



**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
DOUTORADO EM AGRONOMIA**

**ANÁLISE TEMPORAL DA SINALIZAÇÃO ELÉTRICA EM PLANTAS DE SOJA  
SUBMETIDAS A DIFERENTES PERTURBAÇÕES EXTERNAS**

**GUSTAVO FRANCISCO ROSALIN SARAIVA**

Presidente Prudente - SP  
2017

**ANÁLISE TEMPORAL DA SINALIZAÇÃO ELÉTRICA EM PLANTAS DE SOJA  
SUBMETIDAS A DIFERENTES PERTURBAÇÕES EXTERNAS**

**GUSTAVO FRANCISCO ROSALIN SARAIVA**

Tese apresentada a Pró-Reitoria de Pesquisa e Pós-Graduação, Universidade do Oeste Paulista, como parte dos requisitos para obtenção do título de Doutor em Agronomia – Área de concentração: Produção Vegetal.

Orientador: Dr. Gustavo Maia Souza

571.2  
S243a

Saraiva, Gustavo Francisco Rosalin.

Análise temporal da sinalização elétrica em plantas de soja submetidas a diferentes perturbações externas / Gustavo Francisco Rosalin Saraiva – Presidente Prudente, 2017.

82 f.: il.

Tese (Doutorado em Agronomia) - Universidade do Oeste Paulista – Unoeste, Presidente Prudente, SP, 2017.

Bibliografia.

Orientador: Dr. Gustavo Maia Souza.

1. Algoritmos de classificação, 2. Complexidade, 3. Eletrofisiologia Vegetal, 4. Entropia Aproximada. I. Título.

**GUSTAVO FRANCISCO ROSALIN SARAIVA**

**ANÁLISE TEMPORAL DA SINALIZAÇÃO ELÉTRICA EM PLANTAS DE SOJA  
SUBMETIDAS A DIFERENTES PERTURBAÇÕES EXTERNAS.**

Tese apresentada a Pró-Reitoria de Pesquisa e Pós-Graduação, Universidade do Oeste Paulista, como parte dos requisitos para obtenção do título de Doutor em Agronomia - Área de Concentração: Produção Vegetal.

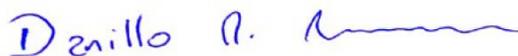
Presidente Prudente, 31 de Março de 2017

**BANCA EXAMINADORA**



---

Prof. Dr. Gustavo Maia Souza  
Instituição: Universidade Federal de Pelotas - UFPEL  
Local: Pelotas - RS



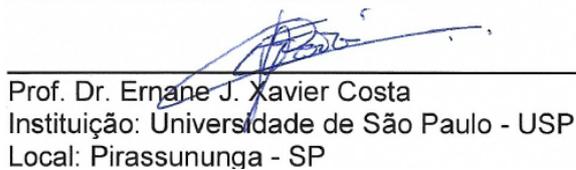
---

Prof. Dr. Danilo Roberto Pereira  
Instituição: Universidade do Oeste Paulista - UNOESTE  
Local: Presidente Prudente - SP



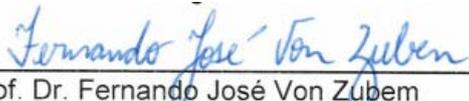
---

Prof. Dr. Luiz Gonzaga Esteves Vieira  
Instituição: Universidade do Oeste Paulista - UNOESTE  
Local: Presidente Prudente - SP



---

Prof. Dr. Ernane J. Xavier Costa  
Instituição: Universidade de São Paulo - USP  
Local: Pirassununga - SP



---

Prof. Dr. Fernando José Von Zuben  
Instituição: Universidade Estadual de Campinas - UNICAMP  
Local: Campinas - SP

## DEDICATÓRIA

Dedico, primeiramente, essa tese de Doutorado às plantas, objeto de estudo e profunda admiração do autor.

Dedico esse trabalho a todos os amantes das plantas, que buscam compreender e conhecer mais a respeito desses seres vivos.

Dedico a minha esposa Regina, meu pai Francisco, minha mãe Marli, minha irmã Grazielle e meu cunhado Pedro.

Dedico ao meu amigo e orientador Dr. Gustavo Maia Souza e meu grande amigo Gabriel de Toledo.

## AGRADECIMENTOS

Agradeço em primeiro lugar a meu orientador, professor **Dr. Gustavo Maia Souza**, por aceitar o desafio de me orientar e possibilitar o acontecimento desta tese.

Agradeço à **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES**, pela bolsa de estudos de doutorado.

Agradeço ao antigo **Laboratório de Inteligência em Plantas e Ecofisiologia “Ulrich Lutgge” – LIPEUL**, pela acolhida e suporte técnico.

Agradeço ao **Centro de Estudos em Ecofisiologia do Oeste Paulista – CEVOP**, pela acolhida e suporte técnico.

Agradeço aos **amigos de laboratório** que contribuíram com conhecimento, experiências, trocas de informações e principalmente pela amizade.

Agradeço à minha esposa **Regina**, por sua paciência e ajuda em momentos difíceis e por estar sempre ao meu lado.

Agradeço ao meu grande amigo **Gabriel**, que fez parte dessa tese com sua contribuição intelectual e filosófica, e sobretudo com sua amizade e companheirismo.

*“O Mestre na arte da vida faz pouca distinção entre o seu trabalho e o seu lazer, entre a sua mente e o seu corpo, entre a sua educação e a sua recreação, entre o seu amor e a sua religião. Ele dificilmente sabe distinguir um corpo do outro. Ele simplesmente persegue sua visão de excelência em tudo que faz, deixando para os outros a decisão de saber se está trabalhando ou se divertindo. Ele acha que está sempre fazendo as duas coisas simultaneamente.”*

*(Sidarta Gautama - Buda)*

## RESUMO

### **Análise temporal da sinalização elétrica em plantas de soja submetidas a diferentes perturbações externas**

As plantas são organismos complexos com processos dinâmicos que, devido ao seu modo sésil de vida, sofrem influência das condições ambientais todo o tempo. Plantas podem perceber e responder com precisão a diferentes estímulos ambientais de forma inteligente, mas para isso se faz necessário um complexo e eficiente sistema de sinalização. A sinalização elétrica em plantas já é conhecida há muito tempo, mas vem ganhando destaque recentemente com seu entendimento em relação aos processos fisiológicos das plantas. O objetivo desta tese foi testar as seguintes hipóteses: séries temporais de dados obtidos da sinalização elétrica de plantas possuem informação não aleatória, com padrão dinâmico e oscilatório, sendo tal dinâmica afetada por estímulos ambientais e que há padrões específicos nas respostas a estímulos. Em ambiente controlado, foram aplicados estímulos ambientais estressantes em plantas de soja, e captados os dados de sinalização elétrica antes e após a aplicação dos mesmos. As séries temporais obtidas foram analisadas utilizando ferramentas estatísticas e computacionais para se determinar o Espectro de Frequências (FFT), Autocorrelação dos valores e Entropia Aproximada (ApEn). Para se verificar a existência de padrões nas séries, foram utilizados algoritmos de classificação da área de aprendizado de máquina. A análise das séries temporais mostrou que os sinais elétricos coletados de plantas apresentaram dinâmica oscilatória com padrão de distribuição de frequências em lei de potência. Os resultados permitem diferenciar com grande eficácia séries coletadas antes e após a aplicação dos estímulos. As análises de PSD e autocorrelação mostraram grande diferença na dinâmica dos sinais elétricos antes e após a aplicação dos estímulos. A análise de ApEn mostrou haver diminuição da complexidade do sinal após a aplicação dos estímulos. Os algoritmos de classificação alcançaram valores significativos na acurácia de detecção de padrões e classificação das séries temporais, mostrando haver padrões matemáticos nas diferentes respostas elétricas das plantas. Conclui-se que as séries temporais de sinais bioelétricos de plantas possuem informação discriminante. Os sinais possuem dinâmica oscilatória, tendo suas propriedades alteradas por estímulos ambientais. Há ainda padrões matemáticos embutidos nas respostas da planta a estímulos específicos.

**Palavras-chave:** Algoritmos de classificação, Complexidade, Eletrofisiologia Vegetal, Entropia Aproximada.

## ABSTRACT

### **Temporal analysis of electrical signaling in soybean plants subjected to different external disturbances**

Plants are complex organisms with dynamic processes that, due to their sessile way of life, are influenced by environmental conditions at all times. Plants can accurately perceive and respond to different environmental stimuli intelligently, but this requires a complex and efficient signaling system. Electrical signaling in plants has been known for a long time, but has recently gained prominence with the understanding of the physiological processes of plants. The objective of this thesis was to test the following hypotheses: temporal series of data obtained from electrical signaling of plants have non-random information, with dynamic and oscillatory pattern, such dynamics being affected by environmental stimuli and that there are specific patterns in responses to stimuli. In a controlled environment, stressful environmental stimuli were applied in soybean plants, and the electrical signaling data were collected before and after the application of the stimulus. The time series obtained were analyzed using statistical and computational tools to determine Frequency Spectrum (FFT), Autocorrelation of Values and Approximate Entropy (ApEn). In order to verify the existence of patterns in the series, classification algorithms from the area of machine learning were used. The analysis of the time series showed that the electrical signals collected from plants presented oscillatory dynamics with frequency distribution pattern in power law. The results allow to differentiate with great efficiency series collected before and after the application of the stimuli. The PSD and autocorrelation analyzes showed a great difference in the dynamics of the electric signals before and after the application of the stimuli. The ApEn analysis showed that there was a decrease in the signal complexity after the application of the stimuli. The classification algorithms reached significant values in the accuracy of pattern detection and classification of the time series, showing that there are mathematical patterns in the different electrical responses of the plants. It is concluded that the time series of bioelectrical signals of plants contain discriminant information. The signals have oscillatory dynamics, having their properties altered by environmental stimuli. There are still mathematical patterns built into plant responses to specific stimuli.

**Keywords:** Classification algorithms, Complexity, Plant Electrophysiology, Approximate Entropy.

## LISTA DE FIGURAS

FIGURA 1 -	Planta modelo utilizada nos experimentos e Inserção dos eletrodos no caule da planta.....	30
FIGURA 2 -	Plantas conectadas aos eletrodos dentro da Gaiola de Faraday (a). Vista externa do Fitotron com equipamentos de coleta de dados ao fundo (b). Biopac MP-36 (c) e eletrodos de agulha utilizados (d).....	31
FIGURA 3 -	Esquema representando o funcionamento dos algoritmos de classificação na busca por padrões nas séries temporais.....	34
FIGURA 4 -	Exemplo de Classificador Bayesiano ( <i>BAYES</i> ).....	35
FIGURA 5-	Exemplo de Rede Neural Multicamadas.....	36
FIGURA 6 -	Exemplo de Máquinas de vetores de suporte (SVM).....	37
FIGURA 7-	Exemplo de Floresta de caminhos ótimos (OPF).....	38
FIGURA 8-	Exemplo de K-vizinhos mais próximos (KNN).....	39
FIGURA 9-	Séries temporais (20 minutos) obtidas de plantas de soja antes e após a aplicação de estímulos: osmótico (a) e (b), térmico (c) e luminoso (d).....	42
FIGURA 10-	Padrões de Espectro de Frequências (FFT) encontrados em séries temporais obtidas de plantas de soja antes da aplicação dos estímulos.....	46
FIGURA 11-	Padrões de Espectro de Frequências (FFT) encontrados em séries temporais obtidas de plantas de soja após a aplicação dos estímulos.....	47
FIGURA 12-	Padrões de Espectro de Frequências (FFT) encontrados em todas as séries temporais coletadas de eletrodos abertos.....	48
FIGURA 13-	Padrões de Autocorrelação encontrados nas séries temporais coletadas antes da aplicação dos estímulos.....	50
FIGURA 14-	Padrões de Autocorrelação encontrados nas séries temporais coletadas após da aplicação dos estímulos.....	51
FIGURA 15-	Padrão de Autocorrelação encontrado nas séries temporais coletadas pelos eletrodos abertos.....	52

FIGURA 16-	Valores de Entropia Aproximada (ApEn) das séries temporais coletadas de plantas de soja e de eletrodos abertos.....	53
FIGURA 17-	Acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas antes e após a aplicação dos estímulos (sem distinção de classes).....	56
FIGURA 18-	Acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas após a aplicação dos estímulos (sem distinção de classes).....	56
FIGURA 19-	Acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas antes e após a aplicação de estresse osmótico.....	57
FIGURA 20-	Acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas antes e após a aplicação de estresse térmico.....	57
FIGURA 21-	Acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas antes e após a aplicação de estresse luminoso.....	58

## LISTA DE TABELAS

TABELA 1-	Média dos valores dos picos de voltagem encontrados nas séries temporais antes e após aplicação dos estímulos (em $\mu\text{V}$ ) .....	43
-----------	---	----

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	13
1.1	Justificativa e Perspectivas do Estudo.....	14
1.1.2	Hipóteses.....	16
<b>2</b>	<b>EMBASAMENTO TEÓRICO</b> .....	17
2.1	A Soja, o estresse e a produtividade.....	17
2.1.2	Eletrofisiologia Vegetal.....	18
2.1.2.3	Estudo de Séries Temporais.....	24
<b>3</b>	<b>MATERIAL E MÉTODOS</b> .....	28
3.1	Planta modelo e local dos experimentos.....	28
3.1.2	Experimentos.....	28
3.1.2.3	Captação dos dados.....	29
3.1.2.3.4	Análises dos dados.....	31
<b>4</b>	<b>RESULTADOS</b> .....	42
4.1	Análise Descritiva das Séries Temporais.....	42
4.1.2	Espectro de Frequências (FFT).....	44
4.1.2.3	Análises de Autocorrelação.....	49
4.1.2.3.4	Entropia Aproximada.....	53
4.1.2.3.4.5	Técnicas de Classificação.....	54
<b>5</b>	<b>DISCUSSÃO</b> .....	59
<b>6</b>	<b>CONCLUSÕES E PERSPECTIVAS</b> .....	65
	<b>REFERÊNCIAS</b> .....	67
	<b>ANEXO</b> .....	76

## 1 INTRODUÇÃO

É notável a capacidade que as plantas possuem de interagir com toda a complexidade ambiental em seu entorno. Mais do que simples respostas mecânicas aos estímulos externos, tem se tornado cada vez mais claro que as plantas possuem uma capacidade de lidar com seus problemas de forma realmente "inteligente", através de comportamentos e respostas inteligentes, exibindo ainda evidências de características cognitivas como memória e aprendizado (TREWAVAS, 2003; GARZÓN, 2007; THELLIER; LÜTTGE, 2012). Esse trabalho tem como contexto a ideia de cognição em plantas, que tem como foco principal a percepção e ação dentro do ambiente e, dentro desta perspectiva, a maioria dos animais e até mesmo o comportamento de bactérias pode ser considerado cognitivo, de maneira limitada. Essa perspectiva enfatiza o fato de que os comportamentos das criaturas não são simples, mas, com reflexos autômatos, incorporam meios flexíveis e adaptáveis para organizar seu comportamento de forma coerente. Até agora a noção de cognição não foi estendida para as plantas, e isso faz com que a possibilidade de plantas exibirem um comportamento cognitivo atinja muitas pessoas de forma surpreendente (GARZÓN; KEIJZER, 2009). Entretanto, as plantas podem exibir um comportamento muito mais complexo e ao mesmo tempo maleável, do que a maioria das pessoas imaginam (TREWAVAS, 2003).

O comportamento complexo das plantas se mostra na grande capacidade e complexidade com que as plantas se relacionam com o ambiente e todos os seres ao seu redor. Em seu ambiente natural, organismos vivos enfrentam diversas situações que requerem respostas eficientes, garantindo assim a sobrevivência da espécie. Tais situações têm natureza complexa e imprevisível, tornando a existência de respostas pré-programadas inadequadas e pouco prováveis (SOUZA, 2009). A capacidade de compreensão dos sinais ambientais é imprescindível para que a planta esteja apta a sobreviver e competir por recursos em seu ambiente natural. As plantas são seres vivos que não dependem da existência de qualquer tecido cerebral específico para executar operações complexas. Portanto, as plantas exibem condições para a posse de uma arquitetura cognitiva que a auxilia na interação com o meio ambiente (STENHOUSE, 1974; TREWAVAS, 2003; THELLIER; LÜTTGE, 2012).

Em sistemas cognitivos, um processo essencial é a transmissão de informação no sistema por meio de uma eficiente rede de sinalização. Dentro do contexto que procura entender como as plantas interagem e respondem aos estímulos ambientais, os estudos envolvendo a sinalização elétrica, uma nova área de sinalização em plantas, vem ganhando espaço (VOLKOV, 2012). Revisões detalhadas sobre os potenciais de ação têm sido publicadas, indicando que as plantas superiores podem utilizar sinais elétricos para regular uma variedade de funções fisiológicas. Já existem evidências de que os sinais elétricos desempenham um papel importante em muitos processos vitais da planta, como a respiração mitocondrial (FILEK, 1997), a absorção de água (DAVIES et al., 1991), o descarregamento do floema (FROMM, 1991), a translocação no floema (FROMM; BAUER, 1994), bem como a fertilização (FROMM et al., 1995). Desta forma, os sinais bioelétricos derivados das plantas têm sido bem estudados a fim de se demonstrar que potenciais elétricos provenientes de suas atividades podem carregar informações em respostas a estímulos ambientais específicos.

### **1.1 Justificativa e Perspectivas do Estudo**

Há um conhecimento já estabelecido na literatura sobre os tipos de sinais elétricos em plantas e suas relações na sinalização de curta e longa distância de algumas respostas fisiológicas. Contudo, nada ainda foi estudado sobre a dinâmica temporal dos sinais elétricos e possíveis relações com o estado fisiológico das plantas. Os estudos atuais estão restritos às análises da amplitude e velocidade das respostas elétricas, sem considerar a complexidade da dinâmica temporal do sinal em si durante eventos de sinalização pós-estimulação.

Em animais, particularmente em seres humanos, estudos com análise da dinâmica temporal de sinais elétricos obtidos com eletroencefalogramas (EEG) têm permitido estabelecer relações consistentes entre o estado de saúde de indivíduos com medidas de complexidade em dados de EEG. Por exemplo, reduções drásticas na complexidade de sinais de EEG foram associadas a ataques epiléticos em humanos (SRINIVASAN et al., 2007). Similarmente, análises da dinâmica temporal de EEG em pacientes com Alzheimer identificaram uma redução na irregularidade (complexidade) no EEG (ABÁSALO et al., 2005). Além disso, foi possível estabelecer uma correlação consistente entre a variação da complexidade

da dinâmica em EEG e os estados de consciência (sono e vigília) em humanos (BURIOKA et al., 2005). Gardiner (2012, 2013) tem sugerido a possibilidade de utilizar EEG em plantas para relacionar possíveis padrões de dinâmica temporal com diferentes estados fisiológicos. No entanto, até a realização desse trabalho, não havia nenhum tipo de estudo nessa direção.

Relações entre a variação na complexidade de séries temporais de parâmetros fisiológicos e o estado de plantas submetidas a diferentes condições ambientais foram previamente estudadas por Souza et al. (2004a, 2004b, 2005). Aumentos no déficit de pressão de vapor do ar induziram um aumento na complexidade da dinâmica da assimilação de CO<sub>2</sub> e uma redução na dinâmica da condutância estomática em *Citrus sinensis* (SOUZA et al., 2004a). Em duas espécies tropicais arbóreas, a deficiência hídrica tendeu a aumentar a complexidade da dinâmica do fluxo de seiva (SOUZA et al., 2004b). Segundo Souza et al. (2005), existe uma relação entre a complexidade dos processos fisiológicos e a capacidade da planta de lidar com as flutuações ambientais, sugerindo que sistemas com dinâmicas mais complexas tenderiam a suportar melhor as variações ambientais por conferir ao sistema maior estabilidade.

Com a evolução da tecnologia para se captar os sinais elétricos emitidos pelas plantas, poderemos identificar padrões temporais nesses sinais, associando-os aos estímulos que os causaram. Essa relação pode auxiliar no entendimento da sinalização e comunicação estabelecida pelas plantas em situações adversas. Ao entendermos a sinalização elétrica e seus efeitos, estaremos a um passo de poder dar um diagnóstico sobre o estado em que uma planta se encontra, de forma segura e minimamente invasiva, de modo similar ao sistema utilizado na medicina humana ao se realizar um EEG ou um ECG. Mais que isso, o advento dessa nova área de estudos, acerca da dinâmica temporal da sinalização elétrica em plantas, nos faz sugerir que uma relação planta-computador, como recentemente explorada em animais e humanos, seria viável. Estudos aprofundando o significado fisiológico da dinâmica elétrica em plantas poderiam vir a permitir o desenvolvimento de uma tecnologia totalmente nova de sensores ambientais, baseados nas respostas elétricas das próprias plantas ao seu meio. Com o auxílio da tecnologia computacional e modelos matemáticos, há a possibilidade de capturar, interpretar e codificar a sinalização elétrica de uma planta, transformando-a numa “linguagem”, que nos ajudará a compreender melhor a fisiologia vegetal a nível de

indivíduo, e sua complexa interação com o ambiente. Assim, um estudo mais amplo e detalhado da interpretação dos sinais bioelétricos das plantas pode contribuir para uma melhor compreensão dos princípios e normas que regem o comportamento organizacional das plantas e, a partir disso, estabelecer modelos bioelétricos.

A partir do contexto apresentado, foram definidas as seguintes Hipóteses:

## **1.2 Hipóteses**

1. Séries temporais de dados obtidos da sinalização elétrica em plantas de soja apresentam informação “não aleatória”, ou seja, possuem padrões informacionais embutidos na variação temporal dos potenciais de voltagem.

2. A dinâmica temporal da sinalização elétrica em plantas de soja apresenta dinâmica oscilatória.

3. Estímulos ambientais afetam a dinâmica da sinalização elétrica em plantas de soja.

4. Há padrões bem caracterizados nas respostas elétricas evocadas por estímulos ambientais específicos.

## **2 EMBASAMENTO TEÓRICO**

### **2.1 A Soja, o Estresse e a Produtividade**

Devido ao grande potencial para a produção e consumo em larga escala, a soja (*Glycine max* (L.) Merrill) é um importante produto agrícola mundial, tendo papel de destaque entre as culturas produtoras de grãos. O Brasil é o segundo maior produtor mundial de soja, atrás apenas dos EUA, contudo a produção e a produtividade da soja brasileira crescem a cada safra. Na safra 2015/2016, a cultura ocupou uma área de 33,17 milhões de hectares, o que totalizou uma produção de 95,63 milhões de toneladas. A produtividade média da soja brasileira foi de 2.882 kg por hectare (Embrapa Soja). Quando se observa em escala mundial a produção chega a 312,362 milhões de toneladas, ocupando uma área plantada de 119,732 milhões de hectares.

No entanto, essa cultura enfrenta problemas em várias regiões do Brasil e do mundo, onde o cultivo pode ser limitado ou ter sua produtividade afetada devido às condições climáticas, como elevadas temperaturas, baixa precipitação média anual e baixa capacidade de retenção de água no solo, entre outros (FRANCHINI et al., 2009). Dentre os grandes fatores para a limitação da máxima produtividade em soja, podemos colocar o estresse vegetal como um grande fator, que ao mesmo tempo recebe pouca atenção.

Quando em condições de estresse, as plantas apresentam uma série de mudanças morfológicas, fisiológicas e bioquímicas, que afetam negativamente a produção de óleo e enchimento de grãos (SOUZA et al., 2013; LISAR et al., 2012) e como consequência levando a perda de produtividade.

Os fatores ambientais abióticos têm grande importância para a cultura da soja, visto que ela possui características peculiares na sua adaptação aos locais de cultivo. Isso especialmente pois o fotoperíodo e a temperatura do ar são fatores de extrema importância na regulação do fotoperíodo. Nesse sentido os estudos que visem entender como o estresse atua nas plantas vem ganhando espaço e importância, visto que as plantas são organismos complexos e dinâmicos que respondem e interagem com o ambiente todo o tempo.

## 2.2 Eletrofisiologia Vegetal

Recentemente, os estudos envolvendo Eletrofisiologia Vegetal, especialmente a sinalização elétrica, uma nova área de sinalização em plantas, vêm ganhando espaço (VOLKOV, 2012). No entanto, há muito tempo se conhece esse tipo de fenômeno em plantas. Os estudos realizados por Bose em 1913 trazem a ideia de que as plantas, e até mesmo órgãos vegetais, são excitáveis e respondem a estímulos ambientais, evocando uma resposta elétrica. A onda de perturbação molecular em um tecido animal vivo sob estímulo é acompanhada por uma onda de perturbação elétrica, e em certos tipos de tecido a região estimulada é relativamente positiva a uma área não perturbada, enquanto em outras ocorre o inverso. Portanto, a característica de exibir uma resposta elétrica sob estímulo não está confinada apenas a animais, mas se estende também para os tecidos vegetais, tendo as mesmas variações elétricas obtidas em nervos e músculos (BOSE, 1902). Burdon-Sanderson em 1873, baseados em estudos de Charles Darwin com plantas carnívoras, foram os primeiros a descobrir a existência de potencial de ação em plantas, após a estimulação de folhas de *Dionaea* sp., reforçando o fato de que sinais elétricos não pertencem exclusivamente ao reino animal. Em 1926, Bose utilizou feixes vasculares isolados de um embrião vegetal para mostrar que a excitação pode ser transmitida como um impulso elétrico, que parecia ser controlado por eventos fisiológicos, semelhantes aos nervos de animais. Em 1930, foram registrados potenciais de ação (PA) com microeletrodos inseridos em células de *Nitella* sp., antes mesmo do primeiro registro intracelular de um PA em células animais (UMRATH, 1930). Revisões detalhadas sobre os potenciais de ação têm sido publicadas, indicando que as plantas superiores podem utilizar sinais elétricos para regular uma variedade de funções fisiológicas. Em 1984, foram descobertos em plantas os canais iônicos, que são a base para os potenciais de ação (Schroeder, HEDRICH; FERANDEZ, 1984). Desde então, umas das questões mais importantes tem sido se a excitabilidade em animais e plantas tem base em conjunto similar de canais iônicos.

Os estudos com sinalização elétrica vêm fornecendo pistas sobre vários processos fisiológicos e, através de ferramentas moleculares, foi possível detectar as rápidas mudanças na expressão de genes (DAVIES; SCHUSTER, 1981; STANKOVIC; DAVIES, 1997), bem como a ativação de genes inibidores de

proteases sobre o fermento, mesmo a longas distâncias (BOWLES, 1990; RYAN, 1990; WILDON et al., 1992). Além de tradução e transcrição, existem evidências de que os sinais elétricos desempenham um papel importante em muitos processos vitais da planta, como a respiração mitocondrial (FILEK, 1997), a absorção de água (DAVIES et al., 1991), o descarregamento do floema (FROMM, 1991), a translocação no floema (FROMM; BAUER, 1994), bem como a fertilização (FROMM et al., 1995).

A literatura mostra que sinais elétricos em plantas podem ser divididos em quatro tipos: potencial elétrico local (LEP), potencial de ação (PA) e potencial de variação (VP), como descrito por Wang et al. (2009) e o sinal oscilatório, descrito por Costa et al. (2011). Esses sinais podem estabelecer modelos para a geração de sinais bioelétricos em plantas. O LEP é uma resposta sub-limiar induzida por mudanças nos fatores ambientais, como solo, água, fertilidade, luz, temperatura e umidade. Embora o LEP seja apenas gerado localmente e não seja transferido para outras partes de uma planta, tem um grande impacto no estado fisiológico da planta. O PA é um sinal que se auto propaga, baseado na atividade de canais voltagem-dependentes, que respondem (e causam) alterações no potencial de membrana (DAVIES, 1993). Em contraste, o VP não pode ser evocado eletricamente e varia de acordo com o grau do estímulo, e parece “viajar” com a diminuição da magnitude e da velocidade ao se afastar do local do estímulo. O VP é um sinal não auto-perpetuante, baseado em sistema sensitivo mecânico, assim o VP é uma mudança local (em canais iônicos de células vivas), requerendo um aumento na pressão hidráulica ou produtos químicos transportados em células mortas do xilema (DAVIES, 1993; STANKOVIC et al., 1997).

Atualmente, a maior parte dos trabalhos sobre sinalização elétrica tem como foco as respostas evocadas por ferimentos ou herbívoros de parte aérea. Wildon et al. (1992) demonstraram que os sinais químicos, evocados por ferimentos, são significativamente mais lentos que as rápidas mudanças no potencial de membrana. Portanto, os sinais elétricos gerados e transmitidos a partir de partes distantes da planta podem chegar aos tecidos receptores bem antes do início da transcrição gênica se expressar, podendo atuar assim na sinalização em resposta a estímulos externos de forma sistêmica nas plantas.

Entender a sinalização sistêmica em plantas é um dos grandes desafios da ciência vegetal. Em princípio, a sinalização sistêmica induzida por

ferimentos e/ou patógenos ou ataque de herbívoros pode ser feita por sinais químicos ou elétricos. Sinais químicos são conhecidos por participar da sinalização à longa distância, propagando-se de órgão para órgão até o sistema vascular ou como compostos voláteis liberados para a atmosfera, carregando a mensagem não apenas para o próprio vegetal, mas também para as plantas vizinhas (HEIL; SILVA BUENO, 2007; HEIL; TON, 2008; HOWE; JANDER, 2008; MITHÖFER et al., 2009). Outros estudos sugerem que sob ferimento, sinais elétricos podem viajar através do floema e/ou xilema (DAVIES, 1997). Curiosamente, tais sinais elétricos também podem afetar as folhas sistemicamente, por exemplo, atuando na regulação de genes (WILDON et al., 1992; STANKOVIC; DAVIES, 1997). Entre outros genes, inibidores de proteases (IP) e RNAm para Calmodulina são estimulados em tomate (*Solanum lycopersicum*) sob ferimento e queimadura (STANKOVIC; DAVIES, 1997).

