

Supplementary Information

Point-of-Use Sensors and Machine Learning Enable Low-Cost

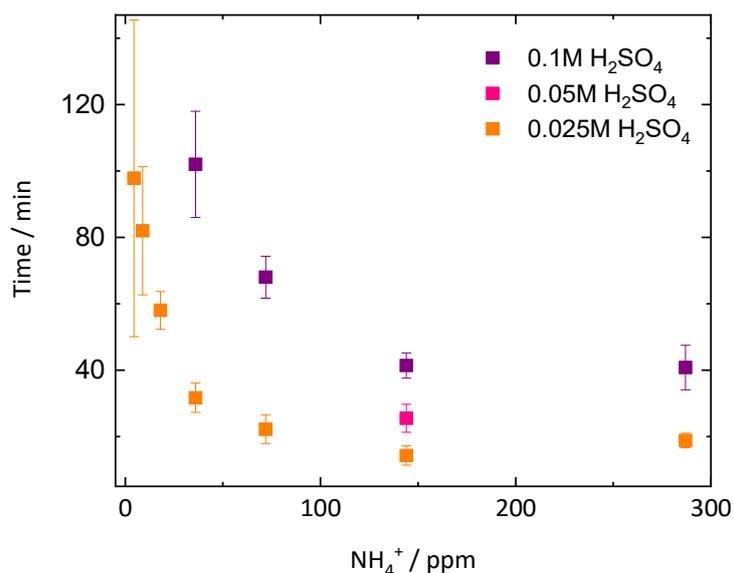
Determination of Soil Nitrogen

Max Grell, Giandrin Barandun, Tarek Asfour, Michael Kasimatis, Alex Collins, Jieni Wang, Firat Güder*

Department of Bioengineering, Imperial College London, London SW7 2AZ, UK

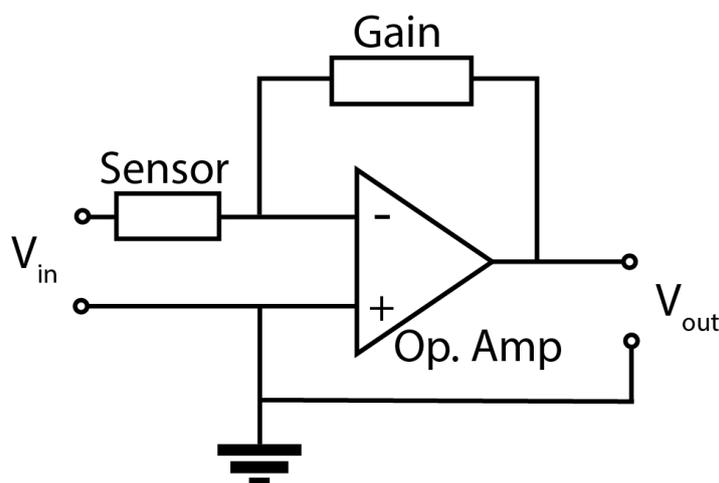
* Corresponding Author

Dr. F. Güder, e-mail: guder@ic.ac.uk



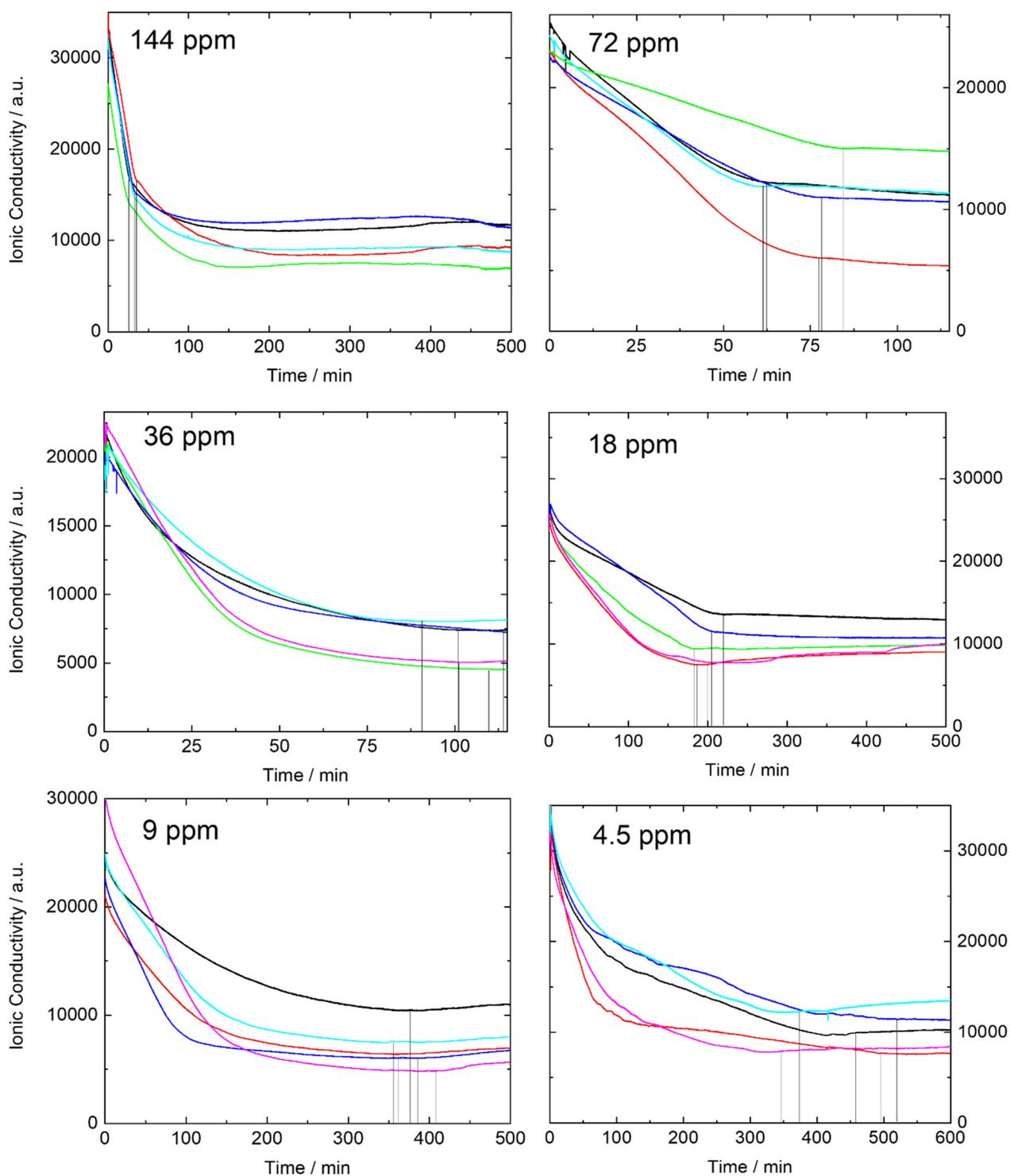
Supplementary Figure 1

We measured calibration curves for NH_4NO_3 in water (no soil), to verify our soil measurements are from NH_4^+ alone. On the chemPEGS we tested 3 concentrations of H_2SO_4 functionalization. Higher concentration H_2SO_4 takes longer to neutralize by the same amount of NH_4^+ . The $10\mu\text{l}$ 0.025M H_2SO_4 functionalization was the best compromise between precision and measurement time.



