

Contents:

S1. Performance of the Candidate Machine Learning Models and Predictor Variable Importance Rankings

S2. Site Characteristics and Meteorological Conditions During the Study Period

S3. Validation of Measurement-based T:ET at Sibbald, Burstall and Bonsai

S1. Performance of the Candidate Machine Learning Models and Predictor Variable Importance Rankings

The machine learning framework developed by Stapleton et al. (2022) was adopted with only minor modifications. The original workflow, data imputation, feature selection procedure, predictor variable importance ranking, and model ranking were retained to ensure consistency with the published framework. Only two minor modifications were made to accommodate differences in the available datasets. First, because winter observations were unavailable, the winter validation step described by Stapleton et al. (2022) was omitted. Second, the F25 predictor set was identical to that of Stapleton et al. (2022), except that the vegetation greenness index (GCC), mean water temperature (TW_mean), ecosystem respiration derived from the linear method (er_linear), speed of sound (sos), degree of instationarity for mean methane concentration (stat_m), and methane flux (wm_gf) were excluded because these variables were unavailable for the present study. All other predictor variables, machine learning algorithms, training, calibration, validation procedures, and model selection criteria were identical to those described by Stapleton et al. (2022). The results of the model comparison, recursive feature elimination (RFE) for predictor selection, feature importance rankings, and the best-performing machine learning model for each study site are presented in Figs. S1 and S2 and Tables S1 and S2.

