

**UNIVERSIDADE DO VALE DO RIO DOS SINOS**  
**UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA**  
**NÍVEL MESTRADO**

**GLAUBER KISS DE SOUZA**

**AVALIAÇÃO DO EMPREGO DE RNAs PARA ESTIMATIVA DA IRRADIÂNCIA  
DIFUSA HORIZONTAL HORÁRIA**

**SÃO LEOPOLDO**  
**2018**

Glauber Kiss de Souza

AVALIAÇÃO DO EMPREGO DE RNAs PARA ESTIMATIVA DA IRRADIÂNCIA  
DIFUSA HORIZONTAL HORÁRIA

Dissertação apresentada como pré-requisito  
parcial para a obtenção do título de Mestre em  
Engenharia Mecânica, pelo Programa de Pós-  
Graduação em Engenharia Mecânica da  
UNISINOS

Orientador: Prof. Dr. Mario Henrique Macagnan  
Coorientador: Prof. Dr. Benjamin Diego François Pillot

Banca Examinadora:

Prof. Dr. Enio Bueno Pereira (INPE)  
Prof. Dr. Paulo Roberto Wander (PPGEM – UNISINOS)  
Prof. Dr. João Batista Dias (PPGEM – UNISINOS)

SÃO LEOPOLDO  
2018

S729a Souza, Glauber Kiss de.  
Avaliação do emprego de RNAs para estimativa da irradiância difusa horizontal horária / Glauber Kiss de Souza – 2018.  
95 f. : il. ; 30 cm.

Dissertação (mestrado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Engenharia Mecânica, 2018.  
“Orientador: Prof. Dr. Mario Henrique Macagnan; Coorientador: Prof. Dr. Benjamim Diego François Pillot.”

1. Energia solar. 2. Modelos difusos. 2. Radiação solar - medição. 3. Redes neurais (computação). I. Título.  
CDU 621

Dados Internacionais de Catalogação na Publicação (CIP)  
(Silvana Teresinha Dornelles Studzinski – CRB 10/2524)

## **AGRADECIMENTOS**

Agradeço ao Prof. Dr. Mario Henrique Macagnan, pela confiança depositada em meu trabalho durante todos estes anos de pesquisa conjunta; ao Prof. Dr. Benjamin D. F. Pillot, pela disponibilidade para me auxiliar durante a pesquisa, pela empolgação com que tratou o tema deste trabalho e pela sua capacidade de encorajamento; e ao Prof. Dr. Enio Bueno Pereira, cujas críticas construtivas foram importantes para a qualidade final desta dissertação.

Também deixo um grande agradecimento ao meu irmão Claiton Kiss de Souza, por ter disponibilizado várias horas de seu tempo para me auxiliar na adaptação realizada nas imagens utilizadas nesse trabalho, e ao meu irmão Higor Kiss de Souza, por seu apoio.

Finalizo com um último agradecimento sincero a todas as pessoas com as quais tive contato ao longo da vida. Cada conversa, por menor que tenha sido, foi de suma importância para a construção da minha personalidade, de meu conhecimento, da minha criticidade e da minha maneira de pensar/agir.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

*“Toda criança começa como um cientista nato e então nós tiramos isso delas. Só algumas passam pelo sistema com sua admiração e entusiasmo pela ciência intacta.”*

Carl Sagan - Entrevista com o autor, Revista Psychology Today, Jan. 1996.

## RESUMO

A energia solar está, mundialmente, em uma crescente expansão. Apesar de pequena, se comparada a outros meios tradicionais de geração de energia, nunca houve indicadores tão otimistas acerca de sua utilização e popularização. No entanto, mesmo uma energia auto-sustentável como essa acaba tendo impacto ambiental, grande investimento econômico, dificuldade de medição de todas as parcelas da radiação solar e problemas na manutenção adequada das estações solares. Sabendo que a medição ou a estimativa das parcelas da radiação são cruciais para o dimensionamento e estimativa de geração de energia, tanto em grandes plantas de potência por concentração solar como em sistemas de geração fotovoltaica residenciais, o presente trabalho foi idealizado para assegurar que, mesmo em proporções continentais, a obtenção da componente difusa da irradiância, a partir da irradiância global, pode ser realizada de maneira confiável. A utilização de RNAs com generalização geográfica demonstrou melhores resultados em relação a modelos de correlação empíricos, além de serem mais abrangentes do que modelos que utilizam RNAs especializadas a partir de dados locais. Os resultados apresentaram erros RMSEr variando entre 16,00% a 29,57%, contra erros RMSEr de 15,94% a 36,98% em modelos tradicionais empíricos. O presente modelo também apresentou um aumento simultâneo de precisão local e generalização geográfica, algo que modelos anteriores não atingiam, visto que, ao aumentar a especialização e precisão em uma área menor, há um comprometimento ao aplicar o modelo em áreas geográficas distantes e/ou maiores (e vice-versa). Dessa maneira, o presente estudo assegura que pesquisadores poderão utilizar RNAs com precisão e generalização suficientes para realizar pesquisas e estudos diversos que necessitem da parcela difusa e/ou direta da radiação solar, com maior eco-eficiência e custo benefício possíveis, em localidades que não as meçam ou não possuam dados meteorológicos confiáveis.

Palavras-chave: Radiação solar; Modelos difusos; Redes neurais artificiais.

## ABSTRACT

Worldwide, the solar energy is on a growing expansion. Although small, if compared to other traditional ways of generating useful energy, there have never been such optimistic indicators about its use and popularization. However, even a self-sustainable energy like this have environmental impact, great economic investment, difficulty in the measurement of all types of solar radiation and problems in the proper maintenance of solar monitoring stations. Knowing that the measurement or estimative of all solar radiation types are crucial for the design and estimative of energy generation in both large concentrating solar power plants or in residential photovoltaic power generation systems, the present work was developed to ensure that, even in continental proportions, the horizontal diffuse fraction can be obtained from the total irradiance with great reliability. The use of Artificial Neural Nertorks (ANNs) with geographical generalization showed better results in relation to empirical models, as well as being more comprehensive than models that use specialized ANNs based on local data. The results showed RMSEr errors ranging from 16,00% to 29,57%, against RMSEr errors from 15,94% to 36,98% in traditional empirical models. The present model also showed a improvement in both local precision and geographic generalization, something that previous models didn't reach, since when increasing the specialization and precision in a smaller area, there's a significant loss when using the model in distant and/or larger geographic areas (and vice versa). In this way, the present study ensures that researchers can use ANNs with enough precision and generalization to perform different researches and studies that require the diffuse and/or direct types of solar radiation with greater eco-efficiency and cost-effectiveness in locations that don't measure them or don't have reliable meteorological data.

Keywords: Solar radiation; Diffuse models; Artificial neural networks.

## LISTA DE FIGURAS

Figura 1.1 – Crescimento da capacidade de geração de energia solar. ....	19
Figura 1.2 – Percentual de utilização de energias renováveis e não renováveis na geração mundial de energia durante o ano de 2016. ....	20
Figura 1.3 – Planta solar fotovoltaica. ....	21
Figura 2.1 – Relações Terra-Sol. ....	24
Figura 2.2 – Componentes da radiação solar. ....	25
Figura 2.3 – Piranômetro. ....	28
Figura 2.4 – Pireliômetro. ....	28
Figura 2.5 – Piranômetro com anel de sombreamento. ....	29
Figura 2.6 – Fatores de correção para anel de sombreamento. ....	29
Figura 2.7 – Piranômetro com esfera de sombreamento. ....	30
Figura 2.8 – Célula fotovoltaica. ....	31
Figura 2.9 – Sistemas fotovoltaicos em casas geminadas no município de Pentecoste, CE. ....	31
Figura 2.10 – Gemasolar: a maior planta de potência por concentração solar térmica do mundo, cuja potência instalada é de 19,9MW, localizada na Espanha ....	32
Figura 2.11 – Fornalha solar de Lavoisier. ....	33
Figura 2.12 – Ângulo de reflexão para o concentrador. ....	33
Figura 3.1 – Correlações entre índice $f$ e coeficiente de claridade. ....	35
Figura 3.2 – Correlações entre índice $f$ , $Kt$ e $Kt$ . ....	36
Figura 3.3– Correlação entre a fração difusa e o coeficiente de claridade horário. ....	37
Figura 3.4 – Separação das faixas do coeficiente de claridade $k_t$ . ....	38
Figura 3.5 – Exemplo de funcionamento de uma RNA. ....	41
Figura 3.6 – RNA do tipo <i>feedforward</i> . ....	42
Figura 3.7 – RNA do tipo <i>recurrent</i> . ....	42
Figura 4.1 – Estações do projeto SONDA. ....	49
Figura 4.2 – Estações SONDA utilizadas para modelagem e treinamento. ....	49
Figura 4.3 – Diferentes cidades utilizadas nos testes extremos de robustez. ....	52
Figura 5.1 – Erro $RMSEr$ (%) para as diferentes cidades e diferentes tipos de modelos/testes. ....	54
Figura 5.2 – Comparação dos erros $RMSEr$ obtidos pelas RNAs com função de ativação sigmoide nos testes de generalização geográfica em relação aos erros obtidos por meio dos modelos empíricos. ....	55

Figura 5.3 – Correlação entre dados medidos [azul] em relação aos dados estimados com menor erro [vermelho] pela RNA de sete neurônios com função de ativação sigmoide nos testes de generalização geográfica.....	56
Figura 5.4 – Comparação entre estimativas realizadas pela RNA de sete neurônios em testes de generalização geográfica com função de ativação sigmoide e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA. ....	57
Figura 5.5 – Comparação entre estimativas realizadas pela RNA de sete neurônios em testes de generalização geográfica com função de ativação sigmoide e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA. ....	58
Figura 5.6 – Comparação dos erros <i>RMSEr</i> obtidos pelas RNAs com função de ativação sigmoide nos testes de precisão local em relação aos erros obtidos por meio dos modelos empíricos. ....	59
Figura 5.7 – Correlação entre dados medidos [azul] em relação aos dados estimados com menor erro [vermelho] pela RNA de quatro neurônios com função de ativação sigmoide nos testes de precisão local. ....	60
Figura 5.8 – Comparação entre estimativas realizadas pela RNA de quatro neurônios em testes de precisão local com função de ativação sigmoide e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA. ..	61
Figura 5.9 – Comparação entre estimativas realizadas pela RNA de quatro neurônios em testes de precisão local com função de ativação sigmoide e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA. ..	62
Figura 5.10 – Comparação dos erros <i>RMSEr</i> obtidos pelas RNAs com função de ativação linear nos testes de generalização geográfica em relação aos erros obtidos por meio dos modelos empíricos.....	64
Figura 5.11 – Correlação entre dados medidos [azul] em relação aos dados estimados com menor erro [vermelho] pela RNA de cinco neurônios com função de ativação linear nos testes de generalização geográfica.....	65
Figura 5.12 – Comparação entre estimativas realizadas pela RNA de cinco neurônios em testes de generalização geográfica com função de ativação linear e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA. ....	66

Figura 5.13 – Comparação entre estimativas realizadas pela RNA de cinco neurônios em testes de generalização geográfica com função de ativação linear e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA. ....	67
Figura 5.14 – Comparação dos erros <i>RMSE<sub>r</sub></i> obtidos pelas RNAs com função de ativação linear nos testes de precisão local em relação aos erros obtidos por meio dos modelos empíricos. ....	68
Figura 5.15 – Correlação entre dados medidos [azul] em relação aos dados estimados com menor erro [vermelho] pela RNA de dois neurônios com função de ativação linear nos testes de precisão local. ....	69
Figura 5.16 – Comparação entre estimativas realizadas pela RNA de dois neurônios em testes de precisão local com função de ativação linear e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA. ....	70
Figura 5.17 – Comparação entre estimativas realizadas pela RNA de dois neurônios em testes de precisão local com função de ativação linear e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA. ....	71
Figura 5.18 – Testes extremos de robustez: Comparação de erros <i>RMSE<sub>r</sub></i> da RNA em relação à erros do modelo de Erbs et al. (1982), base de dados brasileiros utilizada para treinamento das RNAs e correlações de coeficientes de claridade e coeficientes difusos estimados [vermelho] em comparação com dados TMY3 medidos [azuis].....	73
Figura 6.1 - Sobreposição de dados estimados pelos modelos BRL1, BRL2 e RE2, para a cidade de Brasília.....	76
Figura A.1 – Imagem da distribuição da irradiação global horizontal no território brasileiro. ....	86
Figura B.1 – Distribuição da irradiação difusa horizontal no território brasileiro. ....	87
Figura C.1 – Demonstração da qualidade das medições do projeto SONDA. ....	88
Figura A.1 – Demonstração da qualidade dos dados utilizados neste estudo. ....	89

## LISTA DE TABELAS

Tabela 3.1 – Comparação entre modelos obtidos pelo estudo. ....	45
Tabela 5.1 – Coeficientes de correlação ‘ <i>r</i> ’ para as diferentes variáveis na configuração de generalização geográfica com 7 neurônios na camada escondida e função de ativação sigmoide. ....	63
Tabela 5.2 – Coeficientes de correlação ‘ <i>r</i> ’ para as diferentes variáveis na configuração de precisão local com 4 neurônios na camada escondida e função de ativação sigmoide. ....	63
.Tabela 5.3 – Coeficientes de correlação ‘ <i>r</i> ’ para as diferentes variáveis na configuração de generalização geográfica com 5 neurônios na camada escondida e função de ativação linear. ....	72
Tabela 5.4 – Coeficientes de correlação ‘ <i>r</i> ’ para as diferentes variáveis na configuração de precisão local com 2 neurônios na camada escondida e função de ativação linear. ....	72
Tabela 6.1 – Menores erros <i>RMSE</i> (%) obtidos por meio dos modelos clássicos, modelos neurais generalistas e modelos neurais locais. ....	77
Tabela B.1 – Erro <i>RMSEr</i> (%) dos modelos difusos empíricos obtidos a partir da base de testes de generalização geográfica. ....	90
Tabela B.2 – Erro <i>MBEr</i> (%) dos modelos difusos empíricos obtidos a partir da base de testes de generalização geográfica. ....	90
Tabela B.3 – Erro <i>RMSEr</i> (%) dos modelos difusos empíricos obtidos a partir da base de testes de precisão local. ....	90
Tabela B.4 – Erro <i>MBEr</i> (%) dos modelos difusos empíricos obtidos a partir da base de testes de precisão local. ....	90
Tabela C.1 – Erro <i>RMSEr</i> (%) das RNAs por número de neurônios da camada escondida para testes de generalização geográfica com função de ativação sigmoide. ....	91
Tabela C.2 – Erro <i>MBEr</i> (%) das RNAs por número de neurônios da camada escondida para testes de generalização geográfica com função de ativação sigmoide. ....	91
Tabela D.1 – Erro <i>RMSEr</i> (%) das RNAs por número de neurônios da camada escondida para testes de precisão local com função de ativação sigmoide. ....	92
Tabela D.2 – Erro <i>MBEr</i> (%) das RNAs por número de neurônios da camada escondida para testes de precisão local com função de ativação sigmoide. ....	92
Tabela E.1 – Erro <i>RMSEr</i> (%) das RNAs, por número de neurônios da camada escondida, para testes de generalização geográfica com função de ativação linear. ....	93

Tabela E.2 – Erro <i>MBEr</i> (%) das RNAs por número de neurônios da camada escondida para testes de generalização geográfica com função de ativação linear.....	93
Tabela F.1 – Erro <i>RMSEr</i> (%) das RNAs por número de neurônios da camada escondida para testes de precisão local com função de ativação linear. ....	94
Tabela F.2 – Erro <i>MBEr</i> (%) das RNAs por número de neurônios da camada escondida para testes de precisão local com função de ativação linear. ....	94
Tabela G.1 – Erros <i>RMSE</i> (%) dos testes extremos de robustez do modelo desenvolvido. ....	95

## LISTA DE SIGLAS

ANEEL	Agência Nacional de Energia Elétrica
ANN	<i>Artificial Neural Network</i>
CRESESB	Centro de Referência para as Energias Solar e Eólica Sérgio de S. Brito
CSP	<i>Concentrated Solar Power</i>
TWD	Transformada <i>Wavelet</i> Discreta
IBGE	Instituto Brasileiro de Geografia e Estatística
INMET	Instituto Nacional de Meteorologia
INPE	Instituto Nacional de Pesquisas Espaciais
FBR	Função de Base Radial
GNU	<i>Gnu's Not Unix</i>
GPL	<i>General Public License</i>
RNA	Rede Neural Artificial
SONDA	Sistema de Organização Nacional de Dados Ambientais
MSV	Máquina de Suporte Vetorial
PV	<i>Photovoltaics</i>
WRC	<i>World Radiation Center</i>
TW	Transformada <i>Wavelet</i>
TMY	<i>Typical Meteorological Year</i>
RFB	Rede de Função de Base Radial
SAINF	Sistema Adaptativo de Inferência Neuro Fuzzy
LST	<i>Land Surface Temperature</i>
UR	Umidade Relativa do Ar