Plantas que não evocam sinais elétricos não acumulam RNAm para IP (STANKOVIC; DAVIES, 1997). Em particular, a introdução de genes de IP é impressionante porque esses inibidores de proteases são induzidos sob herbívoros por insetos como uma reação de defesa, pois os IPs podem prejudicar os “atacantes” ou simplesmente evitar que os insetos se alimentem das folhas (KOIWA et al., 1997). Embora, em princípio, reações celulares em plantas também demonstram seguir a liberação de sinais elétricos induzidos por calor, frio, ou voltagem elétrica, o que ainda é contestado é em que medida esses sinais carregam informações específicas em órgãos não especializados das plantas. Em plantas, uma variedade de fenômenos elétricos descritos tem de ser considerados como eventos de transdução de sinais. Transientes locais de voltagem, devido à resistência sistêmica, devem desaparecer após uma distância de alguns milímetros e, portanto, não têm qualquer relevância para a transferência de um sinal sistêmico. Os potenciais de ação (PA) e as variações de potencial (VP) podem carregar informação a longas distâncias de órgão para órgão (DAVIES, 2004). Como demonstrado por Felle e Zimmermann em 2007, mesmo se a ativação dos canais não for suficiente para disparar um PA, despolarizações subliminares podem ser propagadas ao longo da folha estimulada sem preceder às folhas vizinhas, porque canais suficientes não foram ativados ou por não existir ligações condutoras de sinais entre as folhas em questão. Embora tais variações de tensão sofram decréscimo, a informação pode ser carregada mais longe que um simples transiente de voltagem (DAVIES, 2004). Como os sinais elétricos são rápidos, eles podem

atuar como precursores dos sinais químicos, que podem estar localizados em toda a planta. Tais sinais podem ser liberados por injúrias causadas por insetos herbívoros, ainda assim um simples PA, como um fenômeno de sinalização, não contém muita informação a respeito do tipo causa do estresse, servindo apenas como sinal geral de estresse, causando respostas no nível da expressão gênica (WILDON et al., 1992; STANKOVIC; DAVIES, 1998).

Após um sinal elétrico ter sido gerado no simplasto, devido ao fluxo de íons pela membrana plasmática, ele pode ser transmitido via plasmodesma, posteriormente via vasos condutores, a todas às outras células (VAN BEL; EHLERS, 2004). Evidências do acoplamento elétrico das células foram constatadas em 1967 por Spanswick and Costerton, que injetaram uma corrente elétrica em células de *Nitella* e conseguiram rastreá-la a várias células de distância da célula estimulada pela corrente. Adicionalmente, em folhas de *Elodea* e coleóptilos de *Avena*, o acoplamento elétrico também foi encontrado (SPANSWICK, 1972), indicando que os plasmodesmas são transmissores de uma rede de sinalização em um nível local. Se a informação tem de ser transmitida até partes distantes da planta, a sinalização elétrica via floema parece ser utilizada para tal fim. O floema se estende continuamente ao longo de toda a planta, e os elementos de vaso podem ser considerados vias de baixa resistência para a transmissão de potenciais de ação. Devido a seu volume relativamente grande, elementos de tubo crivado com poros e a continuidade da membrana plasmática, os vasos do floema parecem adequados para a sinalização elétrica ao longo de grandes distâncias. Além disso, em muitas espécies, os feixes vasculares são cercados por uma bainha de esclerênquima, a fim de “isolar” a sinalização elétrica no floema (FROMM; LAUTNER, 2006). O baixo grau de acoplamento elétrico na direção lateral causada por poucos plasmodesmas na interface entre células companheiras e células do parênquima do floema também facilita a sinalização de longa distância (KEMPERS; AMMERLAAN; VAN BEL, 1998).

A transmissão de sinais elétricos através do floema é alcançada por canais iônicos presentes nas membranas plasmáticas, e alguns desses canais já foram identificados, sendo principalmente canais de  $\text{Ca}^{2+}$  e  $\text{K}^+$ . O potencial de membrana dos tubos crivados é determinado pela condutância ao  $\text{K}^+$  (ACHE et al., 2001), enquanto canais correspondentes à expressão dos genes AKT2/3 no floema foram encontrados em *Arabidopsis*, milho e fava (BAUER et al., 2000; DEEKEN et al., 2000; LACOMBE et al., 2000). Adicionalmente, canais de  $\text{Ca}^{2+}$  foram localizados

no floema das nervuras foliares de tabaco e Pistia, indicando que elementos de tubo crivados são ricos em canais iônicos, que podem estar envolvidos na transmissão dos sinais elétricos a longas distâncias (VOLK; FRANCESCHI, 2000). Tal fato pode ser verificado em experimentos utilizando estiletes de afídeos, onde sinais elétricos foram medidos nos elementos de tubo de *Mimosa pudica* (FROMM; ESCHRICH, 1989), *Zea mays* (FROMM; BAUER, 1994) e *Populus trichocarpa* (LAUTNER et al. 2005) em distâncias superiores a 20 centímetros de distância.

Além da sinalização de curta distância, a transmissão de sinais a longas distâncias via floema é um mecanismo conhecido em muitas plantas. Em milho, potenciais de ação gerados pelo retorno da irrigação de plantas em solos secos causa aumento nas trocas de CO<sub>2</sub> e H<sub>2</sub>O nas folhas (FROMM; FEI, 1998), enquanto que potenciais de ação disparados pelo choque por baixa temperatura na folha causa redução no transporte pelo floema entre folhas distantes (FROMM; BAUER, 1994). Em *Mimosa pudica*, ambos AP e VP causam um dobramento em conjunto dos folíolos, fazendo com que a folha pareça morta e deixe de ser atraente para um herbívoro. Além do papel da sinalização elétrica na regulação do movimento foliar, já foi demonstrado que sinais elétricos podem afetar a fotossíntese em Mimosa (KOZIOLEK et al., 2004). O estresse por flamejar uma folha emana uma VP que viaja rapidamente pela planta a fim de eliminar a taxa de absorção de CO<sub>2</sub> e reduz a taxa de transporte de elétrons pelo fotossistema II (PSII). Adicionalmente, pesquisas em árvores de “poplar” mostraram que a queima de uma folha pelo fogo também evoca sinais elétricos que viajam através da parte aérea pelas folhas adjacentes, onde a taxa de absorção de CO<sub>2</sub> e a taxa de transporte de elétrons pelo fotossistema II (PSII) são temporariamente reduzidos (LAUTNER et al., 2005). Recentemente, evidências foram encontradas in *Chara corallina* de que sinais elétricos vindos do plasmalema são transmitidos para as membranas dos tilacóides, causando mudanças na fluorescência devido ao aumento do gradiente de pH na membrana dos tilacóides (BULYCHEV; KAMZOLKINA, 2006). Mais pesquisas necessitam ser feitas a respeito das moléculas envolvidas no transporte de elétrons e absorção de CO<sub>2</sub> durante a sinalização elétrica. Claramente, o fluxo iônico envolvido e/ou a amplitude e a duração do sinal elétrico desempenham um papel chave na geração da resposta fotossintética. Por exemplo, sob chamas, plantas cálcio-eficientes apresentam amplitudes reduzidas de sinais elétricos e nenhuma resposta nas trocas gasosas foliares (LAUTNER et al., 2005).

Existem duas diferentes técnicas usadas para captar os sinais elétricos em plantas: o método invasivo (medida intracelular) e o método não invasivo (medida extracelular). O método intracelular apresenta a vantagem de captar os potenciais de membrana e sinais elétricos diretamente das células, livres de artefatos e ruídos, mas, em contrapartida, apresenta o inconveniente de ter um eletrodo inserido dentro da célula, fato que pode produzir mudanças na condição elétrica original (VOLKOV, 2012). Devido à impedância da membrana celular, a medição elétrica de uma planta pode ser gravada utilizando dispositivos de amplificação usados para captar sinais bioelétricos em animais, como os usados em eletroencefalograma (EEG) e eletrocardiograma (ECG). Segundo Volkov et al. (2000), sinais bioelétricos de plantas podem ser captados de forma não invasiva utilizando eletrodos de Ag/AgCl com uma relação sinal/ruído baixa.

Sinais bioelétricos derivados das plantas têm sido bem estudados a fim de se demonstrar que potenciais elétricos provenientes de suas atividades podem ser realmente captados utilizando um sistema de captação eletrônico (FROMM; LAUTNER, 2007), e esses sinais podem estabelecer modelos para geração de sinais bioelétricos em plantas. Segundo Mancuso (1999), um eletrodo de Ag/AgCl pode ser conectado à superfície da planta, em diferentes posições, por meio de um gel condutor do tipo comumente usado em ECG. Eletrodos podem ser conectados por cabos a um *input* eletrônico de alta impedância com vários canais e um eletrodo idêntico pode ser colocado em uma região distal da planta ou no solo, para servir como eletrodo de referência. Quando todos os canais apresentarem potenciais estáveis, a planta pode ser estimulada eletricamente em seu ápice (ex. 3V por 2s), ou por outro estímulo qualquer (calor, choque térmico, luz) aplicado sobre a folha, e as respostas elétricas de um estímulo apical podem ser percebidas por todos os eletrodos, do ápice à base da planta, indicando que a transmissão de um sinal elétrico ocorre através da planta. No geral, o conhecimento da sinalização elétrica em plantas pode ajudar a desvendar a natureza e o funcionamento da troca de informações dentro de células vegetais e órgãos, e podendo assim dar origem a questões novas e fascinantes a respeito da dinâmica da sinalização em plantas e suas possíveis utilidades para a produção de conhecimentos e tecnologias, baseadas em eletrofisiologia vegetal. Estudos futuros devem ser direcionados para um melhor entendimento de como a sinalização elétrica controla mecanismos, a

interligação entre fluxo de íons e respostas fisiológicas, além da composição de canais iônicos que participam da sinalização elétrica.

### **2.3 Estudo de Séries Temporais**

Uma grande marca dos organismos vivos é que seus processos são inconstantes ao longo do tempo. Processos em todas as escalas celulares, como ciclo de crescimento, divisão celular, flutuações de voltagem em membranas celulares excitáveis, respiração, regulação da pressão sanguínea, e ciclos de sono-vigília são exemplos de processos complexos encontrados nos organismos vivos. O interesse no estudo das flutuações ao longo do tempo vem ganhando força recentemente, devido ao grande reconhecimento de que sistemas determinísticos podem apresentar dinâmica caótica: ritmos aperiódicos sensíveis às condições iniciais (DEGN, 1986; MAYER-KRESS, 1986; KRASNER, 1990). Assim, a análise de modelos teóricos e experimentos em situações controladas têm fornecido boas informações de que o caos e outras dinâmicas podem ser encontrados em sistemas biológicos (GUEVARA, 1981; CHIALVO, 1990).

O reconhecimento da importância de fenômenos não lineares em sistemas fisiológicos tem uma longa história. O seu começo é marcado pelo trabalho de Van der Pol e van der Mark no começo do século. O trabalho ganhador do prêmio Nobel, desenvolvido por Hodgkin e Huxley, relatou a dinâmica das membranas de células excitáveis para um sistema acoplado a equações diferenciais não lineares. Contudo é recente a percepção de que sistemas com dinâmicas não lineares podem apresentar caos determinístico. Isto causou forte impacto nas pesquisas em análises de séries temporais em fisiologia e medicina (GLASS, 1991). No passado, os cientistas eram levados a conjecturar que séries temporais complexas, como as de um neurônio, eram bem estabelecidas com um processo aleatório (GERSTEIN, 1964). No entanto, agora há uma forte inclinação a interpretar séries temporais fisiológicas em termos de outras dinâmicas, como o caos por exemplo. Às vezes, essa interpretação envolve o uso de técnicas de análise de séries temporais motivadas pela dinâmica caótica de sistemas.

Uma série temporal consiste em um conjunto de observações, medidas ou valores sob um determinado intervalo de tempo. Se estas observações consecutivas são dependentes uma das outras, é possível conseguir-se uma

previsão e, assim, fornecer bases para compreender o comportamento do evento analisado. De acordo com Morettin e Tolo (2004), ao se realizar um estudo de séries temporais, pode-se estar interessado em investigar o mecanismo gerador da série temporal, fazer previsões de valores futuros da série, descrever o comportamento da série, como tendências e sazonalidades, por exemplo, ou procurar periodicidades relevantes nos dados. Esses autores também ressaltam que os modelos devem ser simples, com o menor número de parâmetros possível, e sua utilização não deve apresentar dificuldades às pessoas interessadas em manipulá-los, entendendo-se por modelos uma descrição matemática de uma série temporal. A análise de séries temporais de dados fisiológicos em plantas já foi utilizada para compreender os efeitos dos distúrbios ambientais no metabolismo vegetal (LÜTTGE; BECK, 1992; SHABALA et al., 1997; HÜTT; LÜTTGE, 2002; SOUZA et al., 2004).

Séries temporais obtidas de seres vivos tendem a apresentar uma dinâmica não linear, variando em complexidade e regularidade. Recentemente, uma nova abordagem matemática denominada Entropia Aproximada (ApEn) tem sido utilizada como medida de quantificação da regularidade em séries temporais de dados. Matematicamente, ApEn é parte de um desenvolvimento teórico geral, como um parâmetro teórico de informação natural (PINCUS, 1992). Em aplicações para uma variedade de configurações médicas, tem se evidenciado que séries temporais de indivíduos doentes e envelhecendo apresentam diminuição significativa nos valores de ApEn. A literatura científica associa sistemas que apresentam fisiologia comprometida a uma maior regularidade no tempo, enquanto eventos fisiológicos normativos apresentam grande irregularidade (aleatoriedade, a complexidade) (PINCUS et al., 1991, 1992, 1993). Uma quantificação adequada da regularidade dos conjuntos de dados tem o potencial de fornecer informações em uma ampla gama de ambientes clínicos e fisiológicos, no campo do diagnóstico e da previsão. ApEn atribui um número não negativo a uma sequência ou série temporal, onde maiores valores correspondem a uma maior aleatoriedade aparente ou irregularidade da série, e menores valores correspondem a mais casos de previsibilidade, ou uma menor irregularidade (PINCUS, 1991; KAPLAN, 1991).

A relação entre complexidade e estabilidade fisiológica já foi observada em diferentes tipos de sistemas biológicos. Uma grande irregularidade (complexidade) de ritmos cardíacos, dinâmica respiratória e medidas de pressão sanguínea já foram associadas a um status de fisiologia saudável em humanos.

Reciprocamente, a redução da complexidade nos mesmos parâmetros fisiológicos também é associada com processos de envelhecimento e com o aumento do risco de morbidade e mortalidade (KAPLAN, 1991; PINCUS, 1995). Analogamente em plantas, há registros prévios de que uma dinâmica temporal mais complexa em parâmetros como a fotossíntese, reações enzimáticas e uma grande classe de fluxos é associada com a maior capacidade de manter a homeostase do sistema (HÜTT; LÜTTGE, 2002; SOUZA et. al., 2004a). As plantas, como sistemas biológicos, são sistemas dinâmicos, termodinamicamente em não equilíbrio. Isso implica que esses sistemas e especialmente seus processos fisiológicos, evoluam no tempo, apresentando uma grande variedade de comportamentos dinâmicos. Contudo, sabe-se que tais comportamentos podem ser significativamente influenciados ou alterados em função de estímulos ambientais específicos (HÜTT; LÜTTGE, 2002; SOUZA et al., 2004a).

A quantificação da complexidade de sinais fisiológicos em organismos saudáveis e doentes já é foco de considerável atenção (PINCUS, 2001; GOLDBERGER, 2002, RICHMAN; MOORMAN, 2000). Tais métricas são potencialmente importantes em aplicações que objetivam a avaliação de ambos os modelos de dinâmicas em controle de sistemas biológicos e também diagnósticos. Sabe-se que uma vasta classe de doenças, assim como o envelhecimento, parece degradar o conteúdo de informação fisiológica e reduzir a capacidade de adaptação do indivíduo. Por conseguinte, a perda de complexidade tem sido proposta como uma característica genérica de dinâmicas patológicas (PINCUS, 2001; RICHMAN; MOORMAN, 2000). Sistemas doentes, quando associados à emergência de um comportamento mais regular, apresentam valores reduzidos de entropia quando comparados a sistemas saudáveis (RICHMAN; MOORMAN, 2000). Adicionalmente, podemos verificar que as plantas tendem a se manter sob condições de baixa entropia, sendo esse um requisito fundamental à manutenção da vida (KSENZHEK; VOLKOV, 1998).

Finalmente, há a suposição básica que norteia a análise de séries temporais, que cita que um sistema causal mais ou menos constante, relacionado com o tempo, exerce influência sobre os dados no passado e pode continuar a fazê-lo no futuro. Este sistema causal costuma atuar criando padrões não aleatórios que podem ser detectados em um gráfico de série temporal, ou mediante algum outro processo estatístico. O objetivo da análise de séries temporais é identificar padrões

não aleatórios na variável de interesse, e a observação desse comportamento passado pode permitir fazer previsões sobre o futuro, no caso de séries temporais com dinâmica caótica.

## 3 MATERIAL E MÉTODOS

### 3.1 Planta Modelo e Local dos Experimentos

Para todos os experimentos, foram utilizadas plantas de soja Intacta® cultivar AS3730IPRO (BT + RR). As sementes foram germinadas e cultivadas em substrato composto por vermiculita expandida, em copos de isopor de 180 mL em estufa climatizada tipo Fitotron® (lâmpadas mistas Osram E=40, 250 Watts, 220 Volts e Lâmpadas quentes Osram 250 Watts4Y, na proporção de 8 X 12), que forneciam uma DFFFA de  $500 \mu\text{mols s}^{-1}$ . A temperatura média durante toda a condução dos experimentos foi mantida em  $\pm 26,5 \text{ }^\circ\text{C}$  no período de luz e  $\pm 22 \text{ }^\circ\text{C}$  no período de escuro e umidade relativa do ar em 60%. O fotoperíodo adotado foi de 12h/12h de luz/escuro. O suprimento de água e nutrientes era fornecido pela irrigação diária das plantas com 50 mL de solução nutritiva de Hoogland.

### 3.2 Experimentos

Os experimentos visavam a comparação dos sinais bioelétricos evocados das plantas em situação controle (condição padrão do ambiente, antes do estímulo), com os sinais obtidos de plantas após a aplicação dos estímulos. Toda a captação de dados foi realizada dentro de um dispositivo de blindagem eletrostática (Gaiola de Faraday) de dimensões 0,90 m de comprimento por 0,90 m de largura e 1m de altura. A gaiola faz com que os efeitos da interferência dos campos elétricos sejam nulos em seu interior. Isso possibilita uma captação mais limpa dos sinais elétricos evocados nas plantas e evitando assim interferências e ruídos externos. No dia anterior a cada sessão experimental de coleta de dados, os eletrodos de agulha foram inseridos espaçados 1 cm nos caules das plantas para aclimação das mesmas. Isso se faz necessário pois, segundo o trabalho de Volkov (1995), a inserção dos eletrodos induz potenciais de ação e flutuações locais na variação de potencial, que são estabilizados em poucas horas, com o desaparecimento dos potenciais de ação (VOLKOV, 1995).

Concomitantemente à coleta de dados das plantas, era feita a coleta de dados de eletrodos “abertos”, eletrodos que não estavam conectados as plantas, captando assim somente dados do ambiente. Tal procedimento foi adotado em

todas as coletas de dados realizadas visando a posterior comparação dos sinais de eletrodos “abertos” com os sinais captados das plantas. Devido aos cuidados e procedimentos necessários em experimentos envolvendo eletrofisiologia, foram determinados os estímulos que poderiam ser aplicados sem manusear diretamente as plantas. De acordo com a possibilidade de execução foram selecionados os seguintes estímulos:

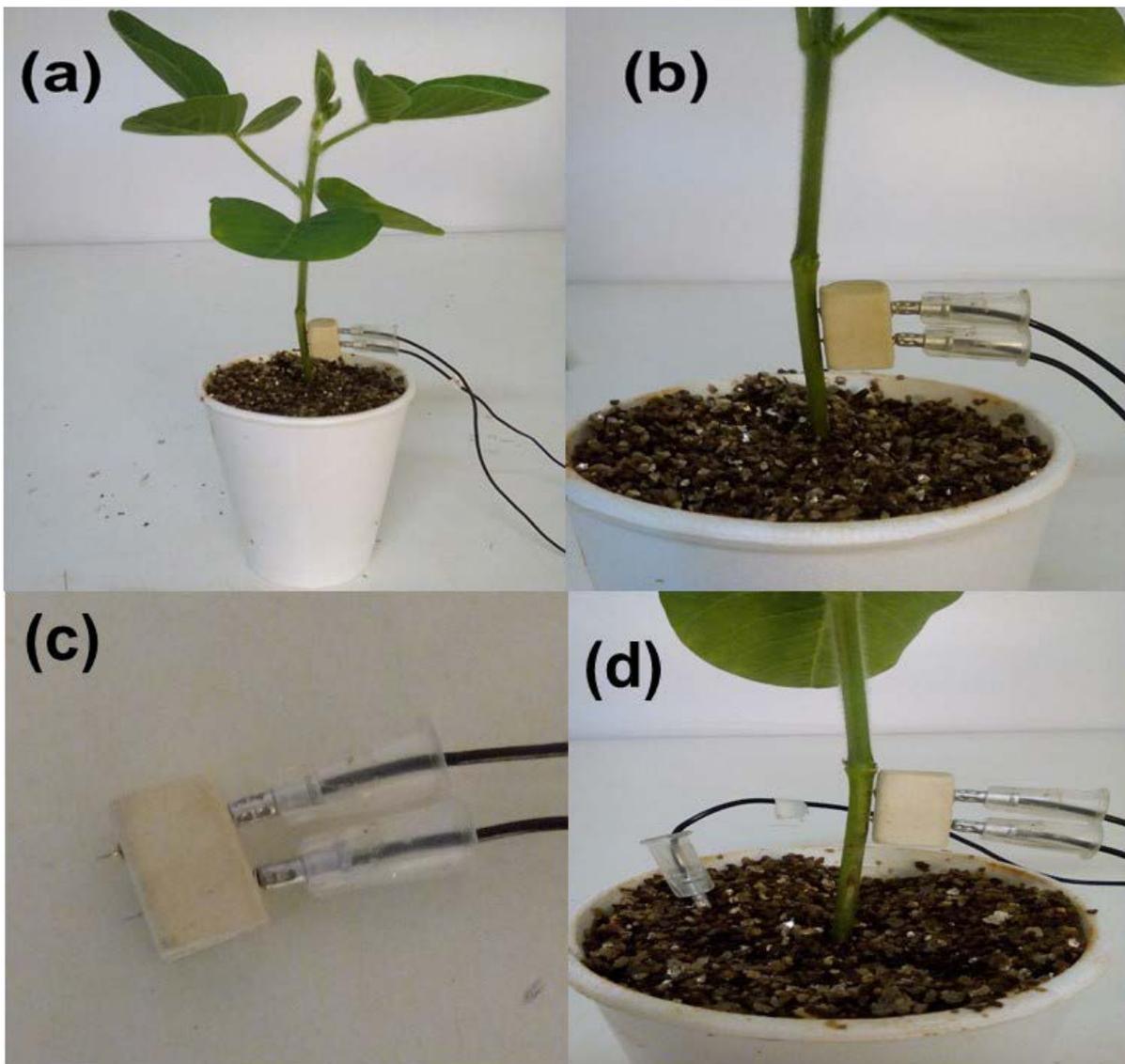
1. Estímulo Osmótico: Foi realizada a aplicação de 50 ml solução de manitol com potencial osmótico de -2 Mpa. Foram utilizadas 134 plantas para coleta de dados.
2. Estímulo térmico: A temperatura média do Fitotron® onde as plantas se encontravam foi reduzida de 26°C para 15°C. Foram utilizadas 72 plantas para coleta de dados.
3. Estímulo luminoso: Foi realizada a redução da Densidade de Fluxo de Fótons Fotossinteticamente ativos de 500  $\mu\text{mol s}^{-1}$  para 50  $\mu\text{mol s}^{-1}$ , através da cobertura da gaiola de Faraday com um pano preto. Foram utilizadas 135 plantas para coleta de dados.

### **3.3 Captação dos Dados**

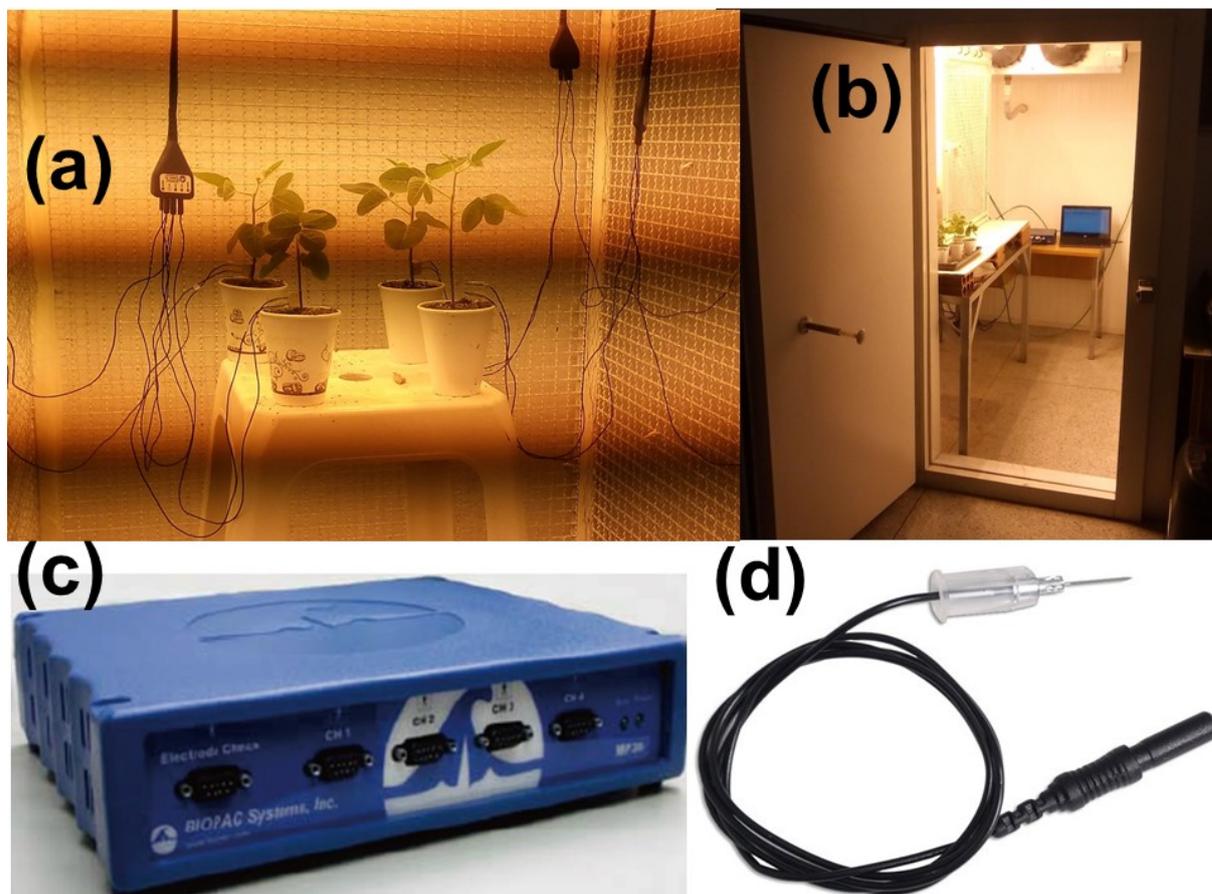
Para a captação dos sinais bioelétricos evocados das plantas, foi utilizado o sistema eletrônico de Aquisição de dados MP36 da Biopac Systems. O sistema possui alta impedância de entrada e filtros “passa-alta” (corte de frequência de 0,5 Hz) e “passa-baixa” (frequência de corte de 1,5 KHz). O sinal é amplificado para garantir um sinal limpo e forte. Os sinais foram adquiridos em uma frequência de amostragem de 125 Hz, com um intervalo de ganho de 20.000x. O MP36 tem um microprocessador interno para controlar a aquisição de dados e de comunicação (via USB) com o computador. O MP36 converte os sinais de entrada em sinais digitais que podem ser processados em um computador. O sistema possui quatro canais de entrada analógicos. Os eletrodos de agulha (+) e (-) (modelo - EL452) foram ligados ao MP36 através de transdutores (módulo SS2LA). Para coleta de dados, foi utilizado o protocolo EEG que varia de 0,5 a 100 Hz, com “frequência de corte” em

60 Hz, que é a frequência das redes elétricas, evitando assim ruído. Para todos os experimentos, captaram-se os sinais de plantas antes da aplicação dos estímulos (controle) e após a aplicação dos estímulos (plantas estimuladas).

**Figura 1-** (a) Planta modelo utilizada nos experimentos. (b) Inserção dos eletrodos no caule da planta. (c) Detalhe da ponta dos eletrodos. (d) Detalhe do aterramento e inserção dos eletrodos.



