A dataset of microclimate and radiation and energy fluxes from the Lake Taihu Eddy Flux Network

Zheng Zhang[a,b], Mi Zhang[a,b,d], Chang Cao[a,b], Wei Wang[a,b], Wei Xiao[a,b,d], Chengyu Xie[a,b], Haoran Chu[a,b], Jiao Wang[a,b], Jiayu Zhao[a,b], Lei Jia[a,b], Qiang Liu[a,b], Wenjing Huang[a,b], Wenqing Zhang[a,b], Yang Lu[a,b], Yanhong Xie[a,b], Yi Wang[a,b], Yini Pu[a,b], Yongbo Hu[a,b], Zheng Chen[a,b], Zhihao Qin[a,b], Xuhui Lee[c*]

a Yale-NUIST Center on Atmospheric Environment, International Joint Laboratory on Climate and Environment Change (ILCEC), Nanjing University of Information Science and Technology, Nanjing, Jiangsu Province, China

b Key Laboratory of Meteorological Disaster, Ministry of Education and Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing, Jiangsu Province, China

c School of Forestry and Environmental Studies, Yale University, New Haven, CT, USA

d NUIST-Wuxi Research Institute, Wuxi, Jiangsu Province, China

* Corresponding author: xuhui.lee@yale.edu

Abstract

Eddy covariance data are widely used for the investigation of surface-air interactions. Although numerous datasets exist in public depositories for land ecosystems, few research groups have released eddy covariance data collected over lakes. In this paper, we describe a dataset from the Lake Taihu Eddy Flux Network, a network consisting of seven lake sites and one land site. Lake Taihu is the third largest freshwater lake (area 2,400 km^2) in China, under the influence of subtropical climate. The dataset spans the period from June 2010 to December 2018. Data variables are saved half-hourly averages and include micrometeorology (air temperature, humidity, wind speed, wind direction, rainfall, and water/soil temperature profile), the four components of surface radiation balance, friction velocity, and sensible and latent heat fluxes. Except for rainfall and wind direction, all other variables are gap-filled, with each datapoint marked by a quality flag. Several areas of research can potentially benefit from the publication of this dataset, including evaluation of mesoscale weather forecast models, development of lake-air flux parameterizations, investigation of climatic controls on lake evaporation, validation of remote sensing surface data products, and global synthesis on lake-air interactions. The dataset is publicly available at https://yncenter.sites.yale.edu/data-access and from Harvard Dataverse https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/HEWCWM&version=DRAFT&faces-redirect=true (Zhang et al., 2020).

1. Introduction

Inland lakes and reservoirs are a vital freshwater resource for the society. Globally, there are more than 27 million water bodies with size greater than 0.01 km^2, occupying a total of 3.5% of the Earth's land surface area (Downing et al., 2006; Verpoorter et al., 2014). Accurate observation of the lake microclimate and lake-air interactions will help to better manage this water resource and to better predict how it may be affected by environmental changes. Towards that end, an increasing number of studies have employed the eddy covariance (EC) methodology to monitor physical state (temperature, wind, humidity) and process variables (momentum flux, and radiation and energy fluxes) in the lake environment (Vesala et al., 2006; Blanken et al., 2011; Nordbo et al., 2011; Wang et al., 2014; Li et al., 2015; Yusup and Liu, 2016; Du et al., 2018; Hamdani et al., 2018; Xiao et al., 2018; Wang et al., 2019). Unlike EC studies in land ecosystems, however, data from these lake studies are rarely published as data papers or are archived in public data depositories accessible by the broader scientific community. For example, of the nearly 500 sites that have contributed EC and micrometeorological data to AmeriFlux, a public data depository (https://ameriflux.lbl.gov/data/data-availability/), none is a lake site. Although a few scientific groups have provided data supplements to their scientific papers on lake-air fluxes (e. g., Charusombat et al., 2018; Franz et al., 2018; Zhao and Liu, 2018), we are not aware of a data paper devoted to systematic description and archival of EC lake observations.

In this paper, we describe the dataset from the Lake Taihu Eddy Flux Network (Lee et al., 2014). Established in 2010, the network currently consists of six active lake sites, one

64   inactive lake site, and one active land site. Lake Taihu is the third largest freshwater lake

65   (area 2,400 km^2) in China. Data variables are recorded at half-hourly intervals and the

66   measurement has continued for over eight years. Several areas of research can potentially

67   benefit from the publication of this dataset, including evaluation of mesoscale weather

68   forecast models, development of lake-air flux parameterizations, investigation of climatic

69   controls on lake evaporation, validation of remote sensing surface data products, and global

70   synthesis on lake-air interactions.

71

72   This paper is organized as follows. Section 2 is a brief overview of the sites and the

73   instruments used by the network. This is followed, in Section 3, with a description of data

74   quality measures employed during the field monitoring. Section 4 provides the essential

75   information about the dataset, including data variables, gap-filling methods, and data quality

76   flags. Results of post-field evaluation of the data quality are given in Section 5.

77

78   Users of this dataset may be interested in the relevant papers published by our group. Lee et

79   al. (2014) gave an overview of the Lake Taihu Eddy Flux Network. Using the data collected

80   at a subset of the sites and during the early phase of the network, Wang et al. (2014)

81   investigated the spatial variability of energy and momentum fluxes across the lake. Xiao et al.

82   (2013) improved the bulk parameterizations of heat, water and momentum fluxes for shallow

83   lakes. Deng et al. (2013) and Hu et al. (2017) modified the CLM lake simulator (Subin et al.,

84   2012) to improve its prediction of the lake evaporation. Wang et al. (2017) and Zhang et al.

85   (2019b) evaluated the performance of two mesoscale models of the lake-land breeze. More

86   recently, Xiao et al. (2020, manuscript in review) investigated drivers of the interannual

87   variability of the lake evaporation observed at one of the lake sites (BFG). The value of the

88   dataset is enhanced by these peer-reviewed publications because they have helped us to

89   continuously improve our measurement and data processing protocols. For example, we have

90   used the locally-calibrated bulk parameterizations of Xiao et al. (2013) to gap-fill the flux

91   variables.

92

93   **2. Sites and Instrumentation**

94   **2.1 Sites and data periods**

95   Table 1 shows the basic site information and Figure 1 is a map that gives the relative position

96   of Lake Taihu in China and locations of the EC measurement sites. Also shown in Figure 1

97   are WMO baseline weather stations around the lake, whose data can be obtained from

98   National Meteorological Information Center in China (http://data.cma.cn/site/index.html).

99   The lake, located between the latitudinal range of 30˚5′40″ N to 31˚32′58″ N and

100   longitudinal range of 119˚52′32″ E to 120˚36′10″ E, has a total area of 2400 km^2 and an

101   average depth of 1.9 m. The climate is subtropical monsoon, with an annual mean

102   temperature of 16.2°C and annual total precipitation of 1122 mm. The lake is ice-free

103   throughout the year.