## LISTA DE SIMBOLOS

$E_0$	Correção da excentricidade solar [-]
$f$	Fração de tempo [-]
$G_0$	Irradiância extraterrestre horizontal [W/m <sup>2</sup> ]
$G_a$	Irradiância global horizontal horária [W/m <sup>2</sup> ]
$G_b$	Irradiância direta normal [W/m <sup>2</sup> ]
$G_d$	Irradiância difusa horizontal horária [W/m <sup>2</sup> ]
$G_{sc}$	Constante solar [W/m <sup>2</sup> ]
$G_{on}$	Irradiância extraterrestre horizontal horária [W/m <sup>2</sup> ]
$\bar{H}$	Irradiância global horizontal média mensal [W/m <sup>2</sup> ]
$\bar{H}_0$	Irradiância extraterrestre normal média mensal [W/m <sup>2</sup> ]
$k_b$	Coeficiente direto horário [-]
$k_{bc}$	Coeficiente direto máximo para uma dada massa de ar [-]
$k_d$	Coeficiente difuso horário [-]
$k_t$	Coeficiente de claridade horário [-]
$K_t$	Coeficiente de claridade diário [-]
$\bar{K}_t$	Coeficiente de claridade médio mensal [-]
$m$	Massa de ar atmosférico [-]
$T_{amb}$	Temperatura Ambiente [°]

## LISTA DE SIMBOLOS GREGOS

$\alpha$	Elevação solar [°]
$\Delta k_b$	Varição do coeficiente direto horário [-]
$\sigma$	Brilho relativo do Sol [-]
$\theta_z$	Ângulo de zênite [°]

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>17</b>
1.1 OBJETIVOS.....	18
<b>1.1.1 Objetivo Geral .....</b>	<b>18</b>
<b>1.1.2 Objetivos Específicos.....</b>	<b>18</b>
1.2 JUSTIFICATIVA .....	19
1.3 ESTRUTURA DA PESQUISA.....	22
<b>2 CONCEITOS BÁSICOS.....</b>	<b>24</b>
2.1 RADIAÇÃO SOLAR.....	24
2.2 COMPONENTES DA RADIAÇÃO SOLAR.....	25
<b>2.2.1 Irradiância Extraterrestre Normal e Horizontal .....</b>	<b>25</b>
<b>2.2.2 Irradiância Global Horizontal .....</b>	<b>26</b>
<b>2.2.3 Irradiância Direta Normal .....</b>	<b>26</b>
<b>2.2.4 Irradiância Difusa Horizontal.....</b>	<b>26</b>
<b>2.2.5 Irradiância de Albedo .....</b>	<b>27</b>
2.3 MEDIÇÃO DAS COMPONENTES DA RADIAÇÃO .....	27
<b>2.3.1 Irradiância Global Horizontal .....</b>	<b>27</b>
<b>2.3.2 Irradiância Direta Normal .....</b>	<b>28</b>
<b>2.3.3 Irradiância Difusa Horizontal.....</b>	<b>29</b>
2.4 IMPORTÂNCIA E UTILIZAÇÃO DE CADA COMPONENTE.....	30
<b>2.4.1 Energia solar fotovoltaica .....</b>	<b>30</b>
<b>2.4.2 Energia solar térmica.....</b>	<b>32</b>
<b>3 MODELAGEM DE COMPONENTES DA RADIAÇÃO SOLAR .....</b>	<b>34</b>
3.1 DIFICULDADES DE MEDIÇÃO DAS COMPONENTES .....	34
3.2 MODELAGEM DA IRRADIÂNCIA TOTAL DIÁRIA DISPONÍVEL .....	35
3.3 MODELOS HORÁRIOS UTILIZANDO CORRELAÇÕES .....	37
<b>3.3.1 Modelo O&amp;R de Orgill e Hollands (1977).....</b>	<b>39</b>
<b>3.3.2 Modelo ER de Erbs et al. (1982) .....</b>	<b>39</b>
<b>3.3.3 Modelo RE1 e RE2 de Reindl et al. (1990).....</b>	<b>39</b>
<b>3.3.4 Modelo BRL1 e BRL2 de Ridley et al. (2010).....</b>	<b>40</b>
3.4 MODELOS HORÁRIOS UTILIZANDO RNA .....	41
<b>3.4.1 Modelo de Mellit et al. (2010) .....</b>	<b>43</b>
3.5 MODELOS DIFERENCIADOS .....	43

<b>3.5.1 Modelo MAX de Maxwell (1987)</b> .....	<b>44</b>
<b>3.5.2 Shamshirband et al. (2016)</b> .....	<b>44</b>
<b>3.5.3 Paulescu e Blaga (2016)</b> .....	<b>45</b>
<b>3.5.4 Basseto et al. (2018)</b> .....	<b>46</b>
<b>4 MODELO GERAL COM UTILIZAÇÃO DE REDES NEURAIS ARTIFICIAIS</b> .....	<b>47</b>
4.1 MODELO BASE PARA O ESTUDO .....	47
4.2 VARIÁVEIS DE ENTRADA .....	48
4.3 BASE DE DADOS PARA TESTE .....	48
4.4 TESTES DE PRECISÃO LOCAL E GENERALIZAÇÃO GEOGRÁFICA.....	50
4.5 TESTES EXTREMOS DE ROBUSTEZ .....	51
4.6 QUANTIFICAÇÃO DO DESVIO.....	52
4.7 COEFICIENTE DE CORRELAÇÃO LINEAR .....	53
<b>5 RESULTADOS</b> .....	<b>54</b>
<b>6 ANÁLISE DOS RESULTADOS</b> .....	<b>75</b>
<b>7 CONCLUSÃO</b> .....	<b>79</b>
<b>REFERÊNCIAS</b> .....	<b>81</b>
<b>ANEXO A – RADIAÇÃO SOLAR GLOBAL HORIZONTAL NO BRASIL</b> .....	<b>86</b>
<b>ANEXO B – RADIAÇÃO SOLAR DIFUSA HORIZONTAL NO BRASIL</b> .....	<b>87</b>
<b>ANEXO C – AMOSTRAS DE QUALIDADE DE DADOS (SONDA)</b> .....	<b>88</b>
<b>APÊNDICE A – DEMONSTRAÇÃO DA QUALIDADE DOS DADOS MEDIDOS PELAS ESTAÇÕES DO PROJETO SONDA</b> .....	<b>89</b>
<b>APÊNDICE B – ERROS RMSE E MBE OBTIDOS A PARTIR DOS MODELOS EMPÍRICOS</b> .....	<b>90</b>
<b>APÊNDICE C – ERROS RMSE E MBE OBTIDOS A PARTIR DA RNA EM TESTES DE GENERALIZAÇÃO GEOGRÁFICA I</b> .....	<b>91</b>
<b>APÊNDICE D – ERROS RMSE E MBE OBTIDOS A PARTIR DA RNA EM TESTES DE PRECISÃO LOCAL I</b> .....	<b>92</b>
<b>APÊNDICE E – ERROS RMSE E MBE OBTIDOS A PARTIR DA RNA EM TESTES DE GENERALIZAÇÃO GEOGRÁFICA II</b> .....	<b>93</b>
<b>APÊNDICE F – ERROS RMSE E MBE OBTIDOS A PARTIR DA RNA EM TESTES DE PRECISÃO LOCAL II</b> .....	<b>94</b>
<b>APÊNDICE G – ERROS RMSE E MBE OBTIDOS A PARTIR DE TESTES EXTREMOS DE ROBUSTEZ</b> .....	<b>95</b>

## 1 INTRODUÇÃO

A sociedade está se importando cada vez mais com a sustentabilidade do planeta, culminando em um crescente - e recente - interesse em energias auto-sustentáveis. Uma delas é a energia solar e suas duas formas de obtenção de energia útil: sistemas fotovoltaicos e sistemas térmicos. Para o estudo da viabilidade econômica e ambiental, as componentes direta e difusa horizontais da irradiância global horizontal devem ser utilizadas. Essas componentes representam os raios solares que atingem o solo sem sofrer alteração em seu curso e os raios solares que sofrem interação com a atmosfera. A grande maioria das bases de dados contam somente com instrumentos de medição para a irradiância global horizontal, que é a soma das duas componentes comentadas anteriormente, o que torna necessária a utilização de modelos para a estimativa da irradiância (ou irradiação) difusa para uma correta estimativa da geração de energia a partir da energia solar.

O Brasil possui uma grande quantidade de estações que realizam a medição da irradiância global horizontal por meio do Instituto Nacional de Meteorologia (INMET). No entanto, nosso país também possui uma carência de dados de irradiância direta e difusa. Essa carência é suprida em parte pelas estações de medição do projeto SONDA, distribuídas em diversas localidades, onde se efetuam medições dessas componentes da radiação solar. No entanto, poucas bases possuem mais de sete anos contínuos de dados com alta confiabilidade, conforme mostrado pelo próprio portal SONDA (2018).

Uma outra alternativa para tentar suprir essa carência é a utilização de modelos de estimativa da radiação difusa clássicos (horário ou mensais), como os desenvolvidos por Orgill e Hollands (1977), Erbs et al. (1982) e Reindl et al. (1990). Esses modelos apresentam correlações realizadas entre dados de determinadas localidades e possuem generalização geográfica razoável. Outras soluções abrangem redes neurais artificiais (RNAs) e sua capacidade de especialização. Nesse caso, a precisão é alta apenas para a localidade cujos dados foram utilizados para o aprendizado da rede. Outros modelos que utilizam diferentes metodologias também sofrem do problema de generalização, devido às características dos dados utilizados para obtenção dos modelos.

Com proporções continentais, o território brasileiro possui seis grandes biomas: Amazônia, Caatinga, Cerrado, Pantanal, Mata Atlântica e Pampa, cada um destes apresentando características únicas de flora, fauna e condições climáticas (IBGE, 2017). O país também possui cinco condições climáticas: Equatorial, Tropical Zona Equatorial, Tropical Nordeste Oriental, Tropical Brasil Central e Temperado (IBGE, 2017). Cada uma

delas apresenta características que, em conjunto com os biomas, podem alterar a disponibilidade de radiação em suas componentes direta e difusa.

Desta maneira, torna-se muito difícil o desenvolvimento de um modelo difuso horário baseado somente em correlações de coeficiente de claridade para obtenção do coeficiente difuso, como realizado no estudo de Orgill e Hollands (1977). Uma nova abordagem deve ser utilizada para um modelo dessa característica em um território com proporções como as do Brasil, sendo ao mesmo tempo precisa e suficientemente generalizada para obter as parcelas difusa e direta no país como um todo.

A maior generalização possível em um modelo difuso é a partir da utilização de variáveis físicas. No entanto, quanto mais geral um modelo é, menos precisão ele possui. Portanto, o presente estudo propõe a utilização de uma RNA treinada a partir de dados que, além do tradicional coeficiente de claridade, incluam variáveis físicas que influenciam as parcelas direta e difusa. Assim, as particularidades e não linearidades características de cada localidade podem ser interpretadas. Com essa abordagem, a precisão de um modelo físico pode ser aumentada consideravelmente, conforme demonstrada em estudos como o de Pillot (2012), realizado com metodologia parecida para estimar as variáveis de temperatura na República do Djibouti.

## 1.1 OBJETIVOS

A seguir, são apresentados o objetivo geral e os objetivos específicos, os quais orientarão o caminho que será percorrido em busca de respostas à questão problema da pesquisa.

### 1.1.1 Objetivo Geral

Avaliar o emprego de redes neurais artificiais para estimativa da irradiância difusa horária horizontal em grandes extensões territoriais, a partir de dados medidos no Brasil.

### 1.1.2 Objetivos Específicos

- Analisar modelos difusos existentes no Brasil;
- Estabelecer a estrutura da RNA proposta neste estudo;
- Avaliar o grau de generalização geográfica da RNA proposta;

- Comparar quantitativa e qualitativamente o modelo proposto em relação a modelos existentes.

## 1.2 JUSTIFICATIVA

Os três principais pontos comumente levantados como justificativas, ao pensar em energia solar, são o interesse crescente de pesquisadores na área, o aumento da utilização dessa energia na geração energética global e a crença de que se trata de uma energia totalmente limpa. No entanto, todos os pontos citados podem ser discutidos, e a verdadeira importância do presente trabalho está um passo à frente desses levantamentos.

Ao analisar pesquisas realizadas em datas muito distantes uma da outra, como Whillier (1953) e Scalco (2016), pode-se perceber que os campos de estudo ainda continuam muito parecidos, transitando principalmente em sistemas fotovoltaicos (PV) e de potência solar concentrada (*Concentrated Solar Power*), CSP. A principal evolução apresentada é a melhoria da eficiência, da técnica e o aumento da precisão de modelos. Whillier (1953) afirma, por exemplo, que a parcela de irradiância direta diária a ser utilizada para estimativa em concentradores é de cerca de 60%. Hoje em dia, modelos diretos diários muito mais precisos, como o obtido por Shamshirband et al. (2016), demonstram que há muito mais do que uma breve verificação de dados para afirmar uma quantia de irradiância direta normal diária.

A utilização das energias renováveis, principalmente a solar, está realmente em crescimento, porém esse aumento da capacidade instalada só aconteceu, significativamente, a partir do ano de 2009, como mostra o gráfico apresentado por Vaughan (2017), na Fig. 1.1.

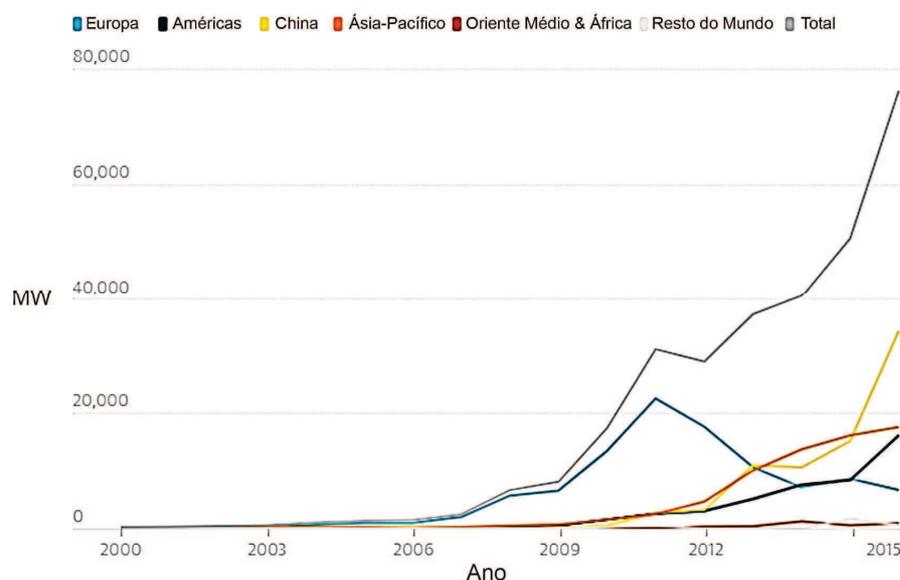


Figura 1.1 – Crescimento da capacidade de geração de energia solar.

Fonte: Adaptado de The Guardian: Solar Power (2017).

Essa matéria mostra que a capacidade instalada para geração de energia solar em 2016 era de 305 GW, valor muito superior aos 50 GW instalados até 2010. No entanto, mesmo com todo este crescimento da capacidade de geração de energia solar, ela representa apenas 2% da geração total de energia mundial, número menor, inclusive, do que os 5% gerados por energia eólica, segundo o Global Energy Statistical Yearbook (ENERDATA, 2017). O gráfico da comparação entre energias renováveis (eólica e solar) e energias não renováveis, em relação à produção total de energia mundial em 2016, pode ser visto na Fig. 1.2.

