scientific reports

OPEN

Quantification of salt stress in wheat leaves by Raman spectroscopy and machine learning

Ibrahim Kecoglu[1,6], Merve Sirkeci[2,6], Mehmet Burcin Unlu[1,3], Ayse Sen[4], Ugur Parlatan[1✉] & Feyza Guzelcimen[5]

The salinity level of the growing medium has diverse effects on the development of plants, including both physical and biochemical changes. To determine the salt stress level of a plant endures, one can measure these structural and chemical changes. Raman spectroscopy and biochemical analysis are some of the most common techniques in the literature. Here, we present a combination of machine learning and Raman spectroscopy with which we can both find out the biochemical change that occurs while the medium salt concentration changes and predict the level of salt stress a wheat sample experiences accurately using our trained regression models. In addition, by applying different machine learning algorithms, we compare the level of success for different algorithms and determine the best method to use in this application. Production units can take actions based on the quantitative information they get from the trained machine learning models related to salt stress, which can potentially increase efficiency and avoid the loss of crops.

Wheat (*Triticum aestivum* L.) is one of the cereal products that can quickly adapt to environmental conditions. Besides, it is one of the most widely cultivated and produced herbal products globally thanks to the high amount of protein and carbohydrates it contains[1]. Wheat is a nutritious plant that meets 20–80% of the energy and protein needs of humans[2]. Turkey produced approximately 23 million tons of wheat in 2015. This production decreased to 20 million tons in 2018.

One of the significant reasons for this decrease is climate change, which affects the weather and plant growth via stressor factors such as salinity. Salinity occurs due to the accumulation of water-soluble salts in the upper part of the soil[3]. Salinity affects the morphology, anatomy[4], growth, root length, and osmotic pressure of the plants[5,6].

To determine the amount of malondialdehyde, chlorophyll, and carotenoids under optimum conditions and stress factors, biochemical methods are frequently used. The most commonly practiced technique in this type of biochemical analysis is UV-visible spectrophotometry. However, this method has some limitations in preparing, measuring, and analyzing a sample, which led scientists to seek alternatives. Recent studies utilized Raman spectroscopy to reveal the effects of salinity stress on the microstructure of wheat plants.

Raman spectroscopy is a vibrational spectral technique that gives information about the samples' internal molecular structure illuminated with a coherent source. Since inelastically scattered photons change their frequencies as much as the molecules' vibrational frequencies, the peaks in the spectra have intrinsic information about the sample, such as the type and amount of molecular bond vibrations[7,8]. Raman spectroscopy has a broad use in biology due to its advantageous properties like non-invasiveness and rapidness[9,10]. Schulz and Baranska documented the use of Raman spectroscopy in plant biology with a detailed list of tentative assignments[11]. Studies have been carried out using Raman Spectroscopy on many plants, including fennel fruits, chamomile clusters, turmeric roots, untreated summer wheat leaves, wheat grains, and fresh soybeans[12–15].

Raman spectroscopy provides rapid results with little or no sample preparation. Specifically, drought stress has already been investigated employing benchtop Raman spectroscopy[13,15,16] with the help of statistical methods.

In addition to the contributions mentioned above, the chemometric analysis helped intensely interpret the plant Raman signals[17–19]. Cebi et al. detected the amino acid L-Cysteine in the wheat floor using chemometric methods such as principal component analysis (PCA) and hierarchical cluster analysis (HCA)[17]. In a food

