

A gap filling method for active surface heat balance structure

Jacek Leśny^{1,*}, Monika Panfil²

¹Poznań University of Life Sciences, Meteorology Department, 60-649 Poznań, Polska

²University of Warmia and Mazury in Olsztyn, Department of Water Resources, Climatology and Environmental Management, 10-719 Olsztyn, Polska

Abstract. The paper describes gap filling procedures for active surface heat balance structure data recorded for fields of rape, maize, spring and winter wheat and an apple orchard. The balance components were determined based on the Bowen ratio requiring direct measurements of net radiation, soil heat flux, temperature and water vapour pressure profiles. The latter is used to determine vertical gradients and the Bowen ratio, with sensible and latent heat fluxes calculated from the heat balance equation. Missing data are filled in from regression dependencies between individual balance components at various measurement sites. The regression data set comprised results recorded over 24 h, before the gap in measurements and after 24 h. Multiple regressions were determined from a 48-h measurement set. Regression was applied to establish missing values of net radiation (R_n), soil heat flux (G) and latent heat (LE), while sensible heat was calculated from the active surface heat balance equation. Relatively the greatest differences were found for latent heat and soil fluxes, with both estimated values deviating by 13% from the measured daily average, for net radiation the relative difference was 10% and for sensible heat – 6%. This method successfully filled gaps in measured heat balance data from April to September.

1 Introduction

When conducting any research, particularly requiring field measurements, we need to assume that for various reasons errors or gaps in measurements will appear in data series. Typical field measurements include all types of weather observations. In this case, apart from gaps in measurement series, we face a much more serious problem related with insufficient density of the measurement grid. Thus in order to obtain the field of values based on single measurement points it is necessary to apply interpolation [1–3]. For some types of data techniques using satellite positioning (GNSS) may be used [4]. In the cases when the variability of values is well known, such as e.g. fluctuations in solar radiation intensity, they may be approximated using the pre-determined formulas [5] or – if statistical parameters are known – data may be generated [6]. Then an artificial set of data is obtained, which still retains all properties of the set originating from actual measurements.

Our paper describes procedures filling gaps in measurement data in the heat balance structure for the active surface, recorded in the fields of rape, maize, spring and winter wheat as well as an apple orchard [7]. The heat balance structure very well presents a synthetic picture of the functioning of the active surface, while it describes not only the exchange of energy, but also mass (latent heat flux is directly related with evapotranspiration) and further on this basis the exchange of carbon dioxide may also be

approximated [8–10]. Various measurement methods may be applied to assess fluxes of mass and energy: the profile method, the Bowen ratio heat balance, eddy covariance and the chamber method [7, 11, 12, 13]. It is also possible to model [13, 15, 16] or estimate fluxes, e.g. evapotranspiration based on satellite photographs [17]. In the case of flux measurements, it is practically impossible to obtain a continuous set of data, both for technical reasons and the requirements of these methods. For example, eddy covariance requires for the flow of air to be turbulent and if the test results do not confirm it, such results are considered unreliable and they are rejected.

2 Field studies

Errors and gaps were also observed in the analyses of the heat balance structure, which results were used in this study. They were related both with technical causes (equipment failure, cultivation operations), weather conditions (e.g. data loggers switching off due to strong atmospheric discharges) and the human factor (errors of the observers and equipment being stolen).

As it has been mentioned above, measurements were taken for 5 different crops and they were planned to be conducted continuously and cover possibly the longest part of the vegetation season. The heat balance components were identified using the Bowen ratio method, which requires direct measurements of net each site were given. Since prototype equipment was radiation

