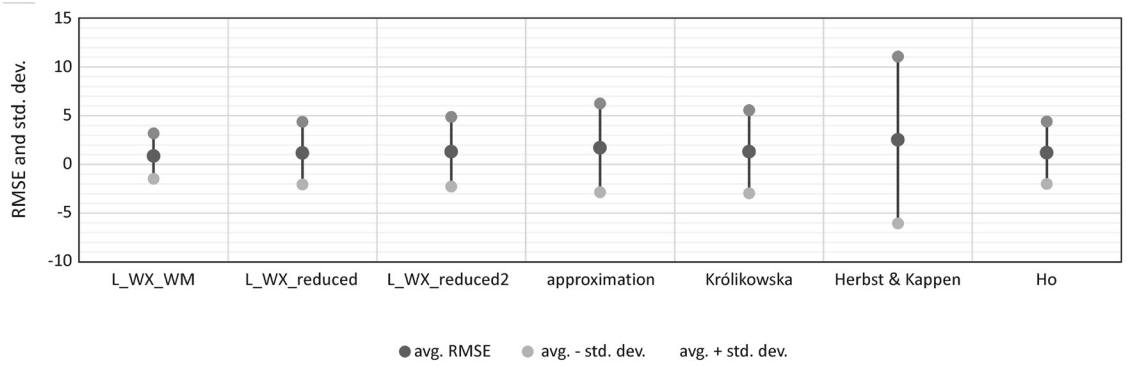

Fig. 3. RMSE with standard deviation for functions returning the surface of common reed leaves

collected research sample very well, and the differences between them – below 1%, might be due to the inaccuracy of measurements and the method of data processing, i.e. foliage surface for their digital numerical value.

Conclusion

To sum up, the formula determined in this study (L_WX_WM) best reflects the real surface of common reed foliage. It is based on three morphometric characteristics: leaf length, mid-width and maximum width. One of the formulas calculated by the MLR method ($L_WX_{reduced}$, $L_WX_{reduced2}$), which uses less morphometric characteristics, i.e. maximum length and width, can be used to determine the surface, which simplifies measurement procedures. All three formulas obtained in this study show a very high consistency of prediction. In addition, the results obtained with the functions $L_WX_{reduced}$ and L_WX_WM, representing the lowest RMSE, can have a significant impact on the estimation of leaf area when there is limited data (small research sample).

References

Allen R.G., Pereira L.S., Raes D. Smith M., 1998, Crop evapotranspiration: Guidelines for computing crop water requirements, FAO Irrigation and drainage paper No. 56, Food and Agriculture Organization (FAO), Rome, 300 pp.

Bernatowicz S., Leszczyński S., Tyczyńska S., 1976, The influence of transpiration by emergent plants on the water balance in lakes, Aquat. Bot. 2: 275–288.

Bresciani M., Stgia C., Fila G.L., Musanti M., Bolpagni R., 2011, Assessing common reed bed health and management strategies in Lake Garda (Italy) by means of Leaf Area Index measurements, Italian Journal of Remote Sensing 43(2): 9–22.

Clevering O.A., Lissner J., 1999, Taxonomy, chromosome numbers, clonal diversity and population dynamics of Phragmites australis, Aquat. Bot. 64(3-4): 185-208.

Czerwiński W., 1980, Fizjologia roślin (Plant physiology), Wydaw. Nauk. PWN, Warszawa, 604 pp (in Polish).

Den Hartog C., Květ J., Sukopp H., 1989, Reed. A common species in decline, Aquat. Bot. 35(1): 1–4.

Grabarczyk S., 1976, Polowe zużycie wody a czynniki meteorologiczne (Field water utilization and meteorological factors), Zesz. Probl. Post. Nauk Rol. 181: 495–511 (in Polish).

Hargreaves G.H., Samani Z.A., 1985, Reference crop evapotranspiration from temperature, Appl. Eng. Agric. 1(2): 96–99.

Herbst M., Kappen L., 1999, The ratio of transpiration versus evaporation in reed belt as influenced by weather conditions, Aquat. Bot. 63(2): 113–125.

Ho Y.B., 1979, Shoot development and production studies of phragmites Australis (Cav.) Trin. Ex Steudel in Scottish Lochs, Hydrobiologia 64(3): 215–222.

