

Spatial and temporal variability of CO₂, N₂O and CH₄ fluxes from an urban park in Denmark

Xiao Bai¹, Tom Cripps¹, João Serra¹, Klaus Butterbach-Bahl^{1,2}, and Zhisheng Yao^{1,3}

¹Pioneer Center Land-CRAFT, Department of Agroecology, Aarhus University, Aarhus, 8000, Denmark

5 ²Institute for Meteorology and Climate Research, Institute for Atmospheric Environmental Research (IMK-IFU), Karlsruhe Institute of Technology, Garmisch-Partenkirchen, 82467, Germany

³State Key Laboratory of Atmospheric Environment and Extreme Meteorology, Institute of Atmospheric Physics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, P. R. China

10 *Correspondence to:* Zhisheng Yao (zhishengyao@mail.iap.ac.cn)

Table S1. Soil nitrous oxide(N₂O) oversampling performance and hyperparameters. * = Selected model.

Oversampling rate (x)	Performance					Hyperparameters			
Hot spot	Hot spot accuracy (%)	Normal accuracy (%)	Overall accuracy (%)	AUC	Log Loss	max depth	mtries	ntrees	sample rate
1	35	98	85	0.84	0.4	7	7	200	0.7
1.5	44	95	84	0.85	0.4	7	7	100	0.9
2*	65	88	83	0.85	0.4	7	10	200	0.7
2.5	58	86	80	0.83	0.4	7	10	200	0.8
3	65	84	80	0.84	0.4	7	10	200	0.8
3.5	67	82	79	0.84	0.5	7	15	200	0.7
4	75	73	73	0.84	0.6	7	15	50	0.7

Note: The oversampling rate (x) is the factor by which minority-class samples are duplicated during training. Hot spot accuracy is the proportion of correctly classified hot spot cases. Normal accuracy is the proportion of correctly classified normal cases. Overall accuracy is the total proportion of correctly classified cases across all classes. AUC is the area under the curve, indicating classification performance across thresholds. Log loss is the cross-entropy loss, measuring the quality of the predicted probability distribution. Hyparameter max depth is the max depth of the trees, mtries is the number of variables in each tree split, ntrees is the number of trees, sample rate is the fraction of training data sampled to build each tree.

Table S2. Soil methane (CH₄) oversampling performance and hyperparameters. * = Selected model.

Oversampling rate (x)		Performance					Hyperparameters			
Hot Spot	Cold Spot	Hot spot accuracy (%)	Cold spot accuracy (%)	Normal accuracy (%)	Overall accuracy (%)	Log Loss	max depth	mtries	ntrees	sample rate
1	1	47	63	95	86	0.34	7	7	200	0.9
2	2	47	83	89	86	0.34	7	7	100	0.8
3	3	47	83	89	86	0.35	7	10	100	0.8
4	4	67	85	85	84	0.52	7	10	50	0.9
5	2	67	85	90	87	0.34	7	8	200	0.8
6	2	67	83	91	88	0.33	7	8	100	0.8
6	3	73	85	87	86	0.36	7	7	100	0.7
7*	2*	67	85	91	88	0.33	7	10	200	0.9
8	2	67	81	89	86	0.33	7	10	50	0.9
9	2	73	81	90	87	0.35	7	8	100	0.7
9	3	73	85	86	85	0.38	7	10	100	0.8
10	2	60	83	90	87	0.35	7	10	100	0.8

Note: The oversampling rate (x) is the factor by which minority-class samples are duplicated during training. Hot spot accuracy is the proportion of correctly classified hot spot cases. Normal accuracy is the proportion of correctly classified normal cases. Overall accuracy is the total proportion of correctly classified cases across all classes. Log loss is the cross-entropy loss, measuring the quality of the predicted probability distribution. Hyperparameter max depth is the max depth of the trees, mtries is the number of variables in each tree split, ntrees is the number of trees, sample rate is the fraction of training data sampled to build each tree.