104

105   The EC network consists of seven lake sites and one land site. The lake sites (Meiliangwan

106   (MLW), Dapukou (DPK), Bifenggang (BFG), Xiaoleishan (XLS), Pingtaishan (PTS),

107   Dongtaihu (DTH), Meiliangwan2 (MLW2)) are distributed according to biological

108   characteristics and across eutrophication gradients of the lake. The MLW site, located in

109   Meiliangwan Bay near the north shore of Lake Taihu, was the first site in operation; the

110   measurement began in June 2010 and was replaced by MLW2 in 2018, at 10 km southwest of

111   MLW. Both MLW and MLW2 sites are located in the lake eutrophic zone. BFG is located in

112   the east part of Lake Taihu in relatively clean water inhabited by submerged vegetation with

113   a growth season from April to November. DTH is located in the shallow water (mean depth

114   of 1.3 m) in the southeast part of the lake. After more than 20 years of crab aquaculture, this

115   zone was returned to unmanaged state in December 2018 in order to improve water quality.

116   The observation at DTH enables the examination of lake-air exchange processes in the

117   transition from human management to a natural state. PTS is situated in the middle of Lake

118   Taihu where occasional algal blooms occur and no aquatic vegetation is present. DPK is

119   located near the west shore, in a relatively deep (depth 2.5 m) super eutrophic zone due to

120   heavy influence of agricultural and urban runoffs. XLS is located in the relatively clean and

121   vegetation-free zone in the southeast. Finally, DS is a land site surrounded by rice agriculture,

122   serving as a land reference for the lake sites. The MLW site is situated at a distance of 200 m

123   from the north shore of the lake. All the other lake sites in the lake are at a distance of more

124   than 1 km away from the land.

125

126   The lake water level is monitored daily by the Taihu Basin Authority at five locations around

127   the lake (http://www.tba.gov.cn/). Using the water-level time series, we have constructed the

128   water depth for our eddy covariance sites (Figure 2).

129

2.2 Instrumentation

Each site is equipped with an EC system for long-term, continuous monitoring of the surface momentum, sensible heat, latent heat and carbon dioxide fluxes. The EC system consists of a sonic anemometer/ thermometer (Model CSAT3A; Campbell Scientific, Logan, UT, USA) and a CO_2/H_2O infrared gas analyzer (Model 7500A, LI-COR, Inc., Lincoln, NE, USA at DS, MLW, MLW2 and DPK; Model EC150, Campbell Scientific at other sites). The EC system is at a height of 3.5 to 9.4 m above the water surface at the lake sites and at a height of 20 m above the ground at the land site.. Other measurements include air humidity and air temperature (Model HMP45D/HMP155A; Vaisala, Inc, Helsinki, Finland), wind speed and wind direction (Model 03002; R. M. Young Company, Traverse City, MI, USA) and four components of the net radiation (Model CNR4; Kipp & Zonen B. V., Delft, the Netherlands). At the lake sites, water temperature profile was measured with temperature probes (Model 109-L; Campbell Scientific) at the water depth of 20, 50, 100, and 150 cm and in the sediment at about 5 cm below bottom of the water column. The top four temperature sensors were tied to a nylon rope hanging from a buoy to ensure that they were at the designed depths regardless of water level fluctuations. At the DS land site, soil temperature profile was measured with the same type of probes at the depths of 5, 10, 20 cm. The MLW and the DS sites are supported by A/C power and other sites are powered by battery packs connected to solar panels. Measurements at the lake sites were made on fixed platforms. Readers are referred to Lee et al. (2014) and Xiao et al. (2017) for photographs of the platform and the instruments.

152　All the variables are reported as 30-min averages. The EC data are expressed in the natural

153　coordinate system (Lee et al., 2004). In this coordinate system, the longitudinal coordinate

154　axis is aligned with the 30-min mean velocity vector so that the 30-min mean lateral and

155　vertical velocity components are zero and the magnitude of the mean velocity is equal to the

156　mean longitudinal component, and the covariance between the lateral and the vertical

157　velocity components is zero. Additionally, a small density correction has been applied to the

158　water vapor flux according to Webb et al. (1980).

159

3 Data Quality Control during Field Monitoring

161　Every site in the Lake Taihu Eddy Flux Network is equipped with a wireless transmission

162　module for real-time monitoring and for data transmission. Time series of all 30-min

163　variables are examined weekly and abnormal behaviors are flagged for site operators. Each

164　site is visited every one to two months to perform instrument repair and maintenance and to

165　download 10 Hz EC data. The data coverage rates are summarized in Table 2, where the

166　percentage values represent the proportions of data with quality flag 0, which indicates

167　high-quality original measurement (Table 3).

168

169　The four-way net radiometers at MLW and XLS were compared in the field against a

170　laboratory standard of the same type in the summer of 2018 to check their long-term stability

171　(Figure 3). These two sites were chosen because they have been in operation for more than

172　five years. Additionally, the radiometer at MLW was relocated to MLW2 after MLW had

173　been discontinued. The laboratory standard, which had been calibrated at the manufacturer

174   prior to this performance evaluation, was mounted next to the field instrument for about 10

175   days at each site, covering overcast to clear-sky conditions. The mean bias error was smaller

176   than 1 W m^{-2} for all the radiation components. It was -0.81, -0.81, 0.79 and -0.44 W m^{-2} for

177   the downward shortwave, upward shortwave, downward longwave and upward longwave

178   radiation flux at MLW, respectively. The corresponding values were 0.91, 0.40, 0.69 and

179   0.77 W m^{-2} for XLS. (Comparison experiments are being planned for the other sites.)

180

181   The EC gas analyzers were calibrated every one to two years. The zero-point calibration was

182   carried out with high-purity nitrogen gas, the CO_2 span calibration was made with standard

183   carbon dioxide gases (in the concentration range of 389 to 525 ppm) provided by the National

184   Institute of Meteorology (NIM), China and certified to an accuracy of 1%, and the H_2O span

185   calibration was made with a portable dew-point generator (LI-610; LI-COR, Inc.).

186

187   **4. Gap-filling Methods and Data Quality Flags**

188   We use five-point moving average to screen outliers. If the deviation from the moving

189   average is greater than two standard deviations, the data point is discarded. If a gap length is

190   30 min to 1 h, the gap is filled by linear interpolation. Larger gaps in meteorological variables,

191   radiation components and water temperature are filled with linear regression involving

192   observation of the same variable at another site. This spatial interpolation consists of three

193   steps. First, linear correlation is calculated using the valid data at the target site and at all

194   other sites for the month during which the data gap occurred. Second, the observation at the

195   site with the highest linear correlation is used to establish a linear regression equation. Third,

196   the gap at the target site is filled with the linear regression and the observation at that site.

