Supplemental information for “Soil drying with experimental warming depends on ecosystem type and warming method: First results of the Soil Warming to Depth Data Integration Effort (SWEDDIE)”

by J. Beem-Miller et al.

Methods

Database

Experiment

The *experiment* table provides information about the physical infrastructure installed in each experiment. This includes the method of warming (infrared, cables, rods, open topped chambers, etc.), sensors installed, data types collected, beginning and end dates of warming, and additional treatments imposed (elevated CO_2, precipitation manipulation, nutrient additions, etc.). The ‘exp_name’ variable provides the key linkage between this and other core tables.

Site

The *site* table provides information about the climate, soils, and vegetation for each experiment. Experiments may have one or more sites, each of which is uniquely defined by the combination of the ‘exp_name’ and ‘sit_name’ variables.

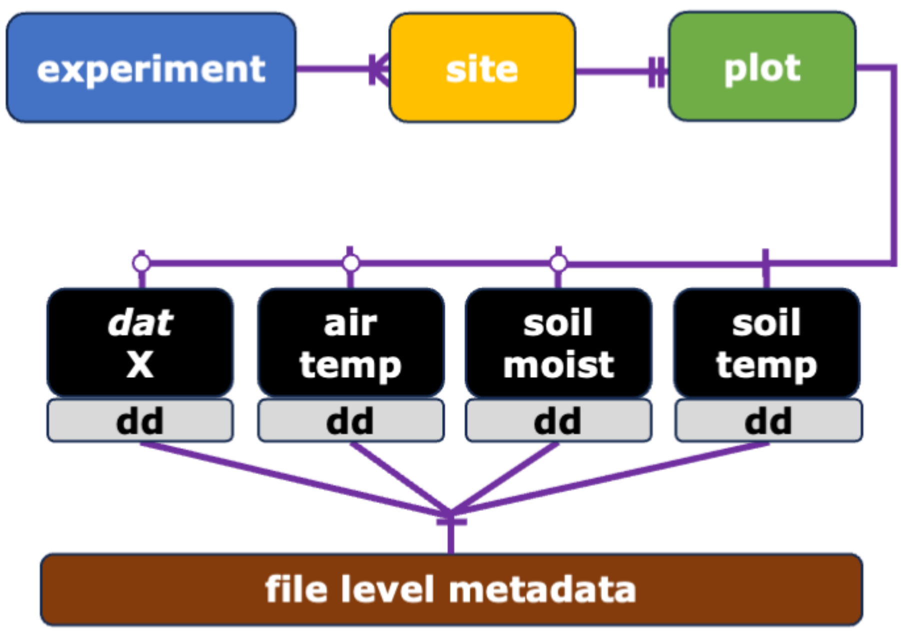
Plot

The *plot* table defines the experimental design, and each record (row) of this table is defined by a unique combination of the ‘exp_name’, ‘sit_name’, and ‘plt_name’ variables. The *plot* table serves to link imposed treatments with measured data,. therefore the values for ‘plt_name’ in all *dat* tables must match with those recorded in the *plot* table. In order to accommodate the diversity of experimental design choices across the network, as well as reflect the naming conventions employed within individual experiments, this table contains additional columns for blocks and sub-blocks as applicable. Imposed heating is encoded in the ‘plt_heat_level’ variable, defined as the difference (in °C) from the control. Additional treatments are encoded in the variables ‘plt_treat_add_name’, ‘plt_treat_add_level’, and ‘plt_treat_add_units’, which define the treatment name, treatment level, and units respectively. Plots with multiple treatments are accommodated through semi-colon separated lists. Ex., for a plot with elevated CO_2 at 100 ppm above ambient and 50 percent reduced precipitation, the values in the 'plt_treat_add_name' column would be "elevated CO2; reduced precipitation", the values in the 'plt_treat_add_level' column would be "100; 50", and the values in the 'plt_treat_add_units' column would be "ppm above ambient; percent of ambient".

Queries and reports

Data are retrieved from the database by creating a query that yields a custom report. A report can include all available data or a specific subset. The query specifies the desired target variables and optional filters for depth, temporal range, imposed treatments, site level characteristics, and experiments. Queries must be generated by the user, although an interactive R tool in the *sweddie* package is currently under development. The typical report output in the R environment is in the format of a hierarchical R list object mirroring the data model (Fig. SI 1), i.e., experiment level data linked to site level data linked to plot data linked to observed data. The R list object can be further flattened and exported as a CSV file by the user if desired.