* Corresponding author: jacek.lesny@up.poznan.pl

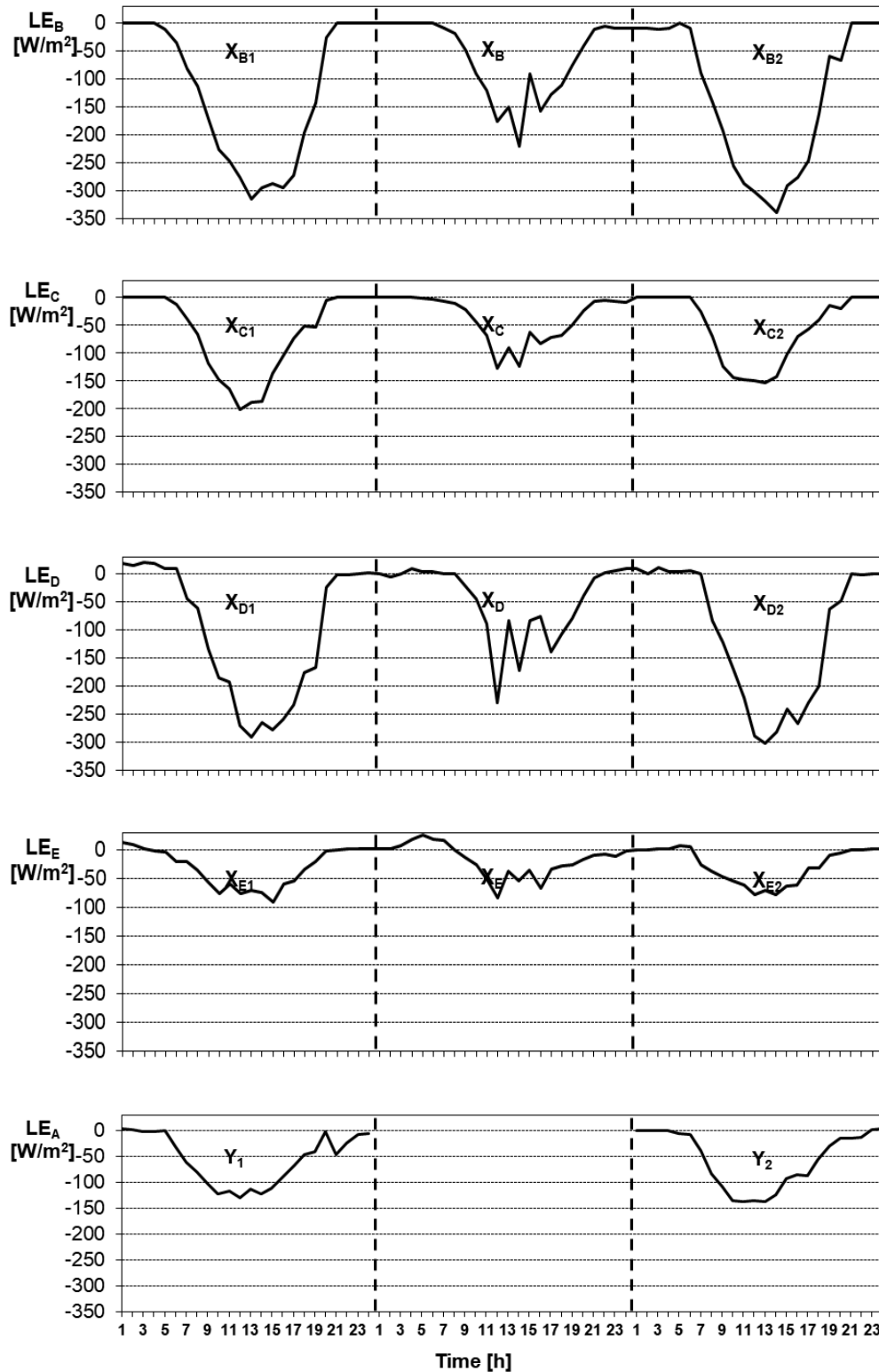


Fig. 1. The course of evapotranspiration in the following sites B, C, D, E and A from 3.07 to 5.07 2002, with the assumed gap in measurements in site A on 4.07.2002.

soil heat flux as well as temperature and watervapour pressure profiles. Based on the latter, vertical gradients and the Bowen ratio are established, while sensible and latent heat fluxes are calculated from the heat balance equation [7, 18, 19, 20, 21, 22, 23]. In order to determine the energy flux, measurements of all the above-

mentioned elements need to be conducted properly. Table 1 presents dates for the beginning and end of measurements at each of the sites and the percentages of the number of correct data obtained from used and recorded results were compared with literature information on the number of missing observations in

comparable field measurements, it seems that the obtained number of corrected data is satisfactory (particularly as these gaps typically lasted up to several hours, with the longest not exceeding 24 h).

Table 1. Dates of beginning and end of measurements in respective fields and the percentage of correct measurement data [7].