**Figura 2-** Plantas conectadas aos eletrodos dentro da Gaiola de Faraday (a). Vista externa do Fitotron com equipamentos de coleta de dados ao fundo (b). Biopac MP-36 (c) e eletrodos de agulha utilizados (d).



### 3.4 Análises dos Dados

Os dados bioelétricos coletados foram analisados na forma de séries temporais  $\Delta V = \{\Delta v_1, \Delta v_2, \dots, \Delta v_N\}$ , onde  $\Delta v_1$  é a diferença de potencial captados pelos eletrodos inseridos nas plantas, como descrito anteriormente. As séries analisadas correspondem a amostras com um comprimento total  $N = 75.000$  pontos, correspondendo a 600 segundos (10 minutos) de coleta de dados de plantas em situação controle e plantas estimuladas. Foram obtidas 268 séries temporais dos experimentos envolvendo estímulo osmótico, 144 séries dos experimentos envolvendo estímulo térmico e 270 séries dos experimentos envolvendo estímulo luminoso. As séries temporais foram analisadas por diferentes métodos, buscando caracterizar a dinâmica temporal dos sinais, focando na comparação entre a dinâmica antes e após a aplicação dos estímulos.

### 3.4.1 Análises descritivas e quantitativas

Foram realizadas análises descritivas das séries temporais, utilizando ferramentas simples como o Microsoft Excel®, para melhor caracterizar a análise visual das séries. De cada série temporal, obteve-se um valor de voltagem mínima e máxima atingidas e uma média de todos os valores da série. Com esses valores, foram feitas médias de todas as séries de um mesmo experimento. Também foi verificado se as séries temporais são estacionárias ou não. Uma série temporal é dita estacionária quando ela se desenvolve no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável. Na prática, a maioria das séries que encontramos apresentam algum tipo de não estacionariedade, por exemplo, tendência, como no caso das séries temporais de dados bioelétricos de plantas.

Para melhor caracterização da dinâmica das séries temporais, foram realizadas análises do espectro de frequências presente nos sinais coletados. Para isso foi utilizada a Transformada Rápida de Fourier (FFT) nas séries temporais, utilizando o software Biopac Student Lab. A análise de FFT torna possível determinar o espectro de frequências presentes nas séries temporais, pois ela “decodifica” um sinal e o descreve na forma de espectro de frequências, com valores comumente expressados em watts por Hertz (W/Hz). Ela leva um vector que representa os dados como uma função de tempo e calcula o conteúdo dos dados de frequência. Nota-se que o resultado da função FFT é um complexo de dados, o que representa dados de magnitude e fase em uma determinada frequência.

Outra técnica utilizada para verificar a presença de informação nas séries temporais coletadas foi a análise de autocorrelação dos valores das mesmas. A autocorrelação possibilita verificar a correlação linear dos valores de uma série temporal, defasados por um lag  $K$ , onde  $K$  é a diferença entre o tempo presente e o tempo passado. Para essas análises utilizou-se o software Statistix 9. As análises foram feitas utilizando um lag  $K$  variando de 1 até 150. Em geral, para uma série temporal com  $n$  elementos, a autocorrelação com lag  $k$  é dada pela expressão:

$$r_k = \frac{\sum_{t=1}^{N-k} (s_t - \bar{s})(s_{t+k} - \bar{s})}{\sum_{t=1}^n (s_t - \bar{s})^2} \cdot$$

Onde  $-1 \leq r_k \leq 1$  e  $\bar{s}$  é a média da série original de  $n$  elementos, admitida estacionária.

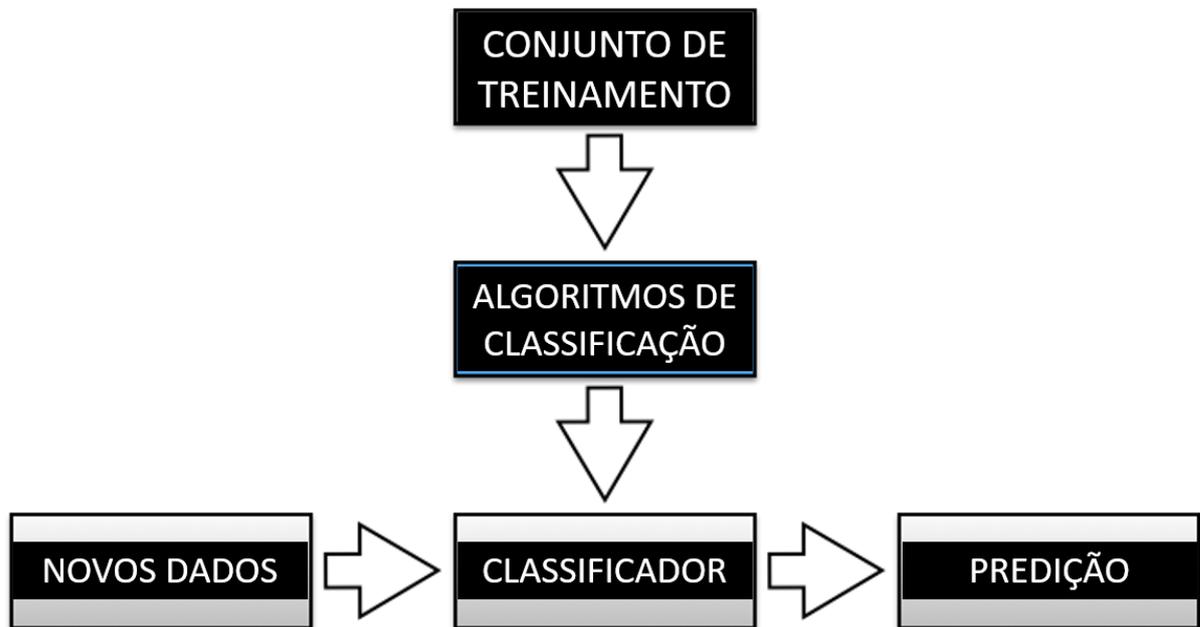
### 3.4.2 Reconhecimento de padrões: Técnicas de Classificação

O uso de algoritmos se faz pela necessidade de um método robusto para a classificação automática das séries temporais, entre antes e após a aplicação dos estímulos (geral e para cada estímulo). Foram comparados o desempenho de classificação de cinco diferentes algoritmos. O objetivo foi encontrar um algoritmo que possa com sucesso ser treinado em um pequeno conjunto de dados das séries temporais, para o reconhecimento das mesmas. Foram utilizados algoritmos de classificação utilizados em técnicas de redes neurais de aprendizagem, que calculam a probabilidade de que séries temporais pertençam a cada uma das classes possíveis, ou seja, predizer a classe mais provável.

Para o correto funcionamento da técnica, os algoritmos requerem um conjunto de dados que já esteja previamente classificado (conjunto de treinamento), ou seja, um conjunto com amostras em que as classes já estejam determinadas. Para cada conjunto de séries analisadas foram usados 8 diferentes tamanhos de janela (0, 100, 500, 1000, 5000, 10000, 20000, 25000). As janelas correspondem ao número de amostras que são geradas a partir das séries originais, após o uso da aritmética intervalar.

Como conjuntos de treinamento foram utilizadas séries temporais que representavam de 30% a 80% do número total de séries utilizadas no experimento. Cada método de classificação foi executado 20 vezes para realizar o teste estatístico Wilcoxon. Nas páginas abaixo segue uma descrição dos algoritmos de classificação utilizados nesse trabalho.

**Figura 3-** Esquema representando o funcionamento dos algoritmos de classificação na busca por padrões nas séries temporais.



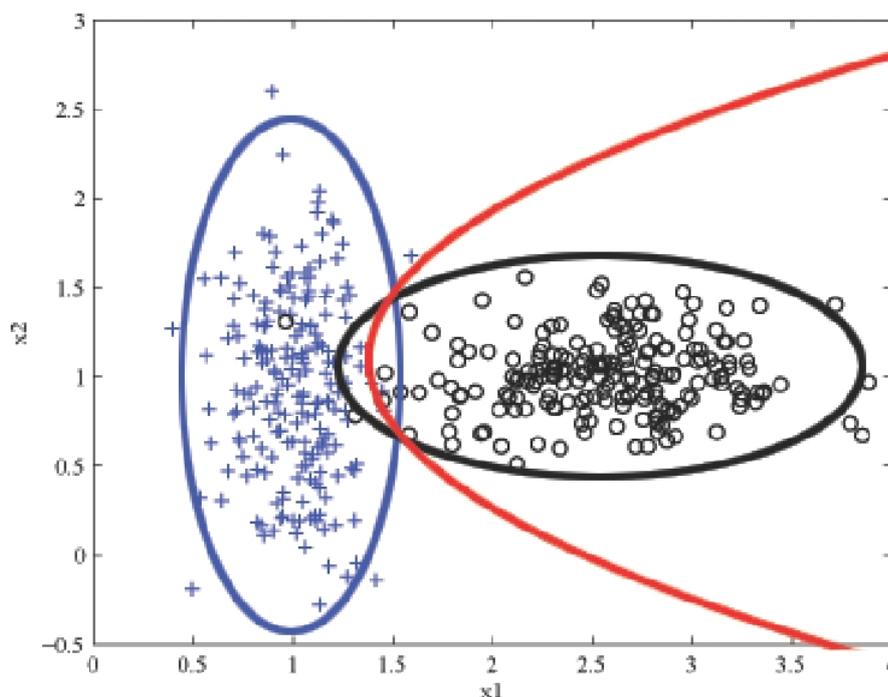
#### 3.4.2.1 Classificador Bayesiano (BAYES)

Um Classificador Bayesiano estima a probabilidade de que uma dada série pertença a certa classe. Essa probabilidade pode ser derivada do Teorema de Bayes (Jaynes, 2003):

$$p(\omega_i | \mathbf{x}) = \frac{p(\mathbf{x} | \omega_i) P(\omega_i)}{p(\mathbf{x})} .$$

Onde  $p(\mathbf{x} | \omega_i)$  denota a probabilidade de se observar o vetor característica  $\mathbf{x}$  dada a classe  $\omega_i$ ,  $P(\omega_i)$  é a propriedade a priori da classe  $\omega_i$ , e  $p(\mathbf{x})$  é a probabilidade de  $\mathbf{x}$ . A fim de estimar  $p(\mathbf{x} | \omega_i)$ , considerou-se que a função de verossimilhança é Gausiana, podendo assim estimar os parâmetros dos conjuntos de dados (DUDA, 2000). Na figura abaixo pode-se visualizar a distribuição estimada desenhando um contorno da densidade. O limite da decisão onde a probabilidade de cada classe de entrada é igual a saída é mostrado em vermelho.

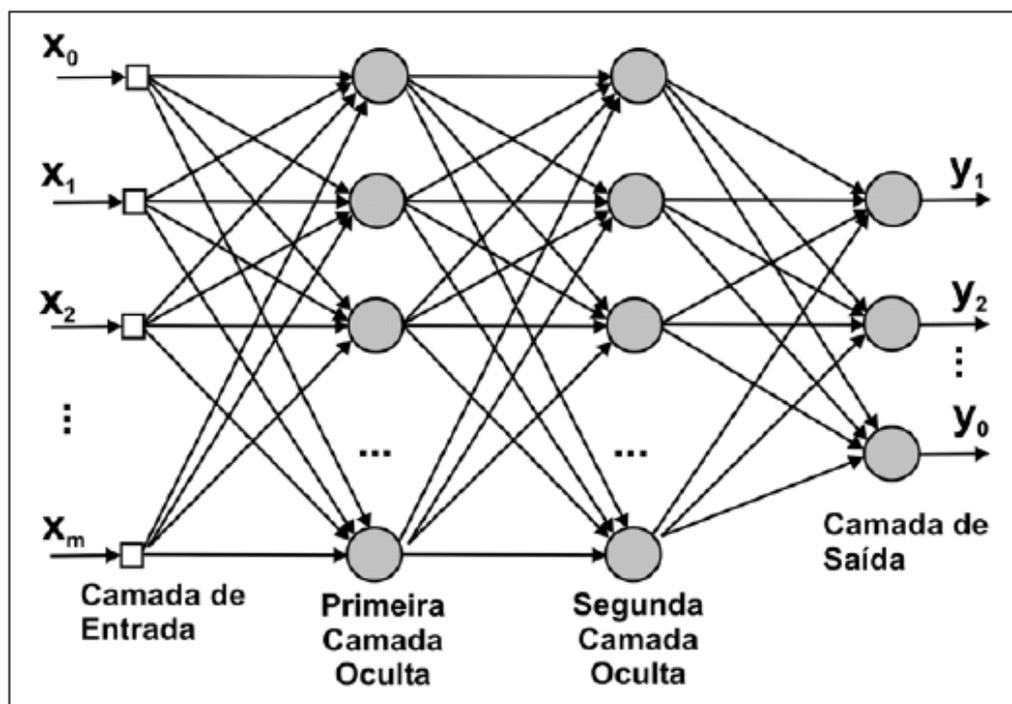
**Figura 4-** Exemplo de Classificador Bayesiano (BAYES)



### 3.4.2.2 Rede Neural Artificial Multicamadas (ANN)

O Classificador ANN é uma rede neural composta de várias camadas de neurônios com o objetivo de resolver problemas multi-classe (HAYKIN, 1999). A Entrada para cada camada é uma soma ponderada da saída a partir da camada anterior. O número de neurônios da primeira camada é igual ao número de características de entrada, enquanto que o número de neurônios da camada de saída é igual ao número de classes. A rede neural atribui um vetor de características extraídas de uma série  $x$  a classe  $\omega_q$  se o neurônio de saída  $q$ -th tem a maior ativação. Foi utilizada a biblioteca “Fast Artificial Neural Network (FANN)” disponibilizada em Nissen (2003), com duas camadas ocultas cada constituída por 8 neurônios. A rede foi treinada com o algoritmo “back propagation” (HAYKIN, 1999). O valor de épocas máximo foi definido com 70.000, e o erro desejado na forma de 0.0001. A taxa de aprendizagem foi ajustada para 0,1.

**Figura 5-** Exemplo de Rede Neural Multicamadas.

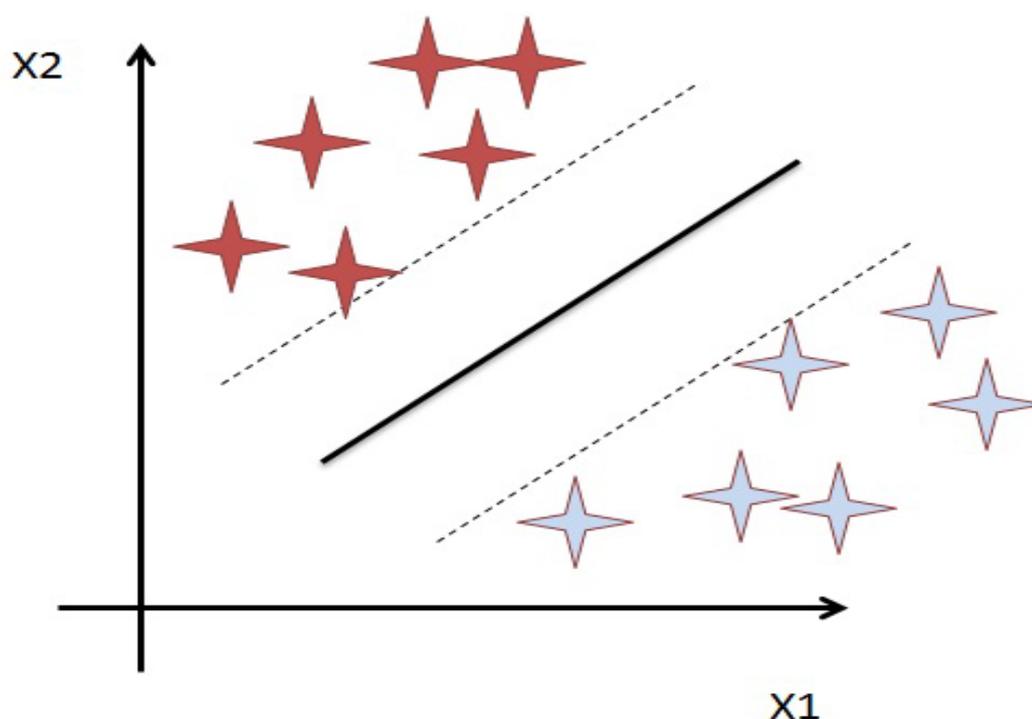


### 3.4.2.3 Máquinas de Vetores De Suporte (SVM)

Enquanto em ANN a aprendizagem é baseada no princípio de minimização do risco empírico, SVM está enraizado no princípio de minimização do risco estrutural (VAPNIK, 1999; SCHÖLKOPF, 2002; CORTES, 1995). Ele visa o estabelecimento de uma função discriminativa ideal entre duas classes de padrões ao realizar o compromisso entre a generalização e sobreajuste. O algoritmo de treinamento SVM constrói o hiperplano ideal que separa as duas classes (VAPNIK, 1999). A fim de estender a classificação de linear para não linear, o truque do Kernel é utilizado (SCHÖLKOPF, 2002), em que as funções de não linearidade de Kernel mapeiam os dados de entrada em características espaciais de alta dimensão, de forma computacionalmente eficiente. Para problemas de classificação com múltiplas classes, duas abordagens são comumente utilizadas para SVM binárias, um-contra-um que trata-se de uma generalização do SVM utilizado para mais de duas classes e um-contra-todos, que tomando  $M$  classes, constrói  $M$  superfícies de decisão ( $g_1, \dots, g_M$ ), onde cada superfície de decisão é treinada para separar uma classe das

classes restantes (WITTEN, 2005). Ambas as estratégias tendem a conduzir a resultados semelhantes em termos de precisão de classificação, contudo o primeiro, que foi adotado aqui, geralmente requer tempo de treinamento menor, apesar de incorrer em maior número de decomposições binárias. Para os experimentos desta tese, foi adotada a biblioteca LibSVM (CHANG, 2011). Em outras palavras, o que uma SVM faz é encontrar uma linha de separação, mais comumente chamada de hiperplano entre dados de duas classes. Essa linha busca maximizar a distância entre os pontos mais próximos em relação a cada uma das classes, ver imagem:

**Figura 6-** Exemplo de Máquinas de vetores de suporte (SVM).

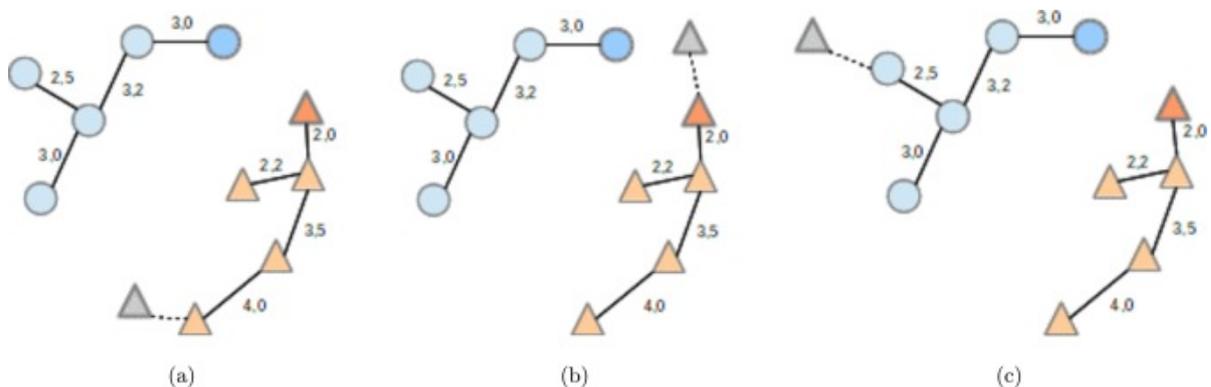


#### 3.4.2.4 Floresta de Caminhos Ótimos (OPF)

O Classificador OPF modela os problemas de reconhecimento de padrões como uma tarefa de partição em grafo, em que um conjunto pré-definido de amostras de cada classe (i.e. protótipos) competem por um caminho de custo mínimo para o resto das amostras. Isso resulta em uma coleção de árvores de caminhos “ótimos” enraizadas nos nós de protótipos, construindo uma “floresta de caminhos ótimos” considerando todas as amostras de treinamento. As amostras de

teste são classificadas por meio de avaliação gradativa dos caminhos ideais dos protótipos, como se eles fossem parte da floresta, e atribuindo rótulos as raízes mais fortemente ligadas. A noção de ótima via de conectividade vem da minimização de uma função do caminho pelo custo (FALCÃO, 2004). Um classificador OPF pode ser projetado usando uma função caminho-custo suave. Apesar de existirem duas versões diferentes de classificadores OPF (PAPA, 2008, 2009, 2012), nesse trabalho foi usada a versão mais amplamente utilizada. A classificação é baseada em florestas supervisionadas para modelar as amostras como nós de um gráfico completo. Os elementos mais representativos de cada classe usada como conjunto de treinamento são selecionados como protótipos, correspondendo a elementos pertencentes às regiões fronteiriças entre as classes. Os protótipos participam de um processo de competição contra outras amostras, oferecendo caminhos de baixo custo e seus respectivos rótulos. No final deste processo, é obtido um conjunto particionado de treinamento em árvores de trajetórias ótimas, e a união delas leva a uma floresta de caminho ótimo.

**Figura 7-** Exemplo de Floresta de Caminhos Ótimos (OPF)

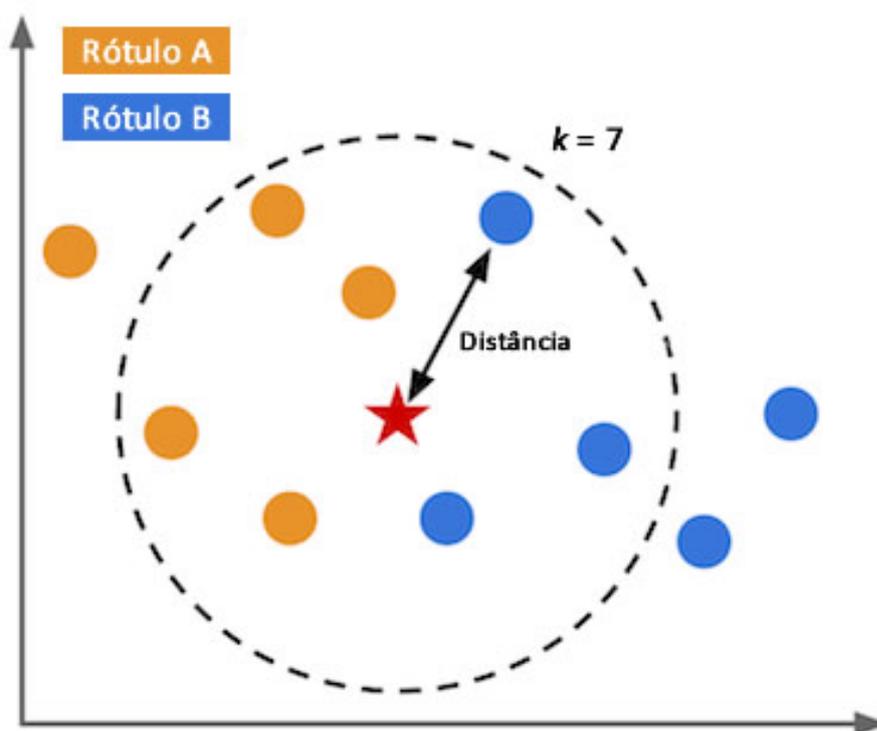


#### 3.4.2.5 K-vizinhos mais próximos (KNN)

KNN é um algoritmo muito simples que trabalha muito bem em diferentes aplicações. Em contraste com o OPF, KNN usa todas as amostras de treinamento como protótipos. O KNN requer um parâmetro de entrada definindo o número de vizinhos  $k$  que contribuem para a classificação de uma amostra (COOMANS, 1982; HAL, 2008). A fim de se classificar uma amostra de teste  $t$ , a

maioria dos rótulos em uma região contendo amostras de treinamento  $k$  e centrado em  $t$ , determinam o rótulo de  $t$ . Nota-se que, para  $k=1$ , a amostra de teste  $t$  é classificada como a classe de formação da amostra mais próxima. Para os experimentos desse trabalho, foi definido o valor de  $k$  como o melhor valor de uma grade de pesquisa no intervalo  $[1, \lfloor m/5 \rfloor]$  em passos de dois; em que  $m$  é o número de amostras de treino.

**Figura 8-** Exemplo de K-vizinhos mais próximos (KNN)



#### 3.4.2.6 Métricas de Avaliação Estatística

Em relação à taxa de reconhecimento dos algoritmos, foi utilizada a medida de acurácia proposta por Papa et. al (2009), onde se houver duas séries, por exemplo, com tamanhos muito diferentes e um classificador sempre atribui o rótulo da maior classe, a sua acurácia irá cair drasticamente devido à alta taxa de erros na menor classe. A acurácia é medida tendo em conta que as classes podem ter diferentes tamanhos em  $D_2$ . Definindo:

$$e_{i,1} = \frac{FP_i}{|D_2| - |D_2^i|} \quad \text{e} \quad e_{i,2} = \frac{FN_i}{|D_2^i|}, \quad i = 1, 2, \dots, K,$$

Onde  $K$  representa o número de classes,  $|D_2^i|$ , preocupações com o número de amostras em  $D_2$  que vêm da série  $i$ , e  $FP_i$  e  $FN_i$  significam falsos positivos e falsos negativos para a série  $i$ , respectivamente. Isto é  $FP_i$  é o número de amostras de outra série que foi classificada como sendo da classe  $i$  em  $D_2$  e  $FN_i$  é o número de amostras da série  $i$  que foram incorretamente classificadas como sendo de outras classes em  $D_2$ . Os termos de erro  $e_{i,1}$  e  $e_{i,2}$  são então usados para definir o erro total da classe  $i$ :

$$E_i = e_{i,1} + e_{i,2}.$$

Finalmente, a acurácia ( $Acc$ ) é então definida como segue:

$$Acc = 1 - \frac{\sum_{i=1}^K E_i}{2K}.$$

Sensibilidade ( $Se$ ), muitas vezes chamada de recolha, é a razão entre o número de séries corretamente classificadas de uma determinada classe o número total de séries nessa classe (incluindo séries classificadas incorretamente).

$$Se = \frac{\text{Positivos verdadeiros}}{\text{Positivos verdadeiros} + \text{Falsos negativos}} \cdot$$

Onde positivos verdadeiros e falsos negativos são o número de séries a partir de uma dada classe que foram corretas ou incorretamente classificadas, respectivamente. O Valor Preditivo Positivo ( $PPV$ ), muitas vezes chamado de precisão, é a relação entre as séries classificadas corretamente a partir de uma dada classe.

$$PPV = \frac{\text{Positivos verdadeiros}}{\text{Positivos verdadeiros} + \text{Falsos negativos}} \cdot$$

Onde falsos positivos indicam o número de séries classificadas incorretamente como pertencentes a classe considerada. Finalmente, como uma métrica de desempenho global, foi calculada a média de F1-score de todas as classes. O F1-score de uma determinada classe é calculado como a média aritmética dos valores de Se e PPV para essa classe:

$$F_1\text{-score} = 2 \left( \frac{Se \times PPV}{Se + PPV} \right) \cdot$$

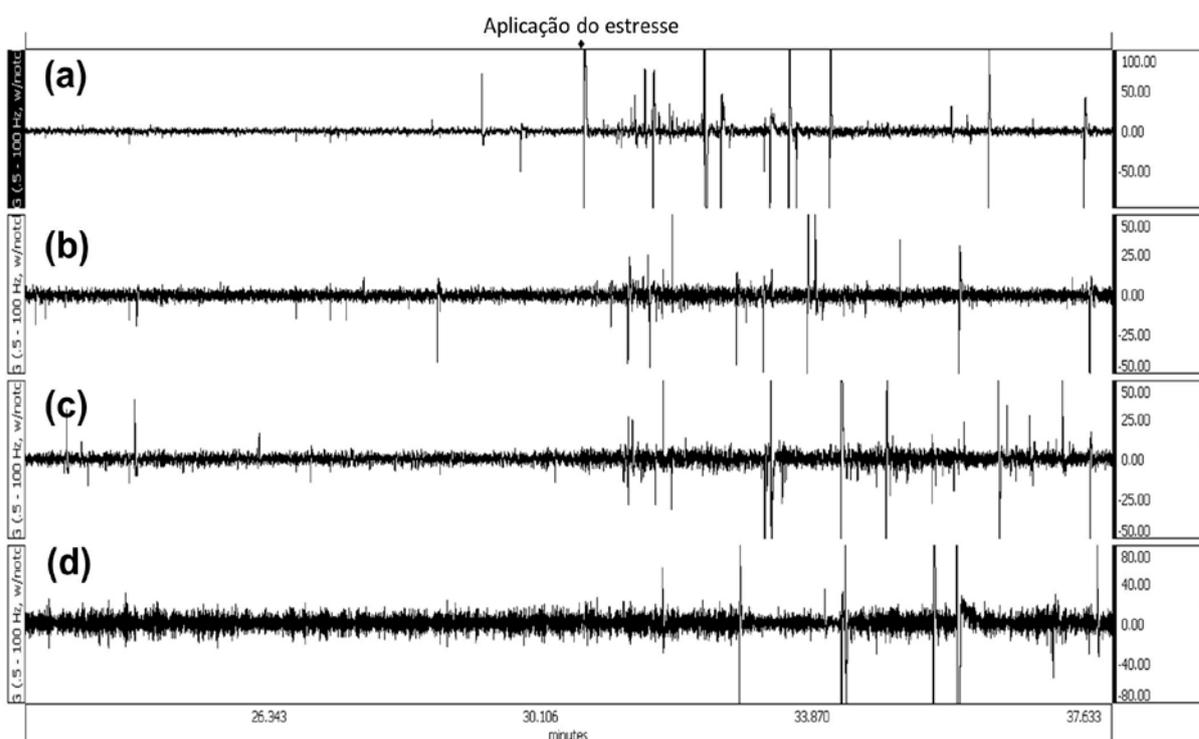
Essas quatro métricas permitem avaliar de forma confiável o desempenho da classificação dos algoritmos considerados nesse trabalho. As métricas de desempenho são relatadas como as médias de mais de 10 repetições de treinamento do classificador e de teste. Todo o treinamento e teste dos algoritmos de classificação foram feitos em um computador com processador Intel i7 5500U, com 8GB de RAM, utilizando o Linux como sistema operacional.

