1 **Averaging over spatiotemporal heterogeneity substantially biases evapotranspiration rates in a mechanistic**

2 **large-scale land evaporation model**

3 Elham Rouholahnejad Freund[1,2,3], Massimiliano Zappa[4], James W. Kirchner[3,4,5]

5 [1]Laboratory of Hydrology and Water Management, Ghent University, Ghent, Belgium

6 [2]Chair of Hydrology, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany

7 [3]Department of Environmental Systems Science, ETH Zurich, CH-8092 Zürich, Switzerland

8 [4]Swiss Federal Research Institute WSL, CH-8903 Birmensdorf, Switzerland

9 [5]Department of Earth and Planetary Science, University of California, Berkeley, CA 94720 USA

11 Correspondence to: Elham Rouholahnejad Freund, elham.rouholahnejad@gmail.com

13 **Abstract**

14 Evapotranspiration (ET) influences land-climate interactions, regulates the hydrological cycle, and contributes

15 to the Earth's energy balance. Due to its feedback to large-scale hydrological processes and its impact on

16 atmospheric dynamics, ET is one of the drivers of droughts and heatwaves. Existing land surface models differ

17 substantially, both in their estimates of current ET fluxes and in their projections of how ET will evolve in the

18 future. Any bias in estimated ET fluxes will affect the partitioning between sensible and latent heat, and thus

19 alter model predictions of temperature and precipitation. One potential source of bias is the so-called

20 "aggregation bias" that arises whenever nonlinear processes, such as those that regulate ET fluxes, are

21 modeled using averages of heterogeneous inputs. Here we demonstrate a general mathematical approach to

22 quantifying and correcting for this aggregation bias, using the GLEAM land evaporation model as a relatively

23 simple example. We demonstrate that this aggregation bias can lead to substantial overestimates in ET fluxes

24 in a typical large-scale land surface model when sub-grid heterogeneities in land surface properties are

25 averaged out. Using Switzerland as a test case, we examine the scale-dependence of this aggregation bias and

26 show that it can lead to an average overestimation of daily ET fluxes by as much as 10% across the whole

27 country (calculated as the median of the daily bias over the growing season). We show how our approach can

28 be used to identify the dominant drivers of aggregation bias, and to estimate sub-grid closure relationships

29 that can correct for aggregation biases in ET estimates, without explicitly representing sub-grid heterogeneities

30 in large-scale land surface models.

31 **Plain Language Summary**

32 Evapotranspiration (ET) is the largest flux from the land to the atmosphere and thus contributes to Earth's

33 energy and water balance. Due to its impact on atmospheric dynamics, ET is a key driver of droughts and

34 heatwaves. In this paper, we demonstrate how averaging over land surface heterogeneity contributes to

35 substantial overestimates of ET fluxes. We also demonstrate how one can correct for the effects of small-scale

36 heterogeneity without explicitly representing it in land surface models.

37

1. Introduction

Earth's surface and subsurface are characterized by spatial heterogeneity spanning wide ranges of scales, including scales that cannot be explicitly resolved by large-scale Earth System Models (ESMs), which are typically run at resolutions of 10-100 kilometers. Averaging over this finer-scale heterogeneity can bias model estimates of water and energy fluxes and hence alter future temperature predictions. Earth system model estimates of global terrestrial evaporation differ substantially from atmospheric reanalyses based on in-situ and satellite remote sensing observations (Mueller et al., 2013), but it is unclear how much of these differences could be attributed to errors in capturing sub-grid heterogeneity.

Several recent studies (e.g., Fan et al., 2019; Shrestha et al., 2018) have emphasized the need to account for land surface heterogeneity in large-scale ESMs. Despite recent community efforts in refining ESMs' spatial resolution (Huang et al., 2016; Rauscher et al., 2010; Ringler et al., 2008; Skamarock et al., 2012; Zarzycki et al., 2014), the grid resolution of present-day ESMs is still too coarse to explicitly capture important effects of surface heterogeneity. Whether the solution lies in hyper-resolution large-scale land surface modeling remains an open question, because heterogeneities that are important to land-atmosphere fluxes will not be fully resolved even at scales of 100 m (Beven and Cloke, 2012).

The effects of aggregating over spatial heterogeneity in land surface models have been assessed using several approaches. Most of these approaches compare grid-cell-averaged energy and water fluxes with flux estimates for finer-resolution grids, or for grid cells that are subdivided into mosaics of several surface types which separately exchange momentum, energy, and water vapor with the overlying atmosphere (e.g., Giorgi, 1997). Several studies have reported increases in average evapotranspiration (ET) (e.g., Kuo et al., 1999; Boone and Wetzel, 1998; Hong et al., 2009; McCabe and Wood, 2006; El Maayar and Chen, 2006), and at least one has reported decreases in grid-cell average ET (Ershadi et al., 2013), as model grids are coarsened and less spatial heterogeneity is accounted for. Shrestha et al. (2018) studied the effects of horizontal grid resolution on ET partitioning in the TerrSysMP Earth system model and found that the aggregation of topography decreases average slope gradients and obscures small-scale convergence and divergence zones, directly impacting surface and subsurface flow. They observed 5 and 8 percent decreases in the transpiration/evapotranspiration ratio for a dry and a wet year, respectively, when their model grid cells were coarsened from 120 m to 960 m. All these studies calculate the effects of land surface heterogeneity on ET fluxes using numerical experiments that refine the model's spatial resolution, either directly or through the use of land-surface mosaics.

Quantifying the effect of sub-grid scale heterogeneity on grid-cell-averaged fluxes is especially important when highly nonlinear processes are involved. Regardless of scale, the main challenge is not to explicitly represent the heterogeneity in all its details, but instead to define an appropriate scale-dependent sub-grid closure relationship that recognizes the important heterogeneities within the grid elements and the nonlinearities in

74 the processes (Beven, 2006). Such a sub-grid closure scheme would capture the effects of sub-grid

75 heterogeneity in large-scale land surface models without forcing them to run at finer spatial resolutions.

76

77 We have recently proposed a general theoretical framework, based on Taylor series expansions, that

78 quantifies the "aggregation bias" that results from averaging over sub-grid heterogeneity when grid-cell-

79 averaged ET is estimated (Rouholahnejad Freund and Kirchner, 2017; Rouholahnejad Freund et al., 2020a). In

80 contrast to the numerical experiments described above, this theoretical framework does not depend on a

81 particular evapotranspiration model or grid scale. Our previous work demonstrated this framework using

82 Budyko curves as a see-through "toy" model, leaving open the question of how strongly ET estimates would be

83 affected by sub-grid heterogeneity in a more typical mechanistic evapotranspiration model. Here we use the

84 mechanistic evapotranspiration model GLEAM to quantify how aggregation biases vary across a range of

85 scales, using Switzerland as a case study. We show how our Taylor expansion framework can be used to

86 quantify the sensitivity of ET fluxes to heterogeneity in their individual drivers. We further demonstrate how

87 this framework can be used to estimate correction factors (i.e., sub-grid closure relationships) that account for

88 the effects of sub-grid heterogeneity without explicitly modeling it, and show how these correction factors can

89 be used to improve grid-scale ET estimates. Because our framework is not model-specific, the analysis

90 presented here could also be applied to many other evapotranspiration algorithms.

91

92 **2.   Methods and results**

93 **2.1. A common mechanistic framework for predicting evapotranspiration**

94 Most large-scale land surface models calculate ET as a function of available water and energy at daily time

95 steps. They typically multiply an estimate of potential evapotranspiration (PET) by a conversion factor to

96 calculate actual evapotranspiration. PET is generally understood as the maximum rate of evapotranspiration

97 from a large area (to avoid the effect of local advection) covered completely and uniformly by actively growing

98 vegetation with adequate moisture at all times (Brutseart, 1984). Models typically estimate PET using the

99 Penman equation (Penman, 1948; intended for open water surfaces), the Penman-Monteith equation

100 (Monteith, 1965, Monteith and Unsworth, 1990; intended for reference crop evapotranspiration by adding

101 atmospheric transport processes and stomatal resistance to Penman's open water evaporation), or the

102 Priestley-Taylor equation (Priestley and Taylor, 1972; intended for open water and water-saturated crops and

103 grasslands). The conversion factor that is used to estimate ET from PET typically depends on plant physiology

104 and on the water that is available for evaporation.