30 **Table S3. Ecosystem respiration (carbon dioxide, CO₂) oversampling performance and hyperparameters. * = Selected model.**

Oversampling rate (x)	Performance					Hyperparameters			
	Hot spot accuracy	Normal accuracy	Overall accuracy	AUC	Log Loss	max depth	mtries	ntrees	sample rate
Hot spot	(%)	(%)	(%)						
1	50	96	90	0.86	0.29	7	6	200	0.9
2	57	96	91	0.88	0.27	7	6	200	0.9
3	57	95	90	0.88	0.28	7	7	200	0.9
4	60	95	90	0.88	0.28	7	7	200	0.8
5	56	97	91	0.88	0.29	7	10	100	0.8
6*	63	95	90	0.88	0.29	7	10	200	0.8
7	63	93	89	0.88	0.42	7	10	50	0.9

Note: The oversampling rate (x) is the factor by which minority-class samples are duplicated during training. Hot spot accuracy is the proportion of correctly classified hot spot cases. Normal accuracy is the proportion of correctly classified normal cases. Overall accuracy is the total proportion of correctly classified cases across all classes. AUC is the area under the curve, indicating classification performance across thresholds. Log loss is the cross-entropy loss, measuring the quality of the predicted probability distribution. Hyparameter max depth is the max depth of the trees, mtries is the number of variables in each tree split, ntrees is the number of trees, sample rate is the fraction of training data sampled to build each tree.

35

Table S4. Predictor variables used to develop a random forest (RF) model to identify hot and cold spots of soil GHG fluxes. Soil moisture: volumetric water content, SOC: soil organic carbon, TN: soil total nitrogen, BD: soil bulk density, SDW: soil dry weight.

Category	Variables	Temporal	Note
Soil physio-chemical properties	Soil temperature (°C)	Y	Recorded for each site and measuring date
	Soil moisture (%)	Y	Recorded for each site and measuring date
	pH	N	Measured in February 2024
	NO ₃ ⁻ (mg N kg ⁻¹ SDW)	N	Measured in February 2024
	NH ₄ ⁺ (mg N kg ⁻¹ SDW)	N	Measured in February 2024
	SOC (g C kg ⁻¹ SDW)	N	Measured in February 2024
	TN (g N kg ⁻¹ SDW)	N	Measured in February 2024
	BD (g cm ⁻³)	N	Measured in February 2024
	Sand (%)	N	Measured in February 2024
	Silt (%)	N	Measured in February 2024
	Clay (%)	N	Measured in February 2024
Vegetation	Grass height (cm)	Y	Recorded for each site and measuring date
Topography parameters	Distance to nearest tree (m)	N	Spatial analysis using ArcGIS Pro and the Danish Datafordeler (2023)
	Distance to nearest water (m)	N	
	Slope (%)	N	

Table S5. Hyperparameter values used to tune each model

Hyperparameters	Values
ntrees	50, 100, 200
max depth	3, 5, 7
sample rate	0.7, 0.8, 0.9
mtries	1, 2, 3, 4, 5, 6, 7, 8, 10, 15

45 Note: ntrees: number of trees, max depth: max depth of the trees, sample rate: the fraction of training data sampled to build
each tree, mtries: the number of variables in each tree split.

Table S6. Soil nitrous oxide(N₂O) stepwise elimination performance and hyperparameters. * = Selected model.

