



PREVISÃO DE UMIDADE DE SOLOS ATRAVÉS DE REDES NEURAIIS SEM PESO

Paulo Felipe Salviano Brandt

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Civil, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Civil.

Orientadores: Valmir Carneiro Barbosa
Felipe Maia Galvão França

Rio de Janeiro
Setembro de 2015

PREVISÃO DE UMIDADE DE SOLOS ATRAVÉS DE REDES NEURAIS SEM
PESO

Paulo Felipe Salviano Brandt

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA CIVIL.

Examinada por:

Prof. Valmir Carneiro Barbosa, Ph.D.

Prof. Felipe Maia Galvão França, Ph.D.

Prof. Priscila Machado Vieira Lima, Ph.D.

Prof. Marcos Bacis Ceddia, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
SETEMBRO DE 2015

Salviano Brandt, Paulo Felipe

Previsão de Umidade de Solos Através de Redes Neurais Sem Peso/Paulo Felipe Salviano Brandt. – Rio de Janeiro: UFRJ/COPPE, 2015.

XV, 102 p.: il.; 29, 7cm.

Orientadores: Valmir Carneiro Barbosa

Felipe Maia Galvão França

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Civil, 2015.

Referências Bibliográficas: p. 101 – 102.

1. Rede Neural Sem Peso. 2. Umidade. 3. Solo. 4. Geotécnica. I. Carneiro Barbosa, Valmir *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Civil. III. Título.

Agradecimentos

Inicialmente agradeço a minha Mãe Eny Góes Salviano, por todo apoio, carinho e atenção, por estar sempre comigo nos momentos mais difíceis tornando mais fácil se sacrificando por mim, a ela eu devo não só este trabalho como também grande parte do meu sucesso que conquistei até agora e ainda almejo conquistar.

Agradeço a minha Irmã, Carolina Maria Salviano Brandt por todo amor e zelo que tem por mim e com toda certeza sem ela em minha vida tudo seria um pouco menos colorido.

Agradeço ao meu Filho, Daniel Karonte de Albuquerque Salviano Brandt por ser o norte que me guia quando tudo parece escurecer, basta lembrar dele para tudo fazer mais sentido.

Agradeço a minha Noiva, Tatiana Eucário de Araújo por estar ao meu lado nos momentos conturbados me levantando e mostrando que sou capaz, mais do que eu penso e acredito.

Agradeço à Mãe do meu filho e Amiga, Tamires de Albuquerque dos Santos por se doar completamente sendo a mãe zelosa e carinhosa para o nosso filho me deixando despreocupado com o seu bem estar, e por ser a amiga de sempre.

Agradeço à minha Tia Solange Honório Passos, por ter o amor de mãe que tem por mim e me fazer sentir o primeiro filho dela me ajudando sempre que pode e algumas vezes até mesmo quando não pode.

Agradeço ao meu Primo, André Luiz de Souza Salviano pelas várias trocas de ideias, companheirismo de irmão, abrigo e todo apoio que pôde me dar durante o período do mestrado.

Agradeço ao grande amigo da família Eduardo Paiva, que plantou uma semente antes mesmo de iniciar esta luta, me apresentando ao CEDERJ.

Agradeço á todos os meus familiares e espero que todos se sintam mesmo agradecidos, pois me sinto agraciado por ter tantos parentes e poder me sentir bem ao lado de todos sem exceção.

Agradeço aos Professores e Orientadores, Valmir Carneiro Barbosa e Felipe Maia Galvão França por ter acreditado em mim desde a iniciação científica e mesmo quando pouco mereci tal confiança ainda estavam lá para me ajudar e orientar no que fosse preciso para o meu sucesso.

Agradeço ao Mario Guilherme Garcia Nacinovic por toda paciência, ajuda e por me apresentar o mundo fantástico da Geotecnia.

Agradeço à todos os colegas de laboratório que tornaram esta jornada muito mais divertida e produtiva, em especial ao Douglas Cardoso grande amigo que fiz neste período com todo apoio técnico e científico dado, Hugo Carneiro com toda sua eloquência abrilhantando ainda mais o laboratório, Saulo Oliveira outro amigo que se mostrou ser um grande companheiro, a dupla Daniel Alves e Daniel Nunes dois pólos dentro do laboratório que traziam momentos únicos, Danilo e João Antônio outros dois colegas que sempre serão lembrados mesmo estando tão longe em distância.

Agradeço à UFRJ, PESC, Geotecnia e CNPq por todo suporte dado durante a pesquisa a mim e aos meus orientadores.

Por fim agradeço a mim, por não ter desistido mesmo quando tudo parecia tão distante de concretizar, por ter respirado fundo e decidido seguir em frente lembrando de tudo que já havia passado.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

PREVISÃO DE UMIDADE DE SOLOS ATRAVÉS DE REDES NEURAI SEM PESO

Paulo Felipe Salviano Brandt

Setembro/2015

Orientadores: Valmir Carneiro Barbosa
Felipe Maia Galvão França

Programa: Engenharia Civil

A umidade é um atributo importante do solo a ser definido e descoberto dentro da Geotécnica e Agronomia. Conhecer este atributo faz parte de muitas pesquisas importantes envolvendo erosão, deslizamentos de encostas e muitos outros estudos e pesquisas agronômicas. Este trabalho se dedica a estudar a relação entre pluviosidade e umidade do solo gerando um sistema de previsão para esta umidade utilizando redes neurais sem peso, mais especificamente o modelo WiSARD.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

PREDICTION OF SOILS MOISTURE BY WEIGHTLESS NEURAL
NETWORK

Paulo Felipe Salviano Brandt

September/2015

Advisors: Valmir Carneiro Barbosa
Felipe Maia Galvão França

Department: Civil Engineering

Humidity is an important attribute of the soil to be defined and discovered within the Geotechnical and Agronomy. Understand this attribute is very important to research involving erosion, lands slopes and other studies involving agricultural research. This work dedicated to study the relationship between rainfall and soil moisture, generating a system for preview this moisture using a wheightless neural networks, more specifically a WiSARD model.

Sumário

Lista de Figuras	xi
Lista de Tabelas	xv
1 Introdução	1
1.1 Motivação	1
1.2 Objetivo	2
1.3 Metodologia	3
1.4 Organização da Dissertação	4
2 Estudo do Solo	5
2.1 Parcelas Sensoriadas	5
2.2 Estrutura do Solo	6
2.3 Dinâmica da Umidificação do Solo	6
2.4 Aquisição dos Dados	7
2.5 TDR e Pluviômetro	8
3 Redes Neurais	9
3.1 Redes Neurais Sem Peso	10
3.2 WiSARD	11
4 Tratamento dos Dados	17
4.1 Dados do TDR	17
4.2 Dados do Pluviômetro	18
4.3 Atributos	18
4.3.1 Hora do Dia	18
4.3.2 Temperatura	19
4.3.3 Período de Seca	20
4.4 Padrão I30	20
5 Sistema de Previsão de Umidade	22
5.1 Janela de Entrada	29

5.2	Previsor de Umidade WiSARD	30
5.3	Quantidade e Tamanho das RAMs	30
5.4	Discriminadores	31
5.4.1	Divisão em seis Classes	32
5.4.2	Divisão Dinâmica das Classes	33
5.5	Treinamento e Testes	34
5.6	Validação Cruzada	34
6	Resultados	36
6.1	Métricas	36
6.2	Previsão com Classes Dinâmicas	38
6.3	Variações nas Escalas do Termômetro	39
6.3.1	Priorizando Hora do Dia	39
6.3.2	Priorizando Temperatura	42
6.3.3	Priorizando Período de Seca	45
6.3.4	Balanceando Dados de Entrada	47
6.3.5	Descarte de Atributos	48
6.4	Variações no Tamanho das RAMs	48
6.4.1	RAMs com 16 bits	49
6.4.2	RAMs com 32 bits	56
6.4.3	RAMs com 48 bits	62
6.4.4	RAMs com 64 bits	69
6.5	Variações no Tamanho das Janelas	76
6.5.1	Janela de Cinco Minutos	77
6.5.2	Janela de Trinta Minutos	85
6.5.3	Janela de Duas Horas	92
7	Conclusões e Trabalhos Futuros	98
7.1	Trabalhos Futuros	99
7.1.1	Sistema de Predição de Deslizamentos em Encostas	99
7.1.2	Previsão de Erosão	99
7.1.3	Inclusão de Novos Atributos	100
	Referências Bibliográficas	101

Lista de Figuras

2.1	Perfil da Parcela de Pastagem e Localização dos TDRs	6
3.1	Neurônio Biológico	9
3.2	Modelo McCulloch e Pitts	10
3.3	Amostras de Setas para Esquerda	13
3.4	Amostras de Setas para Direita	13
3.5	Amostras, Mapeamento e Discriminadores	15
3.6	Classificação	16
4.1	Teor de Umidade e Temperatura no Quadrante Superior na Mata . .	19
4.2	Janela Deslizante de Dados para compor o <i>input</i>	21
5.1	Fluxograma do Sistema de Previsão de Umidade	22
5.2	Teor de Umidade e Temperatura Quadrante Superior x Inferior com precipitação na Pastagem	23
5.3	Teor de Umidade e Temperatura Quadrante Superior x Inferior sem precipitação na Pastagem	24
5.4	Teor de Umidade e Temperatura Quadrante Superior x Inferior com precipitação na Mata	25
5.5	Teor de Umidade e Temperatura Quadrante Superior x Inferior sem precipitação na Mata	26
6.1	Alto coeficiente de Confiança com Baixo acerto da Previsão	38
6.2	Hora do Dia: Quadrante Superior Esquerdo na Pastagem A	40
6.3	Hora do Dia: Quadrante Superior Esquerdo na Pastagem B	41
6.4	Hora do Dia: Quadrante Superior Esquerdo na Mata A	41
6.5	Hora do Dia: Quadrante Superior Esquerdo na Mata B	42
6.6	Temperatura: Quadrante Superior Esquerdo na Pastagem A	43
6.7	Temperatura: Quadrante Superior Esquerdo na Pastagem B	43
6.8	Temperatura: Quadrante Superior Esquerdo na Mata A	44
6.9	Temperatura: Quadrante Superior Esquerdo na Mata B	44
6.10	Período de Seca: Quadrante Superior Esquerdo na Pastagem A . . .	45

6.11	Período de Seca: Quadrante Superior Esquerdo na Pastagem B	46
6.12	Período de Seca: Quadrante Superior Esquerdo na Mata A	46
6.13	Período de Seca: Quadrante Superior Esquerdo na Mata B	47
6.14	Gráfico A Quadrante Superior Esquerdo na Mata (RAM de 16 Bits) .	50
6.15	Gráfico B do Quadrante Superior Esquerdo na Mata (RAM de 16 Bits)	50
6.16	Gráfico A Quadrante Inferior Esquerdo na Mata (RAM de 16 Bits) .	51
6.17	Gráfico B do Quadrante Inferior Esquerdo na Mata (RAM de 16 Bits)	51
6.18	Gráfico A Quadrante Superior Esquerdo na Pastagem (RAM de 16 Bits)	52
6.19	Gráfico B do Quadrante Superior Esquerdo na Pastagem (RAM de 16 Bits)	53
6.20	Gráfico A Quadrante Inferior Esquerdo na Pastagem (RAM de 16 Bits)	54
6.21	Gráfico B do Quadrante Inferior Esquerdo na Pastagem (RAM de 16 Bits)	55
6.22	Gráfico A Quadrante Superior Esquerdo na Mata (RAM de 32 Bits) .	56
6.23	Gráfico B do Quadrante Superior Esquerdo na Mata (RAM de 32 Bits)	57
6.24	Gráfico A Quadrante Inferior Esquerdo na Mata (RAM de 32 Bits) .	57
6.25	Gráfico B do Quadrante Inferior Esquerdo na Mata (RAM de 32 Bits)	58
6.26	Gráfico A Quadrante Superior Esquerdo na Pastagem (RAM de 32 Bits)	59
6.27	Gráfico B do Quadrante Superior Esquerdo na Pastagem (RAM de 32 Bits)	60
6.28	Gráfico A Quadrante Inferior Esquerdo na Pastagem (RAM de 32 Bits)	61
6.29	Gráfico B do Quadrante Inferior Esquerdo na Pastagem (RAM de 32 Bits)	62
6.30	Gráfico A Quadrante Superior Esquerdo na Mata (RAM de 48 Bits) .	63
6.31	Gráfico B do Quadrante Superior Esquerdo na Mata (RAM de 48 Bits)	64
6.32	Gráfico A Quadrante Inferior Esquerdo na Mata (RAM de 48 Bits) .	64
6.33	Gráfico B do Quadrante Inferior Esquerdo na Mata (RAM de 48 Bits)	65
6.34	Gráfico A Quadrante Superior Esquerdo na Pastagem (RAM de 48 Bits)	66
6.35	Gráfico B do Quadrante Superior Esquerdo na Pastagem (RAM de 48 Bits)	67
6.36	Gráfico A Quadrante Inferior Esquerdo na Pastagem (RAM de 48 Bits)	68
6.37	Gráfico B do Quadrante Inferior Esquerdo na Pastagem (RAM de 48 Bits)	69
6.38	Gráfico A Quadrante Superior Esquerdo na Mata (RAM de 64 Bits) .	70
6.39	Gráfico B do Quadrante Superior Esquerdo na Mata (RAM de 64 Bits)	71
6.40	Gráfico A Quadrante Inferior Esquerdo na Mata (RAM de 64 Bits) .	71

6.41	Gráfico B do Quadrante Inferior Esquerdo na Mata (RAM de 64 Bits)	72
6.42	Gráfico A Quadrante Superior Esquerdo na Pastagem (RAM de 64 Bits)	73
6.43	Gráfico B do Quadrante Superior Esquerdo na Pastagem (RAM de 64 Bits)	74
6.44	Gráfico A Quadrante Inferior Esquerdo na Pastagem (RAM de 64 Bits)	75
6.45	Gráfico B do Quadrante Inferior Esquerdo na Pastagem (RAM de 64 Bits)	76
6.46	Gráfico A Quadrante Superior Esquerdo na Mata (Janela 5 Minutos)	78
6.47	Gráfico B do Quadrante Superior Esquerdo na Mata (Janela 5 Minutos)	79
6.48	Gráfico A Quadrante Inferior Esquerdo na Mata (Janela 5 Minutos) .	79
6.49	Gráfico B do Quadrante Inferior Esquerdo na Mata (Janela 5 Minutos)	80
6.50	Gráfico A Quadrante Superior Esquerdo na Pastagem (Janela 5 Minutos)	81
6.51	Gráfico B do Quadrante Superior Esquerdo na Pastagem (Janela 5 Minutos)	82
6.52	Gráfico A Quadrante Inferior Esquerdo na Pastagem (Janela 5 Minutos)	83
6.53	Gráfico B do Quadrante Inferior Esquerdo na Pastagem (Janela 5 Minutos)	84
6.54	Gráfico A Quadrante Superior Esquerdo na Mata (Janela 30 Minutos)	85
6.55	Gráfico B do Quadrante Superior Esquerdo na Mata (Janela 30 Minutos)	86
6.56	Gráfico A Quadrante Inferior Esquerdo na Mata (Janela 30 Minutos)	87
6.57	Gráfico B do Quadrante Inferior Esquerdo na Mata (Janela 30 Minutos)	87
6.58	Gráfico A Quadrante Superior Esquerdo na Pastagem (Janela 30 Minutos)	88
6.59	Gráfico B do Quadrante Superior Esquerdo na Pastagem (Janela 30 Minutos)	89
6.60	Gráfico A Quadrante Inferior Esquerdo na Pastagem (Janela 30 Minutos)	90
6.61	Gráfico B do Quadrante Inferior Esquerdo na Pastagem (Janela 30 Minutos)	91
6.62	Gráfico A Quadrante Superior Esquerdo na Mata (Janela 2 Horas) . .	92
6.63	Gráfico B do Quadrante Superior Esquerdo na Mata (Janela 2 Horas)	93
6.64	Gráfico A Quadrante Inferior Esquerdo na Mata (Janela 2 Horas) . .	93
6.65	Gráfico B do Quadrante Inferior Esquerdo na Mata (Janela 2 Horas)	94
6.66	Gráfico A Quadrante Superior Esquerdo na Pastagem (Janela 2 Horas)	94
6.67	Gráfico B do Quadrante Superior Esquerdo na Pastagem (Janela 2 Horas)	95

6.68 Gráfico A Quadrante Inferior Esquerdo na Pastagem (Janela 2 Horas)	96
6.69 Gráfico B do Quadrante Inferior Esquerdo na Pastagem (Janela 2 Horas)	97

Lista de Tabelas

5.1	Conversão binária pelo Método Matemático e Distância de Hamming para o valor 5 (com 10 bits)	28
5.2	Conversão binária pelo Método do Termômetro e Distância de Hamming para o valor 5 (com 10 bits)	29
5.3	Quantificadores Linguísticos e intervalos de permissividade.	33
6.1	Resultados para RAMs com 16 Bits	49
6.2	Resultados para RAMs com 32 Bits	56
6.3	Resultados para RAMs com 48 Bits	63
6.4	Resultados para RAMs com 64 Bits	70
6.5	Resultados para Janelas de Cinco Minutos	78
6.6	Resultados para Janelas de Trinta Minutos	85
6.7	Resultados para Janelas de Duas Horas	92

Capítulo 1

Introdução

Avaliar o impacto no solo provocado por chuvas intensas ou contínuas se mostram de grande importância. Observando eventos catastróficos que ocorrem com frequência não somente no Rio de Janeiro como em várias cidades do Brasil e até mesmo no mundo, não é difícil pensar em possibilidades de soluções e viabilizá-las para minimizar estes danos. Eventos de intensa precipitação pluvial são corriqueiros no estado do Rio de Janeiro durante o período do verão, principalmente em localidades como a Região Serrana do Rio e Angra dos Reis, o que gera uma sensação de urgência na criação de meios que possam ajudar a minimizar danos provocados por estes eventos.

Este trabalho é abrangido por duas áreas, Geotécnica e Computação. A primeira trazendo todo o estudo do solo, da dinâmica hídrica e um importante conjunto de dados adquiridos com o monitoramento em duas parcelas de terra com coberturas diferentes localizadas na cidade de Itaboraí. A Computação contribuiu com o emprego de Redes Neurais Sem Peso e a modelagem da informação contida no *Data Set* para atingir os melhores resultados possíveis.

A união destas duas áreas neste trabalho abre um grande leque de oportunidades para estudos que envolve o solo, a umidade contida nele e inclusive futuros estudos que não se limitam apenas à previsões de umidade em um solo mas também mostram os primeiros passos de um possível sistema que possa prever com exatidão deslizamentos de encostas.

1.1 Motivação

Na Agronomia, a pesquisa de melhorias e qualidade da irrigação de plantações de forma adequada e eficaz é levado em consideração visando a redução de custos e melhor uso de recursos hídricos. O solo deve ser estudado e constantemente

monitorado, quando o teor de umidade estiver abaixo dos níveis adequados para o cultivo acionará os sistemas de irrigação e desligar caso a umidade esteja no padrão desejado, porém, perdas podem acontecer e desta forma gerar custos que podem ser reduzidos utilizando-se sistemas inteligentes para ligar e desligar a irrigação.

Geralmente utilizam-se Tensiômetros para averiguar a umidade no solo. Todavia para este método é preciso calibrar com frequência o equipamento. Caso contrário acontecerão erros de leituras e falhas ao ligar ou desligar os sistemas de irrigação ocasionando perdas sejam elas recursos hídricos ou na própria safra.[1]

Além da irrigação artificial das plantações, existe a possibilidade de precipitação nos campos, diminuindo os custos do uso dos sistemas artificiais, Faz-se necessário prever o quão úmido o solo ficará e se será necessário ou o quanto se fará necessário irrigar o campo. Pensar em formas de irrigar as plantações prevendo a umidade do solo torna-se importante sendo um dos pontos motivadores para o presente trabalho.

Outra motivação importante para esta pesquisa se refere às perdas sociais no estado do Rio de Janeiro envolvendo deslizamentos de encostas como os visto na Região Serrana do Rio e em Angra dos Reis, podendo se estender a quaisquer outras localidades. As chuvas frequentes no Rio de Janeiro são noticiadas ano após ano e sempre temos perdas sociais, sejam pessoas desabrigadas ou no pior dos casos vidas perdidas com os deslizamentos.

Quando ocorrem precipitações, dependendo de sua intensidade o solo pode rapidamente saturar, com o solo saturado a resistência ao cisalhamento diminui podendo ocasionar o desmoronamento[2]. Prever o teor de umidade de um solo específico dado um índice pluviométrico pode ajudar num futuro sistema de evacuação para as áreas em risco. Ainda que esta dissertação não conceba tal sistema de alarme, certamente pode levar a um futuro sistema deste tipo.

1.2 Objetivo

Para descobrir a necessidade de ligar um sistema de irrigação ou criar um sistema de alarme para desmoronamento é preciso estudar um gatilho eficiente para estes sistemas, ambos se tratam de umidade no solo ou **Teor de Umidade**. Averiguar os teores de umidade dado um índice pluviométrico previsto, ou real no caso de alimentação do sistema, pode ser um gatilho adequado.

Este trabalho objetiva um sistema de previsão do teor de umidade do solo, aprofundando-se nos sistemas de Redes Neurais Sem Peso. Criando um histórico com base no conhecimento adquirido pela Rede Neural a partir do conjunto de dados do solo que se deseja monitorar, utilizando os alguns métodos de aquisição destes dados.

1.3 Metodologia

Neste trabalho utilizou-se um Pluviômetro e TDRs, sondas utilizadas para determinar teor de umidade. Utilizou-se destes equipamentos, Pluviômetro e TDR, para obter dados de índice pluviométrico e umidade do solo respectivamente. No decorrer do trabalho serão detalhados estes métodos para um melhor entendimento do leitor. Estes equipamentos foram instalados em Itaboraí no Estado do Rio de Janeiro.

O Experimento foi feito em duas parcelas com coberturas diferentes, ambos de $1000m^2$. Numa das parcelas a cobertura vegetal rasteira, basicamente toda em gramíneas com inclinação média de 39% foram instalados oito sensores TDR sendo quatro deles a 10cm de profundidade e os outros quatro a 30cm de profundidade. Apesar de ter sido escavada uma trincheira para a instalação dos sensores o solo manteve sua estrutura intacta.

A outra parcela se trata de uma zona de mata com cobertura densa de árvores com pouca vegetação rasteira e inclinações entre 9% e 39%. Com a mesma distribuição dos sensores TDR dispostos na parcela anterior. Pluviômetros localizados em Itaboraí foram utilizados, assim como pluviômetros específicos para as parcelas, este método foi utilizado para determinar a precipitação no solo e gerar um histórico formado por teor de umidade e índice pluviométrico. Ambos alimentando diariamente um banco de dados a cada cinco minutos com as medidas, num período de cinco meses.

Para o sistema de previsão utilizou-se uma Rede Neural Sem Peso WiSARD a qual fora alimentada com os dados históricos para sua convergência e formação de base de conhecimento, e para previsão foram inseridos dados do índice pluviométrico em mm^3 que podem ser determinados por qualquer sistema meteorológico. É possível e se faz necessário alimentar o sistema com dados reais de precipitação para gerar histórico na base de conhecimento da rede ajudando a mesma a não tornar-se defasada.

1.4 Organização da Dissertação

No capítulo dois do presente trabalho é apresentado um estudo sobre o solo e das duas parcelas que tiveram dados recolhidos em Itaboraí. Além de fazer breves considerações a respeito da umidificação do solo e sobre aparelhos que foram utilizados para aquisição dos dados. O capítulo três é referente ao estudo das Redes Neurais Artificiais mais precisamente a versão em peso destas redes detalhando especificamente o modelo WiSARD.

Em seguida é apresentado o quarto capítulo do trabalho onde é tratado a modelagem dos dados contidos no *Data Set*, como foram as escolhas dos atributos mais importantes deste conjunto. O capítulo cinco descreve o funcionamento do sistema proposto neste trabalho, um sistema de previsão de umidade no solo, além de mostrar como cada atributo pode influenciar nos resultados.

Os resultados serão apresentados no sexto capítulo, esmiuçando as possibilidades de modificações dos atributos da Rede Neural assim como do *input* produzido para esta a partir da modelagem do conjunto de dados. Finalmente é apresentada a conclusão do trabalho no capítulo sete e trabalhos futuros mostrando as possibilidades abertas por este estudo.

Capítulo 2

Estudo do Solo

Este trabalho se dá em uma união entre duas áreas de estudo, Geotécnica e Computação por intermédio da Inteligência Artificial. Foi importante além de analisar os dados, entender a importância do sistema **Pluviosidade** x Umidade do Solo. Apesar de inicialmente clara relação entre os dois vetores, existem muitos fatores ocultos que podem interferir diretamente, e por este motivo se fez necessário um capítulo sobre o Estudo do Solo neste trabalho. Este capítulo contará com o apoio da Tese do Doutor Mário Nacinovic como referência para algumas informações técnicas do solo, medidas tomadas em campo e laboratório.

2.1 Parcelas Sensoriadas

Os dados utilizados neste trabalho são frutos do monitoramento de duas parcelas de terra na cidade de Itaboraí, ambas tem o mesmo tamanho físico, com uma área de $1000m^2$ ($40m \times 25m$), porém com característica diferente no que diz respeito ao perfil do solo e à cobertura do solo. Uma das áreas é coberta com mata nativa com grande variedade de espécies, nesta área foi detalhado três perfis de solo que podem ser vistos em [3] e falaremos em poucos detalhes neste trabalho. Na área de pastagem a cobertura é basicamente de vegetação rasteira e capim podendo chegar até $1m$ de altura, possuindo ainda mudas de reflorestamento na ordem de 1660 mudas / h com apenas um perfil de solo detalhado.[3]

Foram instalados oito sondas TDR divididos em duas profundidades do solo, a $10cm$ e a $30cm$, foi instalado também um Pluviômetro no local, todos os dados do Pluviômetro e de cada sonda TDR são armazenados em intervalos de cinco minutos em um *Data Log* instalado também em campo. Sobre as sondas TDR e Pluviômetros serão abordados mais adiante neste trabalho.

2.2 Estrutura do Solo

Este trabalho teve colaboração do Doutor Mario Nacinovic, toda a parte que diz respeito à análise do solo contou com a perícia do supracitado Doutor e para qualquer informação técnica sobre o solo poderá ser conferido em sua Tese, no presente trabalho apenas será informado alguns aspectos a título de informação, todos os dados aqui verificados foram retirados da Tese e devidamente referenciado.

A parcela de mata foi dividida em três partes porém diferenciando somente em altitude e inclinação média. No terço superior desta parcela a inclinação média é de 9%, aproximadamente 5° com altitude de $55m$ do nível do mar. O terço intermediário a altitude é de $35m$ com inclinação média de 20%, 11° . No terço inferior com $20m$ de altitude e inclinação de 39% ou 21° . todos as três partes desta parcela tem a classificação de Latossolo Amarelo Distrocoeso típico com textura muito argilosa, relevo ondulado e acentuadamente drenado. A parcela de pastagem tem apenas um perfil de solo anotado, com altitude média de $25m$ inclinação de 39% ou 21° , sua classificação é de Latossolo Amarelo Distrófico típico, textura argilosa, relevo forte ondulado e acentuadamente drenado.[3]

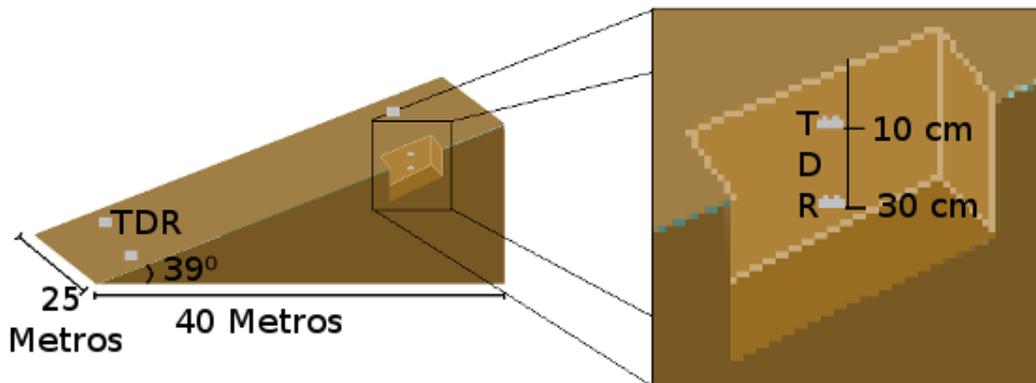


Figura 2.1: Perfil da Parcela de Pastagem e Localização dos TDRs

2.3 Dinâmica da Umidificação do Solo

Muitos fatores interferem na umidificação do solo, desde a temperatura da precipitação, textura do solo, sua inclinação, sua cobertura e até mesmo fatores biológicos como formigueiros.[3][4][5][6] As variáveis são tantas que tornam mensura de dados do solo um trabalho a ser feito utilizando equipamentos tais

quais, Tensiômetros ou mesmo o TDR, para a previsão da umidade dado um índice pluviométrico este trabalho se tornaria ainda mais árduo, levando em consideração que existe ainda a permeabilidade do solo fazendo com que a água leve minutos e até mesmo horas para atingir determinadas camadas do solo [7], como exemplo neste trabalho foram verificados em inúmeras vezes a alteração da umidade no TDR localizado a 30cm duas horas após os eventos de chuva terem iniciado.

A importância de verificar estas discrepâncias no tempo se dá pelo fato de que a precipitação ocorre em um determinado horário e a percepção da alteração da umidade no medidor só se dará algum tempo depois tornando o sistema dessincronizado, mostrando a importância de uma janela de tempo de informação maior que o período em que a água leva para infiltrar até a posição do TDR. Adiante serão informados testes que mostram a evolução do acerto nas previsões do sistema proposto durante este trabalho, em relação ao tamanho da janela de tempo utilizada.

Existe ainda a intensidade da precipitação medida em 30 minutos que é conhecido no meio Geotécnico como Padrão I30 [8]. Esta mensura é importante ser avaliada quando a intensidade da precipitação é muito grande nem toda a água será infiltrada, formará um pequeno espelho de água ocorrendo o escoamento, ou seja, mesmo com um volume de água muito grande num solo muito seco se este volume precipitar num período muito curto de tempo o aumento da umidade verificada no solo pode ser menor do que num período em que a precipitação foi bem inferior mas que ocorreu durante o dia inteiro. Por este motivo esta medida se faz necessária e entra como informação de forma indireta no sistema de janelas de tempo e atributo o `textbf`Período de Seca, ambos serão melhores detalhados mais adiante.

2.4 Aquisição dos Dados

Os dados dos dois terrenos foram recolhidos utilizando uma sonda TDR e Pluviômetros, os sensores de TDR foram instalados em duas profundidades e em quatro locais diferentes distribuídos no terreno. Trincheiras foram abertas para a instalação das sondas a 10cm e 30cm de profundidade. A cada cinco minutos os dados destes sensores são recolhidos e computados pelo *Data log*, arquivando a Permissividade do solo dentre outros dados importantes para o sistema tais quais Temperatura do solo, Data e Hora do recolhimento dos dados e índice pluviométrico recolhido pelo Pluviômetro instalado e conectado ao mesmo *Data log*. A aquisição destes dados foi ininterrupta no período monitorado, não foi constatado presença de ruídos que pudessem comprometer o treinamento e os resultados das previsões

da rede neural.

2.5 TDR e Pluviômetro

O TDR, acrônimo para *Time Domain Reflectometry*, em português Reflectometria no Domínio do Tempo, é um equipamento que tem seu uso aceito para avaliar a umidade do solo com boa precisão obtendo leituras variadas. O funcionamento do equipamento baseia-se na velocidade de propagação das ondas eletromagnéticas em suas hastes condutoras que ficam enterradas no solo, medindo a constante dielétrica aparente Ka do solo e usando como base as constantes dielétricas do ar, do solo e da água estima-se a umidade volumétrica do solo [9], porém seu uso depende de calibrações e seu custo alto o torna um equipamento de uso restrito a pequenas parcelas de terra [10]. O TDR não é a única forma de detectar o teor de água no solo, existem vários outros métodos para estas medições, todos com suas vantagens e desvantagens tendo assim a necessidade de escolher o método mais adequado a ser aplicado.[3][11]

Os dados que foram cedidos a este trabalho pelo Mario Nacinovic da Geotecnia/COPPE, foram tomados a partir de um TDR e por esse motivo aqui se focará apenas neste método para a aquisição dos dados e classificação da umidade no solo no que diz respeito à Previsão de Umidade. Serão utilizadas estas classes atribuídas de forma dinâmica para este fim, o que será abordado no capítulo cinco deste trabalho.

O Pluviômetro é um equipamento simples utilizado para medir a quantidade de chuva precipitada em dado local num período de tempo, o mesmo pode ser utilizado para outros tipos de precipitações, porém, não vem ao caso neste trabalho. Este equipamento mede a chuva em milímetros por minuto, o funcionamento do mesmo é simples e aceito no meio acadêmico, baseia-se em coletar água num aparato em forma de gangorra que suporta dois pequenos receptáculos com uma medida pré-definida em volume de água em cada um no valor de $0,256mm^3$, quando o receptáculo é preenchido, o mesmo se vira entornando o seu conteúdo e marcando acionando um contador em seus circuitos, enquanto o outro receptáculo voltado para cima agora está coletando a precipitação do ambiente.[3] Por ser um equipamento simples e de fácil manuseio sem a necessidade de grandes conhecimentos técnicos para fazê-lo funcionar, é um equipamento largamente utilizado e de custo muito baixo, podendo ser distribuído em grande quantidade em áreas diversas.