197

198   Radiation data gaps at the DS land site require special treatment. The radiometer at DS eddy

199   flux site ended in January 2013. Subsequent measurements of the radiation component are

200   provided by a radiometer belonging to the Dongshan WMO weather station at a distance of

201   50 m from the eddy covariance tower (Figure 1). While large gaps in meteorological

202   variables (air temperature, relative humidity, wind speed and air pressure), downward solar

203   radiation and downward longwave radiation are filled with the spatial interpolation method,

204   large gaps in upward shortwave radiation and upward longwave radiation cannot be filled

205   with data from other lake sites even with linear regression. In the case of the upward

206   shortwave radiation, the data gaps were filled using the relationship between downward

207   shortwave radiation and the monthly mean albedo. In the case of upward longwave radiation,

208   the data gaps were filled by a regression equation between the upward longwave radiation

209   and the fourth power of soil temperature at 5-cm depth. Compared to the original data, the

210   gap-filled data do not capture the full diurnal variations because the 5-cm soil temperature

211   has smaller diurnal amplitudes than the soil surface temperature, but the daily-mean upward

212   longwave radiation flux seems reasonable.

213

214   Large data gaps in the EC variables (sensible heat flux, latent heat flux and friction velocity)

215   are filled with a hybrid method. First, if observations exist for the relevant state variable, the

216   gap is filled with the bulk transfer relationship using a transfer coefficient tuned locally for

217   each site (Xiao et al., 2013). For example, the relationship for filling gaps in the sensible heat

218   flux is

219   $$H = \rho_a c_p C_H U (T_s - T_a)$$

220   where ρ_a is air density, c_p is specific heat of air at constant pressure, C_H is the transfer

221   coefficient for sensible heat, T_a is air temperature and T_s is water surface temperature. The

222   transfer coefficient C_H is determined from the observed H and the state variables (U, T_a and

223   T_s) outside gap periods. The missing data on H is then filled with the above relationship using

224   the tuned C_H the observed U, T_a and T_s. Second, if data for the state variable is missing, the

225   spatial interpolation method is used to fill the gaps in these EC variables.

226

227   The spatial interpolation method described above occasionally causes a sudden jump at the

228   beginning or end of a data gap. To harmonize the data, we apply a 5-point moving averaging

229   to the gap-filled time series. If a data point deviates by 2 times of the standard deviation from

230   the moving average, it is replaced by linear interpolation using the two adjacent data points.

231

232   Each data variable is assigned a quality flag to distinguish original measurements and

233   gap-filled values and gap-filling methods (Table 3). The data flags employed here should not

234   be confused with quality flags commonly assigned to the EC methodology in the literature.

235   Specifically, Flag 0 indicates high-quality original data. Other flag values indicate gap-filled

236   data or missing values. Flag 1 indicates that the data was filled by temporal interpolation.

237   Flag 2 indicates that the data was filled by the spatial interpolation method. Flag 3 for the EC

238   variables indicates that the data was filled by the bulk relationship. We also use Flag 3 to

239   mark the upward shortwave and longwave radiation data filled with the albedo and the

surface temperature relationship, respectively, for the DS land site. Missing values occur on

some situations, which are marked with Flag 4. Figure 4 is an example showing the gap-filled

time series of several variables at BFG along with the flag status.

Rainfall data has not been quantity-controlled or gap-filled. Because of the episodic nature of

rainstorms and high spatial variability of rainfall, it is not appropriate to fill data gaps with

the time or spatial interpolation method. The total rain amount is likely biased low because

no wind screens are used to protect the rain gages from the influence of wind which is much

higher on the lake than on land (Figure 5 below). On several site visits, the drain opening to

the tipping bucket was found to be partially blocked by debris. Rain amount at a constant and

low rate and excessively long rain duration are evidence of such blockage. The flag status of

0 for the rainfall variable simply indicates that the field measurement is available, but it does

not guarantee high data quality.

The data coverage begins from the start time of each site (Table 1) and ends in December

2018. The time resolution is 30 min. The dataset includes microclimate variables (air pressure,

air temperature, relative humidity, wind speed, wind direction and rainfall), radiation fluxes

(upward and downward shortwave radiation, upward and downward longwave radiation),

water temperature at depth of 0.2 m, 0.5 m, 1.0 m and 1.5 m, and in the 5-cm sediment) and

eddy fluxes (friction velocity, sensible heat and latent heat fluxes; Table 4). The time stamp is

Beijing time (UTC + 8 h) given by data columns 1 to 5 as year, month, day, hour, and minute,

and marks the end of the observation period. For example, time stamp "2012, 1, 1, 12, 00"

262   indicates that the data acquisition period is from 11:30 to 12:00 on January 1, 2012.

263

264   Although the data table does not include the radiative surface temperature T_s, the user can

265   easily calculate it from the two longwave radiation fluxes, as

266
$$T_s = (\frac{L_\uparrow - (1-\varepsilon)L_\downarrow}{\varepsilon\sigma})^{\frac{1}{4}}$$

267   where σ is the Stefan-Boltzmann constant, ε is emissivity, and $L_\uparrow$ and $L_\downarrow$ are upward and

268   downward longwave radiation flux, respectively. We use a value of 0.97 for lake surface

269   emissivity in this calculation (Deng et al., 2013; Wang et al., 2014).

270

271   **5. Data Consistency Evaluation**

272   Figure 5 compares the annual mean air temperature, relative humidity, and wind speed at the

273   Taihu eddy flux sites with those at the four WMO weather stations (Wuxi, Liyang, Huzhou

274   and Dongshan) around the lake (Figure 1). The error bars represent the maximum and

275   minimum values among the four WMO stations and the lines represent the mean values of

276   the four station measurements. The annual mean air temperature at DTH is 0.3°C higher than

277   the station mean. At other sites, air temperature is in close agreement of the weather station

278   data, in terms of both magnitude and inter-annual variability. The annual mean wind speed at

279   MLW, a site near the shoreline, is comparable with the station data. At other more exposed

280   sites, the wind speed is much higher than observed at the WMO stations. The annual mean

281   relative humidity RH shows a larger spread among the eddy flux sites than among the WMO

282   stations partly because the measurement height at the eddy flux sites is not standardized

283   (Table 1). The upward trends in RH over time at DPK and XLS seem to be related more to

284   aging of the sensor than to a real inter-annual variability. We have not fully investigated this

285   aging problem, but it is possible to rectify it by doing a detailed regression analysis against

286   the station data.

287

288   Consistency of the energy flux variables can be evaluated with the energy balance closure.

289   Using observations made at a subset of the sites in the earlier years of the flux network,

290   Wang et al. (2014) reported a closure rate of 70 % to 110 % on the monthly basis, meaning

291   that the sum of the measured monthly sensible and the latent heat flux $H + \lambda E$ is 70 % to

292   110 % of the monthly available energy $R_n -$ G, where R_n is net radiation and G is heat storage

293   in the water column. By selecting days without data gaps, we found that the daily energy

294   balance closure is in the range between 66 % and 78 % for all the lake sites and all the years.