No. of predictors	Performance					Hyperparameters			
	Hot spot accuracy (%)	normal accuracy (%)	overall accuracy (%)	AUC	Log loss	max depth	mtries	ntrees	sample rate
15	65	88	83	0.85	0.39	7	10	200	0.7
14	58	89	83	0.84	0.4	7	8	50	0.8
13	71	84	81	0.84	0.41	7	10	50	0.7
12	60	87	82	0.83	0.41	7	10	200	0.7
11	68	84	81	0.84	0.41	7	6	200	0.8
10	65	86	82	0.83	0.41	7	6	200	0.8
9	65	84	80	0.82	0.43	7	6	200	0.8
8	60	87	81	0.81	0.45	7	4	50	0.9
7	56	89	83	0.82	0.42	7	5	100	0.7
6	46	89	80	0.81	0.44	7	5	50	0.9
5*	54	87	80	0.82	0.42	7	3	100	0.7
4	54	85	79	0.80	0.58	7	3	50	0.9
3	60	84	79	0.81	0.58	7	2	200	0.9
2	50	90	81	0.82	0.84	7	2	200	0.8
1	35	94	82	0.80	0.7	7	1	50	0.7

Note: No. of predictors is the number of predictor variables used to train the model. Hot spot accuracy is the proportion of correctly classified hot spot cases. Normal accuracy is the proportion of correctly classified normal cases. Overall accuracy is the total proportion of correctly classified cases across all classes. AUC is the area under the curve, indicating classification performance across thresholds. Log loss is the cross-entropy loss, measuring the quality of the predicted probability distribution. Hyperparameter max depth is the max depth of the trees, mtries is the number of variables in each tree split, ntrees is the number of trees, sample rate is the fraction of training data sampled to build each tree.

Table S7. Soil methane (CH₄) stepwise elimination performance and hyperparameters. * = Selected model.

No. of predictors	Performance					Hyperparameters			
	Hot spot accuracy (%)	Cold spot accuracy (%)	Normal accuracy (%)	Overall accuracy (%)	Log loss	max depth	Mtries	ntrees	sample rate
15	67	85	91	88	0.33	7	10	200	0.9
14	73	83	89	87	0.33	7	7	50	0.9
13	67	83	90	87	0.33	7	10	200	0.9
12	60	83	89	87	0.33	7	7	100	0.9
11	67	83	90	87	0.33	7	7	200	0.8
10	53	83	90	87	0.34	7	8	100	0.9
9	60	83	91	87	0.34	7	7	200	0.8
8	53	83	92	88	0.35	7	6	200	0.8
7	53	83	91	87	0.36	7	4	100	0.9
6	53	81	91	87	0.37	7	4	200	0.8
5*	53	78	91	86	0.35	7	3	200	0.7
4	53	76	91	86	0.35	7	2	200	0.9
3	57	76	85	81	0.42	7	2	50	0.9
2	47	81	85	81	0.41	7	1	100	0.9
1	40	44	79	71	0.92	7	1	200	0.7

55 Note: No. of predictors is the number of predictor variables used to train the model. Hot spot accuracy is the proportion of correctly classified hot spot cases. Normal accuracy is the proportion of correctly classified normal cases. Overall accuracy is the total proportion of correctly classified cases across all classes. Log loss is the cross-entropy loss, measuring the quality of the predicted probability distribution. Hyperparameter max depth is the max depth of the trees, mtries is the number of variables in each tree split, ntrees is the number of trees, sample rate is the fraction of training data sampled to build each tree.

Table S1. Ecosystem respiration (carbon dioxide, CO₂) stepwise elimination performance and hyperparameters. * = Selected model.

No. of predictors	Performance					Hyperparameters			
	Hot spot accuracy (%)	Normal accuracy (%)	Overall accuracy (%)	AUC	Log loss	max depth	mtries	ntrees	sample rate
15	63	95	90	0.88	0.29	7	10	200	0.8
14	67	91	87	0.89	0.30	7	10	100	0.8
13	60	96	91	0.88	0.29	7	8	200	0.8
12	63	94	90	0.88	0.29	7	7	200	0.8
11	57	97	91	0.86	0.32	7	8	50	0.8
10	63	94	90	0.88	0.29	7	6	200	0.8
9	67	92	88	0.87	0.30	7	5	200	0.8
8	63	92	88	0.88	0.29	7	5	200	0.8
7	63	95	90	0.88	0.42	7	4	50	0.7
6	60	95	90	0.89	0.28	7	4	100	0.9
5	60	96	91	0.88	0.28	7	3	200	0.9
4*	67	94	90	0.88	0.30	7	3	100	0.8
3	63	88	85	0.85	0.34	7	2	50	0.8
2	53	95	89	0.87	0.35	7	1	50	0.8
1	33	87	80	0.80	0.75	7	1	200	0.7