[1]Department of Physics, Bogazici University, 34342 Istanbul, Turkey. [2]Graduate School of Engineering and Science, Istanbul University, 34116 Istanbul, Turkey. [3]Global Station for Quantum Medical Science and Engineering, Global Institution for Collaborative Research and Education (GI-CoRE), Hokkaido University, Sapporo, Japan. [4]Department of Biology, Istanbul University, 34134 Istanbul, Turkey. [5]Physics Department, Istanbul University, 34134 Istanbul, Turkey. [6]These authors contributed equally: Ibrahim Kecoglu and Merve Sirkeci. ✉email: ugur.parlatan@boun.edu.tr
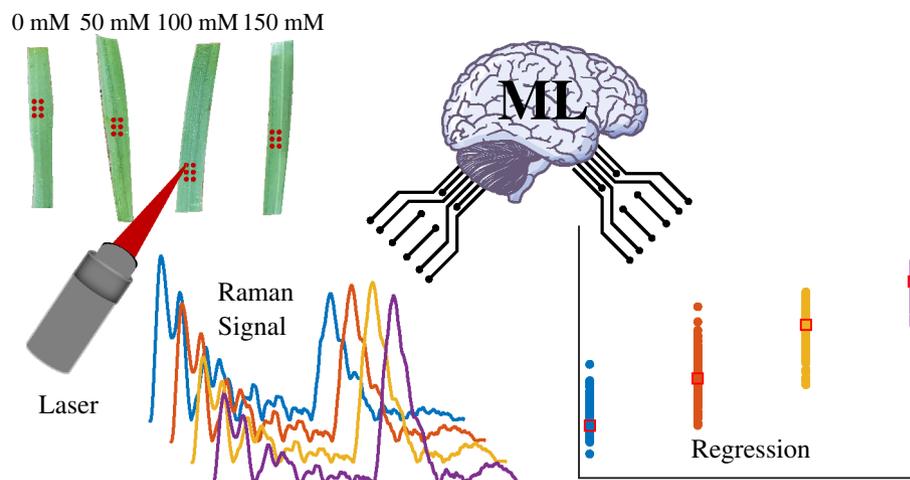
Figure 1. Outline of the Raman-based machine learning procedure for predicting salt stress.
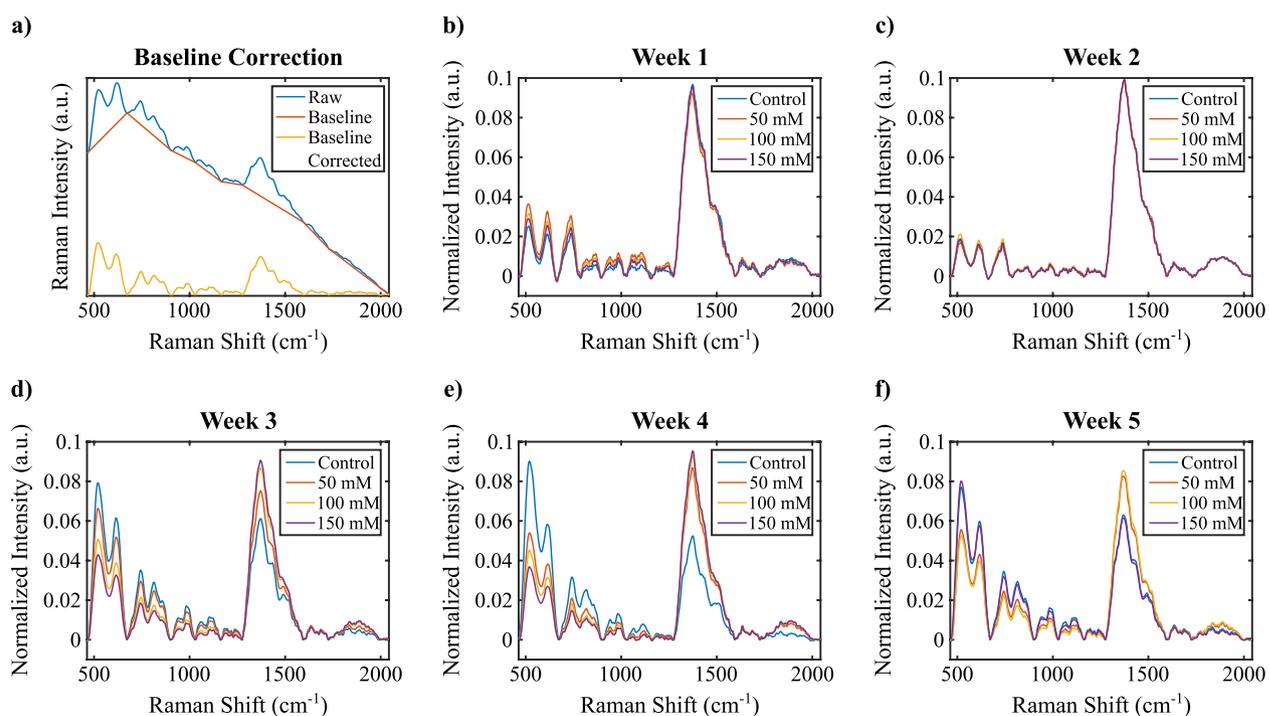
Figure 2. (**a**) Application of baseline correction on an example spectrum. (**b**–**f**) Averages of normalized Raman spectra of each week.

chemistry application, Raman spectroscopy and PCA were used to predict malting barley husk adhesion quality[19]. In another application, Farber et al. measured the pathogen Raman spectra from Maize Kernels and distinguished healthy and diseased states with 100% accuracy using PCA and discriminant analysis (DA)[18].

We employed a machine learning-supported Raman Spectroscopy approach to determine the amount of salt stress factor the plant was exposed to. The outline of the procedure of our study is shown in Fig. 1. Our study employed a range of machine learning algorithms and compared their training times, prediction speeds, and success rates.

Results and discussion

Biochemical analysis using Raman spectroscopy. To determine the biochemical change per week due to different levels of salt stress, we employed Raman spectroscopy. In Fig. 2 we compared the average of normalized Raman spectra corresponding to each medium salt concentrations (0, 50, 100, 150 mM) for each week. There was an expected general trend in the spectra regarding the concentration change.

Although this anticipated structure was observed in the period of maturity (Fig. 2c,d), there were some discrepancies in the early and late developmental periods. In the first 2 weeks (Fig. 2a,b) plants were not yet fully

2

Raman shift (cm^{-1})	Tentative assignments
522	ν(C–O–C), cellulose[21]
747	ν(C–O–H), pectin[22]
855	ω(C$_\epsilon$–H), serine[23]
1515	ν(C=C), carotenoid[24]
1563	Chlorophyll b[11]
1619	δ(H–N–H), methionine[25]; δ_a(NH$_3^+$), lysine[23]; ν(C=C), tyrosine
1646	δ(H$_2$O), water[22]; ν(C=C), ethylene[11]
1657	ν(C=O) + ω(N–H), amide I[11]
1669	ν(C=O) + ω(N–H), amide I[11]
1686	ν(C=O) + ω(N–H), amide I[11]