Capítulo 3

Redes Neurais

Warren McCulloch e Walter Pitts, um Neurologista e um Matemático respectivamente, pesquisaram e criaram as primeiras redes neurais na década de 40, iniciando a ideia de Redes Neurais no campo da Inteligência Artificial. Seus modelos consistiam em criar neurônios artificiais que se assemelhassem ao neurônio natural e suas funcionalidades, este modelo matemático de neurônio artificial apresenta capacidade de aprendizado através de treinamentos semelhante a um neurônio natural sendo um modelo matemático simples. O treinamento é um processo que deve ser feito em várias etapas até o resultado convergir ao que se espera, este processo pode levar um tempo considerável.[12]

As redes neurais artificiais propostas por McCulloch e Pitts, utilizam pesos que são ajustados a cada iteração durante o treinamento, uma função é atribuída na camada oculta da rede e quando as saídas adequadas ao resultados são obtidas tem-se a convergência da rede.

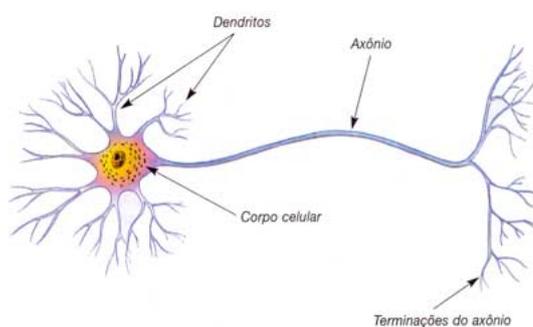


Figura 3.1: Neurônio Biológico

A Figura 3.1 ilustra um neurônio biológico. Neste o impulso elétrico ocorre da direita para esquerda, a célula é estimulada em seus Dendritos e o estímulo viaja

pelo axônio seguindo para suas terminações excitando outras células neurais através das sinapses, as ramificações são tratadas como memória.[12]

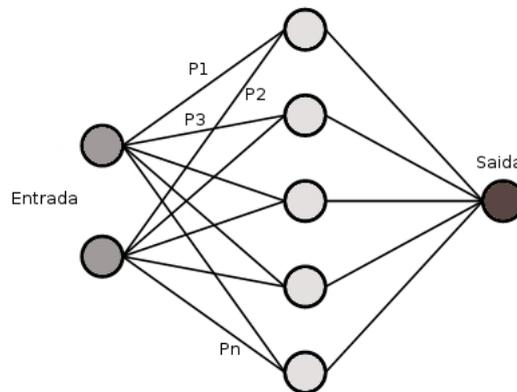


Figura 3.2: Modelo McCulloch e Pitts

A Figura 3.2 ilustra o modelo de McCulloch e Pitts. Neste modelo o fluxo de informação ocorre da esquerda para a direita, as linhas marcadas de $P1$ até Pn são os pesos representando o aprendizado da rede, apenas estes valores são armazenados, a seguir tem-se as funções e na parte final os valores de saída para a rede.

3.1 Redes Neurais Sem Peso

Existe um outro tipo de Rede Neural que não utiliza os pesos ajustáveis, este tipo de rede utiliza memória de acesso randômico (RAM - *Randon Acess Memory*) onde são armazenadas pequenas parcelas de informações das entradas, para o treinamento desta rede um processo direto de escrita é realizado em cada RAM definindo um padrão de entrada com seu rótulo e obtendo convergência imediata, significando que a rede não precisará ser realimentada como no modelo Neural de McCulloch e Pitts, após o treinamento, a rede estará pronta para ser testada e avaliada, uma nova entrada sem rótulo poderá ser fornecida e conforme os acessos ao espaço de memória nas RAMs são verificados como acerto tem-se um somatório destes e esta nova entrada terá o rótulo da RAM com maior escore de acertos [13].

Como definido anteriormente, este será o tipo de rede utilizada para o presente trabalho, mais precisamente o modelo WiSARD o qual será detalhado e melhor definido adiante, a escolha foi motivada pela agilidade no aprendizado visto que o sistema de previsão de umidade precisa ser em tempo real, como os dados chegam dos sensores a cada cinco minutos o tempo de aprendizado e de resposta da rede

precisa ser otimizado.

O sistema também é fortemente baseado num histórico de eventos, cada cinco minutos de dados coletados é tratado como um evento, é importante cada período ser tratado como um evento mesmo que não ocorra precipitação, veremos mais adiante que o período de seca tem impacto para a precisão do sistema, visto isso e assumindo a importância de cada evento de cinco minutos, podemos pensar em como o sistema se manterá atualizado, para manter sua precisão, o sistema terá de ser alimentado com dados do pluviômetro e com suas próprias previsões de umidade, é necessário que se use métodos ágeis para sua manutenção e execução contínua e este é outro forte motivo para escolha das RNSPs para este trabalho.

3.2 WiSARD

Este modelo de RNSP tem como característica principal o uso de endereçamento binário e Discriminadores, como uma RNSP a WiSARD (*Wilkie Stonhan and Aleksander's Recognition Device*) tem o mesmo princípio de guardar informações em RAMs, Porém aqui teremos conjuntos de RAMs ou mesmo Neurônios Artificiais que serão escritos através de um endereçamento aleatório gerando um mapeamento para acessos destas escritas, a este conjunto de RAMs dar-se-á o nome de Discriminadores, o qual receberá um rótulo que servirá para classificar futuras entradas [13].

Na fase de treinamento desta rede os dados de entrada previamente rotulados devem ser em valores binários, esta entrada será dividida e espalhada num mapeamento aleatório sendo distribuída pelos n neurônios da rede, para cada rótulo diferente no conjunto de dados a ser usado para o treinamento será criado um Discriminador com seu conjunto específico de neurônios, o espalhamento dos dados binários será agrupado e transformado em **endereço** para acesso de uma posição do neurônio mapeado, nesta posição será escrito o valor 1 e todas as outras posições terão valores 0, note que podemos ter tamanhos variados para os neurônios e quantidades variadas de neurônios em cada Discriminador, as quantidades e o tamanho de cada neurônio está diretamente ligada ao tamanho da entrada, é importante ter amostras variadas de cada classe (rótulo) para que a rede possa melhor reconhecer instâncias que se diferem umas das outras em alguns aspectos porém podem pertencem à mesma classe.

Para conjunto de dados desbalanceados é preciso uma modelagem nos dados

e mesmo em atributos da WiSARD para contornar estes problemas que são encontrados em conjuntos de dados naturais, veremos mais adiante que este é o caso a ser abordado neste trabalho, os dados reunidos com o monitoramento das parcelas na cidade de Itaboraí se caracterizam por possuir muitas classes e algumas delas em proporções tão pequenas que reduzem drasticamente os resultados das previsões da WiSARD.

Prosseguindo com o treinamento da rede, logo após seu término, poderá passar pelos testes, estes serão feitos com entradas do mesmo padrão das usadas durante o treinamento da rede, a entrada será submetida à mesma divisão e será mantido o mapeamento do treinamento para a fase de teste, porém aqui a operação não será de escrita, será um acesso direto em cada neurônio, e para cada valor 1 encontrado neste acesso será somado ao final como um acerto, este processo se repetirá para cada neurônio de cada Discriminador, o Discriminador que possuir o maior escore neste somatório será o escolhido para rotular a entrada, esta é a **Medida de Similaridade** que rege o processo de Classificação da rede, para este trabalho o processo de Classificação será conhecido como Previsão.

Abaixo um conjunto de figuras para melhor explicar estes processos de Treinamento e Classificação, neste conjunto será treinado uma pequena RNSP WiSARD para reconhecer Setas Direcionais a partir de uma imagem, uma pequena grade com 25 *pixels* possíveis de serem preenchidos, cada *pixels* será representado por um espaço no vetor denominado *input* ou *In1*, *In2* e *In3*, neste pequeno experimento será apresentado à rede seis amostras de setas, três setas rotuladas como **Direita** e outras três como **Esquerda**.

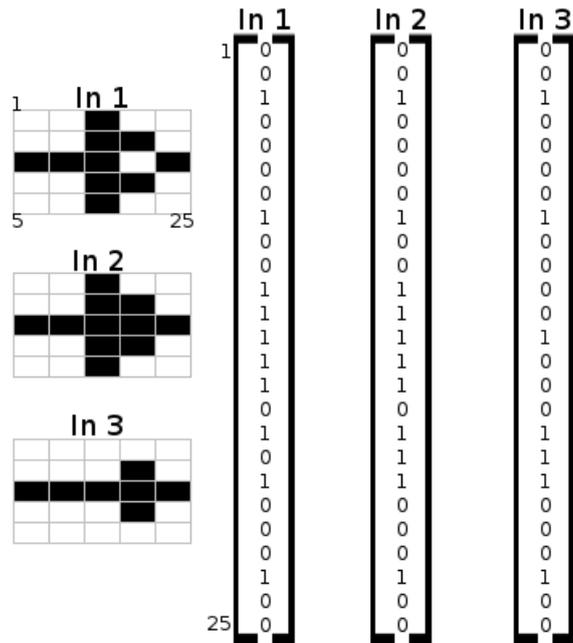


Figura 3.3: Amostras de Setas para Esquerda

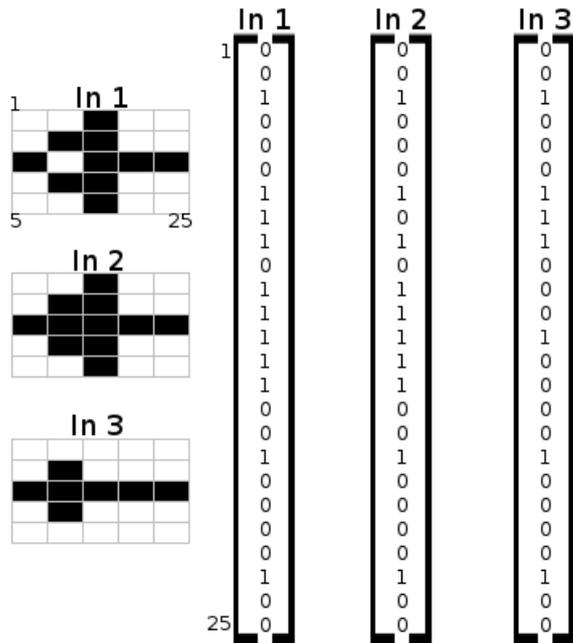
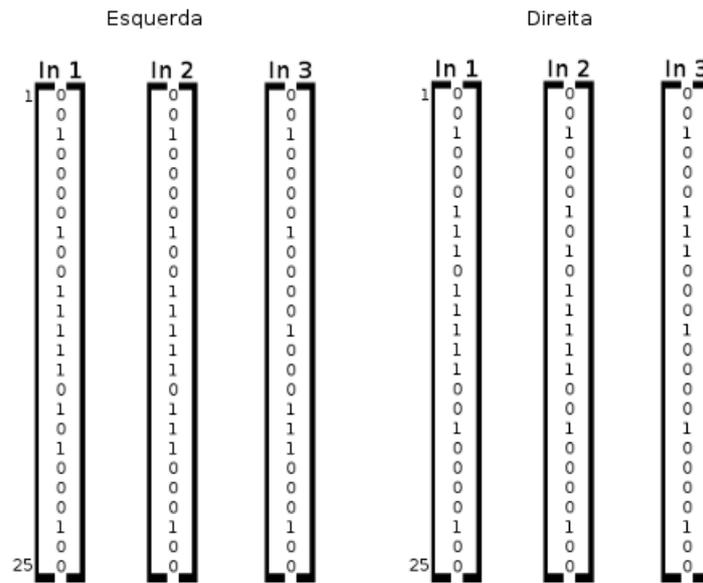


Figura 3.4: Amostras de Setas para Direita

Note que foi convertido a posição do pixel para binário, um *pixel* preenchido em cor preta será equivalente ao valor binário 1 do contrário será 0, com os vetores preenchidos submete-se estes à rede, esta WiSARD terá apenas dois Discriminadores, um para cada conjunto de setas, e cada Discriminador possuirá 9 RAMs de 8 bits, como a grade comporta $25 \text{ pixels} \frac{25}{9} = 2,78$, como não se pode trabalhar com endereços fracionados arredonda-se para 3 bits de endereço, podendo alcançar os 8 bits de cada RAMs, foi escolhido esta quantidade de RAMs para simplificar e

melhor explicar o modelo apresentado, note também que faltarão dois bits a serem preenchidos $9 \times 3 = 27$, estes dois bits podem ser preenchidos com 0 ou 1 sem perda na precisão da WiSARD, neste exemplo estes bits excedentes foram preenchidos com 0.

O **Mapeamento** proposto no exemplo foi simplificado para melhor entendimento, geralmente é feito de forma aleatória como citado anteriormente, aqui foi escolhido de forma sequencial nos vetores alternando nas RAMs, p_0 , p_9 , p_{19} por exemplo, significa que o primeiro bit deste mapeamento é a posição 0 do vetor de entrada, para o segundo bit será a posição 9 e no terceiro bit a posição 18, seguindo esta mesma regra para os endereçamentos seguintes.



RAMs - Neurônios

Mapeamento	000	001	010	011	100	101	110	111
p0-p9-p18	1	1	0	0	0	0	0	0
p1-p10-p19	1	0	1	0	0	0	0	0
p2-p11-p20	0	0	0	0	1	0	1	0
p3-p12-p21	0	0	1	0	0	0	0	0
p4-p13-p22	0	1	0	1	0	0	0	0
p5-p14-p23	1	0	1	0	0	0	0	0
p6-p15-p24	1	0	0	0	1	0	0	0
p7-p16-p25	1	0	0	0	1	0	1	0
p8-p17-p26	1	0	1	0	0	0	1	0

Legenda
 Exclusivo Esquerda
 Exclusivo Direita
 Ambos

Figura 3.5: Amostras, Mapeamento e Discriminadores

Na Figura 3.5, as amostras preenchem a memória de cada discriminador, mais uma vez para simplificar foi feito apenas um conjunto com 9 RAMs, utilizou-se um esquema de coloração para criar o efeito de dois Discriminadores, quando um endereço de memória for acessado numa escrita no Discriminador rotulado como **Esquerda** o bit será apresentado na RAM com a cor **Azul**, para escrita no Discriminador rotulado **Direita** a cor será **Verde** e para o caso do espaço da RAM ser acessado por **Ambos** a cor será **Vermelha**.

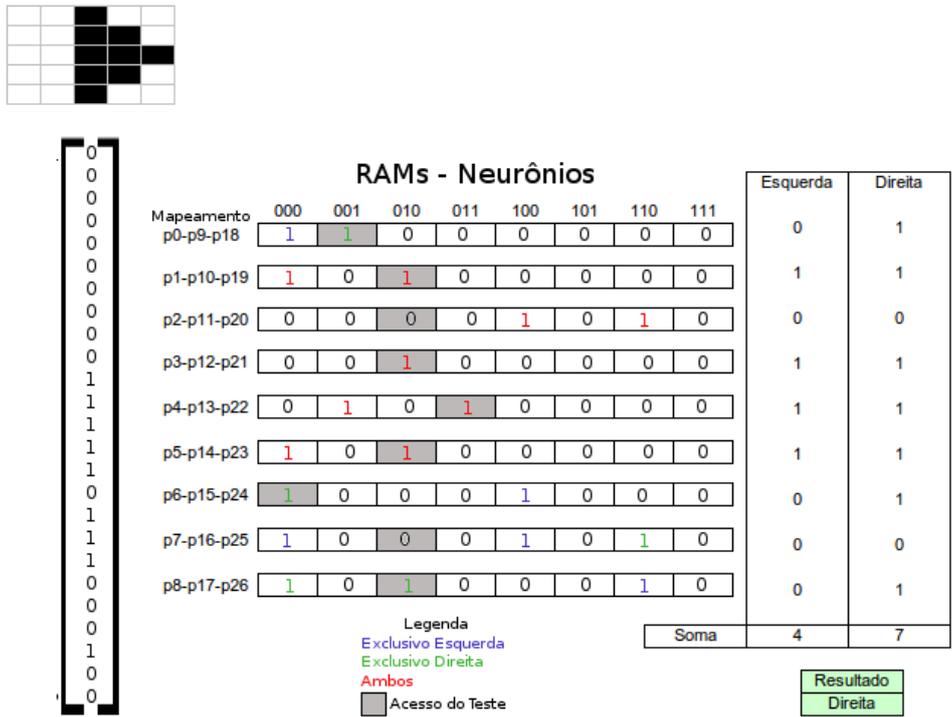


Figura 3.6: Classificação

Após o treinamento, pode-se apresentar uma nova amostra diferente das utilizadas na fase de treino para rede, neste momento se fará um processo de busca e não mais de escrita na WiSARD, a Figura acima ilustra esta nova amostra e seu respectivo vetor de entrada, as RAMs com suas bases de conhecimento preenchidas serão acessadas usando o mesmo **Mapeamento** proposto anteriormente, os acessos foram ilustrados com o preenchimento em cinza do espaço de memória endereçado, e para cada acesso atingido no respectivo Discriminador foi somado 1 ponto, ao final os escores foram somados, verificou-se 4 pontos para o Discriminador rotulado de **Esquerda** e 7 pontos para o discriminador rotulado de **Direita**, assim o resultado para a amostra apresentada foi **Direita**.

Capítulo 4

Tratamento dos Dados

Este capítulo abordará como os dados são tratados após a sua aquisição nos receptores e sensores das terrenos e suas parcelas. Será abordado também como foram feitas as escolhas de cada atributo para mais adiante poder entender qual o nível de correlação de cada um deles na previsão de umidade do solo.

4.1 Dados do TDR

Os dados do TDR recolhidos e armazenados no *Data log* são transmitidos pelos sensores instalados no solo, o *Data log* organiza-os em formato de uma tabela unindo aos coletados a partir do Pluviômetro, esta tabela de dados possui vinte e sete atributos dispostos nas colunas e os valores respectivos de cada atributo são dispostos nas linhas em intervalos de cinco minutos, este intervalo é configurável, porém, no presente trabalho foi utilizado desta forma por conveniência no que diz respeito a espaço em disco.

Dos vinte e sete dados apenas a voltagem da bateria que mantém os sensores ligados foi descartada para fim de treinamento na rede, os outros são distribuídos em Timestamp contabilizando a data e hora em que a leitura foi feita, valores de VWC (*Volumetric Water Content*), Permissividade e Temperatura de cada um dos oito sensores TDR totalizando vinte e quatro entradas, a ultima entrada de dados se refere ao índice pluviométrico, estão separados os dados dos diferentes terrenos monitorados, foram treinadas duas redes neurais distintas para cada tipo de terreno, visto que as características intrínsecas de cada localidade está expressa entremeada na base de conhecimento de cada rede.

4.2 Dados do Pluviômetro

O Pluviômetro é um equipamento simples, mede o índice pluviométrico da região monitorada transmitindo os dados para o *Data log* ligado ao TDR, computando no intervalo configurado o índice pluviométrico no momento em que os dados do solo são verificados, apesar de ser um equipamento simples é amplamente utilizado, existe um pluviômetro em cada terreno monitorando a precipitação local.

Foi necessário um pluviômetro em cada terreno para melhor treinar a rede, os momentos de precipitação são diferentes na ordem de minutos e até mesmo horas o que prejudicaria a rede neural a convergir de forma coerente visto que os valores de teor de umidade colhidos pelos sensores do TDR estaria com tempo defasado em relação a um pluviômetro localizado muito distante do terreno monitorado, apesar dos índices pluviométricos serem praticamente os mesmos.

Para fins de previsão no sistema, pode-se inserir dados de índices pluviométricos previstos por uma estação meteorológica que o sistema fará a inferência retornando a umidade prevista para os valores inseridos, porém será necessário a entrada de dados reais de precipitação para manter a base de conhecimento da rede neural.

4.3 Atributos

O *data log* registra e formata uma tabela com vinte e sete atributos, os quais foram exaustivamente testados buscando a melhor utilização dos mesmos para o sistema de previsão, notou-se a necessidade de não utilizar alguns atributos tal qual a voltagem da bateria que mantém os sensores TDR ligados e o VWC que é o conteúdo volumétrico de água no solo, este valor é o mesmo que a Permissividade, poderia ser escolhido qualquer um dos valores, porém a escala dos valores de Permissividade são mais próximas da Temperatura, que é um dos mais importantes atributo para o sistema, além dos atributos gerado pelos sensores TDR e pelo Pluviômetro, outro atributo foi gerado a partir dos existentes, foi o Período de Seca, atributo este que se tornou muito importante para acurácia da rede.

4.3.1 Hora do Dia

Este é um atributo que em principio seria descartado por se tratar apenas de um rótulo do *Data log*, porém tornou-se muito necessário para tratar a ambiguidade da temperatura do solo nos períodos noturnos e nos períodos diurnos

chuvosos, nos quais as temperaturas do solo se assemelhavam mesmo sendo a umidade do solo diferente para ambos os períodos, gerando assim incoerências nas observações. Testes para entender esta ambiguidade serão mostrados mais adiante evidenciando a importância deste atributo para a previsão gerada pelo sistema.

4.3.2 Temperatura

A temperatura é um importante atributo com alto grau de correlação com a umidade do solo, como podemos ver em [14]. A água que se aloja nos micropóros e nos macropóros após os eventos de precipitação [4] [15], evaporam-se com as altas da temperatura, em um solo com pouca cobertura ocorrerá naturalmente com os raios solares incidindo diretamente no solo. Quando cobertos com Mata, ocorre o aumento da temperatura do solo em menor escala, como vistos no gráfico ilustrado na figura 4.1, porém ainda ocorre a diminuição da umidade em decorrência da Evapotranspiração. A própria vegetação se encarrega de absorver a umidade do solo para seus processos fotossintéticos.[16][17]

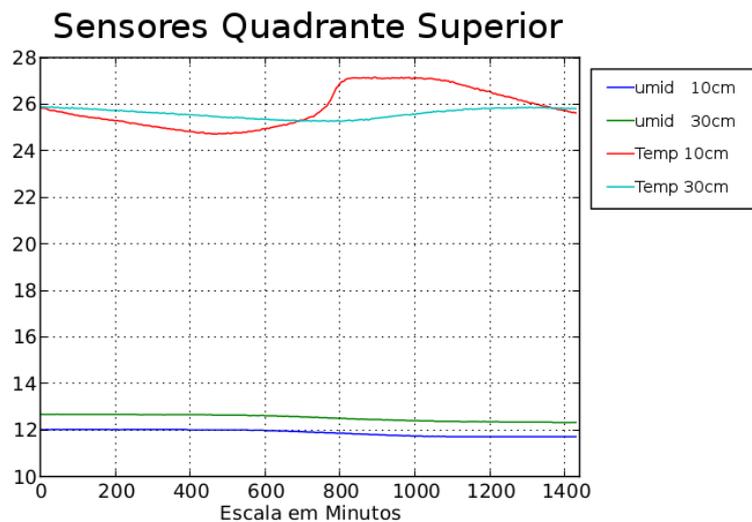


Figura 4.1: Teor de Umidade e Temperatura no Quadrante Superior na Mata

A figura 4.1 ilustra um gráfico onde é apresentado o desenvolvimento diário da Umidade a 10cm e 30cm de profundidade, linhas azul e verde respectivamente, valores estes medidos em Permissividade, que é o índice que este trabalho propoes-se a prever, com este valor será possível descobrir o volume de água no solo sensoriado. Ainda no gráfico, é possível ver o desenvolvimento no mesmo período de tempo das temperaturas, em graus célsius, que se encontram o solo nas profundidades de 10cm e 30cm, linhas vermelha e ciano respectivamente.

4.3.3 Período de Seca

Verificou-se a importância de criar um atributo com a soma do período em que não ocorrem eventos de precipitações, os dados coletados não possuíam este atributo diretamente em sua tabela, porém foi possível obter estes dados de forma indireta, verificando o período de não precipitação e quantificando este novo atributo, a este atributo deu-se o nome de Período de Seca.

Para quantificação destes valores foi um estudo a parte, como o espaço de tempo de recolhimento dos dados no *Data log* é de cinco minutos, como já especificado nos capítulos anteriores, os valores deste atributo ficavam significativamente destoantes do restante dos dados, e os períodos de seca são variados, podendo passar meses sem chuva o que resultaria num valor muito grande em minutos, com inúmeros experimentos chegou-se a um valor que melhor representa este atributo.

Uma observação importante a este atributo, ele é responsável para uma melhor previsão da umidade do solo, gerando acertos mais precisos para as faixas das classes de solo úmido e molhado além das faixas de solo seco, isto acontece principalmente para janelas de tempo maiores que trinta minutos, como este atributo tem correlação direta com o índice pluviométrico e chuvas intensas em espaços curtos de tempo tendem a escoar pela superfície antes mesmo de infiltrar e chegar à impactar na umidade do solo atingindo os sensores TDRs, podemos notar que este comportamento fará parte do aprendizado da rede neural tornando-se parte de sua base de conhecimento, eventos similares serão entendidos futuramente como fazendo parte da mesma classe como é de se esperar.

4.4 Padrão I30

Tomou-se por base o uso deste padrão e os primeiros testes focou o uso de uma janela de eventos em trinta minutos, toda a programação para modelagem dos dados foi baseada neste padrão porém, deixou-se aberto a uma nova configuração para testes, após os primeiros resultados com a janela de trinta minutos, foram feitos inúmeros testes com outras janelas, desde janelas menores com no mínimo cinco minutos já que este é o tempo mínimo de recolhimento dos dados do TDR, até janelas maiores, o limite superior é dado por poder computacional, com o aumento da janela de dados aumenta o tamanho em bits de cada entrada aumentando assim o custo computacional para o processamento dos dados.[8]

Para compor um *input* com este padrão, e qualquer outra janela maior que os cinco minutos de dados coletados, são utilizados o conjunto de dados para o tempo desejado, neste caso seis partes de cinco minutos de informação recolhida pelo `textitdata` log para poder prever a umidade dos cinco minutos seguintes. Após realizada a previsão, para prever os cinco minutos seguintes todo o procedimento é realizado novamente e a janela de dados 'Deslizará' cobrindo outros trinta minutos, a leitura atual dos sensores adicionados de cinco leituras utilizadas na previsão anterior, este processo está ilustrado na figura abaixo.

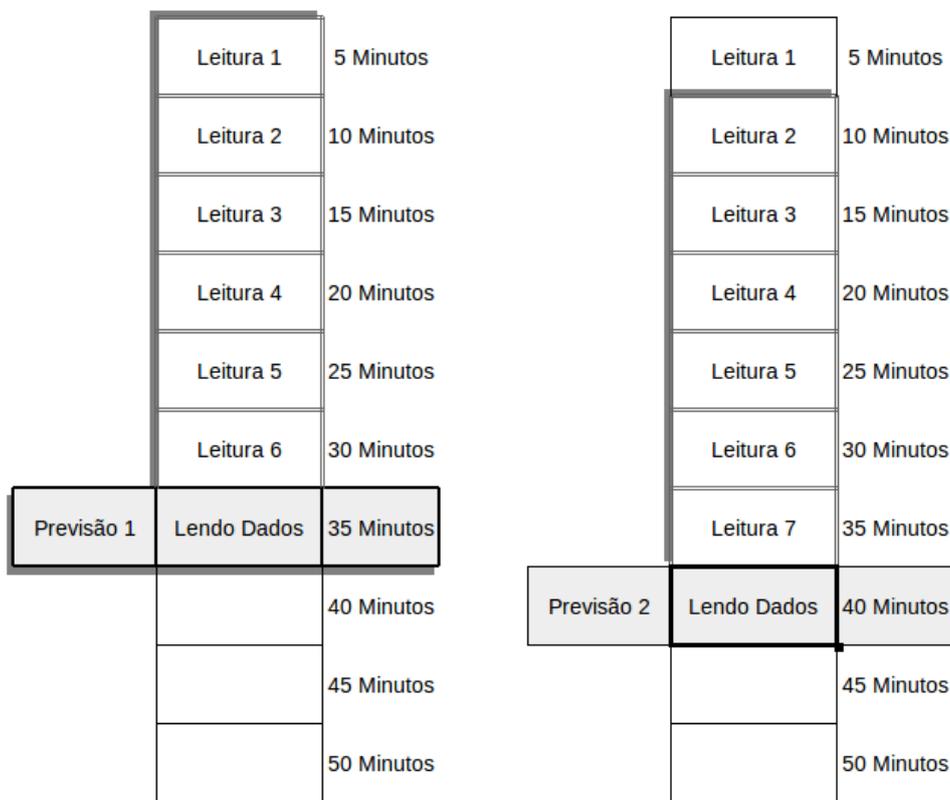


Figura 4.2: Janela Deslizante de Dados para compor o *input*

Capítulo 5

Sistema de Previsão de Umidade

Este trabalho tem como um dos seus objetivos produzir um sistema para prever o teor de umidade do solo dado um índice pluviométrico, algumas etapas se fazem necessárias antes de obter-se a resposta desejada, o fluxograma na Figura 5.1 ilustra as etapas e suas funções serão detalhadas no decorrer deste capítulo.

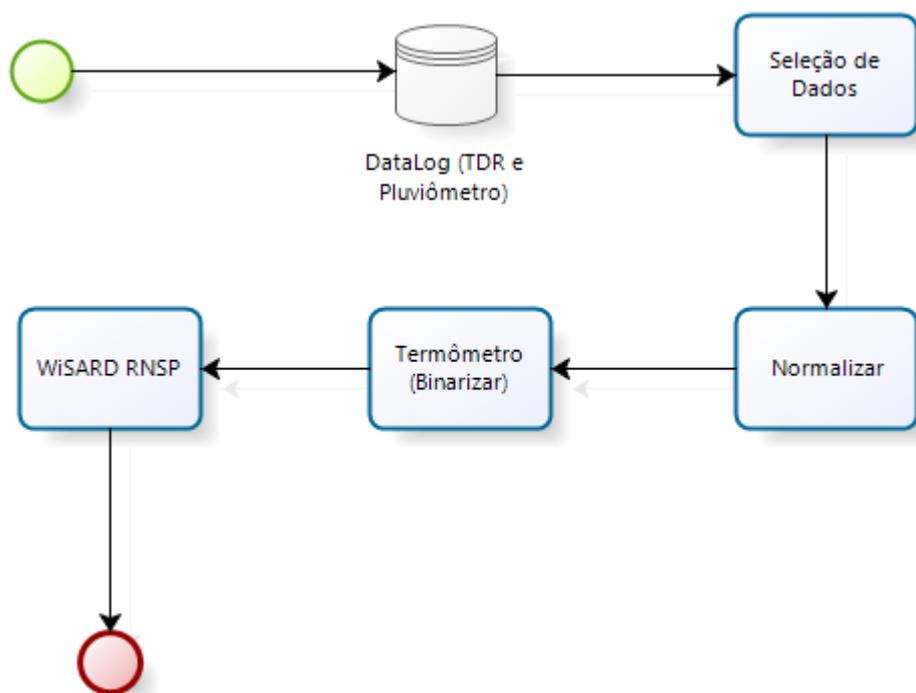


Figura 5.1: Fluxograma do Sistema de Previsão de Umidade

O TDR gera uma grande quantidade de dados com atributos que servirão para o treinamento da rede, existem oito sensores em cada parcela de terreno monitorado, onde quatro estão a dez centímetros da superfície e outras quatro

a trinta centímetros da superfície, existe também a localização geográfica destes sensores onde quatro estão no topo do declive e outras quatro em sua base, esta informação se faz relevante pelo fato que os sensores tendem a medir valores diferentes dado esta posição geográfica medindo maior teor de umidade nos sensores localizados na base, outra característica relevante é que a umidade nos sensores situados à 10cm de profundidade localizado no topo do terreno geralmente é maior do que os sensores localizados à 30cm nesta, ocorrendo o inverso nos sensores localizados na base do terreno como pode-se observar nas Figuras 5.2 e 5.3.