Supplementary Figure 2

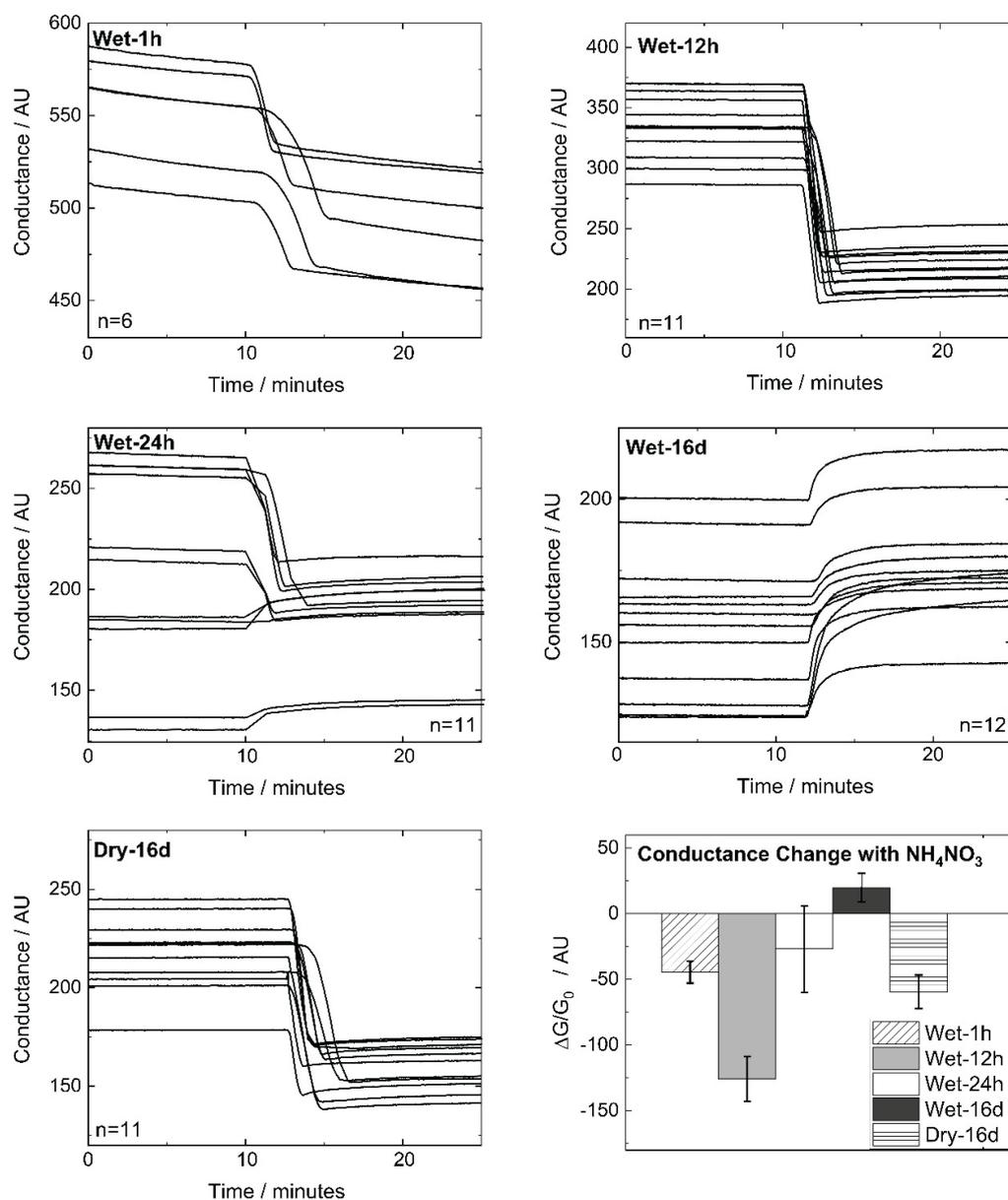
Schematic showing home-made electronics for measuring ionic impedance during neutralization on chemPEGS (Sensor). An alternating voltage was supplied (V_{in}) across the chemPEGS, and the current passing through was measured as a voltage (V_{out}) with a transimpedance amplifier (implemented with an Operational Amplifier), amplified with a resistor (Gain).



Supplementary Figure 3

Raw data showing decrease in ionic conductivity (measured as a voltage), as H_2SO_4 in the paper scrubber is neutralized by NH_3 gas. The moment at which the rate of neutralization slows is recorded (marked in vertical black lines on the Time axis). The time used as analytical signal was A) when the gradient was zero or B) when there was a step change in gradient ($>80\%$ change in <10 minutes),