## 4 RESULTADOS

### 4.1 Análises Descritivas das Séries Temporais

Uma forma de analisar os dados é a observação visual das séries temporais. Comparando visualmente as séries temporais, verifica-se nítida diferença entre séries obtidas antes da aplicação dos estímulos e séries coletadas após a aplicação dos estímulos (Figura 3). A figura mostra séries temporais de 4 plantas diferentes, indicando na parte superior o momento da aplicação do estímulo.

**Figura 9-** Séries temporais (20 minutos) obtidas de plantas de soja antes e após a aplicação de estímulos: osmótico (a) e (b), térmico (c) e luminoso (d).



Pode-se observar que a aplicação dos estímulos osmótico, térmico e luminoso, provocaram uma disparada de picos de voltagem em sequência. Esses picos apresentavam amplitudes variáveis, demonstrando a intensa atividade bioelétrica evocada em resposta a essas situações adversas. A análise visual das séries temporais, portanto, revela intensa atividade bioelétrica evocada pelos estímulos abióticos aplicados. Portanto, a análise visual proporciona uma clara

diferenciação entre séries temporais coletadas antes e após aplicação dos estímulos. Contudo tal análise não permite a diferenciação entre cada tipo de estímulo aplicado, visto que os picos de voltagem causados pelos estímulos não possuem um padrão visual facilmente detectável, a ponto de possibilitar a discriminação dos estímulos.

Visando quantificar melhor os resultados observados pela análise visual, foram calculadas as médias dos valores de picos de voltagem e média de todos os valores de voltagem de cada série temporal. Dessa forma, pode-se visualizar numericamente as diferenças entre as séries temporais coletadas antes e após a aplicação dos estímulos. Com esses dados é possível verificar que a aplicação dos estímulos (térmico, luminoso e osmótico) causou alterações significativas nos valores médios dos parâmetros avaliados antes e após a aplicação dos mesmos (Tabela 1).

**Tabela 1-** Média dos valores dos picos de voltagem encontrados nas séries temporais antes e após aplicação dos estímulos (em  $\mu\text{V}$ ).

	Estímulo Térmico		Estímulo Luminoso		Estímulo Osmótico	
	Antes do estímulo	Após estímulo	Antes do estímulo	Após estímulo	Antes do estímulo	Após estímulo
Mínima Voltagem	- 1,72	-16,55	-48,84	-75,04	-51,28	-287,54
Máxima Voltagem	1,17	33,16	43,76	89,31	38,88	235,79
Amplitude	2,89	49,71	92,60	164,35	90,16	523,33
Média dos valores da série	-0,30	0,30	-0,14	-0,14	-0,10	-0,07

O valor da voltagem mínima atingida foi maior antes do estresse e o valor máximo de voltagem atingido aumentou muito após a aplicação do estímulo. Com isso, verificamos um aumento da amplitude do sinal após o estímulo. Nota-se que o valor das médias dos valores de voltagem obtidos das séries antes e após o estímulo é próximo a zero. Nos experimentos envolvendo a aplicação de estímulo

luminoso, o valor da voltagem mínima atingida foi maior antes do estímulo e o valor máximo de voltagem aumentou após a aplicação do estímulo. Contudo, verificamos que o aumento da amplitude do sinal após o estímulo foi menos acentuado que o aumento verificado com estímulo térmico. Nota-se também que o valor das médias dos valores coletados das séries temporais antes e após o estímulo é próximo a zero, semelhante ao encontrado com estresse térmico.

Nos experimentos envolvendo a aplicação de estímulo osmótico, o valor da voltagem mínima atingida foi maior antes do estímulo e o valor máximo de voltagem atingido aumentou consideravelmente após a aplicação do estímulo. Verificou-se ainda que o aumento da amplitude do sinal ocorrido após a aplicação do estímulo foi consideravelmente maior que o observado com a aplicação dos estímulos térmico e luminoso. Nota-se também que a média dos valores coletados das séries temporais antes e após o estímulo é próximo a zero, semelhante ao encontrado com os estímulos anteriores.

## **4.2 Espectro de Frequências (FFT)**

Conforme o trabalho de Costa et al. (2011) havia reportado, os sinais bioelétricos em plantas apresentam um padrão oscilatório. Um padrão oscilatório é característico por apresentar frequências determinadas pela dinâmica do sistema em questão. Para confirmar a natureza oscilatória do sinal elétrico evocado das plantas, as séries temporais de dados bioelétricos foram analisadas utilizando a técnica da Transformada Rápida de Fourier (FFT).

A FFT fornece a distribuição de energia espectral encontrada por unidade de tempo, uma vez que a energia total de tal sinal durante todo o tempo, em geral, seria infinita. As Transformadas rápidas de Fourier são de grande importância em uma vasta gama de aplicações, de Processamento digital de sinais para a resolução de equações diferenciais parciais a algoritmos para multiplicação de grandes inteiros.

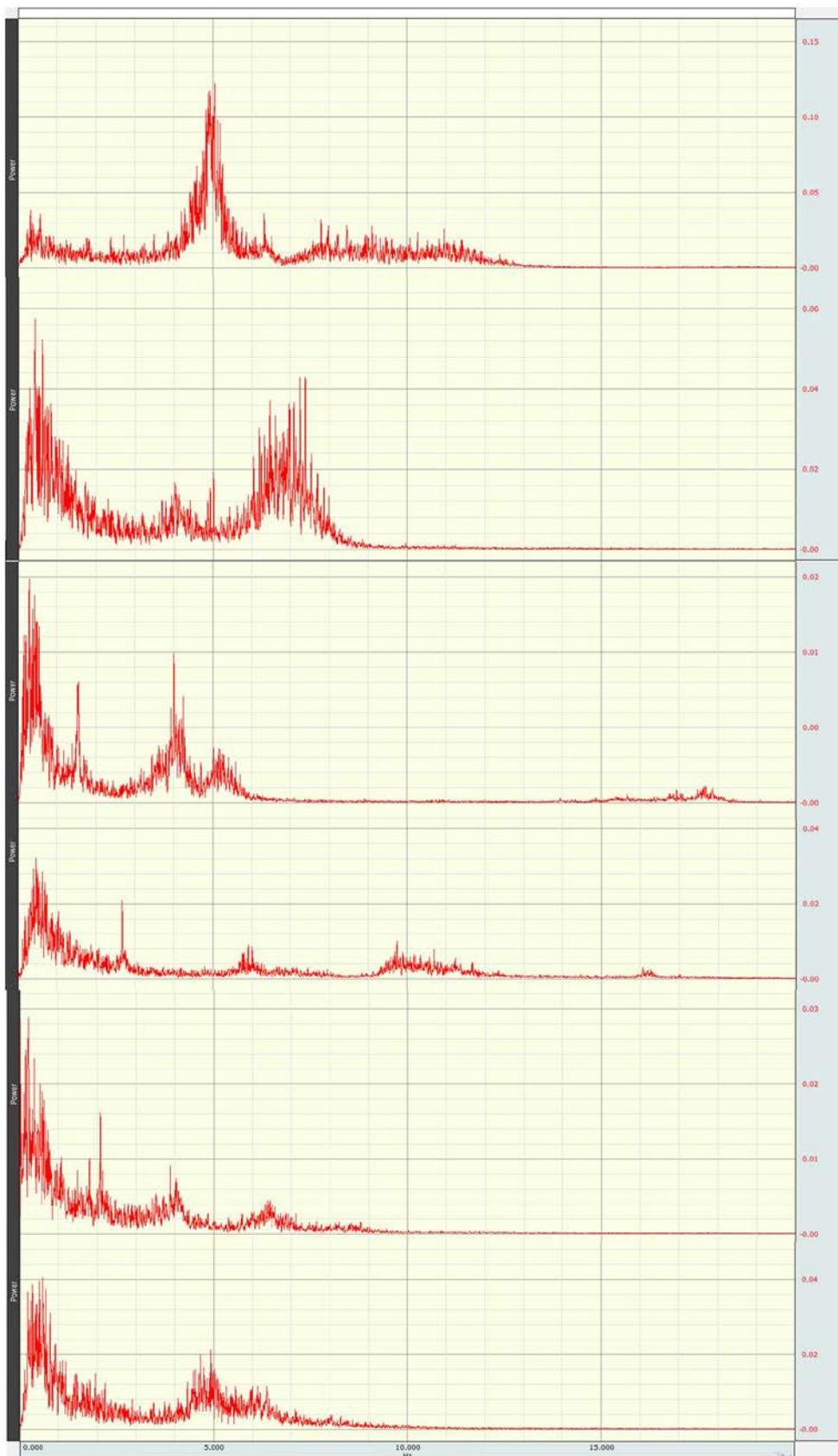
Para comprovar a existência de informação nos sinais coletados, em todos os experimentos os dados das séries temporais de plantas foram comparados com dados de eletrodos “abertos”, ou seja, estavam dentro da Gaiola de Faraday, contudo não estavam inseridos em plantas. Dessa forma pode-se verificar o quanto a planta influenciou na distribuição de frequências nas séries temporais.

Analisando as séries temporais de plantas antes e depois da aplicação dos estímulos (osmótico, luminoso e temperatura) verificamos que ambas as condições apresentaram distribuição de frequências seguindo uma lei de potência, indicando uma dinâmica com padrão livre de escala. Contudo, as séries temporais coletadas antes da aplicação dos estímulos (Figura 10) apresentam conjuntos de frequências variando de 0,15 a 10 Hz, seguindo uma tendência de lei de potência. Já as séries coletadas após a aplicação dos estímulos apresentam conjunto de frequências que variavam de 0,10 a 5 Hz (Figura 11).

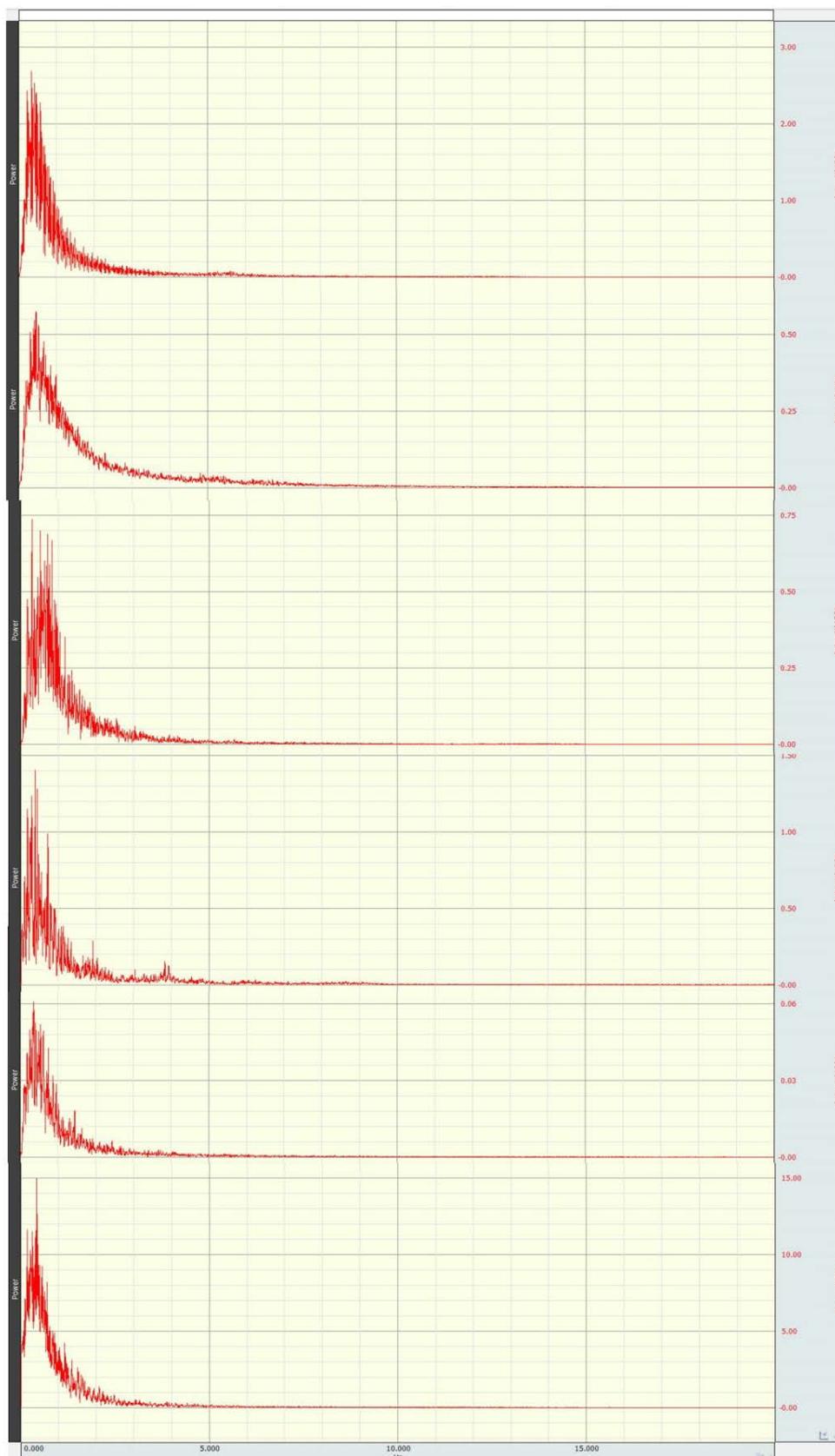
Mais ainda, as séries coletadas antes da aplicação dos estímulos apresentavam um conjunto secundário de frequências variando entre 5 a 10 Hz. Tal conjunto não aparece em nenhuma das séries coletadas após a aplicação dos estímulos. As análises de FFT não permitiram a diferenciação entre o tipo de estímulo aplicado nas plantas, pois os padrões de distribuição de frequências encontrados nas séries temporais coletadas após a aplicação dos estímulos eram iguais. Os resultados dos espectros de frequências encontrados indicam tendência de diminuição da complexidade dos sinais elétricos presentes na comunicação raiz-parte aérea de plantas de soja após a aplicação de um estímulo estressante. A aplicação dos estímulos causou diminuição da complexidade do sinal, com consequente desaparecimento de frequências que estavam presentes antes da aplicação dos estímulos.

Outro ponto importante de se destacar é o padrão de espectro de frequência encontrado nas séries temporais obtidas nos eletrodos abertos (Figura 12). Pode-se verificar grande diferença na distribuição de frequências quando comparadas a séries de plantas, seja antes ou após a aplicação dos estímulos. Em contraste às séries coletadas de plantas, as séries de eletrodos abertos apresentam um sinal com a distribuição homogênea de frequências, sem domínio de nenhum conjunto de frequências específicas, caracterizando assim um sinal típico de ruído branco.

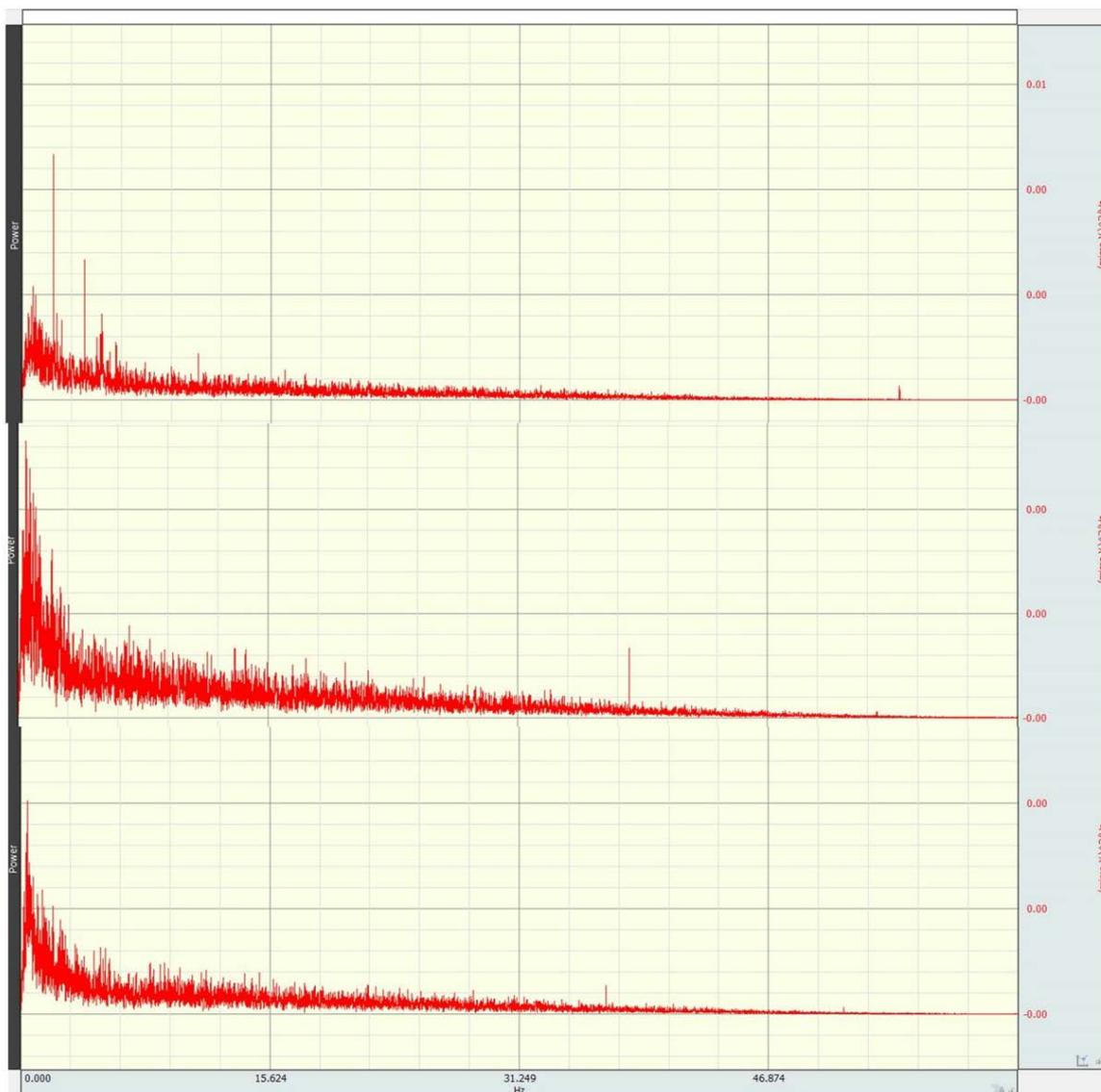
**Figura 10-** Padrões de Espectro de Frequências (FFT) encontrados em séries temporais obtidas de plantas de soja antes da aplicação dos estímulos



**Figura 11-** Padrões de Espectro de Frequências (FFT) encontrados em séries temporais obtidas de plantas de soja após a aplicação dos estímulos.



**Figura 12-** Padrões de Espectro de Frequências (FFT) encontrados em todas as séries temporais coletadas de eletrodos abertos.



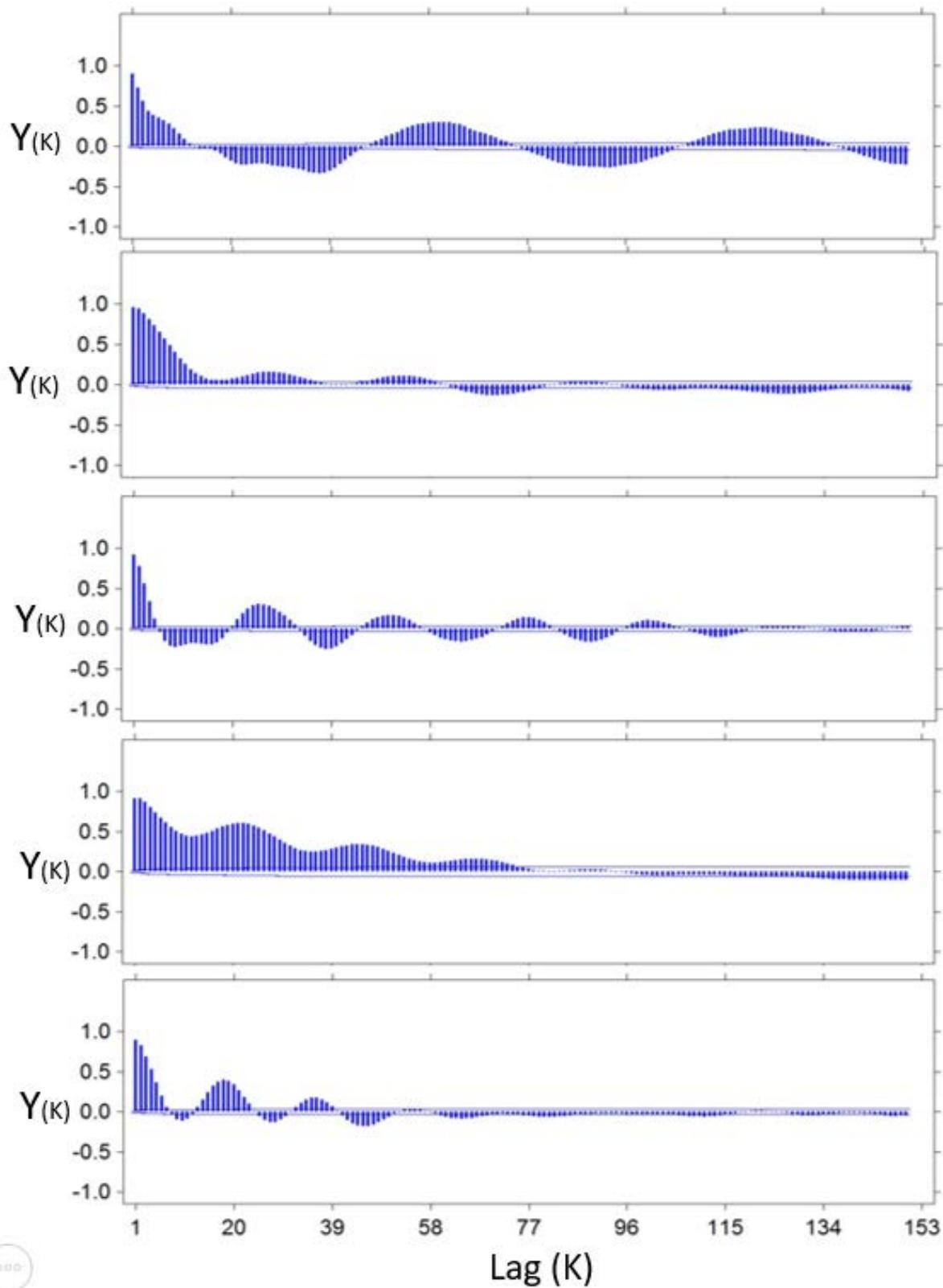
Testes foram feitos com plantas secas, já mortas, cujos dados não foram apresentados nesse trabalho. Tais plantas apresentavam padrão de distribuição de frequências semelhante ao padrão encontrado em eletrodos abertos. Tal fato reforça o papel da planta na informação oscilatória contida nos sinais bioelétricos coletados.

### 4.3 Análises de Autocorrelação

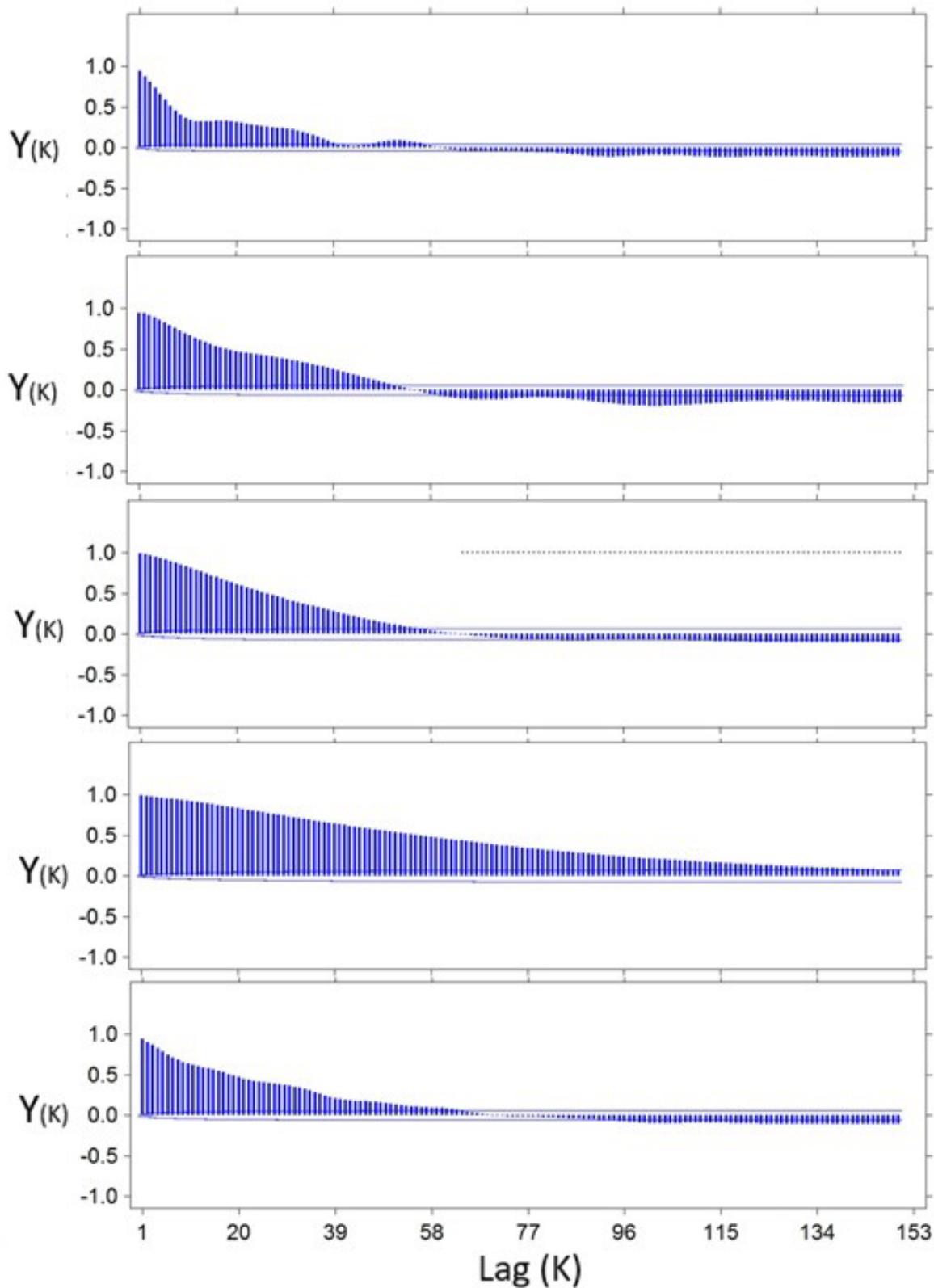
Na adoção de um modelo para uma série temporal, é necessário conhecer-se a relação entre as observações atuais e as anteriores. Uma forma de avaliar tal relação é através das funções de autocorrelação. As séries temporais foram analisadas utilizando autocorrelação com *lag* máximo de 150. Comparando as séries temporais coletadas nas plantas, antes ou após a aplicação dos estímulos, com séries de eletrodos abertos pode-se verificar a existência de informação discriminante nas séries temporais de sinais elétricos coletadas das plantas.

Quando comparados os padrões de autocorrelação encontrados nas séries temporais obtidas nas plantas antes da aplicação dos estímulos (Figura 13), com séries obtidas após a aplicação dos estímulos (Figura 14) e as séries obtidas com eletrodos abertos (Figura 15), verifica-se um padrão bem diferente entre as séries. Observa-se que nas séries com eletrodos abertos os dados da série podem ser considerados descorrelacionados. Tal fato indica que o sinal evocado das plantas possui informação discriminante, não sendo assim um sinal aleatório e nem ruído. As análises de autocorrelação deixam bem claro que há diferença visual significativa no sinal coletado antes da aplicação dos estímulos com os sinais coletados após a aplicação dos estímulos. Contudo, as análises não permitem uma diferenciação do tipo de estímulo aplicado, pois os padrões de autocorrelação encontrados se repetem em todas as séries coletadas após a aplicação dos estímulos. As séries temporais coletadas de plantas antes da aplicação dos estímulos apresentam autocorrelação com um padrão oscilatório bem marcado, havendo diminuição da amplitude da autocorrelação com o aumento da defasagem (Figura 13). Já nas séries temporais obtidas após a aplicação dos estímulos, a autocorrelação apresenta um padrão linear seguindo uma tendência de queda com o aumento da defasagem (Figura 12), atingindo valores negativos a grandes defasagens.