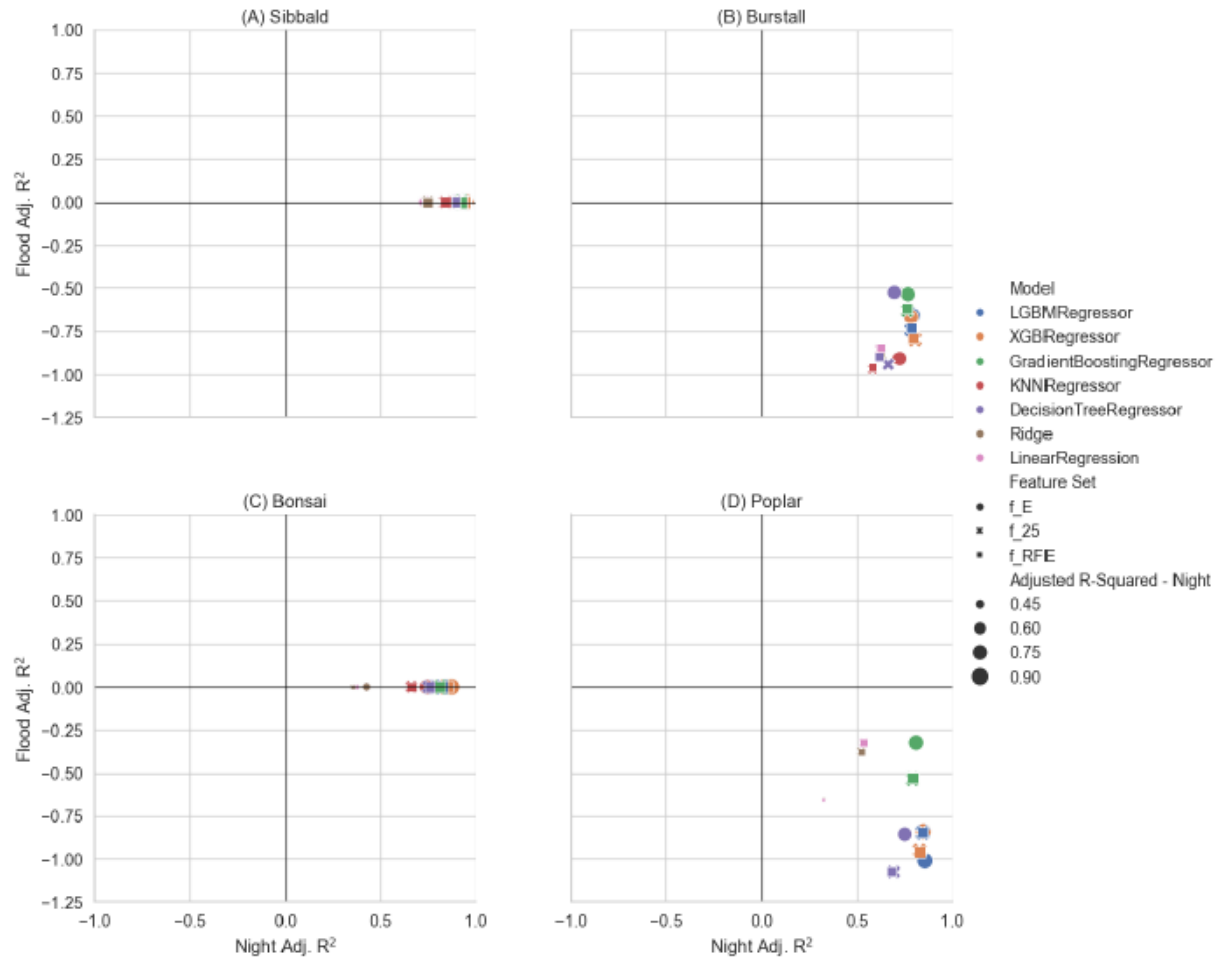


Figure S1: Results of model comparison using the adjusted R^2 defined in Stapleton et al. (2022). The x-axis shows the adjusted R^2 (Adj. R^2) for predictions under non-flooded (night-time validation) conditions, while the y-axis shows the adjusted R^2 for predictions during flooded conditions. Because flooding did not occur at Sibbald and Bonsai during the study period, their candidate models are plotted along the x-axis only. Marker colour indicates the machine learning algorithm, whereas marker shape denotes the predictor variable set used for model training. Models located farther to the right and higher on the plot exhibit better predictive performance. Negative adjusted R^2 values indicate poor generalization of the candidate models to the flooding validation dataset.

Table S1. Feature importance ranked in order of importance for each site where the features obtained by the Recursive Feature Elimination process are denoted by F followed by the site label and the relative importance of that feature at that site is given by I followed by the site label. The notation for the predictor variables follows Stapleton et al. (2022). Detailed definitions of the predictor variables can be found therein.

Rank	F (Sibbald)	I (Sibbald)	F (Burstall)	I (Burstall)	F (Bonsai)	I (Bonsai)	F (Poplar)	I (Poplar)
1	H_gf	0.629	H_gf	0.555	DOY	0.514	RH	0.295
2	TA	0.132	VPD	0.103	H_gf	0.295	ww	0.116
3	DOY	0.084	RH	0.042	TA	0.038	RNET	0.092
4	RNET	0.078	RNET	0.039	RNET	0.036	H_gf	0.089
5	TS_mean	0.026	TA	0.038	RH	0.030	TA	0.082
6	er_Reichstein	0.018	PA_Met	0.033	VPD	0.021	vv	0.041
7	RH	0.017	ustar	0.026	TS_mean	0.018	DOY	0.041
8	VPD	0.004	ze	0.024	er_Reichstein	0.015	VPD	0.034
9	ze	0.003	rho_mol_Met	0.018	time	0.009	tsbar	0.032
10	uu	0.002	WD	0.018	ze	0.009	tbar	0.030
11	stat_wt	0.002	uu	0.014	ustar	0.004	er_Reichstein	0.028
12	uw	0.002	ubar	0.013	WD	0.003	ustar	0.020
13	stat_wm	0.001	TS_mean	0.012	cbar	0.002	TS_mean	0.015
14	tbar	0.001	er_Reichstein	0.011	PA_Met	0.002	ze	0.013
15	ustar	0.001	tbar	0.011	uw	0.002	cbar	0.012
16	time	0.001	stat_wc	0.006			WD	0.010
17			DOY	0.006			rho_mol_Met	0.009
18			cbar	0.004			uu	0.009
19			vv	0.004			PA_Met	0.009
20			stat_wt	0.004			ubar	0.007
21			time	0.003			stat_q	0.004
22			uw	0.003			stat_wm	0.004
23			stat_q	0.003			time	0.003
24			tsbar	0.003			stat_wc	0.001
25			WT	0.002				
26			ww	0.002				