	Experimental surface				
	Rape	Maize	Orchard apple	Winter wheat	Spring wheat
Data	A	B	C	D	E
Beginning	26.04.02	26.04.02	30.04.02	30.04.02	28.04.02
Finish	1.10.02	1.10.02	1.10.02	1.10.02	1.10.02
Correct [%]	92%	88%	85%	95%	84%

3 Gap filling method

Analysis of collected data series was possible only after the missing information had been supplemented. For this purpose we used equations of regression taking place between values of individual heat balance components at various measurement sites.

The methodology was described based on evapotranspiration in the days from 3.07 to 5.07.2002, i.e. for 72 hours of measurements. During those days a complete set of data was collected; however, to ensure the reliability of the picture it was assumed that 24 measurements are missing at site A from 4.07.02, i.e. in the middle of that period (Fig. 1). It was decided that the gaps will be filled in using multiple linear regression of the following form:

$$y(x_B, x_C, x_D, x_E), \quad (1)$$

which parameters were established treating values (x_B, x_C, x_D, x_E) from set $\{X_{B1} \cup X_{B2}, X_{C1} \cup X_{C2}, X_{D1} \cup X_{D2}, X_{E1} \cup X_{E2}\}$ as an independent variable, while value (y) from set $Y_1 \cup Y_2$ as a dependent variable (Fig. 1). After determination of regression parameters values for the missing data were estimated:

$$y^{\wedge}(x_B, x_C, x_D, x_E), \quad (2)$$

for $(x_B, x_C, x_D, x_E) \in \{X_B, X_C, X_D, X_E\}$ (Fig 1).

Sets $X_{B1}, X_{B2}, X_{C1}, X_{C2}, X_{D1}, X_{D2}, X_{E1}, X_{E2}, Y_1, Y_2$ were always selected so that they comprised 24 values each. Sets $X_B, X_C, X_D, X_E, Y^{\wedge}$ had as many elements as there were in the gap in data. There may have been several or around a dozen values. As it was mentioned earlier, the gaps did not exceed 24 values.

In order to confirm the applicability of the described method, values obtained using multiple regression were compared with measured values. The presented procedure was applied for days from 3.07 to 5.07.2002, assuming that on 4.07 a gap occurred in the continuity of data and very high coefficients of determination were obtained for multiple regression coefficients. In order to approximate latent heat, soil heat and net radiation they were 0.94, 0.97 and 0.99, respectively. Figure 2 presents

a complete course of latent heat in the above-mentioned period, while for 4.07.02 additionally a course approximated using multiple regression is presented.

Table 2 presents values of absolute and relative errors for the estimated course of heat balance components on that day (4.07.02). For some hours relative errors reached very high values, particularly in the periods when fluxes themselves were slight in absolute terms.

Table 2. Values of absolute and relative errors for estimated course of heat balance components on 4.07.02.

	Successively hour lack of measurement								
	1	2	3	4	5	6	7	8	
Rn _A	9	1	11	3	6	3	1	4	
Rn _W %	37%	5%	142%	16%	52%	58%	7%	17%	
G _A	1	1	3	2	1	1	0	0	
G _W %	8%	9%	100%	30%	13%	8%	2%	1%	
LE _A	3	3	8	18	25	20	18	3	
LE _W %	109%	163%	189%	100%	87%	98%	131%	26%	
S _A	5	2	17	17	20	16	18	7	
S _W %	27%	18%	-	213%	253%	532%	180%	733%	
Next hours									
	9	10	11	12	13	14	15	16	
Rn _A	15	45	97	7	35	3	3	54	
Rn _W %	22%	31%	37%	2%	15%	1%	2%	22%	
G _A	0	0	0	0	4	10	2	6	
G _W %	3%	2%	1%	0%	25%	25%	31%	52%	
LE _A	3	9	87	42	33	11	13	64	
LE _W %	10%	16%	56%	31%	29%	10%	19%	41%	
S _A	12	37	11	48	2	25	18	4	
S _W %	59%	53%	14%	29%	2%	11%	27%	5%	
Next hours									
	17	18	19	20	21	22	23	24	Day
Rn _A	4	6	5	8	8	2	3	3	9
Rn _W %	2%	5%	7%	30%	98%	16%	79%	56%	10%
G _A	6	7	6	4	2	4	4	8	1
G _W %	32%	72%	198%	78%	25%	40%	32%	65%	13%
LE _A	22	4	6	10	14	14	10	8	5
LE _W %	28%	7%	16%	57%	-	-	149%	108%	13%
S _A	24	18	4	2	5	11	18	20	3
S _W %	28%	27%	10%	14%	454%	-	602%	992%	6%