Fig. SI 1 Entity relationship diagram for a single *entity* (experiment) in SWEDDIE.



Models

Eq. 1 Generalized likelihood formula for Bayesian linear mixed effect models

$$y_{ij} \mid u_{ij}, \sigma_j \sim N\left(u_{ij}, \sigma_j^2\right)$$

where y_{ij} is the *i*-th response of soil moisture difference or soil temperature difference at site *j* given the fixed effects u_{ij}, and site-specific residual variance σ_j. We assumed Gaussian normal distributions for both fixed effects and the standard deviation of the site-specific residual variances (σ_j^2). In both models we included a temporal autocorrelation structure to account for memory effects in the residuals at a daily, i.e., first order, timestep. The autocorrelation was assessed at the level of site, warming treatment, and depth. We also included an additional cyclical spline function term in the ΔSM model to account for site-specific seasonal trends in the response of ΔSM (SI Fig. 3).

Eq. 2 Soil temperature model fixed effect matrix u_{ij}

$$\Delta ST_{ij} = \beta_0 + \beta_H \cdot H_i + \beta_{depth} \cdot D_i + \sum_w \beta_{wrm_w} \cdot wrm_{iw} + \sum_w \beta_{wrmxD_w} \cdot wrm_{iw} \cdot D_i$$

Where ΔST_{ij} is the ith observation of observed temperature difference (treatment - control) at site j, H_i is the warming target (in degrees of heating), and D_i is depth. The coefficients for the categorical variables of warming method *wrm* are given as sums of the product of the coefficients and a dummy variable for each level *w* of *wrm* (above ground, below ground, above + below ground). The coefficients for the interaction terms of wrm by D are given as $wrmxD_w$.

Eq. 3 Soil moisture model fixed effects matrix u_{ij}

$$\Delta SM_{ij} = \beta_0 + \beta_{ST} \cdot \Delta ST_i + \beta_{ambSM} \cdot ambSM_i + \beta_{ambST} \cdot ambST_i + \beta_{depth} \cdot D_i + \beta_{ambSMxST} \cdot \Delta ST_i \cdot ambSM_i + \beta_{ambSTxST} \cdot \Delta ST_i \cdot ambST_i + \sum_e \beta_{eco_e} \cdot eco_{ie} + \sum_g \beta_{growS_g} \cdot growS_{ig} + \sum_e \beta_{ecoxST_e} \cdot \Delta ST_i \cdot eco_{ie}$$

Where ΔSM_{ij} is the i-th difference (treatment - control) in volumetric soil water (VWC) content at site j, and ambSM and ambST are the ith observation of observed temperature difference , mean ambient soil temperature, mean ambient soil moisture**,** and depth, respectively; β values are the corresponding model coefficients. The coefficients for the categorical variables of ecosystem (eco) and growing season (growS) are given as sums of the product of the coefficients and a dummy variable for each level *e* of eco (boreal forest, temperate forest, tropical forest, tundra, grassland, cropland) and *g* of growS (growing season/non-growing season). The coefficients for the interaction terms of ambSM by ΔST, and ambST by ΔST, are denoted with the subscripts ambSMxST and ambSTxST, while those for each level *e* of eco are denoted with the subscript ecoxST.

