

Supporting Information

Artificial neural networks to investigate the bioavailability of selenium nanoparticles in soil-crop system

Jing Li^{ab}, Le Yue^{ab}, Feiran Chen^{ab}, Xuesong Cao^{ab}, Bingxu Cheng^{ab}, Chuanxi Wang^{*ab}, Zhenyu Wang^{ab}, and Baoshan Xing^c

^a Institute of Environmental Processes and Pollution Control, and School of Environment and Civil Engineering, Jiangnan University, Wuxi, Jiangsu, 214122, China

^b Jiangsu Engineering Laboratory for Biomass Energy and Carbon Reduction Technology, Jiangnan University, Wuxi, Jiangsu, 214122, China

^c Stockbridge School of Agriculture, University of Massachusetts, Amherst, MA 01003, USA

*Corresponding author:

Email address: wangcx2018@jiangnan.edu.cn (Dr. Chuanxi Wang)

Number of pages: 19

Number of texts: 1

Number of tables: 6

Number of figures: 16

Text S1 The architecture of the ANN

For yield prediction (GN₁₂₀ as output), the effective structure was designed as follows: D30 ANN: $5 \times 4 \times 1$ using Bias, P_Ch1_30, P_Ht_30, S_Eh_30, and S_Ec_30 in the input layer (Figure 3a); D60 ANN: $5 \times 5 \times 1$ using Bias, P_Gs_60, P_Pn_60, S_Ec_60, and S_Eh_60 in the input layer (Figure 3b); D90 ANN: $6 \times (4 \times 3) \times 1$ using Bias, S_{N-NH₄⁺}_90, P_Ch1_90, S_Eh_90, P_Pn_90, as well as S_{pH}_90 in the input layer (Figure 3c); D120 ANN: $7 \times (5 \times 4) \times 1$ using Bias, P_Ch1_120, P_Fv/Fm_120, P_Pn_120, S_{N-NH₄⁺}_120, P_ADW, as well as P_RFW in the input layer (Figure 3d).

The chosen architecture ANNs for quality prediction (Se_F as output) were in general relatively simple (Figure 3). They were designed as follows: D30 ANN: $5 \times 2 \times 1$ using Bias, P_Ch1_30, S_Eh_30, S_TOC_30, and S_{N-NH₄⁺}_30 in the input layer (Figure 3e); D60 ANN: $5 \times (3 \times 2) \times 1$ using Bias, P_Gs_60, P_PanN_60, P_Pn_60, and S_TOC_60 in the input layer (Figure 3f); D90 ANN: $5 \times 3 \times 1$ using Bias, P_Gs_90, S_{pH}_90, P_Pn_90, and P_Ch1_90 in the input layer (Figure 3g); D120 ANN: $7 \times (5 \times 4) \times 1$ using Bias, P_Ch1_120, P_RFW, S_TOC_120, P_ADW, P_AFW, and P_Fv/Fm_120 in the input layer (Figure 3h).

Table S1 the definitions of abbreviations

Abbreviations	Definitions
<i>Plant parameters</i>	
LFW	Leaf fresh weight
AFW	Aboveground fresh weight
RFW	Root fresh weight
LDW	Leaf dry weight
ADW	aboveground dry weight
RDW	Root dry weight
Fe_L	Fe content in leave
Fe_S	Fe content in shoot
Fe_P	Fe content in whole plant
Fe_F	Fe content in fruit
Fe_R	Fe content in root
Chl	Chlorophyll
Ht	Plant Height
Fv/Fm	Optimal/maximal photochemical efficiency of PS II in the dark
Pn	Net photosynthetic rate
Gs	Stomatal conductance
PanN	Panicle number
<i>Soil parameters</i>	
S_Ec	Soil electrical conductivity
S_Eh	Soil redox potential
S_TOC	Soil organic matter
S_Fe	Soil Fe content
S_Se	Soil Se content
S_NH ₄ ⁺ -N	Soil ammonium nitrogen
<i>Plant yield and quality</i>	
GN_D120	Grain number
Se_F	Rice Se content

Table S2 Descriptive statistic of the plant physiological activities data.

Parameters	CK					Ion					NP				
	Mean±SD	Min	Max	Skewness	Kurtosis	Mean±SD	Min	Max	Skewness	Kurtosis	Mean±SD	Min	Max	Skewness	Kurtosis
P_LFW	1.9±0.29	1.32	2.32	-0.43	-0.38	2.26±0.18	2.06	2.60	0.59	-1.14	3.05±0.32	2.64	3.60	0.25	-1.5
P_AFW	24.92±4.11	18.04	30.72	-0.58	-1.13	22.51±4.11	17.31	29.21	0.5	-1.35	35.08±5.81	25.58	44.15	0.08	-1.14
P_RFW	23.36±5.01	15.85	29.31	-0.29	-1.58	23.24±4.64	15.07	28.55	-0.7	-1.37	25.94±5.92	20.59	38.32	0.97	-0.47
P_LDW	1.55±0.16	1.24	1.82	-0.22	-0.17	1.68±0.18	1.41	1.95	0.2	-1.42	2.23±0.22	1.87	2.56	0.24	-1.32
P_ADW	12.81±2.43	8.90	15.37	-0.71	-1.56	12.86±3.37	8.90	17.71	0.3	-1.7	19.57±4.13	13.70	25.87	0.08	-1.44
P_RDW	2.74±0.88	1.48	4.15	-0.07	-1.35	2.38±0.72	1.26	3.29	-0.16	-1.78	2.93±0.85	1.91	4.03	0.02	-1.97
P_Ht_30	50.22±3.57	46.00	57.00	0.8	-0.94	48.17±2.28	44.00	51.00	-0.47	-1.41	56.22±3.13	51.00	62.00	0.44	-0.38
P_Ht_60	80.99±2.88	75.00	84.50	-0.76	-0.48	78.82±4.22	71.00	85.40	-0.12	-0.73	90.51±2.81	87.00	97.00	0.99	0.62
P_PanN_60	3.08±0.86	2.00	4.00	-0.18	-1.87	3.25±0.92	2.00	5.00	0.14	-0.77	5.42±2.18	3.00	9.00	0.49	-1.03
P_FvFm_60	0.81±0.01	0.79	0.83	-0.39	-0.05	0.81±0.01	0.79	0.83	-0.38	-0.82	0.83±0.01	0.81	0.84	-1.15	2.72
P_FvFm_90	0.83±0.01	0.82	0.84	0	-0.86	0.83±0.01	0.81	0.84	-0.14	-0.77	0.84±0	0.83	0.84	-1.33	-0.33
P_FvFm_120	0.79±0.02	0.76	0.82	0.11	-1.45	0.8±0.02	0.78	0.83	0.24	-1.09	0.82±0.02	0.79	0.84	0.00	-1.19
P_Gs_60	140.22±16.92	116.00	163.00	0.11	-1.68	172.44±10.99	153.00	189.00	-0.15	-0.97	204.22±36.65	165.00	268.00	0.68	-1.37
P_Gs_90	122.22±20.08	91.00	146.00	-0.5	-1.46	94.56±11.15	74.00	112.00	-0.51	-0.84	157.44±9.69	145.00	180.00	1.12	0.78
P_Gs_120	107.92±27.15	79.00	153.00	0.76	-1.49	105.44±7.63	89.00	115.00	-0.9	0.02	142.78±22.79	118.00	176.00	0.68	-1.63
P_Pn_60	8.61±1.08	6.90	10.90	0.33	0.35	8.91±0.77	7.80	10.20	0.58	-1.17	11.18±0.54	9.90	12.10	-0.81	1.51
P_Pn_90	4.92±0.56	4.00	5.60	-0.43	-1.36	5.03±0.41	4.50	6.20	1.74	5.19	8.09±1.25	5.70	9.80	-0.62	-0.8
P_Pn_120	4.66±1.1	3.30	6.20	-0.03	-1.58	4.63±0.18	4.30	5.00	0.27	0.14	7.12±0.65	6.10	7.90	-0.42	-1.72