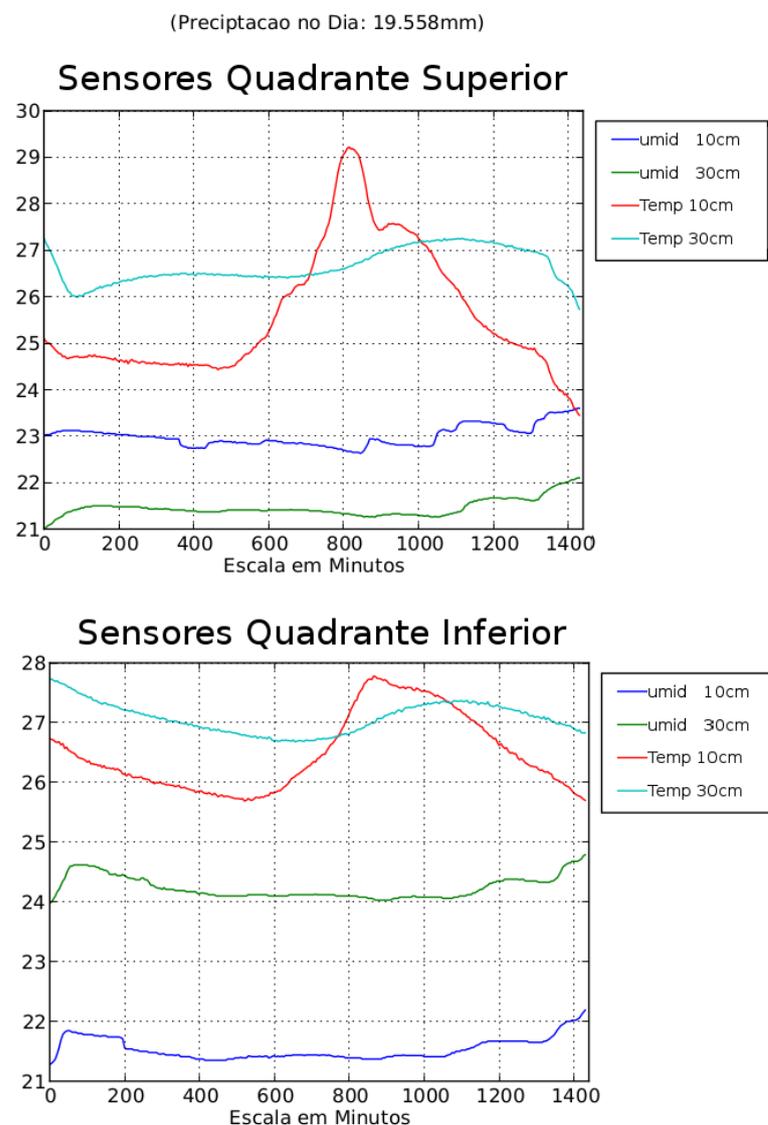


Figura 5.2: Teor de Umidade e Temperatura Quadrante Superior x Inferior com precipitação na Pastagem

(Precipitacao no Dia: 0.0mm)

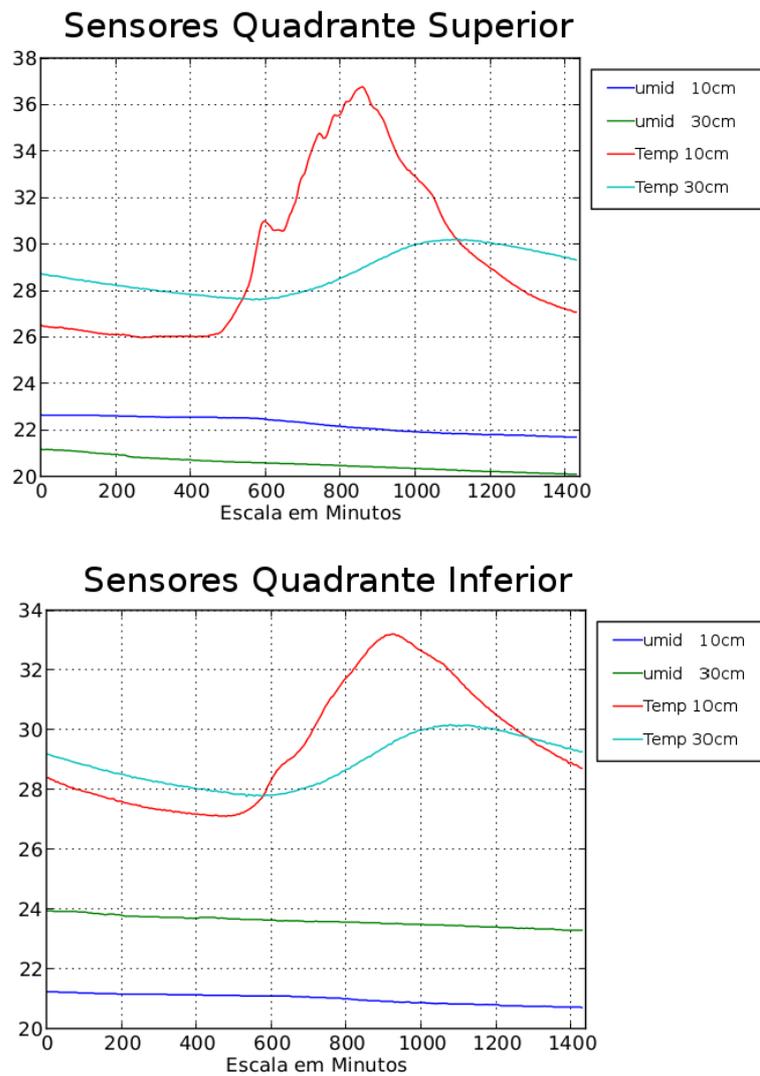


Figura 5.3: Teor de Umidade e Temperatura Quadrante Superior x Inferior sem precipitação na Pastagem

Este comportamento geralmente se repete na mata também como podemos ver nos gráficos ilustrados pelas Figuras 5.4, mas existe algumas exceções como a ilustrada na figura 5.5, neste caso existe quase um empate nas duas profundidades em decorrência de um longo período de seca, aproximadamente 27 dias sem chuva para este período.

(Precipitacao no Dia: 2.54mm)

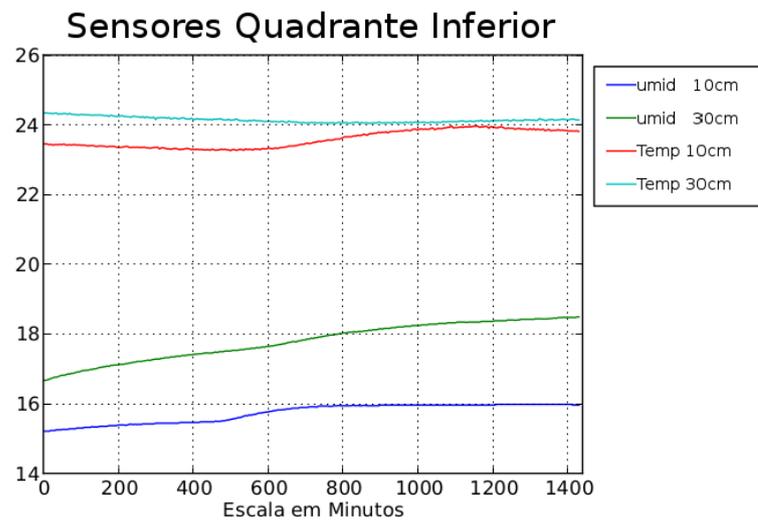
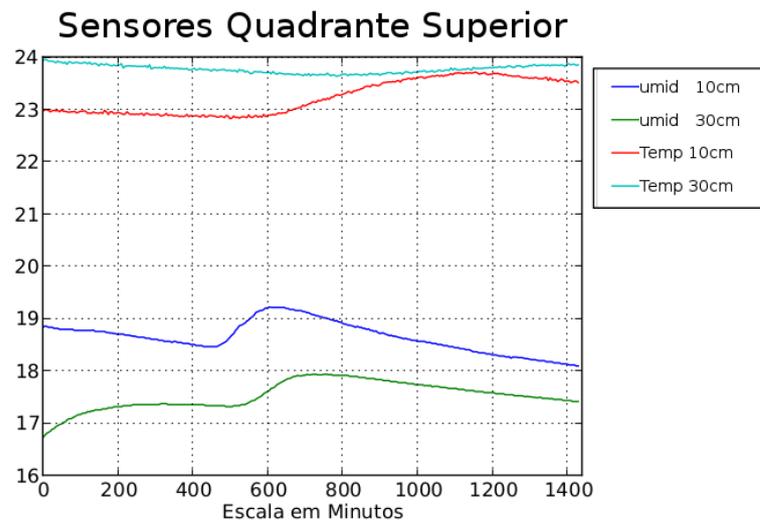


Figura 5.4: Teor de Umidade e Temperatura Quadrante Superior x Inferior com precipitação na Mata

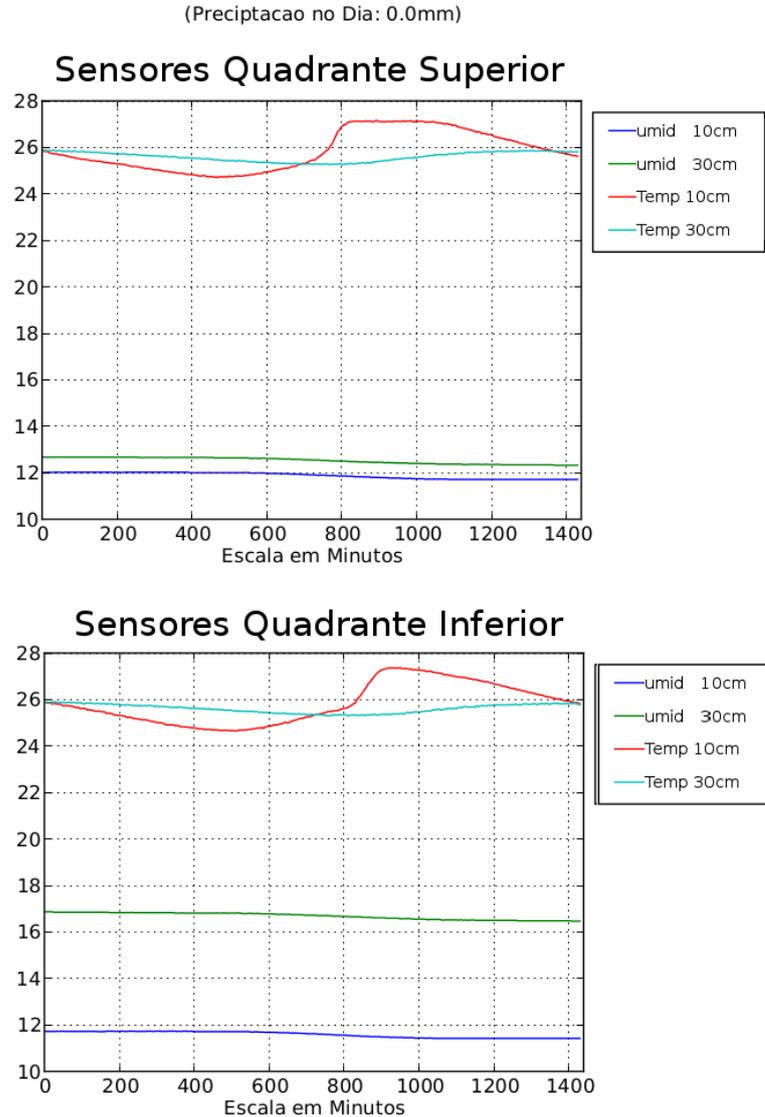


Figura 5.5: Teor de Umidade e Temperatura Quadrante Superior x Inferior sem precipitação na Mata

Isso se dá pelo fato do escoamento da água precipitada ocorrer antes da infiltração, e mesmo em dias em que não ocorrem precipitação este quadro se repete por conta da umidade prévia, a água escoada tende a infiltrar na base do talude aumentando assim a umidade desta região, mantendo as partes mais profundas mais úmidas que as demais regiões, isto evidencia a importância do histórico das medições para uma melhor preditividade da WiSARD.

Esta quantidade de dados tomada pelo TDR consiste na primeira etapa do fluxograma, a retirada de dados do TDR durante este trabalho não foi feita em tempo real, os dados eram retirados em campo na cidade de Itaboraí e entregue

em mãos na UFRJ sem se valer de sistemas *online*, porém, para um futuro treinamento da rede em outras áreas é perfeitamente viável o uso de sistemas deste tipo.

Os dados recolhidos do TDR em geral contém alguns atributos desnecessários para o sistema, desta forma a segunda etapa ilustrada no fluxograma, é preciso fazer a seleção criteriosa dos dados importantes para o treinamento da rede WiSARD, assim como visto no Capítulo 4 deste trabalho, os dados selecionados foram: Permissividade, Temperatura e o Horário que os mesmo são coletados, além dos dados do TDR é selecionado o único dado gerado pelo Pluviômetro, o Índice Pluviométrico.

Com os dados selecionados o sistema ainda não poderá utilizá-los, estes valores estão em escalas diferentes, onde a temperatura pode chegar a $40^{\circ}C$ e o índice pluviométrico não passar de valores entre 0.256 e 6.144, estes valores precisam ter suas escalas harmonizadas e esta será a terceira etapa do sistema que é a normalização, que consiste em enquadrar os valores dos atributos entre zero e um, para normalizar os valores utilizou-se a seguinte formula:

$$\text{Input}_n = \left\{ \frac{v - \text{min}}{\text{max} - \text{min}} \ ; \text{ onde } v \in \text{Input} \right\}$$

Sendo v o valor a ser normalizado, min e max são os valores de mínimos e máximos respectivamente encontrado no atributo.

Com a normalização os valores dos atributos estarão no intervalo contínuo entre 0 e 1, porém estes ainda não podem ser usados pela WiSARD, como visto no capítulo três a rede precisa receber os *inputs* em valores binários, neste momento a quarta etapa do processo transformará estes valores em binários, podendo-se convertê-los utilizando a sua representação na base binária, todavia para avaliar o que acontece quando transforma-se um numero decimal em binário no que diz respeito à **Distância de Hamming**, já que estamos tratando de valores discretos, analisaremos a Tabela 5.1.

Decimal	Binário	Distância de Hamming
1	0000000001	1
2	0000000010	3
3	0000000011	2
4	0000000100	1
5	0000000101	0
6	0000000110	2
7	0000000111	1
8	0000001000	3
9	0000001001	2

Tabela 5.1: Conversão binária pelo Método Matemático e Distância de Hamming para o valor 5 (com 10 bits)

No capítulo três discutiu-se que as RNSPs e por sua vez a WiSARD, utilizam os bits de cada valor em separado, transformando-os em endereço para formar sua base de conhecimento gerando os padrões a serem reconhecidos por seus discriminadores, assim observa-se nos valores exemplificados nesta tabela que existe uma ligeira incoerência na Distância de Hamming destes valores que afetará as respostas obtida na rede, ao fazer a conversão de binário para decimal e tomado por base o valor 5, mediu-se a distância de todos os valores, note que o valor 6 mais próximo do valor 5 do que o valor 1 ou mesmo o valor 7 não expressa esta proximidade aritmética na distância de Hamming, por este tipo de incoerência e pela importância destes dados na eficiência da WiSARD, se faz necessário utilizar métodos alternativos para converter valores decimais em binários.

Outra forma de conversão é o Método do Termômetro, neste utiliza-se uma quantidade de bits para fazer a conversão, esta quantidade é muito importante e denomina-se neste trabalho de **Escala do Termômetro** e será melhor detalhada na seção **Janela de Entrada**, por hora o mais importante a salientar aqui é que a quantidade de bits escolhida deve poder alcançar os maiores valores do conjunto de dados sem perda de dados por *overflow*, assim escolhido o tamanho em bits pode-se converter os valores, para números decimais inteiros, cada acrescimo de valor em 1, um bit 0 será alterado para 1, podendo ser feito da esquerda para direita ou vice-versa, porém uma vez escolhida a forma de alteração a mesma deve ser mantida sempre, visto o método podemos olhar a Tabela 5.2 e analisar as distâncias de Hamming, a Escala utilizada foi de 10 bits como no exemplo da Tabela 5.1:

Note que as Distâncias de Hamming mantém coerência para todos os valores no que diz respeito a distância aritmética e diferença bit a bit, por este motivo o método de conversão de binário para decimal escolhido foi o Termômetro, porém

Decimal	Termômetro	Distância de Hamming
1	0000000001	4
2	0000000011	3
3	0000000111	2
4	0000001111	1
5	0000011111	0
6	0000111111	1
7	0001111111	2
8	0011111111	3
9	0111111111	4

Tabela 5.2: Conversão binária pelo Método do Termômetro e Distância de Hamming para o valor 5 (com 10 bits)

este método funciona com valores inteiros, e os valores dos atributos no *input* estão normalizados no intervalo entre 0 e 1, para valores **Reais**, após escolher a Escala do Termômetro, faz-se a multiplicação de cada valor do *input* por esta escala T :

$$\text{Input}_t = v * T | v \in \text{Input}_n$$

A quarta etapa do sistema finaliza entregando para a etapa final os valores que neste trabalho chama-se endereços, uma alusão ao endereçamento utilizado pela RNSP para consolidar seus conhecimentos e padrões dos discriminadores, neste momento que a rede neural irá utilizar cada linha gerada pelas etapas anteriores para treinar e testar seus discriminadores, as linhas são compostas de um rótulo que é a classe a qual pertence a amostra do endereço na linha, e toda a amostra com os dados tratados previamente nas etapas do sistema, a classe da amostra é a permissividade encontrada pelo TDR e posteriormente poderá ser obtido o teor de umidade do solo com este valor.

5.1 Janela de Entrada

Cada amostra do *Data log* que é composta com os dados do TDR e do Pluviômetro é tomada a cada cinco minutos, nomeou-se neste trabalho este período de coleta dos dados de Janela de Entrada, podendo-se trabalhar com janelas de tamanhos variados compondo um histórico de eventos na própria entrada da RNSP. A importância do histórico de eventos se dá pelo fato do estado atual do solo no que diz respeito a umidade ser influenciado pela cadeia de eventos que ocorreu nos períodos anteriores.

Existe ainda o fator de proximidade temporal do evento, quanto mais próximo do estado da amostra que pretende treinar ou prever, mais importante será, reduzindo este peso com o distanciamento temporal. Compor o *input* da RNSP variando o tamanho da janela de entrada foi importante para chegar aos resultados neste trabalho. Quanto maior a janela de entrada, maior será o *input* e, veremos mais adiante, maior será a quantidade de RAMs e o tamanho de cada uma delas.

5.2 Previsor de Umidade WiSARD

Foi apresentado a WiSARD no capítulo três deste trabalho, mostrando sua praticidade de treinamento e agilidade deste processo sendo feito uma única iteração por amostra a ser treinada, por esta praticidade decidiu-se o uso desta RNSP no presente trabalho e a partir deste momento apenas será referenciado como WiSARD. O Previsor de Umidade WiSARD é o sistema que envolve o processo descrito no início deste capítulo, desde a captação dos dados para treinamento do sistema, até o procedimento de previsão, onde o Previsor receberá apenas dados de Temperatura do solo e Índice Pluviométrico gerando um *input* com o tamanho adequado da janela de entrada para prever qual umidade o solo apresenta sem a necessidade de um equipamento sofisticado para detectá-la. Outra possibilidade é utilizar dados de previsão do tempo para um determinado período completando a lacuna do índice pluviométrico e sistematicamente ir prevendo a umidade futura do solo com estes dados, porém, para este uso do previsor é necessário que as previsões do tempo sejam precisas.

5.3 Quantidade e Tamanho das RAMs

O capítulo quatro deste trabalho detalhou os atributos considerados importantes para serem colocados na entrada da WiSARD: a Hora do Dia, o Período de Seca e a Temperatura do Solo. Notou-se a diferente importância de cada atributo para a previsão e testes serão apresentados no capítulo seis para entender como chegar em valores adequados utilizando o conceito de Escala de Termômetro já abordado neste capítulo, esta escala tem impacto direto no tamanho do *Input*

É possível calcular o tamanho de um *input* para adequar este comprimento e a quantidade de RAMs e o tamanho de cada uma delas, para isso é possível utilizar a seguinte fórmula:

$$b = \sum_{i=1}^{T_{jan}} EJan_i (EHour_i + ESeca_i + ETemp_{10cm_i} + ETemp_{30cm_i})$$

Onde b é o tamanho do *Input* em bits, T_{Jan} é o tamanho da janela em amostras, $EJan_i$ é a escala da i ésima janela, onde a proximidade temporal da janela terá maior escala em relação à históricos mais antigos, $EHour_i$ é a escala da Hora do Dia em que a que a medição i foi tomada, $ESeca_i$ é o Período de Seca até o momento desta medição e $ETemp_{10cm_i}$ e $ETemp_{30cm_i}$ são temperaturas da profundidade de dez centímetros e trinta centímetros respectivamente, tomadas na janela i .

Observado o tamanho do *input* é possível calcular o tamanho e quantidade de RAMs nos discriminadores, seja b o tamanho do *input* como visto na fórmula anterior e ER o espaço endereçável da RAM e QR a quantidade de RAMs dentro de um discriminador temos:

$$QR = \frac{b}{ER}$$

No capítulo três deste trabalho mostrou como um endereçamento é feito em cada RAM de cada Discriminador, é possível notar que estas RAMs se tornam esparsas conforme as amostras são apresentadas para o treinamento da WiSARD algumas regiões tenderão a ser acessadas com maior frequência conforme similaridade das amostras, é possível perceber que quanto maior for o espaço de uma RAM, mais difícil será de um mesmo endereço ser acessado por uma outra amostra o que pode tornar a rede especialista demais, o contrário também será observado, uma RAM muito pequena irá tornar a WiSARD generalista demais, buscar o equilíbrio neste ponto foi importante neste trabalho e uma grande quantidade de resultados será apresentado no capítulo seis para mostrar como chegou-se a um tamanho adequado.

5.4 Discriminadores

Como abordado anteriormente no capítulo três deste trabalho, os discriminadores são estruturas da WiSARD que terá a capacidade de definir o quão uma amostra é similar a uma classe previamente treinada por ele representada, este grau de similaridade, que já foi apresentado neste trabalho, será comparado com todas as respostas dos discriminadores da WiSARD e, como vimos antes, o melhor resultado será apresentado como resposta da previsão da rede.

Verificou a possibilidade de duas abordagens diferentes para criar os discrimina-

dores neste trabalho, em princípio a pesquisa se ateve a usar seis discriminadores onde cada um tinha um quantificador linguístico e uma faixa de valores atribuída a estes quantificadores, com a intensificação dos testes e a dificuldade de calibrar as faixas de umidade para cada tipo de solo experimentou-se a Divisão Dinâmica das Classes.

Estas classes serão as respostas dadas pela WiSARD informando qual o teor de umidade do solo, esta resposta poderá ser em faixas no caso da divisão em seis classes ou um valor numérico inteiro indicando o índice de permissividade do solo como no caso da Divisão Dinâmica.

5.4.1 Divisão em seis Classes

Esta forma de lidar com os discriminadores, requer um trabalho à parte no que diz respeito a entender a dinâmica hídrica do solo, exigindo um conhecimento técnico prévio para calibrar as faixas dentro de cada quantificador linguístico, os muitos tipos de solo alteram a forma de retenção dos líquidos assim um solo com teor de umidade alto pode significar seco para outro tipo solo com o mesmo teor de umidade.

Os dados utilizados neste trabalho se limitam a um tipo de solo apenas com duas coberturas diferentes, Pastagem e Mata, notou-se um teor de umidade menor na parcela de solo com cobertura de mata, como visto no capítulo dois deste trabalho isso deve-se principalmente ao efeito de Evapotranspiração da Vegetação desta parcela[16][17], onde as arvores drenam uma grande quantidade da água precipitada no solo.

Faixas e Quantificadores Linguísticos

As seis classes tiveram quantificadores linguísticos atribuídos a elas, estes quantificadores ajudam a ter uma visão intuitiva da situação do solo a respeito do teor de umidade presente no mesmo e a estes quantificadores foram atribuídos faixas do índice de Permissividade captado pelo TDR, porém estas faixas são diferentes para cada tipo de solo como abordado anteriormente, segue abaixo tabela com os quantificadores e suas faixas dado o tipo de solo:

Exigi-se um trabalho técnico e exaustivo para calibrar estas faixas de valores de permissividade e integrar essas informações de dados com o tipo exato de solo monitorado, devido a dificuldade em chegar nestas faixas e ao conhecimento técnico

	Pastagem	Mata
Seco	Menor que 11	Menor que 9
Muito Baixo	Maior ou igual a 11 e Menor que 18	Maior ou igual a 9 e Menor que 14
Baixo	Maior ou igual a 18 e Menor que 20	Maior ou igual a 14 e Menor que 17
Médio	Maior ou igual a 20 e Menor que 21	Maior ou igual a 17 e Menor que 20
Alto	Maior ou igual a 21 e Menor que 28	Maior ou igual a 20 e Menor que 25
Saturado	Maior que 25	Maior que 25

Tabela 5.3: Quantificadores Linguísticos e intervalos de permissividade.

para chegar em faixas adequadas, optou-se por criar uma forma dinâmica que será abordada a seguir e foi esta abordagem a utilizada durante os testes e para os resultados deste trabalho.

5.4.2 Divisão Dinâmica das Classes

As classes criadas com faixas fixas tem um fator de dificuldade técnica para se manter coerente dentro do tipo de cada solo existente, gerando um desgaste inicial para calibrar o sistema, assim criou-se a Divisão Dinâmica das Classes, que consiste em atribuir as faixas não pelo conteúdo volumétrico de água ou VWC, neste método a WiSARD irá responder com o índice de Permissividade do solo medido pelo TDR e seu teor de umidade em VWC poderá ser obtido em seguida com os cálculos adequados para o solo específico.

Para chegar às classes os dados recebidos do TDR não precisam ser tratados de forma diferente ao que já utilizado para as seis classes definidas anteriormente, basta verificar os valores mínimos e máximos de permissividade que o TDR indica para o solo em cada profundidade desejada, e estes valores são discretizados a números inteiros indo do mínimo ao máximo encontrado. Esta atribuição é feita sem intervenção manual, os mínimos e máximos são atribuídos durante o treinamento da WiSARD

Faixas e Permissividade

Como visto anteriormente ainda que as classes são dinâmicas de acordo com a Permissividade encontrada na leitura do TDR, as mesmas precisam de uma faixa para cada resposta possível da WiSARD, os mínimos e máximos de Permissividade encontrados nos dados recebidos do TDR servirão para ajudar na atribuição das faixas, utilizando a regra de arredondamento matemático para este fim, assim um valor de Permissividade 10,56 recebido do TDR será da classe 11 por estar mais próximo da classe 11 do que da classe 10.

Para a zona de Pastagem os valores obtidos pelo TDR no período chuvoso foi de 9 no mínimo e 25 no máximo, porém este valor máximo de 25 de permissividade é um valor abaixo do esperado, o solo não chegou a ficar saturado em momento algum no período em que os dados foram coletados, porém no momento que um valor diferente for obtido nas medições o sistema poderá gerar uma classe nova automaticamente sem a necessidade de calibrar os valores se a rede for treinada com um conjunto de dados que defina esta nova classe, caso contrário novos resultados que forem recebidos pelo sistema e que fiquem fora de um limiar pré-definido de acerto e confiança poderá ser cadastrado pelo sistema como uma faixa fora da sua base de conhecimento recebendo como rótulo Máximo + 1 ou Mínimo -1, necessitando assim de uma intervenção técnica para cadastro de novas classes. Como os resultados obtidos pela rede será dado em **Permissividade** ficará a cargo do usuário o cálculo do VWC para verificar o teor de umidade relativo ao solo mensurado

5.5 Treinamento e Testes

Dado um conjunto de dados de entrada, o treinamento da WiSARD será efetuado utilizando cada amostra desse conjunto em separado. As amostras serão apresentadas uma a uma e apenas uma vez por treinamento, como abordado anteriormente, não existe a necessidade de iterar mais de uma vez cada entrada. Após o conjunto de dados de entrada que foram escolhidos para o treinamento for todo apresentado para o WiSARD a fase de treino da rede será finalizada e poderá ser realizada a fase de Testes.

Na fase de testes, os resultados de cada previsão é guardado, sendo armazenados os resultados dos discriminadores com maior similaridade e também o segundo melhor resultado, esta abordagem é importante para avaliar os resultados da rede como poderemos verificar no capítulo seis quando verificar-se as métricas de análise dos resultados, mais precisamente a Métrica Confiança. Após os resultados dos testes serem exibidos poderemos averiguar os resultados da WiSARD.

5.6 Validação Cruzada

Durante a fase de treinamento e testes da WiSARD abordados na seção anterior, utilizou-se a técnica de Validação Cruzada, certificando-se da capacidade de

generalização da rede. O modelo k partições foi o adotado, onde o conjunto de amostras foram divididos em dez partes, utilizou-se nove partes para treinamento e uma delas para testes obtendo a precisão de cada iteração, após o término do treinamento outro em seguida é efetuado sendo desta vez a parte utilizada para teste comporá o conjunto de treinamento e uma das partes que ainda não foi utilizada para testes ficará separada para este fim.

O treinamento da WiSARD e os resultados de previsão só serão avaliados após todo o processo que inclui o Validação Cruzada for executado por completo. Os resultados dos testes serão apresentados no próximo capítulo e levam em consideração cada iteração realizada durante o treinamento da rede.

Capítulo 6

Resultados

Será exibido neste capítulo um grande conjunto de testes realizados para chegar ao melhor balanceamento do conjunto de atributos já abordados no capítulo quatro deste trabalho. Foram selecionados os resultados dos testes realizados com os dados da Mata e da Pastagem de apenas um dos quadrantes, ao todo são quatro quadrantes em cada região, sendo dois localizados no topo do terreno e outros dois localizados em sua base. É possível notar que os quadrantes na base apresentam maior índice de umidade, estas diferenças não afetam os resultados da obtidos pela WiSARD.

6.1 Métricas

Foram selecionadas algumas métricas para avaliar as previsões obtidas pela WiSARD neste trabalho, em princípio decidiu-se por averiguar todos os acertos da rede em relação direta ao total de amostras no conjunto, utilizando a Acurácia, $(VP + VN)/A$ onde VP é o total de amostras preditas corretamente para à classe que a mesma se enquadra, VN é o total de amostras preditas corretamente em não pertencer à classe e A é o total de amostras no conjunto de dados. A Acurácia pode apresentar valores incorretos para conjuntos desbalanceados, esta sensibilidade pode ser percebida utilizando outras métricas em conjunto como Revocação, Precisão e Especificidade.

A Revocação, $VP/(VP + FN)$, obtém uma taxa de acertos levando em consideração todas as amostras da classe que deveriam ser preditas pelo sistema, FN é a quantidade de predições que a WiSARD erroneamente classificou como negativa. Precisão leva em consideração os resultados verdadeiramente positivos em relação à todos os resultados positivos obtidos pela rede, $VP/(VP + FP)$, onde FP são os resultados preditos erroneamente pela Rede como verdadeiros. Na métrica

Especificidade, o resultado obtido é a taxa de predições acertadamente negativas em relação ao total de predições que realmente seriam negativas, $VN/(VN + FN)$, mostrando a capacidade da WiSARD responder às amostras negativas onde realmente são negativas.

Além dessas métricas utilizou-se outras duas, Eficiência e Confiança, sendo a Eficiência a média aritmética da Revocação e Especificidade, $(Revocação + Especificidade)/2$, na prática a Revocação e a Especificidade variam em direções opostas onde um classificador sensível a positivos tende a gerar muitos Falsos Positivos. Para a WiSARD, ou qualquer outro classificador, possuir boa eficiência, a Revocação e a Especificidade tenderá a 100% em seus resultados, esta medida ajudará a qualificar como positivo ou negativo os resultados aqui obtidos para em seguida decidirmos quais direções tomar para ajustar os atributos de entrada da WiSARD.

A Confiança é a avaliação da certeza que a WiSARD tem de sua resposta para uma previsão, diferente das outras métricas que tratam diretamente com resultados da WiSARD divididos na Matriz de Confusão, esta métrica tem relação direta com as respostas da WiSARD e como chegou a predição adotada, como mostrado no capítulo três, no momento em que a rede estiver prevendo, cada Discriminador será acionado, o que tiver o maior score será selecionado, porém, o segundo maior score será guardado para efetuar o cálculo da Confiança, $(R1 - R2)/R1$ onde $R1$ é a melhor resposta e $R2$ a segunda melhor resposta [18], é possível que este resultado evidencie uma indecisão da WiSARD fazendo-nos tomar medidas para melhorar a modelagem, ou mostrar o grau de certeza desta, seja para acertos ou erros de fato, visto que a WiSARD pode apresentar um coeficiente de Confiança alto mesmo realizando previsões erradas como podemos observar no gráfico da figura 6.1.

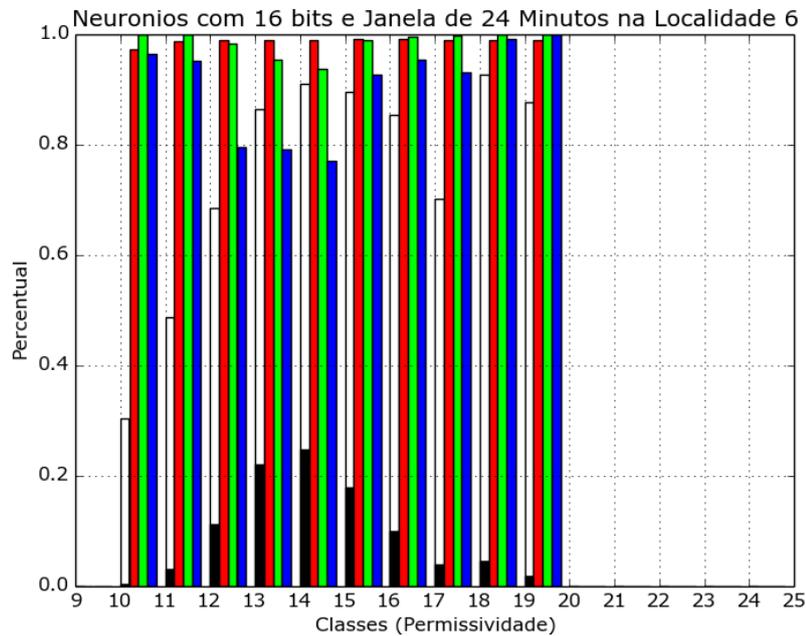


Figura 6.1: Alto coeficiente de Confiança com Baixo acerto da Previsão

Os gráficos dispostos para cada teste é dividido em A e B, para facilitar o entendimento dos resultados evitando poluir a área de plotagem. No gráfico A estão dispostos as Proporções em amostras de cada classe na cor Preta, na cor Branca os resultados da Revocação, em vermelho a Precisão, em verde a Eficiência e por fim em Azul a Especificidade. No gráfico B a Proporção, Revocação e Eficiência são mais uma vez exibidos para facilitar o entendimento dos resultados de Confiança em vermelho e Acurácia em Azul.