Note: No. of predictors is the number of predictor variables used to train the model. Hot spot accuracy is the proportion of correctly classified hot spot cases. Normal accuracy is the proportion of correctly classified normal cases. Overall accuracy is the total proportion of correctly classified cases across all classes. AUC is the area under the curve, indicating classification performance across thresholds. Log loss is the cross-entropy loss, measuring the quality of the predicted probability distribution. Hyparameter max depth is the max depth of the trees, mtries is the number of variables in each tree split, ntrees is the number of trees, sample rate is the fraction of training data sampled to build each tree.

Table S2. Topsoil (0-10 cm) physico-chemical properties as well as cumulative soil nitrous oxide(N₂O), soil methane (CH₄), and ecosystem respiration (carbon dioxide, CO₂) of all sampling sites. For each plot N = 1. SOC: soil organic carbon, TN: soil total nitrogen, SDW: soil dry weight.

Site	Latitude (N)	Longitutde (E)	SOC (g C kg ⁻¹ SDW)	TN (g N kg ⁻¹ SDW)	Clay (<0.002mm)	Silt (0.002- 0.063mm)	Sand (0.063- 0.2mm)	pH (H ₂ O)	NO ₃ ⁻ (mg N kg ⁻¹ SDW)	NH ₄ ⁺ (mg N kg ⁻¹ SDW)	Soil bulk density (g cm ⁻³)	Cumulative N ₂ O fluxes (kg N ha ⁻¹)	Cumulative CH ₄ fluxes (kg C ha ⁻¹)	Cumulative CO ₂ fluxes (Mg C ha ⁻¹ y ⁻¹)
1	56°10.0302'	10°12.2769'	55.7	4.6	9.80	16.44	64.72	7.70	0.91	24.85	0.84	0.41	1.52	8.53
2	56°10.0460'	10°12.2558'	25.4	2.1	7.38	14.66	74.03	5.80	2.26	13.48	1.26	0.33	1.44	8.67
3	56°10.0529'	10°12.2194'	38.9	3.0	11.48	20.73	61.57	7.40	2.80	15.08	1.15	0.16	-0.80	7.10
4	56°10.0601'	10°12.2260'	30.0	2.0	9.20	21.02	64.88	6.10	1.91	14.23	1.27	0.72	-0.28	5.20
5	56°10.0723'	10°12.2022'	26.4	2.1	8.49	21.84	65.53	6.40	0.98	13.38	1.38	0.32	-0.10	7.44
6	56°10.0814'	10°12.1682'	26.3	2.1	11.44	23.29	61.26	5.80	2.39	15.27	1.18	1.02	-0.25	3.93
7	56°10.0799'	10°12.1523'	27.1	2.1	9.69	16.06	69.88	5.60	1.04	12.08	1.35	1.27	-0.19	3.82
8	56°10.1085'	10°12.1384'	22.5	1.7	11.92	22.64	61.94	5.00	0.05	3.40	1.46	0.19	-0.19	1.84
9	56°10.1071'	10°12.1359'	23.8	2.0	10.93	22.56	63.03	6.20	1.82	9.38	1.32	0.16	-0.25	4.13
10	56°10.1025'	10°12.1880'	20.6	1.8	12.45	23.99	60.62	5.40	0.09	18.22	1.35	0.02	0.12	3.68