Parameters	CK					Ion					NP				
	Mean±SD	Min	Max	Skewness	Kurtosis	Mean±SD	Min	Max	Skewness	Kurtosis	Mean±SD	Min	Max	Skewness	Kurtosis
P_Fe_L	0.53±0.08	0.40	0.65	-0.04	-1.4	0.31±0.08	0.22	0.48	0.52	-0.68	0.35±0.1	0.22	0.51	-0.04	-1.35
P_Fe_S	0.51±0.09	0.39	0.65	0.27	-1.39	0.5±0.16	0.33	0.84	0.99	-0.32	0.43±0.06	0.33	0.51	-0.3	-1.38
P_Fe_P	0.11±0.02	0.08	0.13	-0.26	-1.37	0.15±0.04	0.09	0.22	0.18	-0.16	0.1±0.02	0.08	0.13	0.82	-0.26
P_Fe_F	26.92±5.84	18.93	36.55	0.53	-1.2	22.2±2.95	16.30	27.62	-0.43	0.32	23.77±2.7	18.84	27.80	-0.35	-0.88
P_Fe_R	24.78±5.28	17.15	31.80	-0.33	-1.52	24.37±5.73	16.18	32.01	-0.18	-1.58	21.36±1.63	18.26	24.09	0	-0.59
P_Ch1_30	38.33±1.93	35.70	40.90	-0.05	-1.72	39.86±2.31	36.60	43.90	0.69	-0.94	43.99±1.78	41.00	46.90	0.26	-0.91
P_Ch1_60	30.23±1.67	26.90	32.60	-0.26	-0.52	32.15±1.78	28.50	34.50	-0.91	0.02	36.98±1.38	35.60	40.50	1.51	2.32
P_Ch1_90	18.17±4.36	10.70	25.00	0.21	-1.1	16.58±3.83	12.20	24.40	0.79	-0.61	27.57±2.35	23.40	31.90	0.05	-0.41
P_Ch1_120	7.84±0.64	6.30	8.60	-1.3	1.25	8.49±1.2	5.70	9.90	-1.05	0.63	11.19±1.36	9.20	13.50	0.22	-1.09

Table S3 Descriptive statistic of the soil properties and plant response.

Parameters	CK					Ion					NP				
	Mean±SD	Min	Max	Skewness	Kurtosis	Mean±SD	Min	Max	Skewness	Kurtosis	Mean±SD	Min	Max	Skewness	Kurtosis
S_EC_30	0.38±0.08	0.28	0.48	0.01	-1.65	0.4±0.06	0.34	0.48	0.77	-1.62	0.4±0.05	0.35	0.46	0.67	-1.65
S_EC_60	0.32±0.06	0.25	0.40	0.2	-1.62	0.36±0.04	0.32	0.42	0.74	-1.6	0.31±0.06	0.26	0.39	0.8	-1.63
S_EC_90	0.34±0.08	0.23	0.45	-0.01	-1.65	0.32±0.06	0.25	0.41	0.49	-1.48	0.35±0.06	0.27	0.42	-0.03	-1.65
S_EC_120	0.38±0.05	0.30	0.44	-0.45	-1.62	0.43±0.04	0.39	0.49	0.65	-1.64	0.41±0.08	0.30	0.49	-0.68	-1.65
S_Eh_30	259.83±19.73	238.00	285.90	0.26	-1.86	278.4±13.67	259.80	296.00	-0.22	-1.6	272.91±11.1	252.10	284.90	-0.7	-1.12
S_Eh_60	255.68±12.28	235.40	271.10	-0.59	-1.49	265.24±16	237.30	287.00	-0.64	-1.11	272±18.96	242.70	297.10	-0.44	-1.52
S_Eh_90	285.93±29.71	243.00	328.80	0.07	-1.58	287.34±22.04	247.20	316.90	-0.66	-0.9	289.62±19.04	261.30	312.30	-0.46	-1.66
S_Eh_120	304.01±11.34	284.30	320.10	-0.52	-1.18	290.36±7.78	272.70	301.10	-0.68	0.6	293.08±5.41	281.80	302.30	-0.49	0.33
S_N-NH4+_30	1.17±0.39	0.51	1.72	-0.51	-0.71	1.09±0.34	0.73	1.72	0.78	-1.06	1.18±0.48	0.68	2.01	0.69	-1.36
S_N-NH4+_60	2.63±0.36	2.03	3.02	-0.73	-1.5	2.48±0.24	2.11	2.73	-0.8	-1.5	2.63±0.31	2.21	3.05	-0.22	-1.6
S_N-NH4+_90	4.18±1.22	2.86	5.89	0.48	-1.64	4.14±0.78	3.17	5.17	0.19	-1.65	4.13±0.79	3.14	5.16	0.16	-1.64
S_N-NH4+_120	4.59±0.65	3.98	5.54	0.72	-1.64	5.26±0.44	4.82	5.93	0.75	-1.55	4.62±1.17	3.18	6.22	0.25	-1.63
S_Fe_30	24.69±6.64	16.16	33.28	-0.2	-1.63	22.74±6.32	15.13	31.66	0.38	-1.61	22.49±4.36	16.64	27.99	-0.05	-1.6
S_Fe_60	20.29±6.72	11.71	35.01	0.46	0.2	20.54±5.92	11.48	27.11	-0.72	-1.57	19.89±5.15	11.81	25.58	-0.65	-1.6
S_Fe_90	19.45±5.4	10.79	25.19	-0.76	-1.49	19.25±5.15	11.67	25.38	-0.48	-1.61	18.42±4.17	11.42	22.54	-0.79	-1.45
S_Fe_120	20.29±6.72	11.71	35.01	0.46	0.2	20.54±5.92	11.48	27.11	-0.72	-1.57	20.16±5.33	11.81	25.43	-0.67	-1.63
S_pH_30	7.52±0.25	7.16	7.82	-0.65	-1.61	7.53±0.16	7.34	7.74	0.31	-1.65	7.55±0.18	7.29	7.77	-0.24	-1.79
S_pH_60	7.12±0.12	6.97	7.31	0.54	-1.46	7.2±0.14	7.01	7.43	0.43	-1.4	7.22±0.1	7.13	7.40	0.79	-1.17