### Percentual de Produção de Eletricidade (2016)

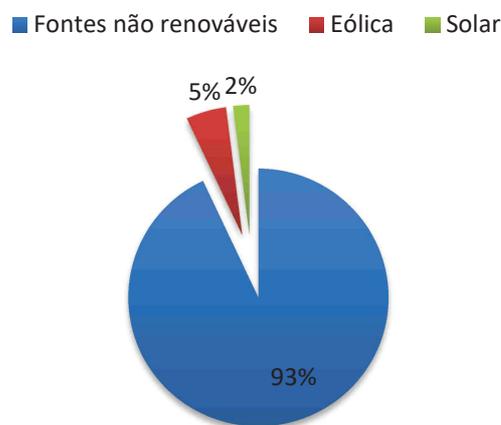


Figura 1.2 – Percentual de utilização de energias renováveis e não renováveis na geração mundial de energia durante o ano de 2016.

Fonte: Adaptado de Global Energy Statistical Yearbook (ENERDATA, 2017).

A energia solar, tanto térmica quanto fotovoltaica, possui um custo ambiental que consome parte da energia obtida de maneira limpa. Há um ótimo levantamento sobre essa relação no estudo realizado por Dai-Prá et al. (2014), cujos autores concluíram que para determinadas condições, na cidade de Porto Alegre - Brasil, um sistema fotovoltaico (Fig. 1.3) se prova auto-sustentável. No entanto, cerca de 25% de sua energia, em 25 anos de vida útil, serão usadas como *payback* da energia utilizada para a fabricação dos módulos de silício monocristalinos. Mesmo em energias sustentáveis, algum impacto ambiental ocorre, ainda que mínimo na maioria dos casos.

Com essa crítica aos três principais pilares que sustentam a utilização da energia solar, onde realmente se encontra a importância deste trabalho? Justamente para auxiliar, com a confiabilidade necessária, no dimensionamento de diferentes meios de obtenção de energia a partir da radiação solar, com o menor custo ambiental e econômico possível.



Figura 1.3 – Planta solar fotovoltaica.

Fonte: Global Energy Statistical Yearbook (ENERDATA, 2017).

Sabe-se que o Brasil possui uma grande extensão de terra e uma grande incidência de energia solar. Ao comparar a incidência da radiação em suas parcelas global e difusa, conforme ANEXO A e B (PEREIRA et al., 2017), percebe-se um grande potencial de geração térmica e fotovoltaica em diversas regiões de nosso país. No entanto, a medição das parcelas direta e difusa é realizada em apenas 17 localidades (SONDA, 2018). Portanto uma correta análise do potencial de energia em uma planta de potência térmica ou fotovoltaica necessita de um modelo adequado para nossas características climáticas e geográficas, principalmente se levado em conta o fato de que uma planta de potência térmica necessita de investimento muito mais alto que uma planta de geração fotovoltaica.

Os meios conhecidos hoje para geração de energia a partir do Sol, mesmo com custo ambiental em seu estágio de produção<sup>1</sup>, possuem uma razoável eficiência na conversão dessa energia em energia elétrica e/ou térmica. No entanto, em muitos casos, a energia solar ainda é a mais eficiente quando variáveis econômicas, legislativas e ambientais são levadas em conta. No Brasil, pode-se citar o exemplo da Resolução Normativa nº 687/2015 (ANEEL, 2015), pela qual é permitida a inserção de energia elétrica produzida por sistemas fotovoltaicos diretamente na rede, possibilitando que consumidores em suas casas produzam energia elétrica por meio de painéis fotovoltaicos, adicionando assim mais uma maneira de geração de energia solar: a micro geração distribuída.

Sabe-se que a produção de eletricidade no Brasil é baseada, principalmente, em hidroeletricidade, por causa de suas bacias hidrográficas grandes. Porém, há potencial de geração solar térmica para plantas de potência em locais com baixa pluviosidade e bacias hidrográficas pequenas; há geração solar térmica de média temperatura em locais com bom

---

<sup>1</sup> Como exemplo, pode-se citar o tratamento do silício para módulos fotovoltaicos.

índice de radiação direta e alguma cobertura do céu; e há muitas localidades com bom índice de radiação difusa, direta e de albedo para geração de energia fotovoltaica.

Para suprir e poder realizar uma estimativa segura e confiável das parcelas difusa e direta no Brasil, em toda a sua amplitude geográfica, é necessário um modelo que obtenha a melhor precisão em relação a todo o nosso território. Para isso, propõe-se uma análise das variáveis que influenciam a alteração da radiação dentro da atmosfera e uma RNA que consiga perceber e interpretar as particularidades de cada região por meio dos dados coletados para este estudo.

A maior contribuição desta pesquisa deve-se ao fato de garantir que a estimativa das parcelas direta e difusa poderá ser realizada com uma maior amplitude geográfica e precisão na equação de disponibilidade das componentes solares, investimento econômico e ecoeficiência para cada tipo de sistema de geração de energia solar. Assim, estudos que necessitem de um modelo confiável poderão verificar quais localidades mais se adequarão aos tipos de geração de energia solar em relação à disponibilidade das componentes da irradiância solar que esses meios utilizem.

Lorenzo (2006) aborda que a radiação solar é um fenômeno aleatório, dessa maneira predições sempre estarão acompanhadas de incerteza. Ele também cita que muitos autores mostram resultados com uma grande precisão ao utilizar ferramentas computacionais para projetos fotovoltaicos, porém sua complexidade aumenta consideravelmente e, pela natureza da radiação solar, não há nenhum sentido estatístico que comprove essa melhora. Além disso, ferramentas mais simples podem fazer esse serviço com resultados de igual confiança.

Esse pensamento, mesmo aplicado a sistemas fotovoltaicos, pode ser abordado em modelos difusos. Os mesmos estão cada vez mais complexos - como os apresentados por Starkveit et al. (1998) - ou altamente precisos para determinadas localidades. Há pouco ou nenhum estudo abordando robustez e generalização de modelos difusos para utilização confiável em outras regiões próximas à localidade cujos dados foram utilizados para adaptar ou desenvolver um modelo difuso. Lorenzo (2006) também cita que em determinados casos, pode-se ou usar correlações universais, ou obter resultados ligeiramente melhores por meio de correlações lineares específicas para a localidade de estudo.

### 1.3 ESTRUTURA DA PESQUISA

O capítulo 2 consiste nos conceitos básicos da radiação solar, as diferentes componentes que formam a irradiância total, como e de que maneira cada uma delas é medida

e a importância de cada uma dessas componentes dentro da geração de energia fotovoltaica e térmica. Também é tratada a contextualização da necessidade de utilização dos modelos difusos para estimativa das componentes direta e difusa em dimensionamento e avaliação de eficiência de plantas de geração de energia.

A modelagem difusa é abordada no capítulo 3. Apresenta o contexto histórico da evolução das correlações até os modelos horários clássicos conhecidos hoje - como os de Orgill e Hollands (1977) e Erbs et al. (1982) - passando por modelos que utilizam RNAs, além de outras metodologias não ortodoxas para a mesma finalidade de estimativa da componente difusa e ou direta.

Em seguida, no capítulo 4, há o detalhamento da metodologia para a obtenção de um modelo difuso horário baseado em redes neurais artificiais para estimativa da irradiância difusa horária horizontal em grandes extensões territoriais. São abordados desde os dados utilizados até as arquiteturas das RNAs e detalhamentos dos testes realizados.

Já nos capítulos 5 e 6 são apresentados os resultados dos testes realizados e a análise dos mesmos, respectivamente. A conclusão do estudo, baseada nos resultados obtidos e na comparação com estudos já realizados, é abordada no capítulo 7.

## 2 CONCEITOS BÁSICOS

Neste capítulo são abordados os conceitos básicos da radiação solar. Serão tratados a origem da radiação (das reações de fusão até a constante solar), suas componentes (oriundas das modificações sofridas pela radiação até incidir em superfícies de conversão para energia térmica ou elétrica), a maneira como cada uma dessas componentes é medida, e a importância e utilização específica de cada componente da radiação total oriunda do Sol.

### 2.1 RADIAÇÃO SOLAR

O Sol é uma grande massa gasosa que gira em seu próprio eixo e possui temperaturas internas e externas altíssimas devido a contínuas reações de fusão que ocorrem em seu interior. Como resultado dessas reações, há uma emissão de radiação onde parte dela atinge o planeta Terra. A radiação emitida pelo Sol que alcança a Terra é função da distância entre ambos e da órbita da Terra em torno do Sol. A radiação, fora da atmosfera terrestre, é aproximadamente fixa. A medição da radiação extraterrestre vem sendo feita por diversos pesquisadores ao longo dos anos, desde balões atmosféricos até satélites. Atualmente, a energia emitida pelo Sol por unidade de tempo e por unidade de área, conhecida como constante solar ( $G_{sc}$ ), foi fixada em  $1367 \text{ W/m}^2$  com 1% de incerteza, pelo *World Radiation Center* (WRC). Essa constante corresponde à potência incidente na Terra a uma distância média Sol-Terra, em função da órbita elíptica da Terra em relação ao Sol. A relação Terra-Sol é mostrada na Fig. 2.1, conforme Duffie e Beckman (2006).

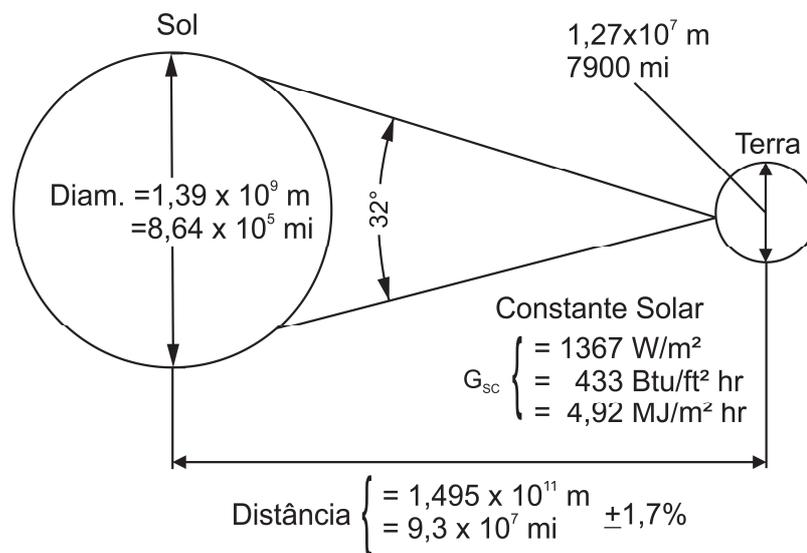


Figura 2.1 – Relações Terra-Sol.

Fonte: Adaptado de Duffie e Beckman (2006).

## 2.2 COMPONENTES DA RADIAÇÃO SOLAR

A constante solar ( $G_{sc}$ ) é uma média arbitrada pela WRC para facilitar cálculos e simulações. Seu valor varia constantemente a partir da atividade da fusão solar, manchas solares, entre outros. A irradiância difusa ocorre quando os raios solares sofrem espalhamento e atenuação na atmosfera terrestre. A irradiância direta ocorre quando os raios solares não sofrem interação com a atmosfera. A parcela da radiação refletida pelo solo é conhecida como radiação de albedo. Essas parcelas da radiação solar são representadas na Fig. 2.2.

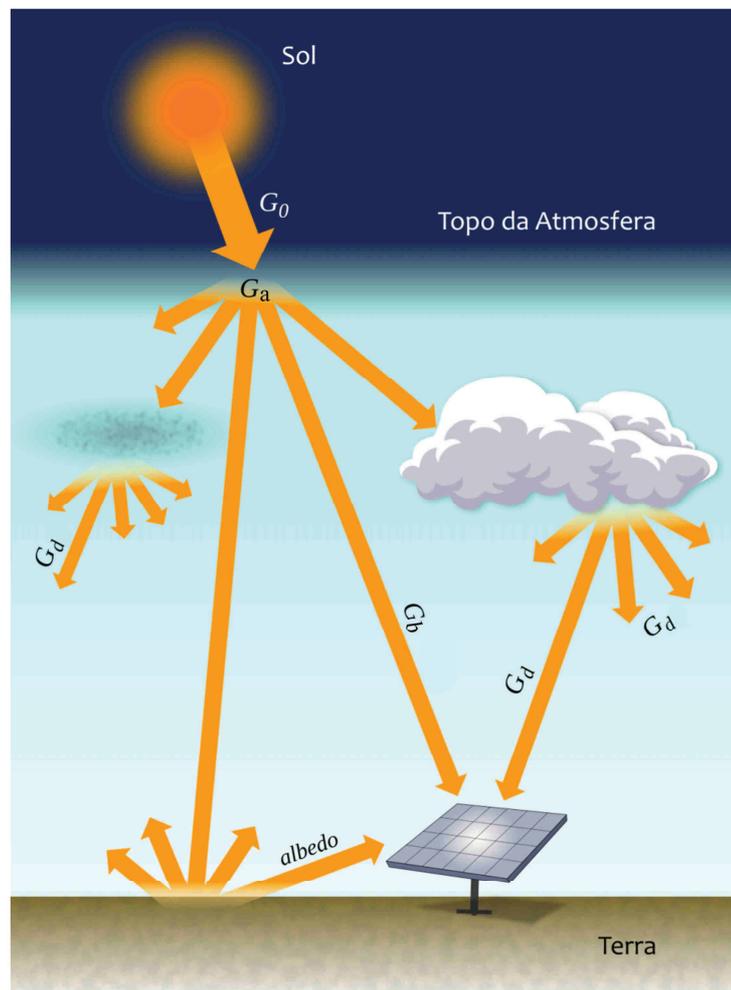


Figura 2.2 – Componentes da radiação solar.

Fonte: Adaptado de Pereira et al. (2017).

### 2.2.1 Irradiância Extraterrestre Normal e Horizontal

Para muitos cálculos de componentes de radiação, é necessário saber o valor teórico máximo de radiação solar disponível no topo da atmosfera de nosso planeta. Como visto, a constante solar é obtida a partir da média entre a menor e a maior distância da Terra em

relação ao Sol. A radiação incidente em um plano transversal ao Sol, localizado fora da atmosfera terrestre, chama-se irradiância solar extraterrestre normal, e é calculada a partir da constante solar definida pela Eq. (2.1), de acordo com Duffie e Beckman (2006):

$$G_0 = G_{sc} E_0 \quad (2.1)$$

em que  $G_{sc}$  é o valor atribuído à constante solar e  $E_0$  é a correção da excentricidade da órbita da Terra em torno do Sol.

A irradiância extraterrestre horizontal, calculada a partir de  $G_0$ , é definida pela Eq. (2.2), também de acordo com Duffie e Beckman (2006):

$$G_{on} = G_0 \cos(\theta_z) \quad (2.2)$$

em que  $\theta_z$  representa o ângulo zenital.

### 2.2.2 Irradiância Global Horizontal

É a medida mais comum da radiação, correspondente à radiação total em uma superfície horizontal, ou seja, a soma das parcelas direta e difusa da radiação disponíveis na atmosfera (IQBAL, 1983). Essa componente também é chamada de radiação solar global.

### 2.2.3 Irradiância Direta Normal

É a radiação solar recebida sem ser espalhada, absorvida e/ou alterada de seu curso pela atmosfera terrestre (DUFFIE e BECKMAN, 2006). Levando em conta que a soma das diferentes componentes da radiação solar na atmosfera corresponde à irradiância global horizontal, pode-se obter a irradiância direta de diferentes formas: subtraindo a irradiância difusa (estimada ou medida) da irradiância global horizontal ou estimando-a através de modelos diretos (MAXWELL, 1987).

### 2.2.4 Irradiância Difusa Horizontal

É a radiação solar recebida do Sol depois de interagir com a atmosfera terrestre. A irradiância difusa horizontal também pode ser obtida de diferentes formas: subtraindo a irradiância direta normal (estimada ou medida) da irradiância global horizontal ou estimando-a a partir de modelos difusos (DUFFIE e BECKMAN, 2006).

### **2.2.5 Irradiância de Albedo**

Quando a radiação atinge o solo, parte dela é absorvida, parte é refletida e parte é transmitida (essa última em caso de superfícies semi-transparentes). Essas propriedades das superfícies são chamadas de absorvidade, refletividade e transmissividade, respectivamente. Cada tipo de superfície possui sua própria refletividade, ou albedo (IQBAL, 1983). Essa refletividade ou albedo é uma propriedade do solo, água e atmosfera.

## **2.3 MEDIÇÃO DAS COMPONENTES DA RADIAÇÃO**

Para quantificar a disponibilidade de energia solar que atinge a Terra, torna-se necessária a medição de suas componentes. Os instrumentos utilizados para medir as componentes da radiação são o piranômetro, o pireliômetro e o pirgeômetro. Esse último é utilizado para medição da radiação de ondas longas, não abordada neste estudo.