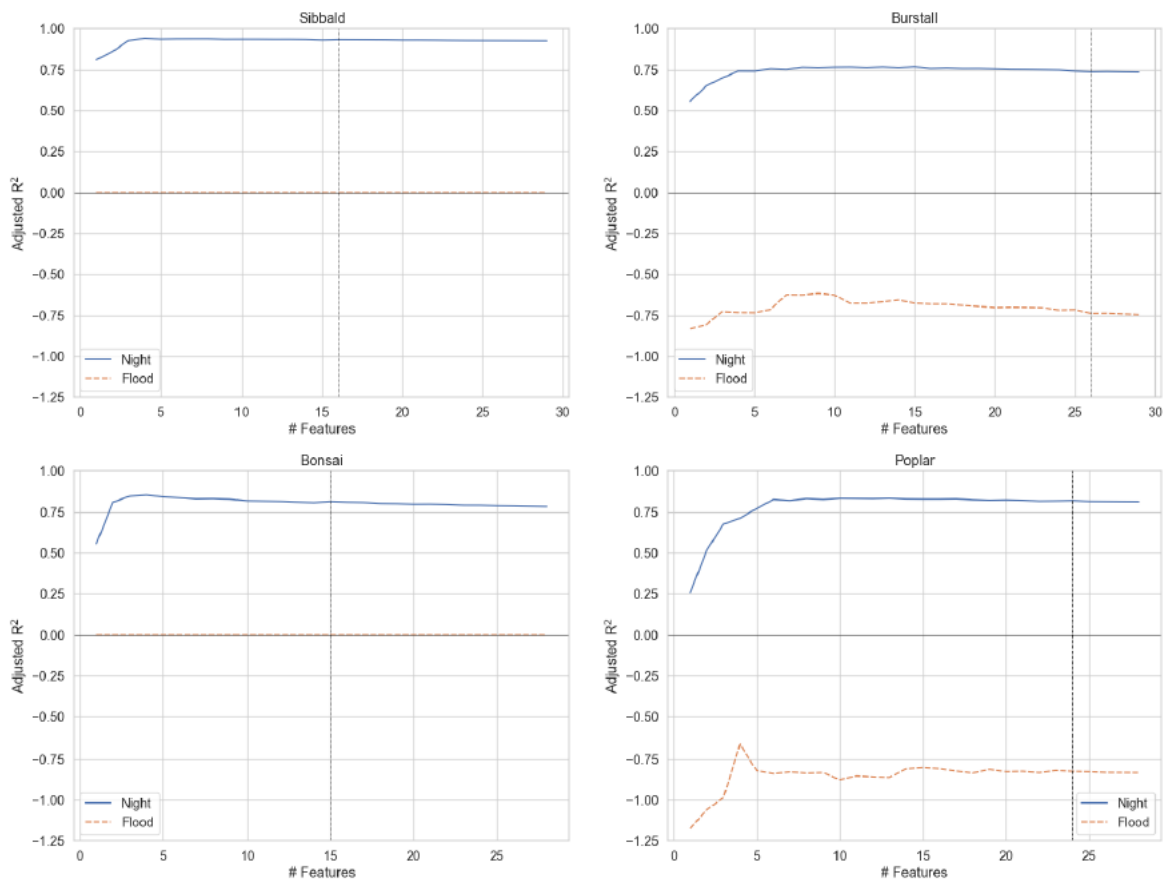


Figure S2: Results of the Recursive Feature Elimination process for each of the 4 sites tested with number of features on the x-axis and adjusted R² results on the y-axis. The iterations start on the right and move towards 0 as Recursive Feature Elimination iteratively decreases the number of features until only 1 feature remains for each of the sites and each of the test sets; Night and Flood. A vertical line on each graph indicates the number of features selected, where the optimal feature set is determined to be the last feature set preceding a 0.1% reduction in adjusted R².