**Figura 13-** Padrões de Autocorrelação encontrados nas séries temporais coletadas antes da aplicação dos estímulos.

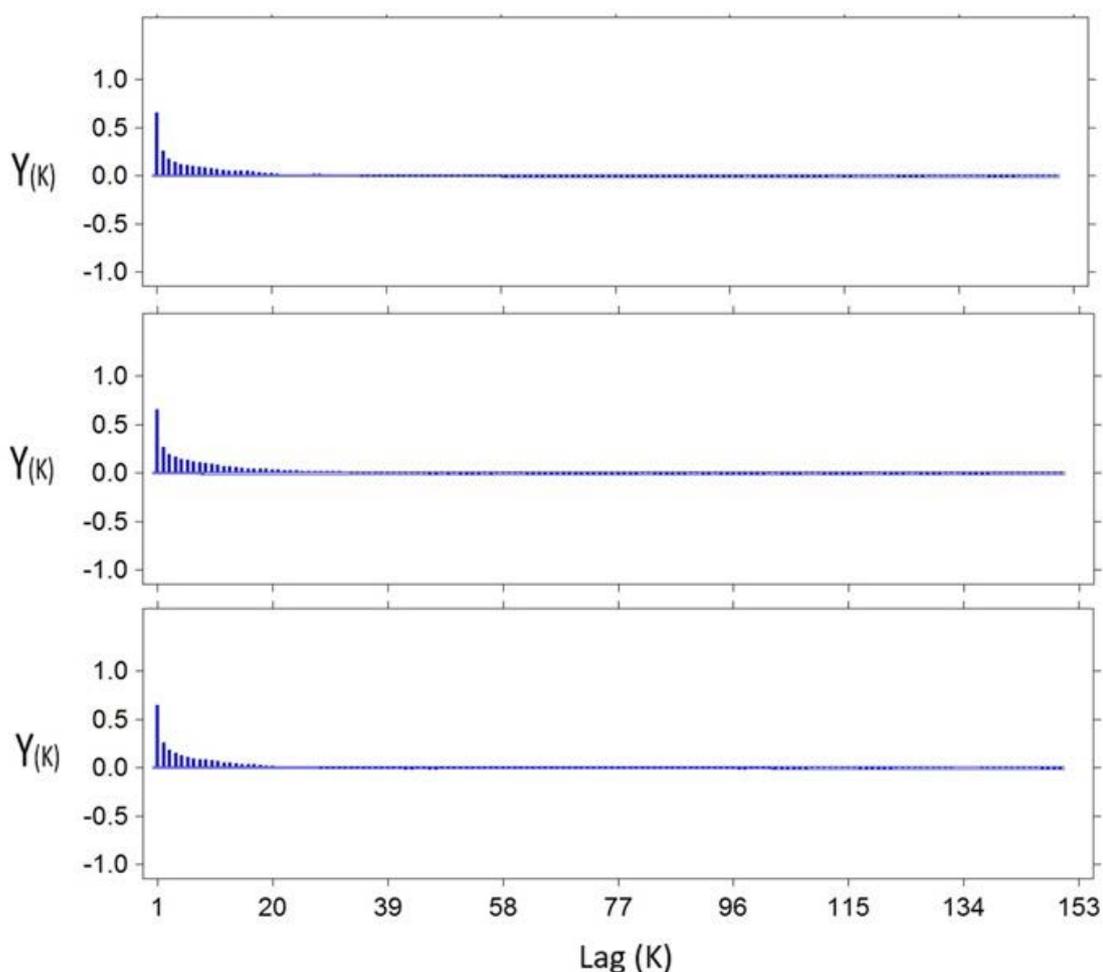


**Figura 14-** Padrões de Autocorrelação encontrados nas séries temporais coletadas após da aplicação dos estímulos.



As séries temporais coletadas dos eletrodos abertos que foram deixados dentro da Gaiola de Faraday (Figura 13) apresentam padrão de autocorrelação bem diferente do encontrado em séries de evocadas de plantas, sendo estressadas ou não. Tal comparação nos permite inferir que as séries temporais coletadas de plantas possuem informação discriminante do sinal, visto que o sinal coletado de eletrodos abertos é um sinal típico de ruído branco, caracterizado como uma sequência de erros (ou choques) aleatórios, sempre que esta tiver média e variância constante e sem autocorrelação.

**Figura 15-** Padrão de Autocorrelação encontrado nas séries temporais coletadas pelos eletrodos abertos.

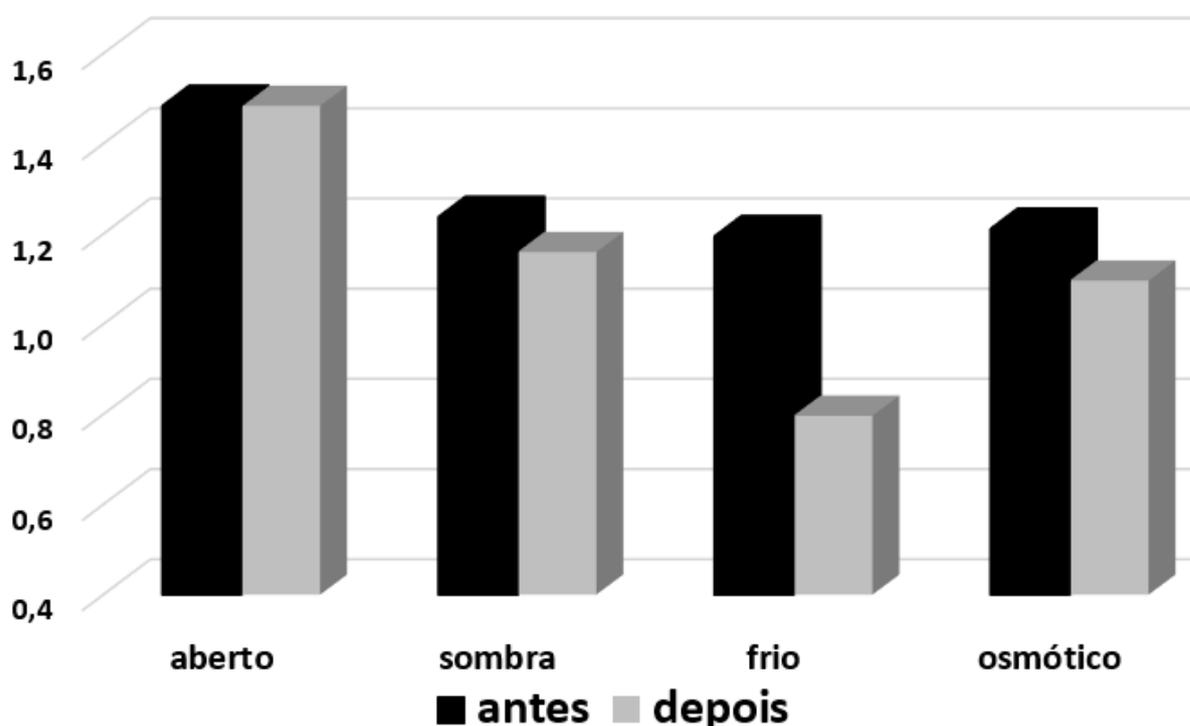


#### 4.4 Entropia Aproximada

Sendo a planta um sistema complexo e dinâmico, faz todo sentido que seus processos e mecanismos também o sejam. Para quantificar a complexidade da sinalização elétrica, através da análise de séries temporais, foi utilizado o método da Entropia Aproximada (ApEn) proposto por Pincus em 1991. Tal método fornece informações a respeito do nível de organização ou desorganização (complexidade) das séries temporais. Os valores de ApEn podem trazer informações importantes para a caracterização e distinção entre séries de plantas antes e após a aplicação dos estímulos. A análise de Entropia Aproximada já foi utilizada previamente em análises de dados em fisiologia de plantas (SOUZA et al. 2004a) e em análises de Eletroencefalograma em animais (ABÁSULO et al., 2005).

Após análise dos valores de ApEn obtidos das séries temporais, pode-se verificar que a aplicação dos estímulos afetou a complexidade dos sinais bioelétricos. A aplicação causou redução nos valores médios de ApEn das séries temporais, para os três estímulos aplicados (Figura 16).

**Figura 16-** Valores de Entropia Aproximada (ApEn) das séries temporais coletadas de plantas de soja e de eletrodos abertos.



Contudo, observa-se também que a aplicação do estímulo térmico (frio) foi a que causou a maior redução nos valores de ApEn. Um ponto importante de se ressaltar é que os valores de ApEn das séries coletadas de eletrodos abertos foram maiores que das séries coletadas de plantas, antes ou após estímulo. Observa-se também que a aplicação dos estímulos nas plantas não causou alteração dos valores de entropia dos eletrodos abertos, que coletavam dados simultaneamente. Isso mostra que os valores encontrados refletem o estado das plantas.

Os valores de ApEn refletem a probabilidade de que padrões similares observados em uma série temporal não serão seguidos por observações posteriores semelhantes (HO, 1997). Sendo assim, a presença de padrões oscilatórios em séries temporais torna-se mais previsível à medida que tais padrões já ocorreram em algum momento em tal série. Uma série temporal contendo diversos padrões repetitivos tem um valor de ApEn pequeno, enquanto que uma série menos previsível, ou mais aleatória, apresenta altos valores de ApEn. Entre todas as séries coletadas, as pertencentes aos eletrodos abertos apresentaram os maiores valores de ApEn, devido sua composição ser basicamente de sinais aleatórios. Já as séries temporais coletadas de plantas possuem padrões que se repetem ao longo do tempo, e, portanto, apresentam menores valores de ApEn.

#### **4.5 Técnicas de Classificação**

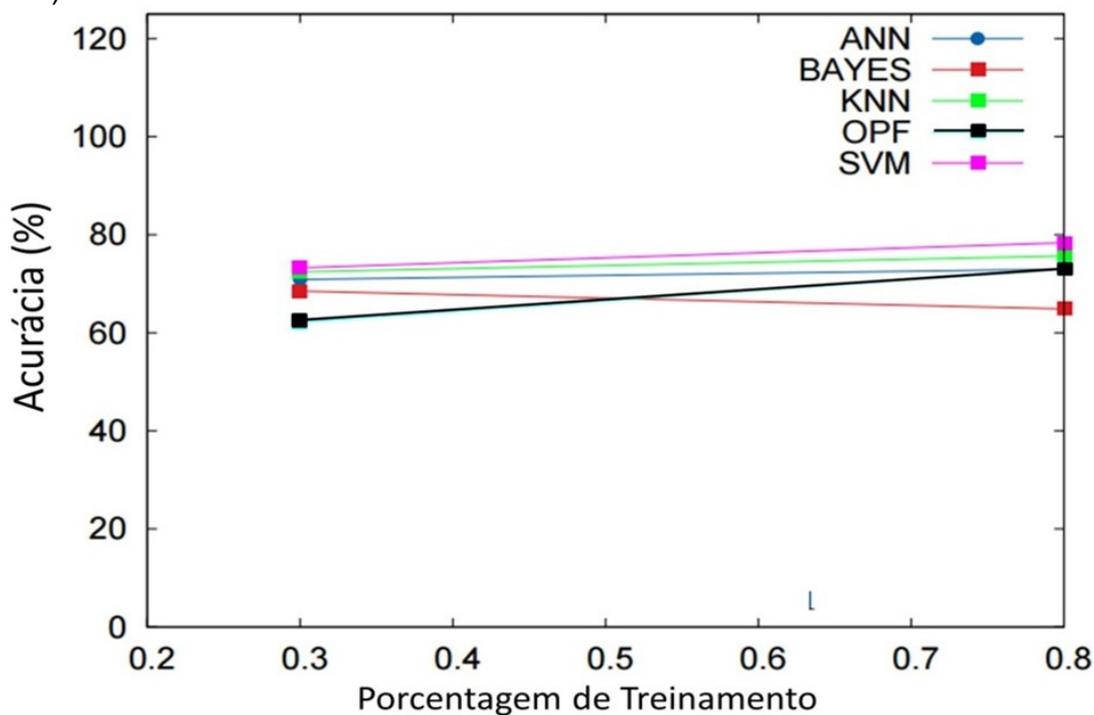
A utilização de técnicas de classificação tem como objetivo calcular a probabilidade de que uma amostra desconhecida (série temporal) pertença a cada uma das classes possíveis, ou seja, prever a classe mais provável. Outra característica de algoritmos de classificação é que eles requerem um conjunto de dados prévio que já esteja classificado, ou seja, um conjunto que já esteja separado em classes. Baseado neste conjunto de dados prévio, que também é chamado de conjunto de treinamento, o algoritmo recebe como entrada uma nova amostra desconhecida, ou seja, que não possui classificação, e retorna como saída a classe mais provável para esta amostra, de acordo com cálculos probabilísticos. Por essa razão, o conjunto de dados original foi dividido em conjuntos de treinamento de tamanhos progressivamente maiores, variando de 30% a 80% do conjunto de dados original. Foram encontrados desempenhos bons e comparáveis entre todos os algoritmos utilizados nesse trabalho, mesmo em conjuntos pequenos de

treinamento, como o de 30%. As Figuras 15-19 mostram a porcentagem de acurácia de detecção de padrões entre as séries temporais para diferentes situações.

Quando comparamos todas as séries temporais relativas a períodos antes e após a aplicação dos estímulos (Figura 17), verificamos que a acurácia na detecção de padrões variou entre 60 a 80%, mesmo utilizando o conjunto de treinamento de 30%. Nota-se que os algoritmos SVM e KNN apresentaram melhores desempenhos nas classificações das séries em antes e após a aplicação dos estímulos. Ao realizarmos as classificações somente entre as séries temporais coletadas após a aplicação dos estímulos (Figura 16), verificamos que os algoritmos apresentaram valores de acurácia de detecção de padrões entre 50 e 80%. Os melhores valores de acurácia foram alcançados com os algoritmos KNN e BAYES, tanto para o conjunto de treinamento de 30% quanto para o de 80% (contudo nos conjuntos maiores os valores alcançaram quase 80% de acurácia).

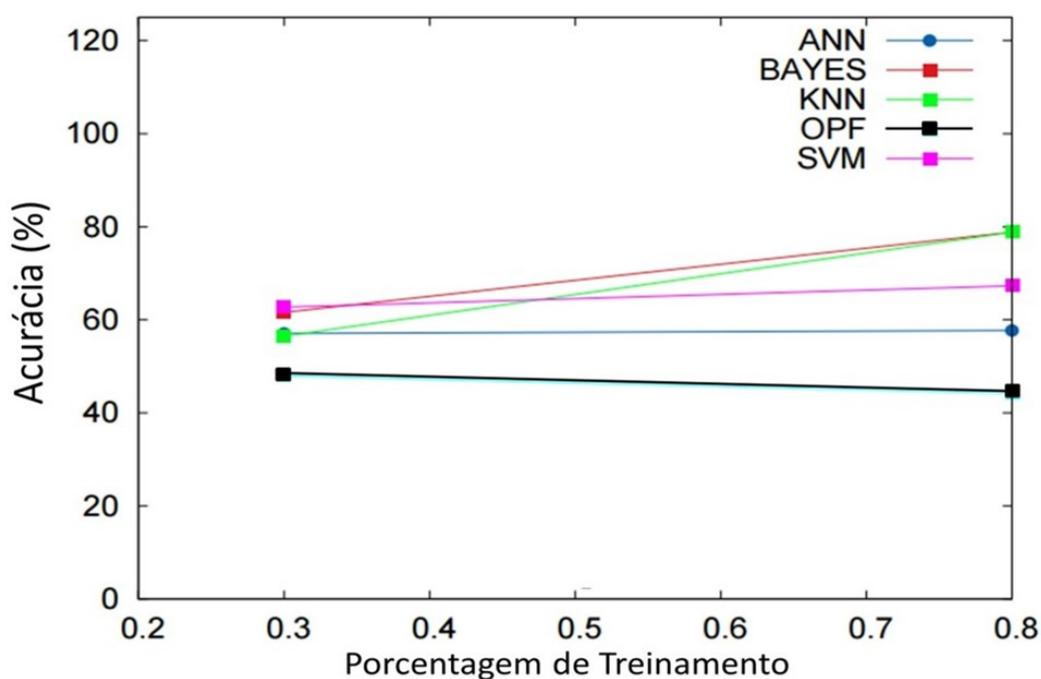
A Figura 19 apresenta os resultados da classificação entre as séries temporais que receberam estresse osmótico, visando a classificação entre antes e após a aplicação do mesmo. A acurácia dos algoritmos na detecção de padrões ficou entre 45 e 85%. Os melhores valores de acurácia na detecção de padrões foram alcançados pelos algoritmos ANN e SVM, que apresentaram acurácia superior a 80% utilizando ambos os conjuntos de treinamento. Já o algoritmo OPF apresentou valores de acurácia entre 45 e 60%, tendo assim o pior desempenho de classificação para as séries temporais em questão. A Figura 18 apresenta os resultados da comparação entre as séries temporais de plantas que receberam estímulo térmico, visando à classificação entre antes e depois da aplicação do mesmo. Os valores de acurácia dos algoritmos na detecção de padrões ficaram entre 70 e 90%, seja quando utilizado o conjunto de treinamento de 30% quanto quando usado o conjunto de 80%. Com exceção do algoritmo OPF os demais apresentaram valores de acurácias na detecção de padrões com valores entre 80 a 90%, indicando alto índice de detecção de padrões. A Figura 21 apresenta os resultados da comparação somente entre as séries que receberam estresse luminoso, visando a classificação entre antes e depois da aplicação do mesmo. Observamos que a acurácia alcançada pelos algoritmos ficou entre 60 a 80%. Podemos aqui verificar que o algoritmo SVM apresentou as maiores acurácias na detecção de padrões, com valores chegando a quase 80% de acurácia.

**Figura 17-** Acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas antes e após a aplicação dos estímulos (sem distinção de classes).



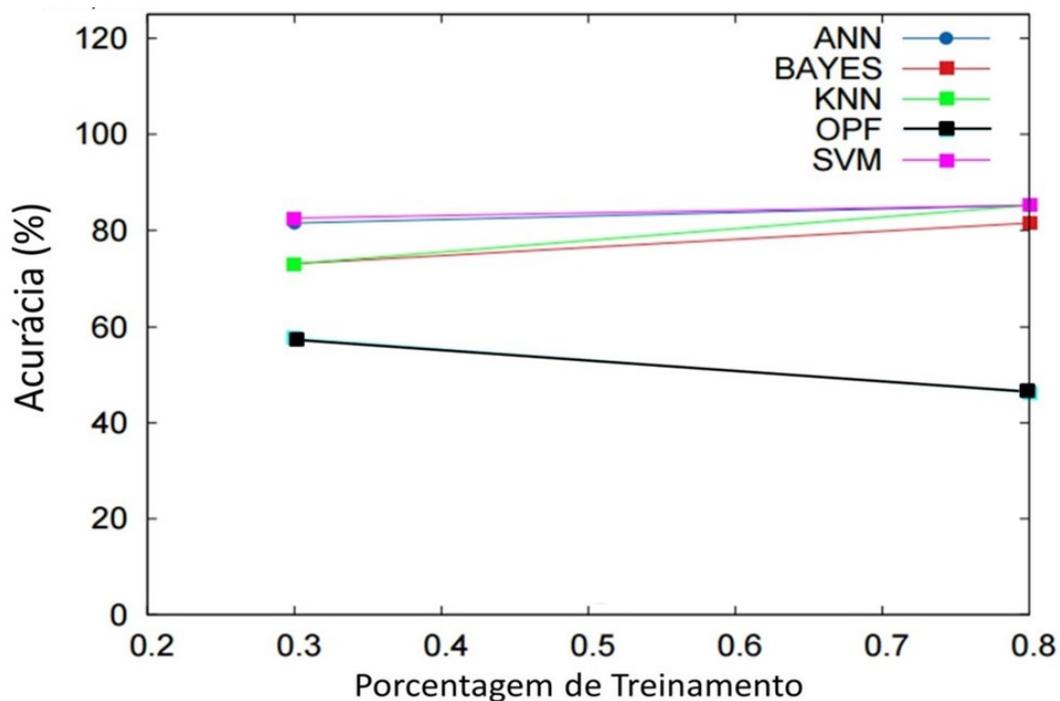
\*Algoritmos de classificação: **ANN**: Rede Neural Artificial; **BAYES**: Classificador Baysiano; **KNN**: K-vizinhos mais próximos; **OPF**: Floresta de caminho ótimo; **SVM**: Máquinas de vetores de suporte.

**Figura 18-** Acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas após a aplicação dos estímulos (sem distinção de classes).



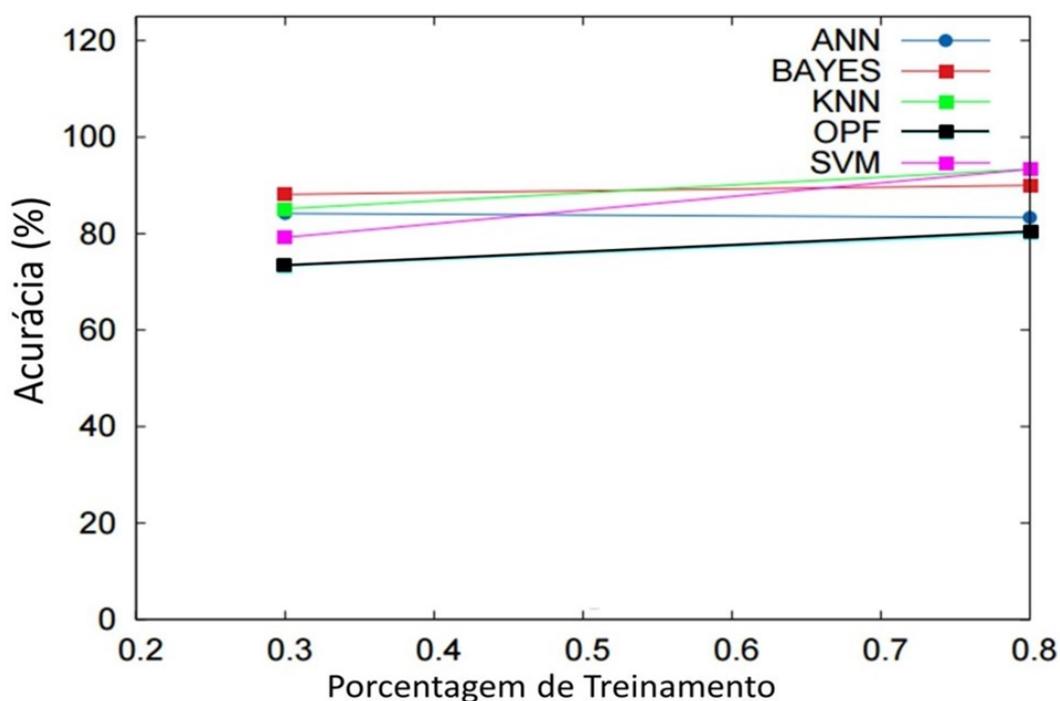
\*Algoritmos de classificação: **ANN**: Rede Neural Artificial; **BAYES**: Classificador Baysiano; **KNN**: K-vizinhos mais próximos; **OPF**: Floresta de caminho ótimo; **SVM**: Máquinas de vetores de suporte.

**Figura 19-** Acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas antes e após a aplicação de estresse osmótico.



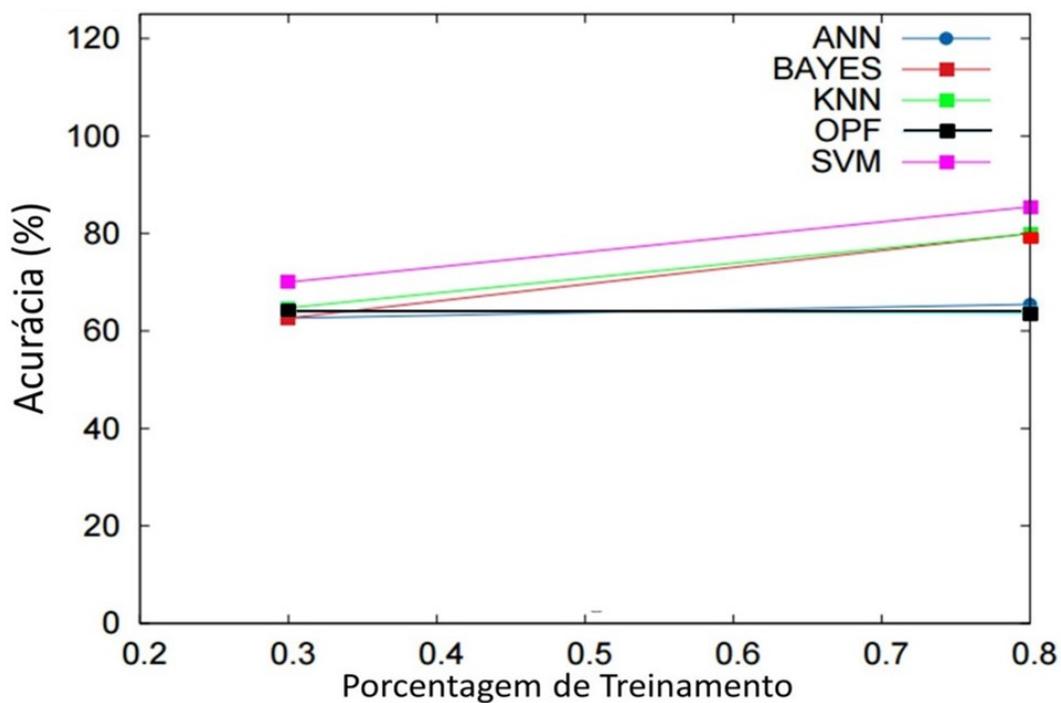
\*Algoritmos de classificação: **ANN**: Rede Neural Artificial; **BAYES**: Classificador Baysiano; **KNN**: K-vizinhos mais próximos; **OPF**: Floresta de caminho ótimo; **SVM**: Máquinas de vetores de suporte.

**Figura 20-** Porcentagem de acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas antes e após a aplicação de estresse térmico.



\*Algoritmos de classificação: **ANN**: Rede Neural Artificial; **BAYES**: Classificador Baysiano; **KNN**: K-vizinhos mais próximos; **OPF**: Floresta de caminho ótimo; **SVM**: Máquinas de vetores de suporte.

**Figura 21-** Acurácia na detecção de padrões nas séries temporais. Classificação entre séries coletadas antes e após a aplicação de estresse luminoso.



\*Algoritmos de classificação: **ANN**: Rede Neural Artificial; **BAYES**: Classificador Baysiano; **KNN**: K-vizinhos mais próximos; **OPF**: Floresta de caminho ótimo; **SVM**: Máquinas de vetores de suporte.

## 5 DISCUSSÃO

O estilo de vida sésstil das plantas faz com que se tornem organismos modulares e com o mínimo de especialização de células e tecidos, se comparadas a outros seres vivos (MANCUSO, 2010). Esse fato faz com que haja a necessidade de um sistema sensorial, de sinalização e de resposta muito bem organizado para proporcionar uma exploração eficiente do ambiente e reagir rapidamente a circunstâncias potencialmente danosas. Isso nos trás uma nova visão de organismos sensitivos, dinâmicos e de comportamentos altamente complexos, que disputam recursos abaixo e acima do solo (BRENNER et al., 2006; BALUSKA; MANCUSO, 2007; TREWAVAS, 2009). A complexidade de um sistema é a quantidade de informação necessária para descrever um sistema. A complexidade depende do nível de detalhes necessários para a descrição do sistema. Em termos básicos, quanto mais detalhes e informações existem sobre um sistema, mais complexo ele é. As hipóteses desta tese partem do princípio de que as plantas são seres vivos complexos e dinâmicos. Sendo assim tais características são esperadas em seus processos subjacentes. Os dados coletados das plantas de soja nos experimentos desta tese corroboram a hipótese inicial de que as séries temporais de dados obtidos da sinalização elétrica em plantas de soja apresentam informação “não aleatória”, ou seja, possuem padrões informacionais embutidos na variação temporal dos potenciais de voltagem. As evidências de informações contidas nas séries temporais nos levam a uma nova era nas análises fisiológicas de plantas. Assim será possível deixar para trás abordagens puramente lineares e deterministas, indo em direção a uma fisiologia vegetal sistêmica e dinâmica. Em eletrofisiologia vegetal os trabalhos têm grande foco nas respostas pontuais das plantas, analisando somente amplitude e picos de voltagem e em pequena escala de tempo. Essa tese traz a perspectiva do estudo das séries temporais, que são coleções de observações feitas em sequência ao longo de um determinado período de tempo. Uma característica muito importante de séries temporais é que as observações vizinhas tendem a ser dependentes. Em muitos trabalhos, o interesse é analisar e modelar tal dependência em séries coletadas de organismos complexos. Outra propriedade de uma série temporal é o fato de que os dados não são gerados independentemente, sua dispersão varia no tempo e são frequentemente regulados por uma tendência, podendo ter componentes cíclicos, também chamados de

componentes sazonais. É dentro desse contexto que as análises de autocorrelação e de FFT das séries temporais podem trazer informações importantes a respeito da sinalização elétrica das plantas.