105

106 Here, we employ an ET algorithm that is used by several land surface models (i.e., Global Land-surface

107 Evaporation: The Amsterdam Methodology (GLEAM); Miralles et al., 2011; Martens et al., 2017), in which

108 actual ET is calculated as a fraction of PET. This fraction is expressed as a multiplicative factor, often called a

109 stress factor, which ranges between 0 and 1 and thus limits ET rates. Under wet conditions, ET can equal PET

110 (stress factor equals one) while under dry conditions, PET is multiplied by a stress factor smaller than one

111 depending on the degree of water stress. This approach is employed by the GLEAM model, among others.

112   GLEAM is a diagnostic satellite-data-driven method that is used to estimate global land evaporation fluxes.

113   GLEAM uses the Priestley-Taylor formula and remotely sensed datasets of radiation and temperature to

114   calculate PET. In GLEAM, actual ET is calculated by constraining PET estimates by a stress factor that is based

115   on estimates of root-zone soil moisture. The root zone soil moisture is derived from a multi-layer water

116   balance module that describes the infiltration of precipitation through the vertical soil profile. ET estimates

117   from GLEAM have been applied in many studies (e.g., Miralles et al., 2013; Miralles et al., 2014; Greve et al.,

118   2014; Jasechko et al., 2013). GLEAM operates on daily time steps at 0.25-degree spatial resolution. 0.25

119   degrees is about 27.6 km in the north-south direction and 18.9 km in the east-west direction at the latitude of

120   Switzerland. To the best of our knowledge, there are no prior studies quantifying the aggregation bias in ET

121   estimates from GLEAM or other models with similar ET formulations.

122

123   GLEAM calculates ET as an explicit function of the stress factor and potential evaporation:

$$ET = S \cdot PET + (1 - \beta)\, I, \tag{1}$$

125   where ET is actual evapotranspiration (mm d^{-1}), S is the evaporative stress factor (-) that accounts for

126   environmental conditions that reduce actual ET relative to potential ET, I is interception loss (mm d^{-1}), and β is

127   a constant (β= 0.07 – Gash and Stewart, 1977) that avoids double-counting of interception losses during hours

128   with wet canopy. The stress factor (S) depends on the soil moisture conditions, and is parametrized separately

129   for tall canopy, short vegetation, and bare soil. GLEAM uses the following soil-moisture-based

130   parameterization to calculate the stress factor (Miralles et al., 2011; Martens et al., 2017):

$$S = 1 - \left(\frac{w_c - w_w}{w_c - w_{wp}}\right)^2, \tag{2}$$

132   where S is the stress factor (-) for tall canopy, w_w is soil moisture saturation at any given time (-), and

133   w_c and w_{wp} are the critical soil moisture saturation level and wilting point. For soil moisture saturation values

134   below the wilting point w_{wp}, the stress is maximal (stress factor equals 0), causing ET to sharply decline to

135   zero. For values above the critical moisture level w_c, there is no water stress (stress factor equals 1) and ET

136   equals PET. Between w_{wp} and w_c the stress increases as soil moisture decreases following a parabolic function

137   (Eq. 2). In the analysis presented below, we set the critical soil moisture level (w_c) and wilting point (w_{wp}) to

138   0.6 and 0.1, respectively. To simplify the analysis presented below, we have used the tall-canopy stress factor

139   (Eq. 2) for all of Switzerland, even though the short-canopy or bare-soil formulations may be better suited to

140   some locations.

141

142   GLEAM uses the Priestley-Taylor approach to calculate PET (Priestley and Taylor, 1972):

$$PET = \frac{\alpha}{\lambda}\,\frac{\Delta}{\Delta + \gamma}(R_n - G), \tag{3}$$

144   where PET is potential evapotranspiration (mm d^{-1}), α is a dimensionless coefficient that parametrizes the

145   resistance to evaporation and is set to 0.8 for tall canopy in GLEAM (Miralles et al., 2011), $\lambda = 2.26$ (MJ kg^{-1}) is

146   the latent heat of vaporization, R_n is net radiation (MJ m^{-2} d^{-1}), G is the ground heat flux, approximated as

147   G=0.05 R_n (MJ m^{-2} d^{-1}) for tall canopy in GLEAM, T is temperature (°C), and Δ is the slope of the

148 temperature/saturated vapor pressure curve (kPa°C^{-1}), which is functionally related to temperature (Tetens,

149 1930; Murray, 1967; Stanghellini, 1987):

$$\Delta = ae^{bT},\tag{4}$$

151 where a= 0.04145 (kPa°C^{-1}), b=0.06088 (°C^{-1}), and γ is the psychrometric constant (kPa°C^{-1}) which can be

152 calculated as (Brunt, 1952):

$$\gamma = \frac{C_{p_{air}} * P}{\lambda * MW_{ratio}},\tag{5}$$

154 where $C_{p_{air}}$ = 0.001013 (MJ kg^{-1}°C^{-1}) is the specific heat of air at constant pressure, P = 101.3 (KPa) is

155 atmospheric pressure, and MW_{ratio} =0.622 (-) is the molecular weight ratio of H$_2$O/air. Substituting the

156 aforementioned constants in Eq. 5 yields γ= 0.073 (kPa°C^{-1}). Expanding Eq. 1 using Eqs. 2-5 yields the ET

157 function as calculated by GLEAM:

158
$$ET_{[mmd^{-1}]} = \left[-4w_{w[m^3m^{-3}]}^2 + 4.8w_{w[m^3m^{-3}]} - 0.44\right] * \frac{\alpha_{[\ ]}}{\lambda_{[MJ\ kg^{-1}]}} * \frac{\Delta_{[kPa°C^{-1}]}}{\Delta_{[kPa°C^{-1}]} + \gamma_{[kPa°C^{-1}]}}$$

$$* 0.95 * \frac{86400}{1000000} * R_{n[Wm^{-2}]} + (1 - \beta)\ I_{[mmd^{-1}]}$$

$$= [-4w_w{}^2 + 4.8w_w - 0.44] * 0.02905 * \frac{a\ e^{bT}}{a\ e^{bT} + 0.073} R_n + (1 - 0.07)\ I_{[mmd^{-1}]}.\tag{6}$$

159

160 In the analysis below, we use the GLEAM evapotranspiration algorithm to demonstrate how aggregation biases

161 can be estimated in land surface modeling schemes. We chose GLEAM because its governing equations are

162 amenable to the analytical solutions derived below. Here we make no particular claim for the accuracy or

163 validity of GLEAM as an evapotranspiration model, nor is our analysis intended to test this. Likewise our

164 analysis should not be interpreted as implying that GLEAM is any more, or less, susceptible to aggregation bias

165 than other evapotranspiration schemes, because this question is beyond the scope of the current paper.

166

167 **2.2. Mathematical framework for predicting aggregation bias**

168 *Nonlinear averaging using second-order Taylor expansions*

169 ET is a nonlinear function of its drivers. An intrinsic property of any nonlinear function is that the average of

170 the function will not equal the function evaluated at the average inputs (e.g., Rastetter et al., 1992; Giorgi and

171 Avissar, 1997; Rouholahnejad Freund and Kirchner, 2017). Thus averaging over sub-grid heterogeneity in ET

172 drivers, as large-scale land surface models do, would be expected to lead to biased ET estimates, even if the

173 underlying equations were exactly correct. For an ET function of three variables, namely R_n, w_w, and T, the

174 mean of the ET function, in terms of the function's value at the mean of its inputs, can be approximated by the

175 second-order Taylor series expansion of the ET function (Eq. 6):

176
$$\overline{ET} \approx \widehat{ET} + \frac{1}{2}\left[\frac{\partial^2 ET}{\partial R_n{}^2}Var(R_n) + \frac{\partial^2 ET}{\partial w_w{}^2}Var(w_w) + \frac{\partial^2 ET}{\partial T^2}Var(T)\right]$$

$$+ \frac{\partial^2 ET}{\partial R_n \partial T}Cov(R_n, T) + \frac{\partial^2 ET}{\partial R_n \partial w_w}Cov(R_n, w_w) + \frac{\partial^2 ET}{\partial w_w \partial T}Cov(w_w, T),\tag{7}$$