11	56°10.1188'	10°12.1606'	29.8	2.4	9.00	21.84	64.31	5.60	0.54	8.02	1.19	0.65	-0.04	7.08
12	56°10.1261'	10°12.0904'	21.5	1.7	11.23	24.28	61.41	5.90	0.17	8.65	1.33	-0.01	-0.26	6.15
13	56°10.1457'	10°12.0543'	28.5	2.2	14.13	26.43	55.40	6.00	0.32	10.35	1.32	0.04	-0.48	6.41
14	56°10.1530'	10°12.0380'	30.2	2.1	11.76	27.78	55.77	5.60	0.09	7.84	1.14	0.02	-0.26	4.33
15	56°10.1821'	10°12.0246'	22.1	1.9	10.51	23.04	62.99	6.00	0.38	5.24	1.45	0.20	-0.13	8.50
16	56°10.1918'	10°12.0176'	33.0	1.8	11.63	24.78	58.16	6.30	1.06	4.05	1.33	0.18	-0.35	4.78
17	56°10.2119'	10°12.0365'	31.4	2.3	14.40	26.57	54.25	5.90	0.69	6.00	1.42	0.39	-0.25	4.00
18	56°10.2011'	10°12.0636'	29.0	2.2	13.90	27.08	54.63	6.10	0.46	6.30	1.29	0.03	-0.16	4.80
19	56°10.2181'	10°12.0680'	30.1	2.6	10.54	22.40	62.30	6.10	1.26	7.43	1.26	0.26	0.03	7.05
20	56°10.2425'	10°12.0956'	64.5	3.3	16.95	15.73	56.57	6.60	3.07	11.55	1.06	0.21	-0.67	6.54
21	56°10.2467'	10°12.1111'	37.2	2.6	12.72	26.25	55.34	6.40	0.55	11.14	1.26	0.41	-0.35	7.88
22	56°10.2161'	10°12.1457'	30.3	2.2	10.24	15.00	70.08	7.50	2.89	18.58	1.11	0.11	-0.60	5.37
23	56°10.2166'	10°12.1313'	44.8	2.7	12.70	24.39	55.58	5.00	0.06	10.93	1.12	0.00	-0.47	6.65
24	56°10.2118'	10°12.1275'	39.1	2.5	11.66	24.55	57.41	5.40	0.62	12.98	1.30	0.54	-0.57	5.69
25	56°10.2193'	10°12.1017'	38.7	2.9	12.51	26.54	54.90	5.30	0.42	18.63	1.06	0.18	-0.79	7.78
26	56°10.1972'	10°12.1089'	65.8	6.4	14.41	25.53	49.82	6.70	0.48	21.71	0.64	-0.01	2.60	10.15
27	56°10.1912'	10°12.1748'	31.6	2.8	9.43	12.37	73.19	6.80	2.57	6.46	1.34	0.06	-0.06	3.09
28	56°10.1827'	10°12.1729'	39.5	3.0	12.21	18.83	62.80	6.70	3.19	6.80	1.25	0.53	-0.41	4.25
29	56°10.1865'	10°12.1930'	23.8	1.8	13.18	21.60	61.85	7.10	2.99	3.43	1.50	0.08	-0.08	5.03
30	56°10.1906'	10°12.2204'	27.8	2.2	12.48	18.26	65.22	7.40	3.48	7.84	1.16	0.53	-0.32	6.39
31	56°10.1698'	10°12.2083'	32.3	2.7	16.49	22.09	56.81	6.10	1.78	15.02	1.28	0.88	-0.40	4.64
32	56°10.1413'	10°12.1802'	92.7	5.9	12.71	11.77	60.11	6.30	0.11	15.66	0.71	0.21	5.44	4.65
33	56°10.1453'	10°12.2147'	50.9	4.0	13.45	24.95	53.40	6.20	1.08	29.06	1.09	0.15	-0.33	10.17
34	56°10.1492'	10°12.2385'	17.0	1.1	8.01	4.90	84.67	8.40	2.49	12.99	1.49	0.04	0.05	4.07
35	56°10.1274'	10°12.2612'	28.5	2.3	13.23	21.49	60.97	5.70	0.28	11.51	1.21	0.06	-0.26	5.53