Rn_A, G_A, LE_A, S_A – absolute errors of estimated fluxes, Rn_W, G_W, LE_W, S_W - relative errors of estimated fluxes, where e.g.:

$$Rn_{Ai} = |Rn_i - Rn_i^{\wedge}| \quad (3)$$

$$Rn_{wi} = Rn_{Ai} / |Rn_i| \cdot 100\% \quad (4)$$

i – hour, Rn_i – measured value, Rn_i[^] – value estimated using the multiple regression equation, analogous formulas were used for the other fluxes.

Thus results from the 24 hours preceding the period of the gap in measurements and 24 h after that period were used as a set of data to establish regression. As a result, multiple regressions were determined based on a set of 48 hours of measurements. Regression was used to determine the missing values of net radiation (Rn), soil heat flux (G) and latent heat (LE), while sensible heat was determined using formula (5) for heat balance of an active surface:

$$S = -LE - G - R_n \quad (5)$$

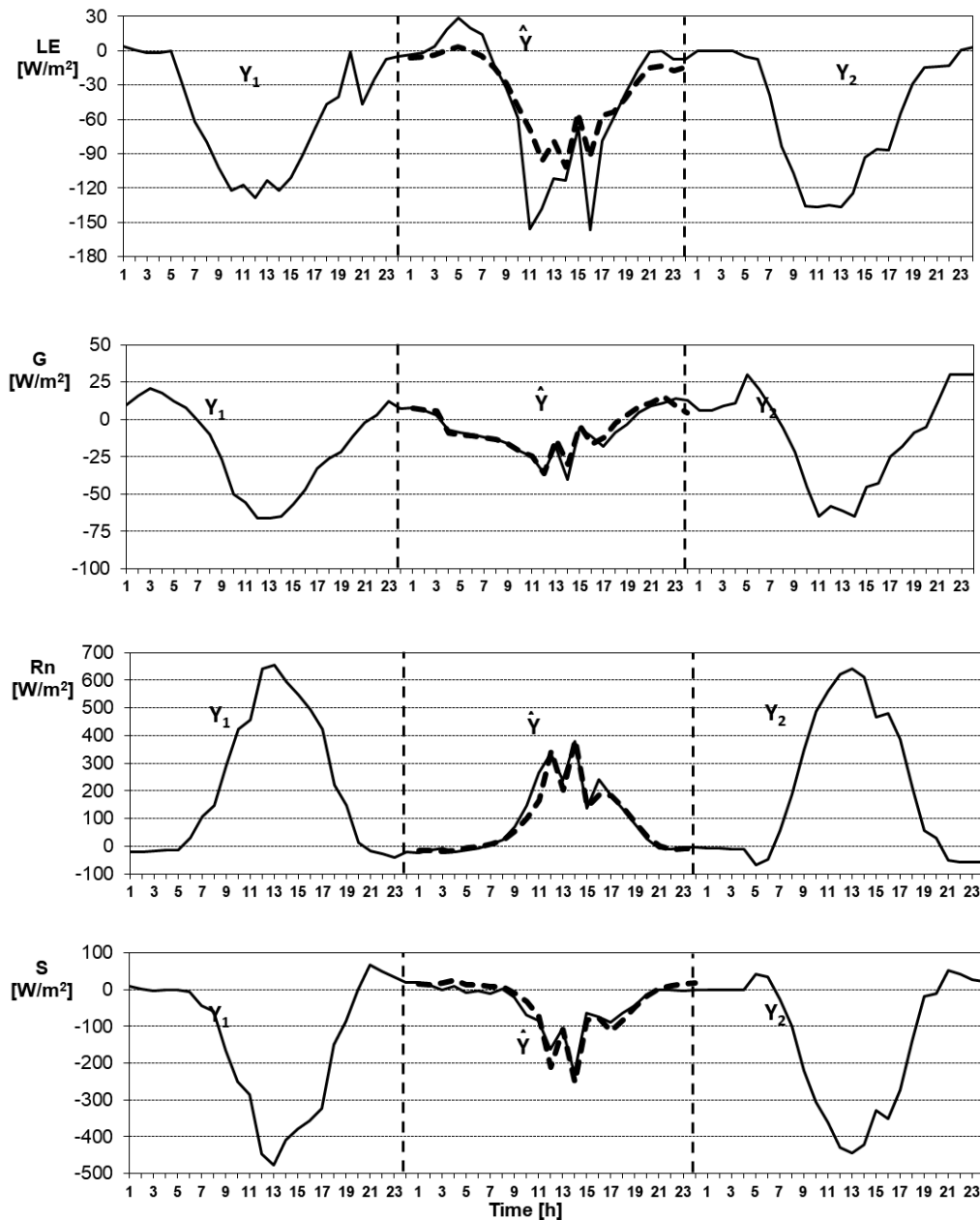


Fig. 2. The course of LE – latent heat flux, Rn – net radiation, G – soil heat flux, S – sensible heat flux quantities in measurements site A in the period from 3.07 to 5.07 2002 together with the course approximated by means of multiple regression (dotted line) on 4.07.02.

4 Discussion

Mean daily values of heat balance components determined on the basis of their estimated fluctuations were close to measured values (Table 2). In the presented examples the greatest relative differences were found between values of latent heat fluxes and the soil flux. Both estimated values diverged from the measured daily mean by 13%, which in absolute numbers gave -40 and -35 W/m² for the latent heat flux as well as -8 and -7 W/m² for the soil

flux. For the other heat balance components the situation was more promising – for net radiation the relative difference amounted to 10%, while for sensible heat it was 6%. Obviously for individual hours of measurements these differences were occasionally much greater. It needs to be stated here all the heat balance components for the active surface in the adopted notation (equation 5) in the 24-h course may take both positive and negative values. This means that at certain times of the day, i.e. around sunrise or sunset, and during rapid weather changes (the sun being obstructed by clouds, rainfall, etc.) they may assume values close to zero. This

in turn means that the value of the relative error of flux estimation (equation 4) may be very large, even if the actual values of the flux are slight. Gap filling using multiple regression reflects momentary fluctuations in filled in data. Figure 2 presents a reliable picture of large changes both in measured and estimated values of evapotranspiration from 11 a.m. to 4 p.m. In this case fluctuations in estimated values are smaller than those in measured values, but they are still detectable. This indicates that such gap filling also reflects situations when a process (e.g. evapotranspiration) is relatively unstable.