### **2.3.1 Irradiância Global Horizontal**

Essa parcela da radiação é medida utilizando um piranômetro, instrumento que mede a radiação hemisférica total de ondas curtas. A medição é realizada a partir da diferença de temperaturas entre uma junção quente (o material sensor da radiação solar) e uma junção fria (a carcaça do instrumento), utilizando termopares em série ou em série/paralelo, em uma montagem conhecida como termopilha de Moll-Gorczyński, conforme Hulstron (1989). Nos piranômetros mais recentes, a termopilha é substituída por células Peltier. Para minimizar a perda de calor por convecção, os medidores possuem uma ou duas cúpulas hemisféricas de 180°, para proteção contra vento e outros fatores externos que possam alterar a precisão do instrumento. Pode-se ver um piranômetro na Fig. 2.3.



Figura 2.3 – Piranômetro.

Fonte: Adaptado de SONDA (2018).

### 2.3.2 Irradiância Direta Normal

Para medir essa componente, utiliza-se um instrumento chamado pireliômetro. Seu funcionamento ocorre por meio da comparação entre a temperatura obtida em um tubo colimador<sup>2</sup> e a temperatura padrão (DUFFIE e BECKMAN, 2006). O tubo colimador deste medidor não permite que a irradiância difusa o atinja, pois essa componente não possui uma orientação preferencial, ao contrário do que acontece com a radiação direta, conforme representado na Fig. 2.4.



Figura 2.4 – Pireliômetro.

Fonte: Adaptado de SONDA (2018).

---

<sup>2</sup> Tubo que permite que apenas raios perpendiculares ou dentro de um determinado ângulo de aceitação atinjam o medidor para que o sensor aqueça em relação a temperatura padrão (DUFFIE e BECKMAN, 2006).

### 2.3.3 Irradiância Difusa Horizontal

Visto anteriormente que um piranômetro (Fig. 2.5) mede a radiação total em uma superfície horizontal - radiação difusa horizontal somada à radiação direta normal - é possível medir a radiação difusa sombreando a radiação direta que incide sobre esse instrumento (DUFFIE e BECKMAN, 2006). O sombreamento da irradiância direta pode ser feito de duas maneiras: com um anel de sombreamento ou com uma esfera de sombreamento.



Figura 2.5 – Piranômetro com anel de sombreamento.

Fonte: Adaptado de SONDA (2018).

O primeiro realiza o bloqueio da radiação direta por meio de um anel circular ou semi-circular em toda a trajetória do Sol. Por acabar bloqueando um pouco da radiação difusa, o valor lido pelo instrumento deverá ser corrigido, conforme a Fig. 2.6. (DUFFIE e BECKMAN, 2006). Além disso, a posição do anel deve ser ajustada em intervalos regulares, em função da variação da declinação solar.

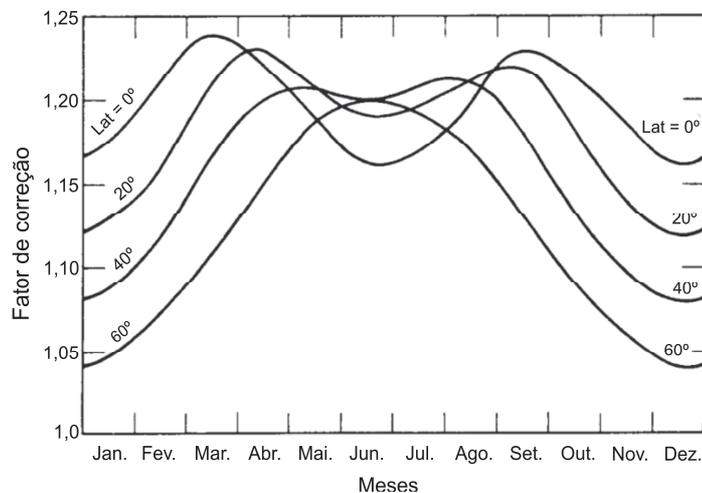


Figura 2.6 – Fatores de correção para anel de sombreamento.

Fonte: Adaptado de Duffie e Beckman (2006).

O piranômetro com anel de sombra ainda é muito utilizado por questões de custo. No entanto, levando em conta o crescente barateamento tecnológico, os piranômetros com seguidores solares possuem um ótimo custo benefício em relação ao aumento de sua precisão. Suas esferas de sombreamento seguem e bloqueiam apenas a irradiância direta normal do Sol - por meio de controle eletrônico - medindo quase somente a irradiância difusa disponível, conforme Fig. 2.7.



Figura 2.7 – Piranômetro com esfera de sombreamento.

Fonte: Adaptado de SONDA (2018).

## 2.4 IMPORTÂNCIA E UTILIZAÇÃO DE CADA COMPONENTE

Todas as componentes da radiação que se formam, a partir da entrada da irradiância extraterrestre em nossa atmosfera, são utilizadas de alguma maneira nos diferentes tipos de conversão de energia a partir do Sol. A energia solar fotovoltaica converte a energia solar em energia elétrica, e a energia solar térmica converte a energia solar em energia térmica. Cada um desses meios de conversão da energia solar utiliza determinadas componentes da radiação e possui suas próprias características.

### 2.4.1 Energia solar fotovoltaica

Mesmo de forma bastante simples e rudimentar, Bequerel (1839) observou que alguns materiais, ao interagirem com a luz, produzem uma corrente elétrica. Esse conceito se chama efeito fotovoltaico (Fig. 2.8), base para toda a utilização de materiais semicondutores na conversão direta de energia solar em energia elétrica. Materiais semicondutores são isolantes em determinadas condições (WENHAM et al., 2007) porém, quando fótons atingem

sua superfície, alguns deles podem gerar corrente elétrica. Dessa maneira, tanto a componente direta como a difusa são cruciais para a geração de energia, bem como a irradiância de albedo - quando essa estiver no campo de visão de um módulo inclinado.

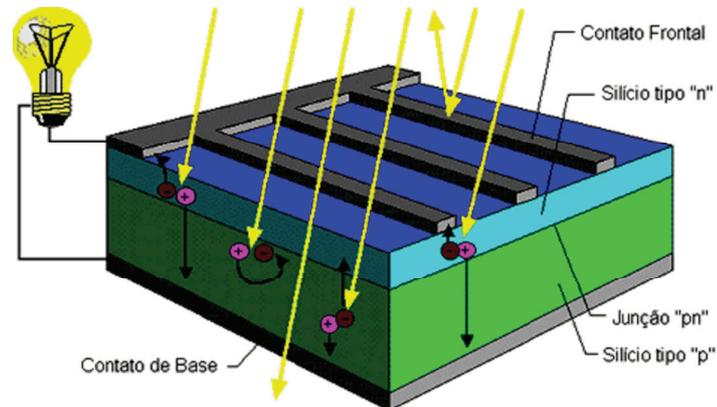


Figura 2.8 – Célula fotovoltaica.

Fonte: CRESESB (2008).

No Brasil, há a nova regulação da ANEEL para que pessoas físicas possam ser produtoras de energia fotovoltaica, possibilitando a criação de uma grande rede de geração de energia distribuída. Pode-se ver, na Fig. 2.9, a aplicação de um sistema fotovoltaico para beneficiamento de uma comunidade rural realizada no estado do Ceará (CRESESB, 2008).



Figura 2.9 – Sistemas fotovoltaicos em casas geminadas no município de Pentecoste, CE.

Fonte: CRESESB (2008).

Os pontos negativos que podem ser levantados são a baixa eficiência de conversão energética e a manipulação do silício contido nas placas solares, processo poluente que, dependendo dos casos, pode consumir até 25% da energia produzida de maneira limpa durante uma vida útil estimada de cerca de 25 anos (DAI-PRÁ et al., 2014). Assim, um bom estudo precisa ser realizado para verificar o *payback* econômico e ambiental tanto para

grandes áreas de produção como para indivíduos que estejam vendendo energia a partir de módulos fotovoltaicos instalados em suas residências.

#### 2.4.2 Energia solar térmica

Os sistemas concentradores solares utilizam uma combinação de lentes e/ou espelhos para concentrar a irradiância solar direta e gerar energia, seja ela elétrica ou térmica. Essa tecnologia teve um primeiro momento de crescimento comercial indo de 1984 até 1995, voltando a gerar interesse comercial somente a partir de 2005. Entre esse tempo, pesquisas, desenvolvimentos e demonstrações aconteceram, o que fomentou essa tecnologia e a fez ganhar muito espaço nos dias de hoje (LOVEGROVE e STAIN, 2012).

No entanto, a mais alta eficiência está nas grandes plantas de potência de concentração solar. Seus maiores benefícios, porém, são potencializados somente com a maior extensão possível da área de espelhos para concentração (STOFFEL et al., 2010), como apresentado na Fig. 2.10. Esses sistemas de geração de energia elétrica, por meio da conversão de calor em energia elétrica, chegam a atingir eficiências entre 15 a 25% e, dependendo da tecnologia empregada, até 30%, como é o caso do concentrador tipo disco (HEATH e BURKHARDT III, 2011).



Figura 2.10 – Gemasolar: a maior planta de potência por concentração solar térmica do mundo, cuja potência instalada é de 19,9MW, localizada na Espanha

Fonte: Torresol Energy (2018).

A história da concentração solar é ainda mais antiga que a do efeito fotovoltaico, pois é sabido que Arquimedes já havia descrito a capacidade de concentração solar com espelhos, em 200 A.C. (LOVEGROVE e STAIN, 2012). Uma das primeiras máquinas de concentração solar a ser largamente comercializada foi a fornalha solar de Lavoisier (Fig. 2.11), datada de 1794, que atingia temperaturas de 1.750 °C (KALOGIROU, 2009). Apesar disso, foi somente nos anos 1980, mais especificamente na Califórnia, que iniciou o caminho comercial dos

concentradores solares de potência tal qual são conhecidos hoje (LOVEGROVE e STAIN, 2012).

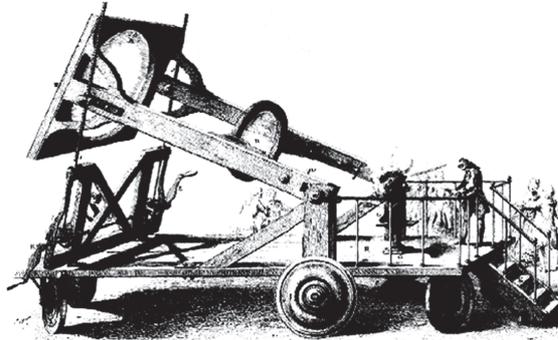


Figura 2.11 – Fornalha solar de Lavoisier.

Fonte: Kalogirou (2009).

Esta geração de energia por concentração depende, no entanto, única e exclusivamente da irradiância solar direta, pois os raios devem atingir os espelhos concentradores em um determinado ângulo que permita sua reflexão correta para a superfície concentradora (Fig. 2.12). Isso faz com que seja necessário um estudo completo da disponibilidade dessa componente da radiação solar e dos três principais elementos de uma grande planta de potência: seleção do sítio de instalação da planta, predição da produção anual de energia e a performance temporal da planta com sua estratégia de operação (STOFFEL et al., 2010).

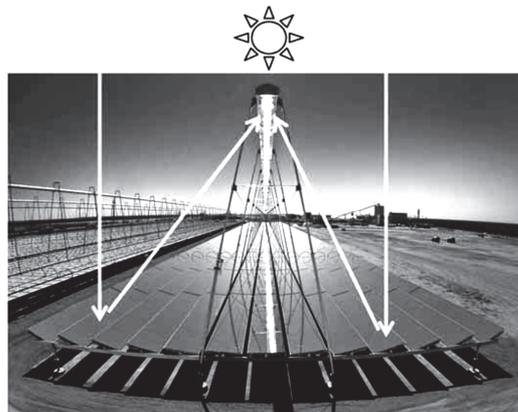


Figura 2.12 – Ângulo de reflexão para o concentrador.

Fonte: Lovegrove and Stain (2012).

Com uma grande eficiência e bom dimensionamento, uma planta de potência de concentração solar pode gerar tanto energia elétrica como energia térmica a altas e médias temperaturas, de acordo com a disponibilidade de irradiância solar direta normal. Assim, a qualidade e confiabilidade dos dados dessa componente da radiação devem ser os melhores possíveis para que a eficiência realmente supere tanto o impacto ambiental quanto o altíssimo investimento econômico.

### 3 MODELAGEM DE COMPONENTES DA RADIAÇÃO SOLAR

A modelagem das componentes da radiação solar provém de estudos datados em mais de sessenta anos. A necessidade de utilização das parcelas separadas da irradiância solar tanto em sistemas fotovoltaicos quanto em sistemas térmicos, a dificuldade de medição de algumas dessas componentes da radiação e até a descoberta de correlações entre as mesmas, tornaram essa uma grande - e importante - área de pesquisa.

#### 3.1 DIFICULDADES DE MEDIÇÃO DAS COMPONENTES

Pireliômetros e piranômetros com seguidores solares são instrumentos muito caros e que necessitam de constante verificação para garantia das variáveis medidas. Piranômetros com anel de sombra são instrumentos mais baratos, porém também necessitam de aferição constante e correção da posição da banda de sombra, de acordo com a declinação solar, para garantia da qualidade de medição.

Essa dificuldade é uma justificativa recorrente em trabalhos de modelagem de componentes de radiação. Desde trabalhos pioneiros, como o de Liu e Jordan (1960), até os mais novos, com metodologias diferenciadas, como os realizados por Paulescu e Blaga (2016), citam o fato das estações solares comumente medirem apenas a irradiância global horizontal. Entre um trabalho e outro há quase 60 anos de evolução tecnológica, porém o problema da medição das parcelas solares ainda permanece, globalmente.

No Brasil, o INPE realiza a medição da irradiância global horizontal em diversas localidades. Para melhorar a disponibilidade de dados solares, o mesmo instituto criou o projeto SONDA. O objetivo desse projeto é alocar recursos físicos e humanos para melhorar a base de dados e realizar validação dos mesmos, por meio da sinalização da confiabilidade de cada dado medido (SONDA, 2018). Pode-se ver, no ANEXO C, um apanhado de confiabilidade de dados de algumas cidades, demonstrando o quão difícil é a manutenção de dados em estações solares remotas.

Tanto no Brasil como no mundo, estações meteorológicas completas que meçam todas as componentes da radiação solar são escassas. Como visto anteriormente, as grandes plantas de potência e até as pequenas unidades domésticas de geração de energia a partir do Sol necessitam de estudos de viabilidade econômica e até ambiental para a melhor eficiência possível na geração de energia útil em relação à disponibilidade das diferentes componentes da radiação solar. Dessa necessidade, cada vez mais crescente, surgiram então os modelos

difusos, cada vez mais precisos e com alta tendência à generalização geográfica de seus resultados.

### 3.2 MODELAGEM DA IRRADIÂNCIA TOTAL DIÁRIA DISPONÍVEL

Um dos primeiros estudos que guiou pensamentos futuros acerca da modelagem é o realizado por Whillier (1953). Nele tem-se a problemática da utilização das parcelas direta e difusa para cálculo efetivo da radiação incidente em coletores inclinados, porém sem a ênfase na modelagem das mesmas. No entanto, algumas das correlações que mais tarde seriam utilizadas em trabalhos do gênero, como comparação entre curvas horárias e diárias de radiação e plotagem em relação ao índice  $f$  ou fração de tempo, podem ser vistas neste trabalho e em trabalhos pioneiros como o de Liu e Jordan (1960), conforme Fig. 3.1. Para Lorenzo (2006), os autores Liu e Jordan foram muito sagazes ao escolher o coeficiente de claridade para caracterizar o clima solar de uma localidade, pois a divisão da irradiância global horizontal pela irradiância extraterrestre horizontal elimina as variações da radiação pelo movimento aparente do Sol, tornando a correlação independente da latitude e, assim, dando ao modelo uma validade universal.

