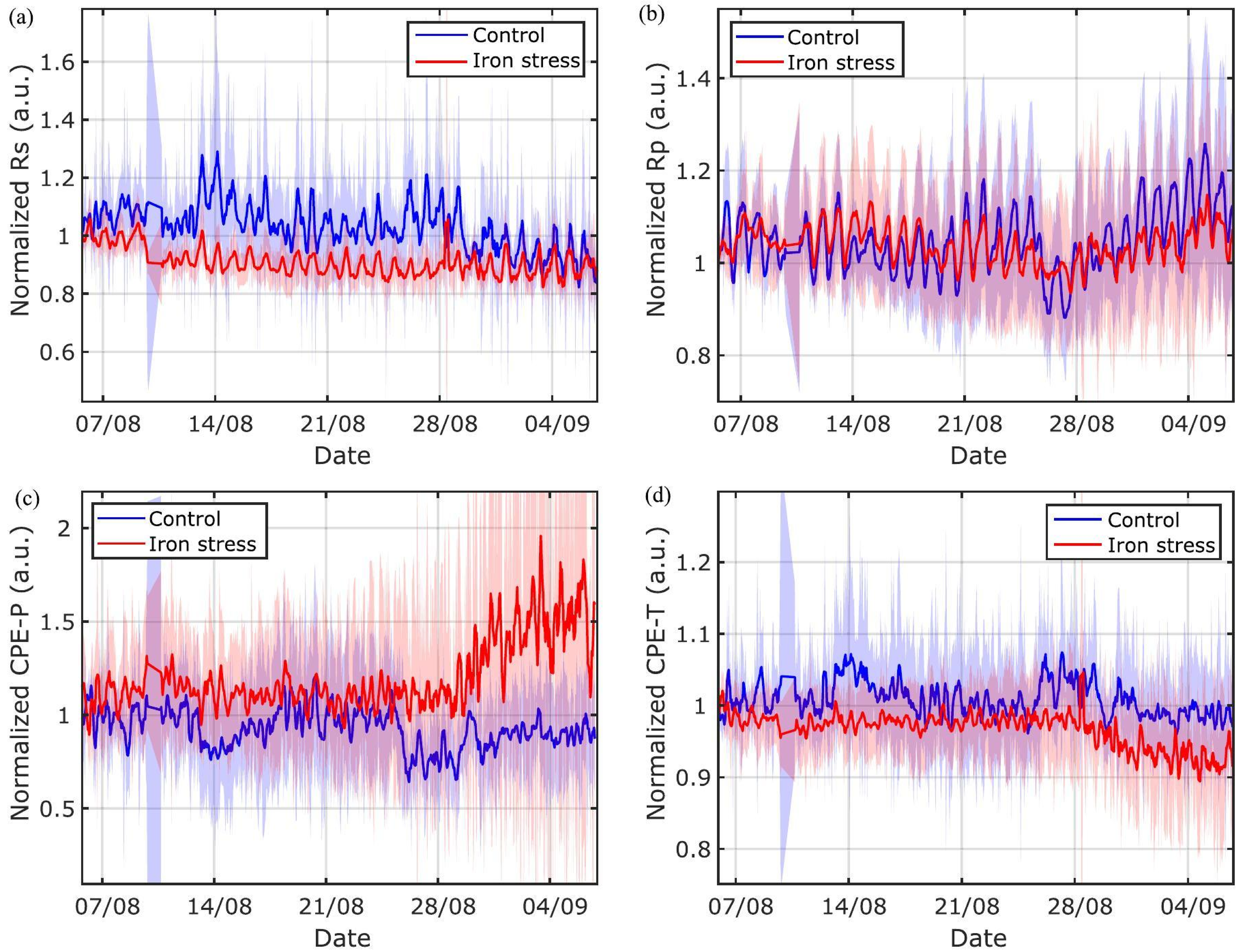

Fig. 4. Time-series plot of normalized circuit parameters, namely Rs (a), Rp (b), CPE-P (c), and CPE-T (d), evolution over a 4-week period. The shaded colored zones represent the range of amplitude changes over time for both groups (blue shadow for control plants and red shadow for iron-stressed plants). The median magnitude values of the control plants represented by a blue line while the iron-stressed plants depicted by a red line.

B. Plant Status Classification

The employed equivalent circuit model, implementing only few electrical components, represents an extreme simplification of the plant physiology and is, therefore, not ideal to accurately represent the complexity of the tissues of the stems. In fact, from the direct interpretation of the extracted parameters it is difficult to find indicators for the identification between a healthy plant and a plant under nutrient stress. Even more complicated would be to identify whether the plant is in early stress condition so that nutrient supply can be intervened before symptoms and consequent loss of yield arise. Nevertheless, the significant difference in circuit parameters observed between the considered classes demonstrates how such a model can be used as a general indicator of the plant status and support the employment of the circuit parameters as discriminant features in the training of classification models, where nonapparent and nonlinear relationships between bioimpedance data and plant health status could be found. For this reason, based on the previously extracted equivalent circuit components of the single dispersion Cole model, various supervised machine learning classification algorithms were trained, validated, and tested, for the discrimination of three distinct classes, respectively representing the control, early stress, and late stress conditions of iron deficiency in tomato plants.

TABLE I

Summary of the Optimized Classification Models F_1-Score Accuracy on the Training and Test Sets, With % Difference Among the Two Phases

Model	Optimized Hyperparameters	F_1 train	F_1 test	% diff
DA	$\Delta = 9.577e^{-6}$, $\gamma = 0.013$	0.497	0.480	-1.65
KM	KS = 0.665, $\lambda = 0.0014$	0.868	0.818	-5.03
KNN	Distance = Mahalanobis, NN = 7	0.942	0.894	-4.79
LM	$\lambda = 1.386e^{-4}$, Learner = SVM	0.543	0.529	-1.34
NBC	Distribution Names = Kernel, Width = 0.0102	0.767	0.639	-12.8
MLP	Activation = Relu, Layer Size = [18, 12, 39]	0.984	0.935	-4.96
SVM	KS = 1.3312, Binary loss = Hinge	0.543	0.531	-1.16
DT	Number of nodes = 317	0.964	0.832	-13.2

KS = Kernel Scale

The accuracy of the seven models developed to classify the plant status is listed in Table I, in terms of F_1-score in training and test, as well as for what concerns the % difference among the two phases. Overall, the algorithms presenting the poorest performance are linear methods such as DA and LM, achieving an F_1-score of around 50%. Such results are expected since the association between bioimpedance data and biological behaviors commonly follows nonlinear patterns. NBC and DT algorithms, on the other hand, present the highest degree of