295   Such closure rates are typical of eddy covariance observations (Tanny et al., 2008; Wilson et

296   al., 2002).

297

298   We have shown that the monthly latent heat flux at the lake sites MLW, BFG and DPK

299   during July 2010 to August 2012 follows the Priestley-Taylor (PT) model prediction with the

300   original PT constant α of 1.26 and that at the DS land site is in agreement with the PT model

301   if the constant is lowered to 1.0 (Lee et al., 2014). Figure 6 demonstrates that the same

302   relationships hold for all the sites and all the observational months, indicating the overall

303   stability of our measurement systems and the robustness of our gap-filling procedure. The

304   reader is reminded that the monthly latent heat flux in Figure 6 has been adjusted to force

305 energy closure following the method recommended by Barr et al. (1994), Blanken et al.

306 (1997) and Twine et al. (2000). (The half-hourly flux data in the data archive have not been

307 adjusted for energy balance.)

308

309 The Stefan-Boltzmann Law offers another way for checking data consistency. Because the

310 lake surface emits longwave radiation like a blackbody and because the annual mean air

311 temperature and the surface water temperature are nearly identical at this lake (Wang et al.,

312 2014), the change in the annual upward longwave radiation $\Delta L_\uparrow$ can be expressed as

313 $$\Delta L_\uparrow = 4\sigma T_a^{\,3} \Delta T_a$$

314 where T_a is annual mean air temperature, and Δ is the difference between the target year and

315 the year with the lowest air temperature observed at the site. All the five long-term lake sites

316 show good consistency between the longwave radiation and the air temperature observations

317 (Figure 7).

318

319 **6 Data availability**

320 All data can be open-accessed online for download and use at https://yncenter.sites.yale.edu/

321 and from Harvard Dataverse (https://doi.org/10.7910/DVN/HEWCWM, Zhang et al., 2020).

322

323 **7 Summary**

324 The dataset described here consists of microclimate variables (air temperature, air humidity,

325 wind speed, wind direction, water or soil temperature profile, and rainfall), four components

326 of the radiation balance, friction velocity, and sensible and latent heat fluxes observed at

seven lake sites and one land site. The period of coverage is from June 2010 to December

2018. The observation interval is 30 min. Except for rainfall and wind direction, all other

variables have been gap-filled. Every data point is tagged with a data quality flag to help the

user determine how to best use the data.

Author contribution

XL, WX and MZ directed the field program, ZZ performed data gap-filling and prepared the
data for public release, CC, WW, CX, HC, JW, JZ, LJ, QL, WH, WZ, YL, YX, YW, YP, YH,
ZC and ZQ participated in field data collection, and ZZ, XL and MZ wrote the manuscript.

Competing interests

The authors declare no conflict of interest.

Acknowledgments

348

Table 1. A list of sites in the Lake Taihu Eddy Flux Network

Site ID	MLW	DPK	BFG	XLS	PTS	MLW2	DTH	DS
Lat (°N)	31.4197	31.2661	31.1685	30.9972	31.2323	31.3818	31.0611	31.0799
Long (°E)	120.2139	119.9312	120.3972	120.1344	120.1086	120.1608	120.4704	120.4346
Start date	Jun 2010	Aug 2011	Dec 2011	Nov 2012	Jun 2013	Feb 2018	Nov 2017	Apr 2011
Biology	Eutrophic	Super eutrophic	Submerged macrophyte	Transitional	Mesotrophic	Eutrophic	Aquaculture	Cropland/ Rural residence
Met height (m)	3.5	8.0	8.5	9.4	8.5	6.0	4.5	10.0
T_w / T_s depths (cm)	20, 50, 100, 150, sediment	20, 50, 100, 150, sediment	20, 50, 100, 150, sediment	20, 50, 100, 150, sediment	20, 50, 100, 150, sediment	20, 50, 100, 150, sediment	20, 50, sediment	5, 10, 20
Radiation height (m)	1.5	1.5	1.5	1.5	1.5	1.5	1.5	3.0
EC height (m)	3.5	8.5	8.5	9.4	8.5	6.0	4.5	20

350

351

352

Table 2. Percent of data coverage. The percentage represents the proportion of high-quality

original measurement.

Variable type	MLW	DPK	BFG	XLS	PTS	DTH	MLW2	DS
Micrometeorology	93.3	81.1	97.6	97.0	97.5	98.1	90.3	91.7
Radiation flux	85.5	90.8	96.9	97.4	98.6	98.2	98.2	82.7
Water/soil temperature	83.4	81.3	94.0	91.1	90.3	87.7	22.4	98.4
Eddy flux	73.3	61.8	82.7	79.1	80.6	85.7	85.5	82.8

355

356

357

358

359

360 **Table 3.** A list of data quality flags

Flag	Data quality description
0	Original data
1	Gap-filled with time interpolation
2	Gap-filled with spatial interpolation
3	Gap-filled with bulk relationship
4	NAN

361

362

363

364 **Table 4.** A list of data columns and variable definitions

Column	Description	Variable name	Unit
1	Year	Year	−
2	Month	Month	−
3	Day	Day	−
4	Hour	HH	−
5	Minute	MM	−
6	Day of Year	DOY	−
7	Air pressure	P	kPa
8	Quality flag of air pressure	P_flag	
9	Air temperature	Ta	°C
10	Quality flag of air temperature	Ta_flag	
11	Relative humidity	RH	%
12	Quality flag of Relative humidity	RH_flag	
13	Wind speed	WS	m s^{-1}
14	Quality flag of wind speed	WS_flag	
15	Wind direction	WD	Degree
16	Quality flag of wind direction	WD_flag	
17	Rainfall	R	mm
18	Quality flag of rainfall	R_flag	
19	Upward shortwave radiation	UR	W m^{-2}
20	Quality flag of upward shortwave radiation	UR_flag	
21	Downward shortwave radiation	DR	W m^{-2}
22	Quality flag of downward shortwave radiation	DR_flag	
23	Upward longwave radiation	ULR	W m^{-2}
24	Quality flag of upward longwave radiation	ULR_flag	
25	Downward longwave radiation	DLR	W m^{-2}
26	Quality flag of downward longwave radiation	DLR_flag	
27	Water temperature at 0.2 m	T_w_20	°C