Table 1. Tentative assignments of Raman shift values. ν—symmetric stretching, ω—wagging, δ—symmetric bending, δ_a—antisymmetric bending.
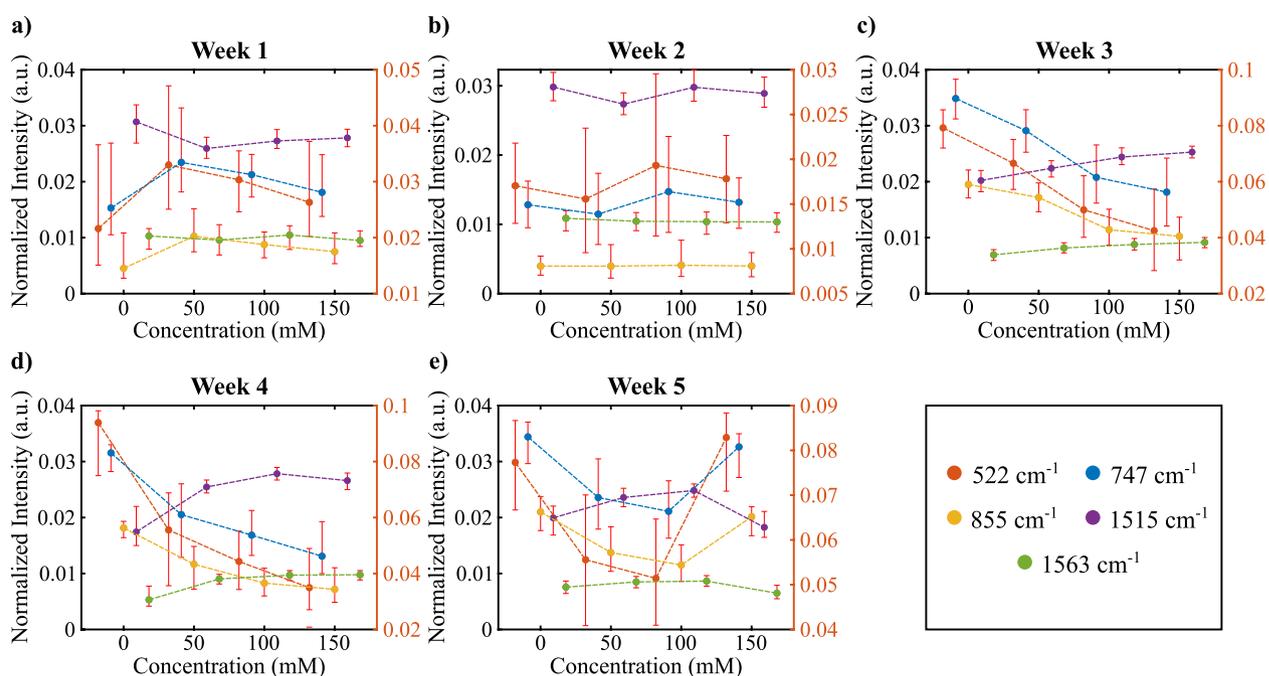
Figure 3. Normalized Raman intensity distributions of different concentrations for each week of plant growth. All colors correspond to different Raman shifts (cm^{-1}) which is given in bottom right.

grown and they showed some heterogeneity when we scanned different areas of the leaf sample. A similar but less prominent effect was also apparent in the fifth-week spectra (Fig. 2e). According to the studies, when plants are exposed to long-term salt stress, they give various responses at vegetative, physiological, and biochemical levels to adapt themselves to this situation and to continue their lives under stress conditions. These responses may differ between genotypes in short-term and long-term exposure to salt stress, as in the strawberry plant[20].

To get a close-up of the important biochemical changes, we plotted the trend of Raman intensities regarding 522 cm^{-1}, 747 cm^{-1}, 855 cm^{-1}, 1515 cm^{-1}, and 1563 cm^{-1} bands (Fig. 3). As given in Table 1, these bands correspond to cellulose, pectin, serine, carotenoid, and chlorophyll b respectively[11,21–24]. We observed that cellulose, pectin, and serine bands tend to decrease as the salt concentration of the medium increases, particularly for the period of maturity, which corresponds to weeks three to five. In week five, these bands have shown an increase for 150 mM concentration which contradicts the general trend. Oligosaccharides such as cellulose, pectin, and amino acids are important components that participate in cell wall formation. Therefore, in our study, cell wall components are negatively affected depending on the duration of the stress exposed to harmful metabolites such as reactive oxygen metabolites that occur due to salt stress[26]. The sudden increase in the fifth week suggests that the plant strives to preserve its integrity. The increase in this week suggests two possibilities. Firstly, by regulating the whole metabolism of the plant for adaptation to salt stress, it moves to a new level from the fifth week, or secondly, the measurements in the fifth week suggest that there may be an experimental error. The change in the amount of chlorophyll and carotenoid depends on the concentration of the applied salt stress and the application

	Week 3				Week 4				Week 5				Mean			
	RMSE	R^2	Train	Pred	RMSE	R^2	Train	Pred	RMSE	R^2	Train	Pred	RMSE	R^2	Train	Pred
Lin. Reg.	22.972	0.81	82.516	1800	23.958	0.81	116.44	3100	23.931	0.81	83.629	2600	23.620	0.810	94.2	2500
Rob. Lin. Reg.	23.127	0.81	1649	2400	24.184	0.81	2713.1	3100	24.21	0.81	1922.4	2900	23.840	0.810	2094.8	2800
GPR																
Exponential	14.615	0.92	2152.2	940	16.139	0.91	2371.9	670	13.149	0.94	1633	850	14.634	0.923	2052.4	820
Rational Quad.	14.326	0.93	2418.4	780	15.378	0.92	3642.9	530	12.873	0.95	2554.1	710	14.192	0.933	2871.8	673
NN																
Narrow	17.337	0.89	2090.9	4400	16.114	0.91	3637.7	5400	13.902	0.94	2329.8	5300	15.784	0.913	2686.1	5033
Medium	18.632	0.88	2556.5	4300	14.411	0.93	3860.5	5300	13.013	0.95	2682.8	4800	15.352	0.920	3033.3	4800
Wide	17.509	0.89	3538.7	6400	14.761	0.93	5167.8	6400	12.241	0.95	3744.5	5700	14.837	0.923	4150.3	6167
Bilayered	16.615	0.9	2752.2	5200	15.338	0.92	4594.1	6900	14.499	0.93	3069.9	5200	15.484	0.917	3472.1	5767
Trilayered	16.175	0.91	2982.1	6000	14.26	0.93	4586	6400	15.321	0.92	3275.6	6500	15.252	0.920	3614.6	6300
SVM																
Linear	21.3	0.84	477.31	140	24.819	0.8	856.21	120	23.328	0.82	565.5	130	23.149	0.820	633.0	130
Cubic	18.661	0.88	652.68	130	18.292	0.89	1175.6	130	16.745	0.91	698.65	160	17.899	0.893	842.3	140
Quadratic	18.704	0.88	569.08	130	18.461	0.89	1254.5	110	18.048	0.89	770.66	140	18.404	0.887	864.7	127
Coarse Gauss	21.382	0.84	908.3	170	26.937	0.76	1695.8	180	23.481	0.82	1068.6	190	23.933	0.807	1224.2	180
Medium Gauss	15.727	0.91	824.46	170	18.357	0.89	1391.5	150	14.155	0.94	895.43	200	16.080	0.913	1037.1	173
Tree																
Coarse	20.472	0.85	118.48	5000	24.295	0.81	169.21	6900	22.253	0.84	131.35	6200	22.340	0.833	139.7	6033
Medium	21.127	0.84	71.227	3700	24.027	0.81	100.75	6700	22.477	0.84	74.088	5400	22.544	0.830	82.0	5267
Fine	22.94	0.81	84.19	3400	25.139	0.79	119.61	6000	22.852	0.83	86.696	5100	23.644	0.810	96.8	4833
Bagged	16.857	0.9	1354.8	4300	18.937	0.88	2201.1	5000	15.278	0.92	1489.4	5100	17.024	0.900	1681.8	4800
Boosted	18.843	0.87	1315.8	4500	21.762	0.84	2236.3	6100	19.173	0.88	1518.1	5400	19.926	0.863	1690.1	5333