Parameters	CK					Ion					NP				
	Mean±SD	Min	Max	Skewness	Kurtosis	Mean±SD	Min	Max	Skewness	Kurtosis	Mean±SD	Min	Max	Skewness	Kurtosis
S_pH_90	7.17±0.1	7.01	7.29	-0.63	-1.32	7.2±0.15	6.99	7.40	-0.28	-1.58	7.16±0.1	7.00	7.34	-0.2	-0.67
S_pH_120	7.07±0.15	6.84	7.24	-0.51	-1.54	7.17±0.18	6.93	7.40	0.06	-1.8	7.07±0.16	6.84	7.27	-0.32	-1.68
S_TOC_30	63.31±1.98	58.70	66.40	-0.88	1.12	63.11±3.01	56.90	69.20	0.03	1.01	61.59±1.02	60.00	63.90	0.87	0.81
S_TOC_60	66.29±2.74	60.20	69.80	-0.74	0.07	60.47±3.52	55.60	66.60	0.37	-1.21	62.87±3.6	56.00	67.80	-0.57	-0.86
S_TOC_90	61.96±2.8	56.70	67.40	0.02	0.39	63.55±2.55	57.10	66.40	-1.51	2.32	60.43±2.85	55.20	65.30	-0.3	-0.3
S_TOC_120	63.61±2.79	59.90	69.30	0.56	-0.63	61.79±3.7	53.70	66.90	-0.59	0.32	60.79±2.99	56.50	65.10	-0.5	-1.48
S_Se_30	0.15±0.04	0.08	0.20	-0.59	-1.38	0.33±0.12	0.18	0.58	0.57	-0.48	0.4±0.18	0.21	0.70	0.48	-1.48
S_Se_60	0.09±0.06	0.02	0.18	-0.18	-1.32	0.21±0.03	0.16	0.25	-0.43	-1.4	0.18±0.03	0.15	0.24	0.76	-0.24
S_Se_90	0.1±0.04	0.05	0.17	0.35	-1.17	0.15±0.08	0.06	0.27	0.32	-1.69	0.19±0.03	0.15	0.24	0.46	-0.34
S_Se_120	0.09±0.06	0.02	0.18	-0.18	-1.32	0.21±0.03	0.16	0.25	-0.43	-1.4	0.12±0.07	0.04	0.24	0.53	-1.48
Se_F	0.11±0.03	0.08	0.17	0.85	-1.22	0.32±0.11	0.19	0.54	0.72	-0.36	0.39±0.18	0.18	0.64	0.57	-1.62
GN_120	112.89±15.3	88.00	133.00	-0.38	-1.66	123.89±11.25	102.00	140.00	-0.2	-0.72	149.78±20.72	117.00	180.00	-0.15	-1.39

Table S4 correlation between features and outcomes

Plant parameters	Grainnumber120d	FruitSe(g/kg)	Soil parameters	Grainnumber120d	FruitSe(g/kg)
P_LFW	0.584**	0.503**	S_pH_30	-0.188	0.269
P_AFW	0.616**	0.676**	S_Ec_30	-0.483**	0.037
P_RFW	0.453**	0.623**	S_Eh_30	0.655**	0.437**
P_LDW	0.500**	0.260	S_TOC_30	-0.368*	-0.358*
P_ADW	0.637**	0.784**	S_Fe_30	-0.423*	0.040
P_RDW	0.324	0.495**	S_Se_30	0.226	-0.020
P_Fe_L	-0.136	-0.309	S_NH4 ⁺ -N_30	0.221	-0.389*
P_Fe_S	-0.410*	-0.077	S_pH_60	-0.263	0.134
P_Fe_P	-0.191	0.127	S_EC_60	-0.666**	-0.063
P_Fe_F	-0.138	-0.486**	S_Eh_60	0.757**	0.461**
P_Fe_R	0.091	-0.315	S_TOC_60	-0.403*	-0.656**
P_Ch1_30	0.795**	0.755**	S_Fe_60	-0.128	0.317
P_Ht_30	0.464**	0.044	S_Se_60	0.124	0.300
P_Fv/Fm_60	0.400*	0.311	S_NH4 ⁺ -N_60	0.603**	0.128
P_Pn_60	0.772**	0.550**	S_pH_90	-0.463**	-0.412*
P_Gs_60	0.727**	0.871**	S_Ec_90	-0.192	0.259
P_Ch1_60	0.617**	0.472**	S_Eh_90	0.629**	0.220
P_Ht_60	0.504**	0.388*	S_TOC_90	-0.381*	-0.193
P_PanN_60	0.619**	0.773**	S_Fe_90	-0.200	0.272
P_Fv/Fm_90	0.487**	0.352*	S_Se_90	0.098	0.378*
P_Pn_90	0.786**	0.567**	S_NH4 ⁺ -N_90	0.583**	0.276
P_Gs_90	0.376*	0.349*	S_pH_120	-0.495**	-0.361*
P_Ch1_90	0.788**	0.520**	S_Ec_120	-0.054	0.533**
P_Fv/Fm_120	0.589**	0.739**	S_Eh_120	-0.001	-0.422*
P_Pn_120	0.689**	0.479**	S_TOC_120	-0.433**	-0.674**
P_Gs_120	0.420*	0.022	S_Fe_120	-0.094	0.346*
P_Ch1_120	0.737**	0.605**	S_Se_120	-0.409*	-0.006
			S_NH4 ⁺ -N_120	0.498**	0.540**

** . P<0.01

* . P<0.05

Table S5 Model results for yield prediction

	Training	Testing	Validation
<i>D30</i>			
MAE	5.142	8.871	6.761
SD	6.293	10.956	7.969
R ²	0.947	0.869	0.785
RMSE	6.12	10.87	7.58
<i>D60</i>			
MAE	6.298	13.692	8.276
SD	7.54	12.071	11.491
R ²	0.893	0.792	0.714
RMSE	8.22	14.96	11.27
<i>D90</i>			
MAE	5.578	7.486	7.056
SD	7.328	8.18	8.724
R ²	0.891	0.935	0.830
RMSE	7.20	7.76	9.38
<i>D120</i>			
MAE	4.381	8.468	7.543
SD	6.875	10.431	10.145
R ²	0.876	0.960	0.893
RMSE	6.91	9.44	10.14

Table S6 Model results for quality prediction

	Training	Testing	Validation
<i>D30</i>			
MAE	0.075	0.089	0.064
SD	0.09	0.102	0.08
R ²	0.709	0.815	0.729
MSE	0.09	0.09	0.08
<i>D60</i>			
MAE	0.107	0.085	0.092
SD	0.123	0.097	0.109
R ²	0.762	0.832	0.867
MSE	0.12	0.10	0.11
<i>D90</i>			
MAE	0.032	0.038	0.065
SD	0.048	0.044	0.115
R ²	0.906	0.904	0.760
MSE	0.05	0.05	0.11
<i>D120</i>			
MAE	0.026	0.113	0.041
SD	0.049	0.092	0.051
R ²	0.933	0.962	0.762
MSE	0.08	0.05	0.05