Table S2. Best-performing machine learning algorithm based on adjusted R^2 from the night time dataset, corresponding predictor feature set, and performance metrics for each of the four study sites

Site	Best ML Model Structure	Best Feature Set	R^2	Adjusted R^2	RMSE	Slope	Number of Features
Sibbald	LightGBM	F_e	0.94638	0.94590	0.00356	0.94304	10
Burstall	XGBoost	F_25	0.81000	0.80446	0.00609	0.81747	29
Bonsai	XGBoost	F_e	0.87160	0.86993	0.00529	0.87158	9
Poplar	LightGBM	F_e	0.85791	0.85689	0.00407	0.85206	10

Note. The F_e feature set includes 10 variables: VPD, ustar, TA, WT, RNET, H_gf, er_Reichstein, year, DOY, time. Bonsai has 9 features because WT is not available for this site. The F_25 set includes 29 variables, and they are: DOY, H_gf, PA_Met, RH, RNET, TA, TA_Met, TS_mean, VPD, WD, WT, cbar, er_Reichstein, rho_mol_Met, stat_q, stat_wc, stat_wm, stat_wt, tbar, tie, tsbar, ubar, ustar, uu, uw, vv, ww, year, and ze. The predictor variable symbols are identical to those used by Stapleton et al. (2022); readers are referred to that study for their definitions.

S2. Site Characteristics and Meteorological Conditions During the Study Period

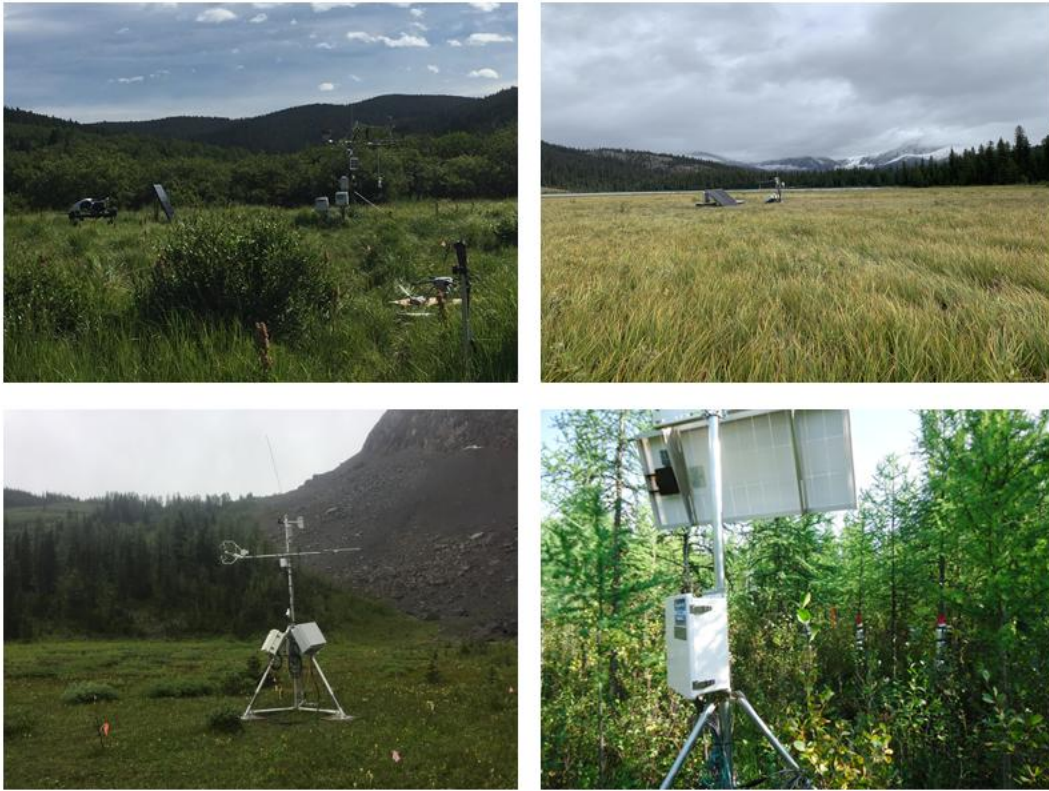


Figure S3: Photographs showing canopy conditions at the four study sites. The upper left image shows Sibbald, characterized by layered canopy with scattered shrubs and dense sedges. The upper right is Burstall, featuring very sparse shrubs and dense sedges. The bottom left is Bonsai, presenting clusters of trees and mixed understory vegetation. The bottom right is Poplar, which has a denser, more closed canopy formed by coniferous trees.