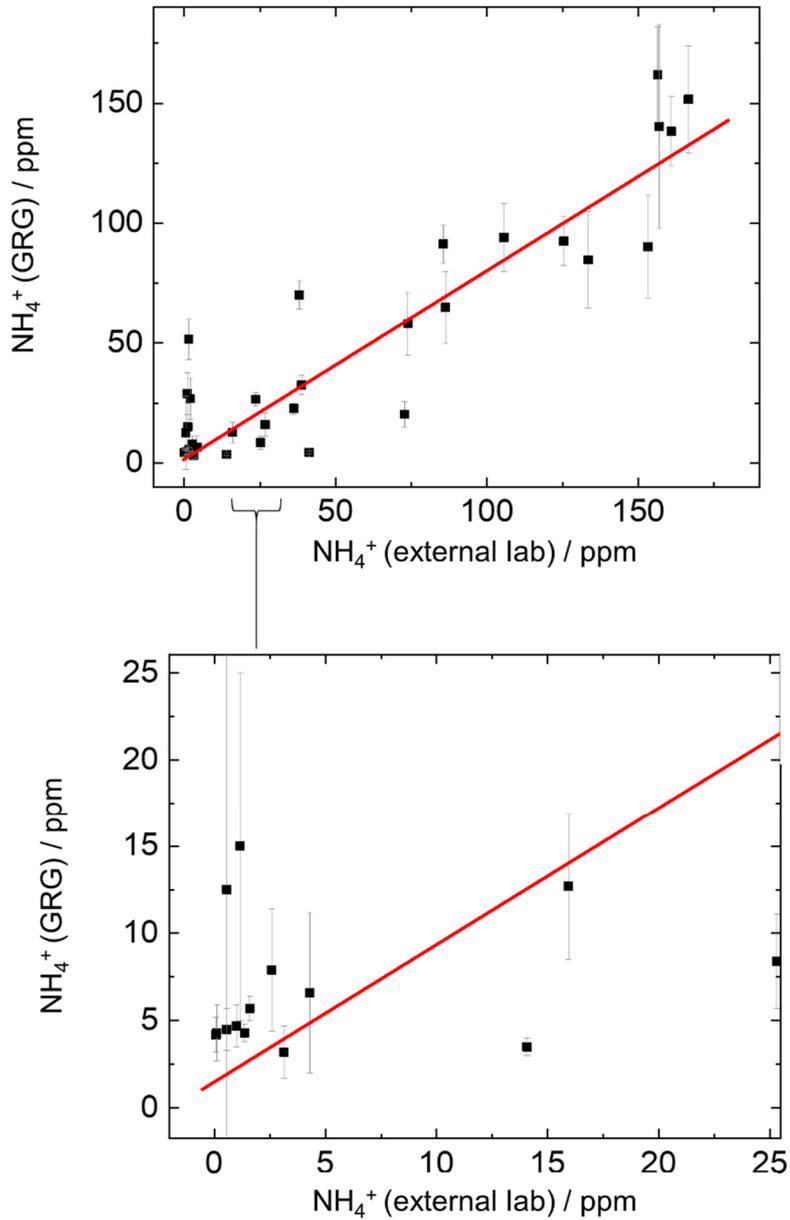
which ever happened sooner. Higher concentrations ($>144\text{ppm}$) are not adequately separated, but could be with higher concentration of H_2SO_4 in the chemPEGS scrubber, however, this would increase measurement time (see **Supplementary Figure 1**).



Supplementary Figure 4

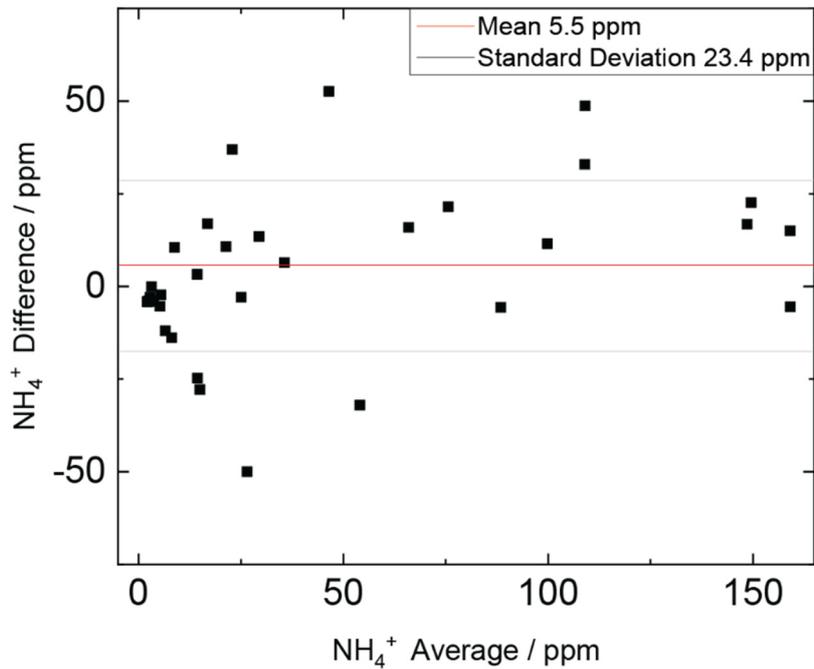
To test the stability of chemPEGS after storage up to 16 days, we have performed the following experiments: chemPEGS were prepared by adding $10\ \mu\text{l}$ 0.025M H_2SO_4 to the paper sensors and placed inside five separate 1-liter boxes. The first three boxes (named Wet-1h, Wet-12h and Wet-24h) contained a solution of $90\ \text{ml}$ $1.7\ \text{M}$ NaOH . The fourth box (Wet-16d) contained only $100\ \text{ml}$ of water

and fifth (Dry-16d) contained nothing (*i.e.* only air). For Wet-1h, Wet-12h and Wet-24h, the sensors were kept in the boxes for 1h, 12h and 24h and a 65.5 mM 10 ml solution of NH_4NO_3 was injected to measure the impact of time on freshly prepared chemPEGS over a short storage period. For Wet-16d and Dry-16d, the boxes were opened at day 16 and the chemPEGS were moved into boxes containing a solution of 90 ml of 1.7 M NaOH and kept for another 24h in the new boxes before injecting a 65.5 mM 10 ml solution of NH_4NO_3 into the boxes. The results of the individual measurements are shown above, in addition to a bar plot (right bottom) summarizing the experiments in percent change in conductance before and after injection of the NH_4NO_3 solution. When the sensors function normally, injection of NH_4NO_3 solution should cause a drop in conductance over time hence the percent change with respect to the baseline should be negative. While for Wet-1h, Wet-12h, we observed the correct behavior (although the magnitude of the response was different, indicating for the purposes of calibration, time is important), Wet-24h showed mixed behavior (*i.e.* some sensors showed a negative and others positive percent change in conductance). This result indicates that, after 24h, freshly prepared sensors no longer produce reliable results, perhaps due to the background reaction between H_2SO_4 and cellulose (or other elements in the sensor) that depletes the H_2SO_4 added. None of the sensors stored under humid conditions for 16 days (Wet-16d) functioned correctly, which was clearly not the case for the sensors stored dry (Wet-16d). In fact, the dry stored sensors performed better than the freshly prepared sensors stabilized for 24h. We can, therefore, conclude that the dry stored chemPEGS remain active for at least 16 days while freshly prepared sensors should be used within 12 hours after insertion into the cartridge.