D30

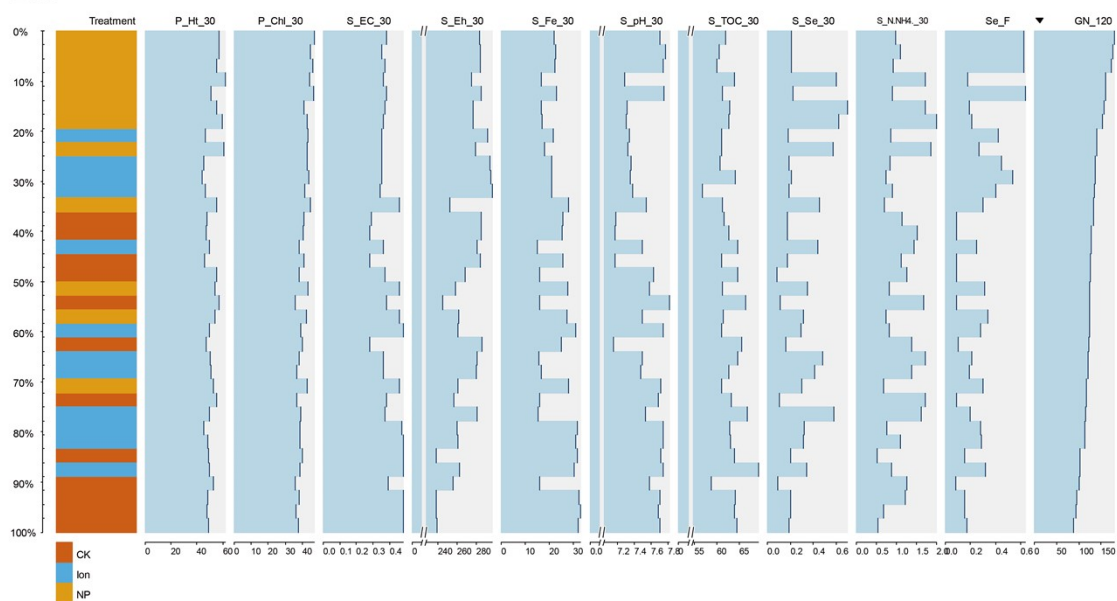


Fig. S1. Visualization of the raw data distribution at D30. Samples are arranged in descending order of the rice yield (GN_120).

D60

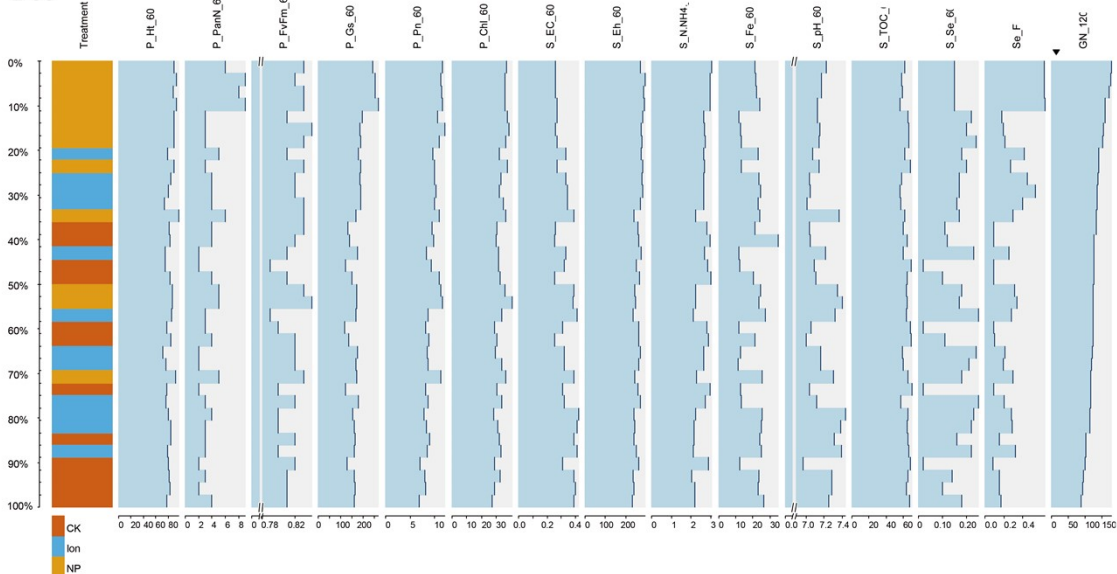


Fig. S2. Visualization of the raw data distribution at D60. Samples are arranged in descending order of the rice yield (GN_120).

D90

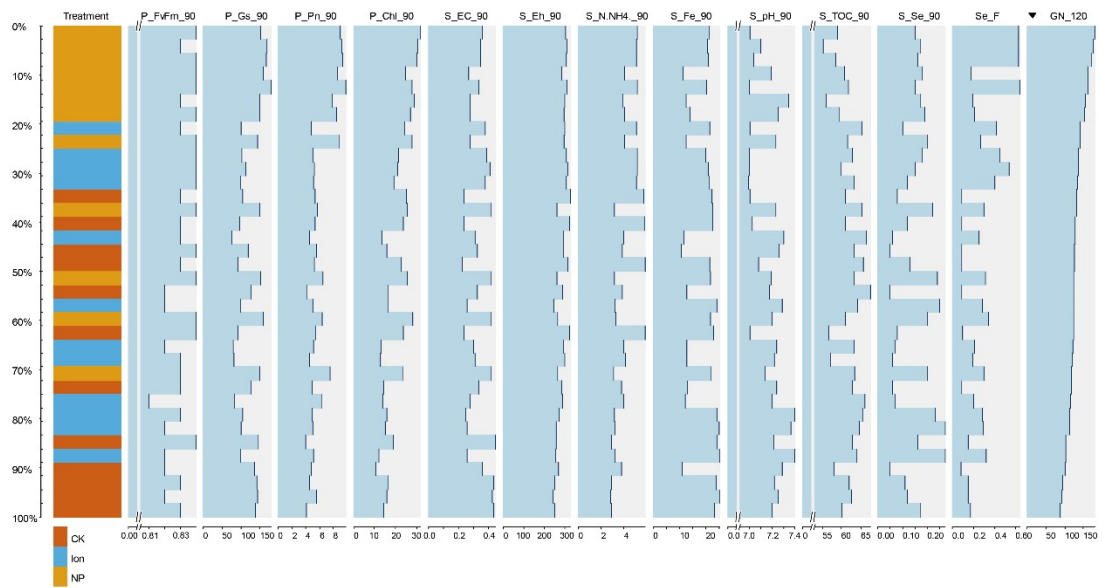


Fig. S3. Visualization of the raw data distribution at D90. Samples are arranged in descending order of the rice yield (GN_120).