We also need to add here that a value close to zero for any heat balance component means that the flow of energy in this range is not observed. In this case for example neither evapotranspiration nor water vapour condensation takes place, with soil neither being heated nor releasing energy, etc. Thus these hours (with values close to zero) will be of limited importance in the total 24-h flow of energy. Such a situation took place in the discussed example of 4.07.2002.

As it was mentioned above, measurements of heat balance components at 5 active surface sites were taken from the end of April to the end of September and before the gap filling method was selected several tests had been performed. It turned out that over the entire measurement period the best similarity was found for values of net radiation, while the greatest divergence was observed for soil heat and sensible heat fluxes. This was the case for several periods, for which preliminary calculations were made to verify the gap filling method [7]. It may have been expected, since the volume of net radiation depends mainly on the amount of radiation reaching the surface and on the surface albedo, which in the described experiment did not differ considerably between crop types. In the example of gap filling given above it was assumed that 24 h of measurements are missing. In practice these gaps were much shorter and their filling was then more accurate than in this example. Nevertheless, it needs to be acknowledged that in the case of analyses of the daily fluctuations in values of heat balance components of an active surface we need to rely on actual measurement data and avoid situations when they had been filled in beforehand. In the case when mean daily values of flux are analysed, we may successfully analyse days with filled in data even when the entire 24-h period of measurements is being filled in. In the presented examples differences were as low as 13%, while in a situation when gap filling is performed for data concerning a shorter period the maximum relative error is even smaller.

5 Summary

The example discussed in this paper for series of measurement data concerning values of fluxes in active surface heat balance shows that the gap filling method well serves its function. In the presented example for the determined mean daily values based on filled in hourly values relative errors did not exceed 13% at a 24-h series of filled in missing data.