36	56°10.1535'	10°12.2963'	28.0	2.1	13.88	20.74	61.05	6.20	2.17	11.65	1.19	5.10	0.06	6.22
37	56°10.1251'	10°12.3218'	41.6	2.8	14.42	15.88	62.92	7.30	7.24	12.14	1.20	0.16	-0.24	2.56
38	56°10.1287'	10°12.3059'	55.0	4.2	17.29	27.13	45.43	7.30	12.37	7.48	1.03	0.37	-0.17	3.68
39	56°10.0820'	10°12.3154'	43.8	3.6	15.92	23.33	54.33	7.20	1.99	18.48	1.10	0.07	-0.34	7.33
40	56°10.1153'	10°12.2329'	55.2	4.6	14.33	26.09	50.87	6.80	0.26	22.58	1.13	1.75	1.36	9.45
41	56°10.0693'	10°12.2798'	53.4	4.0	15.22	18.17	55.95	7.40	6.21	23.09	1.02	0.51	-0.25	9.16
42	56°10.0711'	10°12.3101'	43.1	2.6	10.06	13.28	69.24	8.10	1.19	15.56	1.25	0.01	-0.83	10.14
43	56°10.2074'	10°12.0220'	33.8	2.1	12.33	24.91	57.11	5.80	0.55	10.72	1.47	0.05	-0.46	6.10
44	56°10.1366'	10°12.1934'	79.5	5.4	15.12	14.95	57.39	7.10	4.48	12.61	0.69	0.19	-0.41	4.79
45	56°10.1377'	10°12.1938'	71.4	4.3	14.65	16.87	56.99	7.30	4.37	15.07	0.84	0.10	-0.51	3.24
46	56°10.1345'	10°12.1954'	70.9	4.8	11.27	24.42	52.00	7.50	4.39	16.64	0.99	0.50	-0.24	2.84
47	56°10.1323'	10°12.1909'	48.3	3.1	12.37	20.82	59.21	5.90	1.36	14.00	1.14	0.36	-0.25	4.62
48	56°10.1293'	10°12.1890'	31.6	2.2	12.16	23.91	58.61	5.90	0.80	14.72	1.23	0.04	-0.24	3.11
49	56°10.1267'	10°12.1911'	25.0	2.0	13.71	22.37	60.51	5.90	0.34	8.68	1.29	0.08	-0.10	3.83
50	56°10.1234'	10°12.1931'	28.9	2.2	13.55	21.04	60.81	5.80	0.67	13.20	1.27	0.19	-0.15	2.04
51	56°10.1206'	10°12.1885'	32.1	2.6	13.36	23.20	58.20	6.10	1.10	12.92	1.35	0.47	-0.42	4.17
52	56°10.1465'	10°12.1682'	46.1	4.1	10.49	24.16	57.34	5.20	0.27	3.05	1.08	9.96	0.01	0.71
53	56°10.1455'	10°12.1668'	41.1	3.7	11.02	26.33	56.06	7.20	2.86	11.80	1.28	0.38	0.35	4.75
54	56°10.1434'	10°12.1646'	56.9	4.4	10.69	25.66	54.38	6.50	1.17	20.64	1.01	0.12	-0.06	4.76
55	56°10.1425'	10°12.1614'	47.9	3.6	11.17	27.55	53.65	6.00	4.42	10.34	1.10	0.44	-0.24	3.75
56	56°10.1393'	10°12.1570'	39.0	2.9	10.61	23.18	59.66	6.00	7.17	10.86	1.17	0.56	-0.44	4.33

Table S3. The performance metrics for each greenhouse gases (GHG) of soil nitrous oxide(N₂O), soil methane (CH₄), and soil and ecosystem respiration (carbon dioxide, CO₂) of the final models.