D120

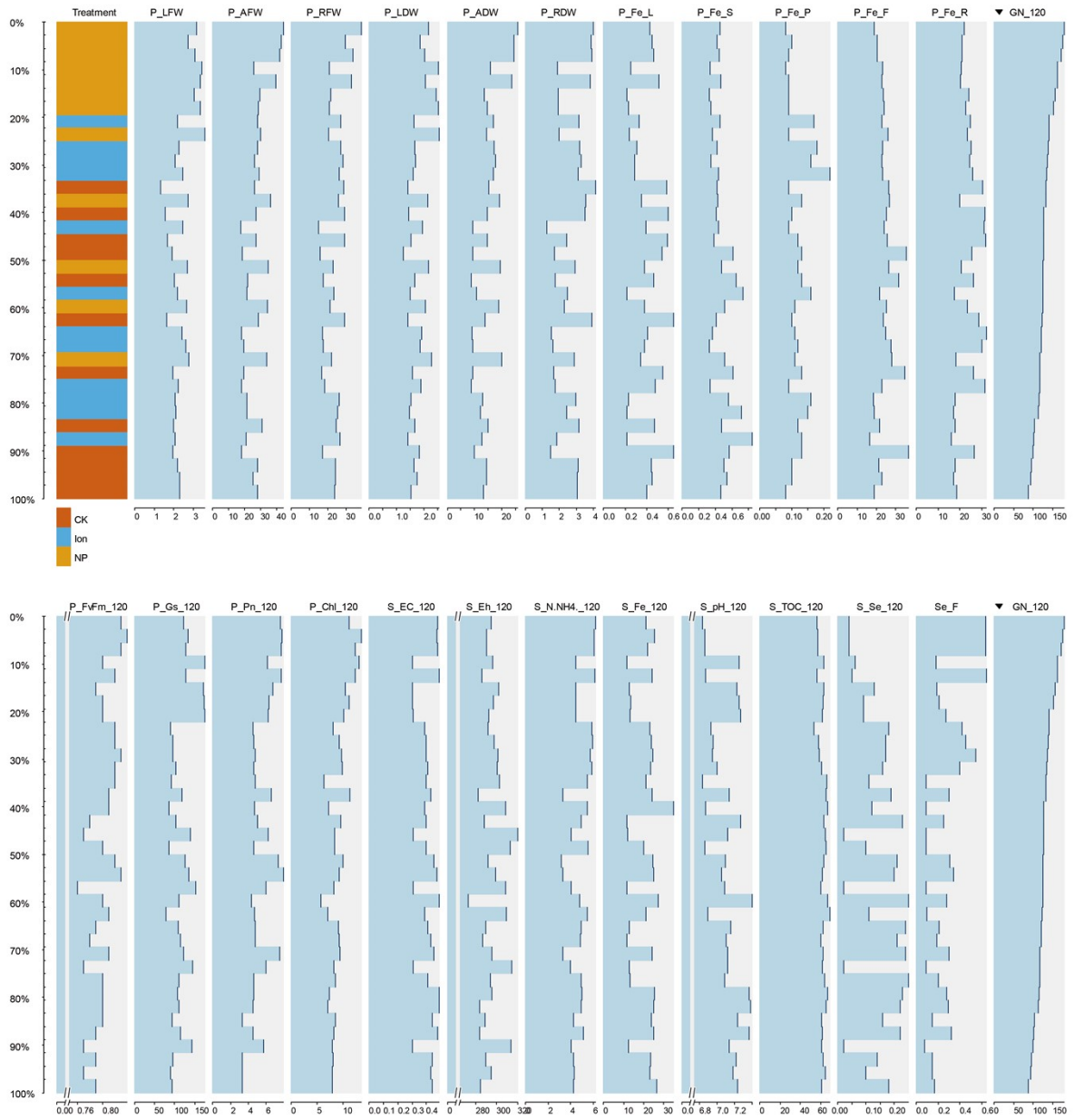


Fig. S4. Visualization of the raw data distribution at D120. Samples are arranged in descending order of the rice yield (GN_120).

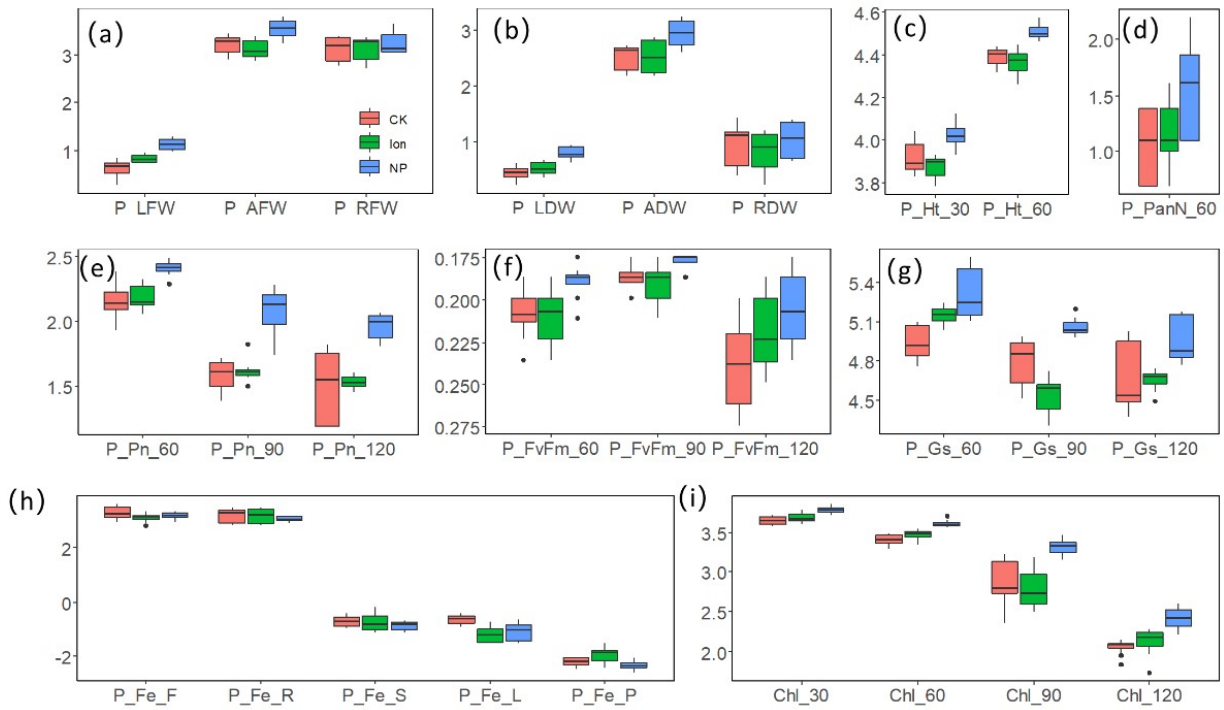


Fig. S5. Boxplot of plant physiological data. To better display the data relationships, the data were Ln transformed.

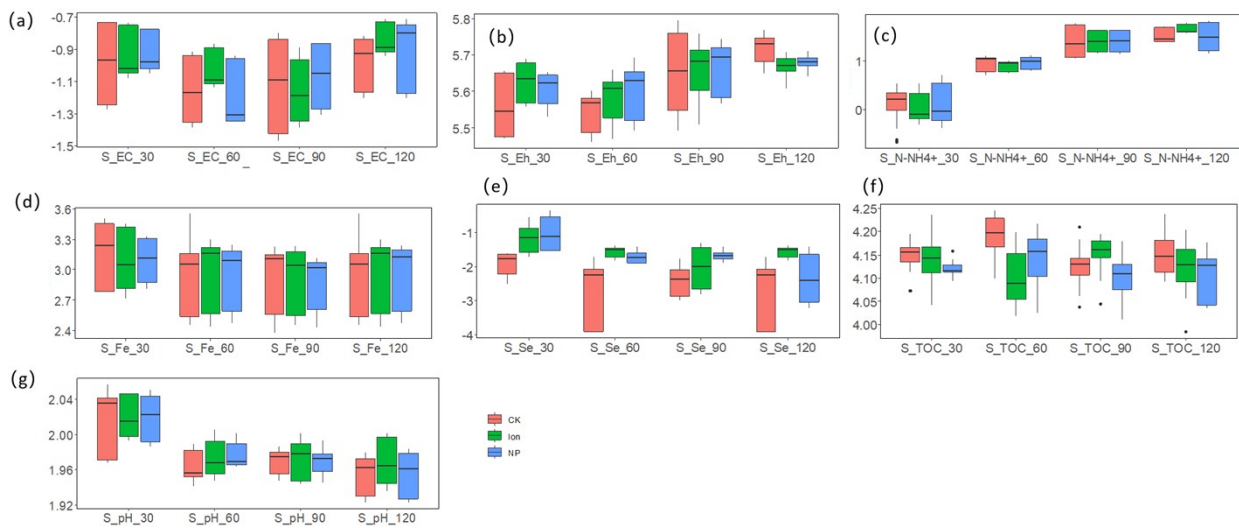


Fig. S6. Boxplot of soil properties data. To better display the data relationships, the data were Ln transformed.