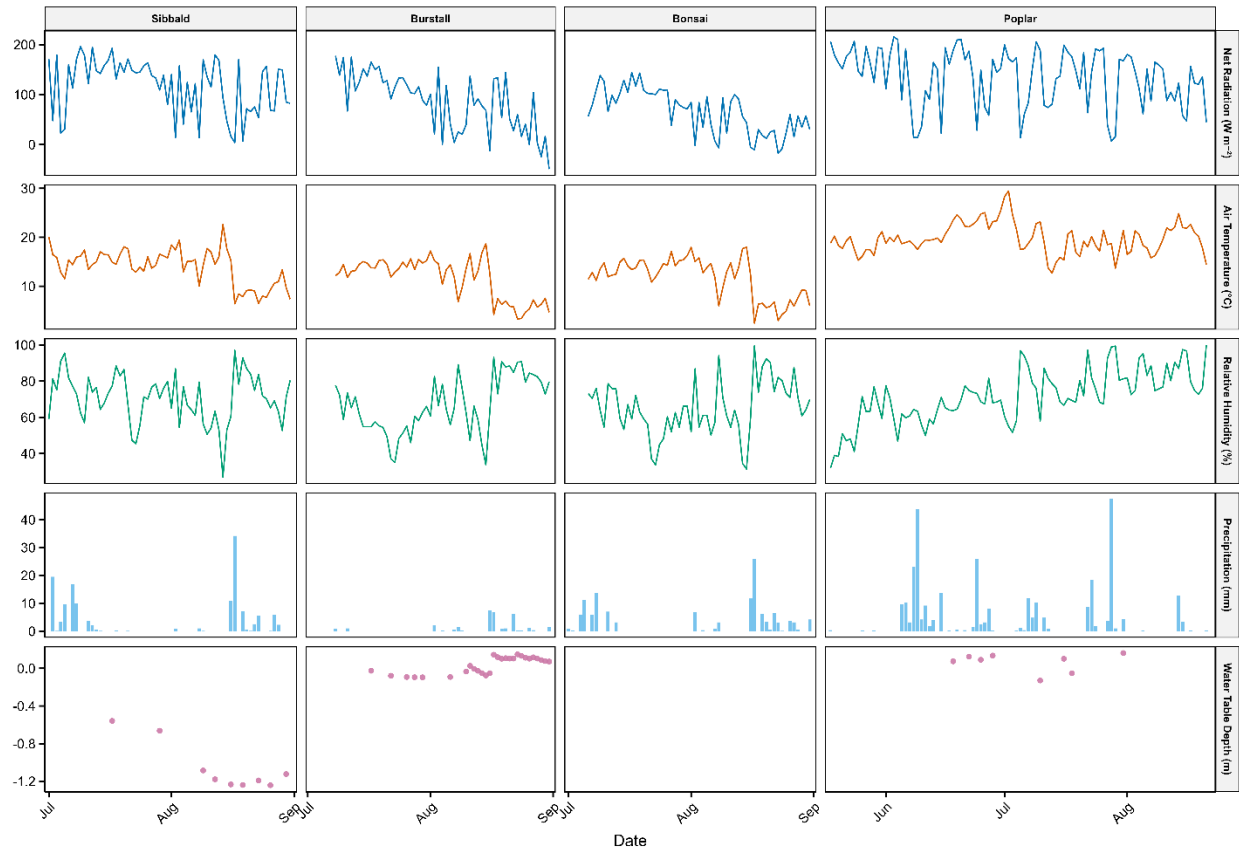


Figure S4: Meteorological conditions, including net radiation, air temperature, relative humidity, precipitation, and water table depth, at the four study sites. Water table depth at Bonsai exceeded 2 m below the ground surface throughout the study period and therefore could not be measured. Positive water table depths indicate flooded conditions (i.e., the water table was above the ground surface).