6.2 Previsão com Classes Dinâmicas

Verificou-se no capítulo cinco do presente trabalho que utilizar seis classes nomeadas com quantificadores linguístico seria demasiado trabalhoso e pouco eficiente. Decidiu-se utilizar apenas o método de Classes Dinâmicas para a previsão da Umidade no solo, deste modo foram feitos exaustivos teste averiguando resultados utilizando as métricas citadas no início deste capítulo, foram feitos testes variando a escala do termômetro e variações no tamanho das RAMs.

As classes são divididas em valores de Permissividade atingido pelo solo

mensurado, podendo ser treinado uma nova classe no momento que for verificado a necessidade. Os intervalos entre uma classe e outra é definido de forma matemática, os valores de permissividade são números reais que serão discretizados para números inteiros gerando assim uma classe. Lembrando o exemplo utilizado no capítulo anterior, uma medida tomada pelo TDR no valor de 10,56 será relativo à classe 11 por estar matematicamente mais próximo do valor de Permissividade 11.

6.3 Variações nas Escalas do Termômetro

Como visto anteriormente no capítulo cinco, utilizou-se para a transformação do *input* em endereços válidos para a WiSARD o modelo de Termômetro, e que esta forma permite uma abordagem adequada para esta finalidade no que diz respeito à distância matemática entre um bit e outro no *input*. Uma quantidade de bits é atribuída no momento da transformação dos valores decimais em binários que chamamos de Escala. Podemos modificar esta escala dimensionando um atributo dentro dos neurônios da rede, aumentando ou diminuindo sua importância pro resultado.

Para descobrir melhores resultados foram feitos testes e estudos na literatura para obter a melhor distribuição desta escala. Nas seções a seguir serão exibidos resultados priorizando cada um dos atributos diferentes que compõe uma janela de dados que representa o *input*.

6.3.1 Priorizando Hora do Dia

Quando ocorre um evento de precipitação no ambiente, a água percorre seu caminho infiltrando no solo reduzindo a temperatura deste, segundo [14] esta mudança de temperatura pode ser utilizada para inferir a umidade do solo com alguma margem de erro tolerável. Observou-se nos estudos do *Data set* de Itaboraí que durante os períodos noturnos a temperatura do solo costumava a equiparar-se à temperatura diurna durante uma precipitação, gerando dúvida nestas ocorrências o que poderia confundir a WiSARD durante o seu treinamento.

Verificou-se a necessidade de integrar ao *input* o horário exato da amostra, esta medida foi importante para tratar a ambiguidade das temperaturas noturnas e diurnas em períodos chuvosos. Outra importância da hora do dia para a previsão da umidade pela WiSARD é mostrada em [16][17], onde evidencia que o aumento

da Evapotranspiração pelo nível de incidência da radiação solar na vegetação, reduzindo a umidade do solo, porém este atributo sozinho não torna a WiSARD capaz de prever a umidade como podemos observar nos gráficos 6.2, 6.3, 6.4 e 6.5

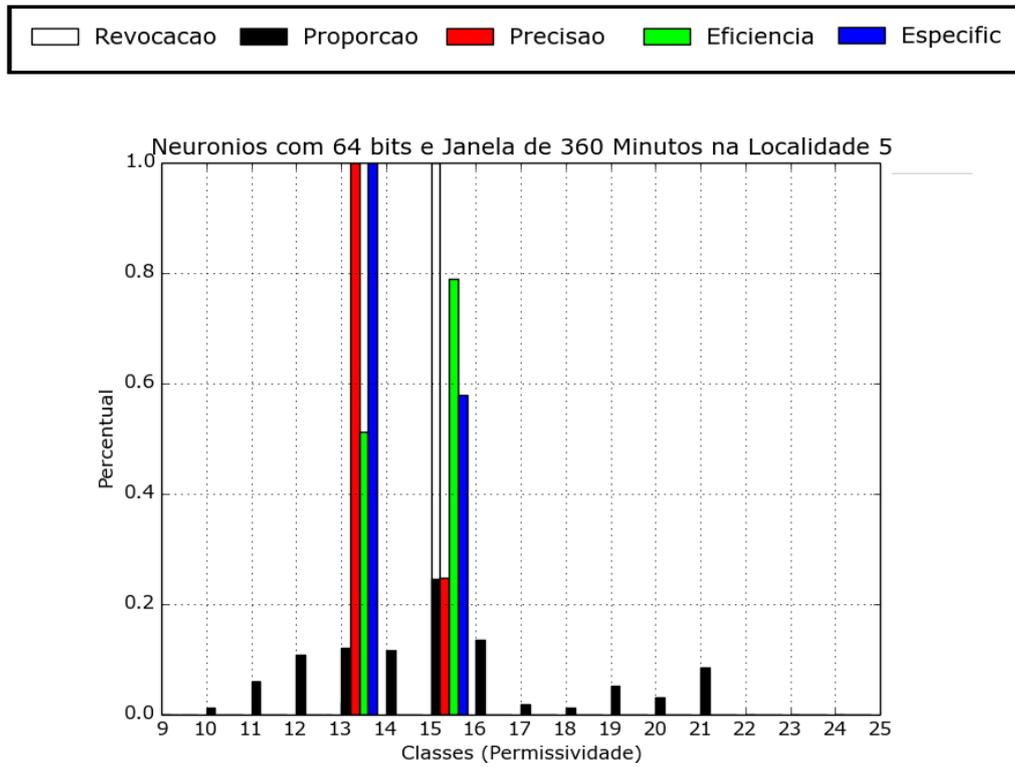


Figura 6.2: Hora do Dia: Quadrante Superior Esquerdo na Pastagem A

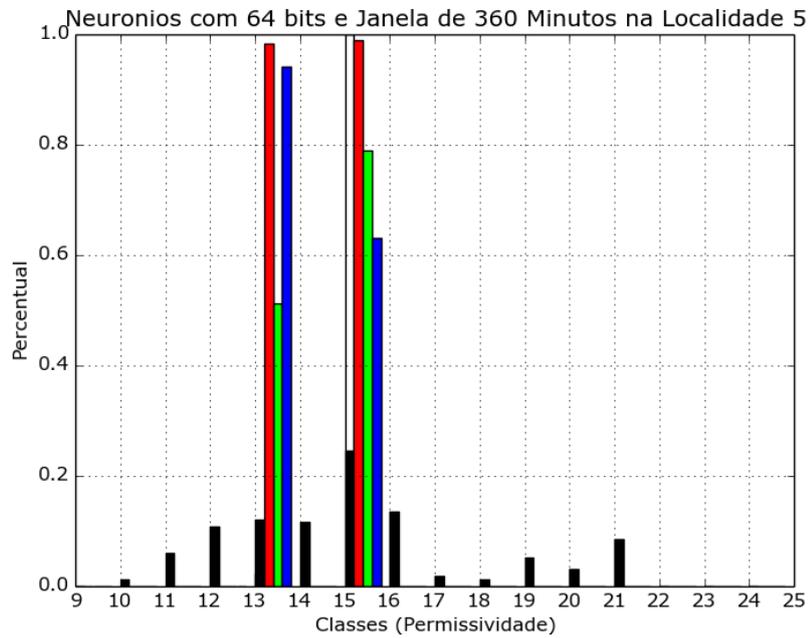


Figura 6.3: Hora do Dia: Quadrante Superior Esquerdo na Pastagem B

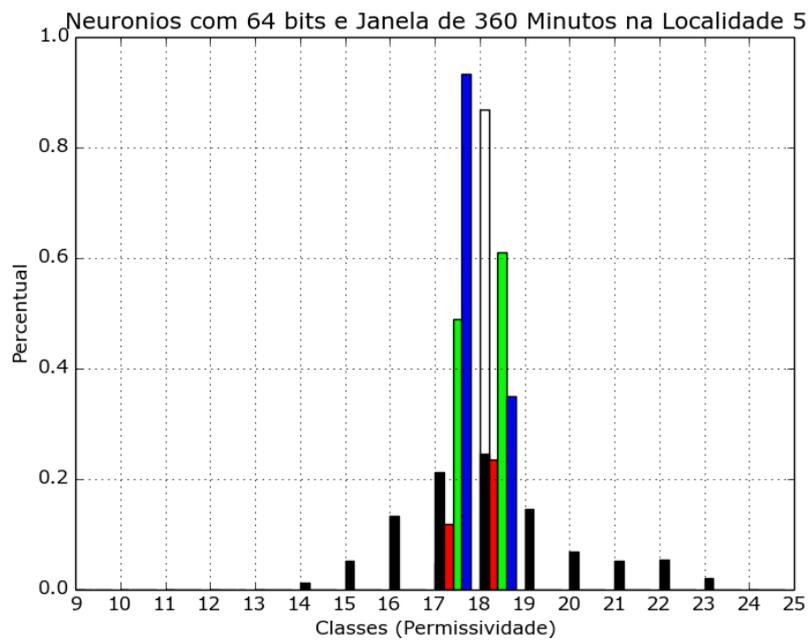


Figura 6.4: Hora do Dia: Quadrante Superior Esquerdo na Mata A

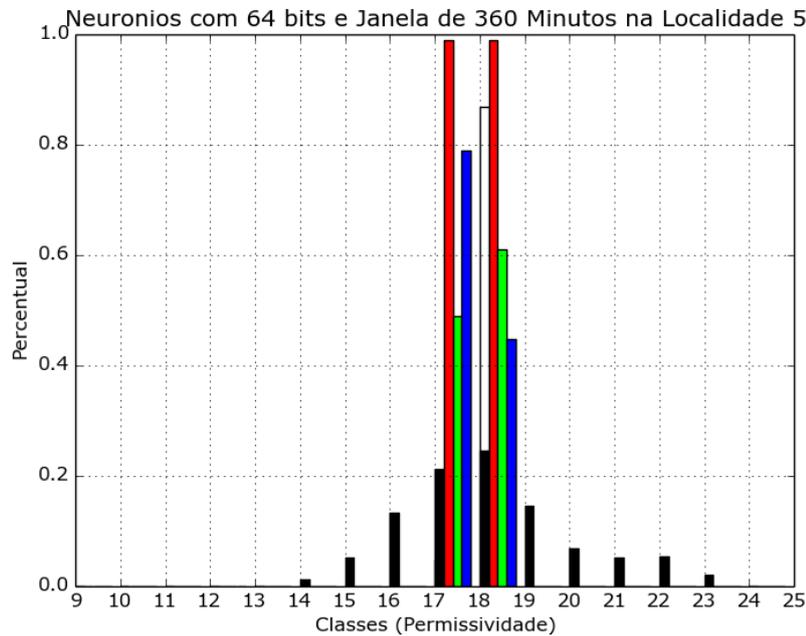


Figura 6.5: Hora do Dia: Quadrante Superior Esquerdo na Mata B

6.3.2 Priorizando Temperatura

Como visto no capítulo cinco, a temperatura mostrou-se um atributo importante desde o início dos estudos para a determinar a umidade no solo. A água que se aloja nos micropóros e nos macropóros após os eventos de precipitação, evaporam-se com as altas da temperatura, em um solo com pouca cobertura ocorrerá naturalmente com os raios solares incidindo diretamente no chão. Em solos cobertos com Mata, existe ainda o aumento da temperatura do solo porém não tão acentuado como vistos no gráfico 5.5, porém ainda ocorre a diminuição da umidade em decorrência da Evapotranspiração, a própria vegetação se encarrega de absorver a umidade do solo para seus processos fotossintéticos[16][17]. Alguns experimentos realizados com a finalidade de entender a correlação da temperatura com a umidade mostrou resultados coerentes no que diz respeito a previsão desta. Os resultados a seguir referem-se aos testes que priorizam o atributo Temperatura aumentando sua Escala e tornando-o 90% do *input* para o treinamento e teste da WiSARD, ficando os outros 10% divididos entre os outros atributos.

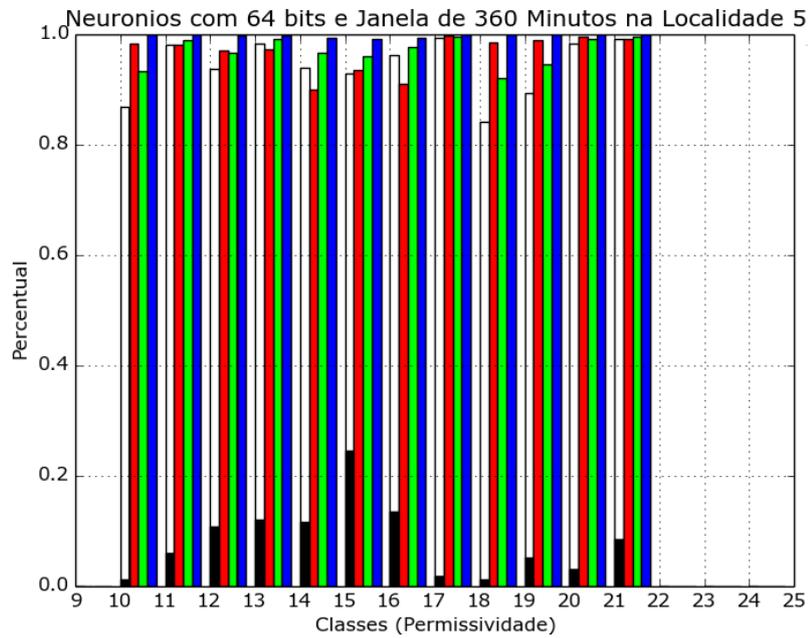


Figura 6.6: Temperatura: Quadrante Superior Esquerdo na Pastagem A

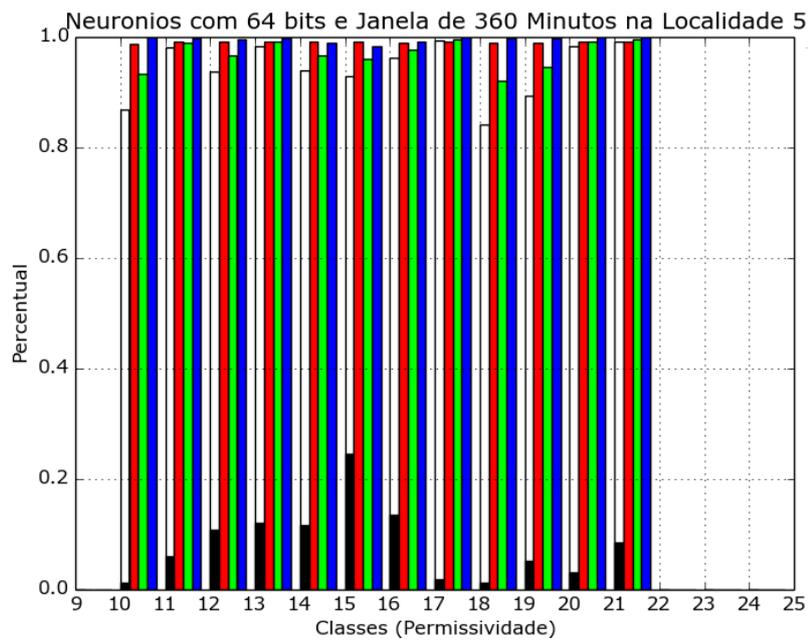


Figura 6.7: Temperatura: Quadrante Superior Esquerdo na Pastagem B

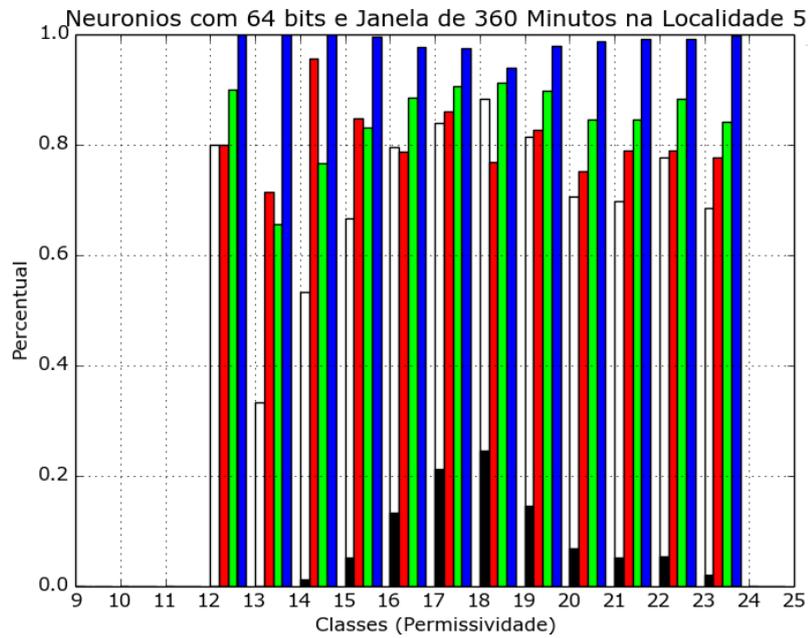


Figura 6.8: Temperatura: Quadrante Superior Esquerdo na Mata A

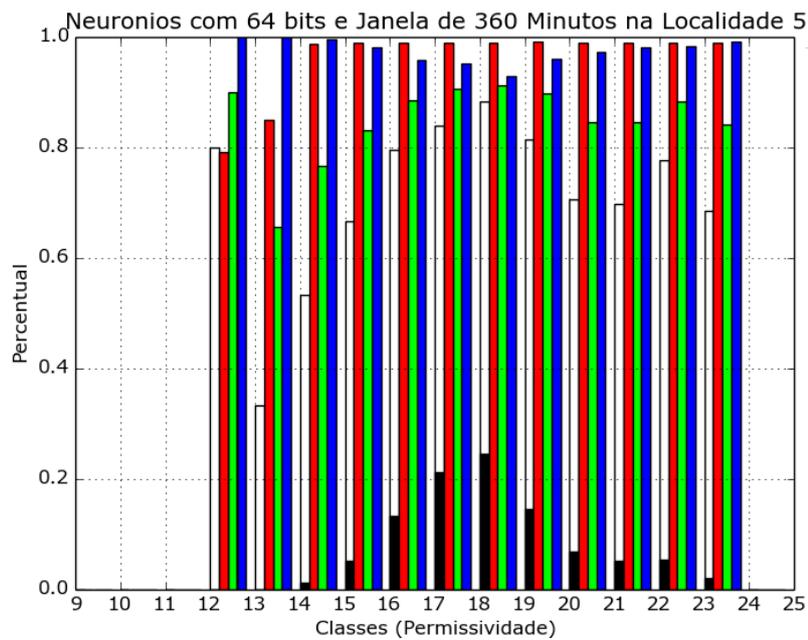


Figura 6.9: Temperatura: Quadrante Superior Esquerdo na Mata B

6.3.3 Priorizando Período de Seca

Priorizou-se este atributo durante alguns experimentos, o que tornou a previsão bastante eficiente, visto anteriormente no capítulo cinco sua importância, nestes resultados ficaram evidentes que a resposta da WiSARD pode ter um incremento de preditividade com o Período de Seca. Nos gráficos a seguir, a Escala deste atributo foi dimensionada para 90% como nos testes anteriores, e os resultados foram os seguintes:

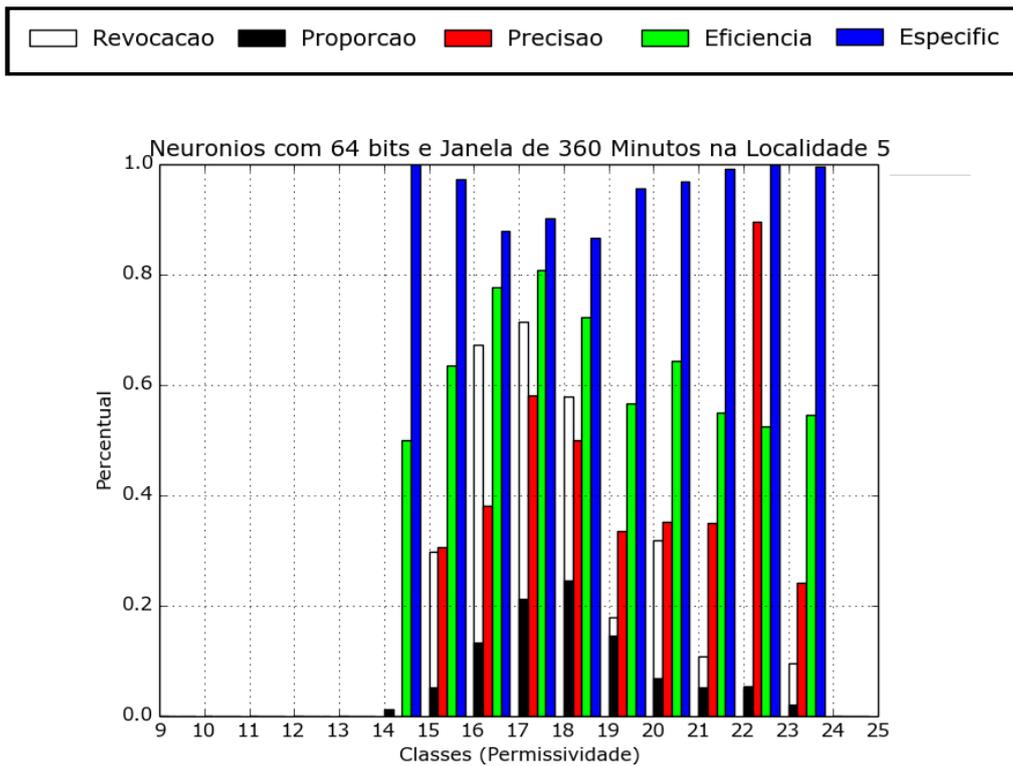


Figura 6.10: Período de Seca: Quadrante Superior Esquerdo na Pastagem A

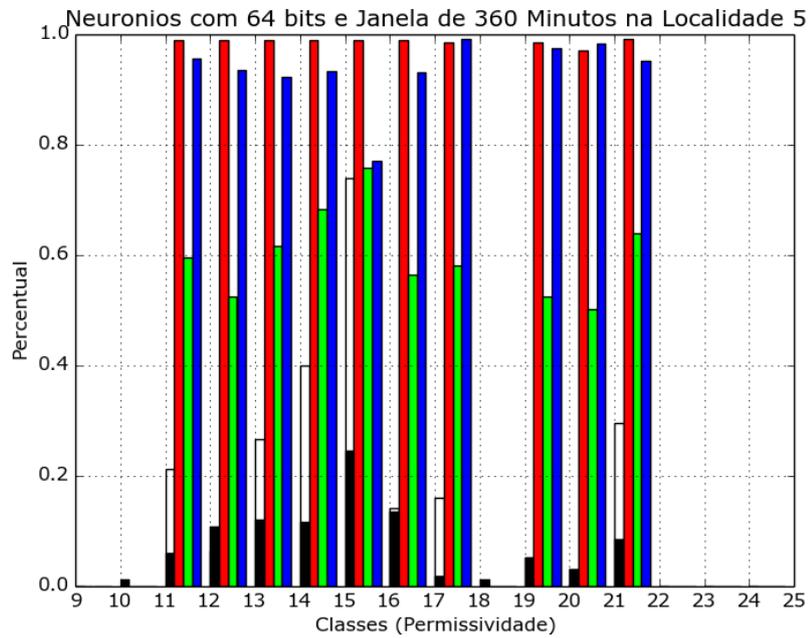


Figura 6.11: Período de Seca: Quadrante Superior Esquerdo na Pastagem B

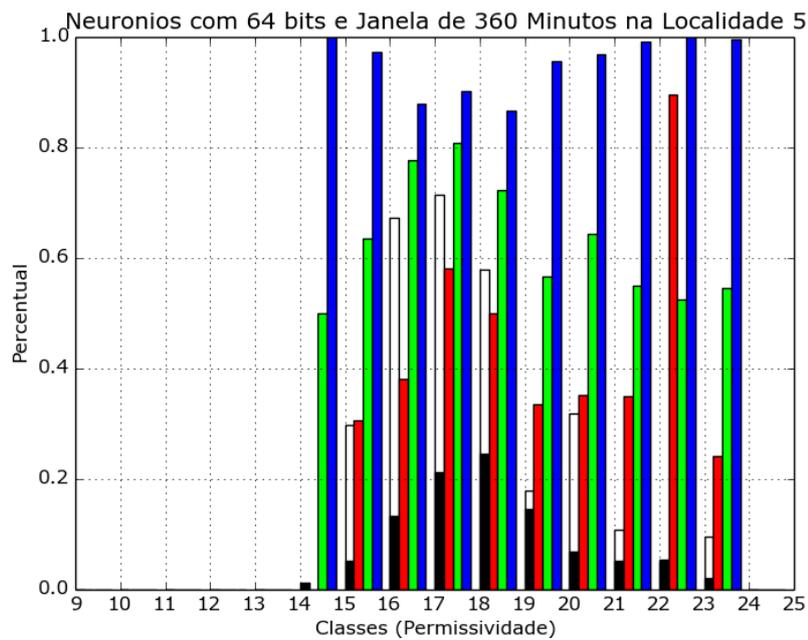


Figura 6.12: Período de Seca: Quadrante Superior Esquerdo na Mata A

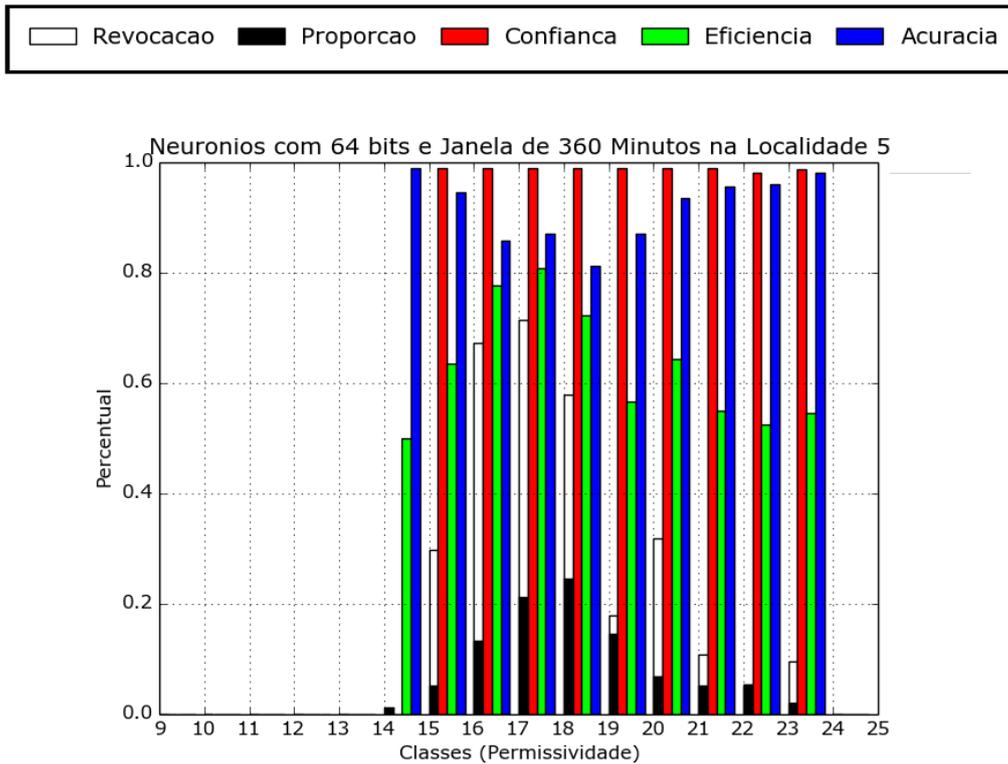


Figura 6.13: Período de Seca: Quadrante Superior Esquerdo na Mata B

6.3.4 Balanceando Dados de Entrada

Seguindo os resultados dos testes priorizando cada um dos atributos do *input* da WiSARD, e todo o estudo realizado durante o desenvolvimento deste trabalho, buscamos balancear os Dados de Entrada, verificando os resultados notou-se uma ordem de correlação dos atributos, a Temperatura responde pelos melhores resultados recebendo assim a maior Escala, logo depois o Período de Seca que também responde por resultados melhores que o atributo Hora do Dia que tem a menor Escala dentre os atributos.

As previsões de umidade mostradas nos resultados são todas referentes ao sensores localizados à trinta centímetros da superfície do solo, com isso outro atributo importante é a profundidade em que é medida a Temperatura, Período de Seca e Hora do Dia, estes três atributos tem duas medidas cada, uma para dez centímetros e outra para trinta centímetros, porém com escalas ligeiramente diferentes, os atributos mais próximos da superfície possuem 80% de sua dimensão em relação aos atributos da profundidade que pretende-se prever a umidade.

Existem outros atributos ainda que indiretos que foram contextualizados no

capítulo cinco, o tamanho da Janela, por exemplo, possui uma escala alternativa, este atributo se trata do histórico do solo que comporá o *input* da WiSARD. Notou-se a importância deste atributo, existe uma correlação direta com os eventos anteriores em relação à umidade do solo, esta correlação diminui com a evolução do tempo, e para evidenciar isto existe ainda uma escala dando maior importância à eventos próximos ao período em que deseja-se obter a previsão da umidade. Esta escala varia com o tamanho da janela, reduzindo a importância de históricos antigos de forma gradativa.

6.3.5 Descarte de Atributos

Alguns atributos do TDR foram destacados, ou não se aplicavam, voltagem da bateria dos sensores, *VWC* ou *Volumetric Water Content* que é o valor da umidade do solo, não usamos este valor nem para o treinamento e teste supervisionado já que se trata de um resultado do volume de água inferido a partir da calibração dos resultados de Permissividade recebidos pelos sensores do TDR, estes valores são muito sensíveis aos diferentes tipos de solos, desta forma utilizou-se apenas os valores de Permissividade, assim a WiSARD prevê apenas este valor e a partir deste poderá ser adquirido o *VWC* para o solo estudado, retirando a necessidade de calibrar e remodelar a WiSARD inteira ao examinar outro tipo de solo.

6.4 Variações no Tamanho das RAMs

Durante o capítulo cinco discutiu-se sobre o tamanho das RAMs e como isso influencia nos resultados da WiSARD, com maiores RAMs obtém-se maior espaço para endereçar o que pode especializar a rede do contrário teremos um poder maior de generalização. Os estudos realizados no *Data Set* utilizado neste trabalho, notou-se que pequenas variações de temperatura durante o dia pode ter grande influencia na umidade do solo, isso combinado ao período de seca e o tamanho da janela com o histórico dos eventos anteriores evidenciam que mudanças aparentemente pequenas podem mudar o índice de permissividade do solo. Analisando este comportamento nos dados e com um grande número de classes que podem compor a WiSARD, decidiu-se por testar diferentes tamanhos de RAMs e decidir qual o espaço de endereçamento mais adequado.

Foram testados quatro tamanhos de RAMs: 16, 32, 48 e 64 bits. Os resultados foram separados em subseções para cada tamanho testado para melhor organização

dos mesmos. Todos os testes foram feitos com Janela de Entrada com o tamanho de seis horas, na próxima seção será apresentado os resultados para diferentes janelas. Como informado no início deste capítulo, será apresentado apenas resultados de quatro quadrantes, dois na parte superior dos terrenos localizados na mata e na pastagem, um de cada tipo de terreno, e outros dois na parte inferior destes terrenos, a escolha destes quadrantes se devem ao fato de abrangerem classes diferentes mesmo fazendo parte de terrenos com a mesma composição e tipo de vegetação, o que muda é apenas a localização, topo do talude ou base do mesmo, mostrando assim a capacidade de previsão do sistema.

Decidiu-se por mostrar os resultados nos gráficos para cada classe em separado ao invés de mostrar o resultado total da rede com as médias dos resultados, para melhor avaliar a precisão da WiSARD em relação a cada classe, como os dados que foram utilizados neste trabalho são dados resultados de medições em um ambiente natural, existe um desbalanceamento natural das classes, algumas destas classes chegam a representar menos 0,01% dos total de amostras, é importante ter certeza de que todas as classes são contempladas com resultados dentro do que esperávamos em suas previsões.

6.4.1 RAMs com 16 bits

Aqui serão apresentados os resultados para tamanhos de RAMs com 16 bits, além destes resultados seguirá uma pequena tabela com informações gerais para estes testes e após a apresentação de todos os gráficos um pequeno resumo para entendimento dos mesmo.