Table 2. Comparison of the outputs of different machine learning regression algorithms. Coloring of the mean section is done for each column separately, and the performance increases as the color shifts from red to green. Abbreviations, Train: Training time (sec), Pred: Prediction speed (predictions/sec), Lin: Linear, Reg: Regression, Rob: Robust, Quad: Quadratic, Gauss: Gaussian.

time, strengthening the first possibility. The decrease of these two components under normal conditions was reported in the literature[26].

We also followed the approach of Huang et al. to investigate further the signals related to vital markers[27]. We modified the approach by investigating the whole Amide I region instead of taking only 1602 cm^{-1} since the samples and their responses are not the same. We utilized band component analysis to reveal the sub-components of the broad peaks measured by Raman spectroscopy. We used Grams AI 8 software for this analysis, a curve-fitting process that takes the type of the line shape from the user such as Gaussian, Lorentz, or Voigt. We chose the line shape as Gaussian since it is more suitable for the broadening profile of the biological samples. In the analysis, first, we determined the number of peaks under the curve using the software's second derivative option. Then, we iterated the curve-fitting procedure until the fitting converged. The calculated peak parameters were band center position, height, width, and area. To find the changes in the applied salt level change in particular weeks, we compared peak heights and locations in the Amide region and showed them in Fig. 4.

Most of the Raman intensities shown in the figure had a common decreasing trend as the weeks progressed. On the other hand, the Raman shift values have a consistent decrease only for the bands at 1619, 1669, and 1686 cm^{-1}. An amide bond is a chemical bond formed between a hydroxyl group of a carboxylic group (–COOH) of one molecule and a hydrogen of an amino group (–NH$_2$) of another molecule. Therefore, amide bonds take part in the formation of the polypeptide chain that makes up the proteins. Proteins play an important role in regulating metabolism to maintain cellular integrity under stress[28]. Therefore, the changes in the bands of amide bonds detected in our study show that the plant struggles with stress. The Raman bands characterize the amino acids methionine, lysine, and tyrosine, all of which participate in the structure of proteins and can be found freely in the cell. Because these amino acids participate in the regulation of the intracellular defense system depending on the stress factors such as salt stress as well as many metabolic processes in the cell[28]. In our study, although a decrease was detected in the amount of these amino acids depending on the duration of the applied salt stress, it was determined that they increased when we evaluated the increase in the applied salt stress concentrations within themselves every week. Farhangi-Abriz and Ghassemi-Golezani[28] obtained similar results to our study on soybean. Alfosea-Simon et al., on the other hand, showed that giving amino acids such as methionine, lysine, and tyrosine exogenously to tomatoes reduced the harmful effects of salt stress in their study[29].

The changes in the Raman shift are a sign of molecular structure change in the plant we measured. As the number of oxidative stress products increases with the increasing salt levels, the Amide I band positions[30] and the level of proteins during stress changes[31]. These changes were in line with the band components analysis results.

Quantification of the acquired Raman spectra using a wide variety of machine learning regression algorithms.

To quantify the salt stress level a plant endures, we trained various machine learning algorithms. We used regression models with output in terms of medium salt concentration. The output of the models can be used to determine the level of stress a plant goes through. We used five subgroups of regression models: linear regression, Gaussian process regression (GPR), neural network (NN), support vector machine (SVM), and regression tree. We gave all these models and their respective success rates, train times, and prediction speeds in Table 2. The first 2 weeks have shown heterogeneity. Their spectra were capriciously changing when we
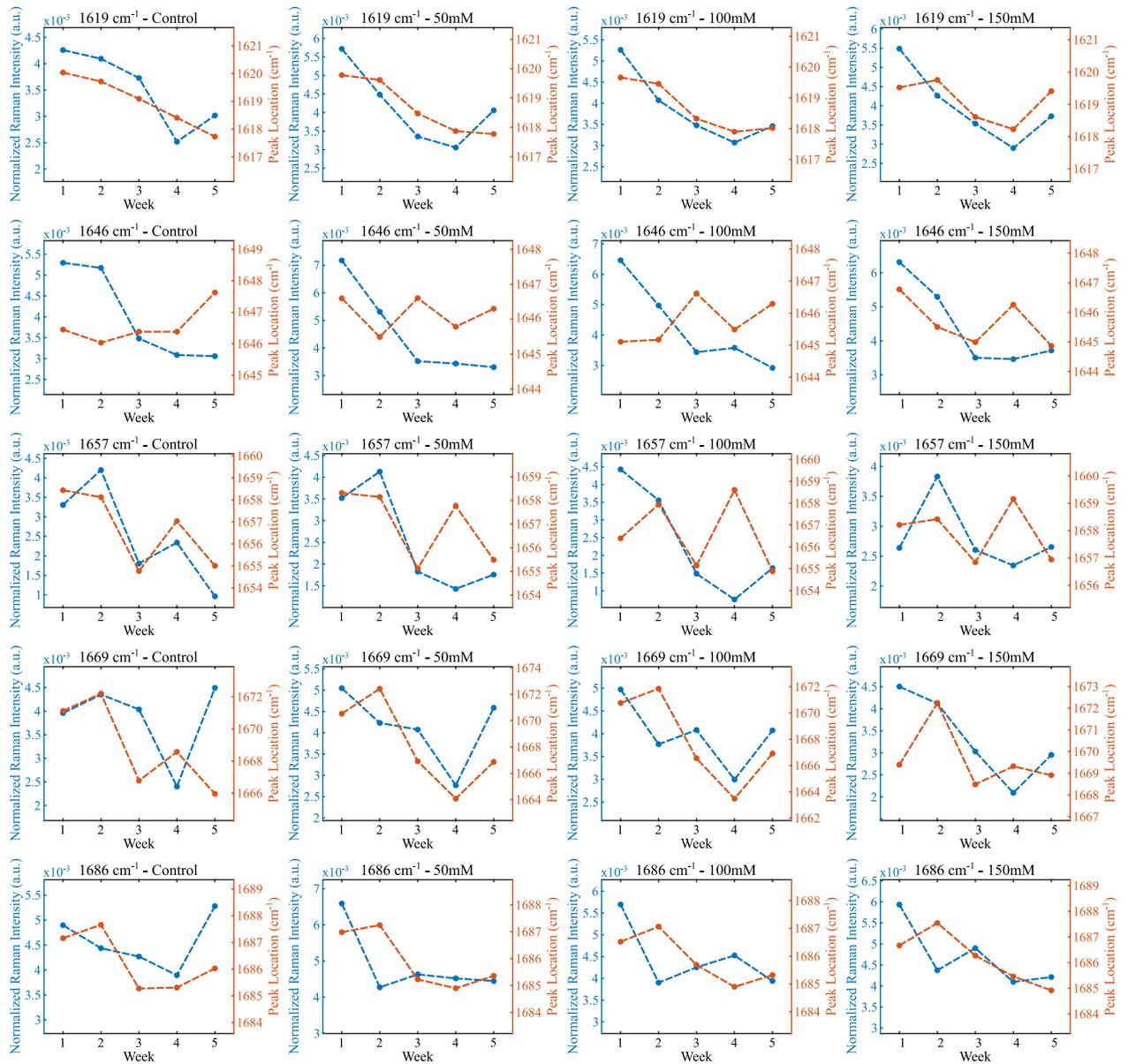
Figure 4. Normalized Raman intensity and peak center location change per week of different concentrations and peaks.