S3. Validation of Measurement-based T:ET at Sibbald, Burstall and Bonsai

Detailed description of the design of the controlled evaporation experiments and supplementary microclimatic measurements has been provided elsewhere (Wang et al., 2023). Briefly, the controlled evaporation experiments involved various ground cover combinations, such as bare soil (S), bryophyte-covered soil (BS), litter and bryophytes covered soil (LBS), and soil covered by ground vascular vegetation, litter, and bryophytes (PLBS). The PLBS treatment was applied only at Bonsai, where the ground vascular vegetation was the understory short vascular plants. Additionally, two levels of canopy treatment were applied: no canopy (NC) and below the canopy of willow trees (BC). The experiments were carried out during July and August of 2021, during which each site underwent 6 measurement rounds, with each round lasting 2 to 4 days (Table S3). Within each round, there were three replicates for every treatment level. At each site, one of the six rounds is unique because ground surface evaporation was measured both in the middle and at the end of that round. As a result, this round contains two evaporation measurements. Specifically, these measurements are labeled as round 1.1 and 1.2 at Sibbald, and round 3.1 and 3.2 at both Burstall and Bonsai.

The detailed calibration of the S–W model for each study site, together with the development of a surface resistance formulation that explicitly accounts for the effects of moss and litter on ground evaporation and its incorporation into the S–W model, are described in previous publications (Wang et al., 2023; Wang, 2025). The errors and uncertainties of the measurements of E, T and ET are summarized in Table S4.

Partitioning results were validated by comparing latent heat (λET) from the modified S-W model with eddy covariance measurements at the daily scale and by comparing modelled λE (modelled evaporation converted to its energy equivalent) with microlysimeter measurements across all 6 rounds of evaporation experiments. As shown in Fig. S5, the modelled λET agrees well with the measured λET , especially with the energy balance-corrected λET , as the slopes of fitted linear regressions range from 0.87 to 1.16, intercepts range from -18.50 to 3.78 W m⁻², and R² values are all greater than 0.78. Better agreement with the energy balance corrected fluxes is anticipated, as the S-W model is based on the energy balance closure.

In Fig. S5, the purple lines linking modelled ground surface evaporation (purple triangles) in the right panel illustrate the temporal variation of the modelled evaporation across 6 rounds of experiments. In these plots, the range and the magnitude of measured evaporation under various ground covers and overstory canopy conditions for each round of evaporation experiment are represented by the black range bars and the colorful dots. Clearly, the modelled ground surface evaporation at the three sites is slightly lower than the lysimeter-based measurements but it is considered a reasonable representation of actual ground surface evaporation. This is because the three wetlands had sedges and other tall vascular plants present, which might exert shading effects that led to lower ground surface evaporation. However, these plants were removed in the controlled evaporation experiments; therefore, the lysimeter-based ground surface evaporation would have been greater than the actual ground surface evaporation. Notably, Sibbald has noticeably denser growth of sedges, as shown in the field photo (Fig. S3). This could explain the larger

intercept of the fitted linear regression between modelled and measured λET , as the actual light availability below the hypothetical canopy at Sibbald may be less than that in the S-W model's parameterization.

Overall, the modelled results show satisfactory consistency with the measurements in terms of both magnitude and temporal dynamics, which substantiates the appropriateness of the modified S-W model for accounting for the impact of bryophytes and litter on ground surface evaporation and evapotranspiration partitioning.

Table S3. The date and duration of evaporation experiments at the three sites.

Site	Round	Start Date	Duration (days)
Sibbald	1-1*	7/12/2021	2
	1-2*	7/14/2021	3
	2	7/29/2021	5
	3	8/5/2021	4
	4	8/12/2021	4
	5	8/19/2021	4
	6	8/26/2021	4
Burstall	1	7/14/2021	3
	2	7/22/2021	4
	3-1*	7/26/2021	2
	3-2*	7/28/2021	2
	4	8/6/2021	4
	5	8/25/2021	2
	6	8/27/2021	4
Bonsai	1	7/9/2021	4
	2	7/15/2021	4
	3-1*	7/19/2021	2
	3-2*	7/21/2021	2
	4	8/13/2021	5
	5	8/20/2021	4
	6	8/25/2021	2

Note. The symbol * indicates the round of the evaporation experiments during which the microlysimeters were weighed midway through the experimental period, but no bryophyte or litter biomass was harvested.