177 where $\overline{ET}$ is the estimate of the true average of the nonlinear ET function over its variable inputs, $\widehat{ET}$ is the ET

178 function evaluated at its mean inputs, and the derivatives are understood to be evaluated at the mean values

179 of the variables ($\overline{R_n}, \overline{w_w}, \overline{T}$) and multiplied by the corresponding variances and covariances among finer-

180 resolution input data. For the specific case of the GLEAM model, the ET function is evaluated at its mean inputs

181 ($\widehat{ET}$) and these derivatives are derived analytically from the ET function described by Eq. 6, directly yielding the

182 following expressions:

$$\widehat{ET} = \left[-4\bar{w}_w{}^2 + 4.8\bar{w}_w - 0.44\right] * 0.02905 * \frac{a\,e^{b\bar{T}}}{a\,e^{b\bar{T}} + 0.073}\bar{R}_n, \tag{8}$$

$$\frac{\partial^2 ET}{\partial R_n{}^2} = 0, \tag{9}$$

$$\frac{\partial^2 ET}{\partial w_w{}^2} = [-8] * 0.02905 * \frac{\Delta}{\Delta + \gamma}R_n \qquad (w_{wp} \le w_w \le w_c), \tag{10a}$$

$$\frac{\partial^2 ET}{\partial w_w{}^2} = 0 \qquad (w_w < w_{wp}, \quad w_w > w_c), \tag{10b}$$

$$\frac{\partial^2 ET}{\partial T^2} = [-4w_w{}^2 + 4.8w_w - 0.44] * 0.02905 * R_n * b^2 * \frac{\gamma^2\Delta - \gamma\Delta^2}{(\gamma + \Delta)^3}, \tag{11}$$

$$\frac{\partial^2 ET}{\partial R_n \partial T} = [-4w_w{}^2 + 4.8w_w - 0.44] * 0.02905 * \frac{\Delta}{\Delta + \gamma} * \frac{b\gamma}{\Delta + \gamma}, \tag{12}$$

$$\frac{\partial^2 ET}{\partial R_n \partial w_w} = [-8w_w + 4.8] * 0.02905 * \frac{\Delta}{\Delta + \gamma} \qquad (w_{wp} \le w_w \le w_c), \tag{13a}$$

$$\frac{\partial^2 ET}{\partial R_n \partial w_w} = 0 \qquad (w_w < w_{wp}, \quad w_w > w_c), \tag{13b}$$

$$\frac{\partial^2 ET}{\partial w_w \partial T} = [-8w_w + 4.8] * 0.02905 * \frac{\Delta}{\Delta + \gamma} * \frac{b\gamma}{\Delta + \gamma} * R_n \qquad (w_{wp} \le w_w \le w_c), \text{ and} \tag{14a}$$

$$\frac{\partial^2 ET}{\partial w_w \partial T} = 0 \qquad (w_w < w_{wp}, \quad w_w > w_c), \tag{14b}$$

193 where Δ depends on temperature as described in Eq. (4). The difference between the average of the functions

194 ($\overline{ET}$) and the function of the averages ($\widehat{ET}$), or, equivalently, the sum of all the other terms in Eq. (7),

195 represents the aggregation bias. The magnitude of this bias can be calculated by combining Eqs. 7-14 with

196 estimates of the variances and covariances of the input variables. Note that the interception term in equation

197 6 is dropped out from the derivatives as the interception loss in GLEAM is a linear function of amount of

198 rainfall necessary to saturate the canopy and therefore has negligible effect when averaged.

199

200 The approach outlined in Eq. (7) is general and could be extended to other land surface modeling schemes.

201 The partial derivatives in Eqs. (8-14), of course, are specific to the GLEAM equations; for other models they

202 would differ. More complex land surface model algorithms may not have such simple analytical derivatives; in

203 that case, the derivatives can be evaluated numerically.

204

205 **2.3. Sub-grid heterogeneity and aggregation bias in ET estimates across Switzerland**

206 Drivers of ET (i.e., soil moisture, net radiation, and temperature) can be highly heterogeneous within the grid

207 cells of typical ESMs. Soil moisture can show pronounced spatial variability, especially in areas where surface

208 roughness, porosity, and permeability vary by orders of magnitude across a variety of length scales (Giorgi and

209 Avissar, 1997). Temperature and incoming radiation vary significantly with season, elevation, altitude, and

210   albedo. Switzerland, for example, shows strong local variations in average annual temperature, soil moisture

211   content, net radiation, and albedo (Fig. 1; albedo values in Fig. S1).

212

213   We quantified how averaging over spatial (and temporal) heterogeneities of ET drivers affects estimated ET at

214   several grid scales across Switzerland, as an example case for which high-resolution data are available. Our

215   analysis is based on 500-m input data of temperature (interpolation of MeteoSwiss data after Viviroli et al.,

216   2009), net radiation (Viviroli et al., 2009), and soil moisture (simulations from the hydrological model PREVAH,

217   Brunner et al., 2019; Speich et al., 2015; Orth et al., 2015; Zappa et al., 2003) at daily time steps for the 2004

218   growing season. Although our soil moisture data are derived from model simulations whose accuracy is

219   difficult to assess due to the scarcity of real-world soil moisture measurements, for our purposes all that is

220   necessary is that the simulated values exhibit realistically complex spatial variability.

221

222   We used the GLEAM equations, as outlined in Sect. 2, to calculate ET for each day at the 500-m resolution of

223   these input data. We use these 500-m ET estimates as virtual "truth" for the purpose of our analysis, because

224   our goal is not to determine whether GLEAM estimates of ET are accurate (compared to direct measurements,

225   for example), but rather to quantify how spatial aggregation affects them.

226

227   To quantify how spatial aggregation affects model estimates of ET, we calculated ET over larger spatial scales

228   in two different ways. We First calculated the arithmetic average of the 500-m ET estimates over 1/32, 1/16,

229   1/8, 0.25, 0.5, 0.75, 1, and 2-degree grid cells across Switzerland, to represent the "true" average ET at those

230   grid scales. Next we calculated the arithmetic average of the 500-m input data (of temperature, soil moisture,

231   and net radiation) over the same grid cells, and then used these grid-cell-averaged input data in the GLEAM

232   equations to calculate the modeled coarse-resolution ET at each grid scale. The deviation of the modeled

233   coarse-resolution ET from the "true" average ET measures the aggregation bias. Because this numerical

234   experiment uses the same model equations, based on the same underlying data, for the ET calculations at

235   each spatial resolution, it isolates spatial aggregation as the only possible cause of the difference between the

236   "true" average ET ($\overline{ET}$ in Eq. 7) and the coarse-resolution modeled ET ($\widehat{ET}$ in Eq. 7) at each grid scale.

237

238   Figure 2a shows that the ET aggregation bias varies considerably across Switzerland, and also varies

239   considerably with grid scale. The average aggregation bias is higher at coarser grid scales, averaging 10% at 2-

240   and 1-degree grid resolution across all of Switzerland (calculated as the median of the daily aggregation biases

241   over the growing season; Fig. 2a). Smaller grid scales typically exhibit smaller aggregation biases (averaging 4%

242   at 1/16-degree grid resolution across all of Switzerland calculated as the median of the daily aggregation

243   biases over the growing season) because they typically average over less spatial heterogeneity, but even at the

244   smallest grid scales, aggregation biases can locally reach 40% as indicated by the scatter plot in Fig. 3. These

245   figures are medians of the daily aggregation biases over the entire growing season of 2004; the aggregation

246   biases of two arbitrarily selected days (May 29th and July 18th, 2004) at several spatial scales lead to much

247 larger overestimation of ET in parts of southern Switzerland (Figs. S2, S3). The two selected days are days 150

248 and 200 of Julian day calendar of year 2004.