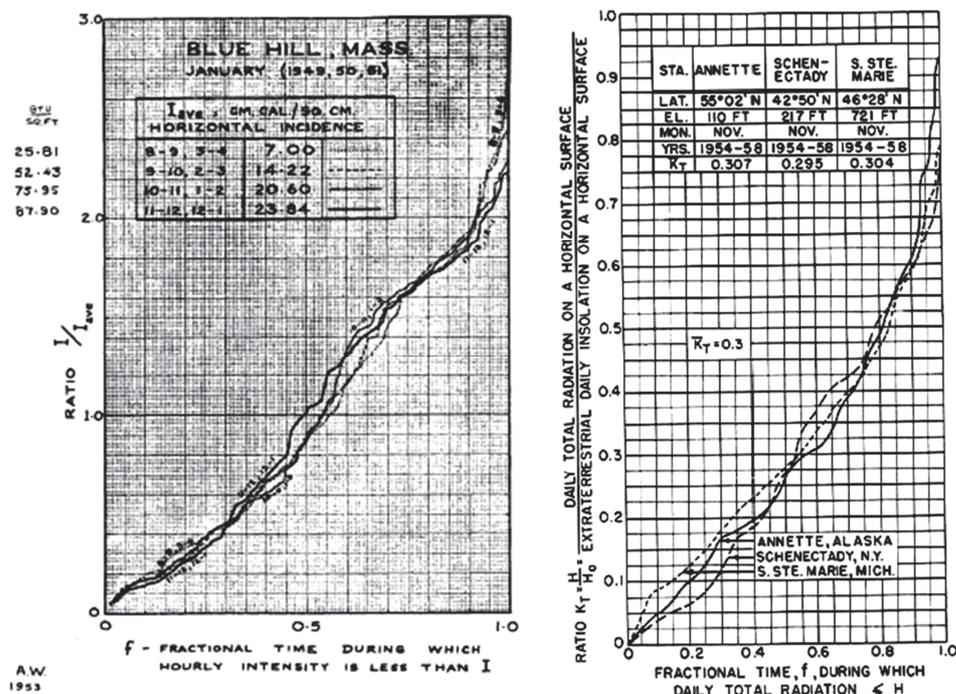


Figura 3.1 – Correlações entre índice  $f$  e coeficiente de claridade<sup>3</sup>.

Fonte: Whillier (1953) e Liu e Jordan (1960).

<sup>3</sup> Como estes gráficos são antigos, optou-se por manter a originalidade dos mesmos e inseri-los no trabalho exatamente da maneira como foram publicados nos artigos originais, evitando assim a perda de informação histórica e o anacronismo gerado a partir da adaptação das mesmas ao padrão atual de publicações científicas.

Pode-se dizer, seguramente, que os primeiros a demonstrar uma relação entre coeficientes de claridade atmosférica, da maneira mais generalizada possível, foram Liu e Jordan (1960). Através de seu estudo, os autores observaram que diferentes localidades, com uma mesma média mensal de coeficiente de claridade, apresentavam distribuições cumulativas de coeficiente de claridade diário em relação à fração de tempo  $f$  semelhantes, independente da altitude e localização. Desta maneira, originou-se o modelo de Liu e Jordan (1960) para obtenção da irradiância global horizontal diária a partir de um coeficiente de claridade médio mensal  $\bar{K}_t$  de acordo com a Eq. (3.1). A Fig. 3.2 mostra que se pode obter a probabilidade de encontrar um valor do coeficiente de claridade diário em relação a um coeficiente de claridade mensal.

$$\bar{K}_t = \frac{\bar{H}}{\bar{H}_0} \quad (3.1)$$

em que  $\bar{H}$  é a irradiância global horizontal diária média mensal,  $\bar{H}_0$  é a irradiância extraterrestre horizontal diária média mensal.

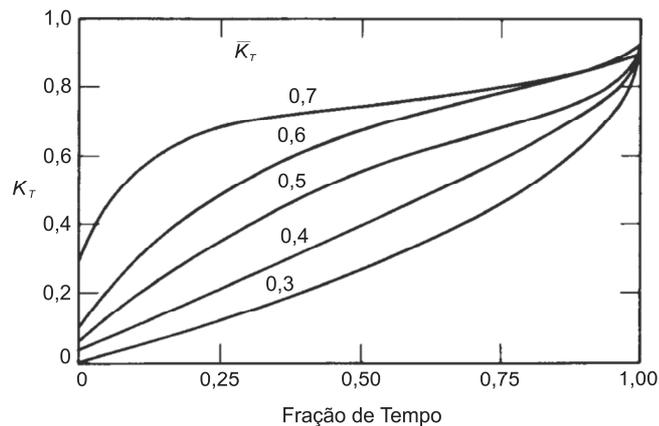


Figura 3.2 – Correlações entre índice  $f$ ,  $K_t$  e  $\bar{K}_t$ .

Fonte: Adaptado de Duffie e Beckman (2006).

Com essa curva é possível saber o percentual de tempo do dia em que haverá disponibilidade de irradiância direta normal. Mais tarde, seu modelo recebeu revisões, como as realizadas nos trabalhos de Bendt et al. (1981), Hollands and Huguet (1983) e Herzog (1985). Estes autores procuraram melhorar problemas de generalização e acurácia do modelo citado anteriormente.

Após esse estudo pioneiro, vários outros surgiram, com modelagens mais aprofundadas e correlações mais complexas, como os estudos de Garrison (1984, 1985), que demonstraram a influência de inúmeros outros parâmetros na relação entre radiação difusa e

global. Subsequentemente, as correlações também evoluíram em escalas de tempo menores, surgindo então os modelos difusos horários.

### 3.3 MODELOS HORÁRIOS UTILIZANDO CORRELAÇÕES

Os modelos apresentados anteriormente, baseados na estimativa de médio e longo prazo, não definiam um comportamento horário das componentes da radiação. A partir das correlações entre o índice  $f$  - fração de tempo - e o coeficiente de claridade dos estudos anteriores, Orgill e Hollands (1977) apresentaram uma metodologia pela qual a fração difusa horária, definida pela Eq. (3.2) e apresentada na Fig. 3.3, e o coeficiente de claridade horário, definido pela Eq. (3.3) são correlacionados para obtenção de um comportamento entre a relação da parcela difusa e da radiação global horizontal (DUFFIE E BECKMAN, 2016).

$$k_d = \frac{G_d}{G_a} \quad (3.2)$$

em que  $k_d$  representa o coeficiente difuso horário,  $G_d$  a irradiação difusa horizontal,  $G_a$  a irradiância global horizontal. O coeficiente de claridade,  $k_t$ , é representado pela Eq. (3.3).

$$k_t = \frac{G_a}{G_{on}} \quad (3.3)$$

em que  $G_{on}$  representa a irradiação extraterrestre horizontal.

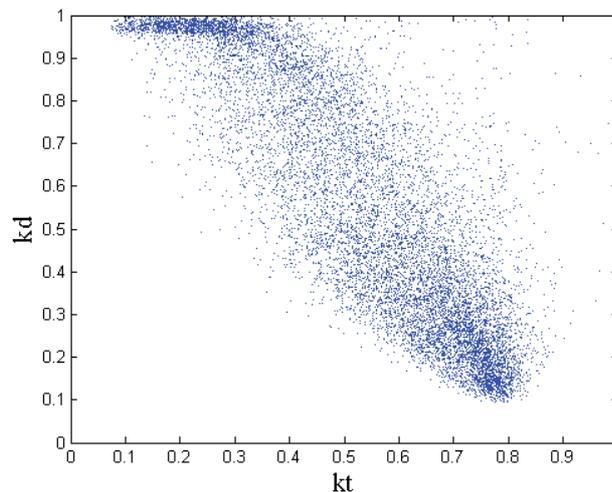


Figura 3.3– Correlação entre a fração difusa e o coeficiente de claridade horário.

A partir da correlação da fração difusa em relação ao coeficiente de claridade, os dados mostrados na Fig. 3.3 são separados em “bins” ou intervalos de valores de  $k_t$ , conforme a Fig. 3.4. Esses valores dos intervalos separados são utilizados como base para a obtenção de equações para o modelo difuso (DUFFIE e BECKMAN, 2006).

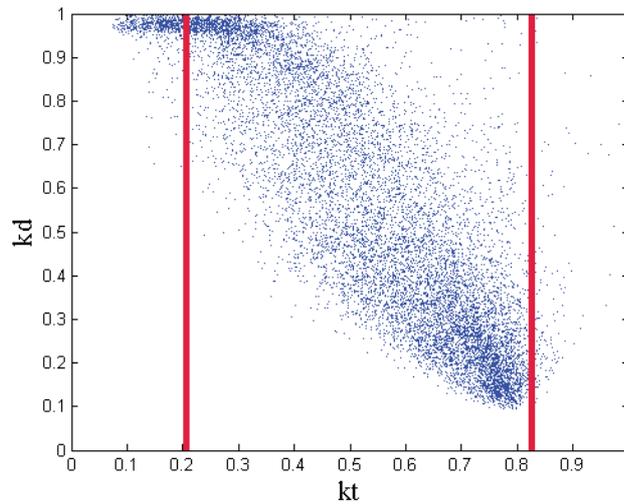


Figura 3.4 – Separação das faixas do coeficiente de claridade  $k_t$ .

Cada um desses intervalos de  $k_t$  representa um tipo de comportamento atmosférico característico. O primeiro representa um céu com maior cobertura de nuvens, resultando em uma fração difusa maior; o terceiro demonstra um comportamento característico de céu claro; já o segundo mostra um comportamento intermediário. Esses diferentes comportamentos resultam em diferentes expressões matemáticas para um modelo difuso horário, como mostrado genericamente pelas Eq. (3.4) a Eq. (3.6). Esse tipo de equação não representa uma hora particular com alta precisão, mas sim o comportamento de um grande número de horas com a maior precisão e generalização possível, de acordo com os dados utilizados para a correlação.

$$k_d = a \pm bk_t \quad (3.4)$$

para o primeiro intervalo de  $k_t$ , em que  $a$  e  $b$  são coeficientes obtidos a partir da primeira faixa de valores da correlação apresentada na Fig. 3.4.

$$k_d = c \pm dk_t \pm ek_t^2 \pm fk_t^3 \dots \quad (3.5)$$

para o segundo intervalo de  $k_t$ , em que  $c$ ,  $d$ ,  $e$ , e  $f$  são coeficientes obtidos a partir da segunda faixa de valores da correlação apresentada na Fig. 3.4.

$$k_d = g \quad (3.6)$$

para o último intervalo de  $k_t$ , em que  $g$  é o coeficiente obtido a partir da terceira faixa de valores da correlação apresentada na Fig. 3.4.

A irradiância difusa é obtida por meio da Eq. (3.7). A partir da irradiância difusa é possível obter a irradiância direta pela Eq. (3.8).

$$G_d = k_d G_a \quad (3.7)$$

$$G_b = G_a - G_d \quad (3.8)$$

em que  $G_b$  é a irradiação direta e  $G_d$  corresponde à irradiação difusa.

Além do modelo apresentado por Orgill e Hollands (1977), outros também utilizaram essa metodologia para seus modelos difusos horários, como Erbs et al. (1982) e Reindl et al. (1990). Esses modelos são detalhadamente apresentados nas próximas seções, juntamente com o modelo de Maxwell (1987), que analisa a fração de radiação direta, não havendo necessidade de obtê-la a partir da radiação difusa.

### 3.3.1 Modelo O&R de Orgill e Hollands (1977)

Os autores desse modelo basearam-se em dados medidos na cidade de Toronto, Canadá. A partir de estudos das correlações entre coeficiente de claridade e fração difusa, chegaram às Eq. (3.9) a Eq. (3.11).

$$k_d = 1,0 - 0,249k_t \quad (3.9)$$

para  $k_t < 0,35$ ;

$$k_d = 1,557 - 1,84k_t \quad (3.10)$$

para  $0,35 \leq k_t \leq 0,75$ ;

$$k_d = 0,177 \quad (3.11)$$

para  $k_t > 0,75$ ;

### 3.3.2 Modelo ER de Erbs et al. (1982)

Modelo proposto a partir de dados de quatro localidades norte-americanas, situadas entre latitudes de 31° N até 42° N. Suas relações polinomiais encontram-se nas Eq. (3.12) a Eq. (3.14):

$$k_d = 1,0 - 0,09k_t \quad (3.12)$$

para  $k_t \leq 0,22$ ;

$$k_d = 0,9511 - 0,1604k_t + 4,388k_t^2 - 16,638k_t^3 + 12,336k_t^4 \quad (3.13)$$

para  $0,22 < k_t \leq 0,80$ ;

$$k_d = 0,165 \quad (3.14)$$

para  $k_t > 0,80$ ;

### 3.3.3 Modelo RE1 e RE2 de Reindl et al. (1990)

Os autores utilizaram dados de cinco localidades, com latitudes entre 28° N e 60° N. Com inúmeras análises, os autores chegaram a mais quatro variáveis climáticas que influem na estimativa dos coeficientes difusos: coeficiente de claridade, seno da altitude solar,

temperatura ambiente e umidade relativa. Dentre os diversos modelos obtidos por Reindl et al. (1990), foram analisadas apenas as versões reduzidas dos mesmos.

O modelo de RE1, que utiliza o seno da elevação solar, é descrito pelas Eq. (3.15) a Eq. (3.17).

$$k_d = 1,020 - 0,254k_t + 0,0123 \sin \alpha \quad (3.15)$$

para  $k_t \leq 0,30$ , com a restrição de  $k_d \leq 1$ ;

$$k_d = 1,4 - 1,749k_t + 0,177 \sin \alpha \quad (3.16)$$

para  $0,30 < k_t < 0,78$ , com a restrição de  $k_d \leq 0,97$  e  $k_d \geq 0,1$ ;

$$k_d = 0,486k_t - 0,182 \sin \alpha \quad (3.17)$$

para  $k_t \leq 0,78$ , com a restrição de  $k_d \geq 0,1$ ;

O modelo de RE2 é descrito pelas Eq.(3.18) a Eq. (3.20):

$$k_d = 1,020 - 0,248k_t \quad (3.18)$$

para  $k_t \leq 0,30$ ;

$$k_d = 1,45 - 1,67k_t \quad (3.19)$$

para  $0,30 < k_t < 0,78$ ;

$$K_d = 0,147 \quad (3.20)$$

para  $k_t \geq 0,78$ .

### 3.3.4 Modelo BRL1 e BRL2 de Ridley et al. (2010)

Citando estudos ainda mais recentes, apresenta-se o realizado por Ridley et al. (2010), na qual os autores partem da pesquisa de Reindl et al. (1990) para adaptar uma função logística com parâmetros físicos para melhor representar o espalhamento de dados reais em correlações de  $k_t$  e  $k_d$ , principalmente em cidades cujas estimativas tradicionais por meio de modelos criados com dados americanos não se mostraram adequadas. Os dados utilizados para encontrar os parâmetros da representatividade das variáveis físicas na sua equação foram de cinco cidades: *Adelaide* (AU), *Darwin* (AU), *Bracknell* (UK), *Lisboa* (PT), *Macau* (MO), *Maputo* (MZ) e *Uccle* (BE). Desta maneira, a forma final do modelo, chamado neste artigo de BRL2, é apresentada na Eq. (3.21).

$$k_d = \frac{1}{1 + e^{-5.38+6.63k_t+0.006AST-0.007\alpha+1.75K_t+1.31\psi}} \quad (3.21)$$

em que  $AST$  é o tempo solar aparente,  $\alpha$  é a altitude solar em graus,  $K_t$  é a média diária do coeficiente horário de claridade e  $\psi$  é a persistência, mostrada na Eq. (3.22):

$$\psi = \begin{cases} (k_{t-1} + k_{t+1})/2 & sr < t < ss \\ k_{t+1} & t = sr \\ k_{t-1} & t = ss \end{cases} \quad (3.22)$$

em que  $sr$  é o ângulo ou horário do nascer do Sol e  $ss$  é o horário ou ângulo do pôr do Sol.

Foram realizados vários testes com dados da cidade de Brasília, onde se procurou conseguir o mesmo comportamento do modelo apresentado nesse artigo. No entanto, o modelo BRL2 não atingiu os resultados esperados. Devido à falta de uma abordagem mais clara na metodologia de obtenção de cada parâmetro físico e de suas constantes (para adaptar o modelo aos dados brasileiros), se optou por incluir somente a primeira equação de Ridley et al. (2010) neste trabalho, conforme a Eq. (3.23). Essa é uma equação logística simples que utiliza somente o coeficiente de claridade horário como variável.

$$k_d = \frac{1}{1 + e^{-5.38 + 6.63k_t}} \quad (3.23)$$

em que  $k_t$  é o coeficiente de claridade horário.