28	Quality flag of Water temperature at 0.2 m	$T_w\_20\_flag$	
29	Water temperature at 0.5 m	T_w_50	°C
30	Quality flag of Water temperature at 0.5 m	$T_w\_50\_flag$	
31	Water temperature at 1.0 m	T_w_100	°C
32	Quality flag of Water temperature at 1.0 m	$T_w\_100\_flag$	
33	Water temperature at 1.5 m	T_w_150	°C
34	Quality flag of water temperature at 1.5 m	$T_w\_150\_flag$	
35	Sediment temperature	T_w_bot	°C
36	Quality flag of sediment temperature	$T_w\_bot\_flag$	
37	Friction velocity	U^*	$m\ s^{-1}$
38	Quality flag of friction velocity	U^*_flag	
39	Sensible heat flux	H	$W\ m^{-2}$
40	Quality flag of sensible heat flux	H_flag	
41	Latent heat flux	LE	$W\ m^{-2}$
42	Quality flag of latent heat flux	LE_flag	

Notes: 1) Time marks end of a half-hourly observation in Beijing time (UTC+8:00); 2) At the DS site, columns 27, 29, and 31 represent soil temperature at 5, 10, and 20 cm, respectively, column 33 represents soil heat flux G ($W\ m^{-2}$) measured at 5-cm depth, and column 34 represents quality flag of soil heat flux.
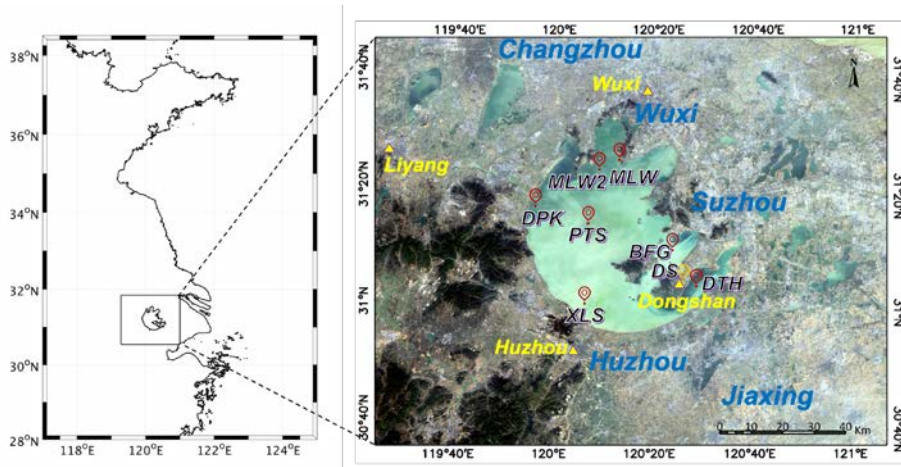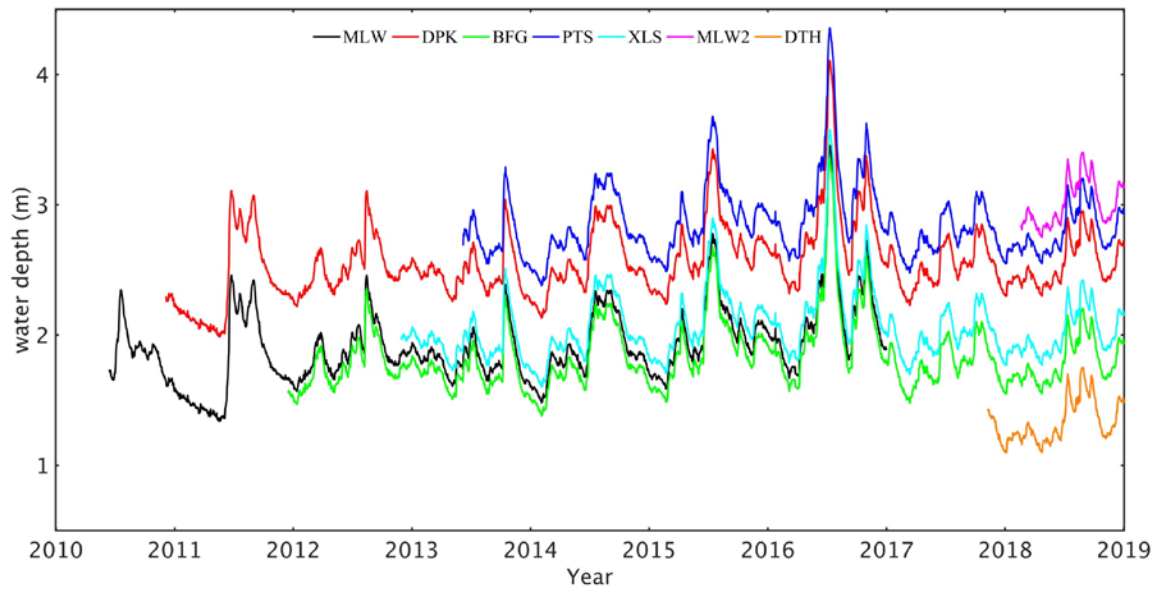
.

372

373

374

Figure 1. Map showing locations of Lake Taihu, eddy covariance sites (red bubbles) and
WMO weather stations (yellow triangles). City names are shown in blue. DS is a land site,
and MLW, MLW2, DPK, PTS, XLS, BFG and DTH are lake sites. The background is a
natural color image from LANDSAT 8 without correction for atmospheric interference.