249

250 Using our 500-m input data, we can test how well Eq. (7) estimates the difference between the "true" average

251 ET and the coarse-resolution modeled ET at each grid scale. We used Eqs. (8-14) to calculate the partial

252 derivatives of the GLEAM equations for each grid cell and time step, using the grid-cell averaged values of the

253 input data. We then multiplied these derivatives by the corresponding variances and covariances among the

254 500-m input data to obtain bias estimates via Eq. (15) for each grid cell and time step:

$$\text{Bias} = \widehat{\text{ET}} - \overline{\text{ET}} \approx -\frac{1}{2}\left[\frac{\partial^2\text{ET}}{\partial R_n^2}\text{Var}(R_n) + \frac{\partial^2\text{ET}}{\partial w_w^2}\text{Var}(w_w) + \frac{\partial^2\text{ET}}{\partial T^2}\text{Var}(T)\right]$$

255

$$-\frac{\partial^2\text{ET}}{\partial R_n\partial T}\text{Cov}(R_n,T) - \frac{\partial^2\text{ET}}{\partial R_n\partial w_w}\text{Cov}(R_n,w_w) - \frac{\partial^2\text{ET}}{\partial w_w\partial T}\text{Cov}(w_w,T), \qquad (15)$$

256 where $\overline{\text{ET}}$ is the true average ET at some grid resolution, $\widehat{\text{ET}}$ is the modeled coarse-resolution ET at the same

257 spatial scale, the right-hand side is the Taylor expansion estimate of the aggregation bias. We then compared

258 these estimated biases against the "true" aggregation biases (the difference between the "true" average ET

259 and the coarse-resolution modeled ET) in the numerical experiment described above. The true bias, in other

260 words, is $\widehat{\text{ET}} - \overline{\text{ET}}$ in Eq. (15), and the estimated bias is the Taylor approximation on the right-hand side.

261

262 Figure 2b shows that the aggregation bias estimated by Eq. (15) is generally similar, in both overall magnitude

263 and spatial distribution, to the "true" aggregation biases calculated by the numerical experiment. This

264 comparison is shown more explicitly in Fig. 3, in which the estimated aggregation bias is compared with the

265 "true" aggregation bias for each grid cell at each grid scale. Figures 2 and 3 show that Eq. (15) is generally a

266 good predictor of aggregation bias. Both the estimated aggregation biases (Fig. 2) and the "true" aggregation

267 biases are markedly higher in regions of greater topographic complexity (Fig. S4).

268

269 **2.4. Correcting for aggregation bias**

270 2.4.1. Identifying drivers of aggregation bias

271 The Taylor expansion in Eq. (15) not only allows one to quantify the aggregation bias; it also allows one to

272 quantify the relative importance of the three input variables (net radiation, soil moisture, and temperature) as

273 drivers of that bias. Each of the terms in Eq. (15) combines a variance or covariance that expresses how

274 variable the input data are, and a second derivative that expresses how sensitive the average ET is to that

275 variability. Each of these terms – a derivative multiplied by a variance or covariance – has the same units as ET,

276 and thus they can be directly compared to one another.

277

278 Table 1 shows each of the aggregation bias terms, calculated over all of Switzerland for the two arbitrarily

279 chosen days mentioned in Sect. 2.3 (May 29th and July 18th, 2004). For these two example days, the

280 aggregation bias is clearly dominated by a single term, associated with the variance of soil moisture. The

281 variance in net radiation (Rn) creates no aggregation bias, because GLEAM ET is a linear function of Rn; thus

282 positive and negative deviations from average Rn will increase and decrease ET by exactly offsetting amounts.

283 Similarly, the variance in temperature (T) also results in little aggregation bias, because GLEAM ET increases

284 nearly linearly with T across a wide range of temperature. The covariance terms similarly lead to little

285 aggregation bias. By contrast, the strong curvature in the quadratic dependence of ET on soil moisture (Eq. 6)

286 implies that positive and negative deviations from mean soil moisture will not have offsetting ET effects, and

287 thus that spatial heterogeneity in soil moisture can significantly alter average ET. On most of the days of the

288 year 2004, the soil moisture variance term is the dominant driver of the aggregation bias. However, there are

289 some days in which other factors such as the T and Rn covariance term are the dominant factors.

290

291 2.4.2. Correcting for aggregation bias using sub-grid closure relationships

292 The Taylor expansion framework in Eq. (7) can be used not only to diagnose aggregation bias, but also to

293 estimate sub-grid closure relationships that correct for the effects of small-scale heterogeneity. The variance

294 and covariance terms in Eq. (7) express how sub-grid heterogeneity affects average ET at the grid scale,

295 implying that these aggregation bias estimates could be used to improve grid-scale ET estimates, without

296 explicitly modeling ET at high resolutions. This approach could be particularly useful in land surface algorithms

297 that are part of coarser-resolution Earth system models; in such cases it may be much more efficient to

298 evaluate Eqs. 7-14 at the coarse grid resolution than to directly evaluate the underlying ET model, Eq. 6, at

299 high resolution. The Taylor expansion approach could also be attractive where we lack spatially explicit high-

300 resolution maps of the ET drivers, but where their variances and covariances can nonetheless be estimated

301 from other sources (such as from the variability of topography, mapped soil units, remote sensing data, etc.).

302

303 It is beyond our scope here to construct such variance and covariance estimates, but we can illustrate how

304 they could potentially be used. The solid red symbols in Fig. 4 show the relationships between "true" average

305 ET and modeled grid-cell-averaged ET, for each grid cell (and one example day, May 31st, 2004) at several

306 different grid scales. For comparison, the open grey symbols in Fig. 4 show average ET estimated by the Taylor

307 expansion approach of Eq. (7), which corrects for sub-grid heterogeneity effects using only grid-cell-averaged

308 estimates of the ET drivers and their small-scale variances and covariances.

309

310 The heterogeneity-corrected ET estimates shown by the open symbols in Fig. 4 cluster much closer to the 1:1

311 line than the modeled grid-cell-averaged ET values shown by the solid red symbols, suggesting that the Taylor

312 expansion approach may substantially improve estimates of grid-cell-averaged ET. Real-world results may be

313 less clear than those shown in Fig. 4, because the heterogeneity-corrected ET estimates (the open symbols in

314 Fig. 4) are calculated using exact values for the variances and covariances of the ET drivers within each grid

315 cell, and in real-world cases these variances and covariances will not be known precisely. Figure 4 nonetheless

316 demonstrates the potential value of knowing, or being able to estimate, those variances and covariances.

317   Efforts to determine those variances and covariances can be focused on the terms that matter the most, if one

318   can identify the main drivers of aggregation bias using the methods described in Sect. 2.2 above.

319
320
321
322
323
324
325
326
327
328   Table 1. Relative importance of different ET drivers in aggregation bias estimates (different terms in Eq. 15). Va
329   lues are calculated for all of Switzerland for the two arbitrarily chosen days (May 29th and July 18th, 2004). The
330   aggregation bias is dominated by the term associated with the variance of soil moisture for these two example
331   days.

	$\widehat{ET}$ mm d^{-1}	$\overline{ET}$ mm d^{-1}	Bias %	Contribution of $Var(R_n)$ term in % aggregation bias (%)	Contribution of $Var(w_w)$ term in % aggregation bias (%)	Contribution of $Var(T)$ term in % aggregation bias (%)	Contribution of $Cov(R_n,T)$ term in % aggregation bias (%)	Contribution of $Cov(R_n,w_w)$ term in % aggregation bias (%)	Contribution of $Cov(R_n,w_w)$ term in % aggregation bias (%)
Calculation	(Eq. 8)	(Eq. 7)	(Eq. 15)	$\dfrac{\frac{1}{2}\frac{\partial^2 ET}{\partial R_n^2}Var(R_n)}{(\widehat{ET}.Bias)}$	$\dfrac{\frac{1}{2}\frac{\partial^2 ET}{\partial w_w^2}Var(w_w)}{(\widehat{ET}.Bias)}$	$\dfrac{\frac{1}{2}\frac{\partial^2 ET}{\partial T^2}Var(T)}{(\widehat{ET}.Bias)}$	$\dfrac{\frac{\partial^2 ET}{\partial R_n\partial T}Cov(R_n,T)}{(\widehat{ET}.Bias)}$	$\dfrac{\frac{\partial^2 ET}{\partial R_n\partial w_w}Cov(R_n,w_w)}{(\widehat{ET}.Bias)}$	$\dfrac{\frac{\partial^2 ET}{\partial w_w\partial T}Cov(w_w,T)}{(\widehat{ET}.Bias)}$
29.05.2004	2.3	1.89	21.7	0	81.65	0.90	1.05	2.80	14.41
18.07.2004	2.11	1.84	14.84	0	83.35	2.34	6.56	1.84	6.01