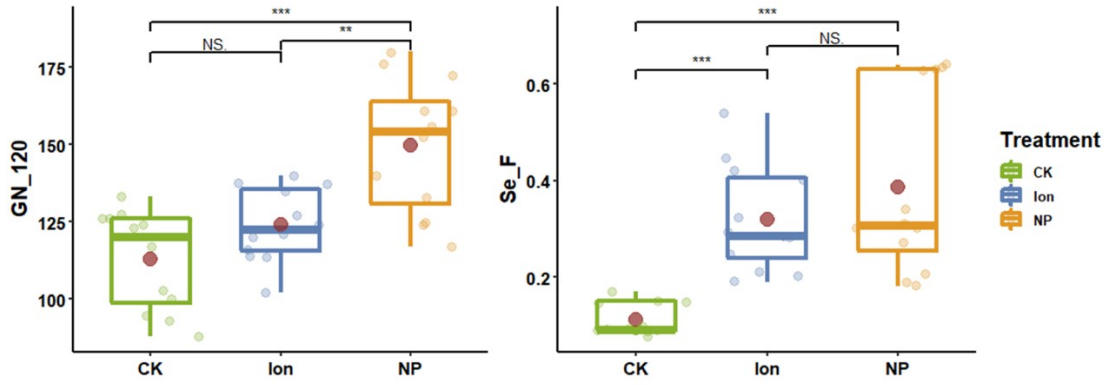


Fig. S7 the effect of treatments on rice yield (left) and quality (right).

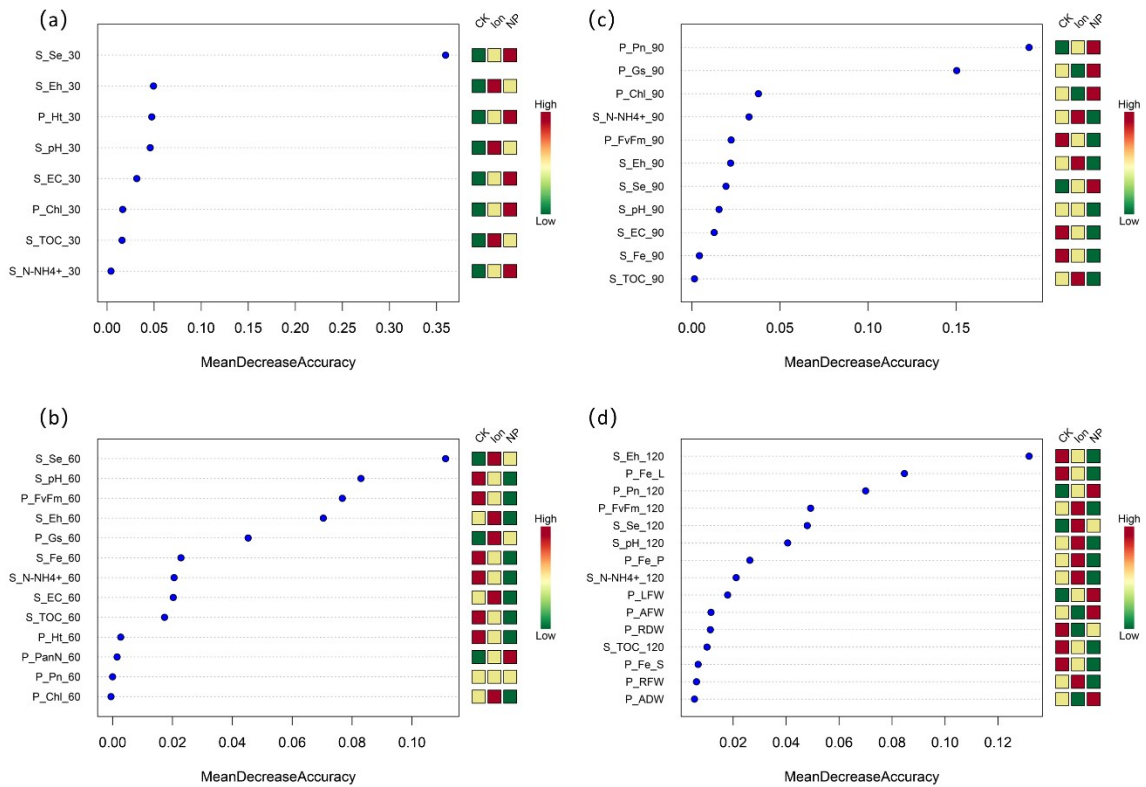


Fig. S8. A random forest (RF) approach was used to select the highest contribution to the variance in the different treatment groups on (a) day 30, (b) day 60, (c) day 90, (d) day 120. The horizontal coordinate is the average importance and the vertical coordinate is the plant and soil index, and the heatmap shows the relative differences between different groups.

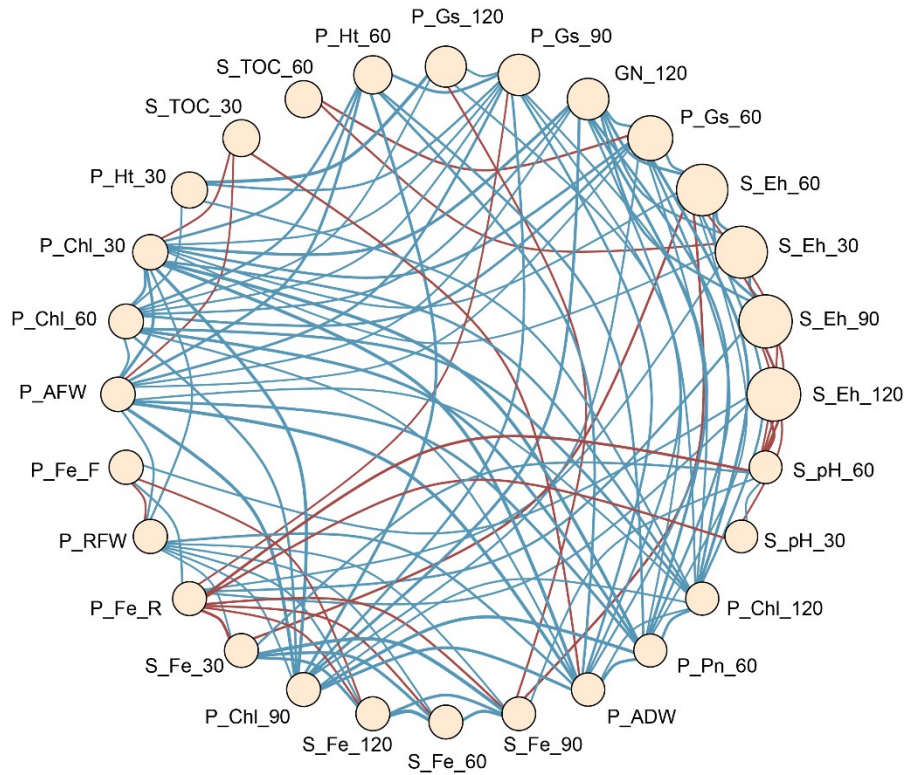


Fig. S9. Co-occurrence networks of plant and soil features. Nodes represent the basic units of a network. Light blue lines represent a significant positive correlation and red lines represent a significant negative correlation.