379

380

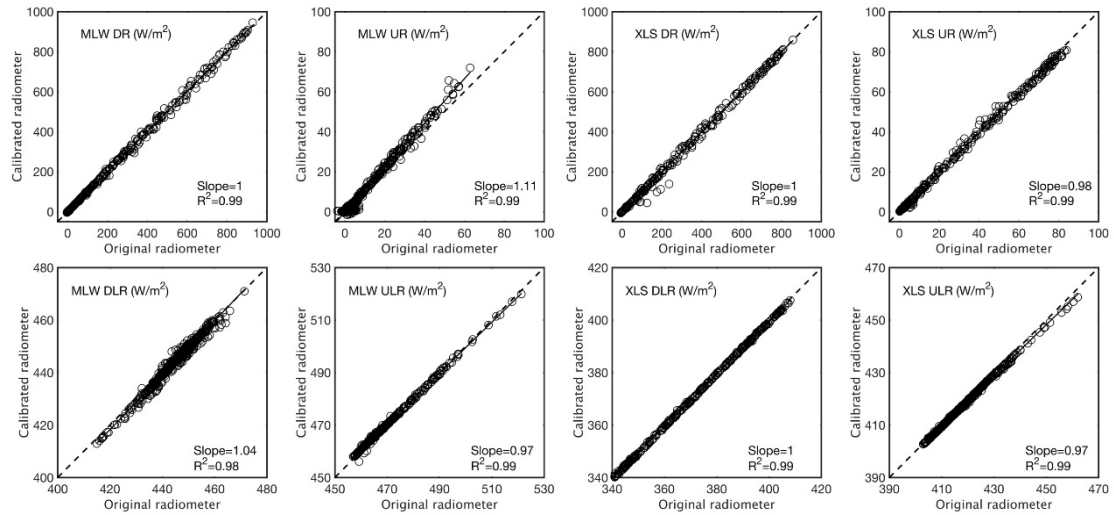381

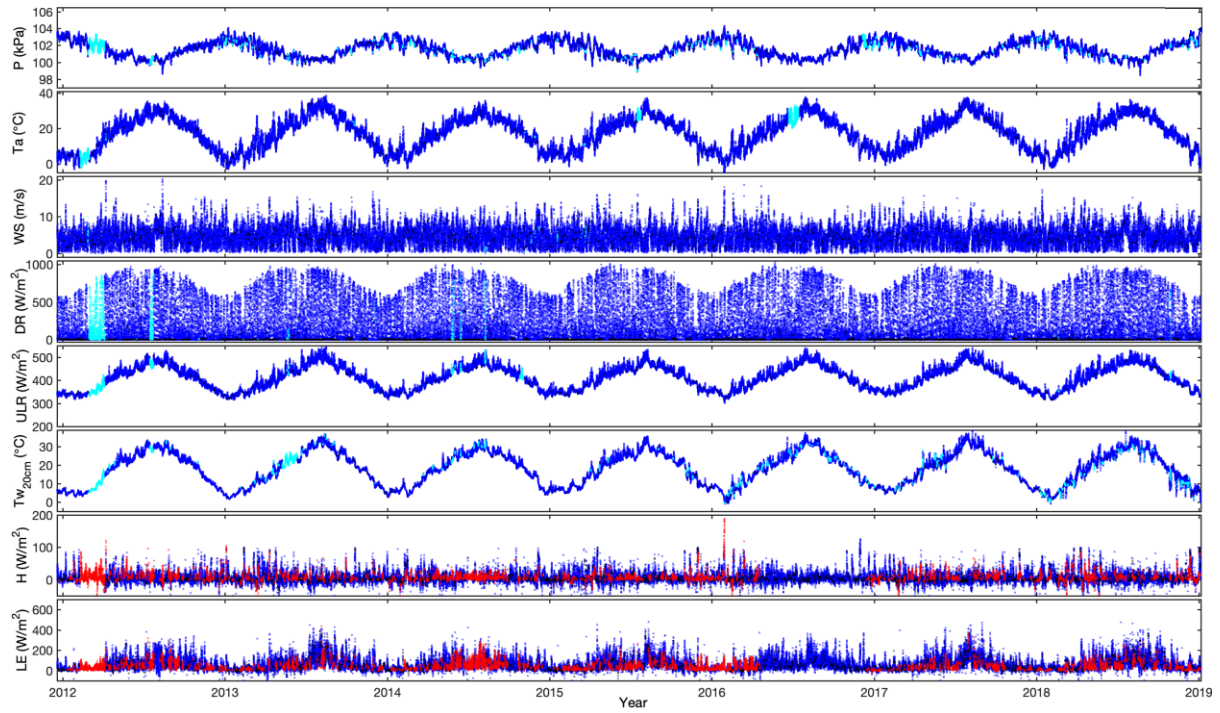**Figure 2.** Water depth at the eddy covariance sites.

383

384

385

386

387

Figure 3. Comparison of four components of the radiation balance between the original radiometer (horizontal axis) and a laboratory standard (vertical axis) at MLW and XLS. Refer to Table 4 for variable definitions.

388

392

393

394

395

396

Figure 4. Complete gap-filled time series for selected variables observed at BFG. Blue, black, cyan and red dots represent quality flag 0, 1, 2, and 3, respectively. Variable definitions are given in Table 4
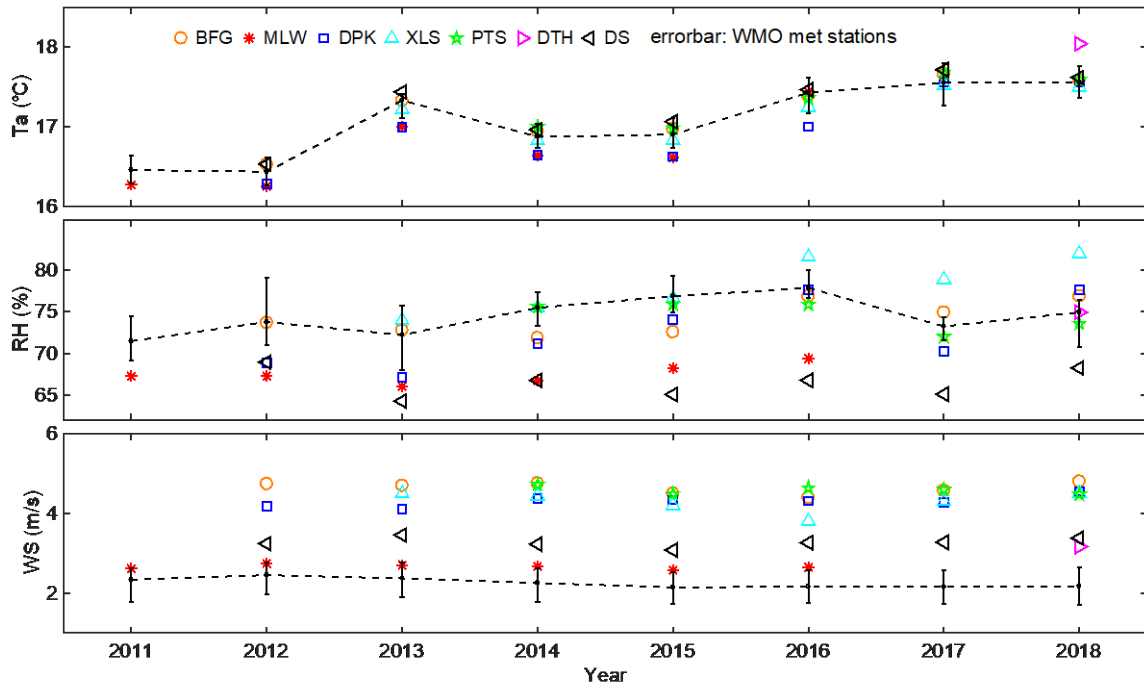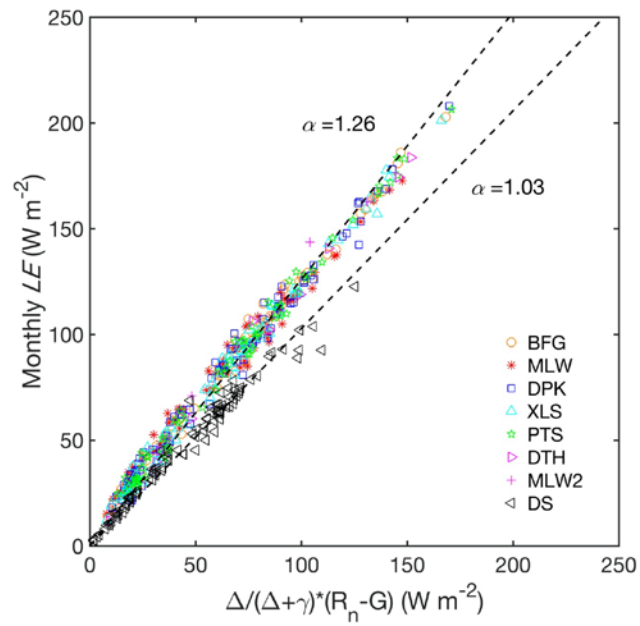
397