332
333

334
335
336

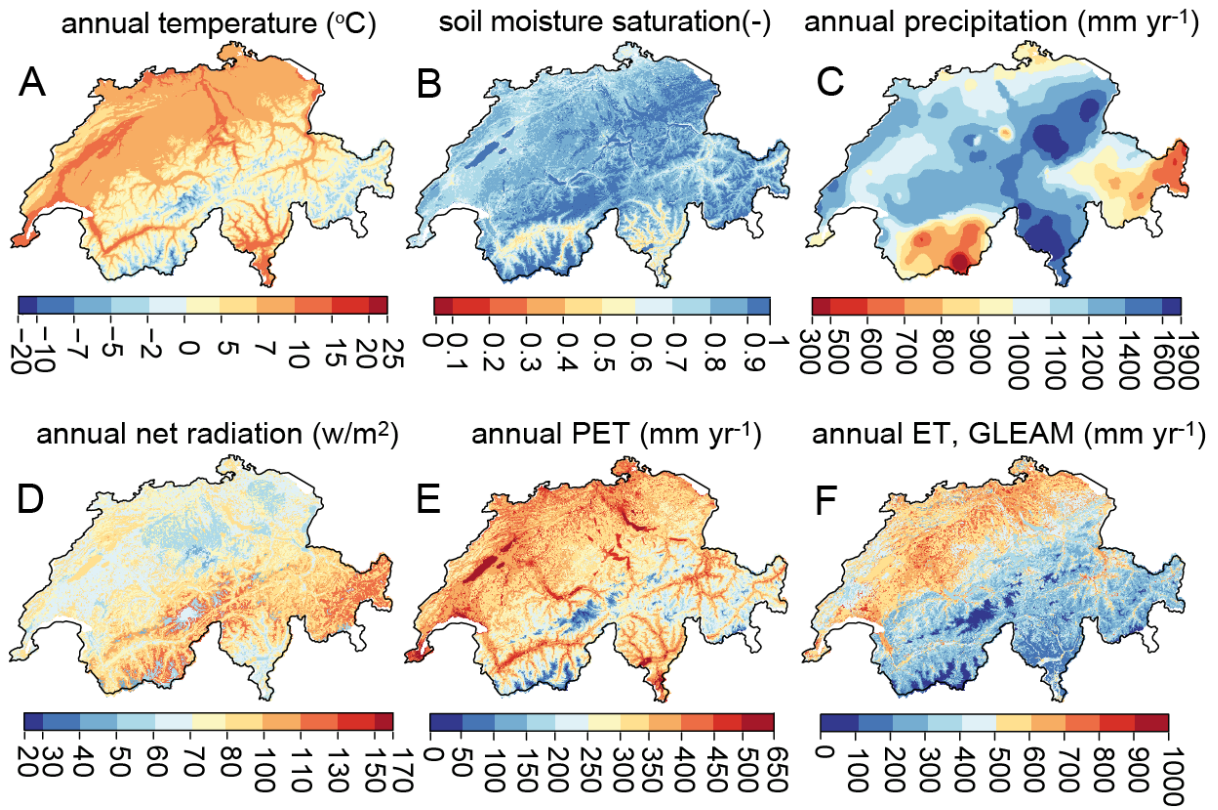



Figure 1. Spatial distribution of input data for the year 2004 at 500-m resolution: Annual mean (A) temperature (°C), (B) soil moisture saturation (-, simulated by the PREVAH hydrological model), (C) precipitation (mm yr^{-1}), (D) net radiation (W m^{-2}), (E) potential evapotranspiration (PET, mm yr^{-1}) using the Priestley-Taylor equation (Eq. 3), and (F) evapotranspiration (ET, mm yr^{-1}) using the approach used in the GLEAM model (Eq. 1). See Table. S1 for references.

343

344

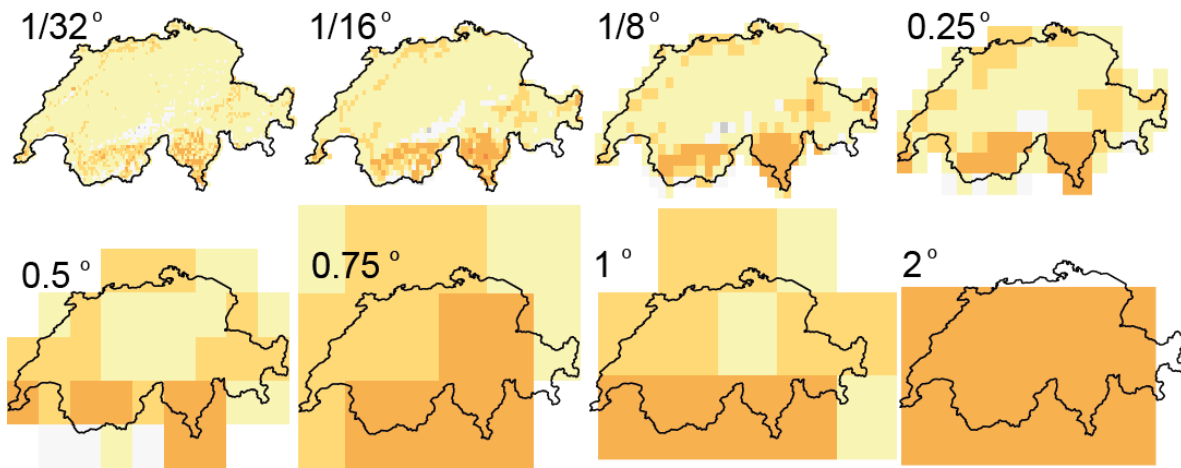
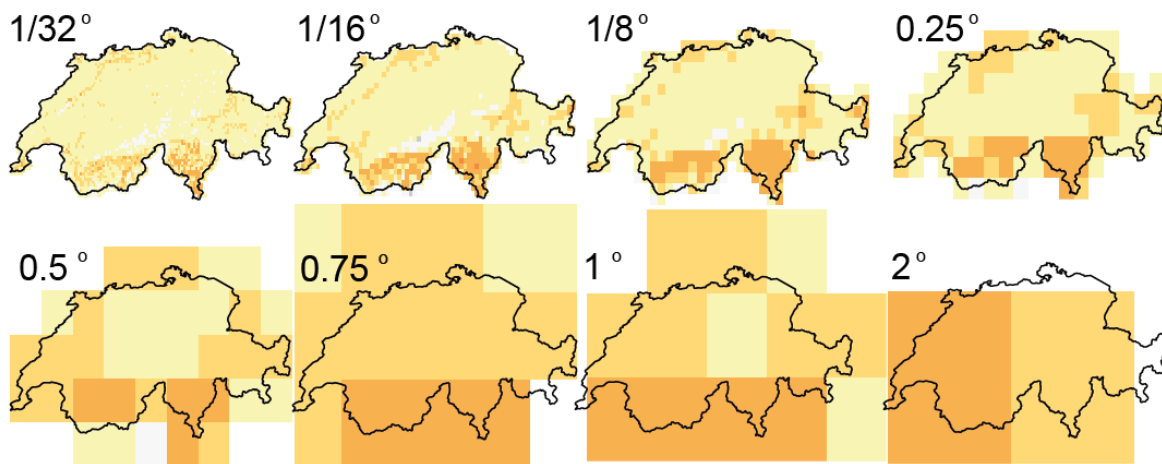
a) True Aggregation Bias (%)

b) Estimated Aggregation Bias (%)

% aggregation bias in ET estimate
(median of daily errors in April-Oct 2004)

Figure 2. a) "True" aggregation bias in ET, as calculated by averaging the 500-m resolution ET estimates using fine-resolution input data in Eq. 6, over 1/32, 1/16, 1/8, 0.25, 0.5, 0.75, 1, and 2-degree grid cells across Switzerland. b) Aggregation bias in ET, as estimated by Eq. 7 from grid-cell averaged temperature (°C), soil moisture (w_w), net radiation (R_n), their variances at each grid scale, and the covariances of all pairs of variables using the 500-m input data. At finer grid scales, the aggregation bias is more localized, and smaller on average. Across Switzerland as a whole, average aggregation bias becomes smaller as grid scales become finer, but never disappears completely.