	Mata	Pastagem
Tempo de Treinamento	866 Segundos	793 Segundos
Média Percentual de Revocação	92,43%	92,39%
Desvio Padrão	0,37%	0,4%

Tabela 6.1: Resultados para RAMs com 16 Bits

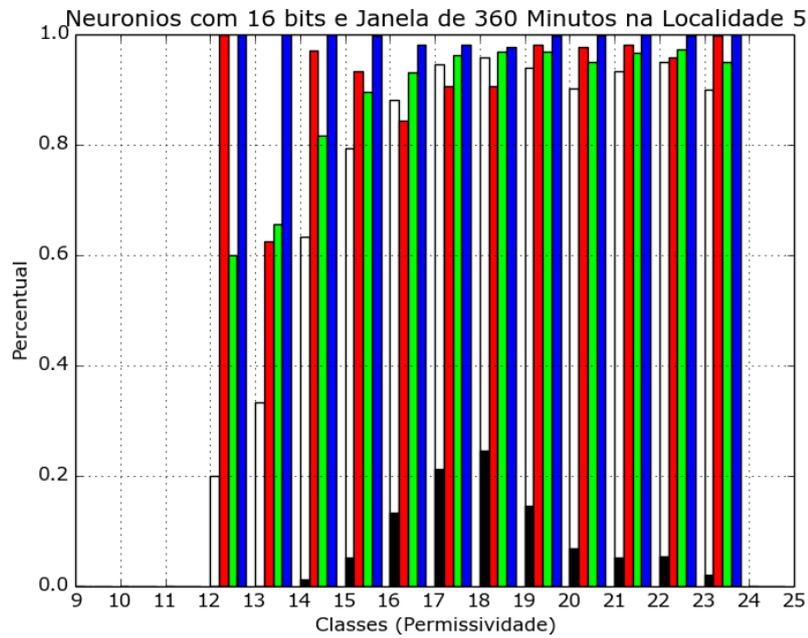


Figura 6.14: Gráfico A Quadrante Superior Esquerdo na Mata (RAM de 16 Bits)

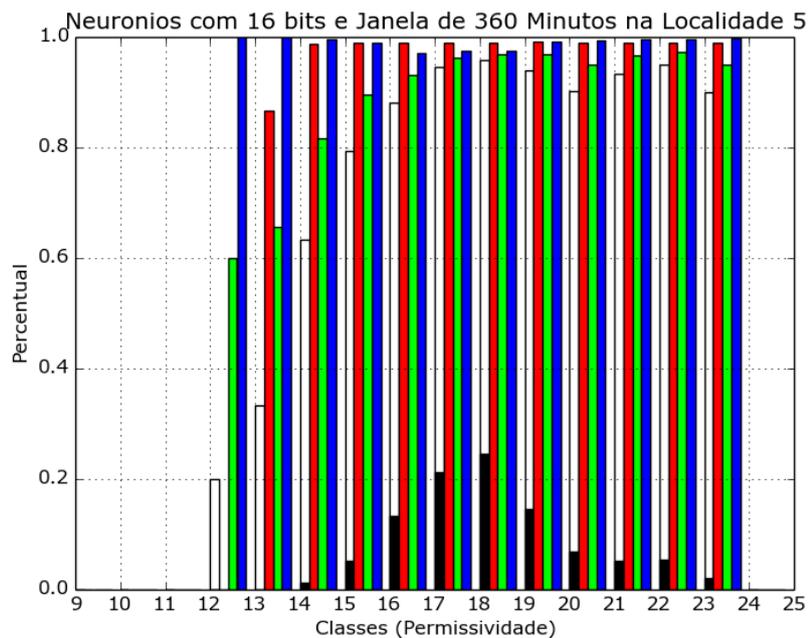


Figura 6.15: Gráfico B do Quadrante Superior Esquerdo na Mata (RAM de 16 Bits)

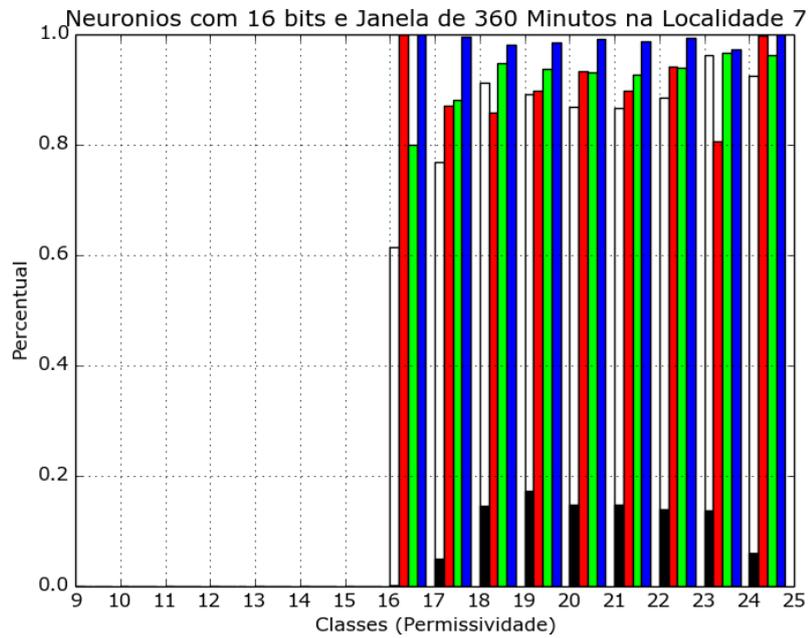


Figura 6.16: Gráfico A Quadrante Inferior Esquerdo na Mata (RAM de 16 Bits)

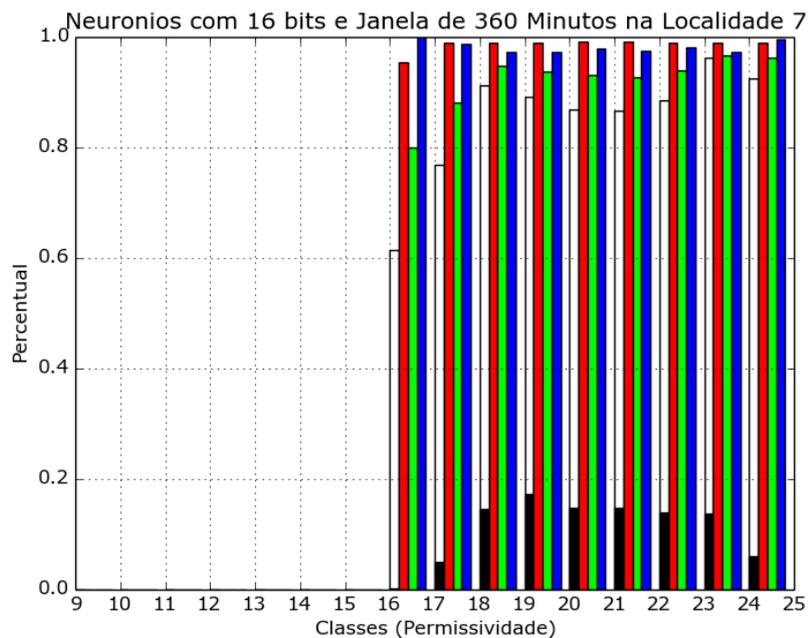


Figura 6.17: Gráfico B do Quadrante Inferior Esquerdo na Mata (RAM de 16 Bits)

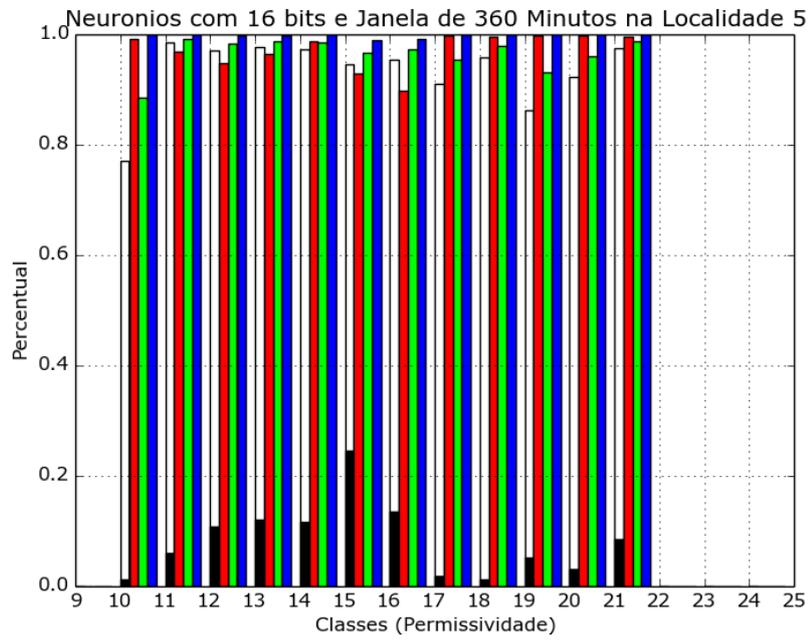


Figura 6.18: Gráfico A Quadrante Superior Esquerdo na Pastagem (RAM de 16 Bits)

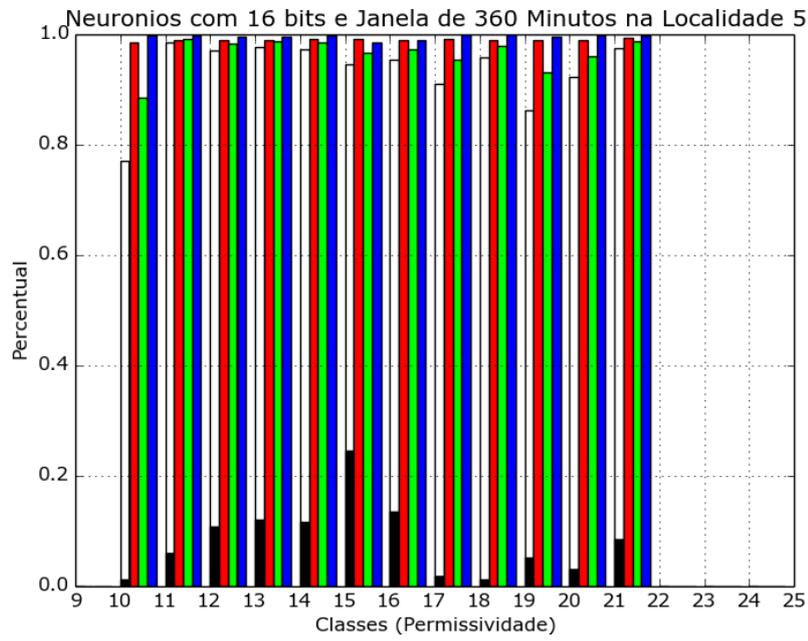


Figura 6.19: Gráfico B do Quadrante Superior Esquerdo na Pastagem (RAM de 16 Bits)

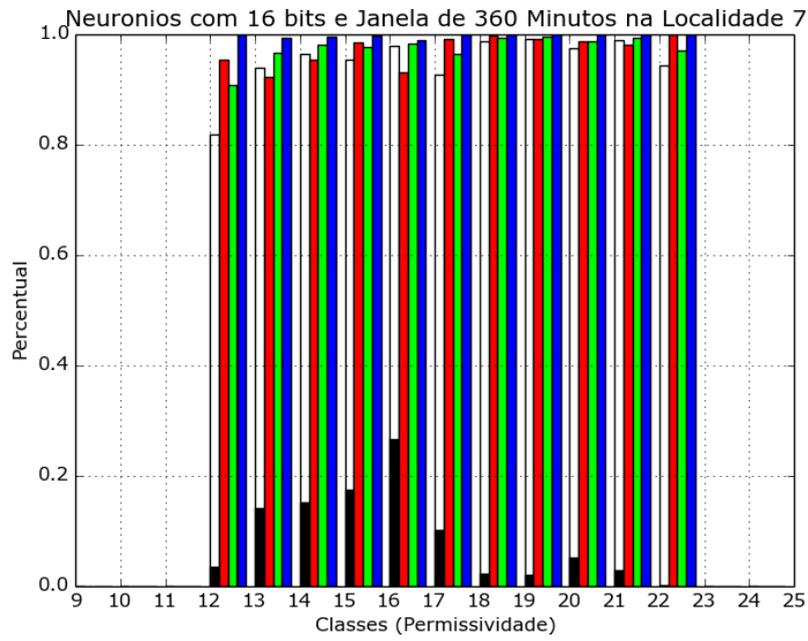


Figura 6.20: Gráfico A Quadrante Inferior Esquerdo na Pastagem (RAM de 16 Bits)

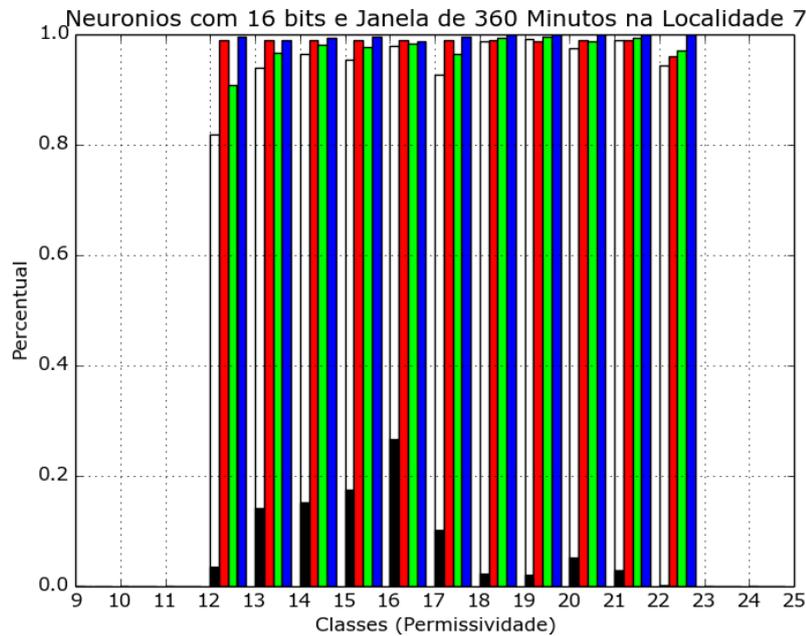


Figura 6.21: Gráfico B do Quadrante Inferior Esquerdo na Pastagem (RAM de 16 Bits)

Resumo

Os resultados dos testes mostram que apesar de a média dos acertos da WiSARD serem elevados chegando a 92% e Desvio Padrão desta média de 0,37%, estes estão longe de serem bons resultados, claramente foram influenciados pelos resultados das classes mais presentes no *data set*. Pôde-se notar os resultados das classes 12 e 13 na Mata, são bem inferiores aos outros e podemos considerar não satisfatórios para a finalidade que desejamos que é Prever a Umidade do solo dado um evento de chuva, estes resultados mostram que a rede está generalizando demais.

Podemos notar também que em algumas métricas os resultados pareceram aceitveis, como no caso da Precisão e Especificidade das classes 12 e 13 na Mata, estas métricas são suscetíveis à conjuntos desbalanceados de dados, que é exatamente o caso em questão. Em Ambos os resultados, Precisão e Especificidade, pode-se notar que bastando algumas poucas boas respostas acertadas da rede para Verdadeiros Positivos com poucos Falsos Positivos ou Verdadeiros Negativos com poucos Falsos Negativos obteve-se resultados aparentemente aceitveis, porm se usarmos somente estas métricas teremos resultados falíveis de previsão.

6.4.2 RAMs com 32 bits

Abaixo serão exibidos resultados para RAMs com tamanho de 32 bits, resultados estes que serão discutidos mais adiante. Pode-se notar na tabela abaixo, que os resultados com esta configuração tem um grande aumento na taxa de acertos das previsões, a redução no tempo de treinamento é esperada, como o comprimento do *input* se mantém o mesmo a quantidade de RAMs por discriminador irá reduzir com o aumento do comprimento de cada RAM, como visto no capítulo cinco na fórmula $QR = \frac{b}{ER}$, com isto existirá um número bem menor de iterações durante o treinamento da WiSARD.

	Mata	Pastagem
Tempo de Treinamento	516 Segundos	539 Segundos
Média Percentual de Revocação	97,55%	97,51%
Desvio Padrão	0,21%	0,32%

Tabela 6.2: Resultados para RAMs com 32 Bits

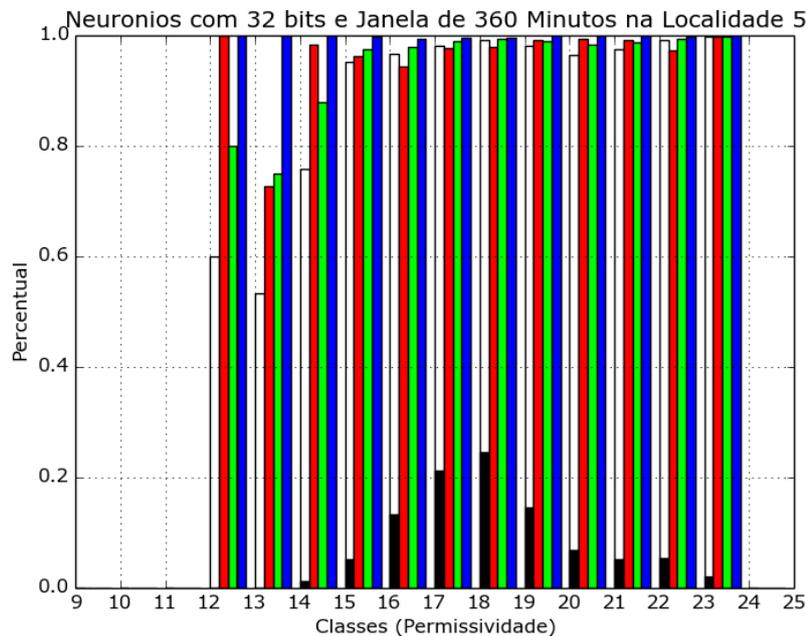


Figura 6.22: Gráfico A Quadrante Superior Esquerdo na Mata (RAM de 32 Bits)

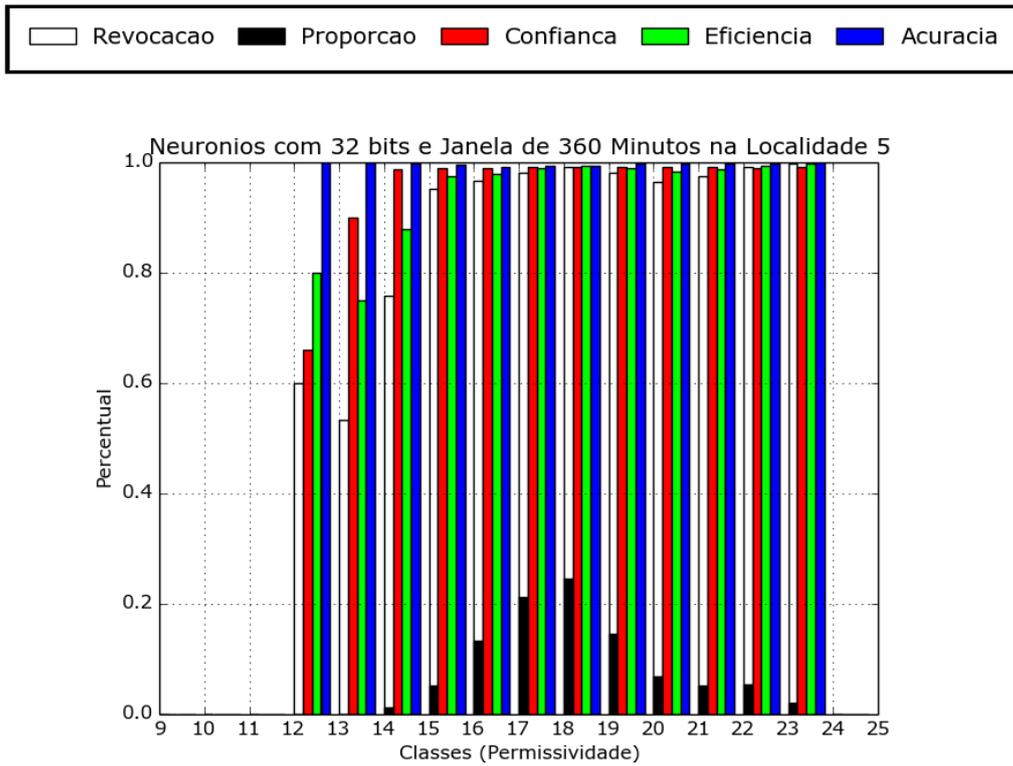


Figura 6.23: Gráfico B do Quadrante Superior Esquerdo na Mata (RAM de 32 Bits)

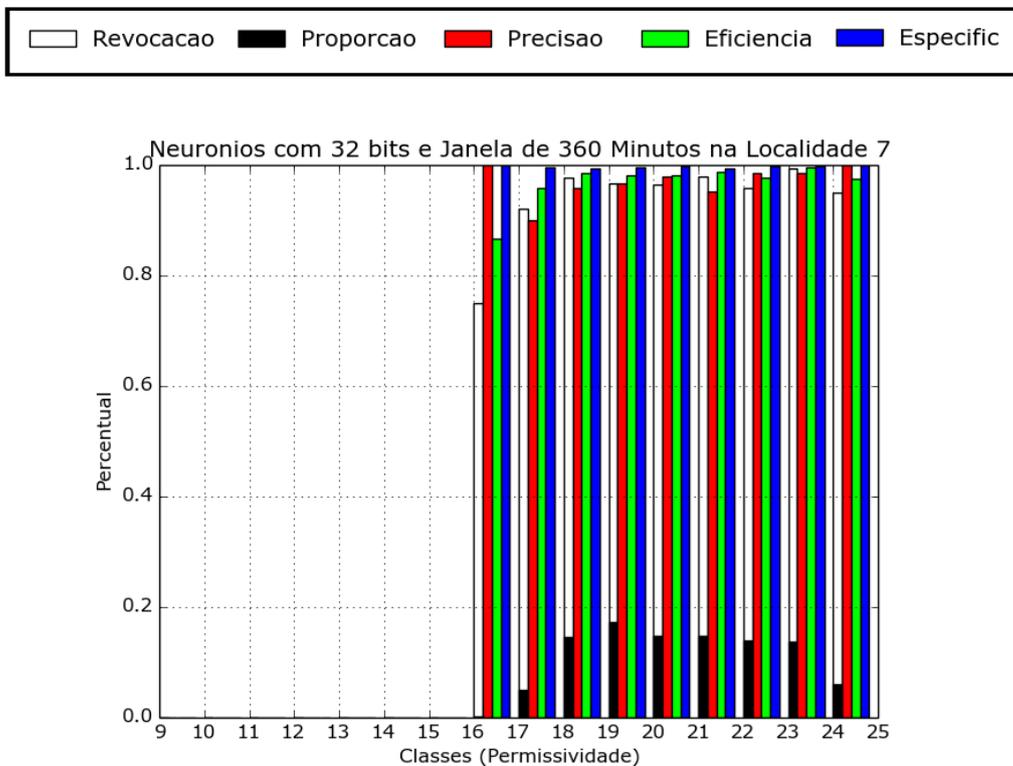


Figura 6.24: Gráfico A Quadrante Inferior Esquerdo na Mata (RAM de 32 Bits)

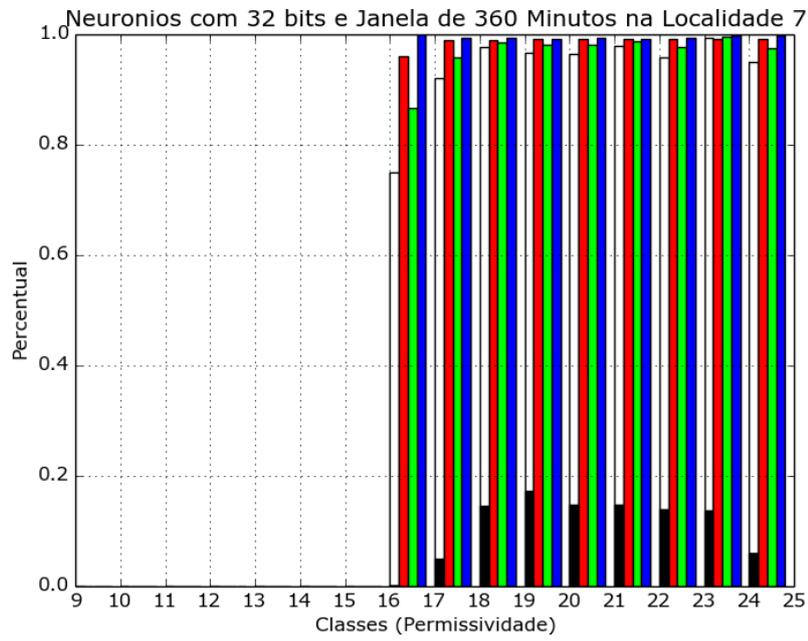


Figura 6.25: Gráfico B do Quadrante Inferior Esquerdo na Mata (RAM de 32 Bits)

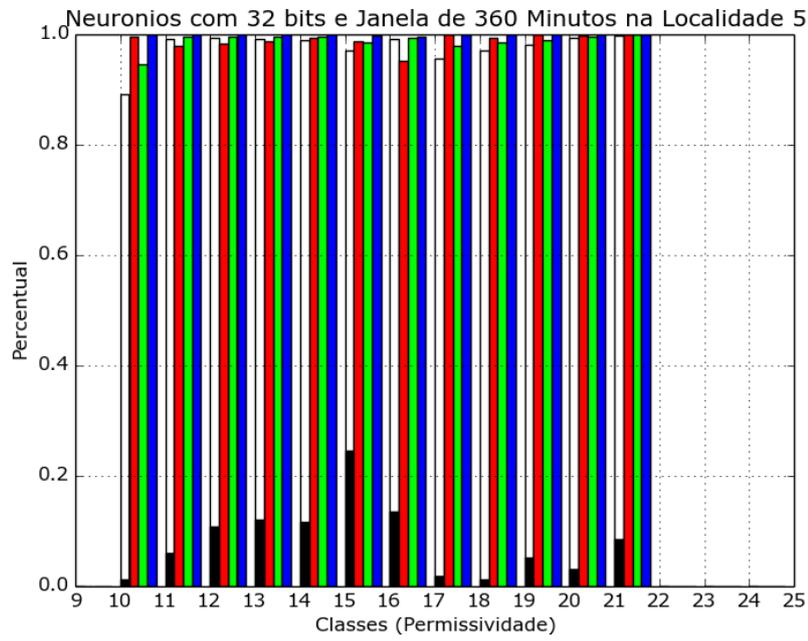


Figura 6.26: Gráfico A Quadrante Superior Esquerdo na Pastagem (RAM de 32 Bits)

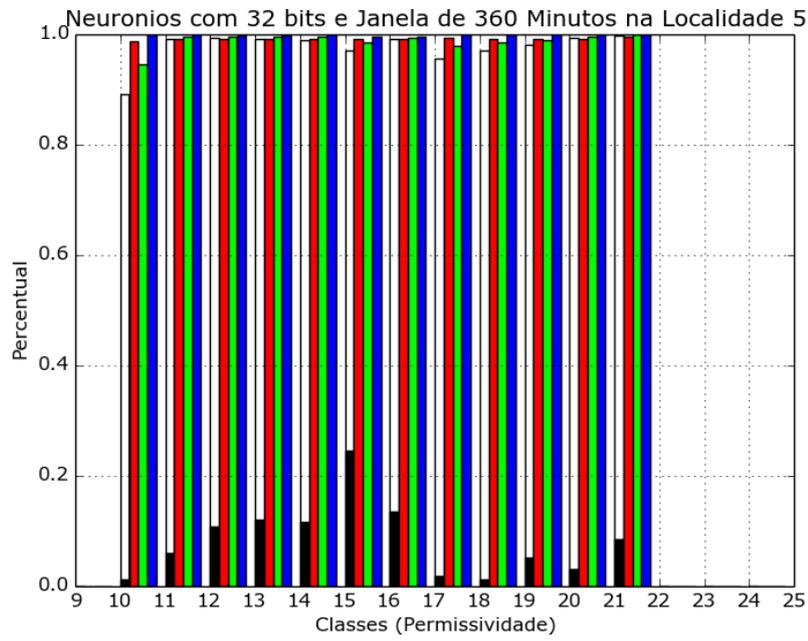


Figura 6.27: Gráfico B do Quadrante Superior Esquerdo na Pastagem (RAM de 32 Bits)

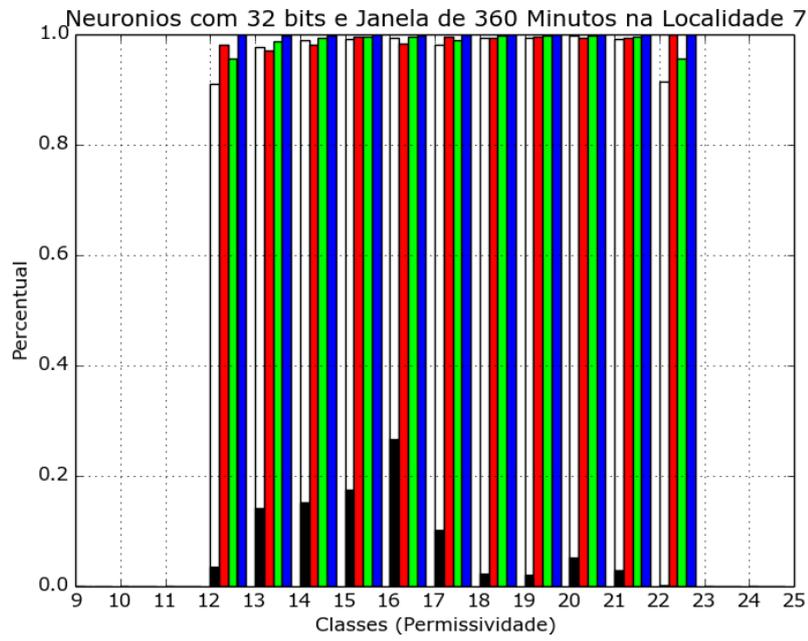


Figura 6.28: Gráfico A Quadrante Inferior Esquerdo na Pastagem (RAM de 32 Bits)

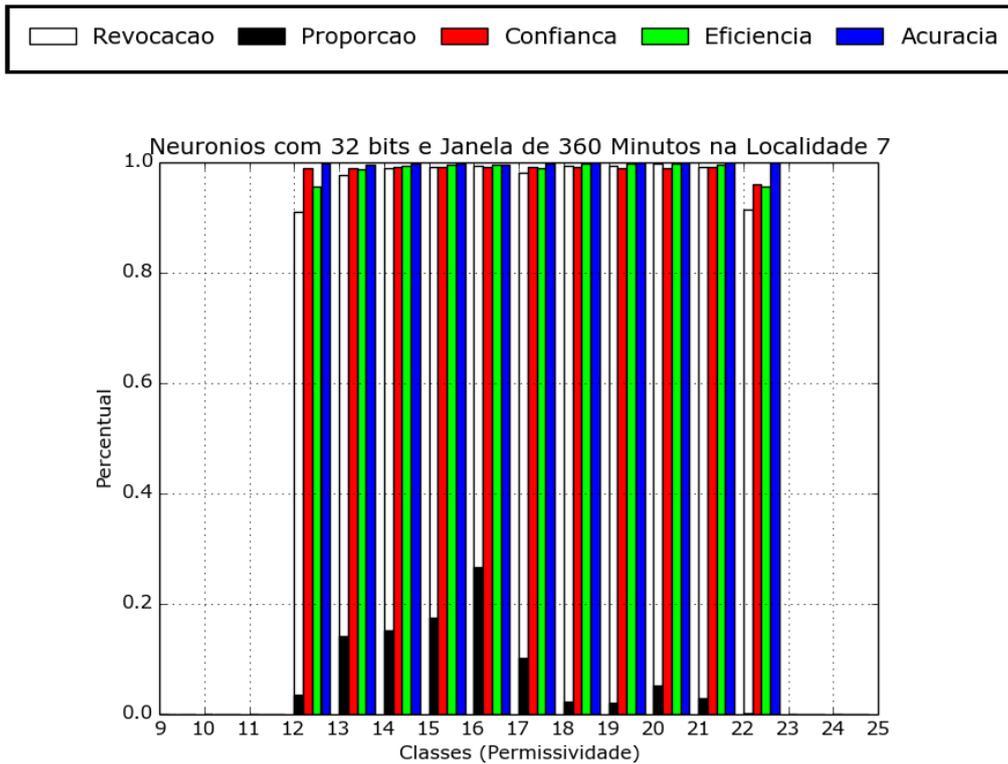


Figura 6.29: Gráfico B do Quadrante Inferior Esquerdo na Pastagem (RAM de 32 Bits)

Resumo

Os resultados para RAMs de 32 bits, mostram um grande aumento nos acertos das previsões como já mencionado antes, porém o ganho mais evidente é observado nos resultados de por classe, onde classes menos presentes no *data set* tiveram percentuais elevados acima de 50% de Revocação consequentemente elevando a Eficiência para estas classes, ainda não é um resultado satisfatório, no entanto mostra um caminho para possíveis melhoras nos resultados.

6.4.3 RAMs com 48 bits

Podemos observar na tabela que os resultados para 48 bits também obtiveram incremento em seus resultados, o tempo de treinamento novamente apresenta uma grande queda, porém, este tempo pode ser consideravelmente menor se o sistema for executado em hardware mais robusto. Seguindo o mesmo formato das subseções anteriores, será exibido aqui a tabela e os gráfico para então discutir os resultados.