398

399

400

401

402

Figure 5. Annual mean air temperature (top), relative humidity (middle) and wind speed (bottom) observed at the eddy flux sites (symbols) and at the four WMO weather stations around the lake (line). Error bars represent the range of the annual means of the four WMO stations.
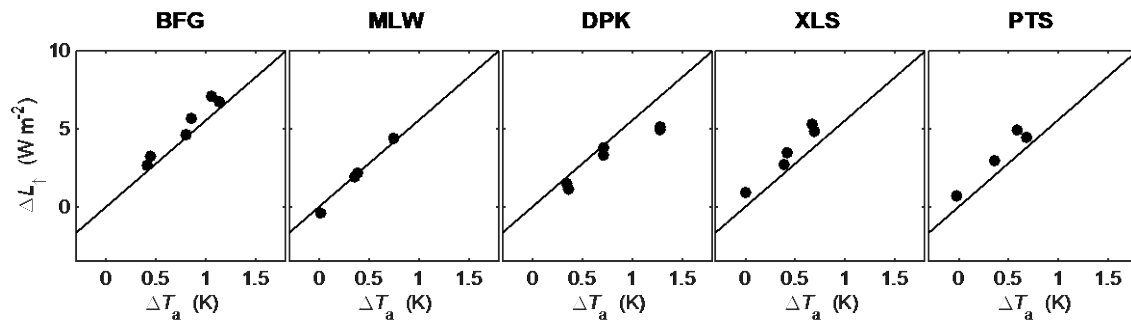
411

412

413

414 **Figure 6.** Comparison of observed monthly latent heat flux with Priestley-Taylor model

415 prediction using the origional α coefficeint of 1.26 and a modified coefficient of 1.03. Here

416 R_n is net radiation, G is heat storage in the water column, Δ is the slope of the saturation

417 vapor pressure curve, and γ is the psychrometric constant.

418

Figure 7. The relatinship between changes in observed annual mean upward longwave radiation flux and annual mean air temperature (dots). Solid lines represent the prediction of the Stefan-Boltzmann Law.

424

425 **References**

426 Barr, A. G., King, K. M., Gillespie, T. J., Den Hartog, G. and Neumann, H. H.: A comparison of Bowen ratio

427 and eddy correlation sensible and latent heat flux measurements above deciduous forest, Boundary-Layer

428 Meteorol., 71(1–2), 21–41, 1994.

429 Blanken, P. D., Black, T. A., Yang, P. C., Neumann, H. H., Nesic, Z., Staebler, R., Den Hartog, G., Novak, M.

430 D. and Lee, X.: Energy balance and canopy conductance of a boreal aspen forest: partitioning overstory and

431 understory components, J. Geophys. Res. Atmos., 102(D24), 28915–28927, 1997.

432 Blanken, P. D., Spence, C., Hedstrom, N. and Lenters, J. D.: Evaporation from Lake Superior: 1. Physical

433 controls and processes, J. Great Lakes Res., 37(4), 707–716, doi:10.1016/j.jglr.2011.08.009, 2011.

434 Charusombat, U., Fujisaki-Manome, A., Gronewold, A. D., Lofgren, B. M., Anderson, E. J., Blanken, P.,

435 Spence, C., Lenters, J. D., Xiao, C. and Fitzpatrick, L. E.: Evaluating and improving modeled turbulent heat

436 fluxes across the North American Great Lakes, Hydrol. Earth Syst. Sci., 22(10), 2018.

437 Deng, B., Liu, S., Xiao, W., Wang, W., Jin, J. and Lee, X.: Evaluation of the CLM4 lake model at a large and

438 shallow freshwater lake, J. Hydrometeorol., 14(2), 636–649, doi:10.1175/JHM-D-12-067.1, 2013.

439 Downing, J. A., Prairie, Y. T., Cole, J. J., Duarte, C. M., Tranvik, L. J., Striegl, R. G., McDowell, W. H.,

440 Kortelainen, P., Caraco, N. F., Melack, J. M. and Middelburg, J. J.: The global abundance and size distribution

441 of lakes, ponds, and impoundments, Limnol. Oceanogr., 51(5), 2388–2397, doi:10.4319/lo.2006.51.5.2388,

442 2006.

443 Du, Q., Liu, H., Xu, L., Liu, Y. and Wang, L.: The monsoon effect on energy and carbon exchange processes

444 over a highland lake in the southwest of China, Atmos. Chem. Phys., 18(20), 15087–15104, 2018.

445 Franz, D., Mammarella, I., Boike, J., Kirillin, G., Vesala, T., Bornemann, N., Larmanou, E., Langer, M. and

446 Sachs, T.: Lake-atmosphere heat flux dynamics of a thermokarst lake in arctic Siberia, J. Geophys. Res. Atmos.,

447 123(10), 5222–5239, 2018.

448 Franz, D., Mammarella, I., Boike, J., Kirillin, G., Vesala, T., Bornemann, N., Larmanou, E., Langer, M. and

449 Sachs, T.: Lake-atmosphere heat flux dynamics of a thermokarst lake in arctic Siberia, J. Geophys. Res. Atmos.,

450 123(10), 5222–5239, 2018.

451 Hamdani, I., Assouline, S., Tanny, J., Lensky, I. M., Gertman, I., Mor, Z. and Lensky, N. G.: Seasonal and

452 diurnal evaporation from a deep hypersaline lake: the Dead Sea as a case study, J. Hydrol., 562, 155–167, 2018.

453 Hu, C., Wang, Y., Wang, W., Liu, S., Piao, M., Xiao, W. and Lee, X.: Trends in evaporation of a large

454 subtropical lake, Theor. Appl. Climatol., 129(1–2), 159–170, doi:10.1007/s00704-016-1768-z, 2017.

455 Lee, X., Massman, W. and Law, B.: Handbook of micrometeorology: a guide for surface flux measurement and

456 analysis, Springer Science & Business Media., 2004.

457   Lee, X., Liu, S., Xiao, W., Wang, W., Gao, Z., Cao, C., Hu, C., Hu, Z., Shen, S., Wang, Y., Wen, X., Xiao, Q.,

458   Xu, J., Yang, J. and Zhang, M.: The taihu eddy flux network: An observational program on energy, water, and

459   greenhouse gas fluxes of a large freshwater lake, Bull. Am. Meteorol. Soc., 95(10), 1583–1594,

460   doi:10.1175/BAMS-D-13-00136.1, 2014.