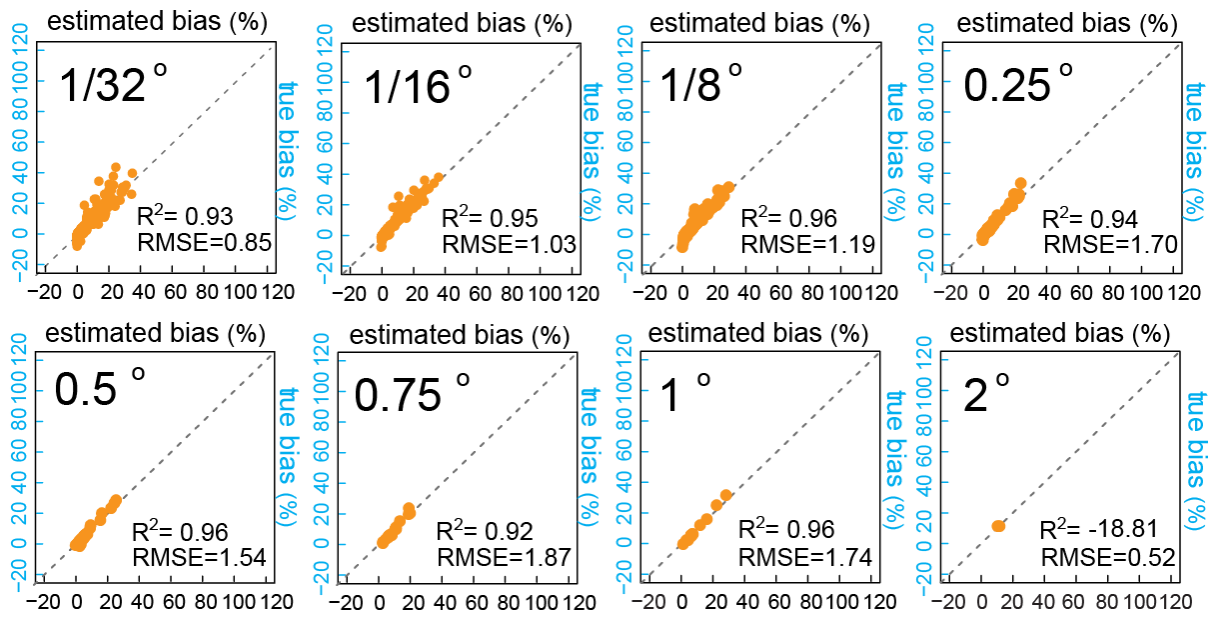



Figure 3. Daily estimated aggregation bias in ET estimates (%, median of daily biases in Apr.-Oct. 2004) versus daily true aggregation bias in ET estimates (%, median of daily biases in Apr.-Oct. 2004) at several spatial scales. Estimated aggregation biases are calculated using Eq. 7. True aggregation biases are calculated as differences between the finer resolution ET estimates from finer resolution input data, averaged over several spatial scales (average of functions) and ET values calculated from average inputs at each spatial scale (function of averages). The coefficients of determination (R^2) between the true and estimated aggregation biases verify the reliability of the Taylor expansion method and Eq. 7 as estimates of the aggregation bias.

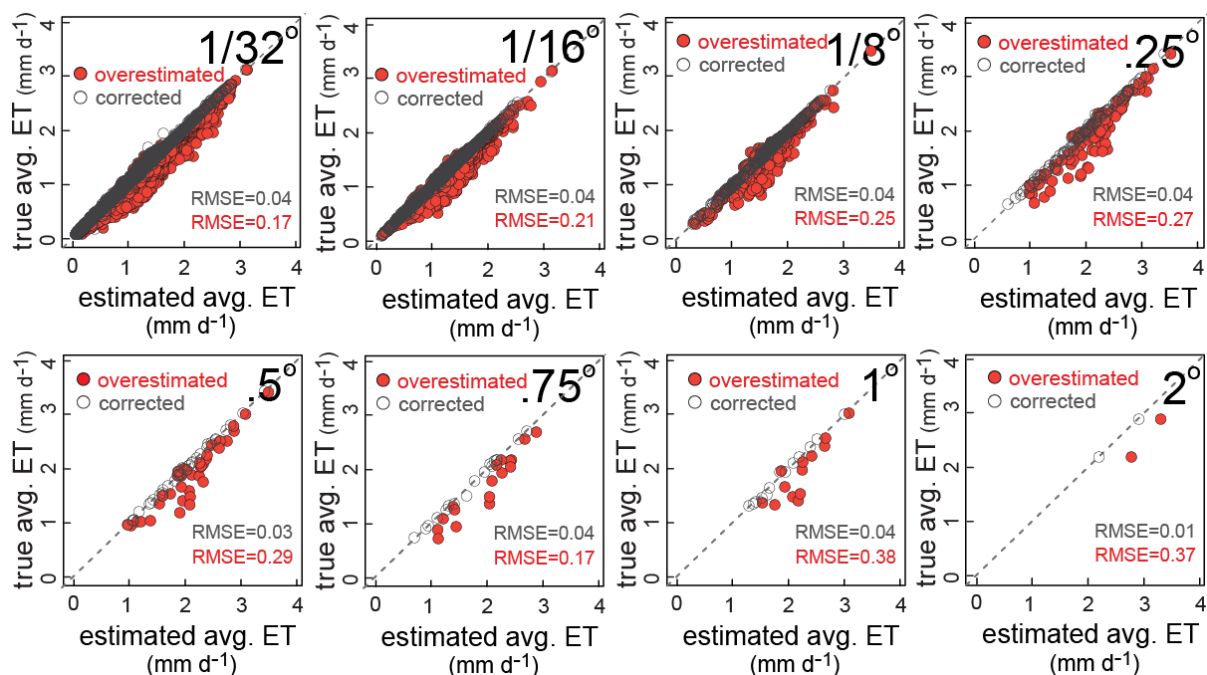



Figure 4. Daily estimated ET rates versus "true" average ET at each grid cell at several different grid scales (example day, May 31st, 2004). The solid red symbols demonstrate the relationships between "true" average ET calculated using fine-resolution data at each grid cell and modeled grid-cell-averaged ET using grid-cell-averaged inputs in Eq.8, for each grid cell at several different grid scales (overestimated). For comparison, the open symbols show true average versus average ET estimated by the Taylor expansion approach of Eq. (7), which corrects for sub-grid heterogeneity effects using only grid-cell-averaged estimates of the ET drivers and their small-scale variances and covariances (heterogeneity-corrected ET estimates, corrected).

3. Discussion

Averaging over spatially heterogeneous ET drivers leads to substantial aggregation biases in ET flux estimates from a typical mechanistic large-scale land surface model. This aggregation bias arises from the inherent nonlinearities in evapotranspiration processes, coupled with the inherent spatial heterogeneity in the driving factors. The joint effects of these nonlinearities and heterogeneities can be estimated using second-order Taylor expansions of the governing equations. Using Switzerland as a test case, we have shown that median aggregation biases of 10-35% are common, even at grid scales substantially smaller than those typically used in land surface models (Fig. 2). These biases can be much larger for individual days (Figs. S2 and S3) and potentially have substantial consequences for water and energy flux estimates in land surface models and consequently for temperature predictions in coupled models. The overestimated evaporative fluxes would lead to overestimated latent heat fluxes and underestimated sensible heat fluxes, and thus potentially to underestimates of expected temperature increases in a changing climate. Unrealistically high evaporation estimates lead to cooler modeled temperatures and wetter modeled climates. Correcting for the aggregation

388 bias in ET fluxes would lead to reduced evaporative cooling and increased atmospheric heating via sensible
389 heat flux.