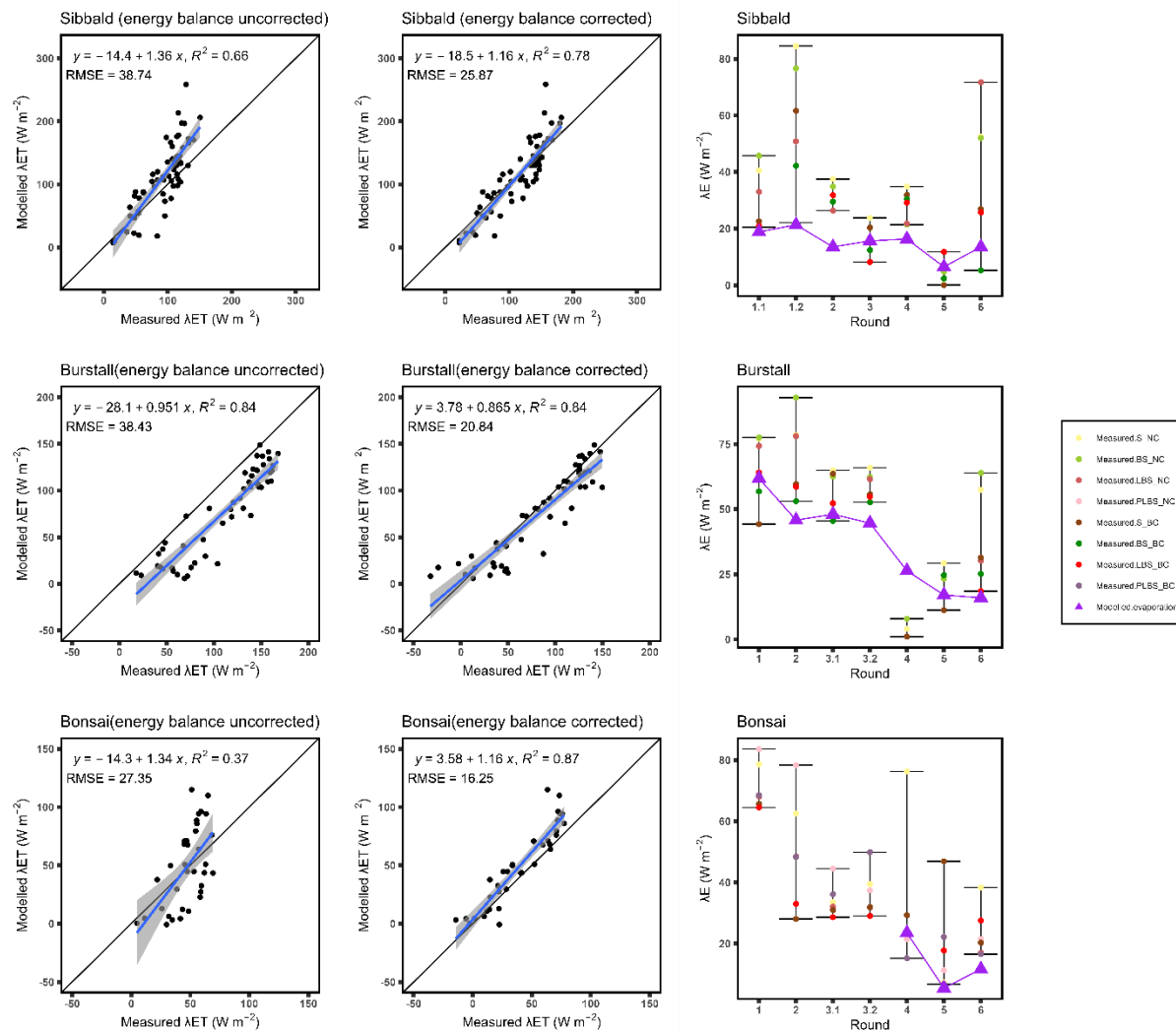


Figure S5: Comparisons between the modelled and measured latent heat (λET) during the study period (left and middle panels), and between the modelled and measured ground evaporation (λE) across all 6 rounds of evaporation experiments (right panel).