scanned different areas of the same sample. Consequently, we did not perform regression analysis on the spectra of the first 2 weeks. As a measure to determine the success rate of the models, we examined R^2 and root mean square error (RMSE). Even though we did not consider training time and prediction speed when we determined the best-performing model, they are significant parameters, especially in real-time applications. Gaussian process regression (RMSE = 14.192–14.634, R^2 = 0.923–0.933) and neural network models (RMSE = 14.837–15.784, R^2 = 0.913–0.923) gave the best results.

In terms of mean RMSE and R^2, the best performing model among all was rational quadratic Gaussian process regression (RQGPR). We used the test portion of the data to determine the model's success rate. In Fig. 5a–c, we showed the RQGPR model's predictions for the test set. The median of the predictions corresponding to each true concentration value is approximately equal to the true concentration value, an indicator of a well-performing model. To make this observation more visible, we also presented residuals for each concentration in Fig. 5d–f. The medians in the residual plots are all around zero. This indicates that if the age of the observed leaf is entered in terms of weeks, the model can predict the salt level of the plant.

To test the overestimation, we took new wheat leaves from different plants from biology department members that we did not measure before. The ages and the concentrations of the leaves were mixed, and initially, we used only the current week information in the predictions. We tried our models and performed the predictions depending on the given information. In Fig. 6 we showed the residuals of the predictions when compared to the actual salt concentrations. Our predictions deviated from the true concentration by around 20 mM.
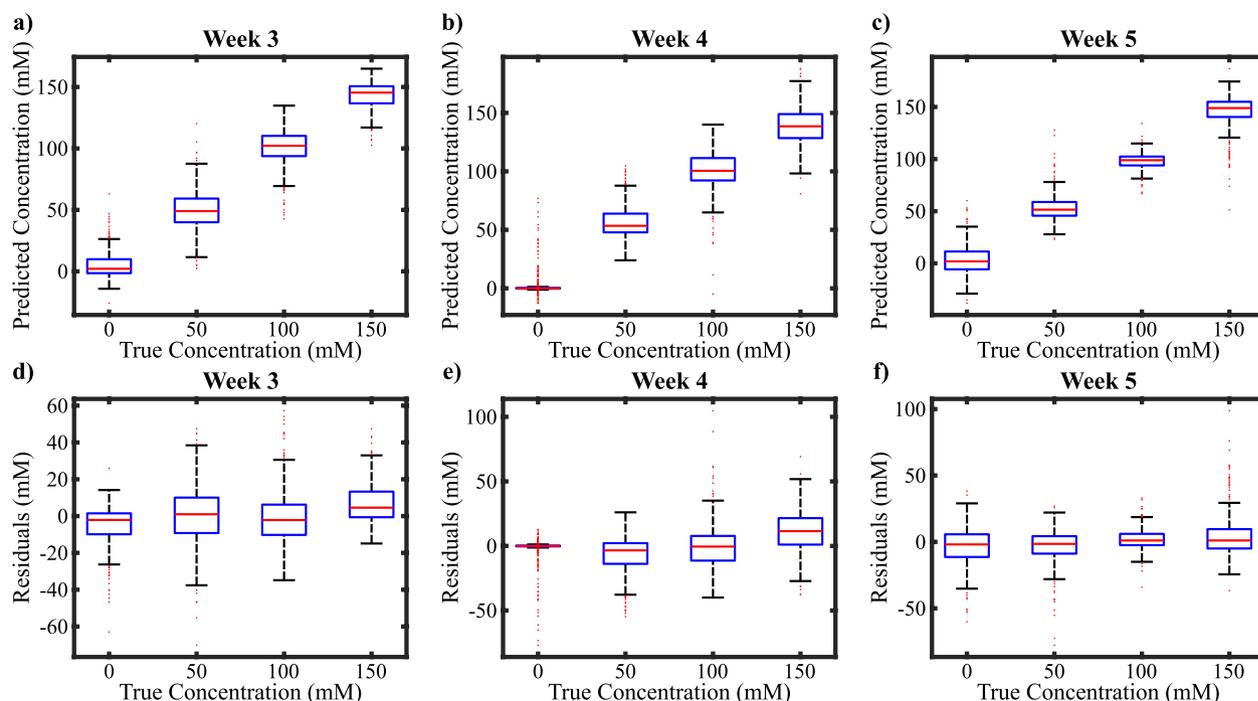
Figure 5. Rational Quadratic Gaussian Process Regression results (RQGPR). Out of all the models trained RQGPR has given the best results. (**a**–**c**) Prediction results of the model on the test set. (**d**–**f**) Residuals of the model predictions on the test set.
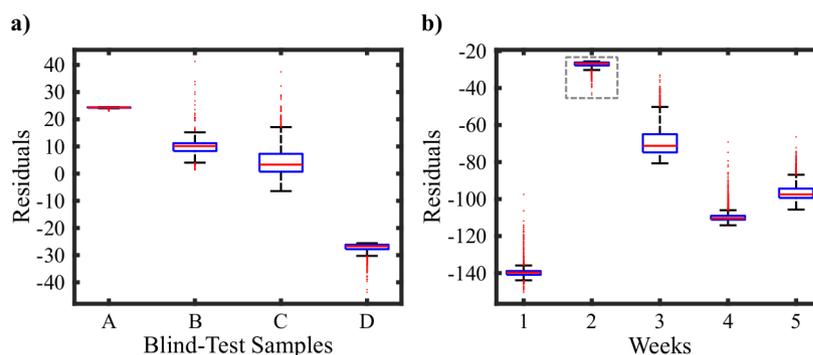
Figure 6. Test results for untrained samples. (**a**) Salt concentration prediction on the untrained blind-test groups. A: Week 3, 100 mM, B: Week 3, Control, C: Week 4, 50 mM, D: Week 2, 150 mM. (**b**) Week prediction example using a week 2 sample. The residual is at its minimum value at week 2.