390

391 In coupled Earth system models, ET fluxes influence how surface temperature, net radiation, and soil moisture
392 evolve through time, and thus influence future values of ET. The analyses shown in Figs. 2-4 are based on static
393 values for each day, and thus do not account for the propagation of aggregation biases forward through time.
394 Estimating the consequences of aggregation biases for dynamic modeling would require fully coupled Earth
395 system model simulations rather than the single ET algorithm analyzed here. In a dynamic model, the Taylor
396 expansion approach can potentially be used to correct for aggregation biases in each time step, using
397 statistical models for the variances and covariances of the ET drivers. Thus, estimating aggregation biases in a
398 dynamic model would not require explicitly simulating sub-grid heterogeneity at every time step. Correcting
399 for aggregation biases at each modeling time step would prevent them from propagating further into future
400 time steps, or into the partitioning of future water and energy fluxes at the land surface. The present paper
401 does not illustrate this dynamic correction for aggregation biases, but establishes the theoretical framework
402 for it.

403

404 The purpose of our analysis was to demonstrate how aggregation bias due to spatial heterogeneity can be
405 quantified (Sects. 2.2-2.3), how its dominant drivers can be identified (Sect. 2.4.1), and how its effects can be
406 efficiently corrected for, using sub-grid closure relationships (Sect. 2.4.2). For this demonstration, we chose
407 GLEAM as an illustrative example, and Switzerland as a topographically complex case study where high-
408 resolution data on the ET drivers are available. Applications of this approach to more complex land surface
409 models may require calculating the necessary derivatives (see Eq. 7) numerically rather than analytically, and
410 applications where high-resolution data are unavailable may require statistically estimating the variances and
411 covariances among the drivers of ET, based on their relationships with topography, soil types, land cover, etc.
412 Using the approach outlined here, one can account for the effects of sub-grid heterogeneity without explicitly
413 modeling ET at fine spatial resolution, which could be impractical due to computational costs, or impossible
414 due to a lack of fine-resolution input data.

415

416 In our analysis, spatial heterogeneity in soil moisture emerged as the dominant driver of aggregation bias in ET
417 estimates. Particularly if this result can also be confirmed in other regions and climates, it points to the
418 importance of improving our understanding of spatial patterns of soil moisture and what controls them. The
419 lower topographic curvature of coarsely gridded landscapes can lead models to predict higher soil moisture at
420 coarser grid scales (Kuo et al., 1999); higher soil moisture at larger grid scales would lead to even higher
421 modeled values of ET, beyond the effects of the aggregation biases analyzed here. Soil moisture may also be
422 substantially influenced by lateral subsurface transfers of water, which are ignored in our analysis and are also

423   ignored by many land surface models. Overlooking lateral transfers could potentially bias ET estimates in large-

424   scale land surface models (Fan et al., 2019), but this is beyond the scope of the present study.

425

426

427   **Acknowledgements**

428   We thank Prof. Ying Fan Reinfelder for numerous insightful discussions and for helpful comments on the

429   manuscript. E.R.F. acknowledges support from the Swiss National Science Foundation (SNSF) under Grant No.

430   P2EZP2_162279.

431   **Data Availability Statement**

432   Temperature data are interpolation of MeteoSwiss data after Viviroli et al., 2009 and are originally from

433   archive data of MeteoSwiss ground level monitoring networks. However, the acquired data may not be used

434   for commercial purposes (e.g. by passing on the data to third parties or by publishing them on the internet). As

435   a consequence, we cannot offer direct access to the data used in this study. Daily soil moisture saturation, net

436   radiation, and temperature data over Switzerland at a 500m resolution for the year 2004 can be retrieved from

437   EnviDat at https://doi.org/10.16904/envidat.176 (Rouholahnejad Freund et al., 2020b).

438   **Author contributions.** ERF and JWK designed the study. ERF ran the analysis. MZ provided the soil moisture

439   simulations and contributed to the discussions. ERF and JWK wrote the paper.

440   **Competing interests.** The authors declare that they have no conflict of interest.

441

442

References

Beven, K. J., and H. L. Cloke: Comment on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al., Water Resour. Res., 48, W01801, https://doi.org/10.1029/2011WR010982, 2012.

Beven, K. J.: The holy grail of scientific hydrology: $Q_t=H(SR)A$ as closure, Hydrol. Earth Syst. Sci., 10, 609–618, https://doi.org/10.5194/hess-10-609-2006, 2006.

Boone, A., and O. J. Wetzel: A simple scheme for modeling sub-grid soil texture variability for use in an atmospheric climate model, Journal of the Meteorological Society of Japan, 77(1), 317–333, https://doi.org/10.2151/jmsj1965.77.1B_317, 1998.

Brunner, M. I., K. Liechti, and M. Zappa: Extremeness of recent drought events in Switzerland: dependence on variable and return period choice, Natural Hazards and Earth System Science, 19(10), 2311–2323, https://doi.org/10.5194/nhess-19-2311-2019, 2019.

Brunt, D.: Physical and dynamical meteorology, 2nd ed., Univ. Press, Cambridge. 428 pp, 1952.

Brutsaert, W.: Evaporation into the atmosphere, ISBN 978-90-481-8365-4, https://doi.org/10.1007/978-94-017-1497-6, 1984.

BFS, Die Bodennutzung der Schweiz: Arealstatistik 1979/85, Bundesamt fuer Statistik, Bern, 1995.

Budyko, M. l.: Climate and life, Academic, New York, 1974.

Bundesamt für Landestopographie: Digitales Höhenmodell RIMINI, Wabern, https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-bundesstatistik/topografie.assetdetail.230215.html, 1991.

El Maayar, M. , and J. M. Chen: Spatial scaling of evapotranspiration as affected by heterogeneities in vegetation, topography, and soil texture, Remote Sensing of Environment, 102, 33–51, https://doi.org/10.1016/j.rse.2006.01.017, 2006.

Ershadi A., M. F. McCabe, J. P. Evans, and J. P. Walker: Effects of spatial aggregation on the multi-scale estimation of evapotranspiration, Remote Sensing of Environment, 131, 51–62, http://dx.doi.org/10.1016/j.rse.2012.12.007, 2013.

Fan, Y., M. Clark, D. M. Lawrence, S. Swenson, L. E. Band, S. L. Brantley, P. D. Brooks, W. E. Dietrich, A. Flores, G. Grant, J. W. Kirchner, D. S. Mackay, J. J. McDonnell, P. C. D. Milly, P. L. Sullivan, C. Tague, H. Ajami, N. Chaney, A. Hartmann, P. Hazenberg, J. McNamara, J. Pelletier, J. Perket, E. Rouholahnejad-Freund, T. Wagener, X. Zeng, E. Beighley, J. Buzan, M. Huang, B. Livneh, B. P. Mohanty, B. Nijssen, M. Safeeq, C. Shen, W. van Verseveld, and J. Volk, D. Yamazaki: Hillslope hydrology in global change research and Earth system modeling, Water Resources Research, 55, https://doi.org/10.1029/2018WR023903, 2019.

Gash, J. H. C.: an analytical model of rainfall interception by forests, Q. J. R. Meteorol. Soc. 105 (433), 43–55, https://doi.org/10.1002/qj.49710544304, 1979.

477 Giorgi, F.: An Approach for the Representation of Surface Heterogeneity in Land Surface Models. Part I:

478 Theoretical Framework, Mon. Wea. Rev., 125, 1885–1899, https://doi.org/10.1175/1520-

479 0493(1997)125<1885:AAFTRO>2.0.CO;2, 1997.

480 Giorgi, F., and R. Avissar: Representation of heterogeneity effects in Earth system modeling: Experience from

481 land surface modeling, Rev. Geophys., 35, 413–437, https://doi.org/10.1029/97RG01754, 1997.

482 Greve, P., B. Orlowsky, B. Mueller, J. Sheffield, M. Reichstein, and S. I. Seneviratne: Global assessment of

483 trends in wetting and drying over land, Nature Geoscience, 7: 716, 2014.

484 Hong, S. H., J. M. H. Hendrickx, and B. Borchers: Up-scaling of SEBAL derived evapotranspiration maps from

485 Landsat (30 m) to MODIS (250 m) scale, Journal of Hydrology, 370, 122–138,

486 https://doi.org/10.1016/j.jhydrol.2009.03.002, 2009.

487 Huang, X., A. M. Rhoades, P. A. Ullrich, and C. M. Zarzycki: An evaluation of the variable-resolution- CESM for

488 modeling California's climate, J. Adv. Model. Earth Syst., 8, 345–369, doi:10.1002/2015MS000559, 2016.