461   Li, Z., Lyu, S., Ao, Y., Wen, L., Zhao, L. and Wang, S.: Long-term energy flux and radiation balance

462   observations over Lake Ngoring, Tibetan Plateau, Atmos. Res., 155, 13–25, 2015.

463   Nordbo, A., Launiainen, S., Mammarella, I., Leppäranta, M., Huotari, J., Ojala, A. and Vesala, T.: Long-term

464   energy flux measurements and energy balance over a small boreal lake using eddy covariance technique, J.

465   Geophys. Res. Atmos., 116(D2), 2011.

466   Subin, Z. M., Riley, W. J. and Mironov, D.: An improved lake model for climate simulations: Model structure,

467   evaluation, and sensitivity analyses in CESM1, J. Adv. Model. Earth Syst., 4(2), 1–27,

468   doi:10.1029/2011MS000072, 2012.

469   Tanny, J., Cohen, S., Assouline, S., Lange, F., Grava, A., Berger, D., Teltch, B. and Parlange, M. B.:

470   Evaporation from a small water reservoir: Direct measurements and estimates, J. Hydrol., 351(1–2), 218–229,

471   2008.

472   Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, Pr., Meyers, T. P., Prueger, J. H., Starks, P. J.

473   and Wesely, M. L.: Correcting eddy-covariance flux underestimates over a grassland, Agric. For. Meteorol.,

474   103(3), 279–300, 2000.

475   Verpoorter, C., Kutser, T., Seekell, D. A. and Tranvik, L. J.: A global inventory of lakes based on

476   high-resolution satellite imagery, Geophys. Res. Lett., 41(18), 6396–6402, doi:10.1002/2014GL060641, 2014.

477   Vesala, T., Huotari, J., Rannik, Ü., Suni, T., Smolander, S., Sogachev, A., Launiainen, S. and Ojala, A.: Eddy

478   covariance measurements of carbon exchange and latent and sensible heat fluxes over a boreal lake for a full

479   open-water period, J. Geophys. Res. Atmos., 111(11), 1–12, doi:10.1029/2005JD006365, 2006.

480   Wang, B., Ma, Y., Wang, Y., Su, Z. and Ma, W.: Significant differences exist in lake-atmosphere interactions

481   and the evaporation rates of high-elevation small and large lakes, J. Hydrol., 573, 220–234, 2019.

482   Wang, W., Xiao, W., Cao, C., Gao, Z., Hu, Z., Liu, S., Shen, S., Wang, L., Xiao, Q., Xu, J., Yang, D. and Lee,

483   X.: Temporal and spatial variations in radiation and energy balance across a large freshwater lake in China, J.

484   Hydrol., 511, 811–824, doi:10.1016/j.jhydrol.2014.02.012, 2014.

485   Wang, Y., Gao, Y., Qin, H., Huang, J., Liu, C., Hu, C., Wang, W., Liu, S. and Lee, X.: Spatiotemporal

486   Characteristics of Lake Breezes over Lake Taihu, China, J. Appl. Meteorol. Climatol., 56(7), 2053–2065, 2017.

487   Webb, E. K., Pearman, G. I. and Leuning, R.: Correction of flux measurements for density effects due to heat

488   and water vapour transfer, Q. J. R. Meteorol. Soc., 106(447), 85–100, 1980.

489   Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R.,

490   Dolman, H. and Field, C.: Energy balance closure at FLUXNET sites, Agric. For. Meteorol., 113(1–4), 223–243,

491   2002.

492   Xiao, K., Griffis, T. J., Baker, J. M., Bolstad, P. V, Erickson, M. D., Lee, X., Wood, J. D., Hu, C. and Nieber, J.

493   L.: Evaporation from a temperate closed-basin lake and its impact on present, past, and future water level, J.

494   Hydrol., 561, 59–75, 2018.

495   Xiao, W., Liu, S., Wang, W., Yang, D., Xu, J., Cao, C., Li, H. and Lee, X.: Transfer Coefficients of Momentum,

496   Heat and Water Vapour in the Atmospheric Surface Layer of a Large Freshwater Lake, Boundary-Layer

497   Meteorol., 148(3), 479–494, doi:10.1007/s10546-013-9827-9, 2013.

498   Xiao, W., Liu, S., Li, H., Xiao, Q., Wang, W., Hu, Z., Hu, C., Gao, Y., Shen, J., Zhao, X., Zhang, M. and Lee,

499   X.: A flux-gradient system for simultaneous measurement of the CH4, CO2, and H2O fluxes at a lake-air

500   interface, Environ. Sci. Technol., 48(24), 14490–14498, doi:10.1021/es5033713, 2014.

501   Xu, J., Lee, X., Xiao, W., Cao, C., Liu, S., Wen, X., Xu, J., Zhang, Z. and Zhao, J.: Interpreting the $^{13}C/^{12}C$ ratio

502   of carbon dioxide in an urban airshed in the Yangtze River Delta, China, Atmos. Chem. Phys., 17(5),

503   doi:10.5194/acp-17-3385-2017, 2017.

504   Yusup, Y. and Liu, H.: Effects of Atmospheric Surface Layer Stability on Turbulent Fluxes of Heat and Water

505   Vapor across the Water–Atmosphere Interface, J. Hydrometeorol., 17(11), 2835–2851, 2016.

506   Zhang, M., Xiao, Q., Zhang, Z., Gao, Y., Zhao, J., Pu, Y., Wang, W., Xiao, W., Liu, S. and Lee, X.: Methane

507   flux dynamics in a submerged aquatic vegetation zone in a subtropical lake, Sci. Total Environ., 672,

508   doi:10.1016/j.scitotenv.2019.03.466, 2019a.

509   Zhang, X., Huang, J., Li, G., Wang, Y., Liu, C., Zhao, K., Tao, X., Hu, X.-M. and Lee, X.: Improving

510   Lake-Breeze Simulation with WRF Nested LES and Lake Model over a Large Shallow Lake, J. Appl. Meteorol.

511   Climatol., 58(8), 1689–1708, 2019b.

512   Zhang, Z., Zhang, M., Cao, C., Wang, W., Xiao, W., Xie, C., Chu, H., Wang, J., Zhao, jiayu, Jia, L., Liu, Q.,

513   Huang, W., Zhang, W., Lu, Y., Xie, Y., Wang, Y., Pu, Y., Hu, Y., Chen, Z., Qin, Z. and Lee, X.: A dataset of

514   microclimate and radiation and energy fluxes from the Lake Taihu Eddy Flux Network, Harvard Dataverse,

515   https://doi.org/10.7910/DVN/HEWCWM, 2020

516   Zhao, X. and Liu, Y.: Variability of surface heat fluxes and its driving forces at different time scales over a large

517   ephemeral lake in China, J. Geophys. Res. Atmos., 123(10), 4939–4957, 2018.

518