On the other hand, to test our ability to predict the week and the salt concentrations together using the trained models, we used the untrained test samples and found the weeks corresponding to the minimum deviation from the actual value. In Fig. 6b, we demonstrate the residuals in all weeks. Since this is a second-week sample, our results are consistent with the current state of the plant. As we also predicted the salt concentration of the plant before, our developed tool can be used to find the salt levels of a wheat leave whose age is not known.

To test our tool's ability to predict an untrained group, we left the 100 mM group out of the training group and trained using others. We tested our model using the left-out group and found that the median of the prediction deviates − 25.35 mM from the expected value. Although this value is smaller for the 80/20 partitioned model where the training set includes data from all the concentrations, the value for the untrained trial is still reasonable and practical.

Methods
Wheat growth. Wheat (*Triticum aestivum* L.) seeds of the Selimiye variety were obtained from Trakya Agricultural Research Institute (Edirne). Murashige and Skoog (MS) medium was used as the basic medium for the creation of a tissue culture medium. 20 g/l sucrose ($C_{12}H_{22}O_{11}$), 50 µg/l 2.4-D (2.4-Dichlorophenoxyacetic

Acid), and 8 g/l Agar regeneration medium were added into the medium as a carbon source, a plant growth regulator, and as a thickener, respectively. To determine the sensitivity of wheat embryos to salt stressors, different concentrations of NaCl (50 mM–100 mM–150 mM NaCl) were added to the regeneration medium whose content was given above. The pH values of the mixtures formed were adjusted to be between about 5.8 and 6.0. The created nutrient media were sterilized in an autoclave at 121 °C at 1.2 atm pressure for 20 min and poured into sterile magendas in an aseptic environment in equal amounts[32].

For sterilization, the wheat seeds were kept in 70% ethyl alcohol for 20 seconds, passed through the sterile distilled water series 3 times, then kept in 20% commercial bleach for 20 min, and passed through the distilled water series three times to remove the bleach. Then, the seeds were incubated in sterile distilled water for 2 h in an oven at 35 °C, and swelling was allowed[32,33]. The mature seed embryos were left to swell after surface sterilization. They were separated from their endosperms under aseptic conditions and inoculated into media prepared for salt stressor application. Cultures were observed at 25±2 °C temperature for 4 weeks in a daily illumination period of 16 h light/8 h dark[32].

Wheat leaf samples were prepared for Raman spectral measurements by cutting leaves from the plant near their tip using a scalpel. For consistency, we cut nearly the same size leaves for each sample. We measured leaves while they are still fresh.

Experimental setup and data acquisition. Leaf specimens were compressed between two slides to make sure that the region of interest of the sample will be in the focal plane during the measurement. The thicknesses of the slides were 1 mm. Prepared specimens were scanned with our hand-built motorized Raman spectral scanning setup.

We utilized a diode laser with 785 nm wavelength and 500 mW output power. The output of the laser was steered to a spatial filtering section to select the portion of the beam with the Gaussian distribution. This Gaussian beam was directed to the microscope objective (MO) by reflecting from a dichroic mirror (Thorlabs, DMLP 805) and some beam steering mirrors. The diameter of the beam was adjusted to fill the numerical aperture of the MO. The magnification and numerical aperture of the MO were 10× and 0.25. Since the working distance was greater than the thickness of the slides, we were able to illuminate the wheat leaves with this configuration. The backscattered photons follow the same path with 180° geometry until the dichroic mirror. The inelastically scattered portion of the beam (Raman beam) transmits through the dichroic mirror since the wavelength of the stokes photons are higher (we selected Stokes photons due to their higher intensity). This transmitted beam was coupled into the fiber coupling port using an achromatic lens with a nine mm focal length. The coupled photons are transferred into the Czerny-Turner type USB spectrometer (Ocean QE Pro) via a multimode fiber (NA 0.22).

We collected the spectra with two seconds of integration time. We used boxcar-width averaging with a window size of 3 pixels to denoise the spectra, which allows us to avoid averaging and saving a significant amount of time. We scanned the samples with 5 μm step sizes along the axes of the 300 μm × 300 μm region of interest. The total number of spectra was 3721 (61 × 61) for each scan. This procedure were repeated for varying regions of different leaves.