### 3.4 MODELOS HORÁRIOS UTILIZANDO RNA

RNAs são um conjunto de unidades chamados neurônios que conseguem realizar processamento não linear através de diferentes camadas por meio de uma função matemática chamada função de ativação (HAYKIN, 1998). Cada neurônio computacional realiza uma simulação de um neurônio biológico, possuindo múltiplas entradas, pesos sinápticos, função de soma, função de ativação, sinais de entrada e sinal de saída, como mostrado na Fig. 3.5. Um maior detalhamento acerca de seu funcionamento pode ser encontrado em Haykin (1998).

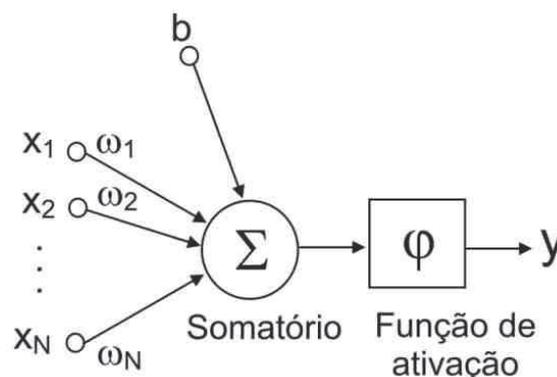


Figura 3.5 – Exemplo de funcionamento de uma RNA.

Fonte: Simas (2018).

A arquitetura das redes neurais consiste em uma camada de entrada, uma ou mais camadas intermediárias (também chamadas de camadas escondidas, as quais são responsáveis

pelo processamento dos dados) e uma camada de saída. Sua estruturação pode ser do tipo *feed forward*, na qual cada camada é ligada à camada posterior e as entradas e saídas seguem um fluxo direto em direção à camada de saída, como mostrado na Fig. 3.6; ou também do tipo *recurrent*, no qual algumas camadas ligam-se às camadas anteriores, mostrada na Fig. 3.7.

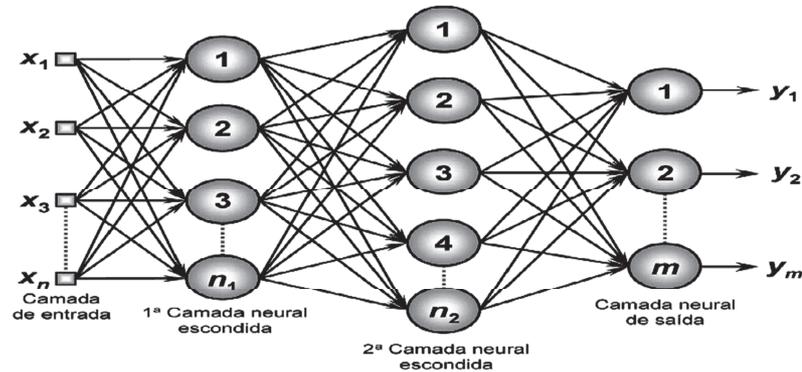


Figura 3.6 – RNA do tipo *feedforward*.

Fonte: Palmiere (2018).

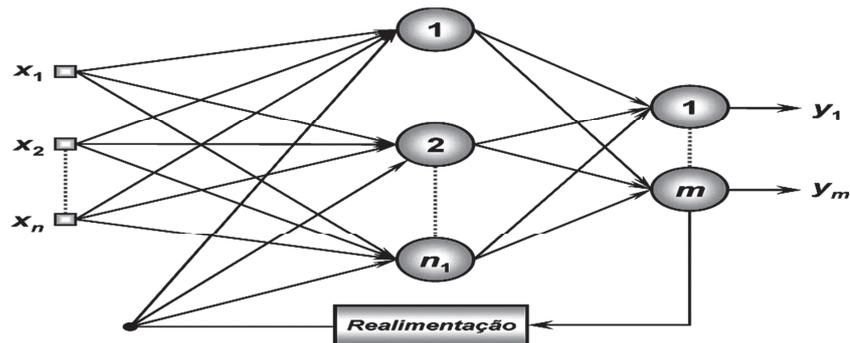


Figura 3.7 – RNA do tipo *recurrent*.

Fonte: Palmiere (2018).

Os principais tipos de treinamento das RNAs são: supervisionado, não-supervisionado e com reforço. O primeiro consiste em utilizar as entradas e saídas esperadas durante o processo de treinamento para ajustes de pesos; o segundo consiste no ajuste automático da camada de saída para representar particularidades do conjunto de dados do treinamento; e o terceiro é um treinamento supervisionado, com um algoritmo que visa a encontrar particularidades no conjunto de dados para reforçar respostas adequadas ao problema em questão.

As RNAs utilizam diferentes arquiteturas e metodologias de treinamento de acordo com o problema em que serão aplicados. Sua maior utilização encontra-se em problemas de previsão e reconhecimento, inclusive das diferentes parcelas da radiação solar. Alguns dos

estudos mais recentes que utilizam RNAs, abordados por este trabalho, são os realizados por Mellit et al. (2010).

### 3.4.1 Modelo de Mellit et al. (2010)

O modelo proposto pelos autores utiliza uma modelagem de RNAs *feed-forward* adaptativa, utilizada anteriormente em aplicações financeiras de negociação para *e-commerce* e previsões de estoque, chamada  $\alpha$ -*model*, como base para prever a radiação global horizontal, difusa horizontal e direta normal para a hora posterior à hora medida. Nas aplicações financeiras, o parâmetro  $\alpha$  representa a fração entre o preço final,  $P_f$ , e o preço inicial,  $P_s$ , por meio da Eq. (3.24):

$$\alpha = \frac{P_f}{P_s} \quad (3.24)$$

Pelo parâmetro  $\alpha$  é possível então decidir pela venda ou compra de um produto. Também é possível calcular o preço  $P_f$  a partir do preço inicial pela Eq.(3.25):

$$P_f = \alpha P_s \quad (3.25)$$

se  $\alpha > 1$ , deve-se realizar a compra, pois o produto irá valorizar;

se  $\alpha < 1$ , deve-se realizar a venda, pois o produto irá desvalorizar.

De forma análoga, o modelo  $\alpha$  pode ser reinterpretado para prever a irradiância  $j+1$ , conhecendo a radiação solar, a duração do dia, a temperatura e a umidade do ar medidos na hora  $j$ . Uma matriz é calculada para obtenção do parâmetro  $\beta$ , que será utilizado para calcular o  $\alpha$  e prever a irradiância a partir desses parâmetros. Os resultados encontrados foram satisfatórios em todas as parcelas da radiação estudadas, demonstrando apenas um pequeno subdimensionamento. Para maiores detalhes, consultar Mellit et al. (2010).

## 3.5 MODELOS DIFERENCIADOS

Além das metodologias anteriores, que já podem ser tratadas como clássicas e muito bem estabelecidas nesta área de pesquisa, há outras que vão além e tentam estabelecer novos conceitos de modelagem de radiação. Esses modelos fogem completamente dos anteriores, pois não utilizam correlações clássicas tampouco RNAs. Suas modelagens não são ortodoxas e tentam definir um terceiro caminho a ser estabelecido e seguido neste tipo de pesquisa. Um bom exemplo é o modelo de Maxwell (1987), que obtém a parcela direta da irradiância global a partir de variáveis físicas, sendo então chamado de modelo quase físico. Seguindo essa

linha, alguns modelos novos e, na falta de adjetivo melhor, inovadores, são os obtidos pelos estudos de Shamshirband et al. (2016), Paulescu e Blaga (2016) e Basseto et al. (2018).

### 3.5.1 Modelo MAX de Maxwell (1987)

É o modelo utilizado para obtenção da fração direta a partir do coeficiente de claridade. Seus dados de entrada são o coeficiente de claridade e a massa de ar. É conhecido como um modelo quase físico por obter o índice direto máximo para uma determinada massa de ar. É representado pela Eq. (3.26) à Eq. (3.31)(3.28).

$$a = 0,512 - 1,56k_t + 2,286k_t^2 - 2,222k_t^3 \quad (3.26)$$

$$b = 0,310 + 0,962k_t \quad (3.27)$$

$$c = -0,280 + 0,932k_t - 2,048k_t^2 \quad (3.28)$$

para  $k_t \leq 0,60$ ;

$$a = -5,743 + 21,77k_t - 27,49k_t^2 + 11,56k_t^3 \quad (3.29)$$

$$b = 41,40 - 118,5k_t + 66,05k_t^2 + 31,90k_t^3 \quad (3.30)$$

$$c = -47,01 + 184,2k_t - 222,0k_t^2 + 73,81k_t^3 \quad (3.31)$$

para  $k_t > 0,60$ ;

A variação de  $k_b$  é dada por:

$$\Delta k_b = a + b \exp(cm) \quad (3.32)$$

em que  $m$  corresponde a massa de ar atmosférico (para maiores detalhes, consultar Duffie e Beckman, 2006).

A fração direta se determina por:

$$k_b = k_{bc} - \Delta k_b \quad (3.33)$$

em que  $k_{bc}$  é o maior valor de  $k_b$  para uma determinada massa de ar  $m$ , conforme Eq. (3.34):

$$k_{bc} = 0,886 - 0,122m + 0,0121m^2 - 0,000653m^3 + 0,000014m^4 \quad (3.34)$$

### 3.5.2 Shamshirband et al. (2016)

Esse estudo utilizou sete anos de dados, de 2006 a 2012, medidos pela estação do aeroporto da cidade de Kernan, capital da província de Kernan, Irã. Essa cidade é localizada na parte mais ensolarada da província, entre 30° 29' N e 57° 06' E, e está situada a 1760 m de altitude. Duas técnicas foram utilizadas para solucionar o problema da obtenção da irradiância direta diária nesta cidade: Máquina de Suporte Vetorial (MSV) e Transformada Wavelet (TW). Os primeiros cinco anos de dados foram utilizados para obtenção do modelo enquanto os dois últimos foram utilizados como dados de teste.

Utilizando uma Transformada *Wavelet* Discreta (TWD), o modelo consegue separar diferentes padrões em cinco ondas, cada uma delas passa por um vetor SVM onde a informação útil acerca do comportamento da base de dados é obtida. Com todas as séries de ondas devidamente analisadas, a saída geral do modelo difuso é obtida, com a previsão da radiação difusa horizontal diária.

Os resultados obtidos foram comparados com outros três modelos aplicados com essa mesma base de dados: MSV-FBR (Função de Base Radial), RNA e um modelo empírico tradicional. Os resultados para este conjunto de dados representativos de uma porção geográfica muito pequena foram muito satisfatórios e superaram todos os outros modelos de confronto, conforme pode ser visto na Tab. 3.1.

Tabela 3.1 – Comparação entre modelos obtidos pelo estudo.

Modelo	MBE (MJ/m <sup>2</sup> )	RMSE (MJ/m <sup>2</sup> )	R
SVM–WT	0,5757	0,6940	0,9631
SVM–RBF	1,0877	1,2583	0,8599
RNA	1,1267	1,3184	0,8392
Modelos Lineares	1,2171	1,4548	0,8156

### 3.5.3 Paulescu e Blaga (2016)

No trabalho de Paulescu e Blaga (2016) foram utilizados dados obtidos da cidade de Timisoara, Romênia, medidos pela estação solar da Universidade do Oeste de Timisoara. Essa cidade fica a 45° 46' N, 21° 25' E e 85 m de altitude, possui uma temperatura média elevada, elevada umidade relativa e um verão muito quente. Os autores utilizaram métodos regressivos para encontrar diferentes variáveis preditoras para obter modelos com maior precisão e maior generalização, como realizado anteriormente por Reindl et al. (1990). Cerca de 75% da base de dados foi utilizada para a obtenção dos modelos, e 25% foi utilizada para verificação e testes.

Diversos modelos com variáveis preditoras estabelecidas por métodos de regressão foram obtidos. Essas variáveis eram compostas por coeficientes de claridade, média diária de coeficientes de claridade horários, elevação, temperatura, brilho relativo do Sol, entre outras. Dentre todos os modelos do estudo, o que mais se destacou foi o modelo RB7, com coeficiente de correlação  $R^2 = 0.930$ . Segundo os autores, esse foi o primeiro modelo a utilizar o brilho relativo do Sol como variável preditora em um modelo regressivo.

Esse modelo teve como base o estudo da Eq. (3.35), que demonstrou a relevância estatística do nascer do sol (modelo PB6). A partir deste, a inclusão das relações entre o coeficiente de claridade culminou na Eq. (3.36), relativa ao modelo PB7.

$$k_d = 0,972 - 0,69\sigma \quad (3.35)$$

em que  $\sigma$  representa o brilho relativo do Sol.

$$k_d = 0,964 - 0,584\sigma - 1,0038(k_t - 0.514)\theta(k_t - 0.514) \quad (3.36)$$

em que  $\theta_z$  representa o ângulo de zênite.

### 3.5.4 Basseto et al. (2018)

Neste estudo, os autores utilizam diversos modelos para estimativa de radiação, entre os quais, uma RNA tipo RFB (Rede de Função de Base Radial) e uma RNA tipo SAINF (Sistema Adaptativo de Inferência Neuro Fuzzy). A primeira utiliza uma rede com uma camada intermediária com função de ativação de base radial, como gaussiana, já a segunda, utiliza lógica fuzzy juntamente com a RNA.

Para treinamento da rede, foram utilizados dados da cidade de Botucatu entre janeiro de 2000 a dezembro de 2006. As variáveis de entrada foram: irradiação global horizontal, irradiação de ondas longas atmosféricas, coeficiente de claridade, irradiação extraterrestre horizontal, declinação solar e nebulosidade. Os dados de teste foram retirados da própria base de teste, através de um ano típico e atípico obtidos com metodologia similar à do Ano Meteorológico da Organização Mundial de Meteorologia. Os melhores resultados das RNAs para o ano típico local variaram entre 26,48 a 29,63% na rede RFB, e na rede SAINF, os mesmos variaram entre 22,63 a 25,13%. No ano atípico, os erros variaram entre 24,69 a 26,93% (RFB) e 21,19 a 22,76% (SAINF). Levando em conta a situação de testes, apenas na própria localidade, a ambiguidade da informação da retirada dos dados da base de teste (não há definição se realmente foram retiradas antes do treinamento), os resultados apresentaram-se como muito melhores para o ano meteorológico atípico do que para o típico. A complexidade de cálculo das variáveis utilizadas, a falta de informação sobre os piores resultados obtidos e a variação do erro conforme novos treinamentos da rede são realizados, demonstram que o modelo necessita de outras combinações e parâmetros, como já fora concluído pelos próprios autores. Também há necessidade de mais testes de robustez e generalização de utilização (principalmente pela escolha de variáveis muito específicas como nebulosidade).

## 4 MODELO GERAL COM UTILIZAÇÃO DE REDES NEURAIS ARTIFICIAIS

O modelo apresentado nesta dissertação é baseado na utilização de variáveis físicas, já definidas como muito influentes na fração difusa pela literatura, utilizando-as como entrada de uma RNA cuja saída é a fração difusa correspondente aos dados de entrada. Assim, o modelo em questão é computacional, mas simples, ainda mais preciso que os paradigmas clássicos, porém generalizável geograficamente.

Para o seu desenvolvimento, limpeza dos dados, treinamento das RNAs, testes, avaliações de erro, etc, foi desenvolvida uma série de códigos no software MATLAB, compilados em um sistema sob licença GNU/GPL, chamado ADASYS (*Atmospheric Data Analysis System*). Esses sistemas auxiliares visam a facilitar o processamento e a reprodutibilidade/expansão da utilização do modelo pela facilidade de codificação e operação dos sistemas em novos estudos que visem a desenvolver ainda mais essa modelagem para obtenção da parcela difusa da radiação, em grandes extensões territoriais.

As variáveis físicas escolhidas para integrar o modelo, além do tradicional coeficiente de claridade, foram: elevação solar, temperatura, umidade relativa e persistência atmosférica. Dessa maneira, a arquitetura das RNAs do presente estudo é formada por cinco neurônios de entrada, representando as variáveis físicas citadas, um neurônio de saída, representando o coeficiente difuso, e neurônios na camada escondida (variados de 2 a 9, de acordo com a necessidade de cada aplicação). O valor do erro  $\epsilon$ , para parada do treinamento, foi de 0,01, a proporção dos valores da entrada foi fixada em 0,6, a proporção dos dados de teste foi fixada em 0,2 e o tipo da camada escondida foi variada entre linear e sigmoide durante cada ciclo de testes. Cada configuração da camada de entrada foi treinada três vezes em cada configuração para verificar a robustez da convergência da RNA.