GHG	Hot spot accuracy (%)	Cold spot accuracy (%)	Normal accuracy (%)	Overall accuracy (%)	AUC	Log loss
N ₂ O	54	n/a	87	80	0.82	0.42
CH ₄	53	78	91	86	n/a	0.36
CO ₂	67	n/a	94	90	0.88	0.30

Note: Hot spot accuracy is the proportion of correctly classified hot spot cases. Normal accuracy is the proportion of correctly classified normal cases. Overall accuracy is the total proportion of correctly classified cases across all classes. AUC is the area under the curve, indicating classification performance across thresholds. Log loss is the cross-entropy loss, measuring the quality of the predicted probability distribution.

75



Fig S1: Soil greenhouse gases fluxes measurement with the fast closed chamber approach

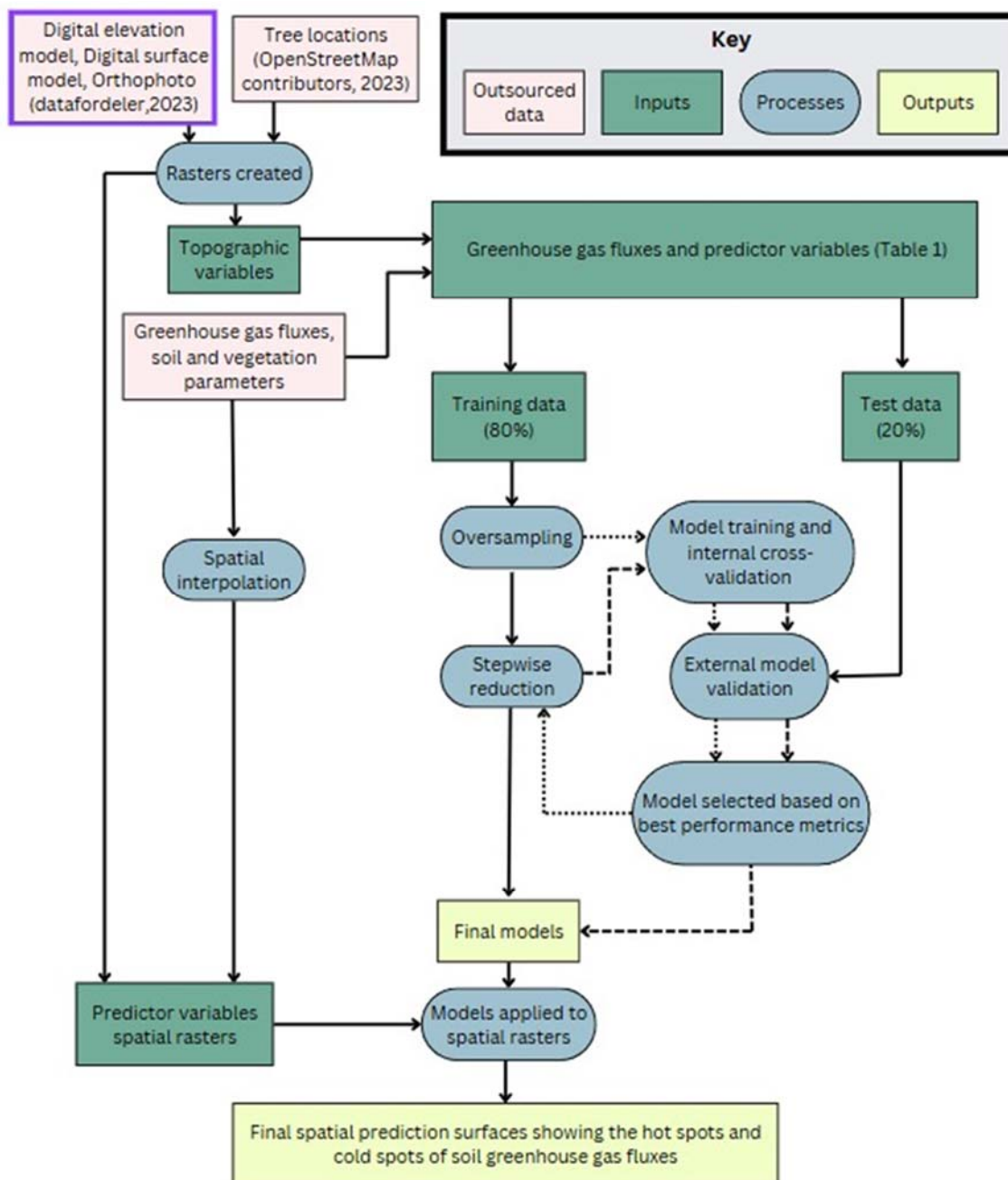


Fig S2: Workflow summary showing the methods used to develop the Random Forest models to predict the spatial distribution of hot spots and cold spots of soil greenhouse gas (GHG) fluxes. The solid