Data analysis. *Raman spectra pre-processing.* Raman spectroscopy measurements include fluorescent profiles, as well. To analyze only the Raman peaks, one should estimate and clean this baseline profile. We determined local minimum points and applied linear fit using the Raman shift of consecutive pairs of minimum points to achieve this task. Concatenating these lines creates the total estimated fluorescence profile. We calculated this profile for each spectrum and subtracted it.

After correction, we applied an L2 normalization, also known as vector normalization. We calculated the vector norm of each spectrum and divided each element of the spectrum by this norm value. This operation provides a normalized spectrum whose vector norm equals one. It was shown that L2 normalization after fluorescent baseline correction is one of the most efficient methods in Raman spectroscopy[34].

Outliers in the data are eliminated by just keeping the portion between quantiles 0.5% and 99.5%. After elimination 56,067 spectra were left in the data set and used in further analysis. The number of spectra left in each group is given in Table 3. Each week was treated as a separate data set and models were trained for each week separately.

Machine learning application. After pre-processing, we trained a collection of machine learning regression models, such as linear regression, Gaussian process regression, neural network, support vector machines, and regression tree algorithms. We used MATLAB (R2021a, The MathWorks, Inc., Natick) and its regression learner app provided under the statistics and machine learning toolbox. Prior to training the models, we randomly split the data into train and test sets with a ratio of 80 to 20, respectively. While training, we used five cross-validation folds. After training, we tested the model using the aforementioned test set. The best result was given by a Gaussian process regression model with a rational quadratic kernel which is both stationary and non-degenerate covariance function (Eq. 1), where r is the Euclidean distance between two values (x, x') in the input domain (Eq. 2), l is the characteristic length scale, and α is the scale-mixture parameter.

$$k_{RQ}(r) = \left(1 + \frac{r^2}{2\alpha l^2}\right)^{-\alpha} \tag{1}$$

$$r = |x - x'| \tag{2}$$

Week	Concentration (mM)	Number of data
1	0	4307
	50	4334
	100	4372
	150	4716
2	0	2869
	50	3341
	100	3368
	150	2537
3	0	1373
	50	1824
	100	2175
	150	1671
4	0	2836
	50	2808
	100	2685
	150	2414
5	0	1954
	50	1879
	100	2356
	150	2248

Table 3. Number of Raman spectra left after quantile elimination for each week and medium concentration.

Kernel function maps the pair (x, x') from input domain to $\mathbb{R}$[35]. Basis function used was constant with a basis matrix H that is an n-by-1 vector of 1s, such that n is the number of observations. Initial values of the parameters were determined automatically by the regression learner app.

Conclusion

We showed that with the help of machine learning, Raman spectroscopy could be used to quantitatively determine the level of salt stress a plant is exposed to. The shifts in the Raman spectra give precise information about how much a plant struggles from the stressor factors and how close it is to dying. We also presented that models show high levels of success rates with this data; therefore, these models can be used with high accuracy. Our proposed models can quickly provide information on leaf salinity levels when used in a hand-held Raman spectrometer.

Data availability

The spectroscopy data and the trained models that support the findings of this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.5831203[36].

References
1. Arzani, A. & Ashraf, M. Cultivated ancient wheats (*Triticum* spp.): A potential source of health-beneficial food products. *Compr. Rev. Food Sci. Food Saf.* **16**, 477–488 (2017).
2. Shiferaw, B. *et al.* Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security* **5**, 291–317 (2013).
3. Chaves, M. M., Flexas, J. & Pinheiro, C. Photosynthesis under drought and salt stress: Regulation mechanisms from whole plant to cell. *Ann. Bot.* **103**, 551–560 (2009).
4. Yasar, F., Uzal, O. *et al.* Investigation of the effect of calcium applications on plant development of pepper plant in salt stress. In *X International Agriculture Symposium, Agrosym 2019, Jahorina, Bosnia and Herzegovina, 3-6 October 2019. Proceedings*, 312–317 (University of East Sarajevo, Faculty of Agriculture, 2019).
5. Büyük, İ., Soydam Aydın, S. & Aras, S. Bitkilerin stres koşullarına verdiği moleküler cevaplar. *Turk. Bull. Hyg. Exp. Biol./Türk Hijyen ve Deneysel Biyoloji* **69**, 97–110 (2012).
6. Sehar, Z., Masood, A. & Khan, N. A. Nitric oxide reverses glucose-mediated photosynthetic repression in wheat (*Triticum aestivum* L.) under salt stress. *Environ. Exp. Bot.* **161**, 277–289 (2019).
7. Raman, C. V. & Krishnan, K. S. A new type of secondary radiation. *Nature* **121**, 501–502 (1928).
8. Agarwal, U. P. Analysis of cellulose and lignocellulose materials by Raman spectroscopy: A review of the current status. *Molecules* **24**, 1659 (2019).
9. Movasaghi, Z., Rehman, S. & Rehman, I. U. Raman spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **42**, 493–541 (2007).
10. Agarwal, U. P., Reiner, R. S. & Ralph, S. A. Cellulose I crystallinity determination using FT-Raman spectroscopy: Univariate and multivariate methods. *Cellulose* **17**, 721–733 (2010).
11. Schulz, H. & Baranska, M. Identification and quantification of valuable plant substances by IR and Raman spectroscopy. *Vib. Spectrosc.* **43**, 13–25 (2007).