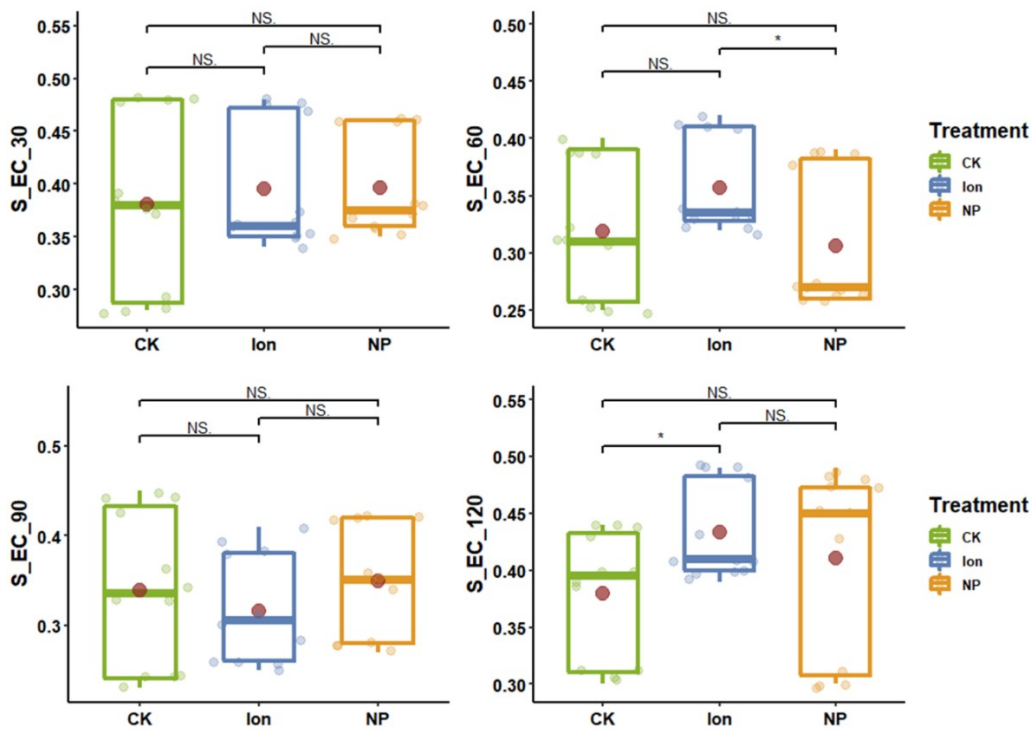


Fig. S10 the effect of treatments on soil Ec.

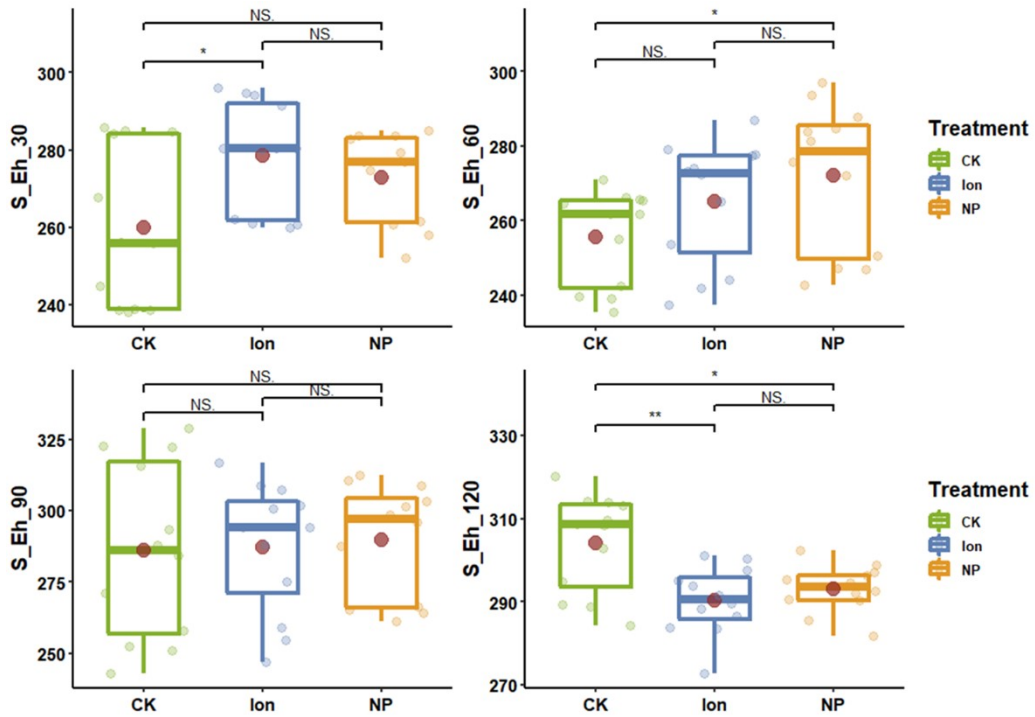


Fig. S11 the effect of treatments on soil Eh.

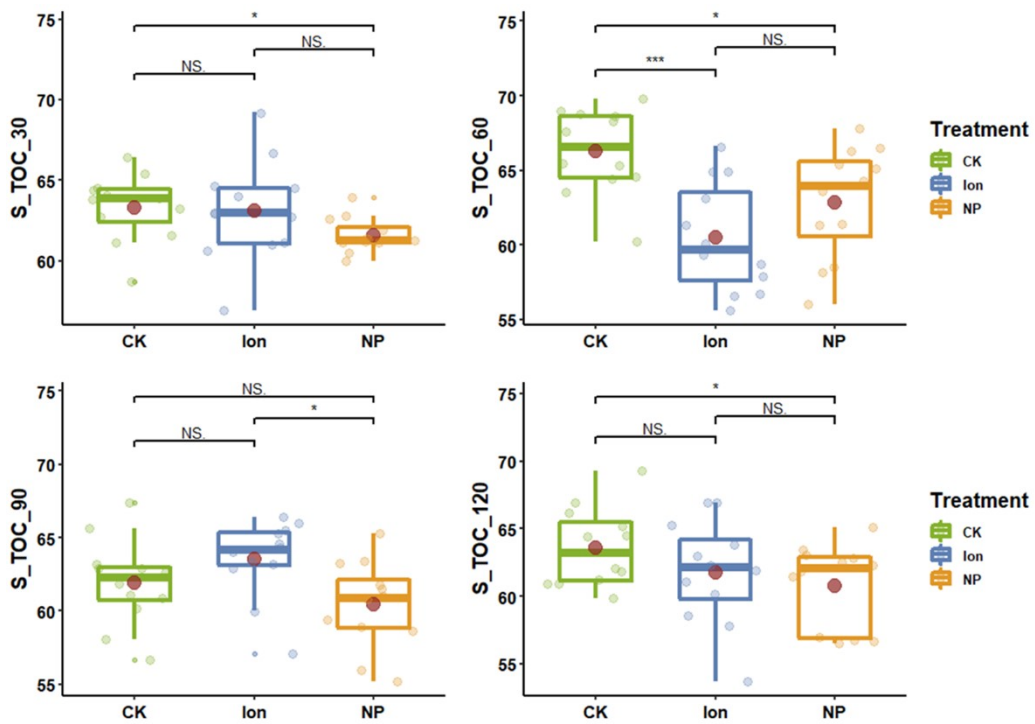


Fig. S12 the effect of treatments on soil TOC.