A autocorrelação é a correlação de um sinal ou valor com ele mesmo em diferentes instantes no tempo. Informalmente, podemos dizer que é a similaridade entre observações como uma função de um “time lag” entre eles. Trata-se de uma ferramenta matemática que também pode ser usada para encontrar padrões repetitivos, como a presença de um sinal periódico escondido por um ruído. As análises de autocorrelação mostraram que, ao contrário do ruído do dispositivo de EEG, todas as séries temporais coletadas das plantas apresentavam correlação de longo alcance. As séries temporais coletadas pelos eletrodos abertos apresentaram padrão de decaimento bem rápido para valores próximos de zero, indicando a ausência de correlação temporal, típico de uma sequência de sinais aleatórios, conhecida como ruído branco (Figura 11). No entanto, observa-se que o padrão de decaimento do sinal medido antes e após a aplicação dos estímulos exibiu diferentes comportamentos. Antes da aplicação dos estímulos, a autocorrelação apresenta um comportamento oscilatório com amplitude decaindo com o tempo (ou Lag  $K$ ). Em contrapartida, as séries coletadas após a aplicação dos estímulos apresentam tendência de decaimento lento para zero a levemente negativo (Figura 13). A literatura indica que, se o valor de autocorrelação estiver próximo de +1 ou -1, encontra-se uma correspondência linear entre os valores de uma e os correspondentes de outra variável. Com valores próximos de zero, onde para a correlação  $k=0$  significa que o presente e o passado ( $k$ ) estão linearmente descorrelacionados, não é perceptível uma relação entre os valores assumidos por uma variável e por outra (BOX, 1970). A aplicação de metodologias de análise de séries temporais, incorporando a correlação existente entre as observações, possibilita o desenvolvimento de modelos de predição e monitoramento mais adequados. Nos últimos anos, considerável atenção tem sido dedicada a métodos alternativos para o estudo de séries temporais com padrões não-lineares e até mesmo caóticos.

As análises de FFT confirmaram a existência de diferenças no espectro de frequências das séries temporais de plantas antes dos estímulos, após a aplicação dos estímulos considerando eletrodos abertos. Os eletrodos abertos apresentaram distribuição de frequências constantes, típica de um ruído branco. Em

contrapartida, a FFT das séries temporais coletadas nas plantas antes e após a aplicação dos estímulos apresentou distribuição de frequências seguindo uma lei de potência, [indicando um padrão de dinâmica livre de escala]. No entanto a análise FFT revelou diferenças significativas entre as séries coletadas antes e após aplicação dos estímulos. Após a aplicação dos estímulos, nota-se uma tendência de diminuição da complexidade dos sinais, pois houve o desaparecimento de picos de frequências que estavam presentes antes dos estímulos. Uma grande variedade de fenômenos físicos e biológicos apresenta distribuição que segue uma lei de potência, em uma grande variação de magnitudes: Indo desde tamanhos de crateras na lua e erupções solares (NEWMAN, 2005), padrão de forrageamento de espécies (HUMPHRIES et al. 2010), padrões de atividades em populações de neurônios (Klaus, 2011), frequências de palavras em idiomas, frequências de nomes de famílias e riqueza de espécies em clados de organismos (ALBERT, 2011). E também é conhecido que a existência de leis de potência em sistemas biológicos pode ser um indicador de complexidade, particularmente no contexto de Criticalidade auto-organizada (SOC). Tal fenômeno pode ser verificado também nos dados em anexo referentes a experimento envolvendo a aplicação de estímulo osmótico em plantas de soja. Os espectros de frequência encontrados nas séries temporais evocadas das plantas apresentam características semelhantes a encontrados em Eletroencefalogramas, onde existe uma correlação temporal de longo alcance e um comportamento de escala em lei de potência das oscilações, variando entre 10 a 20 Hz (LINKENKAER-HANSEN, 2001). As oscilações neuronais *in vivo* exibem flutuações e amplitudes não aleatórias, refletidas em um decaimento lento das autocorrelações temporais que persistem por dezenas de segundos. Curiosamente, o decaimento da autocorrelação é alterado em várias desordens relacionadas com o cérebro, como a epilepsia, a depressão e outras desordens cerebrais (LINKENKAER-HANSEN, 2001). Outro aspecto interessante encontrado nos dados desse trabalho foi o espectro de frequências que varia entre 8-13 Hz, que embora a natureza dessas flutuações seja ainda desconhecida, atraem grande interesse no contexto de eletroencefalografia humana (PALUS, 1996; CERF et al., 1997; STAM et al., 1999).

As oscilações em plantas acontecem em todos os níveis, com grandes variedades de períodos e frequências. As séries temporais coletadas nos experimentos envolvendo a aplicação de estímulos externos corroboram a ideia de

que oscilações na sinalização elétrica fazem parte de um mecanismo de codificação. Em tal mecanismo, as frequências poderiam codificar informações ambientais, como já acontece em relação a picos de  $\text{Ca}^{2+}$  em células-guarda (MCAINSH et al. 1995, MCAINSH; HETHERINGTON, 1998). No entanto, entende-se que mecanismos de codificação baseados em frequências podem ser verdadeiros para vários outros processos relacionados às plantas e que os estímulos ambientais podem alterar a dinâmica desses processos (ERDEI et al., 1998). Os dados coletados reforçam a hipótese inicial de que os estímulos ambientais alteram a dinâmica das séries temporais. Adicionalmente, analisando visualmente e sistematicamente a forma original de cada série temporal coletada antes e após a aplicação dos estímulos, observa-se a presença de picos de voltagem chegando a  $500 \mu\text{V}$  em todas as séries coletadas após a aplicação dos estímulos. Esses picos não foram encontrados nas séries coletadas antes da aplicação dos estímulos, demonstrando que o estímulo causou tais picos como forma de sinalização.

A entropia aproximada (ApEn) é um parâmetro não linear amplamente utilizado para quantificar a imprevisibilidade de séries temporais, tais como as obtidas de respostas fisiológicas, sendo utilizado também para determinar a complexidade da dinâmica de sistemas (PINCUS, 1991, 1994). Atualmente os trabalhos na área médica também fazem uso das análises de ApEn. Isso porque elas também podem ser aplicadas para classificar sinais de EEG em doenças psiquiátricas, como a esquizofrenia (MALIHE, 2009), epilepsia (YUAN, 2011) e alguns tipos de vícios (YUN, 2012). Neste trabalho, os valores de ApEn calculados das séries temporais revelaram importante informação a respeito da dinâmica dos sinais elétricos coletados antes e após a aplicação dos estímulos. As medidas de complexidade mostraram consistentemente que, após a aplicação dos estímulos, houve decréscimo dos valores de ApEn, indicando que o sinal elétrico possuía alta complexidade antes da aplicação dos estímulos, havendo então uma diminuição da complexidade do sinal após a aplicação dos estímulos. ApEn atribui valores positivos a uma sequência ou série temporal, onde os maiores valores correspondem a maior aleatoriedade aparente dos processos ou irregularidade em séries temporais, como verificado nos valores encontrados nos eletrodos abertos, e valores pequenos estão associados a uma maior ocorrência de características reconhecíveis ou padrões nos dados (KAPLAN, 1991; PINCUS, 1991). Os dados analisados revelam então aumento da regularidade do sinal elétrico após a

aplicação dos estímulos. O aumento da regularidade após distúrbios tem ocorrência em diferentes fenômenos da natureza. Sistemas doentes, quando associados à emergência de comportamento mais regular, apresentam valores reduzidos de entropia quando comparados a sistemas saudáveis (INCHAUSTI, 2002). Uma grande classe de patologias e distúrbios dos sistemas metabólicos são associados à perda de complexidade dos sinais, sendo assim proposto que tal característica seja comum às dinâmicas patológicas (INCHAUSTI, 2002; NUNES AMARAL, 1998).

Uma das hipóteses do trabalho é que existem padrões matemáticos que descrevem o comportamento dos sinais bioelétricos em cada condição ambiental e entre o estado estimulado e não estimulado. Os algoritmos de classificação são uma ferramenta muito usada para a identificação de padrões, pois trabalham com um conjunto de treinamento, baseado em sistemas de aprendizado de máquina, onde reconhece a partir do treinamento a que classe pertence determinada série dada. Em linhas gerais os algoritmos apresentaram bons resultados na acurácia de classificação das séries, levando em conta que até conjuntos de treinamento pequenos (30%) obtiveram acurácias variando entre 60 a 80%. Tendo como referência sistemas biológicos, que são muito dinâmicos e complexos, tal acurácia traz a clara noção de que há padrões matemáticos embutidos nas séries temporais. Tais padrões podem ser utilizados para “classificar” as séries temporais em momentos antes e após a aplicação dos estímulos, visto que classificação é um exemplo de reconhecimento de padrões. Na terminologia utilizada em aprendizado de máquina, classificação é considerada uma instância de aprendizado supervisionado, como aconteceu com os algoritmos utilizados que utilizam um conjunto de treinamento corretamente classificado previamente (ALPAYDIN, 2010).

As técnicas de classificação utilizadas nesse trabalho utilizam conceitos da aritmética intervalar para reduzir o tamanho das séries e facilitar a busca por padrões. A aritmética intervalar lida com os dados das séries temporais na forma de intervalos compactos com o objetivo de codificar erros computacionais, ou ruídos no próprio intervalo (MOORE, 1979; MARKOV, 1995; HICKEY, 2001). Devido a esse fator a abordagem intervalar de sinais e transformadas proporciona maior controle de erros nas representações, e também resultados mais confiáveis na execução e prática de cálculos computacionais complexos como os utilizados nas análises de séries temporais (OPPENHEIM, 1989). A aritmética intervalar é uma

ferramenta muito eficiente em fornecer representações de limites de erros e incertezas. Essa propriedade tornou o uso das técnicas de classificação muito apropriado para trabalhar com as variações presentes nos sinais obtidos das plantas. Nas análises realizadas nesse trabalho, foram utilizados diferentes operadores e relações de Inteligência Artificial, mas apenas a representação mais básica e simples apresentou bons resultados.

## 6 CONCLUSÕES E PERSPECTIVAS

A análise da sinalização elétrica em plantas através do estudo de séries temporais apresenta grande potencial de aplicação na detecção e prevenção de situações estressantes para as culturas. A comparação dos resultados obtidos nas análises de séries temporais da sinalização elétrica, coletadas de plantas de soja, em situações controle e após a aplicação de estímulos estressantes, suportam as hipóteses propostas por essa tese. Ficou evidente que as séries temporais apresentam informação não-aleatória como propriedade, ou seja, possuem padrões embutidos na variação temporal dos potenciais de voltagem. Os resultados também suportam a hipótese de que estímulos ambientais específicos afetam a dinâmica da sinalização elétrica em plantas de soja. A utilização dos algoritmos de classificação revelou que possivelmente há padrões matemáticos nas respostas elétricas evocadas por estímulos ambientais específicos.

Essa tese traz uma nova perspectiva ao olharmos para os processos que ocorrem nas plantas. Ela traz a ideia de que as respostas, assim como os mecanismos, possuem dinâmica oscilatória no tempo, sendo todo o tempo influenciados pelo ambiente. Contudo, verifica-se que a grande maioria dos trabalhos mede as respostas vegetais de forma pontual, não levando em conta as oscilações endógenas do organismo. Espera-se que, a partir dos dados apresentados, haja a discussão no meio acadêmico a respeito da introdução de estudos de séries temporais nos experimentos envolvendo plantas. Tais estudos permitirão entender melhor como os mecanismos e processos evoluem no tempo, fato importante se considerarmos as plantas como sistemas complexos e dinâmicos. Portanto, essa tese pode servir como um primeiro passo rumo a uma nova visão sobre as plantas e um novo jeito de se conceber e conduzir a experimentação vegetal. No entanto, faz-se necessário que os experimentos desenvolvidos nessa tese sejam replicados em condições de campo e outros ambientes controlados. Fora do ambiente controlado, as respostas e padrões encontrados podem ser diferentes, visto que o ambiente tem total influência sobre os organismos vegetais e também sobre os instrumentos de medida, os quais precisam eliminar os ruídos ambientais.

Outro ponto importante revelado por esse trabalho é a ocorrência de informações complexas subjacentes aos sinais elétricos evocados em plantas. Isso traz a possibilidade de aumento do entendimento dos significados biológicos dos

sinais elétricos em plantas. Com isso torna-se possível decodificar e atribuir significados fisiológicos para essa informação. Assim, torna-se possível utilizar tal informação para criar protocolos de diagnóstico de condições estressantes e estados fisiológicos em tempo real. Muitos estudos ainda devem ser feitos, mas essa tese tem um papel importante como um primeiro passo na decodificação da “linguagem” utilizada na sinalização elétrica das plantas. Os avanços e estudos em eletrofisiologia vegetal nos dão a perspectiva da construção de “biosensores” e dispositivos de detecção em tempo real, visando a prevenção de situações potencialmente danosas à produtividade das culturas em detrimento do uso inadequado de recursos na agricultura.

## **ARTIGOS CIÊNTIFICOS**

A partir dos dados coletados nessa tese foram feitas análises complementares e trabalhos foram produzidos em parceria com outros pesquisadores, que incorporaram conhecimentos e técnicas aos resultados obtidos nos experimentos desta tese.

**ARTIGO I:** “Osmotic stress decreases complexity underlying the electrophysiological dynamic in soybean”, publicado no periódico **Plant Biology**.

O **ARTIGO II:** “Plant “electrome” can be pushed toward a self-organized critical state by external cues: evidences from a study with soybean seedlings subject to different environmental conditions”, publicado no periódico **Plant Signalling and Behavior**.

O **ARTIGO III:** “Automatic classification of plant electrophysiological responses to environmental stimuli using machine learning and interval arithmetic”, publicado no periódico **Computers and Electronics in Agriculture**.

## REFERÊNCIAS

- ABÁSULO, D. et al. Analysis of regularity in the EEG background activity of Alzheimer's disease patients with Approximate Entropy. **Clinic. Neurophysiol.**, v.116, p. 1826-1834, 2005.
- ACHE, P. et al. Vfk1, a *Vicia faba* K<sup>+</sup> channel involved in phloem unloading. **Plant Journal**, v.27, p. 571-580, 2001.
- ALBERT, J. S.; REIS, R. E. (Eds.). **Historical Biogeography of Neotropical Freshwater Fishes**. Berkeley: University of California Press, 2011.
- ALLÈNE, C. et al. Some links between extremum spanning forests, watersheds and min-cuts. **Image Vision Computing**, v. 28, n.10, p. 1460–1471, 2010.
- ALPAYDIN, E. **Introduction to Machine Learning**. MIT Press, 2010. p.9.
- BALUSKA, F.; MANCUSO, S. Plant neurobiology as paradigm shift not only in plant sciences. **Plant Signaling and Behavior**, v.2, p. 205-207, 2007.
- BAUER C.S. et. al. Differential expression and regulation of K<sup>+</sup> channels in the maize coleoptile: molecular and biophysical analysis of cells isolated from cortex and vasculature. **Plant Journal**, v.24, p.139-145, 2000.
- BOSE J. C. **Researches on irritability of plants**. London: Longmans, Green and Co, 1913.
- BOSE J. C. **Response in the living and non-living**. London: Longmans, Green and Co, 1902.
- BOWLES, D. J. Defense-related proteins in higher plants. **Annu Ver Biochem**, v.59, p. 873–907, 1990.
- BOX, G. E. P.; JENKINS, G. M. **Time Series Analysis: forecasting and control**. San Francisco: Holden-Day, 1970.
- BRENNER, E. D. et al. Plant Neurobiology: an Integrated View of Plant Signaling. **Trends in Plant Science**, v.11, p.413-9, 2006.
- BULYCHEV, A. A.; KAMZOLKINA, N. A. Effect of action potential on photosynthesis and spatially distributed H<sup>+</sup> fluxes in cells and chloroplasts of *Chara corallina*. **Russian Journal of Plant Physiology**, v.53, p. 5–14. 2006.
- BURIOKA, N. et al. Approximate Entropy in the electroencephalogram during wake and sleep. **Clin. EEG and Neurosc.**, v.36, p.21-24, 2005.
- CERF, R. et al. Episodes of low-dimensional self-organized dynamics from electroencephalographic a-signals. **Biol Cybern.**, v.77, p. 235-245, 1997.

CHANG, C.-C.; LIN, C.-J. LIBSVM: A library for support vector machines. **ACM Transactions on Intelligent Systems and Technology**, v.2, p. 1–27, 2011.  
Disponível: < <http://www.csie.ntu.edu.tw/~cjlin/libsvm>>.

CHIALVO, D.; MICHAELS, M. C.; JALIFE, J. Supernormal excitability as a mechanism of chaotic dynamics of activation in cardiac Purkinje fibers. **Circ. Res.**, v.66, p. 525-545, 1990.

COOMANS, D.; MASSART, D. Alternative k-nearest neighbour rules in supervised pattern recognition. **Analytica Chimica Acta**, v.136, p. 15-27, 1982.

CORTES, C.; VAPNIK, V. Support vector networks. **Machine Learning**, v.20, p. 273–297, 1995.

DAVIES, E. Intercellular and intracellular signals in plants and their transduction via the membrane – cytoskeleton interface. **Seminars in Cell Biology**, v.4, p. 139–147, 1993.

DAVIES E. New functions for electrical signals in plants. **New Phytol.**, v.161, p. 607-610, 2004.

DAVIES, W.J.; ZHANG, J. Root signals and the regulation of growth and development of plants in drying soil. **Annu. Rev. Plant Physiol. Plant Mol. Biol.**, v.42, p. 55–76, 1991.

DAVIES, E.; SCHUSTER, A. M. Intercellular communication in plants: evidence for a rapidly generated, bidirectionally transmitted wound signal. **Proc Natl Acad Sci USA**, v.78, p. 2422–2426, 1981.

DEEKEN, R. et al. Development and light-dependent regulation of a phloem-localised K<sup>+</sup> channel of *Arabidopsis thaliana*. **Plant Journal**, v.23, p. 285–290, 2000.

DEGN, H.; HOLDEN, A. V.; OLSEN, L. F. (Eds). **Chaos in Biological Systems**. New York: Plenum, 1986.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**. 2.ed. Wiley-Interscience Publication, 2000.

ERDEI, L. et al. Environmental stress and the biological clock in plants: changes of rhythmic behaviour of carbohydrates, antioxidant enzymes and stomatal resistance by salinity. **Journal of Plant Physiology**, v.152, p. 265-271, 1998.

FALCÃO, A. X.; STOLFI, J.; LOTUFO, R. A. The image foresting transform: Theory, algorithms, and applications. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v.26, p. 19–29, 2004.

FILEK, M.; KOSCIELNIAK, J. The effect of wounding the roots by high temperature on the respiration rate of the shoot and propagation of electric signal in horse bean seedlings (*Vicia faba* L minor). **Plant Sci**, v.123, p. 39–46, 1997.

FRANCHINI, J. C. et al. **Manejo do solo para redução das perdas de produtividade pela seca**. Londrina: EMBRAPA SOJA, 2009.

FROMM J. Control of phloem unloading by action potentials in Mimosa. **Physiol Plant**, v.83, p. 529–533, 1991.

FROMM, J.; BAUER, T. Action potentials in maize sieve tubes change phloem translocation. **J Exp Bot**, v.45, p. 463–469, 1994.

FROMM, J.; ESCHRICH, W. Correlation of ionic movements with phloem unloading and loading in barley leaves. **Plant Physiology and Biochemistry**, v.27, p. 577–585, 1989.

FROMM, J.; FEI, H. Electrical signaling and gas exchange in maize plants of drying soil. **Plant Science**, v.132, p. 203–213, 1998.

FROMM, J.; HAJIREZAEI, M.; WILKE, I. The biochemical response of electric signaling in the reproductive system of Hibiscus plants. **Plant Physiol**, v.109, p. 375–384, 1995.

FROMM, J.; LAUTNER, S. Characteristics and functions of phloem-transmitted electrical signals in higher plants. In: BALUSKA, F.; MANCUSO, S.; VOLKMANN, D. (Eds.) **Communication in Plants** – neuronal aspects of plant life. Germany: Verlag, Berlin and Heidelberg, 2006. p. 321–332.

FROMM, J.; LAUTNER, S. **Electrical signals and their physiological significance in plants**. *Plant, Cell and Environment*, 30: 249–257, 2007.

GARDINER, J. Fractals and the irreducibility of consciousness in plants and animals. **Plant Sign. Behav.**, v. 8; p. e25296, 2013.

GARDINER, J. Insights into plant consciousness from neuroscience, physics and mathematics: a role for quasicrystals? **Plant Sign. Behav.**, v.7, p.1049-1055, 2012.

GARZÓN, F. C. The Quest for Cognition in Plant Neurobiology. **Plant, Sign. and Behav.**, v.2, p. 208-211, 2007.

GARZÓN, P. C.; KEIJZER, F. Cognition in Plants. In: BALUŠKA, F. (Ed.) **Plant – environment interactions: behavioral perspective**. Elsevier, 2009. p. 247-263.

GERSTEIN, G. L.; MANDELROT, M. Random walk models for the spike activity of a single neuron. **Biophysical Journal**, v.4, p. 41-68, 1964.

GLASS, L.; HUNTER, P.; McCULLOCH, A. (Eds.). **Theory of Heart: biomechanics, biophysics, and nonlinear dynamics of cardiac function**. New York, Springer-Verlag, 1991.

GUEVARA, M. R.; SHRIER, A.; GLASS, L. Phase-locking, period-doubling bifurcations and irregular dynamics in periodically stimulated cardiac cells. **Science**, v.214, p. 1350-1353, 1981.

GOLDBERGER, A. L. et al. Fractal dynamics in physiology: Alterations with disease and aging. **PNAS**, v.99, n.1, p. 2466-2472, 2002.

HALL, P.; PARK, R. J. Samworth, Choice of neighbor order in nearest-neighbor classification. **Ann. Statist.**, v.36, n.5, p. 2135–2152, 2008.

HAYKIN, S. **Neural Networks: a comprehensive foundation**. 2. ed. Prentice Hall, 1999.

HEIL, M.; SILVA BUENO, J.C. Within-plant signaling by volatiles leads to induction and priming of an indirect plant defense in nature. **Proc Natl Acad Sci USA**, v.104, n.1, p. 5467–5472, 2007.

HEIL, M.; TON, J. Long-distance signalling in plant defence. **Trends Plant Sci.**, v.13, p. 264–272, 2008.

HICKEY, T.J. et al. Interval arithmetic: from principles to implementation. **Journal of the ACM**, v.48, p. 1038–1068, 2001.

HO, K. K. et al. Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. *Circulation*, v.96, n.3. p. 842–848, 1997.

HODGKIN, A. L.; HUXLEY, A. F. A quantitative description of membrane current and its application of conduction and excitation in nerve. **J. Physiol. (London)**, v.117, p. 500-544, 1952.

HOWE, G. A.; JANDER, G. Plant immunity to insect herbivores. **Annu Ver Plant Biol.**, v.59, p. 41–66, 2008.

HUMPHRIES, N.E. et al. Environmental context explains Lévy and Brownian movement patterns of marine predators. **Nature**, v.465, n.7301, p. 1066–1069, 2010.

HUTT, M. T.; LUTTGE, U. Nonlinear Dynamics as a Tool for Modelling in Plant Physiology. **Plant biology**, v.4, n.3, p. 281-297, 2002.

INCHAUSTI, P.; HALLEY, J. The long-term temporal variability and spectral colour of animal populations. **Evol Ecol Res.**, v.4, p. 1033-1048, 2002.

JAYNES, E. T. **Probability theory: the logic of science**. Cambridge University Press, 2003.

KAPLAN, D. T., M. I. et al. Aging and the complexity of cardiovascular dynamics. **Biophys. J.**, v.59, p. 945-949, 1991.

KEMPERS, R.; AMMERLAAN, A.; VAN BEL A. J. E. Symplasmic constriction and ultrastructural features of the sieve element/companion cell complex in the transport phloem of apoplasmically and symplasmically phloem-loading species. **Plant Physiology**, v.116, p. 271–278, 1998.

KLAUS, A. et al. (Eds.). Statistical Analyses Support Power Law Distributions Found in Neuronal Avalanches. **PLoS ONE.**, v.6, n.5, p. e19779, 2011.

KOZIOLEK, C. et al. Transient knockout of photosynthesis mediated by electrical signals. **New Phytologist**, v.161, p. 715–722, 2004.

KOIWA, H.; HASEGAWA, P. M.; BRESSAN, R. A. Regulation of protease inhibitors and plant defense. **Trends on plant Science**, v.2, n.10, p. 379-384, 1999.

KRASNER, S. (Ed.). **The Ubiquity of Chaos**. Washington, DC: American Association for the Advancement of Science, 1990.

KSENZHEK, O. S.; VOLKOV, A. G. **Plant Energetics**. Academic Press, 1998. 89p.

LACOMBE, B. et. al. A shaker-like K<sup>+</sup> channel with weak rectification is expressed in both source and sink phloem tissues of Arabidopsis. **Plant Cell**, v.12, p. 837–851, 2000.

LAUTNER, S. et al. Characteristics of electrical signals in poplar and responses in photosynthesis. **Plant Physiology**, v.138, p. 2200–2209, 2005.

LINKENKAER-HANSEN et. al. Scaling Behavior in Human Brain Oscillations. **J. Neurosci.**, v.21, n.4, p.1370–1377, 2001.

LISAR, S. Y. S. et al. Water stress in plants: causes, effects and responses. In: RAHMAN, I. M. M.; HASEGAWA, H. **Water stress**. Croatia: Intech, 2012. p.01-14.

LÜTTGE, U.; BECK, F. Endogenous rhythms and chaos in crassulacean acid metabolism. **Planta**, v.188, n.1, p.28-28, 1992.

MALIHE, S.; KATEBI, S.; BOOSTANI, R. Entropy and complexity measures for EEG signal classification of schizophrenic and control participants. **Artificial Intelligence in Medicine**, v.47, n.3, p. 263-274, 2009.

MANCUSO, S. Hydraulic and electrical transmission of wound-induced signals in *Vitis vinífera*. **Australian Journal of Plant Physiology**, v.26, n.1, p. 55–61, 1999.

MARKOV, S. On direct interval arithmetic and its applications. **J. Universal Computer Science**, v.1, n.7, p. 514–526, 1995.

MAYER-KRESS, G. ed. **Dimensions and entropies in chaotic systems: quantification of complex behavior**. Berlin: Springer-Verlag, 1986.

MAZZOLAI, B. et al. Plant signaling & behavior. **Landes Bioscience**, n. 5, p.2, p. 90-93, 2010.

MCAINSH, et al. Stimulus-induced oscillations in guard cell cytosolic free calcium. **The Plant Cell**, v.7, p.8, p. 1207-1219, 1995.

MCAINSH M. R.; HETHERINGTON, A. M. Encoding specificity in calcium signalling systems. **Trends Plant Sci.**, v.3, p. 32-36, 1998.

MITHÖFER, A.; BOLAND, W.; MAFFEI, M. E. Chemical ecology of plant-insect interactions. In: Parker, J. (Ed.) **Annual Plant Reviews: plant disease resistance**. Chichester, UK: Wiley-Blackwell, 2009.p. 261–291.

MOORE, R. E. Methods and Applications of Interval Analysis. **Studies in Applied Mathematics-SIAM**, Philadelphia, 1979.

MORETTIN, P. A.; TELOI, C. M. C. **Análise de séries temporais**. Blucher, 2004.

NEWMAN, M. E. J. Power laws, Pareto distributions and Zipf's law. **Contemporary Physics**, v.46, n.5, p. 323–351, 2005.

NISSEN, S. **Implementation of a Fast Artificial Neural Network Library (FANN)**. 2003. Department of Computer Science University of Copenhagen (DIKU). Software Disponível em: <<http://leenissen.dk/fann/>>

NUNES AMARAL, L. A. et al. Scale-independent measures and pathologic cardiac dynamics. **Phys. Rev. Lett.**, v.81, n.2388, 1998.

OPPENHEIM, A. V.; SCHAFER, R. W.; BUCK, J. R. **Discrete signal processing**. Englewood Cliffs, NJ: Prentice Hall, 1989.

PALUS, M. Nonlinearity in normal human EEG: cycles, temporal asymmetry, nonstationarity and randomness, not chaos. **Biol Cybern**, v.75, p. 389–396, 1996.

PAPA, J. P. et al. Efficient supervised Optimum-Path Forest classification for large datasets. **Pattern Recognition**, v.45, n. 512–520, 2012.

PAPA, J. P.; FALCÃO, A. X. A new variant of the optimum-path forest classifier. **International Symposium on Visual Computing**, v.47, n.1, p. 935–944, 2008.

PAPA, J. P.; FALCÃO, A. X.; SUZUKI, C. T. N. Supervised pattern classification based on Optimum-Path Forest. **International Journal of Imaging Systems and Technology**, v.19, n.2, p. 120–131, 2009.

PINCUS, S. M.; GLADSTONE, I. M.; EHRENKRANZ, R. A. A regularity statistic for medical data analysis. **J. CZin. Monit.**, v.7, p. 335-345, 1991.

PINCUS, S. M.; VISCARELLO, R. R. Approximate entropy: a regularity measure for fetal heart rate analysis. **Obstet. Gynecol.** regularity measure for fetal heart rate analysis. **Obstet. Gynecol.**, v.79, p. 249-255,1992.

PINCUS, S. M. Approximating markov chains. **Proc. N&Z. Acad. Sci.**, v.89, p. 4432-4436, 1992.

PINCUS, S. M.; CUMMINS, T. R.; HADDAD, G. G. Heart rate control in normal and aborted-SIDS infants. **Am. J . Physiol.**, v.264, n.33, p. 638-646, 1993.

PINCUS, S. M. Approximate entropy (ApEn) as a complexity measure. **AIP**, v.5, p. 110, 1995.

PINCUS, S. M. Assessing serial irregularity and its implications for health. **Ann. N.Y. Acad. Sci.**, v.954, p. 245-267, 2001.

RICHMAN, J. S.; MOORMAN, J. R. Physiological time-series analysis using approximate entropy and sample entropy. **Am. J. Physiol.**, v.278, p.H2039-49, 2000.