Supplementary Figure 5

After calibration shown in **Figure 2** in the main text with data from **Supplementary Figure 3**, 35 soil- NH_4^+ measurements were made in soil fertilized with NH_4NO_3 for a variety of weather conditions (see **Figure 3** in the main text). Soil- NH_4^+ measurements from our gas-phase NH_4^+ sensor (GRG) compared to external laboratory measurements with a score of $R^2 = 0.85$.



Supplementary Figure 6

This Bland-Altman plot of data in **Supplementary Figure 5** shows an average bias of 5.5 ppm, meaning the external laboratory measurements were on average 5.5 ppm higher than the electrical PoU sensor measurements of soil-NH₄⁺. The bias is slightly stronger at higher concentrations.

Calculation of Fertilization Rate [kg/ha] from Concentration [ppm] and Soil Parameters

$$\text{Fertilization Rate [kg/ha]} = \text{Hectare Area [m}^2\text{]} \times \text{Sample Depth [m]} \times \text{Density [kg/m}^3\text{]} \times \text{Concentration [ppm]}$$

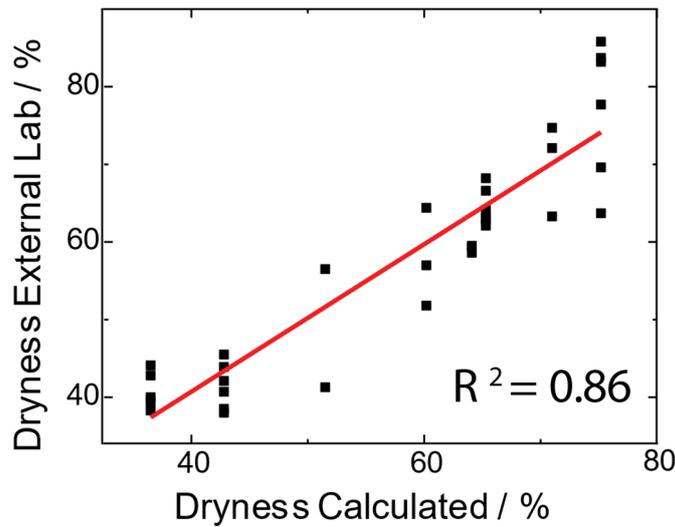
$$\text{Hectare Area} = 10,000 \text{ m}^2$$

$$\text{Sample Depth} = \text{Pot height} = 0.26 \text{ m}$$

$$\text{Soil Density} = 774 \text{ kg/m}^3$$

$$\text{Concentration of NH}_4\text{NO}_3 = 120 \text{ ppm}$$

$$10,000 \times 0.26 \times 774 \times (120/1,000,000) = 241 \text{ kg/ha}$$



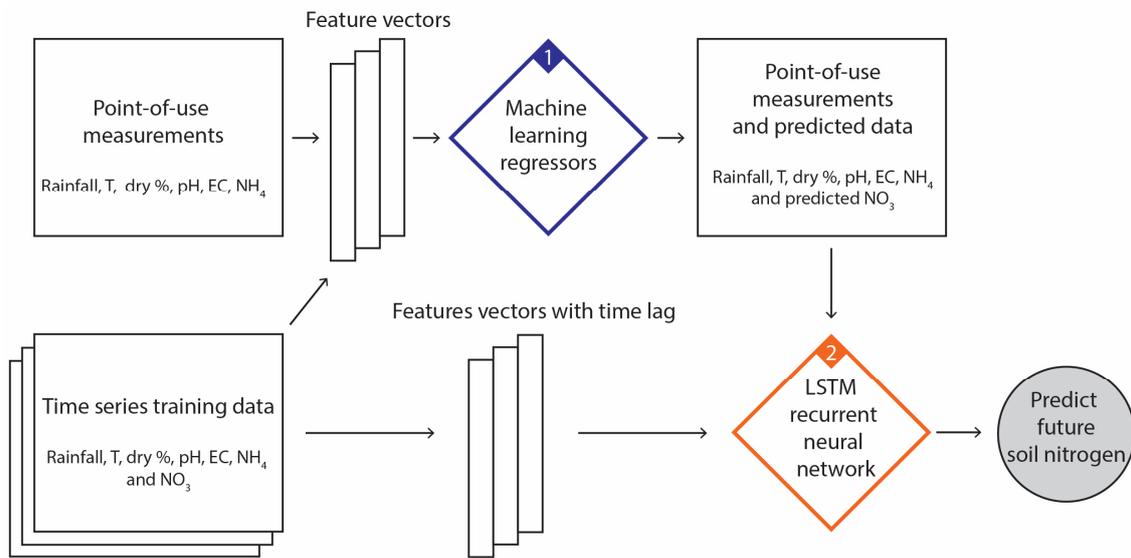
Supplementary Figure 7

In our dataset, dryness is highly correlated with rainfall and temperature. Dryness can be predicted using linear regression with $R^2 = 0.86$ using the equation below, and hence can be estimated using these two metrics without needing further analytical measurements.

$$\text{Dryness}[\%] = 0.8853 \text{ Temperature}[\text{°C}] - 3.0373 \text{ Rainfall}[\text{mm}] + 49.3928$$

Discussion of the discrepancy between our measurements and external laboratory measurements of EC and pH

The EC and pH measurements performed in our laboratory and externally did not correlate to the degree we expected, although the instruments used in our laboratory were calibrated weekly with calibration solutions to produce reliable measurements. Upon investigation, we found out that the difference in sample preparation was the likely culprit behind differences in the results. The external laboratory dried the soil samples before taking a fixed weight and mixing with water for measurements, whereas we took the samples directly from the pots without drying and mixed with water, which caused varied values for EC and pH. In any case, in the context of this work, these differences in sample preparation did not affect the underlying trends in the data generated by the external laboratory and such small errors may happen under real experimental conditions at the point-of-use (hence the entire system should be robust enough to absorb these errors).



Supplementary Figure 8

Machine learning models trained with readily available environmental data and PoU measurements enable prediction of difficult-to-measure soil nitrogen instantaneously with machine learning regressors (**Supplementary Figure 8.1**). Prediction into the future is then enabled by treating data as time series in a LSTM neural network (**Supplementary Figure 8.2**).