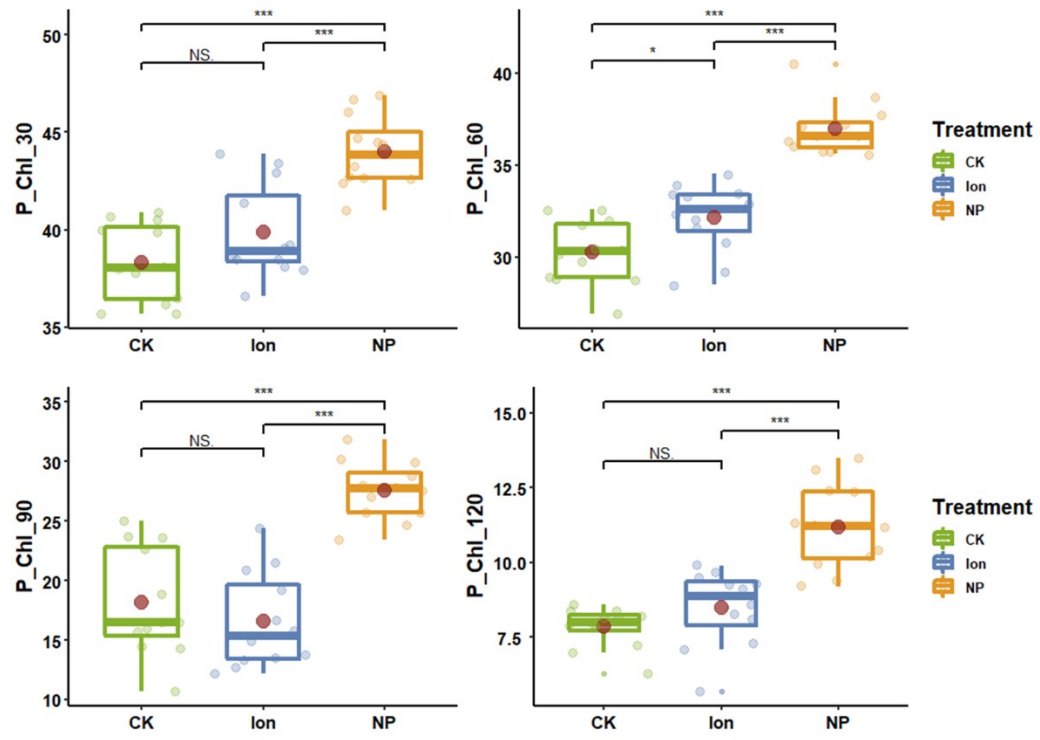


Fig. S13 the effect of treatments on Chl.

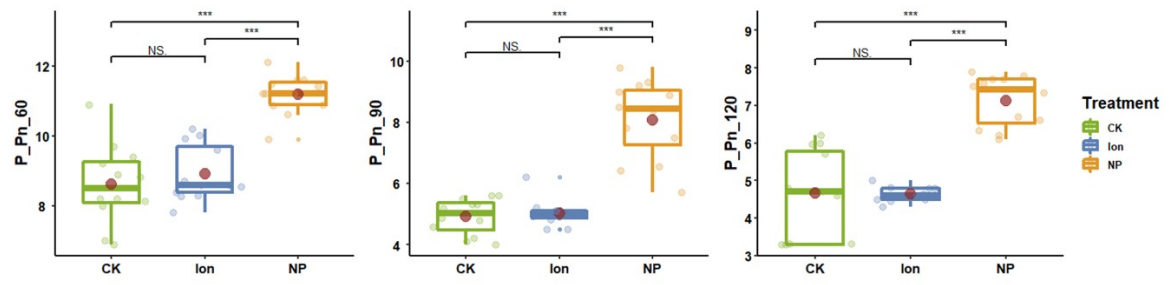


Fig. S14 the effect of treatments on Pn.

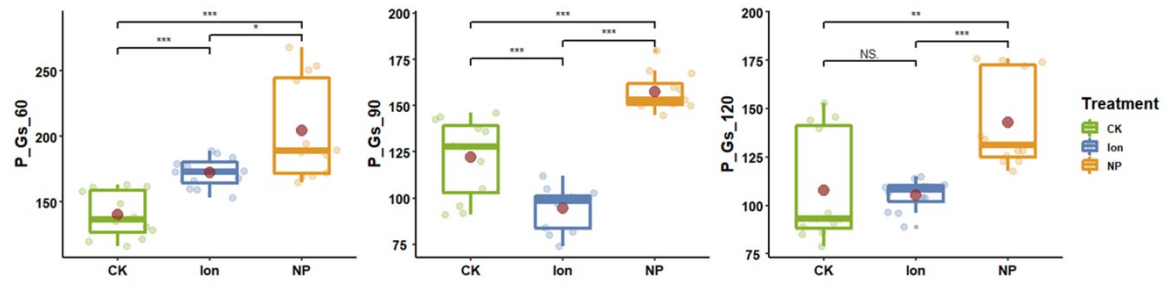


Fig. S15 the effect of treatments on Gs.

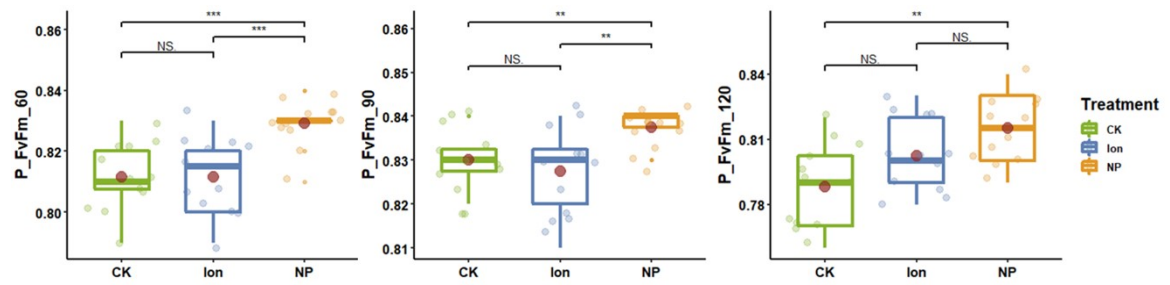


Fig. S16 the effect of treatments on Fv/Fm.