### 4.1 MODELO BASE PARA O ESTUDO

Este modelo teve seu desenvolvimento iniciado a partir da metodologia desenvolvida nos estudos realizados por Pillot et al. (2012) e Pillot (2014), quando da estimativa da temperatura para a cidade de Djibouti, capital da República do Djibouti. A temperatura foi estimada a partir de uma RNA *feed forward*, cuja camada escondida era composta de 20 neurônios do tipo sigmoide e as entradas correspondiam às variáveis de temperatura da superfície terrestre da linha correspondente, valor do erro associado, ângulo de zênite solar e fração do ano. A saída da rede era correspondente à temperatura ambiente medida pela

estação meteorológica local. Para o treinamento da rede foram utilizadas quatro estações na cidade de Djibouti e do distrito de Dikhil, onde as entradas para treinamento foram dados de três estações e os testes foram realizados com dados de outra estação ainda não conhecidos pela rede para avaliação da generalização geográfica do modelo.

## 4.2 VARIÁVEIS DE ENTRADA

Um dos pontos chave deste estudo foi a escolha de variáveis físicas para a entrada da RNA que proporcionasse um aprendizado mais generalista da correlação difusa. Como neste trabalho foram realizados os testes iniciais dessa metodologia aplicada a modelos difusos horários, optou-se por utilizar a literatura existente - Reindl et al. (1990) e Ridley et al. (2010) - e suas justificativas para a seleção final das cinco variáveis do modelo proposto.

O coeficiente de claridade é a base para todos os modelos difusos, a persistência demonstra a inércia atmosférica e a variabilidade da radiação entre duas leituras de radiação, a elevação solar está relacionada diretamente com a massa de ar (quanto mais elevada a posição do Sol menor é a massa de ar atmosférico para interagir com a radiação), a umidade interfere no espalhamento da radiação e a temperatura ambiente (temperatura de bulbo seco) demonstra variações relacionadas a dias sem nuvens, além de relações no aumento da temperatura conforme aumenta a incidência de radiação solar.

## 4.3 BASE DE DADOS PARA TESTE

Os dados utilizados para treinamento das RNAs são os providos pela rede SONDA. O critério de seleção para treinamento das RNAs foram os dados de todas as cidades que estão sinalizados como confiáveis pela avaliação desta mesma instituição responsável. A base de dados completa contém quatorze localidades diferentes: São Martinho da Serra, Sombrio, Chapecó, Florianópolis, Joinville, Curitiba, Ourinhos, Campo Grande, Cachoeira Paulista, Cuiabá, Brasília, Rolim de Moura, Palmas, São Luís, Caicó e Natal. Pode-se ver na Fig. 4.1 que as estações compreendem uma vasta gama de cidades brasileiras localizadas em todas as subdivisões geográficas e climáticas do país.

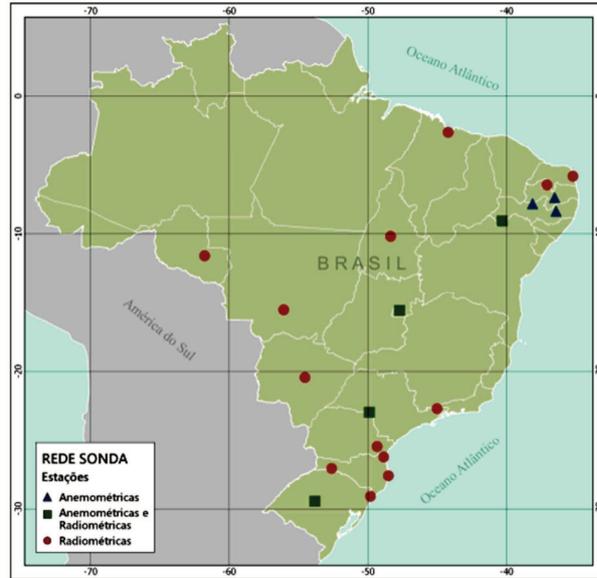


Figura 4.1 – Estações do projeto SONDA.

Fonte: SONDA (2018).

Para este estudo, delimitou-se a utilização de seis cidades para a base de dados, as quais possuíam leituras revisadas pelo algoritmo de limpeza na versão 3.3 ou superior, sendo esse um dos meios de validação mais robustos dentro do sistema de medição do SONDA. As cidades selecionadas para os testes, juntamente com o intervalo de tempo dos dados, foram: Petrolina (2013-2016), Palmas (2013-2016), Brasília (2011-2015), Cachoeira Paulista (2014-2016), Florianópolis (2013-2016) e São Martinho da Serra (2013-2016). Na Fig 4.2 pode-se ver a distribuição geográfica dessas cidades, abrangendo uma boa parte do território e condições climáticas do país.



Figura 4.2 – Estações SONDA utilizadas para modelagem e treinamento.

Antes de sua utilização para testes, todos os dados, medidos minuto a minuto, passaram por um algoritmo de verificação que avaliou e corrigiu dados marcados com grande possibilidade de erro, de acordo com os arquivos DQC<sup>4</sup> fornecidos para cada localidade. Os dados foram preenchidos com a média dos valores válidos do minuto anterior e posterior a sua contagem consecutiva e, passando de 4 horas de medição com incerteza consecutiva, as mesmas eram marcadas como dados faltantes. Os dados utilizados estavam em condições razoavelmente boas, alguns *gaps*, exclusões por falta de qualidade mínima e uma pequena quantidade não passou no algoritmo de qualidade mais leve (o que não os qualifica como ruins). Em seguida, as médias horárias foram obtidas para serem utilizadas nas estimativas somando 46.643 linhas de dados válidos. A qualidade de dados válidos pode ser vista no APÊNDICE A, mostrando as correlações entre os coeficientes de claridade e difusos das cidades utilizadas para testes. Os dados limpos de Petrolina possuem coeficientes de claridade até 1,1 e os de Florianópolis, até 1,01.

Para conversão dos dados de medições minuto a minuto para hora a hora, implementou-se um software escrito em MATLAB que também realiza a limpeza de dados suspeitos e comprova a qualidade a partir das cidades já convertidas para o padrão horário. Durante esse procedimento, descobriu-se uma "defasagem de tempo"<sup>5</sup> nas medições do SONDA para cada cidade em relação ao fuso horário informado (algo facilmente observável pelo pico de radiação solar em um dia claro). Para corrigir isso, foi desenvolvida uma metodologia de avaliação da defasagem de tempo nas medidas dos dados de irradiância solar para cálculo correto da irradiância extraterrestre horizontal horária. Os dados horários selecionados, a partir desse critério, foram primeiramente normalizados e após organizados de maneira aleatória para o treinamento das diferentes arquiteturas de RNA a serem testadas durante a pesquisa. Assim evita-se a influência da continuidade temporal no aprendizado das redes para que as mesmas apenas interpretem a influência das diferentes variáveis físicas escolhidas para o presente modelo, no coeficiente difuso medido para uma mesma hora.

#### 4.4 TESTES DE PRECISÃO LOCAL E GENERALIZAÇÃO GEOGRÁFICA

Neste estudo foram realizados dois testes independentes, um para precisão local e outro para generalização geográfica. No primeiro, 80% dos dados de uma localidade eram

---

<sup>4</sup> *Data Qualification Code*, arquivo que mostra a qualidade dos dados de acordo com testes realizados pelo projeto SONDA.

<sup>5</sup> Algumas cidades ficam com seus dados deslocados em relação à hora UTC informada. Algo observado ao verificar o horário de meio dia solar em dados de dias de céu claro.

utilizados para treinar uma rede que possuía de 2 a 5 neurônios na camada escondida (variando seu número a cada ciclo de testes) e os dados restantes da mesma localidade eram utilizados para obter o coeficiente difuso com a rede treinada. No segundo, uma localidade teste era escolhida, então os dados das outras cinco localidades restantes eram utilizados para treinar uma rede que possuía de 3 a 9 neurônios na camada escondida (variando seu número a cada ciclo de testes). Após este treinamento, os coeficientes difusos eram obtidos a partir dos dados da cidade teste (não utilizada no treinamento da rede).

Esses dois tipos de testes foram realizados para garantir que a metodologia de treinamento das RNAs realmente se mostrasse efetiva para a utilização em outras regiões próximas à localidade cujos dados foram utilizados para teste. Estas metodologias também garantem que, ao treinar uma RNA com dados de mais de uma localidade, as regiões entre elas, distante delas e até localidades de outros países com comportamento climático semelhante, são capazes de utilizar essa mesma RNA e ainda assim estimarem resultados com alta confiança. Os modelos utilizados para comparação dos resultados foram os obtidos por Orgill e Hollands (1977), Erbs et al. (1982), Reindl et al. (1990) e Ridley et al. (2010). Tanto a base de testes de generalização geográfica quanto a de precisão local foram utilizadas em cada um desses modelos para uma melhor eficiência comparativa em relação às metodologias que utilizam RNAs com variáveis físicas em seus neurônios de entrada.

#### 4.5 TESTES EXTREMOS DE ROBUSTEZ

Para testar ao máximo a robustez deste modelo, realizou-se um teste de extremos para averiguar o funcionamento e poder de generalização da RNA. A rede possuía 7 neurônios na camada escondida, função de ativação sigmoide e todos os parâmetros restantes idênticos ao que já foi descrito anteriormente. Ela foi treinada a partir de toda a base de dados horária das seis estações SONDA utilizadas neste estudo, e as cidades teste foram quatro localidades dos Estados Unidos cujos dados meteorológicos estavam disponíveis em formato TMY3 no site da NREL. São elas: *Anchorage (Alaska)*, *Tucson (Arizona)*, *San Francisco (California)*, e *New York City (New York, Aeroporto JFK)*.

Todas as cidades teste representam um tipo de clima que extrapola as características das cidades brasileiras, seja com climas temperados e estações bem definidas, climas desérticos (Tucson possui altos índices de radiação solar no verão e baixa pluviosidade anual) ou climas com mínimas negativas altas (Anchorage possui temperaturas mínimas médias em torno de  $-11^{\circ}\text{C}$ ). Todas as cidades tiveram dados menores que  $30 \text{ W/m}^2$  excluídos, com

exceção de Anchorage, que possui muitos dias do ano com relativamente pouca luminosidade causada pela sua localização geográfica.

Os arquivos TMY3 já possuíam leituras de base horária e somente a temperatura foi convertida para temperatura absoluta (Kelvin). O ângulo horário foi calculado de acordo com cada zona de tempo da hora local (unidade de tempo utilizada pelas estações TMY3). O modelo de confronto escolhido foi o de Erbs et al. (1982) que utiliza as cidades de *Fort Hood (Texas)*, *Maynard (Massachusetts)*, *Raleigh (North Carolina)* e *Livermore (California)*. Todas essas cidades são norte-americanas e localizadas entre latitudes de 31° N até 42° N. Pode-se observar na Fig. 4.3 a relação entre as cidades utilizadas para treinar a RNA (círculos em vermelho), as cidades utilizadas pelo modelo de Erbs et al. (1982) (círculos em roxo) e as cidades cujos TMY foram utilizados para testar os modelos (círculos em verde).



Figura 4.3 – Diferentes cidades utilizadas nos testes extremos de robustez.

#### 4.6 QUANTIFICAÇÃO DO DESVIO

Para quantificar o desvio entre valores calculados e medidos, foram utilizados o desvio médio relativo (*MBEr*), que demonstra se a estimativa está positiva ou negativa, e o desvio quadrático médio relativo (*RMSEr*), que demonstra a dispersão dos pontos em relação à

variável de referência medida. Esses métodos de cálculo de desvio são definidos pelas Eq. (4.1) e (4.2).

$$MBEr = \frac{1}{n} \sum_{i=1}^i \frac{g_{calc} - g_{med,i}}{g_{med,i}} 100 \quad (4.1)$$

$$RMSEr = \frac{1}{n} \sum_{i=1}^i \sqrt{\left(\frac{g_{calc} - g_{med,i}}{g_{med,i}}\right)^2} 100 \quad (4.2)$$

em que  $n$  é o número de dados considerados na análise,  $g_{calc,i}$  é a valor calculado a partir de alguma estimativa e  $g_{med,i}$  é a variável de referência medida.

#### 4.7 COEFICIENTE DE CORRELAÇÃO LINEAR

Coefficientes de correlação são indicadores que demonstram a intensidade entre a correlação de duas variáveis (CORREA, 2003). Há vários indicadores desta correlação, porém neste estudo é utilizado o coeficiente de correlação linear de Pearson, que mede o grau de ajustamento em relação a uma reta e é definido pela Eq. (4.3).

$$r = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2][n \sum y_i^2 - (\sum y_i)^2]}} \quad (4.3)$$

em que  $n$  é o número de observações,  $x_i$  é a variável independente e  $y_i$  é a variável dependente.

O valor do coeficiente  $r$  é limitado ao intervalo [+1, -1] e corresponde a uma correlação positiva ou negativa, respectivamente. Quanto mais próximo dos extremos, mais forte é a correlação, e quanto mais próximo de 0, mais fraca ela é, indicando que o valor pouco influencia linearmente, tanto de maneira crescente quanto decrescente, a variável escolhida. Resultados próximos a zero podem também indicar que a variável independente influencia a variável dependente de alguma outra maneira que não seja linear (CORREA, 2003).

## 5 RESULTADOS

Os resultados obtidos para cada localidade de testes são apresentados nas figuras a seguir. A Fig. 5.1 mostra o erro *RMSEr* dos modelos empíricos. As Fig. 5.2, 5.6, 5.10 e 5.11 mostram a comparação dos erros *RMSEr* obtidos pelas diferentes configurações das RNAs em relação aos erros obtidos por meio dos modelos empíricos. As Fig. 5.3, 5.7, 5.11 e 5.15 mostram a comparação da correlação entre coeficientes de claridade e coeficientes difusos obtidos por meio de algumas RNAs em relação à mesma correlação obtida por meio de dados medidos pelas estações SONDA. As Fig. 5.4, 5.5, 5.8, 5.9, 5.12, 5.13, 5.16 e 5.17 mostram a dispersão dos erros da estimativa dos coeficientes difusos obtidos por meio das RNAs em comparação com os coeficientes difusos obtidos por meio do modelo de Erbs et al. (1982). As Tab. 5.11 e 5.13 mostram o coeficiente de correlação '*r*' para a melhor configuração média de neurônios em testes de generalização geográfica e as Tab. 5.12 e 5.14 mostram o coeficiente de correlação '*r*' para a melhor configuração média de neurônios em testes de precisão geográfica. A Fig. 5.18 mostra os resultados dos testes extremos de robustez. Os resultados detalhados de erros *RMSEr* e *MBEr* podem ser vistos no APÊNDICE B, C, D, E, F e G.

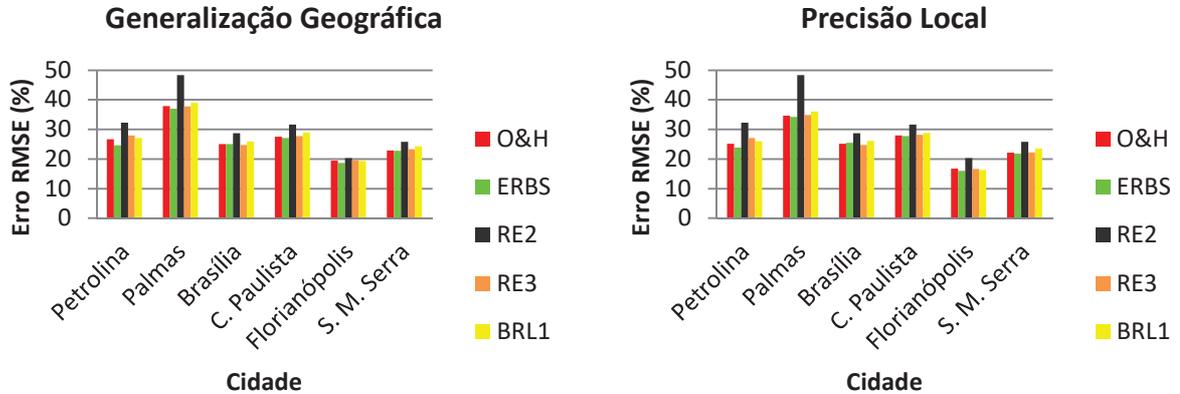


Figura 5.1 – Erro *RMSEr* (%) para as diferentes cidades e diferentes tipos de modelos/testes.