For individual hours, at a longer period of gap filling, these errors may be considerably greater and in the case of analyses of 24-h fluctuations in heat balance fluxes such data should not be filled in (unless the gap filling periods were very short, maximum several hours).

The described method was successfully applied to supplement measurement data for heat balance structure at five differences surfaces. This produced 5-month hourly measurement series of heat balance components for five different crops, which definitely facilitated the performance of several further analyses.

References

1. J. Szturc, A. Jurczyk, K. Ośródk, A. Wyszogrodzki, M. Giszterowicz, *Meteorol. Hydrol. Water Manage.* **6**, 1, 3–12 (2018)
2. L. Katsalova, V. Shpyg, *Meteorol. Hydrol. Water Manage.* **4**, 2, 23–28 (2016)
3. L. Kuchar, S. Iwański, L. Jelonek, W. Szalińska, *Geografie* **119**, 1, 1–25 (2014)
4. N. Dymarska, W. Rohm, J. Sierny, J. Kapłon, T. Kubik, M. Kryza, J. Jutarski, J. Gierczak, R. Kosierb, *Meteorol. Hydrol. Water Manage.* **5**, 1, 3–13 (2017)
5. L.A. Hunt, L. Kuchar, C.J. Swanton, *Agr. Forest. Meteorol.* **91**, 3–4, 293–300 (1998)
6. L. Kuchar, *Math. Comput. Simulat.* **65**, 1–2, 69–75 (2004)
7. J. Leśny, *Zesz. Nauk. UP Pozn.* **429** (2011)
8. J. Leśny, J. Olejnik, *Acta Agrophys.* **179**, 11–25 (2011)
9. F. Eulenstein, J. Leśny, B.H. Chojnicki, A. Kędziora, J. Olejnik, *Int. Agrophys.* **19**, 3 (2005)
10. P. Struzik, M. Kepinska-Kasprzak, *Meteorol. Hydrol. Water Manage.* **4**, 2, 3–13 (2016)
11. M. Panfil, R. Jassal, C. Arevalo, R. Ketler, Z. Nestic, N. Grant, A. Black, J. Bhatti, D. Sidders, J. Leśny, *Przegl. Geofiz.* **57**, 2, 245–254 (2012)
12. M. Acosta, R. Juszcak, B. Chojnicki, M. Pavelka, K. Havránková, J. Leśny, L. Krupková, M. Urbaniak, K. Macháčová, J. Olejnik, *Wetlands* **37**, 3, 423–435 (2017)
13. J. Leśny, J. Olejnik, B. H. Chojnicki, F. Eulenstein, *Acta Agrophysica* **57**, 69–75 (2001)
14. D. T. Mihailović, B. Lalić, J. Lesny, J. Olejnik, *Arch. Biol. Sci.* **64**, 4, 1533–1542 (2012)
15. J. Nyćkowiak, J. Leśny, R. Juszcak, E. Haas, R. Kiese, K. Butterbach-Bahl, J. Olejnik, A. Blecharczyk, *Ann. Set Environ. Prot.* **15**, 1937–1951 (2013)
16. J. Nyćkowiak, J. Leśny, L. Merbold, S. Niu, E. Haas, R. Kiese, K. Butterbach-Bahl, J. Olejnik, *Reg. Environ. Change.* **14**, 3, 1073–1082 (2014)
17. P. Struzik, M. Kempinska-Kasprzak, *Meteorol. Hydrol. Water Manage.* **4**, 2, 3–13 (2016)
18. J. Paszyński, *Geogr. Pol.* **22** (1972)

19. J. L. Monteith, *Vegetation and the atmosphere* (Acad. Press, London, 1975)
20. T.R. Oke, *Boundary layer climates* (Methuen & Co LTD, 372, New York, 1978)
21. A. Kędziora, *Podstawy agrometeorologii* (PWRiL, Poznań, 1999)
22. P. Baranowski, B. Usowicz, R. Walczak, W. Mazurek, *Evapotranspiration into the boundary layer of the atmosphere* (IA PAS, Lublin, 2005)
23. R.T. Walczak, P. Baranowski, W. Mazurek, *Plant Growth in Relation to Soil Physical Conditions* (Agro. Inst. PAS, Lublin, 2004)