RYAN, C.A. Protease inhibitors in plants: genes for improving defenses against insects and pathogens. **Annu Rev Phytopathol.**, v.28, p. 425–449, 1990.

SCHÖLKOPF, B.; SMOLA, A. J. **Learning with Kernels**. Cambridge, MA: MIT Press, 2002.

SCHROEDER, J. I.; HEDRICH, R.; FERANDEZ, J. M. Potassiumselective single channels in guard cell protoplasts of *Vicia faba*. **Nature**, v.312, p.361–362, 1984.

SHABALA, S. N.; NEWMAN, I. A.; MORRIS, J. Oscillations in H<sup>+</sup> and Ca<sup>2+</sup> Ion Fluxes around the Elongation Region of Corn Roots and Effects of External Ph. **Plant Physiology**, v.113, n.1, p. 111-118, 1997.

SOUZA, G. M.; BUCKERIDGE, M. S. Sistemas complexos: novas formas de ver a Botânica. **Revista Brasil. Bot.**, v.27, n.3, p. 407-419, 2004.

SOUZA, G. M.; RIBEIRO, R. V.; PINCUS, S. M. Changes in network connectance and temporal dynamics of gas exchange in *Citrus sinensis* under different evaporative demands. **Braz. J. Plant Physiol.**, v.16, p. 119-130, 2004a.

SOUZA, G. M. et al. Approximate Entropy as a measure of complexity in sap flow temporal dynamics of two tropical tree species under water deficit. **An. Acad. Ciênc.**, v.76, p. 625-630, 2004b.

SOUZA, G. M., PINCUS, S. M., MONTEIRO, J. A. F. The complexity-stability hypothesis in plant gas exchange under water deficit. **Braz. J. Plant Physiol.**, v.17, p. 363-373, 2005.

SOUZA, G. M. Inteligência em plantas como um fenômeno ecofisiológico. In: CLAE; IXCEB, 3. **Anais...** Setembro, São Lourenço – MG, 2009.

SOUZA, G. M. et al. Soybean Under Water Deficit: Physiological and Yield Responses. In: BOARD, E. J. (Org). A comprehensive survey of international soybean research-genetics, physiology, agronomy and nitrogen relationships. **InTech**, v.1, p. 273-299, 2013.

SPANSWICK, R. M. Electrical coupling between cells of higher plants: a direct demonstration of intercellular communication. **Planta**, v.102, p. 215–227, 1972.

SPANSWICK, R. M.; COSTERTON, J. W. F. Plasmodesmata in *Nitella translucens*: structure and electrical resistance. **Journal of Cell Science**, p.2, p. 451–464, 1967.

SRINIVASAN, V., ESWARAN, C., SRIRAAM, N. Approximate Entropy-based epileptic EEG detection using artificial neural networks. **Inform. Tech. in Biomed.**, 11: 288-295, 2007.

STAM, C.J. et al. Dynamics of the human alpha rhythm: evidence for non-linearity? **Clin Neurophysiol.**, v.110, p. 1801–1813, 1999.

STANKOVIC, B.; DAVIES, E. Wounding evokes rapid changes in tissue deformation, electrical potential, transcription, and translation in tomato. **Plant Cell Physiol.**, v.39, p. 268–274, 1997.

STANKOVIC, B. et al. Action potentials and variation potentials in sunflower: an analysis of their relationship and distinguishing characteristics. **Physiologia Plantarum**, v.103, p. 51–58, 1998.

STENHOUSE, D. **The Evolution of intelligence-a general theory and some of its implications**. London: George Allen and Unwin, 1974.

THELLIER, M.; LÜTTGE, U. Plant Memory: a tentative model. **Plant Biology**, German Botanical Society and The Royal Botanical Society of the Netherlands, p.1-12, 2012.

TREWAVAS, A. Aspects of plant intelligence. **Ann. Bot. (Lond.)**, v.92, p.1–20, 2003.

TREWAVAS, A. What is plant behaviour? **Plant Cell & Environment**, v.32, p. 606, 2009.

UMRATH K. Untersuchungen über Plasma und Plasmaströmungen an Characeen. IV. Potentialmessungen an *Nitella mucronata* mit besonderer Berücksichtigung der Erregungserscheinungen. **Protoplasma**, v.9, p. 576–597, 1930.

VAN BEL, A. J. E.; EHLERS, K. Electrical signalling via plasmodesmata. In: OPARKA, K.J. (Ed.). **Plasmodesmata**. Oxford, UK: Blackwell Publishing, 2004. p. 263–278.

VAN DER POL, B.; VAN DER MARK, J. The heartbeat considered as a relaxation oscillator and an electrical model of the heart. **Phil. Mag.**, v.6, p. 763-775, 1928.

VAPNIK, V. N. An Overview of Statistical Learning Theory. **IEEE Transactions on Neural Networks**, v.10, n.5, p. 988–999, 1999.

VOLK, G.; FRANCESCHI, V. R. Localization of a calcium channel-like protein in the sieve element plasma membrane. **Australian Journal of Plant Physiology**, v.27, p. 779–786, 2000.

VOLKOV A. G.; HAACK, R. A. Insect-induced bioelectrochemical signals in potato plants. **Bioelectrochemistry and Bioenergetics**, v.37, n.1, p. 55-60, 1995.

VOLKOV A. G. **Plant Electrophysiology: methods and cell electrophysiology.** Springer, 2012.

VOLKOV, A. G.; COLLINS, D. J.; MWESIGWA, J. Plant electrophysiology: pentachlorophenoil induces fast action potentials in soybean. **Plant Science**, v.153, p. 185–190, 2000.

WANG, Z. Y. et. al. Monitoring system for electrical signals in plants in the greenhouse and its applications. **Biosystems Engineering**, v.10, p. 1–11, 2009.

WILDON, D. C. et. al. **Electrical signalling and systemic proteinase- Inhibitor induction in the wounded plant.** Nature, v.360, p. 62–65, 1992.

WITTEN, I. H.; FRANK, E. **Data Mining: practical machine learning tools and techniques.** 2. ed. Morgan Kaufmann, 2005.

YUAN, Q. Epileptic EEG classification based on extreme learning machine and nonlinear features. **Epilepsy Research**, v.96, n.1-2, p. 29–38, 2011.

YUN, K. Decreased cortical complexity in methamphetamine abusers. **Psychiatry Research: Neuroimaging**, v.201, n.3, p.226–32, 2012.

## ANEXO

Computers and Electronics in Agriculture 145 (2018) 35–42



Contents lists available at ScienceDirect

## Computers and Electronics in Agriculture

journal homepage: [www.elsevier.com/locate/compag](http://www.elsevier.com/locate/compag)

## Original papers

## Automatic classification of plant electrophysiological responses to environmental stimuli using machine learning and interval arithmetic

Danillo Roberto Pereira<sup>a</sup>, João Paulo Papa<sup>b,\*</sup>, Gustavo Francisco Rosalin Saraiva<sup>a</sup>, Gustavo Maia Souza<sup>c</sup><sup>a</sup> Universidade do Oeste Paulista, Presidente Prudente, São Paulo, Brazil<sup>b</sup> Universidade Estadual Paulista "Júlio de Mesquita Filho", Bauri, São Paulo, Brazil<sup>c</sup> Universidade Federal de Pelotas, Pelotas, Rio Grande do Sul, Brazil

## ARTICLE INFO

**Keywords:**  
Plant stress  
Optimum-Path Forest  
Convolutional Neural Networks  
Interval Arithmetic

## ABSTRACT

In plants, there are different types of electrical signals involving changes in membrane potentials that could encode electrical information related to physiological states when plants are stimulated by different environmental conditions. A previous study analyzing traits of the dynamics of whole plant low-voltage electrical showed, for instance, that some specific frequencies that can be observed on plants growing under undisturbed conditions disappear after stress-like environments, such as cold, low light and osmotic stimuli. In this paper, we propose to test different methods of automatic classification in order to identify when different environmental cues cause specific changes in the electrical signals of plants. In order to verify such hypothesis, we used machine learning algorithms (Artificial Neural Networks, Convolutional Neural Network, Optimum-Path Forest, *k*-Nearest Neighbors and Support Vector Machine) together Interval Arithmetic. The results indicated that Interval Arithmetic and supervised classifiers are more suitable than deep learning techniques, showing promising results towards such research area.

## 1. Introduction

Plants as sessile and modular organisms face the challenge to keep their stability growing in environments under constant changing (Souza et al., 2016). Since plants lack a central command to organize the environmental information gathered in each module (e.g., a branch root or a leaf), an efficient communication system has evolved in order to integrate local information (cell-to-cell communication) and to signalize through the plant body (long-distance communication) (Trewavas, 2003; Lüttge, 2012).

Long-distance communication, also referred to as systemic communication, can be triggered by different stimuli, such as biotic ones (e.g., systemic acquired resistance as a response to pathogens) or by abiotic stimuli (e.g., water deficit, heat, and salinity). Therefore, the ultimate goal of these systemic signaling is to activate response mechanisms in remote tissues, improving the ability of the whole plant to prepare its tissues to an upcoming challenge (Gilroy et al., 2014). Among the signals involved in long-distance communication, ROS (reactive oxygen species), calcium and electrical signals perform a central role (Baluska, 2016).

In plants, there are different types of electrical signals, which are

electrical activities involving changes in membrane potential, such as action potential (AP), variation potential (VP, or slow wave - SW), and system potential (Davies, 2006; Sukhova et al., 2017). APs are characterized by spike-like changes of the resting membrane potential and, independent of the stimulus strength, start propagating through the plant with a defined amplitude and velocity. Like in animals, APs seem to be all-or-nothing events (Fromm and Spanswick, 1993; Pyatygin et al., 2008). VPs differ from APs in various ways. VPs do not obey the all-or-nothing law, they are known as slow wave potentials (SWPs) with variable shape, amplitude and time frame. Moreover, the signals are related with the stimulus strength, and last for periods of 10 s up to 30 min (Zimmermann and Mithofer, 2013; Vodeneev et al., 2015). System potentials (SPs), in contrast to APs and VPs, reflect a systemic self-propagating hyperpolarization of the plasma membrane or depolarization of the apoplasmic voltage. Like VPs, SPs have a magnitude and duration that are depended on the stimulus, but they are initiated via membrane hyperpolarization through the sustained activation of the proton pump. SPs are dependent on experimental conditions, and then they may occur under a very specific set of environmental conditions (Zimmermann and Mithofer, 2013; Choi et al., 2016). Furthermore, quite often all these signals are mixed altogether, which impairs a

\* Corresponding author.

E-mail addresses: danilopereira@unoeste.br (D.R. Pereira), papa@fc.unesp.br (J.P. Papa), gustavosaraiva88@gmail.com (G.F.R. Saraiva), gumaia.gms@gmail.com (G.M. Souza).

<https://doi.org/10.1016/j.compag.2017.12.024>Received 15 August 2017; Received in revised form 12 December 2017; Accepted 15 December 2017  
0168-1699/© 2017 Elsevier B.V. All rights reserved.

proper signal analysis (van Bel et al., 2014; Saraiva et al., 2017). Strong evidences have demonstrated that bioelectrical signals play a central role in both cell-to-cell and long-distance communication in plants (Baluska et al., 2006; Zimmermann et al., 2009; van Bel et al., 2014), also supporting the ability to adjust their phenotypes to different environmental conditions (Fromm and Lautner, 2007; Gallè et al., 2015; Ríos-Rojas et al., 2014). For instance, Sukhov et al. (2014) and Magdalena et al. (2017) have demonstrated the role of electrical signals in the regulation of photosynthetic responses to different stimuli.

Very recently, Souza et al. (2017) proposed the concept of “plant electrome” based on the general proposition of “electrome” by De Loof (De Loof, 2016), describing the totality of the ionic currents in different scales of plant organization. By measuring low-voltage electrical signal using electrophotography (EPG) (Debono, 2013; Souza et al., 2017), Souza et al. (2017) showed that different environmental stimuli could change some characteristics of the temporal dynamic of the electrical signaling, including the level of complexity. It was noticed that some specific frequencies, which were observed in non-stimulated plants, have disappeared after stimulation. Moreover, the environmental stimuli changed the type of color noise of the electrical signals. However, it was not clear if the different environmental cues (cold, low light, and osmotic stress) caused distinct effects on the plant signals (Souza et al., 2017).

Therefore, measuring the level of stress in plants is of crucial importance to a better understanding of their working mechanism. However, automatic plant stress identification by means of machine learning techniques has been considered recently only. Singh et al. (2016) presented an overview about machine learning tools and their applications in the context of biotic and abiotic stress traits classification. The main goal of such work is to guide the plant community when using machine learning techniques in the aforementioned situation. Two years earlier, Ma et al. (2014b) also considered a similar study, but in a more general way, and Ma et al. (2014a) employed machine learning to study stress-responsive transcriptomes in *Arabidopsis thaliana*. The experiments highlighted that such tools were able to outperform standard statistical approaches. Shaik and Ramakrishna (2014) used machine learning techniques to identify multiple stress conditions genes for broad resistance in rice, and Behmann et al. (2015) presented a review of different machine learning techniques applied for biotic stress identification in precision crop protection.

Chatterjee et al. (2014) established a relationship between the light stimulus and plant electrical response for different light stimuli intensity considering 19 different plants (17 *Zamioculcas zamiifolia* and 2 *Cucumis sativus* plants). The best results were obtained by Nonlinear Hammerstein-Wiener (NLHW) a good matching over others fitting methods. Later on, Chatterjee et al. (2015) classified three different types of stimuli (NaCl, H<sub>2</sub>SO<sub>4</sub> and O<sub>3</sub>) based on the response of electrical signals of tomatoes and cucumbers. The authors used different machine learning algorithms (FLDA - Fisher Linear Discriminant Analysis, QDA - Quadratic Discriminant Analysis, NB - Naïve Bayes, and Mahalanobis Classifier) considering eleven features extracted from the electrical signal using linear and nonlinear methods. The best result was around 73.67% of recognition rate.

Chen et al. (2016) applied four classifiers (Template Matching, Artificial Neural Networks, Support Vector Machines and Deep Belief Networks) for the recognition of plant stimuli from electrical signals. The aforementioned work combined a waveform-based feature extractor and the Principal Component Analysis (PCA) approach, obtaining around 96% of recognition rate with Template Matching.

In this paper, we propose to use the concept of plant electrome (Souza et al., 2017) to automatically identify whether different environmental cues cause specific changes in the electrical signals of soybean plants. In order to verify such hypothesis, we considered using machine learning algorithms and arithmetic intervalar, a branch of mathematical tools that allows one to extend standard numbers to an interval representation. Therefore, the main contributions of this paper are:

- to use the plant electrome data as input for machine learning-based prediction of plant stress; and
- to employ deep learning techniques for plant stress identification.

The remainder of this paper is organized as follows. Sections 2 and 3 present the theoretical background and methodology used in this paper, as well as the results obtained using the proposed approach, respectively. Finally, Section 4 states conclusions and future works.

## 2. Materials and methods

### 2.1. Data acquisition

All datasets used herein to test the different methods of classification are part of the study published by Souza et al. (2017). The data consist of time series of low-voltage variation ( $\Delta V$  in  $\mu V$ ) measured in soybean plants subjected to different environmental stimuli: cold, low light and osmotic stress. The protocol of data acquisition was defined by Saraiva et al. (2017), using a signal amplifier (model MP36, Biopac Systems, US) inside a grounded Faraday cage. The measurements were carried out with one reference electrode attached to the grounded Faraday cage, and two electrodes inserted in the plants operating in a differential mode, where the instrumental amplifier cuts off the similar frequencies recorded in both electrodes. The sampling rate was 125 Hz with a high-pass filter settled to allow pass higher frequencies (>0.5 Hz), since the objective of that study was investigate the low-voltage noise that underlies the electrical signals (see more details in Saraiva et al. (2017)).

### 2.2. Datasets

The datasets described in the previous section were cropped to contain features per sample (signals obtained from the plants). Besides the large number of features, the signal is not so homogenous, therefore applying classical machine learning methods in the raw data is not advisable. To overcome this weakness, we applied some concepts of Arithmetic Intervalar to map raw data into lower-dimensional feature space.

In our work we consider four different datasets, as follows:

- **cold**: 67 signals obtained from plants in ideal conditions (without stress) and 76 signals obtained after cold stress.
- **low light**: 152 signals obtained in ideal conditions (without stress) and 118 signals after low light stress.
- **osmotic**: 123 signals obtained in ideal conditions (without stress) and 145 signals after osmotic stress.
- **all**: 342 signals obtained in ideal conditions (without stress), 76 signals after cold stress, 118 signals after low light stress, and 145 signals obtained after osmotic stress.<sup>1</sup>

Some examples of the signals from each class are depicted in Fig. 1.

### 2.3. Theoretical background

In this section, we present a brief theoretical background related to the machine learning and feature mapping techniques based on Interval Arithmetic used in this work.

#### 2.3.1. Interval Arithmetic

The Interval Arithmetic (IA) was proposed by Moore in the 1960's, being the main idea to represent values as a range model instead of

<sup>1</sup> This dataset is a merge of the cold, low light and osmotic datasets. The main idea of this dataset is to verify whether the methods are able to differentiate the stress type or not.

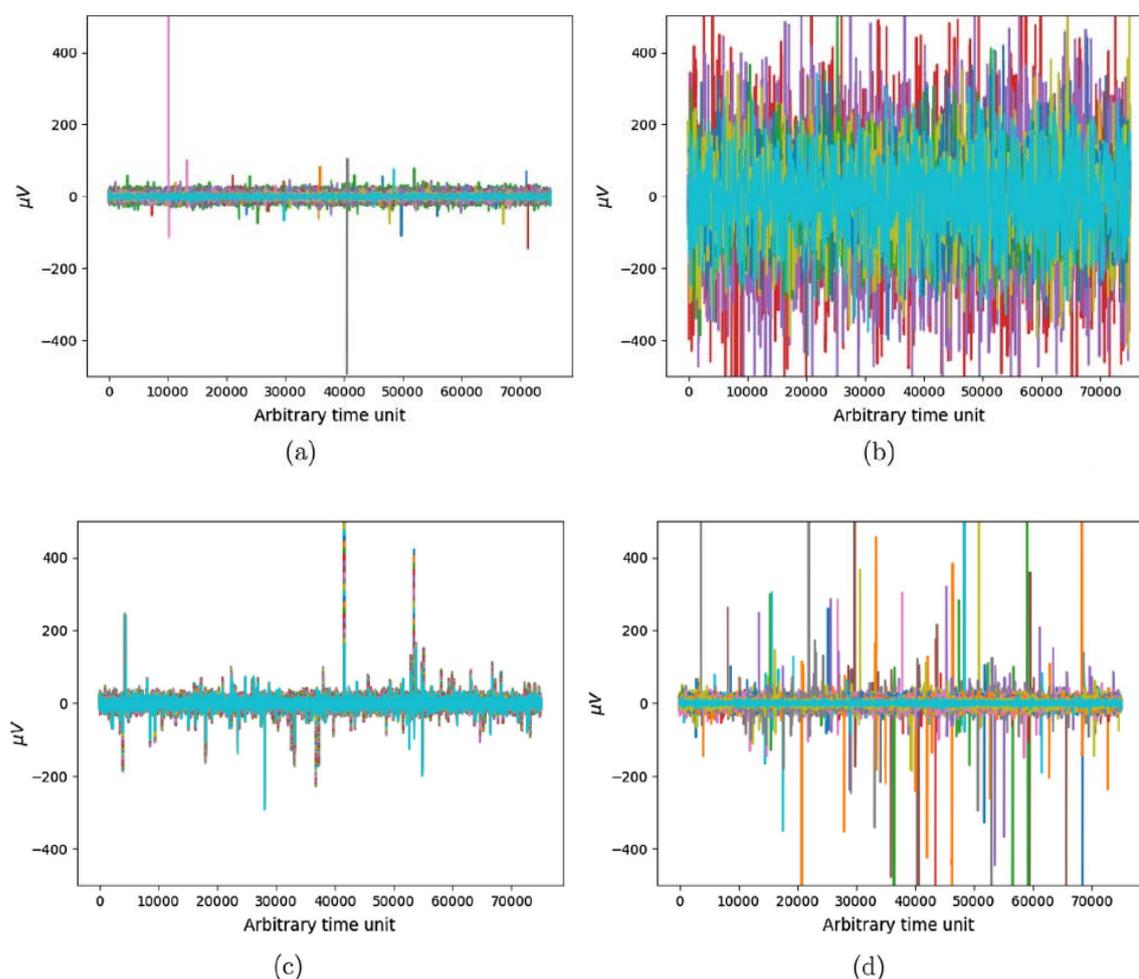


Fig. 1. Examples of some signals obtained from the plants: (a) ideal condition, (b) after cold stress, (c) after low light, and (d) after osmotic stress.

single numbers. In such representation, an interval  $I$  is denoted by an non-empty real-valued range  $[I_l, I_u]$ , such that each interval encodes a subset of real numbers that satisfy the following condition:

$$I = \{x \in \mathbb{R} \mid I_l \leq x \leq I_u\}. \quad (1)$$

As in the traditional arithmetic, the IA background defines a set of relations and operations (comparison, join, sum, and multiply) over the intervals (Moore, 1966; Moore and Yang, 1959; Moore, 1960, 1962, 1979). However, the Interval Arithmetic is more powerful than traditional arithmetic, since any real number  $x$  can be represented by the singular interval  $[x, x]$ . The Interval Arithmetic is a useful apparatus to provide representations of error bounds and uncertainty. This property is extremely interesting when working with the variations present in the EEG signals obtained from plants.

### 2.3.2. Machine learning techniques

**2.3.2.1. Optimum-Path Forest.** The Optimum-Path Forest (OPF) classifier models the problem of pattern recognition as a graph partition task, in which a predefined set of samples from each class (*prototypes*) compete for a minimal path-cost attribution to the rest of the samples. Such process results in a collection of optimum-path trees rooted at the prototype nodes, building an optimum-path forest considering from all training samples. Test samples are classified through incrementally evaluating the optimum paths from the

prototypes, as though they were part of the forest, and assigning the labels of the most strongly connected roots. An OPF classifier can be designed as long as we use a smooth path-cost function (Falcão et al., 2004). Although there are two different versions of the supervised OPF classifier (Papa et al., 2009, 2017), in this paper we make use of the former and most widely used approach proposed by Papa et al. (2009) and further enhanced by Papa et al. (2012). The OPF classifier has been used in a number of applications in the last years, and particularly for leaf-based aquatic weed recognition (Pereira et al., 2012).

**2.3.2.2. Multilayer perceptron.** An Artificial Neural Network with Multilayer Perceptrons (ANN) is a feedforward neural network composed of several neuron layers aiming to solve multi-class problems (Haykin, 1999). The input to each layer is a weighted sum of the output from the previous layer, and the number of neurons in the first layer is equal to the number of features of the input, while the number of neurons in the last layer is equal to the number of classes. Such approaches can learn highly non-linear models that map inputs to their corresponding outputs.

**2.3.2.3.  $k$ -Nearest neighbors.** The  $k$ -nearest neighbors ( $k$ -NN) is a simple but effective technique that works pretty well in many different applications (Coomans and Massart, 1982; Hall et al., 2008). In contrast to the OPF, the  $k$ -NN uses all training samples as prototypes,

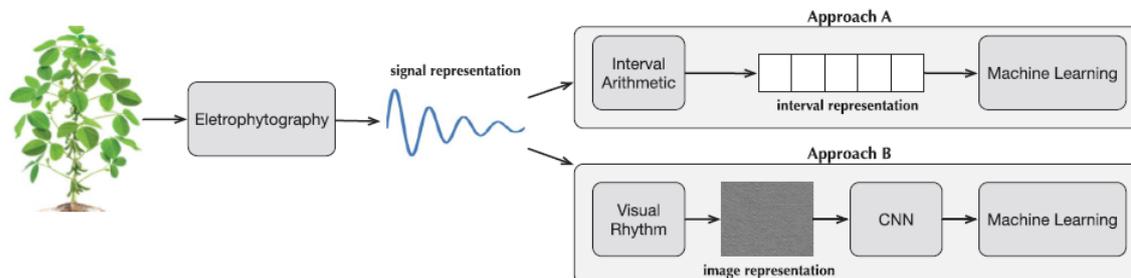


Fig. 2. Proposed pipeline for automatic plant stress identification.

and requires the input parameter  $k$  that establishes the number of neighbors that contribute to the classification of a given sample. For such purpose, a given test sample is labeled with the most frequent label in its  $k$ -neighborhood. Notice that for  $k = 1$ , the testing sample is assigned to the class of its closest training sample.

**2.3.2.4. Support Vector Machines.** The Support Vector Machines (SVM) approach based on the principle of structural risk minimization (Vapnik, 1999; Schölkopf and Smola, 2002; Cortes and Vapnik, 1995), aiming at establishing an optimal discriminative function among two classes of patterns while accomplishing the trade-off between generalization and overfitting. The standard SVM training algorithm constructs the optimal hyperplane separating a two-class feature space (Vapnik, 1999). However, some problems may require a more robust approach. In order to extend from linear to nonlinear classification, the *kernel trick* is used (Schölkopf and Smola, 2002), where kernel functions nonlinearly map input data into higher-dimensional feature spaces in a computationally-efficient manner.

**2.3.2.5. Convolutional Neural Networks.** Convolutional Neural Networks (CNN) are a special type of artificial neural networks that have been extensively used for unsupervised feature learning (Lecun et al., 1998). The main idea is to employ several layers of pooling and convolution operators in order to automatically extract features from input data that are invariant to some geometric operations (e.g., rotations and translations).

## 2.4. Methodology

In this work, we considered two distinct methodologies to learn features from the signal extracted from the plants: (i) the first one, hereinafter called approach A, is based on Interval Arithmetic (IA), and (ii) the second approach (approach B) is based on Visual Rhythm and CNNs. After extracting features, they are used to feed supervised machine learning techniques in order to classify whether the plant has

been affected by stress or not. Fig. 2 illustrates the aforementioned pipeline. The next sections describe in details the aforementioned approaches.

### 2.4.1. Approach A

As depicted in Fig. 2, approach A aims at decomposing the signal into an intervalar representation based on the theory presented in Section 2.3.1. Roughly speaking, the mapping process is straightforward and window-based. Let  $S = \{s_1, s_2, \dots, s_n\}$  be the discrete representation of the signal extracted from the soybean, where  $s_i$  stands for an energy acquisition at time step  $i$ . Additionally, let  $s$  be the window size, such that the whole signal  $S$  is equally divided into  $m = \lfloor n/s \rfloor$  bins  $b_k, k = 1, 2, \dots, m$ .

The idea of the intervalar representation is to represent each bin  $b_k$  as a triplet  $T_k = [min_k, avg_k, max_k]$ , where  $min_k$  and  $max_k$  stand for the minimum and the maximum values present in  $b_k$ , and  $avg_k$  denotes for the average value of the numbers in  $b_k$ . Therefore, the signal  $S$  is finally represented as a set of  $m$  triplets

$$S' = \{[min_1, avg_1, max_1], [min_2, avg_2, max_2], \dots, [min_m, avg_m, max_m]\},$$

which is further used to feed the supervised machine learning techniques. Fig. 3 illustrates the aforementioned procedure.

The proposed approach reduces considerably the number of features that describe the signal, which affects the computational cost as well. We performed tests with four different window sizes, say that 1000, 5000, 15,000 and 25,000, being the window size as of 15,000 the one that obtained the best results. Therefore, all experiments were conducted using such configuration. In this context, the number of features were reduced from to 15 ( $75,000/15,000 = 5$  buckets, being each one represented by a triplet).

### 2.4.2. Approach B

The second approach aims at encoding the signal as an image by using the so-called Visual Rhythm (Almeida et al., 2015), which basically stacks the signal into rows in order to generate the image.

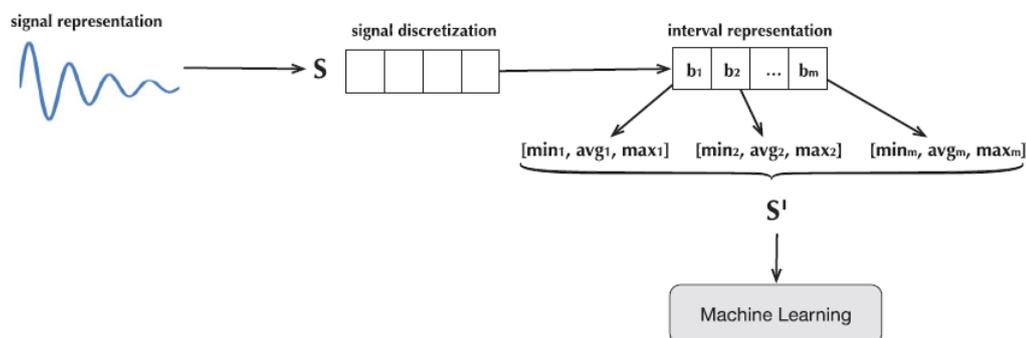


Fig. 3. Approach A used for interval-based signal representation.