line follows the overall workflow, the dotted line follows the oversampling workflow, and the dashed line follows the stepwise reduction workflow.

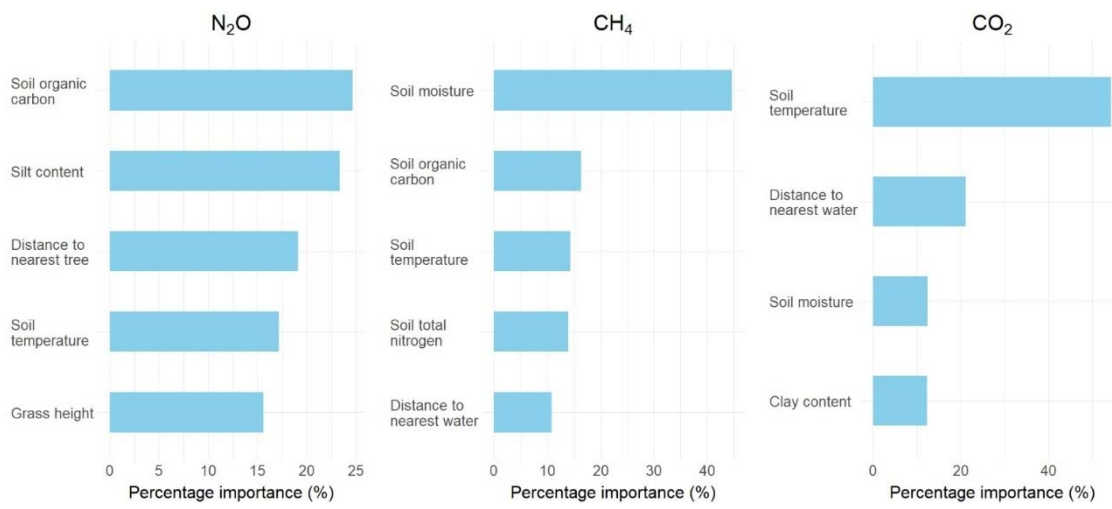


Fig S3: Predictors variables importance plots showing the relative importance assigned by the random forest algorithm to predictor variables in the final models for soil nitrous oxide(N₂O), soil methane (CH₄), and ecosystem respiration (carbon dioxide, CO₂)

Reference

Datafordeler: Digital Elevation Model, Datafordeler [data set], <https://datafordeler.dk/dataoversigt/danmarks-hoejdemodel-dhm/terraen-praedefineret-geotiff/>, 2023.

95 Datafordeler: Digital Surface Model, Datafordeler [data set], <https://datafordeler.dk/dataoversigt/danmarks-hoejdemodel-dhm/overfladepraedefineret-geotiff/>, 2023.

Datafordeler: Orthophoto, Datafordeler [data set], <https://datafordeler.dk/dataoversigt/geodanmark-ortofoto/geodanmark-ortofoto-10-cm-praedefineret/>, 2023.