	Mata	Pastagem
Tempo de Treinamento	477 Segundos	483 Segundos
Média Percentual de Revocação	98,49%	98,48%
Desvio Padrão	0,17%	0,19%

Tabela 6.3: Resultados para RAMs com 48 Bits

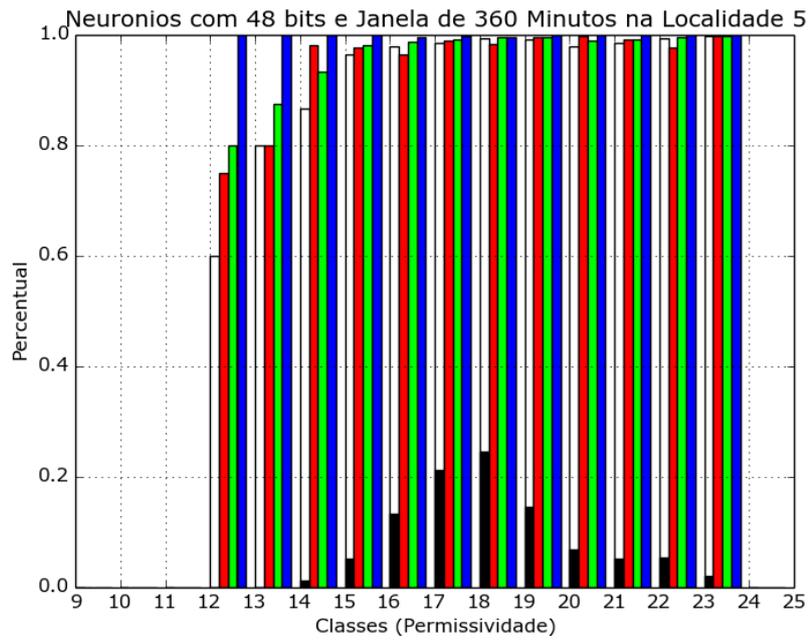


Figura 6.30: Gráfico A Quadrante Superior Esquerdo na Mata (RAM de 48 Bits)

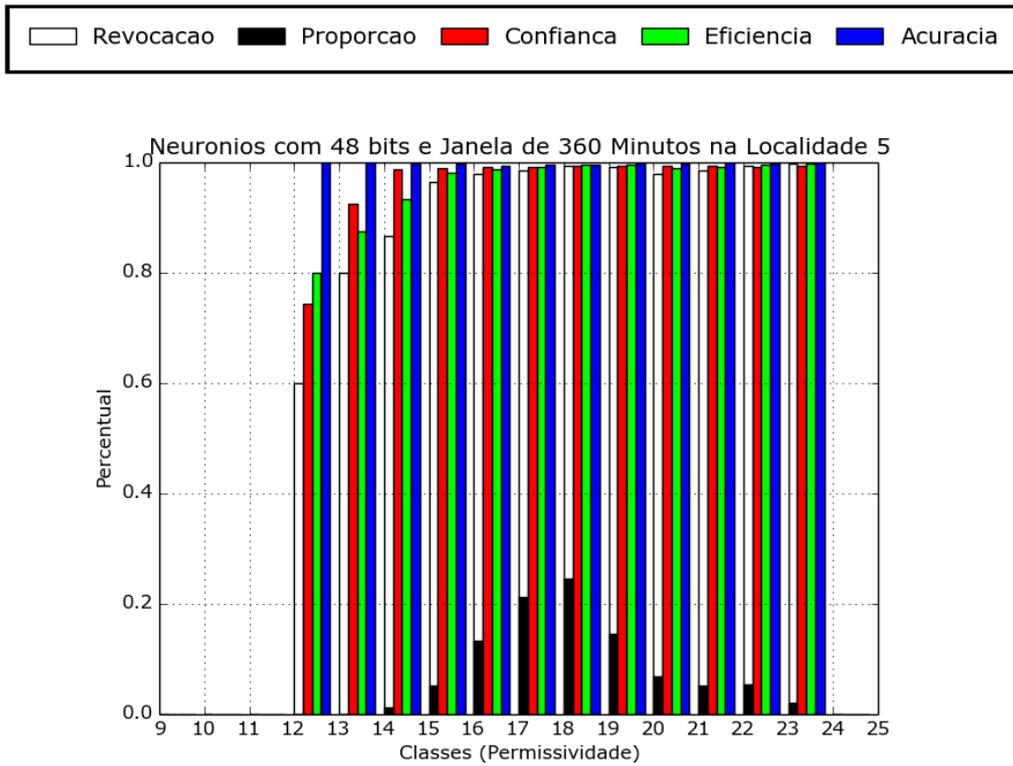


Figura 6.31: Gráfico B do Quadrante Superior Esquerdo na Mata (RAM de 48 Bits)

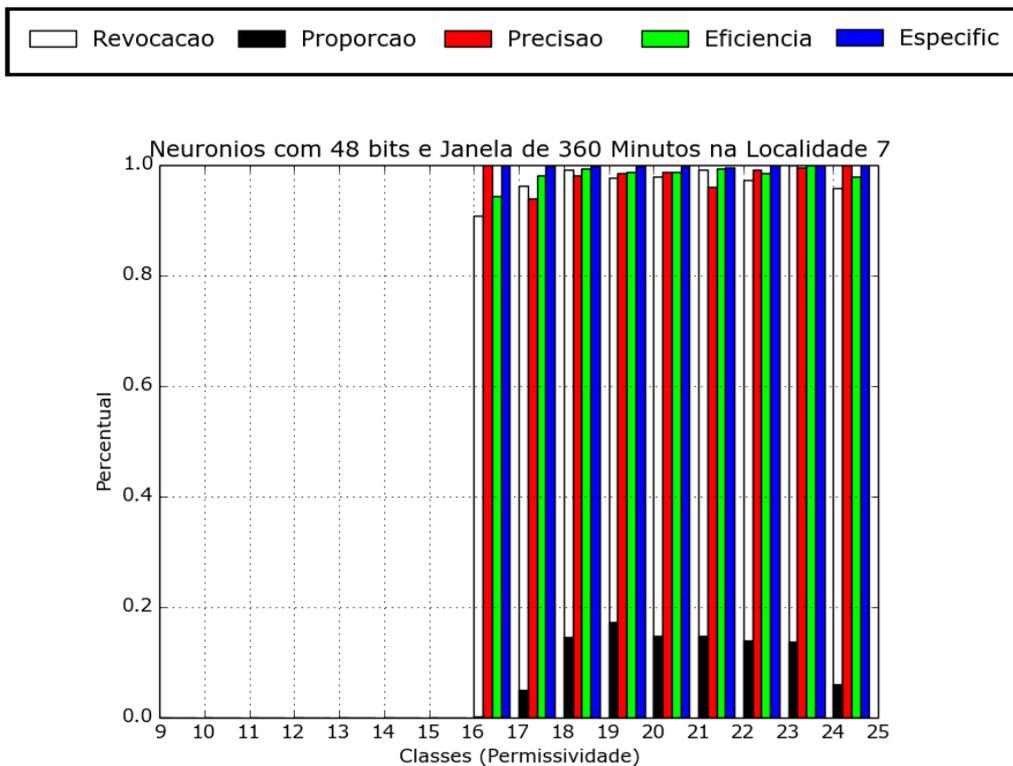


Figura 6.32: Gráfico A Quadrante Inferior Esquerdo na Mata (RAM de 48 Bits)

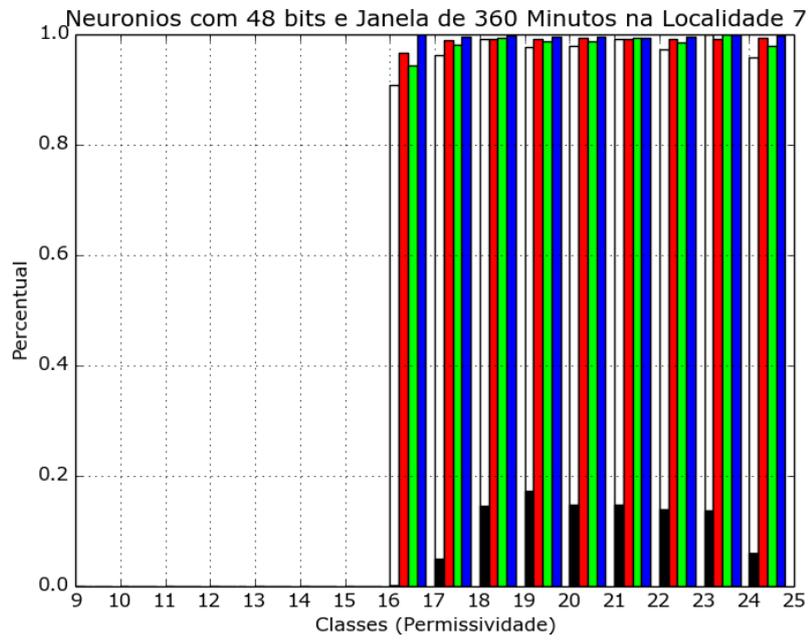


Figura 6.33: Gráfico B do Quadrante Inferior Esquerdo na Mata (RAM de 48 Bits)

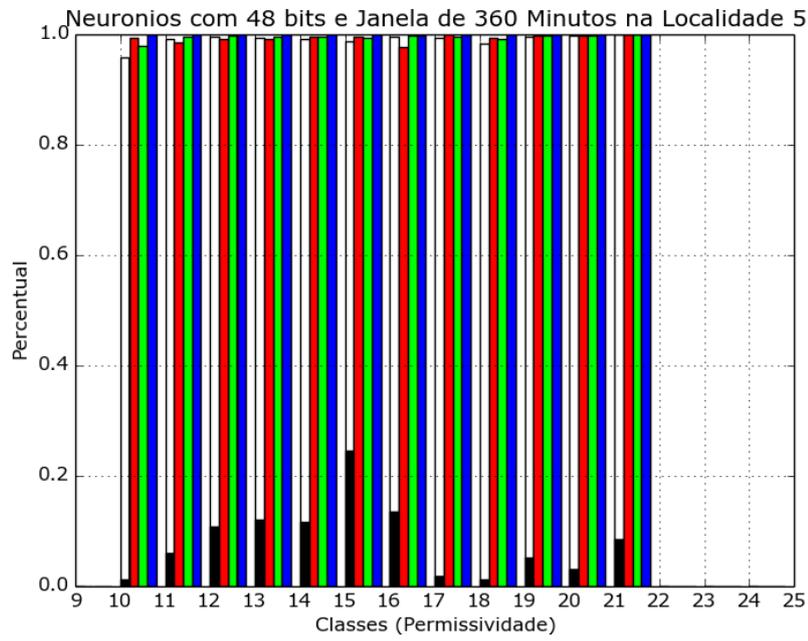


Figura 6.34: Gráfico A Quadrante Superior Esquerdo na Pastagem (RAM de 48 Bits)

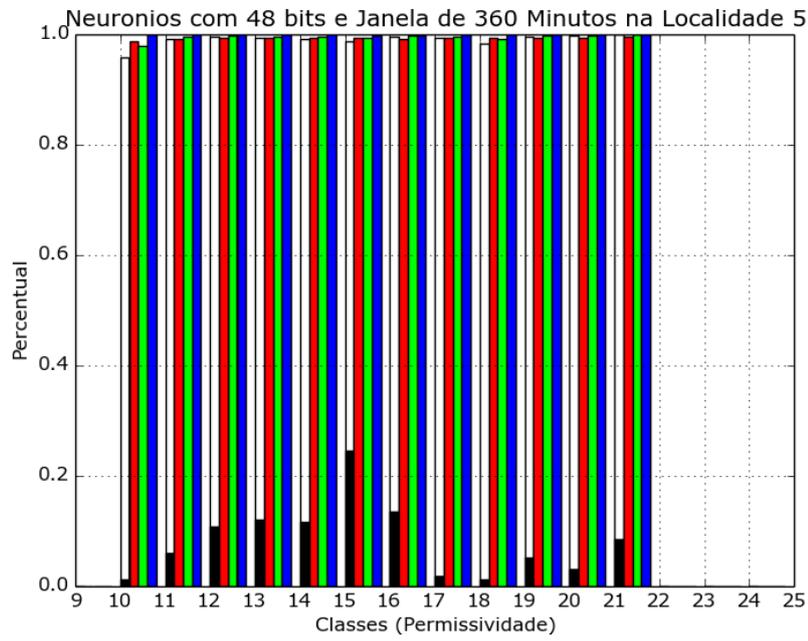


Figura 6.35: Gráfico B do Quadrante Superior Esquerdo na Pastagem (RAM de 48 Bits)

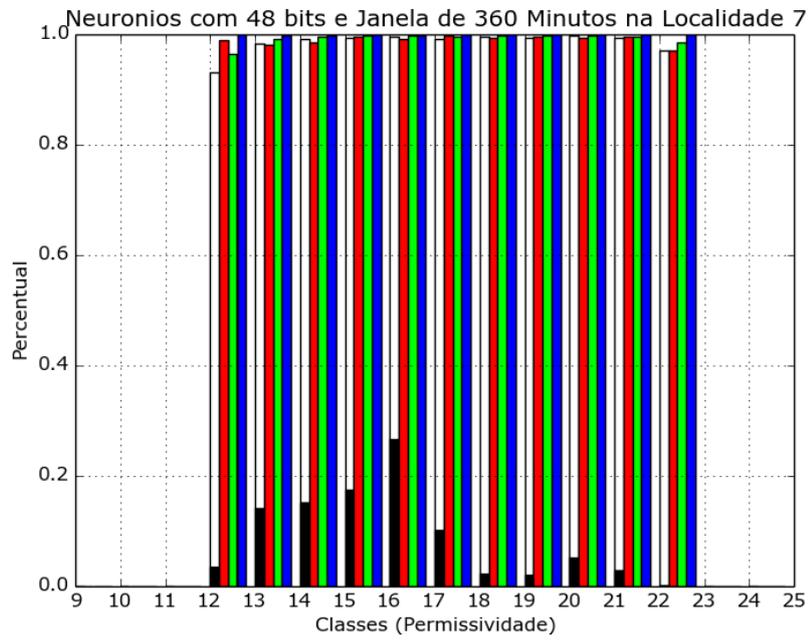


Figura 6.36: Gráfico A Quadrante Inferior Esquerdo na Pastagem (RAM de 48 Bits)

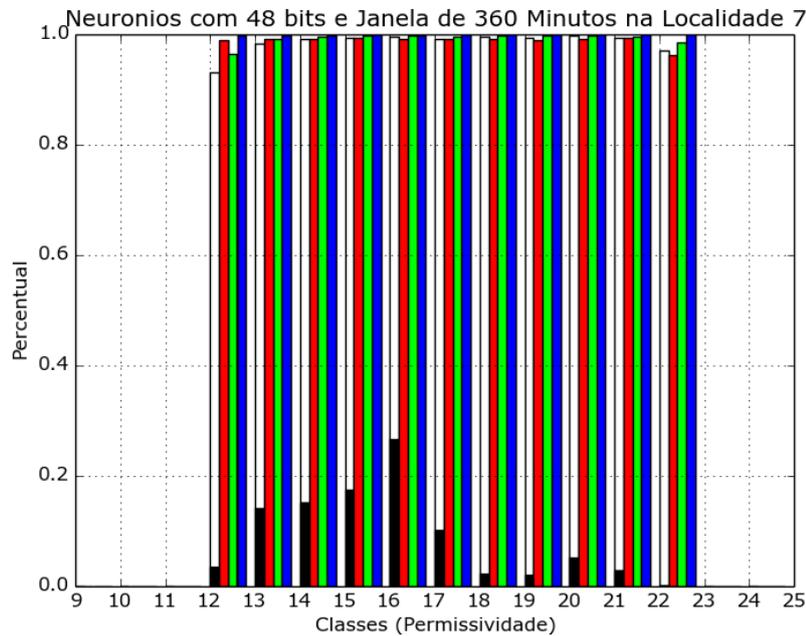


Figura 6.37: Gráfico B do Quadrante Inferior Esquerdo na Pastagem (RAM de 48 Bits)

Resumo

Os resultados ilustrados pelos gráficos para o terreno com cobertura de Pastagem são condizentes ao que buscou-se neste trabalho em todas as métricas utilizadas, exibindo resultados acima de 90% inclusive para classes com baixas proporções no conjunto. Já os gráficos da Mata ilustram que ainda tem margem para melhores resultados, as classes que representam maior umidade no solo em geral tem ótimos resultados mesmo para classes menos presentes no conjunto de dados de treinamento e teste, porém as classes que representam um solo mais seco em geral tem resultados não condizentes como o caso da classe 12 que atinge pouco menos de 60% de acerto em suas previsões.

6.4.4 RAMs com 64 bits

Apesar dos resultados obtidos com as RAMs de 48 bits parecerem dentro do esperado, foram feitos testes com RAMs de 64 bits, os testes como pode-se observar os resultados abaixo tiveram resultados aceitáveis para este trabalho, os incrementos nos resultados foram menores e seguindo uma tendência. O tempo de treinamento

ficou abaixo também como era esperado, com hardware melhor é possível reduzir ainda mais este tempo podendo ficar abaixo dos cinco minutos, o tempo de uma janela de dados, assim se existir a necessidade de de treinar novamente a rede, poderá ser feito no intervalo entre uma medida e outra dos termômetros e pluviômetros.

Abaixo seguem a tabela com os resultados gerais para este teste, os gráficos e logo em seguida o resumo final para esta seção.

	Mata	Pastagem
Tempo de Treinamento	426 Segundos	443 Segundos
Média Percentual de Revocação	98,84%	98,95%
Desvio Padrão	0,14%	0,18%

Tabela 6.4: Resultados para RAMs com 64 Bits

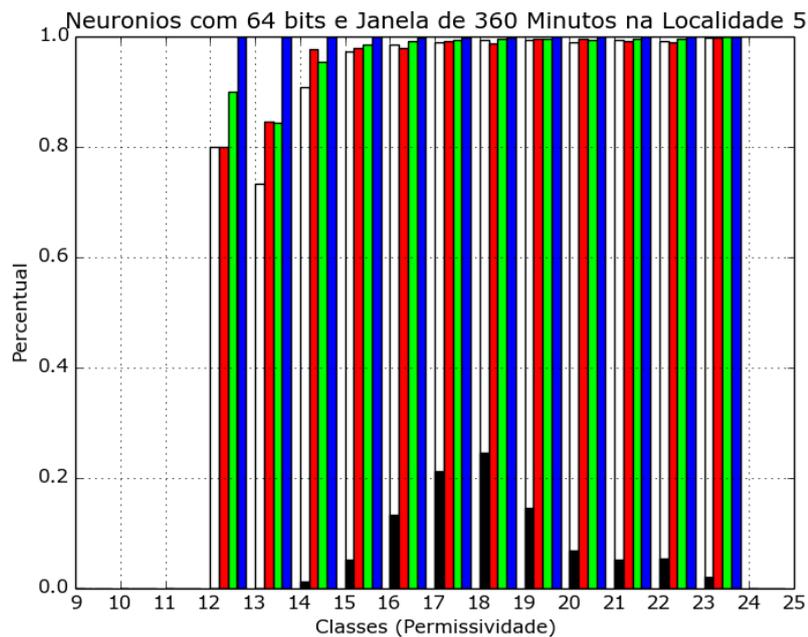


Figura 6.38: Gráfico A Quadrante Superior Esquerdo na Mata (RAM de 64 Bits)

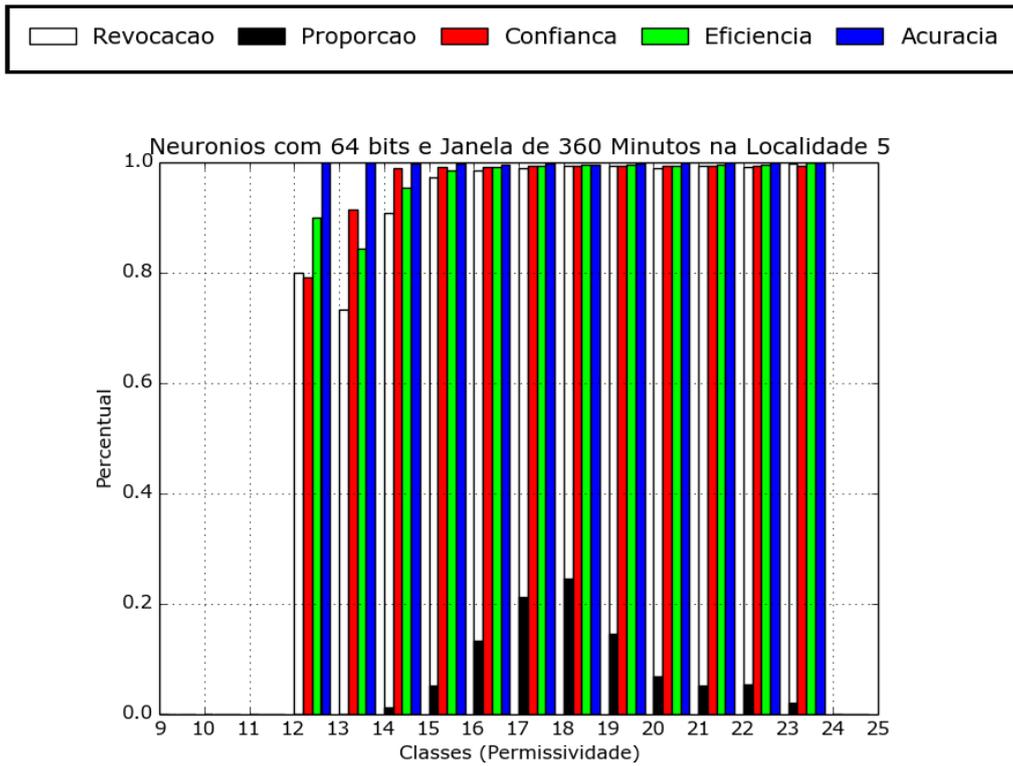


Figura 6.39: Gráfico B do Quadrante Superior Esquerdo na Mata (RAM de 64 Bits)

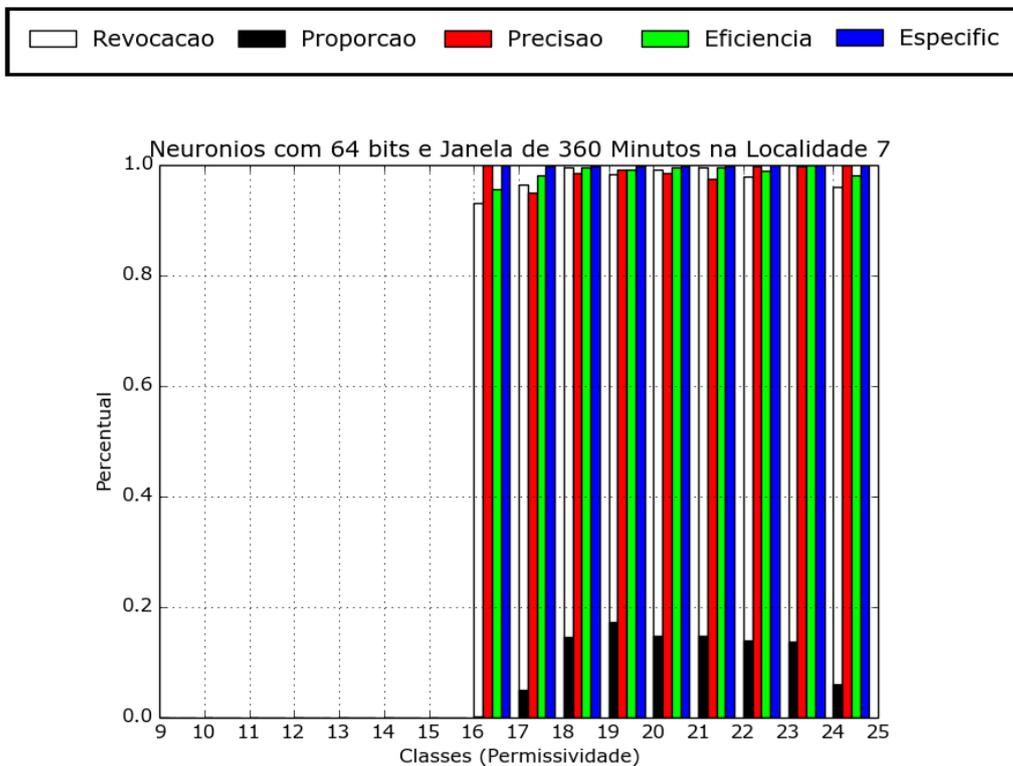


Figura 6.40: Gráfico A Quadrante Inferior Esquerdo na Mata (RAM de 64 Bits)

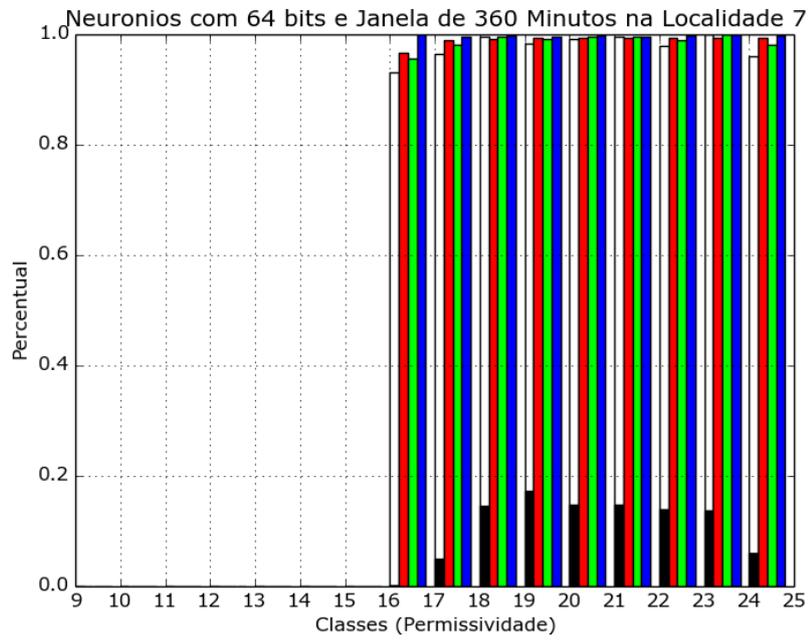


Figura 6.41: Gráfico B do Quadrante Inferior Esquerdo na Mata (RAM de 64 Bits)

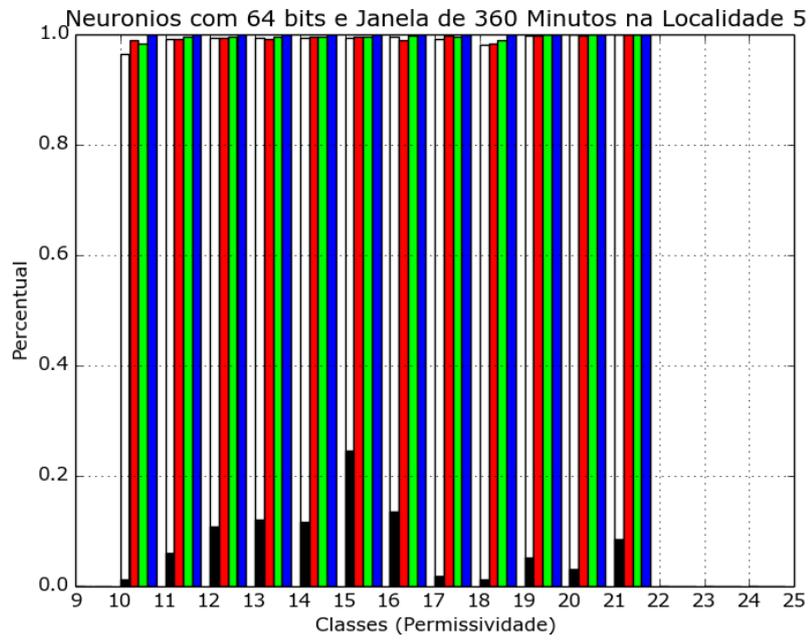


Figura 6.42: Gráfico A Quadrante Superior Esquerdo na Pastagem (RAM de 64 Bits)

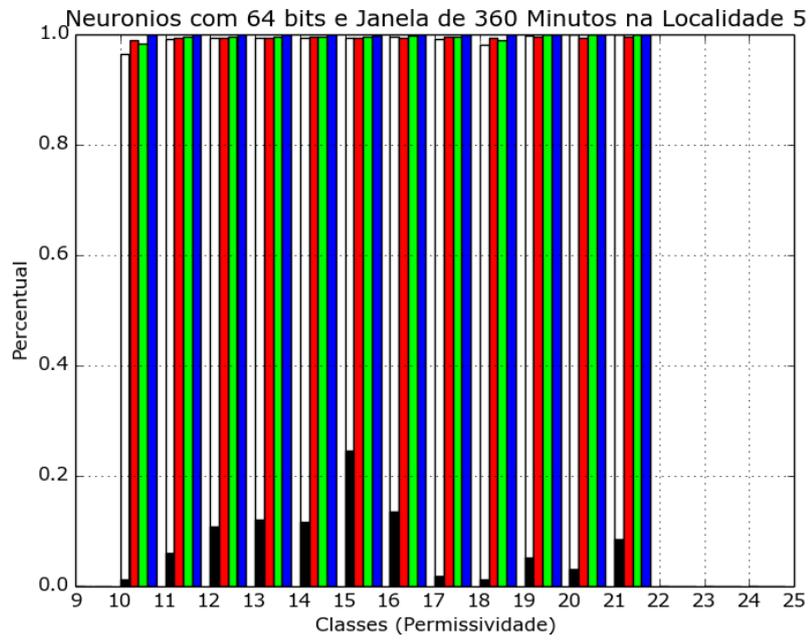


Figura 6.43: Gráfico B do Quadrante Superior Esquerdo na Pastagem (RAM de 64 Bits)

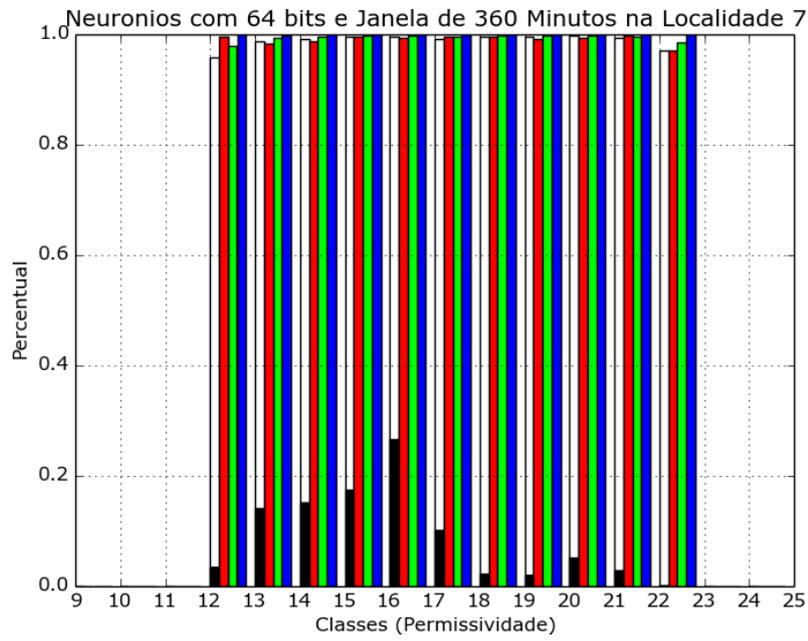


Figura 6.44: Gráfico A Quadrante Inferior Esquerdo na Pastagem (RAM de 64 Bits)

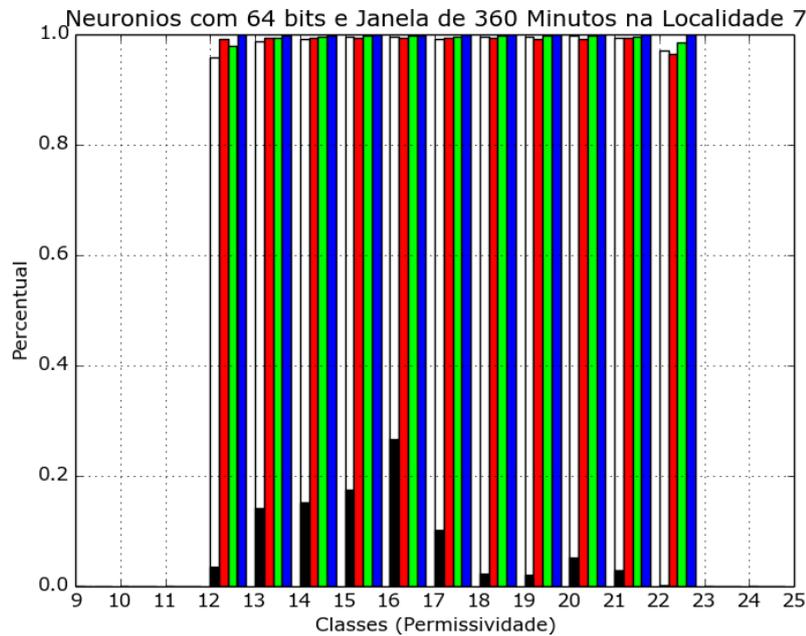


Figura 6.45: Gráfico B do Quadrante Inferior Esquerdo na Pastagem (RAM de 64 Bits)

Resumo

Nos resultados é possível notar uma melhora menor que em resultados anteriores, dessa forma ficou decidido que não havia mais necessidade para testes envolvendo RAMs com outros tamanhos em bits. Os resultados para Pastagem foram todos condizentes com o que buscamos, sendo possível notar que em todas as métricas obteve-se acima de 95%. Nos teste para dados da Mata, na classe 12 ainda obteve-se melhor resultado que os anteriores, pode-se considerar estes mesmo estando pouco acima dos 80%, esta classe possui quantidades tão baixas no conjunto de dados que chegar a um resultado como este é um avanço importante.