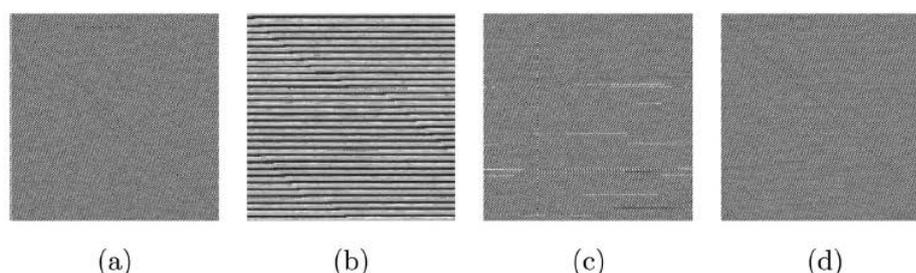


Fig. 4. Some examples of signals mapped into images under: (a) ideal conditions, (b) after cold stress, (c) after low light stress, and (d) after osmotic stress. The images were equalized for the sake of visualization purposes.

Therefore, each signal needs to be normalized into [0,1] for further being resized into a gray-scale squared matrix. Finally, one obtains a  $273 \times 273$  image, where each signal value denotes a pixel in this image. Some examples of images mapped using this methodology are presented in Fig. 4. Clearly, one can observe the different patterns encoded by the images with respect to different stress conditions.

Later on, the images generated using the aforementioned pipeline are then employed to feed an CNN technique. In this work, we performed tests with two different classical architectures, as follows:

- **ImageNet:** composed of 5 convolution layers, 5 pooling layers and 2 normalization layers. It is also constituted by 5 ReLU layers among the convolution ones, 2 inner product layers, 2 dropout layers, 1 softmax loss layer and 1 accuracy layer for testing purposes.
- **Cifar-10:** a quick version is used, composed of 3 convolution layers and 3 pooling layers. It is also constituted by 3 ReLU layers among the convolution ones, 2 inner product layers, 1 softmax loss layer and 1 accuracy layer for testing intentions.

However, as the Cifar-10 model obtained the best results (significantly better than ImageNet), all experiments using CNN were realized using this configuration.

### 2.5. Statistical evaluation

In regard to the comparison assessment, we used an accuracy measure proposed by Papa et al. (2009), which is similar to the Kappa index (Cohen, 1960), but being more restrictive. If there are two classes, for example, with very different sizes and a classifier always assigns the label of the largest class, its accuracy will fall drastically due to the high error rate on the smallest class. The accuracy is measured by taking into account the classes may have different sizes in the test set. Also, we experiments were executed in 20 runs with different training and test sets for the further statistical evaluation by means of the Wilcoxon signed-rank test with significance of 0.05 (Wilcoxon, 1945).

### 2.6. Setting-up machine learning techniques

We considered the LibSVM library (Chang and Lin, 2011) for the implementation of SVM classifier. The hyperparameters  $C$  and  $\sigma$  were determined via a 5-fold cross-validation grid-search in the range  $[2^{-15}, 2^{15}]$  with steps of 2, for each input parameter.

With respect to the ML classifier, we used the Fast Artificial Neural Network (FANN) library with two hidden layers consisting of eight neurons each (Nissen, 2003). The network was trained using the backpropagation algorithm (Haykin, 1999) with 70,000 epochs and the desired error as of  $10^{-4}$ . The learning rate was set to 0.1.

In regard to  $k$ -NN technique, we defined  $k$  as the best value of a grid-search within the range  $[1, \lfloor \frac{z}{5} \rfloor]$  with steps of two, where  $z$  is the number

of training samples. Finally, with respect to OPF, we used the LibOPF library.<sup>2</sup>

## 3. Results

In this section, we present the experimental results concerning the task of automatic identifying stress-like patterns before and after stress environment (i.e., cold, low light and osmotic). As aforementioned, we applied five different supervised classification algorithms: ANN, CNN, OPF,  $k$ -NN and SVM. Additionally, we considered the IA mapping experiment (procedure A) for ANN, OPF,  $k$ -NN, and SVM classifiers.

The main goal is to find an algorithm that can be successfully applied to learn a pattern even when trained on small datasets. Therefore, we partitioned the original dataset into training sets of increasing sizes, ranging from 10% to 90% of the original dataset. For each dataset, we executed each algorithm 20 times using a holdout configuration. Fig. 5 presents the results considering the “cold”, “low light”, “osmotic”, and “all” datasets, respectively.

One can observe that good performances were obtained in the four datasets by all methods, except ANN that achieved the worst results in “cold” and “low light” datasets. As expected, all techniques were able to improve their results upon larger training sets, mainly CNNs, which are complex models and usually need a considerable amount of data for leaning (see Fig. 5d). Since “all” dataset comprises all stress-like patterns (i.e., it seems to pose a greater challenge), CNNs obtained the worst results over small training sets.

Table 1 presents the average results using 90% of the datasets for training purposes, being the most accurate techniques in bold according to the Wilcoxon statistical test. One can observe that SVM obtained the best results for all datasets, followed by  $k$ -NN that obtained statistically similar results in the “cold” and “osmotic” datasets. OPF classifier achieved good accuracies for three datasets, and CNN also obtained good results for “cold” and “low light” datasets, although it obtained the worst results for the “osmotic” dataset. Roughly speaking, one can observe that AI combined with supervised classifiers can provide better results than deep learning techniques. The recognition rates showed one can obtain promising results concerning the task of automatically identifying stress-like patterns.

Fig. 6 depicts the confusion matrices using 90% of the datasets for training purposes with SVM classifier. As one can observe, the signs of plants after cold stress have a peculiar pattern that is easier to differentiate than others stress conditions since it obtained better results. Additionally, the “low light” and “osmotic” datasets presented similar results. The experiment with “all” dataset showed that is possible to detect patterns among different stress signals. However, one can observe a confusion between “low light” and “no stress” conditions (Fig. 6d), which may deteriorate the final classifier’s accuracy.

<sup>2</sup> <https://github.com/jpppsi/LibOPF>.

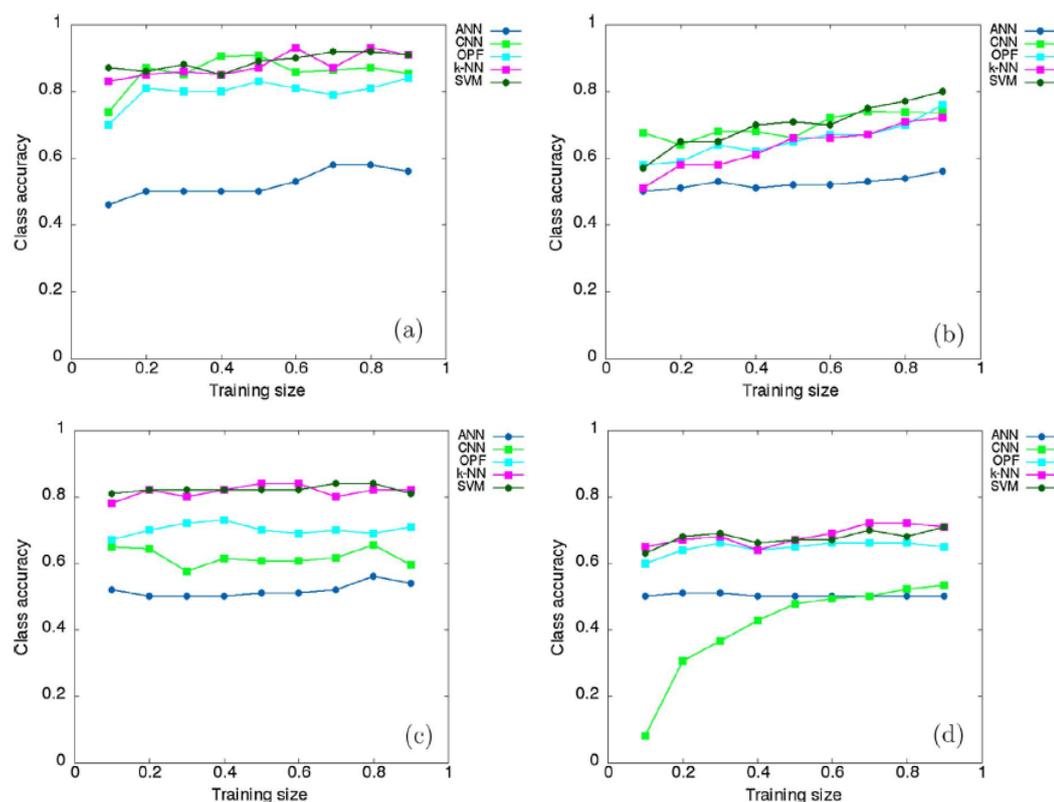


Fig. 5. Effect of different training set sizes for classification purposes over: (a) “cold”, (b) “low light”, (c) “osmotic”, and (d) “all” datasets.

**Table 1**  
Average recognition rates considering 90% of the samples for training purposes.

	Datasets			
	Cold	Low light	Osmotic	“all”
ANN	53.33 ± 0.00	58.57 ± 1.75	53.57 ± 0.00	63.11 ± 2.02
CNN	85.33 ± 9.07	73.57 ± 9.71	49.63 ± 6.22	53.33 ± 2.19
OPF	84.00 ± 5.33	77.86 ± 6.14	67.86 ± 6.39	65.23 ± 3.21
k-NN	90.67 ± 5.33	74.29 ± 5.25	81.43 ± 5.25	71.22 ± 0.65
SVM	90.67 ± 5.33	80.71 ± 7.35	80.97 ± 5.80	70.98 ± 1.01

#### 4. Discussion

Different environmental cues, biotic or abiotic, trigger electrophysiological responses in plants, which are basically an ion imbalance across plasma membranes leading to a voltage transient (Maffei and Bossi, 2006; Huber and Bauerle, 2016). While VPs and SPs are mostly triggered by invasive injurious stimuli, such as wounding, APs are elicited as responses to non-invasive stimuli including irradiation, moisture, and temperature (Fromm and Lautner, 2007; Ríos-Rojas et al., 2014; Sukhov et al., 2017). Herein, the data analysed were prevented from plants subjected to three abiotic stimuli (cold, osmotic, and low light) that affected the temporal dynamic of the low-voltage signals (Souza et al., 2017). The measured signals are supposed to be the resultant of all electrical activity of the cells around the electrodes, likely emerging from the superposition of cells APs propagated via plasmodesmata linking neighbouring cells (Volkov, 2012; Zimmermann and Mithofer, 2013). According to Debono (2013), such microvolted activity represents the resulting macroscopic currents locally sustained

by plant receptor-channels acting through the different membrane compartments of synchronized proneural networks, which was named “plant electrome” by Souza et al. (2017).

The previous spectral analysis (Souza et al., 2017) showed that before stimuli the runs exhibited three ranges of dominant frequencies (0.3–1.5, 3.8–5.7, and 9.3–11.1 Hz) and, after stimuli, only the lower frequencies remained. The  $\beta$  exponents calculated from power spectrum density function indicated that all time series showed long-range temporal correlation. However,  $\beta$  values from non-stimulated plants were lower, suggesting that the electrical signalling after plant stimulation showed more persistence (for  $\beta \neq 0$  and  $\beta \neq 2$ ), enabling long distance signalling (Souza et al., 2017). Besides some changes in the type of noise (reddened to black) and in the spectral analysis, it was observed an interesting appearance of “bursts of spikes” up to 500  $\mu\text{V}$  (considering the baseline signal around 10  $\mu\text{V}$ ) mostly in the runs scored after stimuli. The distribution of these higher voltage variations was different regarding the specific stimuli. While the spikes observed under low temperature showed an exponential distribution, under low light and, mainly, under osmotic stimulus they followed a power law, indicating that the spikes have no characteristic size (Souza et al., 2017). Those bursts of spikes were supposed as resulting from a self-organized collective behaviour among groups of cells Saraiva et al. (2017). The complex nature of these microvolted potential and their coherent responses to environmental cues suggests that  $H^+$  electrogenic pumps, and others ionic receptors channels (e.g.  $\text{Ca}^{2+}$  and  $\text{K}^{1+}$ ), are involved in such responses (Debono, 2013). Following a membrane electrical event, secondary messengers are initiated, leading to a downstream signalling cascade that induce many different metabolic changes, likely mediated by calcium waves and ROS (Maffei et al., 2007; Gilroy et al., 2014; Choi et al., 2016). Evidences have indicated

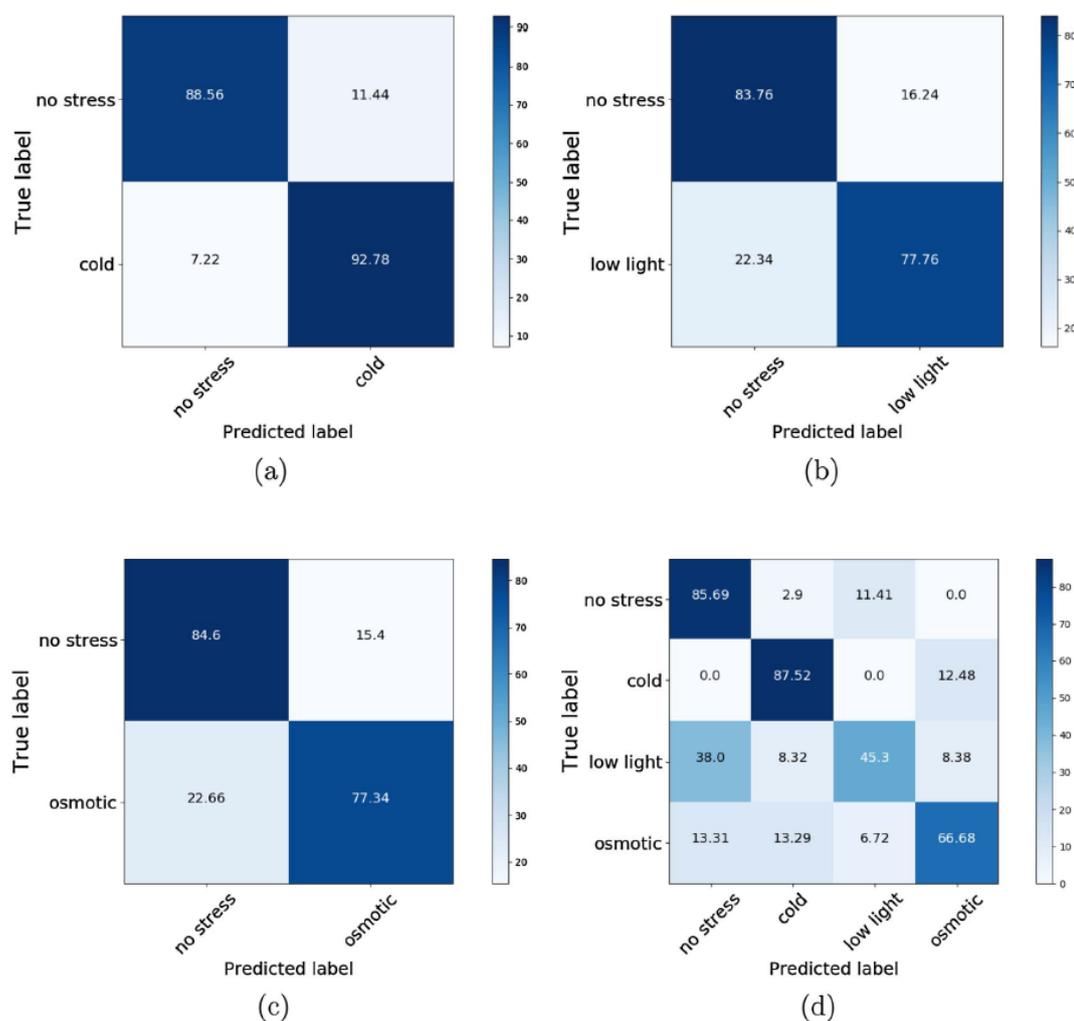


Fig. 6. Confusion matrices considering 90% for training purposes and SVM classifier over: (a) "cold", (b) "low light", (c) "osmotic", and (d) "all" datasets.

that electrical signalling regulates a fundamental physiological process, the well-known photosynthesis. Although the mechanisms of photosynthetic activation are still unclear, two different mechanisms have been proposed for the fast inactivation of photosynthesis. The first is calcium dependent, in which the increase of  $Ca^{2+}$  cytosolic can inactivate the Calvin-Benson cycle; and the second supposes that electrical activity changes the intra- and extracellular pH, decreasing the cytoplasmic pH by inactivation of  $H^+$ -ATPase, which reduces the photochemical yields (Sukhov, 2016).

Accordingly, our hypothesis is that the low voltage variations affected by different abiotic stimuli could be representative of the electrophysiological state of the plants as affected by specific stimulus allowing an algorithmic classification. In this work, we introduced two different methods for the automatic classification of plant stress based on the plant electrome. The first approach makes use of Interval Arithmetic to reduce the dimensionality of the input signal, and the second one employs deep learning for unsupervised feature learning. Chen et al. (2016) have studied the efficiency of some deep learning methods, such as ANN and SVM, comparing with template matching method to classify APs. In that study, after the elimination of some artifacts present in the raw signals, they extracted 19 features from the AP's signals, including time-domain, frequency-domain, statistics

characteristics, and nonlinear features. Then, by reducing the dimensionality with PCA, the different classifiers were tested against non-AP signals. Although the best performance have been reached with template matching algorithm (96.0% accuracy), ANN and SVM reached a maximum accuracy of 84.1% and 75.8%, respectively, which is pretty close to our best results (Table 1). However, differently from Chen et al. (2016) study, our classification algorithms were applied to, basically, continuous time series of raw data of low voltage variation, showing very complex dynamics (Saraiva et al., 2017; Souza et al., 2017). Chatterjee et al. (2015) have explored strategies for the classification of raw non-stationary plant electrical signal after different environmental stimuli (mV) by univariate and bivariate feature-based classification using five different discriminant analysis classifiers. The classification has reached a best average accuracy of 70% using variance and skewness as feature pairs, and an accuracy of 73.67% using variance and IQR as feature pairs in a diagaquadratic classifier. Thus, as far as we know, this is the first time that microvolted runs are classified by machine learning techniques.

## 5. Conclusions

The proposed approaches herein achieved surprisingly good

classification performance on a relatively small set of training data, thus confirming the existence of a clear pattern captured by the plant electrome to identify potential plant stress condition. The experiments also suggested that Interval Arithmetic and supervised classifiers are more suitable than deep learning. Such achievement is very important since the plant stress could be identified before the plant has demonstrated physical injuries, such as leaf fall, and decreased productivity, among other factors. Our future works will be guided to develop better models that can distinguish all three stressful conditions at the same time.

#### Acknowledgments

The authors are grateful to FAPESP grants #2013/07375-0, #2014/16250-9, #2014/12236-1, and #2016/19403-6, as well as CNPq grant #306166/2014-3.

#### Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.compag.2017.12.024>.

#### References

- Almeida, J.G., dos Santos, J.A., Alberton, B., Morellato, L.P.C., Torres, R.S., 2015. Phenological visual rhythms: Compact representations for fine-grained plant species identification. *Pattern Recogn. Lett.* <http://dx.doi.org/10.1016/j.patrec.2015.11.028>.
- Baluska, F., 2016. Long-Distance Systemic Signalling and Communication in Plants. Springer International Publishing, Berlin.
- Baluska, F., Mancuso, S., Volkmann, D., 2006. Springer International Publishing, Berlin.
- Behmann, J., Mahlein, A.K., Rumpf, T., Römer, C., Plümer, L., 2015. A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. *Precision Agric.* 16, 239–260.
- van Bel, A.J.E., Furch, A.C.U., Will, T., Buxa, S.V., Musetti, R., Hafke, J.B., 2014. Spread the news: systemic dissemination and local impact of  $Ca^{2+}$  signals along the phloem pathway. *J. Exp. Bot.* 65, 1761. <http://dx.doi.org/10.1093/jxb/ert425>.
- Chang, C.C., Lin, C.J., 2011. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Chatterjee, S.K., Das, S., Maharatna, K., Masi, E., Santopolo, L., Mancuso, S., Vitaletti, A., 2015. Exploring strategies for classification of external stimuli using statistical features of the plant electrical response. *J. R. Soc. Interface* 12.
- Chatterjee, S.K., Ghosh, S., Das, S., Manzella, V., Vitaletti, A., Masi, E., Santopolo, L., Mancuso, S., Maharatna, K., 2014. Forward and inverse modelling approaches for prediction of light stimulus from electrophysiological response in plants. *Measurement*.
- Chen, Y., Zhao, D.J., Wang, Z.Y., Wang, Z.Y., Tang, G., Huang, L., 2016. Plant electrical signal classification based on waveform similarity. *Algorithms* 9. <http://dx.doi.org/10.3390/a9040070>. <http://www.mdpi.com/1999-4893/9/4/70>.
- Choi, W., Hilleary, R., Swanson, S.J., Kim, S., Gilroy, S., 2016. Rapid, long-distance electrical and calcium signaling in plants. *Annu. Rev. Plant Biol.* 287–307.
- Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educ. Psychol. Measur.* 20, 37–46.
- Coomans, D., Massart, D., 1982. Alternative k-nearest neighbour rules in supervised pattern recognition. *Anal. Chim. Acta* 136, 15–27.
- Cortes, C., Vapnik, V., 1995. Support vector networks. *Machine Learn.* 20, 273–297.
- Davies, E., 2006. *Electrical Signals in Plants: Facts and Hypotheses*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 407–422. [http://dx.doi.org/10.1007/978-3-540-37843-3\\_17](http://dx.doi.org/10.1007/978-3-540-37843-3_17).
- De Loof, A., 2016. The cell self-generating electrome: The biophysical essence of the immaterial dimension of life. *Commun. Integr. Biol.* e1197446.
- Debono, M., 2013. Dynamic proton neural networks in plants: A new approach of spontaneous extracellular potential variations. *Plant Signall. Behav.*
- Falcão, A.X., Stolfi, J., Lotufo, R.A., 2004. The image foresting transform: Theory, algorithms, and applications. *IEEE Trans. Pattern Anal. Machine Intell.* 26, 19–29.
- Fromm, J., Lautner, S., 2007. Electrical signals and their physiological significance in plants. *Plant Cell Environ.* 30, 249.
- Fromm, J., Spanswick, R., 1993. Characteristics of action potentials in willow (*salix viminalis* L.). *J. Exp. Bot.* 1119–1125.
- Gallé, A., Lautner, S., Flexas, J., Fromm, J., 2015. Environmental stimuli and physiological responses: The current view on electrical signalling. *Environ. Exp. Bot.* 114, 15–21. <http://dx.doi.org/10.1016/j.envenpbot.2014.06.013>.
- Hall, P., Park, B.U., Samworth, R.J., 2008. Choice of neighbor order in nearest-neighbor classification. *Ann. Statist.* 36, 2135–2152.
- Haykin, S., 1999. *Neural Networks: A Comprehensive Foundation*, second ed. Prentice Hall.
- Huber, A.E., Bauerle, T.L., 2016. Long-distance plant signaling pathways in response to multiple stressors: the gap in knowledge. *J. Exp. Bot.* 67.
- Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324.
- Lüttge, U., 2012. Modularity and emergence: biology's challenge in understanding life. *Plant Biol.* 14, 865–871. <http://dx.doi.org/10.1111/j.1438-8677.2012.00659.x>.
- Magdalena, S.H., Maria, L., Stanislaw, K., 2017. Electrical signaling, photosynthesis and systemic acquired acclimation. *Front. Physiol.* 8, 684.
- Ma, C., Xin, M., Feldmann, K.A., Wang, X., 2014a. Machine learning-based differential network analysis: A study of stress-responsive transcriptomes in arabidopsis. *Plant Cell Online* 26, 520–537.
- Ma, C., Zhang, H.H., Wang, X., 2014b. Machine learning for big data analytics in plants. *Trends Plant Sci.* 19, 798–808.
- Maffei, E.M., Bossi, S., 2006. *Electrophysiology and Plant Responses to Biotic Stress*. Springer, Berlin.
- Maffei, M.E., Mithofer, A., Boland, W., 2007. Before gene expression: early events in plant-insect interaction. *Trends Plant Sci.*
- Moore, R.E., 1960. Book review: On Numerical Approximation (Proceedings of a Symposium) (R.E. Langer, ed.) 2, 49–50. <http://dx.doi.org/10.1137/1002015>.
- Moore, R.E., 1962. Interval Arithmetic and Automatic Error Analysis in Digital Computing. Ph. D. dissertation. Department of Mathematics, Stanford University, Stanford, CA, USA. Also published as Applied Mathematics and Statistics Laboratories Technical Report No. 25.
- Moore, R.E., 1966. *Interval Analysis*. Prentice-Hall, Englewood Cliff, New Jersey.
- Moore, R.E., 1979. *Methods and Applications of Interval Analysis*. SIAM Studies in Applied Mathematics.
- Moore, R.E., Yang, C.T., 1959. *Interval Analysis I*. Technical Document LMSD-285875. Lockheed Missiles and Space Division, Sunnyvale, CA, USA.
- Nissen, S., 2003. Implementation of a Fast Artificial Neural Network Library (FANN). Department of Computer Science University of Copenhagen (DIKU). Software available at <http://leenissen.dk/fann/>.
- Papa, J.P., Albuquerque, V.H.C., Falcão, A.X., Tavares, J.M.R.S., 2012. Efficient supervised Optimum-Path Forest classification for large datasets. *Pattern Recogn.* 45, 512–520.
- Papa, J.P., Falcão, A.X., Suzuki, C.T.N., 2009. Supervised pattern classification based on Optimum-Path Forest. *Int. J. Imaging Syst. Technol.* 19, 120–131.
- Papa, J.P., Fernandes, S.E.N., Falcão, A.X., 2017. Optimum-path forest based on k-connectivity: Theory and applications. *Pattern Recogn. Lett.* 87, 117–126.
- Pereira, L.A.M., Nakamura, R.Y.M., Souza, G.F.S., Martins, D., Papa, J.P., 2012. Aquatic weed automatic classification using machine learning techniques. *Comput. Electron. Agric.* 87, 56–63.
- Pyatygin, S.S., Opritov, V.A., Vodenev, V.A., 2008. Signaling role of action potential in higher plants. *Russ. J. Plant Physiol.* 285–291.
- Rios-Rojas, L., Tapia, F., Gurovich, L.A., 2014. Electrophysiological assessment of water stress in fruit-bearing woody plants. *J. Plant Physiol.* 171, 799–806. <http://dx.doi.org/10.1016/j.jplph.2014.02.005>. URL: <http://www.sciencedirect.com/science/article/pii/S0176161714000509>.
- Gilroy, S., Suzuki, N., Miller, G., Choi, W.G., Toyota, M., Deviredy, A.R., Mittler, R., 2014. A tidal wave of signals: calcium and ROS at the forefront of rapid systemic signaling. *Trends Plant Sci.*
- Saraiva, G.F.R., Ferreira, A.S., Souza, G.M., 2017. Osmotic stress decreases complexity underlying the electrophysiological dynamic in soybean. *Plant Biol.* 5, 702–708.
- Schölkopf, B., Smola, A.J., 2002. *Learning with Kernels*. MIT Press, Cambridge, MA.
- Shaik, R., Ramakrishna, W., 2014. Machine learning approaches distinguish multiple stress conditions using stress-responsive genes and identify candidate genes for broad resistance in rice. *Plant Physiol.* 164, 481–495.
- Singh, A., Ganapathysubramanian, B., Singh, A.K., Sarkar, S., 2016. Machine learning for high-throughput stress phenotyping in plants. *Trends Plant Sci.* 21, 110–124.
- Souza, G.M., Bertolli, S.C., Lüttge, U., 2016. Hierarchy and Information in a System Approach to Plant Biology: Explaining the Irreducibility in Plant Ecophysiology. *Springer International Publishing, Cham*, pp. 167–186. [http://dx.doi.org/10.1007/978-3-319-25688-7\\_5](http://dx.doi.org/10.1007/978-3-319-25688-7_5).
- Souza, G.M., Ferreira, A.S., Saraiva, G.F.R., Toledo, G.R.A., 2017. Plant electrome can be pushed toward a self-organized critical state by external cues: Evidence from a study with soybean seedlings subject to different environmental conditions. *Plant Signal. Behav.* 12, e1290040.
- Sukhov, V., 2016. Electrical signals as mechanism of photosynthesis regulation in plants. *Photosynth. Res.*
- Sukhov, V., Gaspírovich, V., Mysyagin, S., Vodenev, V., 2017. High-temperature tolerance of photosynthesis can be linked to local electrical responses in leaves of pea. *Front. Physiol.*
- Sukhov, V., Sherstneva, O., Surova, L., Katicheva, L., Vodenev, V., 2014. Proton cellular influx as a probable mechanism of variation potential influence on photosynthesis in pea. *Plant, Cell Environ.* 2532–2541.
- Sukhova, E.R., Akinchits, E., Sukhov, V., 2017. Mathematical models of electrical activity in plants. *J. Membr. Biol.*
- Trevas, A., 2003. Aspects of plant intelligence. *Ann. Bot.* 92, 1. <http://dx.doi.org/10.1093/aob/mcg101>.
- Vapnik, V.N., 1999. An overview of statistical learning theory. *IEEE Trans. Neural Netw.* 10, 988–999.
- Vodenev, V., Akinchits, E., Sukhov, V., 2015. Variation potential in higher plants: Mechanisms of generation and propagation. *Plant Sign. Behav.*
- Volkov, A., 2012. *Plant Electrophysiology: Signaling and Responses*. Plant Electrophysiology. Springer Berlin Heidelberg URL: <https://books.google.com.br/books?id=OjlfvX665EC>.
- Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics Bull.* 1, 80–83.
- Zimmermann, M.R., Maischak, H., Mithofer, A., Boland, W., Felle, H.H., 2009. System potentials, a novel electrical long-distance apoplastic signal in plants, induced by wounding. *Plant Physiol.*
- Zimmermann, M.R., Mithofer, A., 2013. *Long-distance Systemic Signaling and Communication in Plants*. Springer.