Table S4. Examples of Errors and Uncertainties of Flux Measurements

Flux contributions	Method	Source of errors	Normalized Root Mean Square Error (%)	Reference
ET	EC	One-point Uncertainty	8.3	Kroon et al. 210
T	Sap flow system (stem heat balance method)	System measurement resolution, Steady-state assumption, Heat-storage error	15.0	Groot and King, 1992 Shackel et al., 1992 Grime et al., 1995 Perämäki et al., 2001
ET (ground)	Chamber	Lower Q*, Altered Microclimate	6.6	Reicosky et al., 1983 McLeod et al., 2004 Hamel et al., 2015
E	Micro-lysimeter	Soil disturbances, Handling and weighing	17.8	Calculated using micro-lysimeter measurements from this study

(Note: Table adapted from Gabrielli 2016).

References:

- Gabrielli, E. C.: Partitioning Evapotranspiration in Forested Peatlands within the Western Boreal Plain, Fort McMurray, Alberta, Canada, Master's thesis, Wilfrid Laurier University, 2016.
- Grime, V.L., Morrison, J.I.L., Simmonds, L.P. 1995. Including heat storage term in sap flow measurements with the heat storage method. *Agricultural and Forest Meteorology*, 74: 1-25. [https://doi.org/10.1016/0168-1923\(94\)02187-O](https://doi.org/10.1016/0168-1923(94)02187-O).
- Groot, A., King, K.M. 1992. Measurement of sap flow by heat balance method: numerical analysis and application to coniferous seedlings. *Agricultural and Forest Meteorology*, 59: 289-308. [https://doi.org/10.1016/0168-1923\(92\)90098-O](https://doi.org/10.1016/0168-1923(92)90098-O).
- Hamel, P., Mchugh, I., Coutts, A., Daly, E., Beringer, J., Fletcher, T.D. 2015. Automated Chamber System to Measure Field Evapotranspiration Rates. *Journal of Hydrologic Engineering*, 20(2): 04014037. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001006](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001006).
- Kroon, P.S., Hensen, A., Jonker, H.J.J., Ouwersloot, H.G., Vermeulen, A.T., Bosveld, F.C. 2010. Uncertainties in eddy covariance flux measurements assessed from CH₄ and N₂O observations. *Agricultural and Forest Meteorology*, 150: 806-816. <https://doi.org/10.1016/j.agrformet.2009.08.008>.
- McLeod, M.K., Daniel, H., Faulkner, R. Murison, R. 2004. Evaluation of an enclosed portable chamber to measure crop and pasture actual evapotranspiration at small scale. *Agricultural Water Management*, 67: 15-34. <https://doi.org/10.1016/j.agwat.2003.12.006>.
- Perämäki, M., Vesala, T., Nikinmaa, E. 2001. Analyzing the applicability of the heat balance method for estimating sap flow in boreal forest conditions. *Boreal Environment Research*, 6: 29-43. <https://doi.org/10.60910/4tn6-73en>.
- Reicosky, D.C., Sharratt, B.S., Ljungkull, J.E., Baker, D.G. 1982. Comparison of Alfalfa Evapotranspiration Measured By A Weighing Lysimeter And A Portable Chamber. *Agricultural Meteorology*, 28: 205-211. [https://doi.org/10.1016/0002-1571\(83\)90026-2](https://doi.org/10.1016/0002-1571(83)90026-2).
- Stapleton, A., Eichelmann, E., and Roantree, M.: A framework for constructing machine learning models with feature set optimisation for evapotranspiration partitioning, *Applied Computing and Geosciences*, 16, 100105, <https://doi.org/10.1016/j.acags.2022.100105>.
- Wang, Y., Petrone, R. M., and Van Huizen, B.: The dependence of evaporative efficiency of vegetated surfaces on ground cover mass fractions in vegetated soils in mesic ecosystems, *Hydrological Processes*, 37, e15036, <https://doi.org/10.1002/hyp.15036>.
- Wang, Y.: Uncovering the understudied role of microtopography and ground cover on evapotranspiration partitioning in high-elevation wetlands in the Canadian Rocky Mountains, Doctoral dissertation, University of Waterloo, 2025.