Table SI 1 Experiment locations

Experiment	Country	Lat.	Long.	Elev.	MAT	MAP	PET	Ecosystem	Soil Order
		dec. deg.		*m*	*°C*	*mm yr-1*			*USDA*
185 Experiment Station	China	37.88	114.68	50	14.8	741	1468	cropland	Inceptisols
ACBB	France	46.42	0.12	146	12	760	1025	grassland; cropland	Alfisols
Achenkirch	Austria	47.58	11.64	910	6.8	1563	883	temperate forest	Alfisols
B4WarmED	USA	47.03	-91.92	382	3.0; 4.5	807	870; 970	temperate forest; boreal forest	Spodisols; Alfisols
Blodgett	USA	38.91	-120.66	1370	12.5	1774	1681	temperate forest	Alfisols
CiPEHR	USA	63.88	-149.23	700	-1	378	592	tundra	Inceptisols
ForHot	Iceland	64.01	-21.18	126	5.2	1457	564	boreal forest	Inceptisols
FutureClim	Australia	-36.37	148.43	1600	6.3	2144	1111	grassland	Inceptisols
GENX	USA	38.88	-76.55	0	13	1130	1359	salt marsh	Ultisols
Harvard Forest	USA	42.54	-72.18	340	8.2	1141	1007	temperate forest	Inceptisols
Lyon Arboretum	USA	21.33	-157.8	150	25.6	4200	1891	tropical forest	Andisols
KAEFS	USA	34.98	-97.52	330	16.3	914	1773	grassland	Inceptisols
MERIT	Germany	54.6	8.82	0	10.2	1048	679	salt marsh	Entisols
Point Reyes	USA	37.97	-122.73	125	12.5	1774	1384	grassland	Alfisols
Sanming	China	26.17	117.47	250	19.3	1638	1355	subtropical forest	Ultisols
SMARTX	USA	38.88	-76.55	0	13	1130	1359	salt marsh	Ultisols
SPRUCE	USA	47.51	-93.45	418	3.3	768	921	peatland	Alfisols
Stillberg	Switzerland	46.77	9.87	2180	2.1	1155	716	temperate forest	Inceptisols
SWAMP	China	37.5	100.24	3670	-0.1	365	1037	tundra	Mollisols
SWELTR	Panama	9.16	-79.85	60	26	2600	1679	tropical forest	Ultisols
TEAM	China	37.62	101.2	3200	-0.6	460	1090	tundra	Gelisols
TeRaCON	USA	45.4	-93.2	282	7	660	1044	grassland	Entisols

TRACE	USA	18.3	-65.83	100	24	3500	1575	tropical forest	Ultisols

Table SI 2 Residual warming model summary

Family: gaussian
Links: mu = identity; sigma = identity
Formula: stR ~ tDif + wrm2 * depth + (1 | sit_name)
Data: df (Number of observations: 240287)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
total post-warmup draws = 4000

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

Multilevel Hyperparameters:
~sit_name (Number of levels: 23)

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sd(Intercept)	0.76	0.14	0.55	1.08	1.00	846	1323

Regression Coefficients:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	0.19	0.39	-0.58	0.96	1.00	1337	1522
tDif	-0.38	0.00	-0.38	-0.38	1.00	3459	2367
wrm2aboveDbelow	0.04	0.47	-0.93	0.98	1.00	1228	1249
wrm2below	0.79	0.45	-0.11	1.70	1.00	1208	1389
depth	0.00	0.00	0.00	0.00	1.00	3841	26.87
wrm2aboveDbelow:depth	0.01	0.00	0.01	0.01	1.00	3853	2710
wrm2below:depth	0.01	0.00	0.01	0.01	1.00	3822	2595

Further Distributional Parameters:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	1.37	0.00	1.37	1.38	1.00	3486	2146

SI Table 3 Soil moisture model summary

Family: gaussian
Links: mu = identity; sigma = identity
Formula: sm_mean ~ st_mean * eco + st_mean * control_mean.sm + st_mean * control_mean.st + depth + s(doy, sit_name, bs = "fs") + (1 + st_mean | sit_name)
autocor ~ arma(time = day, gr = sit_name:tDif:depth, p = 1, q = 0, cov = FALSE)

Data: df (Number of observations: 61315)
Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
total post-warmup draws = 4000

Smoothing Spline Hyperparameters:

	Estimate	Est.Error	l-95%CI	u-95%CI	Rhat	Bulk_ESS	Tail_ESS
sds(sdoysit_name_1)	20.23	3.19	14.47	26.81	1.01	1205	2232
sds(sdoysit_name_2)	5.28	4.96	0.23	18.49	1.00	3766	2548
sds(sdoysit_name_3)	16.53	4.14	9.86	25.98	1.00		

Correlation Structures:

	Estimate	Est.Error	l-95%CI	u-95%CI	Rhat	Bulk_ESS	Tail_ESS
ar[1]	0.99	0.00	0.99	0.99	1.00	3719	2938

Multilevel Hyperparameters:
~sit_name (Number of levels: 19)

	Estimate	Est.Error	l-95%CI	u-95%CI	Rhat	Bulk_ESS	Tail_ESS
sd(Intercept)	6.71	1.34	4.61	9.83	1.00	2313	2582
sd(st_mean)	0.48	0.10	0.31	0.73	1.00	1929	2697
cor(Intercept,st_mean	0.28	0.25	-0.27	0.71	1.00	1471	2093