Kacperska A., 2009, Gospodarka wodna (Water management), [in:] Kopcewicz J., Lewak S. (eds), Fizjologia roślin (Plant physiology), PWN, Warszawa: 36–57 (in Polish).

Kiendl J., 1953, Zum Wasserhalt des Phragmitetum Communis und Glycerictum aquatiacac, Ber. Deutsch. Bot. Ges. 66: 246–263.

Köppen W., 1936, Das Geographische System der Klimate (The Geographical System of the Climate), [in:] Köppen W., Geiger R. (ed.), Handbuch der Klimatologie. Bd. 1 (Handbook of Climatology. Vol. 1), Gebrüder Bornträger, Berlin, 44 pp (in German).

Królikowska J., 1971, Transpiration of reed (Phragmites Communis Trin.), Pol. Arch. Hydrobiol. 18(4): 347–358.

Lessmann J.M., Brix H., Bauer V., Clevering O.A., Francisco A.C., 2001, Effect of climatic gradients on the photosynthetic responses of four Phragmites australis populations, Aquat. Bot. 69(2–4): 109-126.

Minchinton T.E., Bertness M.D., 2003, Disturbance-mediated competition and the spread of Phragmites Australis in a coastal marsh, Ecol. Appl. 13(5): 1400–1416.

Ostendrop W., 1993, Reed bed characteristics and significance of reeds in landscape ecology, Limnol. Aktuell 5: 149–161.

Otis C.H., 1914, The transpiration of emersed water plants: its measurement and relationships, Bot. Gaz. 58(6): 457–494.

Priestley C.H.B., Taylor R.J., 1972, On the assessment of surface heat flux and evapotranspiration using large scale parameters, Mon. Weather Rev. 100: 81–92.

Rana G., Katerji N., 2000, Measurement and estimation of actual evapotranspiration in the field under Mediterranean climate: a review, Eur. J. Agron. 13(2): 125–153.

Raunkiær C., 1907, Planterigets livsformer og deres betydning for geografien [Plant life forms and their importance to geography], Nordisk Forlag, København and Kristiania, 132 pp (in Danish).

Roscoe J.T., 1975, Fundamental research statistics for the behavioral science, Holt, Rinehart and Winston Inc., New York, 483 pp.

Rösler A., Chmal M., Chmal T., 2012, Evapotranspiration from a reed bed as example of oasis effect, [in:] Grześkowiak A., Nowak B., Grzonka B. (eds), Anthropogenic and natural transformations of lakes. Vol. 6, Wydaw. IMGW-PIB, Poznań: 131–135.

Sekaran U., 2003, Research methods for business: A skill-building approach, Wiley, New York, 450 pp.

Srivastava J., Kalra S.J.S., Naraian R., 2014, Environmental perspectives of Phragmites australis (Cav.) Trin. Ex. Steudel, Appl. Water Sci. 4: 193–202.

Stanisz A., 2007, Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 2: Modele liniowe i nieliniowe (Statistic course using STATISTICA software on examples of medicine. Vol. 2: Linear and nonlinear models), Kraków, StatSoft Polska, 868 pp (in Polish).

Thomaz S.M., Cunha E.R., 2010, The role of macrophytes in habitat structuring in aquatic ecosystems: methods of measurement, causes and consequences on animal assemblages' composition and biodiversity, Acta Limnol. Bras. 22(2): 218–236.

Turc L., 1961, Water requirements assessment of irrigation, potential evapotranspiration: Simplified and updated climatic formula, Ann. Agron. 12: 13–49.

Uryvaev V.A., 1953, Eksperimentalnye issledovania na Valdae (Experimental research in Valdai), Gidrometeoizdat, Leningrad, 232 pp (in Russian).

Wetzel R.G., 2001, Limnology. Lake and river Ecosystems, Academic Press, San Diego, 1066 pp.

Woś A., 1999, Klimat Polski (Climate of Poland), PWN, Warszawa, 301 pp (in Polish).

Xu S., Ma T., 2011, Evapotranspiration observation and data analysis in reed swamp wetlands, IAHS Publ. No. 344, Wallingford: 239–244.