Data Processing for Machine Learning

The following steps were taken to predict instantaneous soil-NO₃⁻ (**Supplementary Figure 8.1**)

1. Anomalous data removed
2. Data for missing days interpolated smoothly with Akima spline
3. Time series length capped at 16 days
4. Data normalized over range 0-1
5. Feature selection: Features ranked for predicting NO₃⁻ calculated by weight and gain using XGBoost – see **Figure 4** in the main text. Combinations are tested starting with all 7 features, then only 6 most important features, then the most important 5, down to the single most important feature

6. Cross-validation: Data from each time series is removed from the training data set sequentially, and used as test data instead
7. Regressor selection: Random Forest, Gradient Boosting, Adaboost, Extra Trees, Knn and XGBoost were tested
8. Hyperparameter tuning: Grid search used to tune up to 3 hyperparameters for each regressor

The following steps were taken to predict soil-NH₄⁺ and soil-NO₃⁻ 1-12 days into the future

(Supplementary Figure 8.2)

1. Concatenate all time series into one multivariate time series
2. Normalize data $-1 < x < 1$ (for a bounded and stable neural network computation)
3. Format time series data for supervised learning (generate additional input features with time lag, generate output, which is soil-NH₄⁺ and soil-NO₃⁻ at future time)
4. Remove each time series sequentially and train model on remaining data
5. Predict the removed time series from its measurements on Day 0
6. Create LSTM model and tune hyperparameters with grid search (time lag, epochs, batch size, number of neurons)
7. Record score (mean squared error) of predictions, repeat each prediction 7 times and compare average scores to determine optimal tuning, train LSTM model with optimal tuning to generate predictions of NH₄⁺ and NO₃⁻ into the future

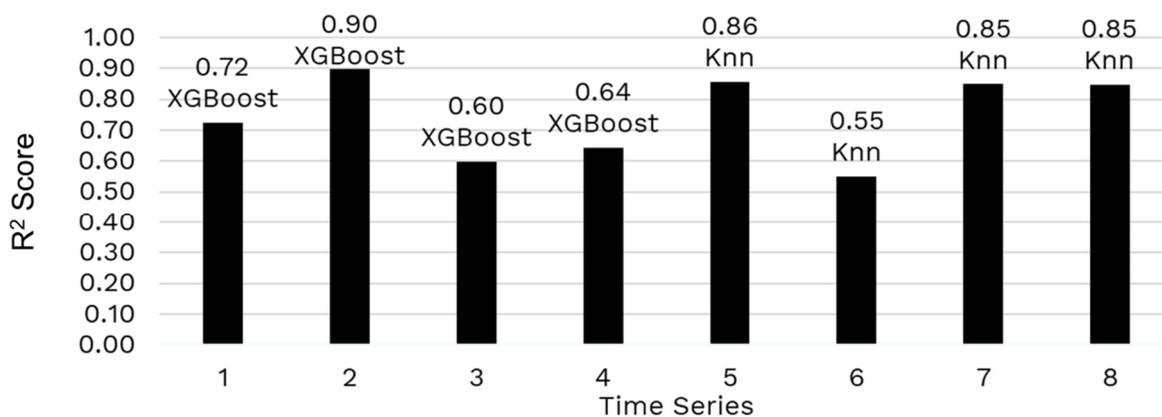
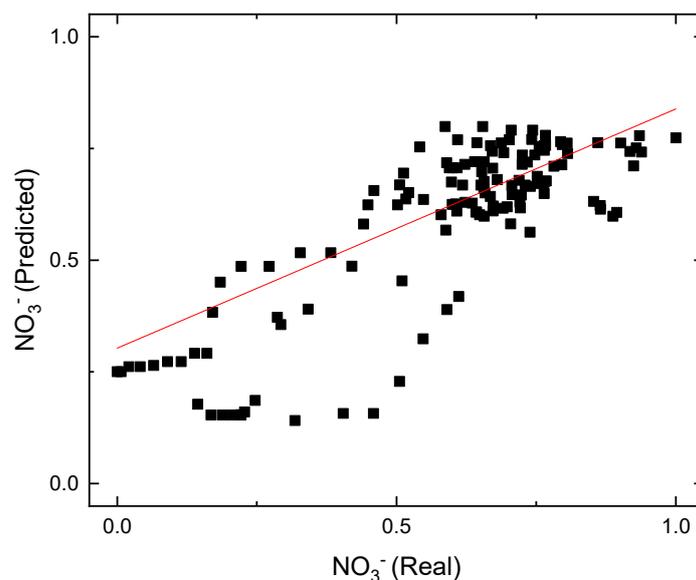


Figure Supplementary 9

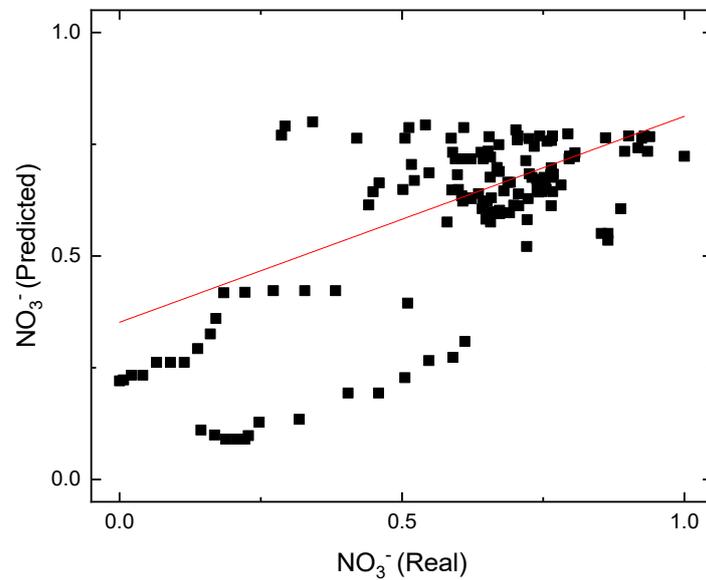
Each time series 1-8 (corresponding to soil under different environmental conditions, by controlling rainfall and temperature - see **Figure 3** in the main text) was sequentially removed from the data, and predicted using the remaining data. Testing all combinations of features, regressors and tuning parameters results in a best-case prediction score for each time series, shown here. XGBoost is gives the best predictions for dryer soils and Knn for wetter soils. Poorer NO_3^- predictions at moderate (3-5mm) rainfalls, where NO_3^- levels were relatively constant over time. The number of features ($n_features$), regressor and optimal tuning parameters for each time series are listed below.