Regression Coefficients:

	Estimate	Est.Error	l-95%CI	u-95%CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	1.34	2.38	-3.44	6.06	1	1233	2058
st_mean	-0.42	0.1	-0.61	-0.23	1	2376	2618
ecocropland	-2.89	3.21	-9.26	3.4	1	2191	2850
ecograssland	-3.62	2.71	-8.85	1.81	1	2199	2961
ecosubtropicalforest	-4.18	2.31	-8.69	0.41	1	2670	3219
ecotemperateforest	-4.48	4.22	-12.58	3.82	1	1916	2445
ecotropicalforest	2.25	1.68	-0.86	5.52	1	2920	3130
ecotundra	-0.1	1.84	-3.68	3.63	1	2461	2874
control_mean.sm	-0.07	0	-0.08	-0.06	1	6886	3286
control_mean.st	0.01	0.01	-0.02	0.04	1	6904	3300
depth	0.01	0.01	0	0.03	1	7310	3164
st_mean:ecocropland	0.55	0.15	0.25	0.84	1	2658	2963
st_mean:ecograssland	0.85	0.22	0.42	1.3	1	3079	2866
st_mean:ecosubtropicalforest	0.53	0.23	0.09	0.98	1	4422	3447
st_mean:ecotemperateforest	0.25	0.27	-0.29	0.79	1	2506	2687
st_mean:ecotropicalforest	0.5	0.18	0.16	0.85	1	4007	3171
st_mean:ecotundra	0.4	0.11	0.19	0.62	1	2734	3013
st_mean:control_mean.sm	0.01	0	0.01	0.01	1	3017	3563
st_mean:control_mean.st	0.01	0	0	0.02	1	6768	3359

Further Distributional Parameters:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	1.94	0.01	1.93	1.95	1.01	4778	2614

Fig. SI 2 Estimated day of year offset in ΔSM after accounting for fixed and random effects in the ΔSM model. Lines and ribbons show posterior mean and 95% credible intervals, respectively. Depth was set to 0.3 m, and days of the year outside of the range of observed data were excluded for sites with partial year warming (B4WarmED sites, FutureClim, SWAMP, and TeRaCON).