6.5 Variações no Tamanho das Janelas

Na seção anterior os testes abordaram a variação no tamanho das RAMs, chegando a uma conclusão dado os resultados destes teste que, para a WiSARD não generalizar os resultados de suas previsões a melhor solução é aumentar o tamanho no espaço de endereçamento de uma única RAM, assim chegou-se ao resultado de 64 bits como coerentes ao que buscamos. Agora é preciso averiguar o melhor tamanho

de uma janela de dados, para isso realizou-se um conjunto de testes avaliando quatro possíveis janelas.

Os testes que se seguirão nesta seção utilizaram parte dos resultados dos testes anteriores, usando apenas RAMs com tamanho de 64 bits, serão apresentados gráficos com resultados para janelas de cinco minutos, janelas de trinta minutos, de duas horas e finalmente de seis horas. A escolha das janelas não foram aleatórias, decidiu-se utilizar a janela mínima de cinco minutos como teste de controle e para entender de fato a importância de históricos maiores, estes tem resultados melhores ou no mínimo iguais. Decidiu-se utilizar janelas de trinta minutos para contemplar a literatura da área da Geotécnica, como pôde ser observado no capítulo quatro deste trabalho, todos os primeiros testes foram utilizado esta janela observando o Padrão I30, intensidade de chuva em 30 minutos.[8]

Para os testes em janelas de duas e seis horas, a decisão ficou a cargo do tempo que leva para a água infiltrar e mudar a condição do solo, como é possível observar em [7][19], em alguns tipos de solos é possível que leve horas para a água atingir níveis mais profundos no solo por conta deste tipo de solo possuir escoamento lateral, em outros casos o escoamento é vertical fazendo com que a água consiga atingir estes níveis mais rapidamente. Para o tipo de solo e cobertura de vegetação que este trabalho teve acesso, considerou-se estas janelas máximas adequada, porém é notório que diferentes tipos de solo podem obter resultados melhores com janelas maiores ou até mesmo intermediárias à estas produzidas nos testes.

6.5.1 Janela de Cinco Minutos

Será tratado aqui um conjunto de testes de controle para averiguar o crescimento em eficiência das previsões realizadas pelo sistema, janelas de cinco minutos foram utilizadas nestes testes, com RAMs de 64 bits, ainda são exibidos teste dos mesmos quatro quadrantes dos testes anteriores, Superior Esquerdo e Inferior Esquerdo na Mata e Superior Esquerdo e Inferior Esquerdo na Pastagem. O Formato de apresentação dos resultados é o mesmo das seções anteriores, uma Tabela mostrando os resultados de Tempo de Treinamento, Média de Revocação e Desvio Padrão desta média, logo em seguida todos os gráficos dos testes para só então discutir os resultados em um resumo.

	Mata	Pastagem
Tempo de Treinamento	213 Segundos	209 Segundos
Média Percentual de Revocação	88,24%	86,56%
Desvio Padrão	0,64%	0,56%

Tabela 6.5: Resultados para Janelas de Cinco Minutos

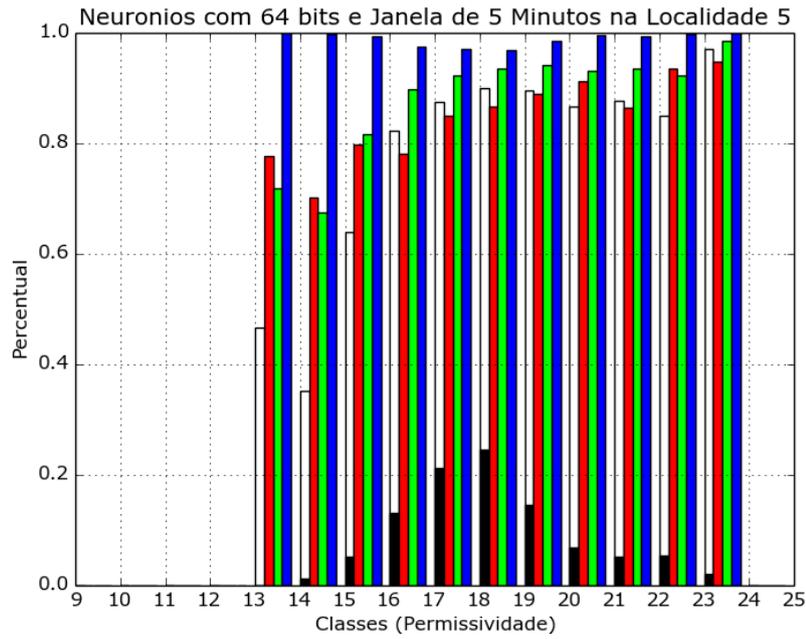


Figura 6.46: Gráfico A Quadrante Superior Esquerdo na Mata (Janela 5 Minutos)

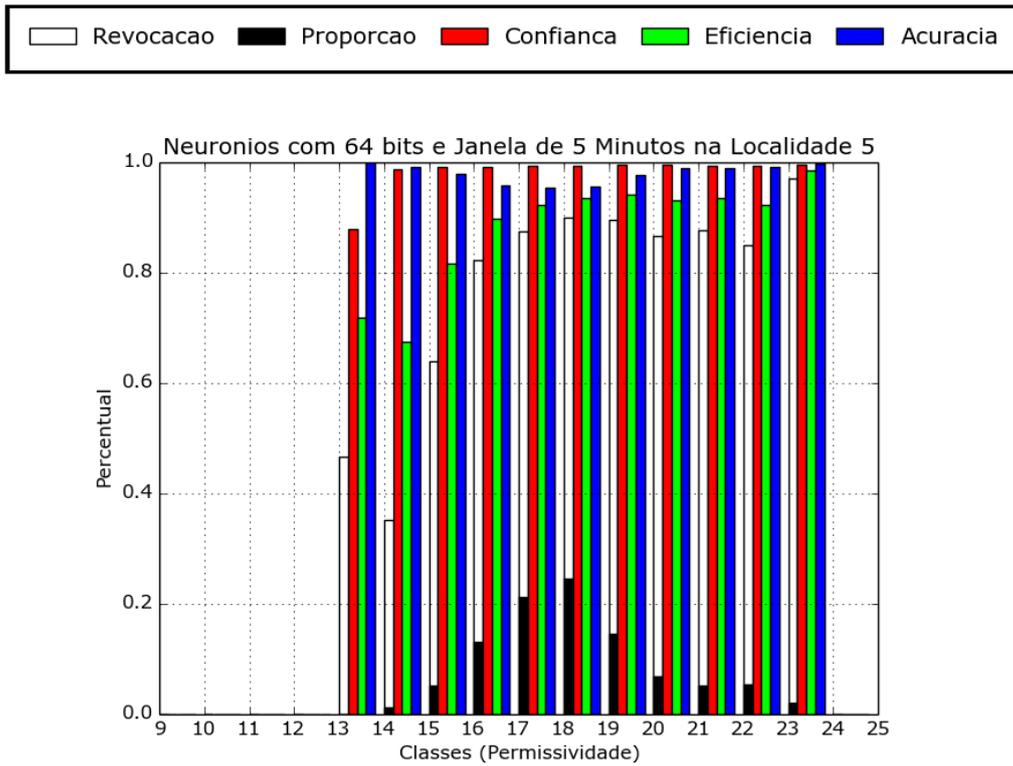


Figura 6.47: Gráfico B do Quadrante Superior Esquerdo na Mata (Janela 5 Minutos)

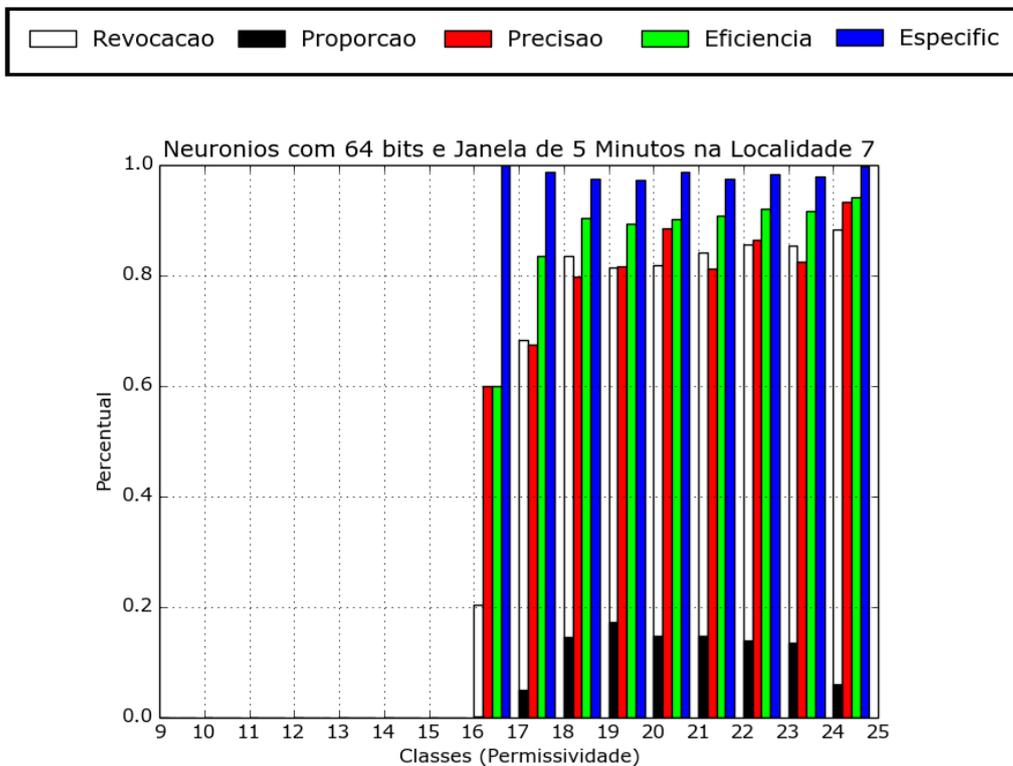


Figura 6.48: Gráfico A Quadrante Inferior Esquerdo na Mata (Janela 5 Minutos)

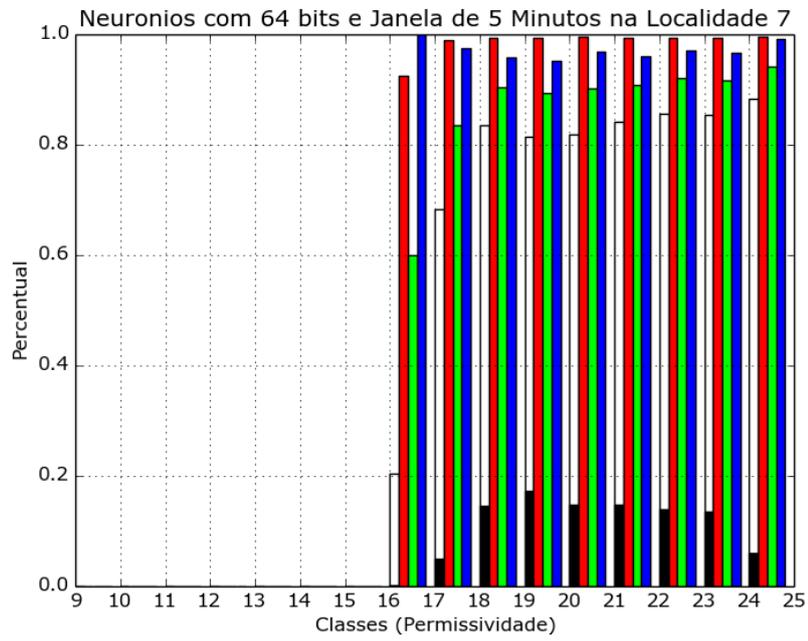


Figura 6.49: Gráfico B do Quadrante Inferior Esquerdo na Mata (Janela 5 Minutos)

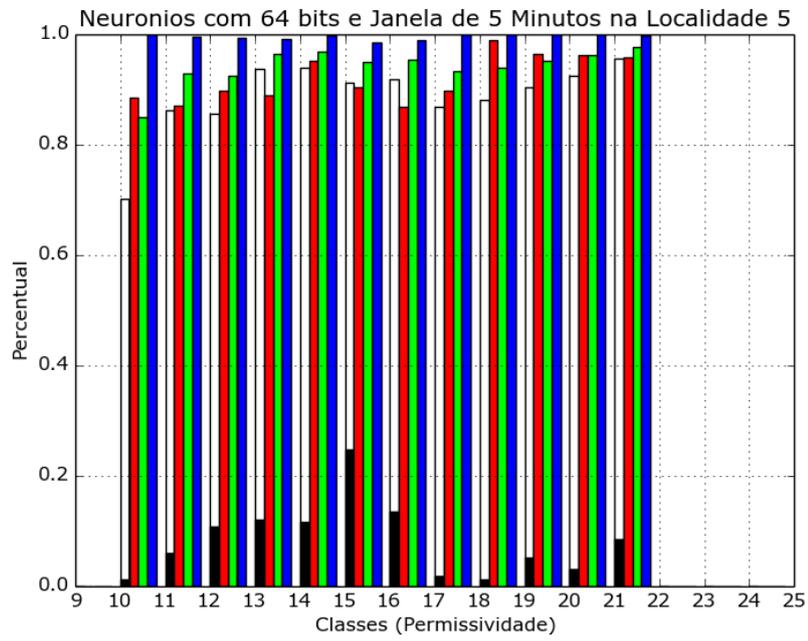


Figura 6.50: Gráfico A Quadrante Superior Esquerdo na Pastagem (Janela 5 Minutos)

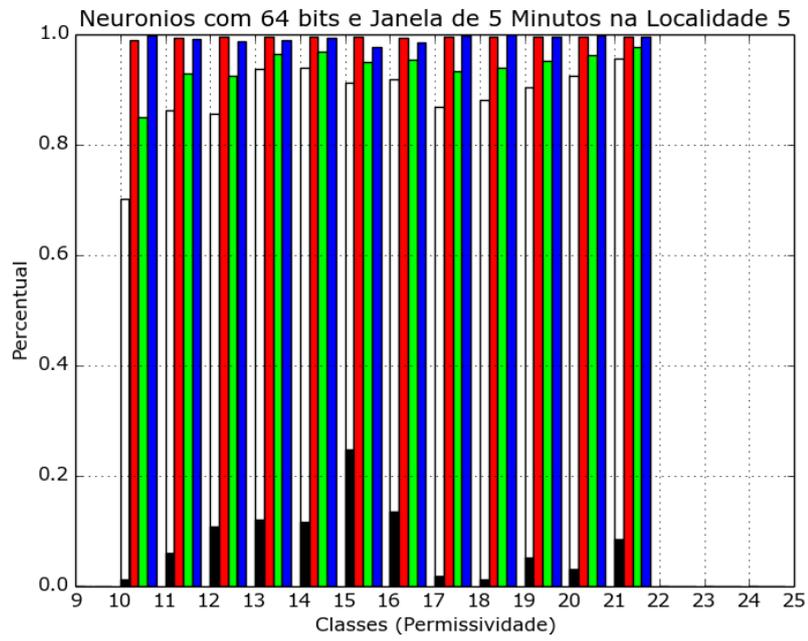


Figura 6.51: Gráfico B do Quadrante Superior Esquerdo na Pastagem (Janela 5 Minutos)

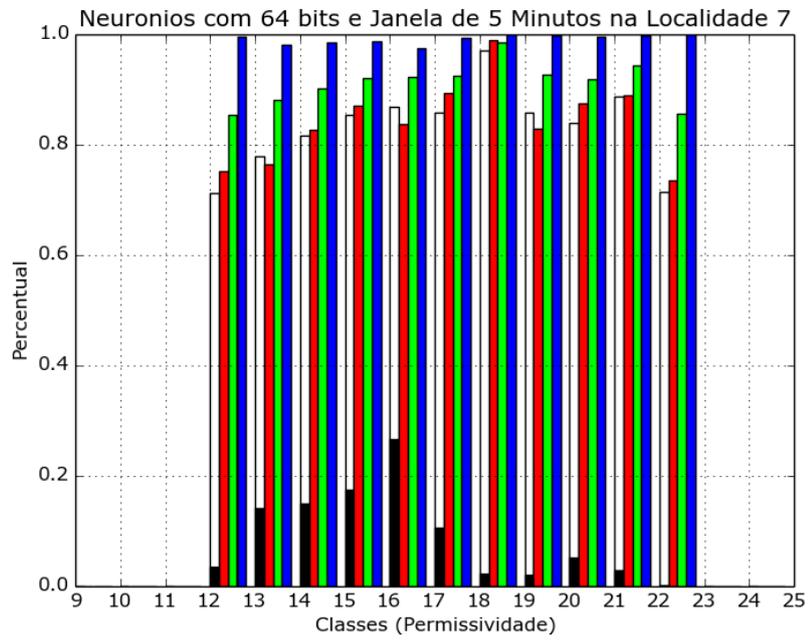


Figura 6.52: Gráfico A Quadrante Inferior Esquerdo na Pastagem (Janela 5 Minutos)

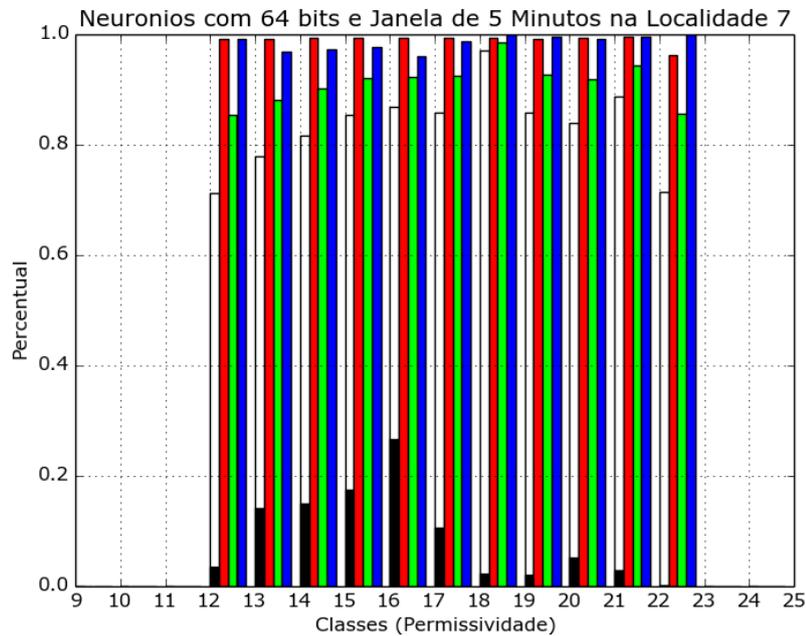


Figura 6.53: Gráfico B do Quadrante Inferior Esquerdo na Pastagem (Janela 5 Minutos)

Resumo

Neste teste já era esperado resultados baixos em relação aos melhores resultados obtidos anteriormente, foi ilustrado como teste de controle para maior percepção de como evoluem as respostas da WiSARD de acordo com o aumento das janelas. É possível observar que mesmo estes resultados podem ser considerados aceitáveis, com altos índices atingidos na maioria das métricas utilizadas, principalmente nos dados referentes a Pastagem, mesmo em classes menos presentes como o caso da classe 10 que atingiu mais de 70% de Revocação, evidenciando a importancia de encontrar um tamanho adequado para as RAMs dos discriminadores WiSARD.

No caso dos testes para o terreno coberto por Mata, os resultados mostram ainda uma grande necessidade de melhora, os vetores que influenciam na umidade, tais quais escoamento e evapotranspiração, tem influencia muito maior neste tipo de terreno, gerando a necessidade de aumentar as janelas para a WiSARD ter maior diferenciação das classes, principalmente as classes mais secas, este terreno tende a perder umidade muito rapidamente o que pode gerar muita ambiguidade dos atributos criando resultados de previsão incorretos.

6.5.2 Janela de Trinta Minutos

Abaixo resultados dos testes realizados com Janela de Trinta Minutos, satisfazendo o Padrão I30 que é o calculo feito para intensidade de chuva nos ultimos 30 minutos, este tipo de calculo ajuda a entender qual a taxa de erosividade da chuva, para o impacto na unidade do solo é importante entender que um valor alto deste indice I30 fará com que a água precipitada não infiltre no solo, apenas escoe o que já foi discutido no capítulo quatro deste trabalho ao falar do Padrão I30. Realizar testes utilizando esta janela tem importancia direta para este trabalho mesmo sendo o Padrão I30 mais utilizado para taxas de erosividade de precipitações. É importante notar também que este conhecimento será passado para rede através do indice pluviométrico de forma indireta no atributo Período de Seca como já mencionado anteriormente.[8]

	Mata	Pastagem
Tempo de Treinamento	244 Segundos	246 Segundos
Média Percentual de Revocação	92,60%	92,53%
Desvio Padrão	0,43%	0,42%

Tabela 6.6: Resultados para Janelas de Trinta Minutos

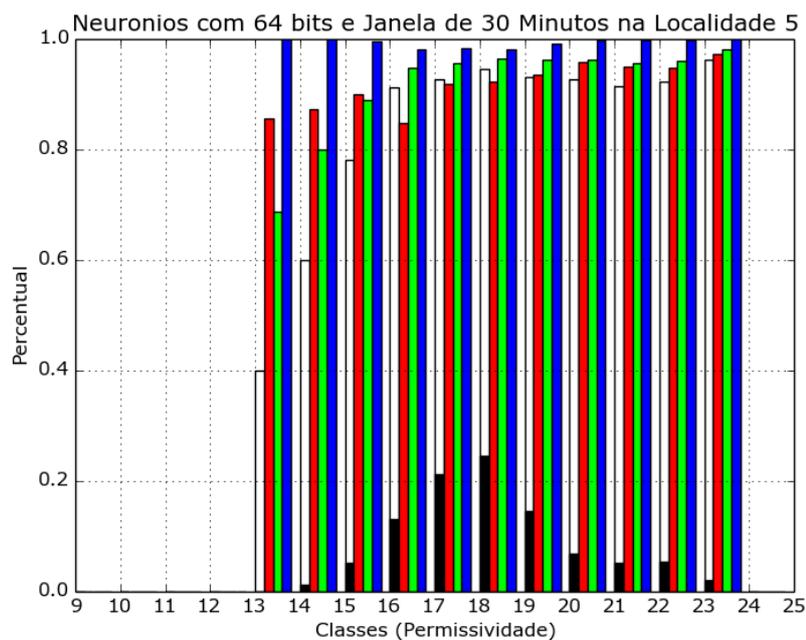


Figura 6.54: Gráfico A Quadrante Superior Esquerdo na Mata (Janela 30 Minutos)

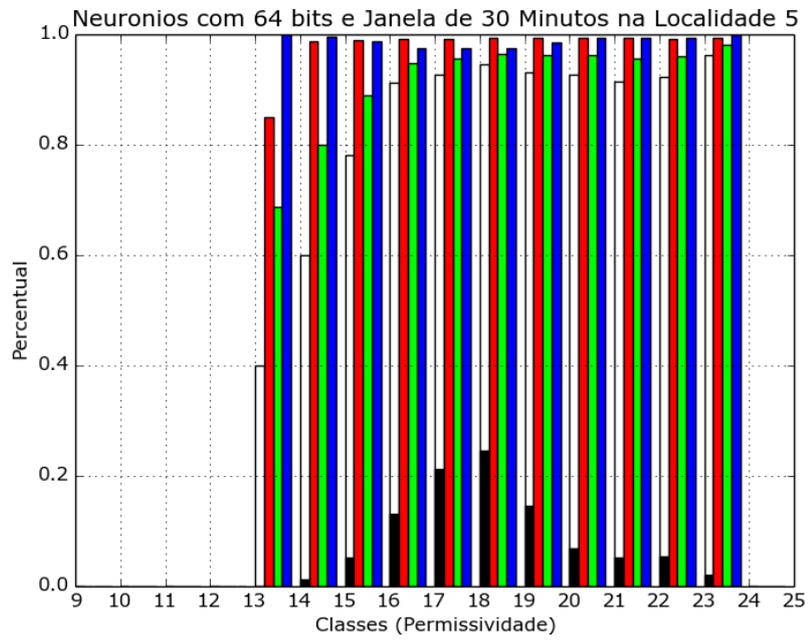


Figura 6.55: Gráfico B do Quadrante Superior Esquerdo na Mata (Janela 30 Minutos)

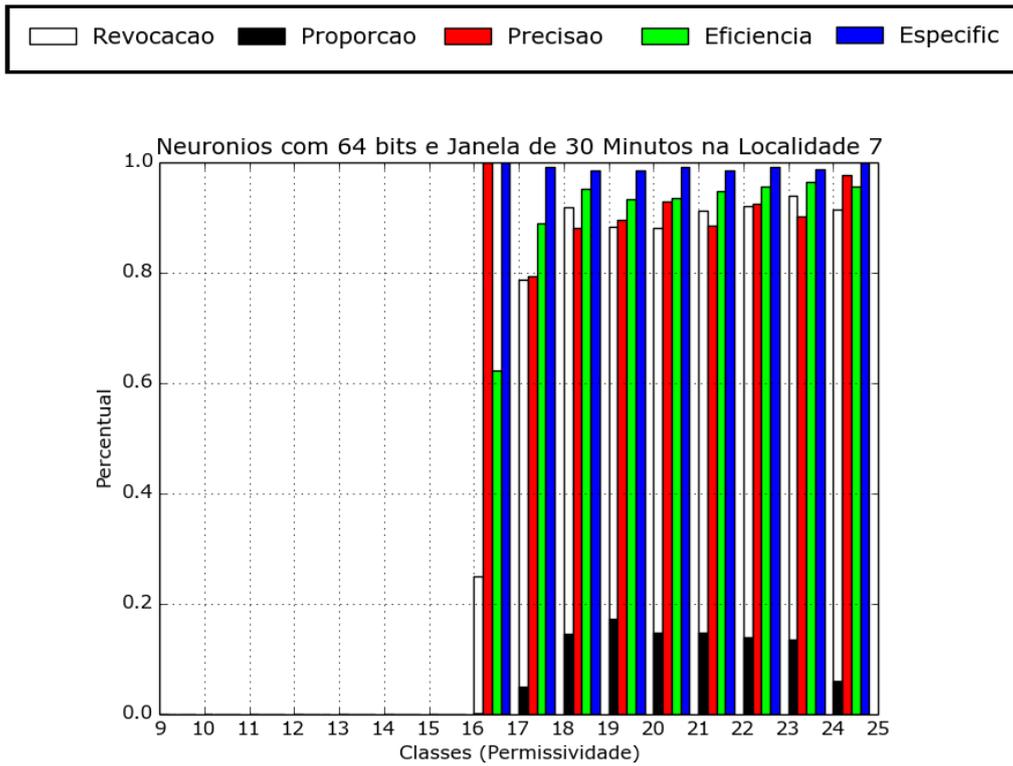


Figura 6.56: Gráfico A Quadrante Inferior Esquerdo na Mata (Janela 30 Minutos)

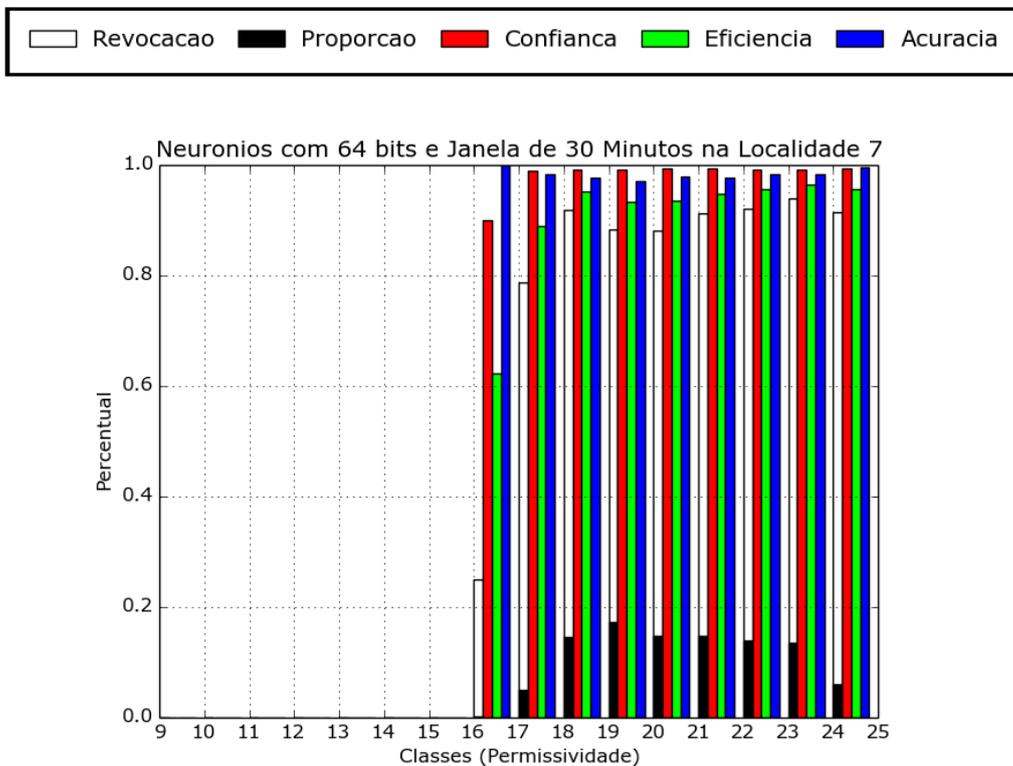


Figura 6.57: Gráfico B do Quadrante Inferior Esquerdo na Mata (Janela 30 Minutos)

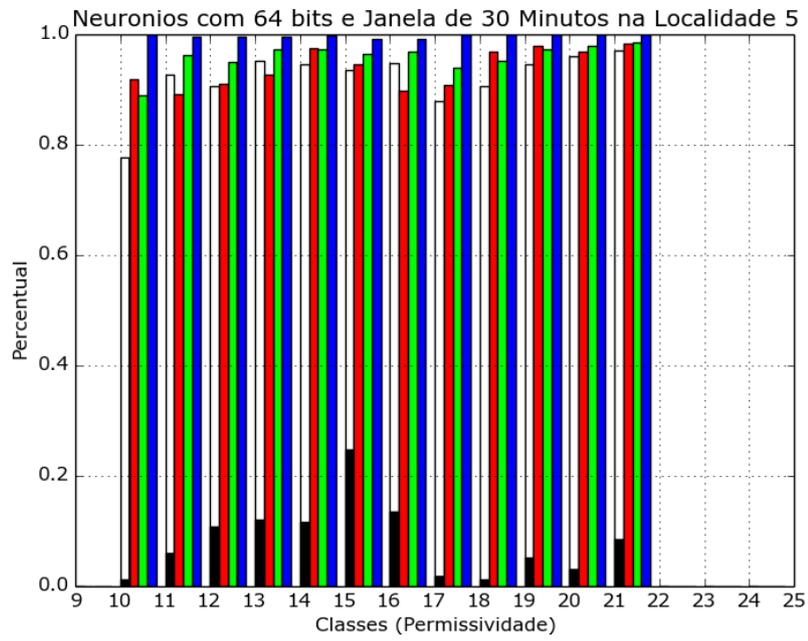


Figura 6.58: Gráfico A Quadrante Superior Esquerdo na Pastagem (Janela 30 Minutos)

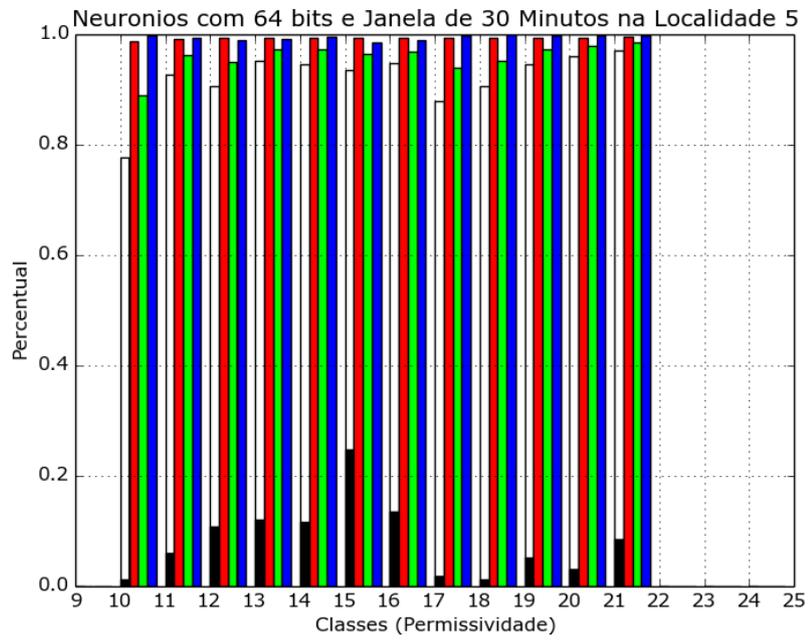


Figura 6.59: Gráfico B do Quadrante Superior Esquerdo na Pastagem (Janela 30 Minutos)

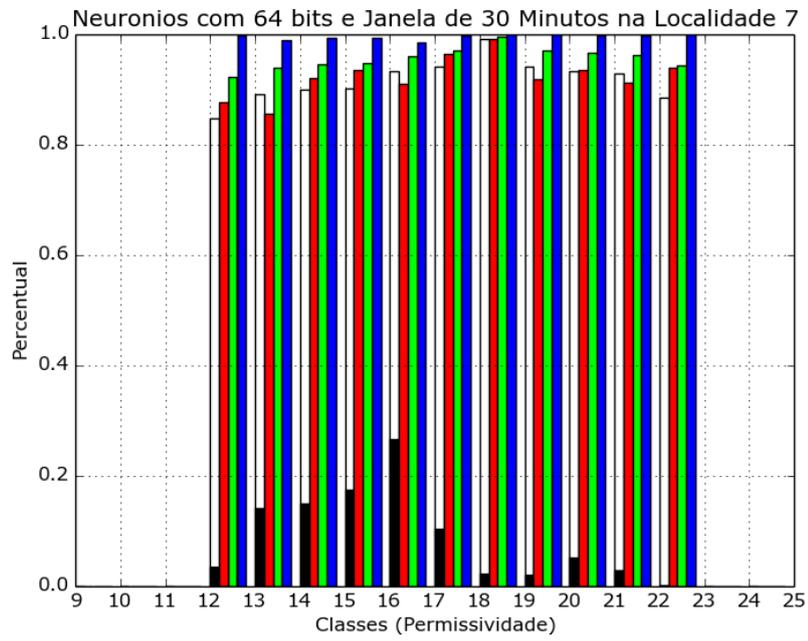


Figura 6.60: Gráfico A Quadrante Inferior Esquerdo na Pastagem (Janela 30 Minutos)

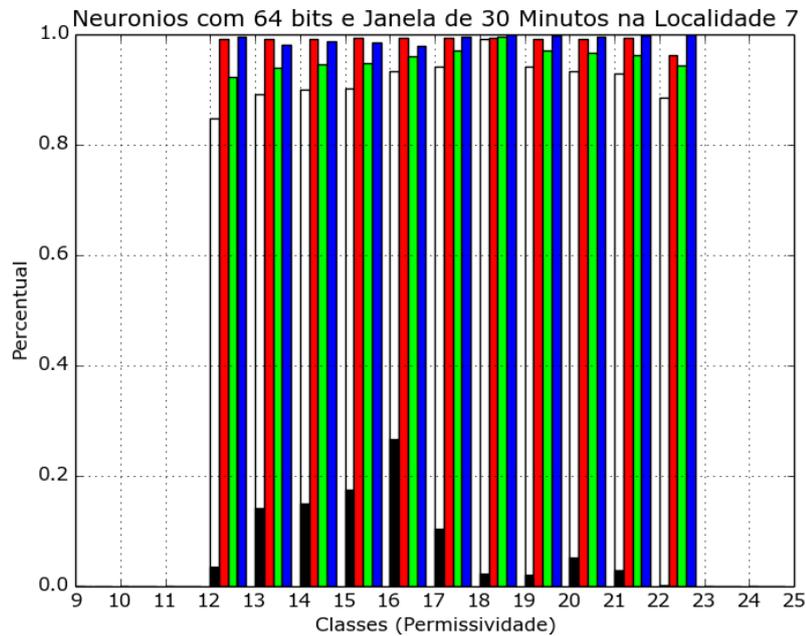


Figura 6.61: Gráfico B do Quadrante Inferior Esquerdo na Pastagem (Janela 30 Minutos)

Resumo

Os resultados para os testes com os dados do terreno com cobertura de Pastagem já mostram resultados dentro do esperado, classes pouco presentes com alto índice de Revocação e Eficiencia, resultado geral acima em média acima dos 90%, baixos valores de Desvio padrão desta média por volta de 0,4%, é possível notar que o tempo de treinamento aumenta com o tamanho da janela, isso é natural visto que estes valores aumentam diretamente o tamanho do *input* o que tem impacto direto na quantidade de RAMs já que o tamanho delas estão fixados em 64 bits, já vimos antes que maior quantidade de RAMs por discriminadores aumentam a quantidade de iterações quando ocorre o treinamento da WiSARD, logo este aumento gradual do tempo de treinamento é compreensivo.