1. $n_features = 5$, XGBoost ($\eta = 0.3$, $\gamma = 0.0$, $\text{max depth} = 6.0$)
2. $n_features = 3$, XGBoost ($\eta = 0.3$, $\gamma = 0.0$, $\text{max depth} = 6.0$)
3. $n_features = 4$, XGBoost ($\eta = 0.3$, $\gamma = 0.0$, $\text{max depth} = 7.0$)
4. $n_features = 7$, XGBoost ($\eta = 0.3$, $\gamma = 0.0$, $\text{max depth} = 1.0$)
5. $n_features = 4$, Knn ($n\ \text{neighbors} = 8.0$, $\text{leaf size} = 1.0$, $p = 1.0$)
6. $n_features = 7$, Knn score ($k = 50.0$, $\text{leaf size} = 1.0$, $p = 3.0$)
7. $n_features = 7$, Knn score ($k = 14.0$, $\text{leaf size} = 1.0$, $p = 1.0$)
8. $n_features = 6$, Knn score ($k = 20.0$, $\text{leaf size} = 1.0$, $p = 10.0$)



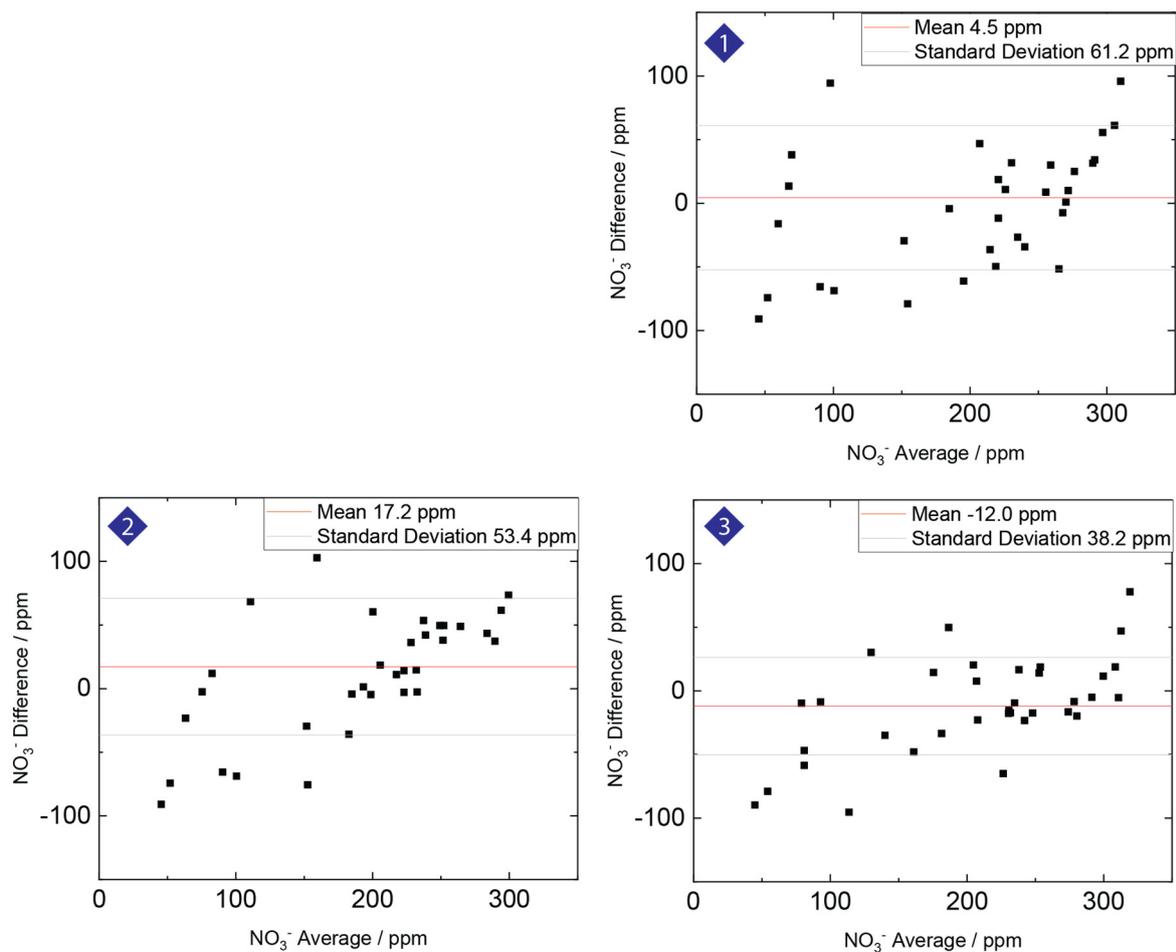
Supplementary Figure 10

The same process was followed as **Figure 4a** (bottom left) in the main text, with the same tuning parameters ($k=14$, leaf size=1, $p=1$), but test inputs were from the external lab rather than PoU sensors from our lab, thus removing the impact of inaccuracy from our lab. The resulting score is $R^2=0.68$ (compared to the $R^2=0.63$ with our lab inputs).