12. Baranska, M., Schulz, H., Rösch, P., Strehle, M. A. & Popp, J. Identification of secondary metabolites in medicinal and spice plants by NIR-FT-Raman microspectroscopic mapping. *Analyst* **129**, 926–930 (2004).
13. Wesełucha-Birczyńska, A., Łabanowska, M., Kurdziel, M. & Filek, M. Resonance Raman and EPR spectroscopy studies of untreated spring wheat leaves. *Vib. Spectrosc.* **60**, 113–117 (2012).
14. Kurdziel, M., Dłubacz, A., Wesełucha-Birczyńska, A., Filek, M. & Łabanowska, M. Stable radicals and biochemical compounds in embryos and endosperm of wheat grains differentiating sensitive and tolerant genotypes-EPR and Raman studies. *J. Plant Physiol.* **183**, 95–107 (2015).
15. Rys, M., Szaleniec, M., Skoczowski, A., Stawoska, I. & Janeczko, A. FT-Raman spectroscopy as a tool in evaluation the response of plants to drought stress. *Open Chem.* **13**, 1091–1100 (2015).
16. Filek, M., Łabanowska, M., Kurdziel, M., Wesełucha-Birczyńska, A. & Bednarska-Kozakiewicz, E. Structural and biochemical response of chloroplasts in tolerant and sensitive barley genotypes to drought stress. *J. Plant Physiol.* **207**, 61–72. https://doi.org/10.1016/j.jplph.2016.09.012 (2016).
17. Cebi, N., Dogan, C. E., Develioglu, A., Yayla, M. E. A. & Sagdic, O. Detection of l-cysteine in wheat flour by Raman microspectroscopy combined chemometrics of HCA and PCA. *Food Chem.* **228**, 116–124 (2017).
18. Farber, C. & Kurouski, D. Detection and identification of plant pathogens on maize kernels with a hand-held Raman spectrometer. *Anal. Chem.* **90**, 3009–3012 (2018).
19. Brennan, M., McDonald, A. & Topp, C. F. Use of Raman microspectroscopy to predict malting barley husk adhesion quality. *Plant Physiol. Biochem.* **139**, 587–590 (2019).
20. Ghaderi, N., Hatami, M. R., Mozafari, A. & Sioschmardeh, A. Change in antioxidant enzymes activity and some morpho-physiological characteristics of strawberry under long-term salt stress. *Physiol. Mol. Biol. Plants Int. J. Funct. Plant Biol.* **24**, 833–843. https://doi.org/10.1007/s12298-018-0535-2 (2018).
21. Edwards, H., Farwell, D. & Webster, D. FT Raman microscopy of untreated natural plant fibres. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **53**, 2383–2392. https://doi.org/10.1016/S1386-1425(97)00178-9 (1997).
22. Synytsya, A., Čopíková, J., Matějka, P. & Machovič, V. Fourier transform Raman and infrared spectroscopy of pectins. *Carbohydr. Polym.* **54**, 97–106. https://doi.org/10.1016/S0144-8617(03)00158-9 (2003).
23. Zhu, G., Zhu, X., Fan, Q. & Wan, X. Raman spectra of amino acids and their aqueous solutions. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **78**, 1187–1195. https://doi.org/10.1016/j.saa.2010.12.079 (2011).
24. Jehlička, J. *et al.* Potential and limits of Raman spectroscopy for carotenoid detection in microorganisms: Implications for astrobiology. *Philos. Trans. R. Soc. A.* **372**, 20140199. https://doi.org/10.1098/rsta.2014.0199 (2014).
25. Kecel-Gunduz, S., Bicak, B., Celik, S., Akyuz, S. & Ozel, A. E. Structural and spectroscopic investigation on antioxidant dipeptide, ʟ-Methionyl-ʟ-Serine: A combined experimental and DFT study. *J. Mol. Struct.* **1137**, 756–770. https://doi.org/10.1016/j.molstruc.2017.02.075 (2017).
26. Zhao, S. *et al.* Regulation of plant responses to salt stress. *Int. J. Mol. Sci.* **22**, 4609. https://doi.org/10.3390/ijms22094609 (2021).
27. Huang, Y.-S., Karashima, T., Yamamoto, M., Ogura, T. & Hamaguchi, H.-O. Raman spectroscopic signature of life in a living yeast cell. *J. Raman Spectrosc.* **35**, 525–526 (2004).
28. Farhangi-Abriz, S. & Ghassemi-Golezani, K. Improving amino acid composition of soybean under salt stress by salicylic acid and jasmonic acid. *J. Appl. Bot. Food Qual.* **89**, 243–248 (2016).
29. Alfosea-Simón, M. *et al.* Application of biostimulants containing amino acids to tomatoes could favor sustainable cultivation: Implications for tyrosine, lysine, and methionine. *Sustainability* **12**, 9729 (2020).
30. Yang, J. & Yen, H. E. Early salt stress effects on the changes in chemical composition in leaves of ice plant and Arabidopsis. A Fourier transform infrared spectroscopy study. *Plant Physiol.* **130**, 1032–1042 (2002).
31. Alam, I. *et al.* Proteome analysis of soybean roots subjected to short-term drought stress. *Plant Soil* **333**, 491–505 (2010).
32. Sen, A. *et al.* Effect of salt stress on growth parameters and antioxidant enzymes of different wheat (*Triticum aestivum* L.) varieties on in vitro tissue culture. *Fresenius Environ. Bull.* **20**, 489–495 (2011).
33. Özgen, M., Türet, M., Altınok, S. & Sancak, C. Efficient callus induction and plant regeneration from mature embryo culture of winter wheat (*Triticum aestivum* L.) genotypes. *Plant Cell Rep.* **18**, 331–335 (1998).
34. Heraud, P., Wood, B. R., Beardall, J. & McNaughton, D. Effects of pre-processing of Raman spectra on in vivo classification of nutrient status of microalgal cells. *J. Chemom. J. Chemom. Soc.* **20**, 193–197 (2006).
35. Rasmussen, C. E. *Gaussian Processes for Machine Learning* (MIT Press, 2006).
36. Kecoglu, I. *et al.* Quantification of salt stress in wheat leaves by Raman spectroscopy and machine learning. *bioRxiv* https://doi.org/10.5281/zenodo.5831203 *(2022)*.

Acknowledgements

Author contributions

U.P., F.G. and, M.B.U. conceived and directed the study. U.P. and I.K. conducted the Raman spectroscopy experiments. M.S. and A.S cultivated and prepared wheat leaf samples. I.K. and U.P. analysed the results and trained the ML models. All authors reviewed the manuscript. I.K. and M.S. contributed equally to this work.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to U.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.