No caso dos resultados para Mata, pôde-se notar que são resultados aquém do esperado, sendo possível melhorar outros testes foram realizados para atingir melhores níveis também para os esta cobertura de terreno.

6.5.3 Janela de Duas Horas

Notou-se durante estudo e alguns experimentos de ensaios de laboratório [3], que os solos monitorados tem baixa taxa de infiltração, chegando a levar de 2 à 6 horas para alcançar os sensores do TDR localizados à trinta centímetros. Com a percepção deste comportamento, ficou clara a importancia em realizar testes com janelas de tempo que abrangesse estes periodos, nesta subseção serão exibidos testes realizados com Janelas de Duas Horas.

	Mata	Pastagem
Tempo de Treinamento	269 Segundos	267 Segundos
Média Percentual de Revocação	96,71%	96,74%
Desvio Padrão	0,25%	0,34%

Tabela 6.7: Resultados para Janelas de Duas Horas

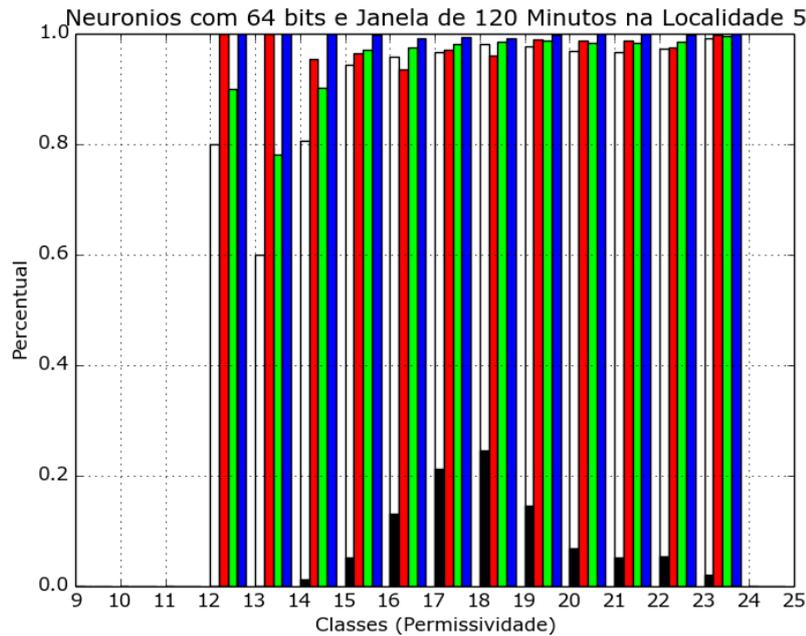


Figura 6.62: Gráfico A Quadrante Superior Esquerdo na Mata (Janela 2 Horas)

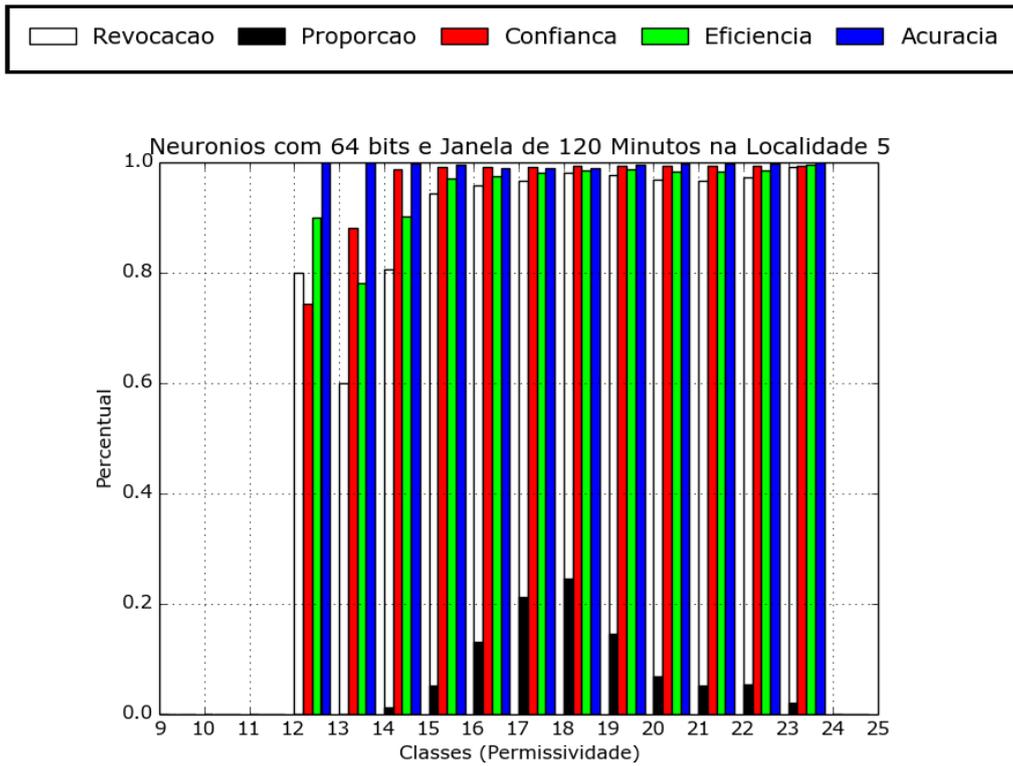


Figura 6.63: Gráfico B do Quadrante Superior Esquerdo na Mata (Janela 2 Horas)

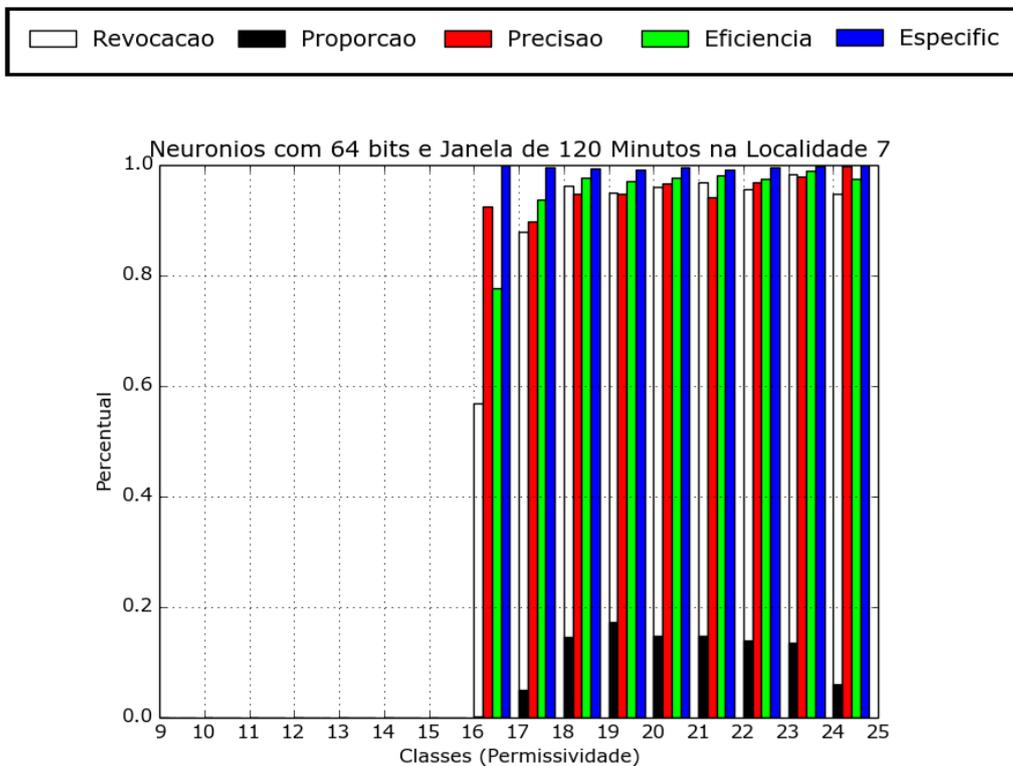


Figura 6.64: Gráfico A Quadrante Inferior Esquerdo na Mata (Janela 2 Horas)

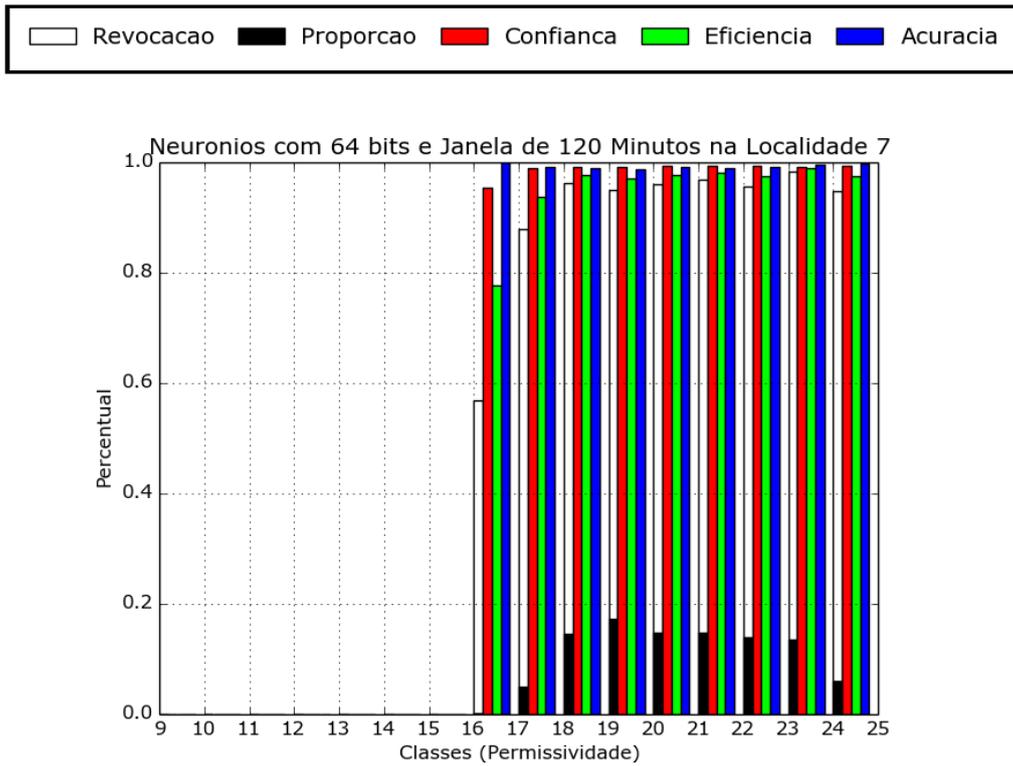


Figura 6.65: Gráfico B do Quadrante Inferior Esquerdo na Mata (Janela 2 Horas)

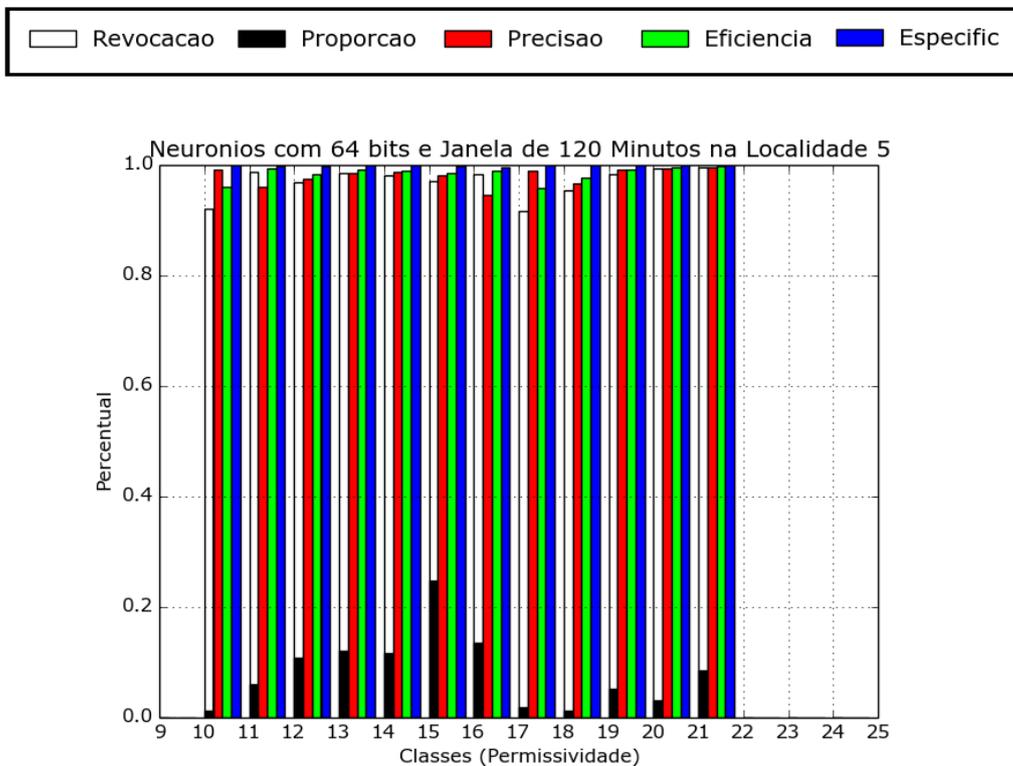


Figura 6.66: Gráfico A Quadrante Superior Esquerdo na Pastagem (Janela 2 Horas)

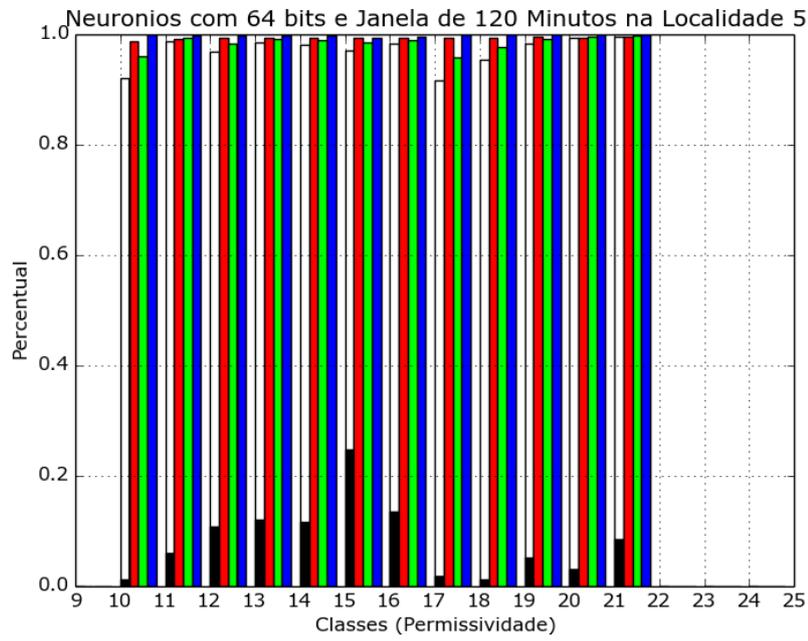


Figura 6.67: Gráfico B do Quadrante Superior Esquerdo na Pastagem (Janela 2 Horas)

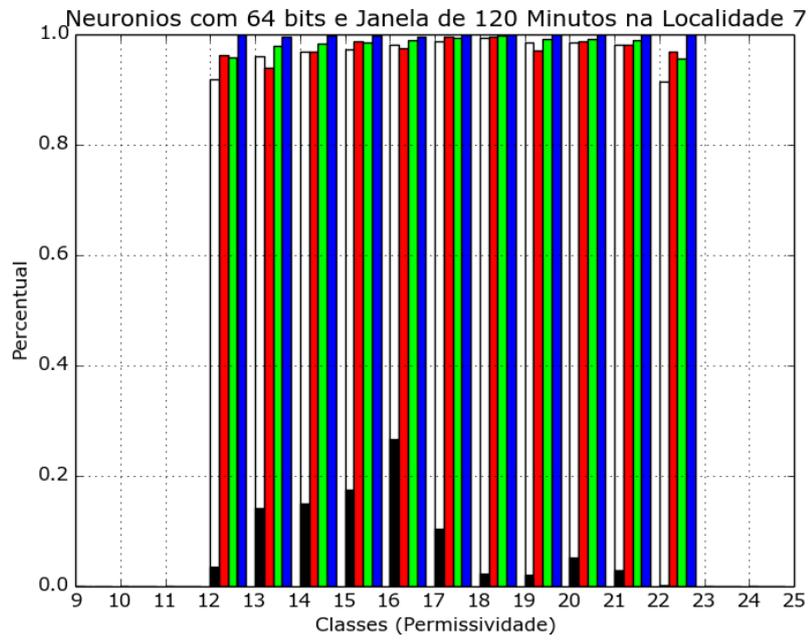


Figura 6.68: Gráfico A Quadrante Inferior Esquerdo na Pastagem (Janela 2 Horas)

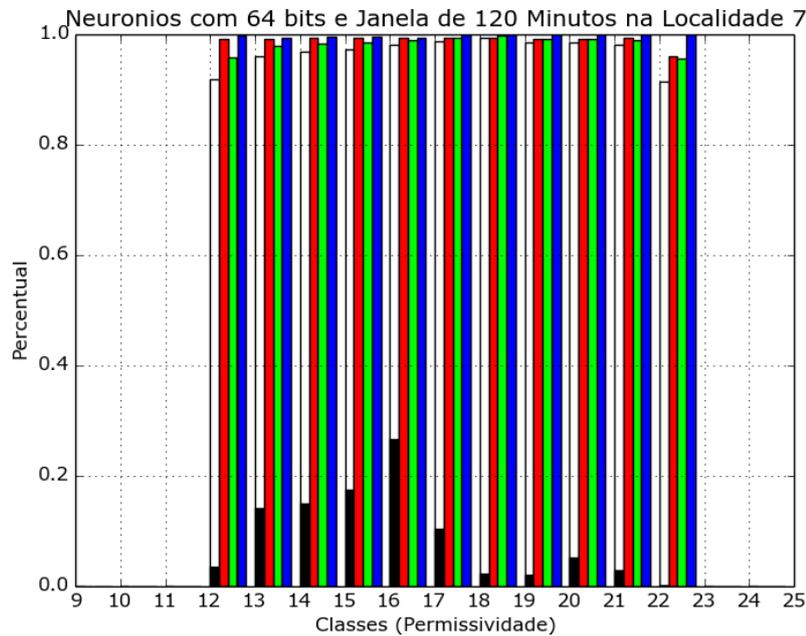


Figura 6.69: Gráfico B do Quadrante Inferior Esquerdo na Pastagem (Janela 2 Horas)

Resumo

Entre os conjuntos de dados, Pastagem e Mata, os resultados para Janelas de Duas Horas para Pastagem como de costume tendem a ser melhores. Ambos os resultados atingiram o que se esperava para este trabalho, resultados com índices acima de 90%. Estes resultados fazem parte dos últimos testes, já que os resultados para Seis Horas já foram exibidos anteriormente nas variações de tamanho de RAMs.

Note, ainda que os resultados da Mata são considerados razoáveis e dentro do que se espera, classes com menor proporção no conjunto de dados ainda apresentam resultados abaixo do ideal, outro bom indício que é possível melhorar incrementando ainda mais os resultados dos testes, comparando os resultados das Janelas de Duas Horas e de Seis Horas, mesmo este último levando em média 60% mais tempo para treinar é mais compatível com os resultados almejados por este trabalho.

Capítulo 7

Conclusões e Trabalhos Futuros

Como visto nos capítulos anteriores, este trabalho teve por finalidade a criação de um sistema de predição de umidade do solo, que possa ser utilizado em qualquer ambiente, em qualquer tipo de solo com qualquer cobertura. Vimos que o sistema se torna viável bastando ter um conjunto inicial de informações para uso no treinamento da rede WiSARD, descartando assim o uso de equipamentos caros para manter o monitoramento.

Os resultados da WiSARD, seguindo métricas adotadas para avaliá-los tiveram resultados dentro do esperado, com índice de acerto geral próximo de 99% e desvio padrão destes acertos abaixo de 0,2%. Notou-se também a facilidade de calibrar a rede no que diz respeito ao índice de permissividade e classes da rede, visto que estes resultados são obtidos pelo sensor que fará o monitoramento inicial para transmitir estas informações à WiSARD, os resultados podem ser utilizados sem recalibrar todo o sistema devido a troca de ambiente, ficando o trabalho de calibrar o resultado em permissividade para VWC delegado à um geotécnico que terá todo ferramental necessário para esta função.

Ficou evidente nos testes que a Temperatura tem correlação direta com a umidade do solo, em condições normais, que o Período de Seca gera um incremento considerável na melhoria da previsão assim como janelas de entradas maiores, pelo fato do atributo Período de Seca expressar melhor a intensidade do índice pluviométrico, intensidade esta que pode determinar que a água precipitada está de fato infiltrando no solo aumentando assim sua umidade ou simplesmente escoando diminuindo o efeito de um evento de chuva. É possível entender o contrário deste efeito, uma evento de chuva que dure o dia inteiro terá um impacto maior na umidade mesmo tendo este evento um volume menor de água do que um evento de chuva que dura apenas dez minutos.

7.1 Trabalhos Futuros

Durante o trabalho notou-se a possibilidade de melhorias do sistema coletando maior quantidade de atributos, utilizando o sistema para prever erosão, ou até mesmo prever Deslizamentos em Encostas. O trabalho mostrou grande potencial aberto para grande variedade de finalidades possíveis, pelo fato de ter um índice de acerto tão elevado é possível gerar outros sistemas confiáveis a partir deste trabalho.

7.1.1 Sistema de Predição de Deslizamentos em Encostas

Antes do sistema se tornar um Previsor de Umidade, intencionou-se criar um Sistema de Predição de Deslizamentos. Um sistema de utilidade pública que pode ajudar a salvar vidas monitorando encostas e acionando alarmes quando houver o risco iminente de deslizamento dado um índice pluviométrico previsto pela meteorologia. Viu-se a necessidade de dados em que estes eventos de deslizamentos ocorrem com dados de índices pluviométricos atrelados a eles. É possível utilizar dados coletados em laboratório para este fim, ou mesmo dados naturais obtidos em localidades de grande incidência destes eventos como Angra dos Reis e Friburgo no Rio de Janeiro. Como trabalho futuro e até mesmo extensão deste trabalho é possível criar um sistema para esta finalidade inclusive utilizando todo este trabalho aqui apresentado. Prever a umidade do solo é um passo importante para prever deslizamentos, é preciso prever a umidade em vários níveis do solo e não somente nas camadas superficiais, o que não é feito neste trabalho, porém bastando obter resultados de outras profundidades é possível ter resultados similares aos apresentados neste trabalho.

7.1.2 Previsão de Erosão

Outro objetivo deste trabalho foi tentar prever a erosão gerada por um evento de chuva, resultados deste tipo de previsão poderá ajudar em várias partes da Geotécnica e Agronomia [20], analisar o desgaste do solo, perda de massa e superfície. Para este tipo de previsão se faz necessário um conjunto de dados adequado, onde eventos de chuva ocasione diferentes tipos de erosão para que possa ser gerado classes realmente diferentes umas das outras, o que não ocorreu nos dados de Itaboraí, estes dados, apesar de ter eventos de chuva com grande índice pluviométrico, não gerou erosão suficiente nos eventos para diferenciação dos resultados. Do contrário seria possível prever quantidade de massa erodida e

inclusive percentuais de material em relação à composição do solo monitorado.

7.1.3 Inclusão de Novos Atributos

Durante o trabalho verificou-se a possibilidade de inclusão de novos atributos os quais poderiam melhorar ainda mais os resultados da rede e inclusive poderiam ser utilizados para os outros sistemas propostos para trabalhos futuros como o caso do Sistema de Predição de Deslizamentos.

- **Temperatura da Chuva:**

Verificou-se durante o trabalho a correlação entre a temperatura do solo e sua umidade, existe ainda a possibilidade de instalar em pluviômetros termômetros para armazenar também dados da temperatura da chuva, a temperatura do solo certamente vai cair mais com chuvas mais geladas o que não está diretamente ligado com um volume maior de água no solo, Este novo atributo pode refinar ainda mais os resultados da WiSARD.

- **Insolação na área sensoriada:**

Quando foram avaliados os dados do *data log*, em dias em que ocorrem eventos de precipitação, notou-se uma elevação da temperatura do solo menos acentuada, o que sugere período nublado. Poder quantificar o nível de radiação solar recebida na área sensoriada ajudaria a entender ainda mais a correlação entre Temperatura e Umidade e pode gerar um incremento considerável nos resultados da previsão.

Referências Bibliográficas

- [1] C. F. A. TEIXEIRA, S. O. MORAES, M. A. S. “Desempenho do Tensiômetro, TDR e Sonda de Nêutrons na Determinação da Umidade e Condutividade Hidráulica do Solo”, fev. 2005.
- [2] K. TERZAGHI, R. B. PECK, G. M. *Soil Mechanics in Engineering Practice*. 3 ed. Nova Iorque, John Wiley Sons, 1996.
- [3] NACINOVIC, M. G. G. *Avaliação de Erosão Hídrica Ssuperficial em Parcelas Experimentais*. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2013.
- [4] LIBARDI, P. L. *Dinâmica da água no solo*. 1 ed. Piracicaba, EdUSP, 1995.
- [5] B.P. MOHANTY, T. S. “Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation”, *Journal of Elsevier*, fev. 2001.
- [6] DALMAGO, G. A. *Dinâmica da Água no Solo em Cultivos de Milho Sob Plantio Direto e Preparo Convencional*. Tese de D.Sc., UFRGS, Porto Alegre, RS, Brasil, 2004.
- [7] W. D. REYNOLDS, E. G. GREGORICH, W. E. C. “Characterization of water transmission properties in tilled and untilled soils using tension infiltrometers”, *Soil Tillage Research*, v. 33, n. 2, pp. 117–131, fev. 1995.
- [8] D. D. DA SILVA, S. B. PEREIRA, F. F. P. R. R. G. F. A. M. Q. L. L. G. N. B. “Equações de Intensidade-Duração-Frequência da Precipitação Pluvial Para o Estado de Tocantins”, *Engenharia de Viçosa*, v. 11, 2003.
- [9] D. R. KAISER, D. J. REINERT, J. M. R. J. P. G. M. “Dielectric Constant Obtained From TDR and Volumetric Moisture of Soils in Southern Brazil”, *SciELO*, v. 34, n. 3, mar. 2010.
- [10] J. D. F. MEDEIROS, N. CASTRO, J. A. G. R. T. C. “Calibração de Sondas do TDR em um Latossolo”, *Revista Brasileira de Recursos Hídricos*, v. 12, n. 2, pp. 19–25, jun. 2007.

- [11] JUNIOR, G. O. J. *Desempenho do Reflectômetro no Domínio do Tempo na Detecção de Variação de Umidade do Solo*. Ms.C. dissertation, USP, Piracicaba, São Paulo, Brasil, 2003.
- [12] W. S. MCCULLOCH, W. P. “A Logical Calculus of the Ideas Immanent in Nervous Activity”, *Bulletin of Mathematical Biophysics*, v. 5, n. 4, pp. 115–137, dez. 1943.
- [13] I. ALEKSANDER, H. M. *An Introduction to Neural Computing*. 1 ed. Londres, Chapman Hall, 1990.
- [14] ALAIN M. B. P. DE SILANS, L. M. W. “Dinmica da umidade de um solo da Caatinga em funo de sua condutividade trmica”, *Agriambi*, jun. 2011.
- [15] REICHARDT, K. *A água em sistemas agrícolas*. 1 ed. São Paulo, Manole, 1986.
- [16] O. T. DENMEAD, R. H. S. “Availability of soil water to plants as affected by soil moisture content and meteorological conditions”, *Agronomy Journal*, v. 54, n. 5, pp. 385–390, dez. 1962.
- [17] V. P. GORDIYENKI, P. V. K. “Patterns of moisture evaporation at variable bulk densities of the upper soil layer”, *Soviet Soil Science*, v. 22, n. 8, pp. 111–122, jan. 1990.
- [18] I. ALEKSANDER, M. DE GREGORIO, F. F. P. L. H. M. “A brief introduction to Weightless Neural Systems”, *ESANN*, pp. 299–305, abr. 2009.
- [19] FRANZLUEBBERS, A. J. “Water infiltration and soil structure related to organic matter and its stratification with depth”, *Soil Tillage Research*, v. 66, n. 2, pp. 197–205, 2001.
- [20] CUNHA, J. E. *Funcionamento hídrico e suscetibilidade erosiva de um sistema pedológico constituído por Latossolo e Argissolo no município de Cidade Gaúcha-PR*. Tese de D.Sc., USP, São Paulo, SP, Brasil, 2002.