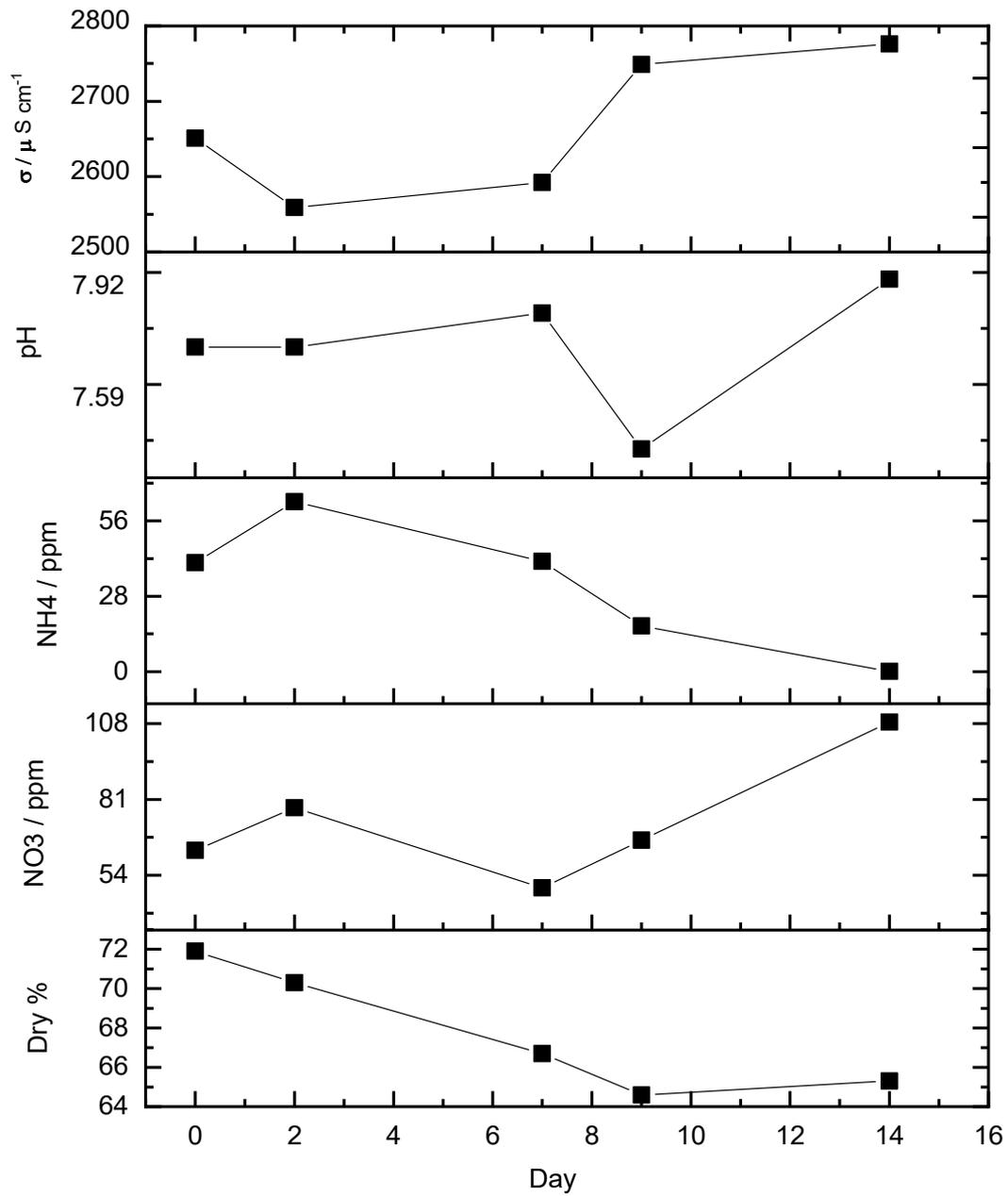
Supplementary Figure 11

Knn predicts NO_3^- using a model tuned ($k=11$, leaf size=3, $p=20$) to only the most basic inputs – days since fertilization, rainfall and temperature (*i.e.* requiring no soil sensors at all.) with $R^2 = 0.54$. One model was used for all environmental conditions.



Supplementary Figure 12

Bland-Altman plots **Supplementary Figure 12.1** and **Supplementary Figure 12.2**, corresponding to **Figure 4a** top right and **Figure 4a** bottom left in the main text respectively, have a positive bias. This means the external laboratory measurements of soil- NO_3^- were greater than the predictions by the Knn models. **Supplementary Figure 12.3**, corresponding to **Figure 4a** bottom right in the main text, has a negative bias, and tighter standard deviation resulting from more precise predictions from the Knn model. The predictions from **Figure 4a** top right in the main text, using PoU measurements and the same model for all weather conditions, were the truest. This is shown in the Bland-Altman plot **Supplementary Figure 12.1**, which has the smallest mean difference of 4.5 ppm.



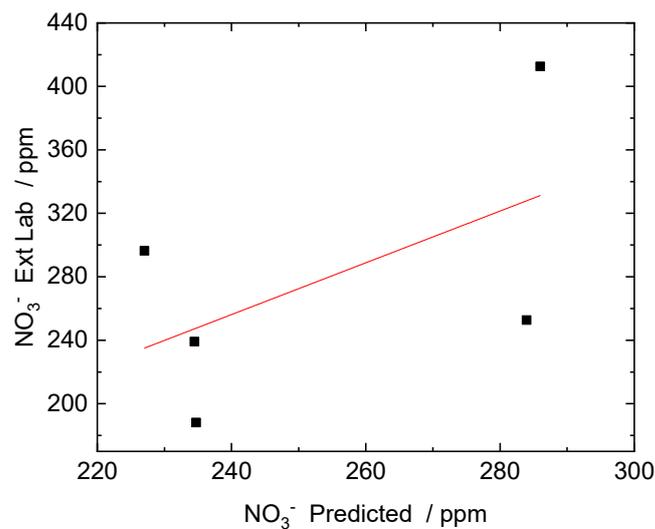
Supplementary Figure 13

To assess suitability for different soil types, soil-NH₄⁺, NO₃⁻, pH and EC were measured in clay soil (1% sand, 44% silt, 55% clay) fertilized with NH₄NO₃. The clay soil received 3 mm/day rainfall at a temperature of 21±1 °C

Day	NH ₄ ⁺ (ppm)	NO ₃ ⁻ (ppm)	NH ₄ ⁺ Scaled (ppm)	NO ₃ ⁻ Scaled (ppm)	NO ₃ ⁻ Predicted (ppm)
0	40	63	154	239	243
2	63	78	240	296	240
7	41	50	156	188	284
9	17	67	65	253	294
14	0	109	0	413	286

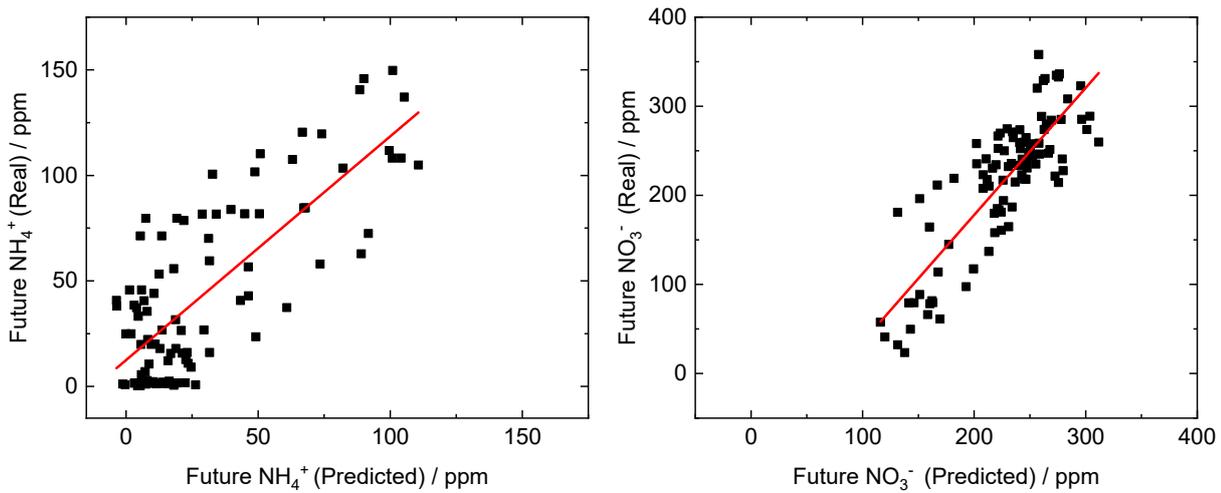
Supplementary Table 1

Soil-N concentrations at Day 0 are lower (NH₄⁺ 40 ppm, NO₃⁻ 63 ppm) than seen in the sandy loam soil (>100 ppm), since less NH₄NO₃ fertilizer was retained in the clay soil than expected. The model failed to give sensible with lower starting values combined with the new soil type. The entire time series was therefore scaled (multiplied by 3.8), bringing the starting soil-N values close (withing 3%) of the starting values for sandy loam soil in similar weather conditions (rainfall 3 mm/day and temperature 22±2 °C).