489 Jasechko, S., Z. D. Sharp, J. J. Gibson, S. J. Birks, Y. Yi, and P. J. Fawcett: Terrestrial water fluxes dominated by

490 transpiration, Nature, 496: 347, https://doi.org/10.1890/ES13-00391.1, 2013.

491 Kuo, W. L., T. S. Steenhuis, C. E. McCulloch, C. L. Mohler, D. A. Weinstein, S. D. DeGloria, and D. P. Swaney:

492 Effect of grid size on runoff and soil moisture for a variable-source-area hydrology model, Water Resour. Res.,

493 35(11), 3419–3428, https://doi.org/10.1029/1999WR900183, 1999.

494 Martens, B., D. G. Miralles, H. Lievens, R. van der Schalie, R. A. M. de Jeu, D. Fernández-Prieto, H. E. Beck, W. A.

495 Dorigo, and N. E. C. Verhoest: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci.

496 Model Dev. 10(5): 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

497 McCabe M., and E. Wood: Scale influences on the remote estimation of evapotranspiration using multiple

498 satellite sensors, Remote Sensing of Environment 105, 271–285, https://doi.org/10.1016/j.rse.2006.07.006,

499 2006.

500 Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman: Global land-

501 surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci. 15(2): 453–469,

502 https://doi.org/10.5194/hess-15-453-2011, 2011.

503 Miralles, D. G., A. J. Teuling, C. C. van Heerwaarden, and J. Vilà-Guerau de Arellano: Mega-heatwave

504 temperatures due to combined soil desiccation and atmospheric heat accumulation, Nature Geosci, 7(5):

505 345–349, https://doi.org/10.1038/NGEO2141, 2014.

506 Miralles, D. G., M. J. van den Berg, J. H. Gash, R. M. Parinussa, R. A. M. de Jeu, H. E. Beck, T. R. H. Holmes, C.

507 Jiménez, N. E. C. Verhoest, W. A. Dorigo, A. J. Teuling, and A. Johannes Dolman: El Niño–La Niña cycle and

508 recent trends in continental evaporation, Nature Climate Change, 4: 122,

509 https://doi.org/10.1038/NCLIMATE2068, 2013.

510 Monteith, J. L., and M. H. Unsworth: Principles of Environmental Physics, Edward Arnold, London, 1990.

511   Monteith, J. L.: Evaporation and environment, the state of and movement of water in living organisms,

512   Proceeding of Soc. for Exp. Biol., 19, 205–234, doi:10.1002/qj.49710745102, 1965.

513   Mueller, B., M. Hirschi, C. Jimenez, P. Ciais, P. A. Dirmeyer, A. J. Dolman, J. B. Fisher, M. Jung, F. Ludwig, F.

514   Maignan, D. G. Miralles, M. F. McCabe, M. Reichstein, J. Sheffield, K. Wang, E. F. Wood, Y. Zhang, and S. I.

515   Seneviratne: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis, Hydrol.

516   Earth Syst. Sci. 17(10): 3707–3720, https://doi.org/10.5194/hess-17-3707-2013, 2013.

517   Murray, F. W.: On the computation of saturation vapor pressure, J. Appl. Meteor. 6: 203–204,

518   https://doi.org/10.1175/1520-0450(1967)006<0203:OTCOSV>2.0.CO;2, 1967.

519   Orth, R., M. Staudinger, S. I. Seneviratne, J. Seibert, and M. Zappa: Does model performance improve with

520   complexity? A case study with three hydrological models, Journal of Hydrology, 523, 147–159,

521   https://doi.org/10.1016/j.jhydrol.2015.01.044, 2015.

522   Penman, H. L.: Natural evaporation from open water, bare soil, and grass, Proc. Roy. Soc. London A193,

523   120–146, 1948.

524   Priestley, C. H. B., and R. J. Taylor: On the assessment of surface heat flux and evaporation using large-scale

525   parameters, Monthly Weather Review, 100, 81–92, https://doi.org/10.1175/1520-

526   0493(1972)100<0081:otaosh>2.3.co;2, 1972.

527   Rauscher, S. A., E. Coppola, C. Piani, and F. Giorgi: Resolution effects on regional climate model simulations of

528   seasonal precipitation over Europe, Clim. Dyn., 35(4), 685–711, https://doi.org/10.1007/s00382-009-0607-7,

529   2010.

530   Ringler, T., L. Ju, and M. Gunzburger: A multiresolution method for climate system modeling: Application of

531   spherical centroidal Voronoitessellations, Ocean Dyn., 58(5–6), 475–498, https://doi.org/10.1007/s10236-008-

532   0157-2, 2008.

533   Rouholahnejad Freund, E., and J. W. Kirchner: A Budyko framework for estimating how spatial heterogeneity

534   and lateral moisture redistribution affect average evapotranspiration rates as seen from the atmosphere,

535   Hydrology and Earth System Sciences, 21(1), 217–233, https://doi.org/10.5194/hess-21-217-2017, 2017.

536   Rouholahnejad Freund, E., Y. Fan, and J. W. Kirchner: Global assessment of how averaging over spatial

537   heterogeneity in precipitation and potential evapotranspiration affects modeled evapotranspiration rates,

538   Hydrol. Earth Syst. Sci., 24, 1927–1938, https://doi.org/10.5194/hess-24-1927-2020, 2020a.

539   Rouholahnejad Freund, E., M. Zappa, and J. W. Kirchner: Daily 500m gridded net radiation and soil moisture for

540   Switzerland, 2004. EnviDat. doi: 10.16904/envidat.176., 2020b.

541   Seneviratne, S. I, T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B., Orlowsky, and A. J. Teuling:

542   Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Science Reviews 99(3–

543   4): 125–161, https://doi.org/10.1016/j.earscirev.2010.02.004, 2010.

544 Shrestha, P., M. Sulis, C. Simmer, and S. Kollet: Impacts of grid resolution on surface energy fluxes simulated

545 with an integrated surface-groundwater flow model, Hydrol. Earth Syst. Sci. 19(10): 4317–4326,

546 https://doi.org/10.5194/hess-19-4317-2015, 2015.

547 Shrestha, P., M. Sulis, C. Simmer, and S. Kollet: Effects of horizontal grid resolution on evapotranspiration

548 partitioning using TerrSysMP, Journal of Hydrology, 557: 910–915,

549 https://doi.org/10.1016/j.jhydrol.2018.01.024, 2018.

550 Skamarock, W. C., J. B. Klemp, M. G. Duda, L. D. Fowler, S.-H. Park, and T. D. Ringler, A multiscale

551 nonhydrostatic atmospheric model using centroidal Voronoi tesselations and C-grid staggering, Mon. Weather

552 Rev., 140(9), 3090–3105, https://doi.org/10.1175/MWR-D-11-00215.1, 2012.

553 Speich, M. J. R., L. Bernhard, A. J. Teuling, and M. Zappa: Application of bivariate mapping for hydrological

554 classification and analysis of temporal change and scale effects in Switzerland, Journal of Hydrology, 523, 804–

555 821, https://doi.org/10.1016/j.jhydrol.2015.01.086, 2015.

556 Stanghellini, C., Transpiration of Greenhouse Crops. PhD thesis, Wageningen University, Wageningen, The

557 Netherlands, 1987.

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

559 Viviroli, D., M. Zappa, J. Gurtz, and R. Weingartner: An introduction to the hydrological modelling system

560 PREVAH and its pre- and post-processing-tools, Environmental Modelling and Software, 24(10), 1209–1222,

561 https://doi.org/10.1016/j.envsoft.2009.04.001, 2009.

562 Zappa, M. and J. Gurtz: Simulation of soil moisture and evapotranspiration in a soil profile during the 1999

563 MAP-Riviera Campaign, Hydrol. Earth Syst. Sci., 7, 903–919, https://doi.org/10.5194/hess-7-903-2003, 2003.

564 Zarzycki, C. M., M. N. Levy, C. Jablonowski, J. R. Overfelt, M. A. Taylor, and P. A. Ullrich: Aquaplanet

565 experiments using CAM's variable-resolution dynamical core, J. Clim., 27(14), 5481–5503,

566 https://doi.org/10.1175/JCLI-D-14-00004.1, 2014.

567

568