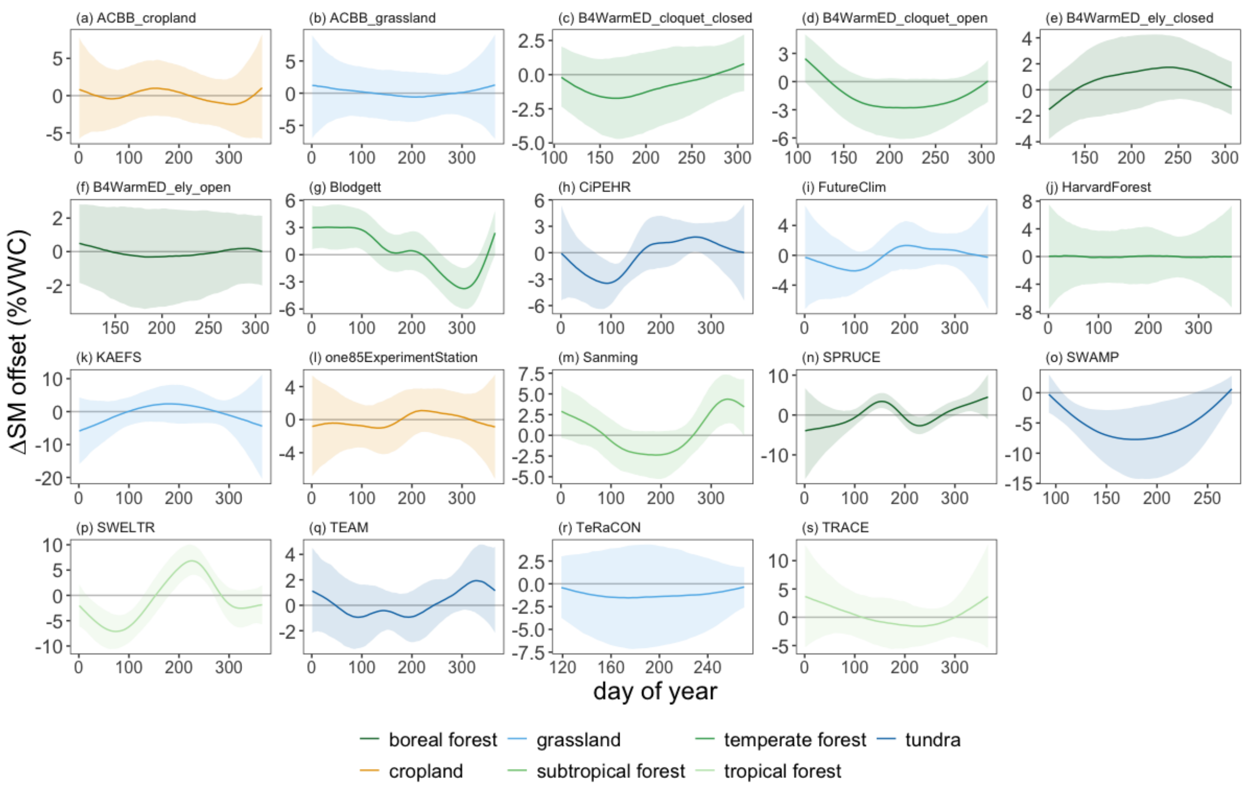


Fig. SI 3 Model predicted ∆SM

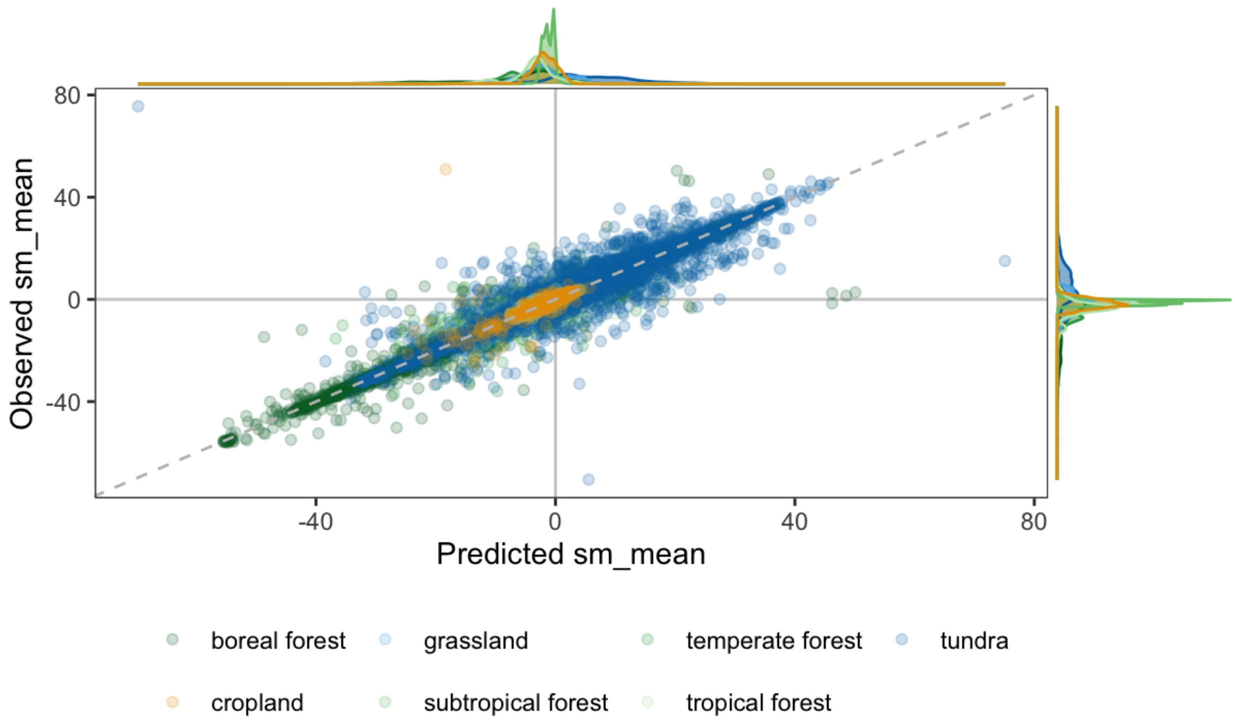


Fig. SI 4 Predicted difference in volumetric soil moisture between warmed and ambient temperature plots as a function of ambient soil moisture and ecosystem type. Lines show estimated marginal means from Bayesian generalized linear model, ribbons show the 75% highest posterior density interval.