Supplementary Figure 14

Testing the Knn model (from **Figure 4a** top right in the main text) with the scaled values from **Supplementary Table 1** gives a weak positive correlation ($R^2 = 0.31$) between predicted and expected NO₃⁻ values in clay soil. More data for different soil types and fertilization rates would improve performance.

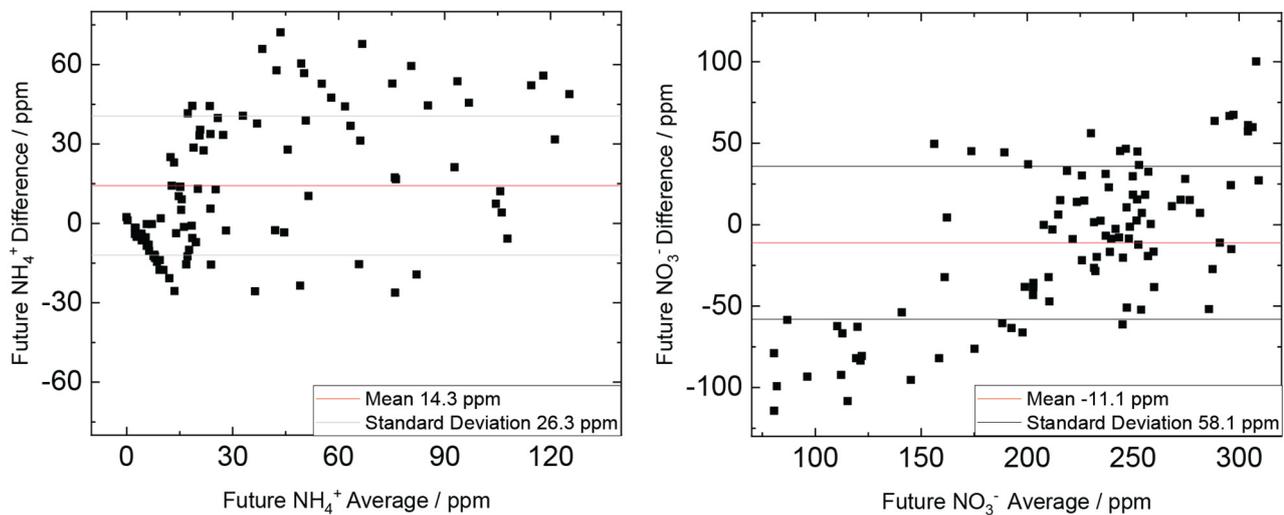


Supplementary Figure 15

R² plots for data in **Figure 4b** in the main text, showing soil-NH₄⁺ and soil-NO₃⁻ predicted by LSTM ML model 1-12 days into the future with R² = 0.60 and 0.70 respectively, using the same model for all environmental conditions.

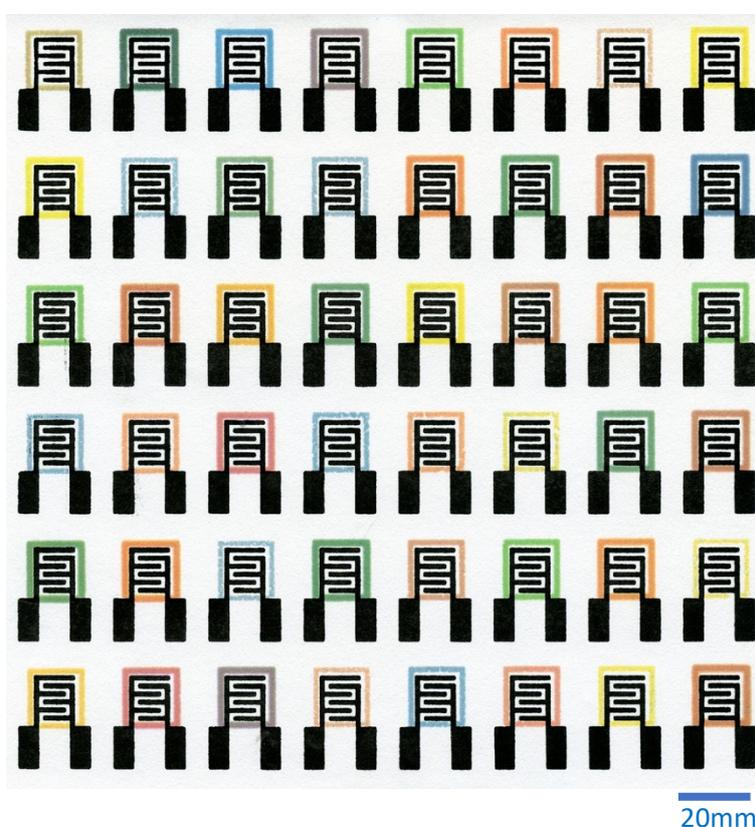
The hyperparameters tuned were the following

1. Number of neurons – the amount of nodes that comprise the neural network. The nodes are mathematical functions of the inputs (training data)
2. Batch size – the amount of data that passes through the LSTM model before it updates
3. Training epochs – the number of the times the LSTM model sees the entire training data set



Supplementary Figure 16

Bland-Altman plots for the LSTM predictions from **Supplementary Figure 13** show a fairly consistent bias of 14.3 ppm in the case of soil-NH₄⁺, meaning measurements from the external laboratory were on average 14.3 ppm higher than predicted by the LSTM neural network. For soil-NO₃⁻ there is a strong bias below 150 ppm, with the external measurements being lower than the LSTM predictions. The bias shifts at higher ppm, with external measurements being higher than LSTM prediction. The result is a flattening of the more extreme values by the LSTM neural network.



Supplementary Figure 17

Array of batch processed chemPEGS, comprising interdigitated carbon electrodes on chromatography paper, surrounded by a colorful printed wax barrier.