Tanto o modelo com o menor desvio, quanto o modelo com o maior desvio em relação às tendências de erro, tanto *MBEr* como *RMSEr*, variaram pouco entre as cidades. O modelo de Erbs et al. (1982) obteve os melhores resultados de *MBEr* e *RMSEr* de acordo com a tendência de desvio dos outros modelos testados e diferiu pouco desta tendência em apenas uma cidade, sendo então o modelo mais generalizável dentre todos, corroborando assim a afirmação de Ricieri et al. (1995), na qual o autor verificou que curvas polinomiais de quarto grau diferem muito pouco umas das outras independente da localidade utilizada para obtenção da curva de aproximação.

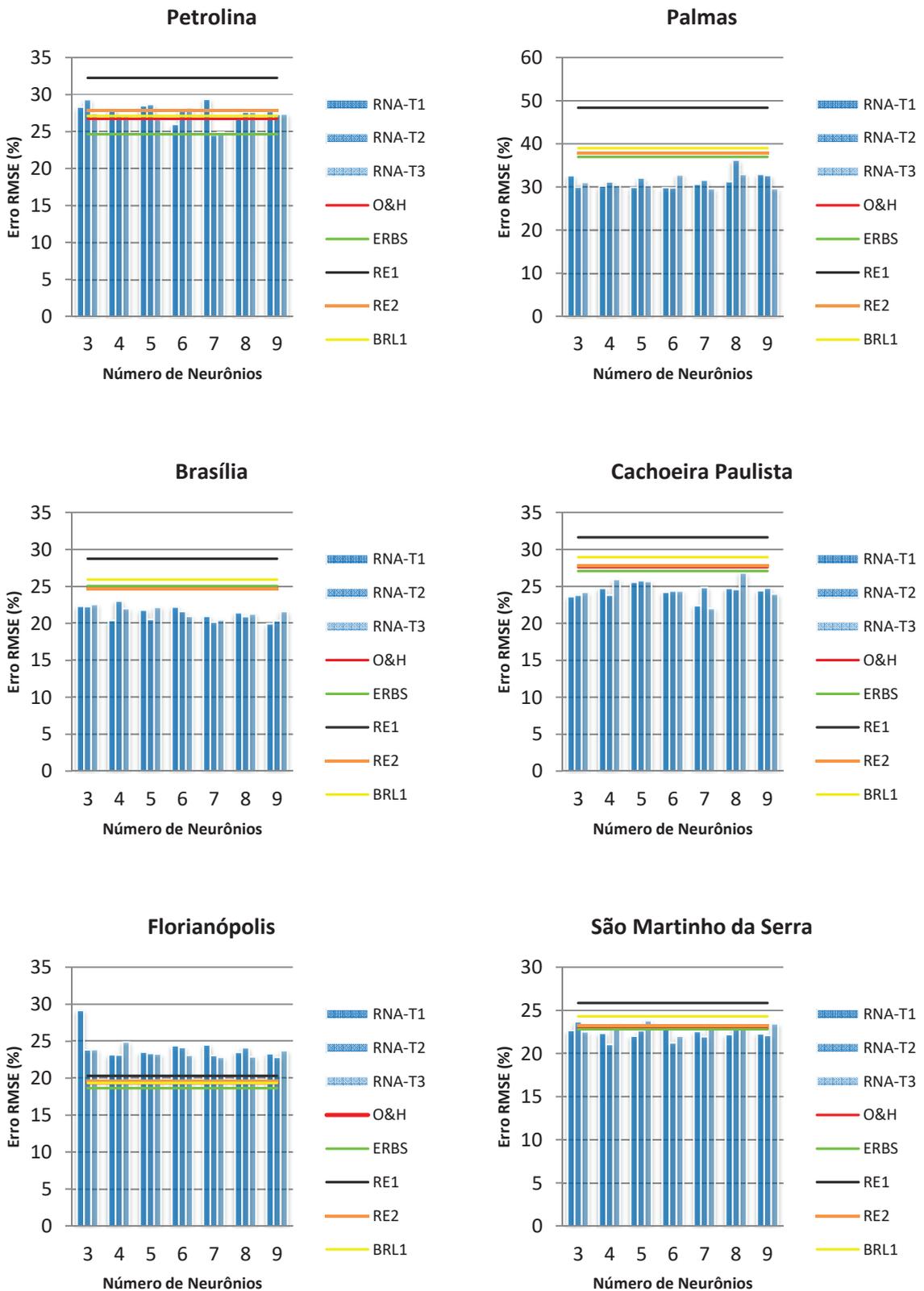


Figura 5.2 – Comparação dos erros  $RMSEr$  obtidos pelas RNAs com função de ativação sigmoide nos testes de generalização geográfica em relação aos erros obtidos por meio dos modelos empíricos.

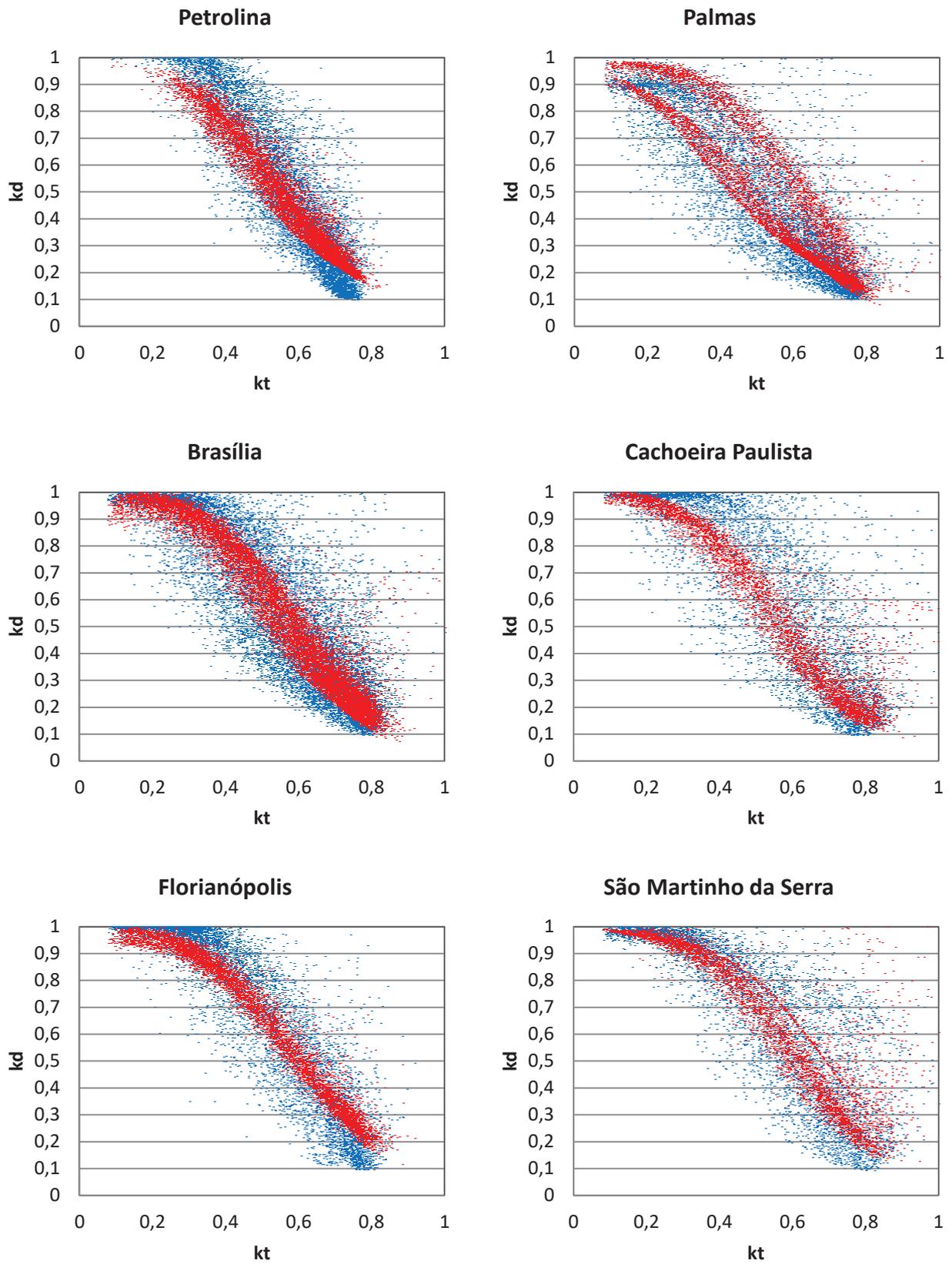


Figura 5.3 – Correlação entre dados medidos [azul] em relação aos dados estimados com menor erro [vermelho] pela RNA de sete neurônios com função de ativação sigmoide nos testes de generalização geográfica.

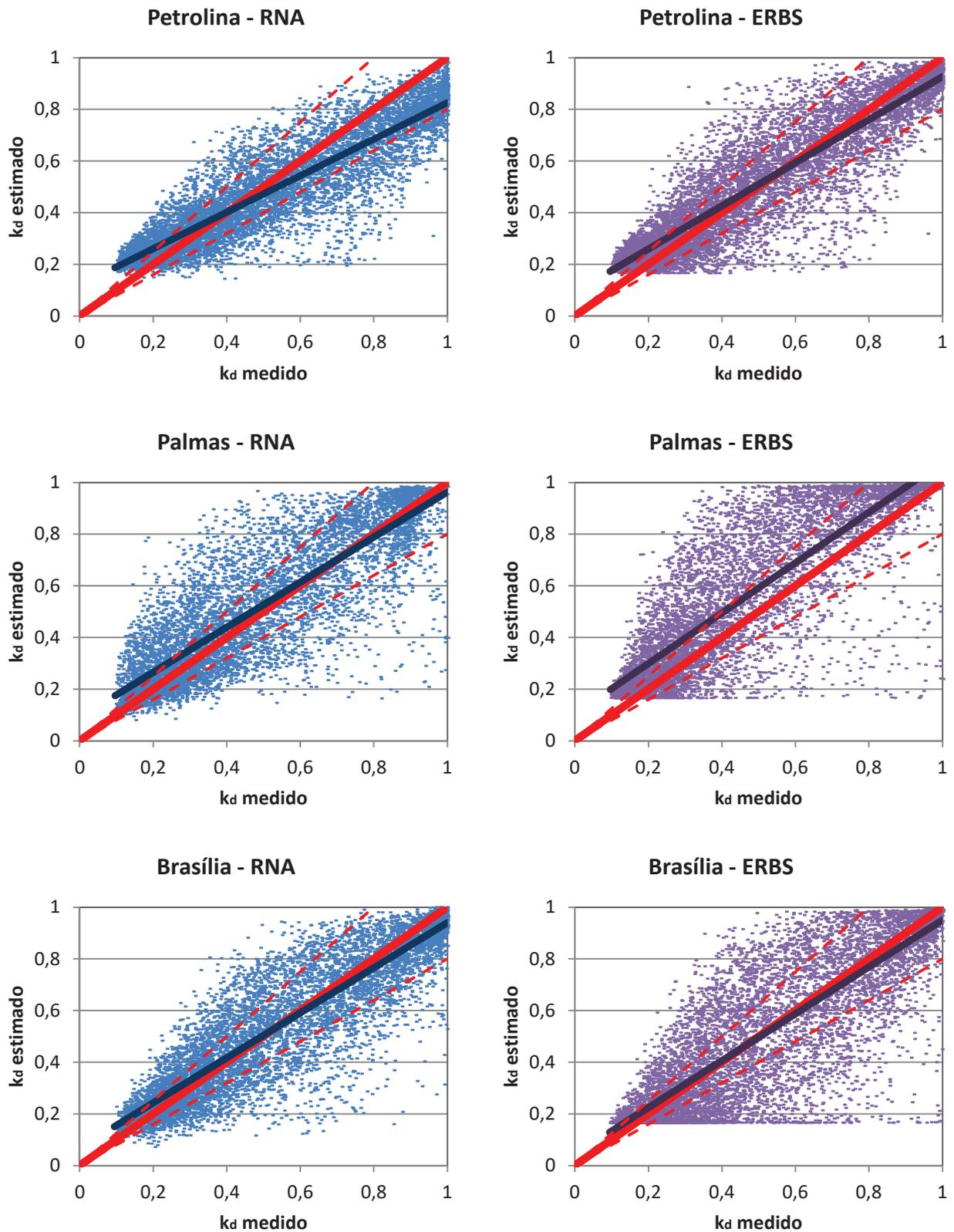


Figura 5.4 – Comparação entre estimativas realizadas pela RNA de sete neurônios em testes de generalização geográfica com função de ativação sigmoide e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA.

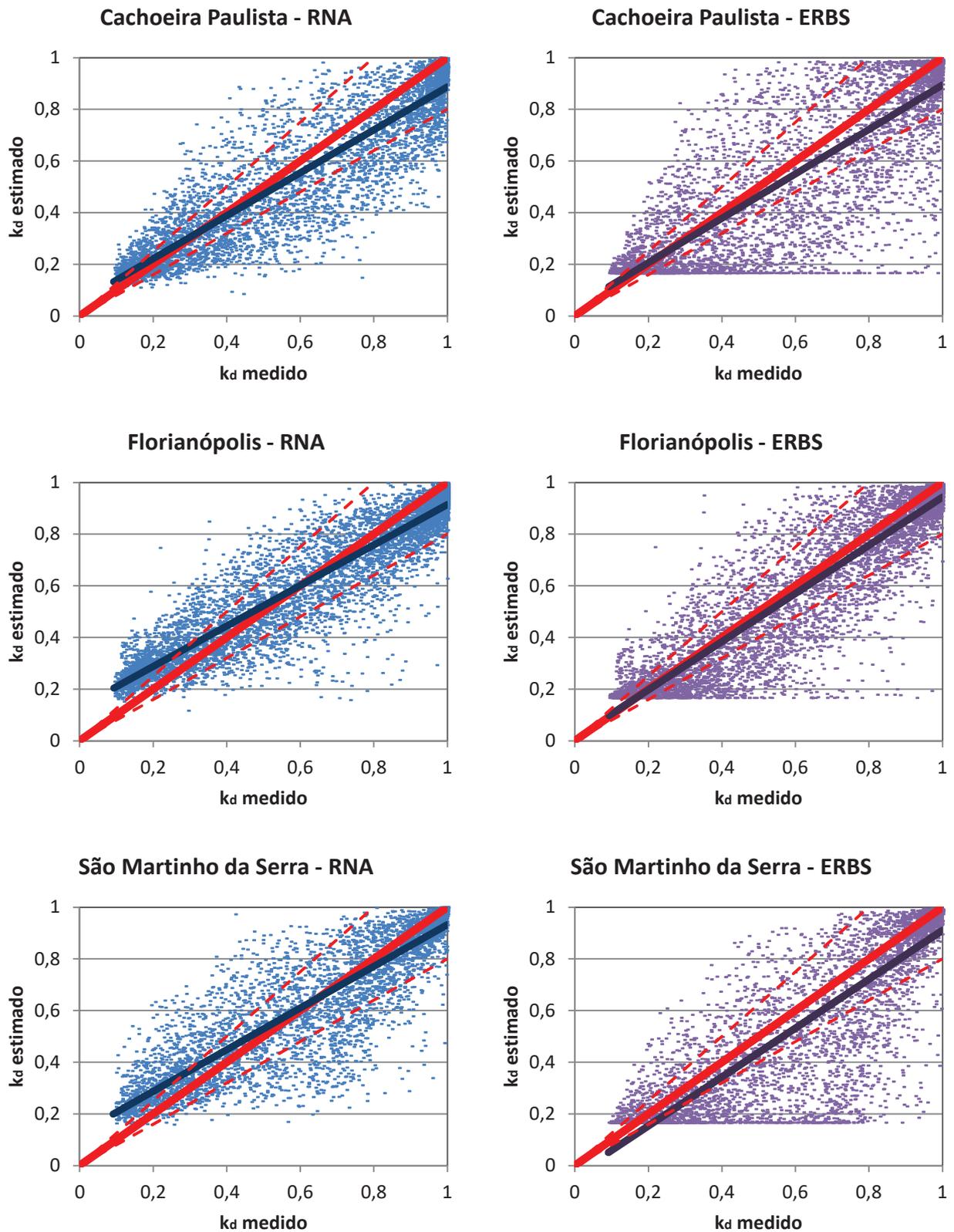


Figura 5.5 – Comparação entre estimativas realizadas pela RNA de sete neurônios em testes de generalização geográfica com função de ativação sigmoide e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA.

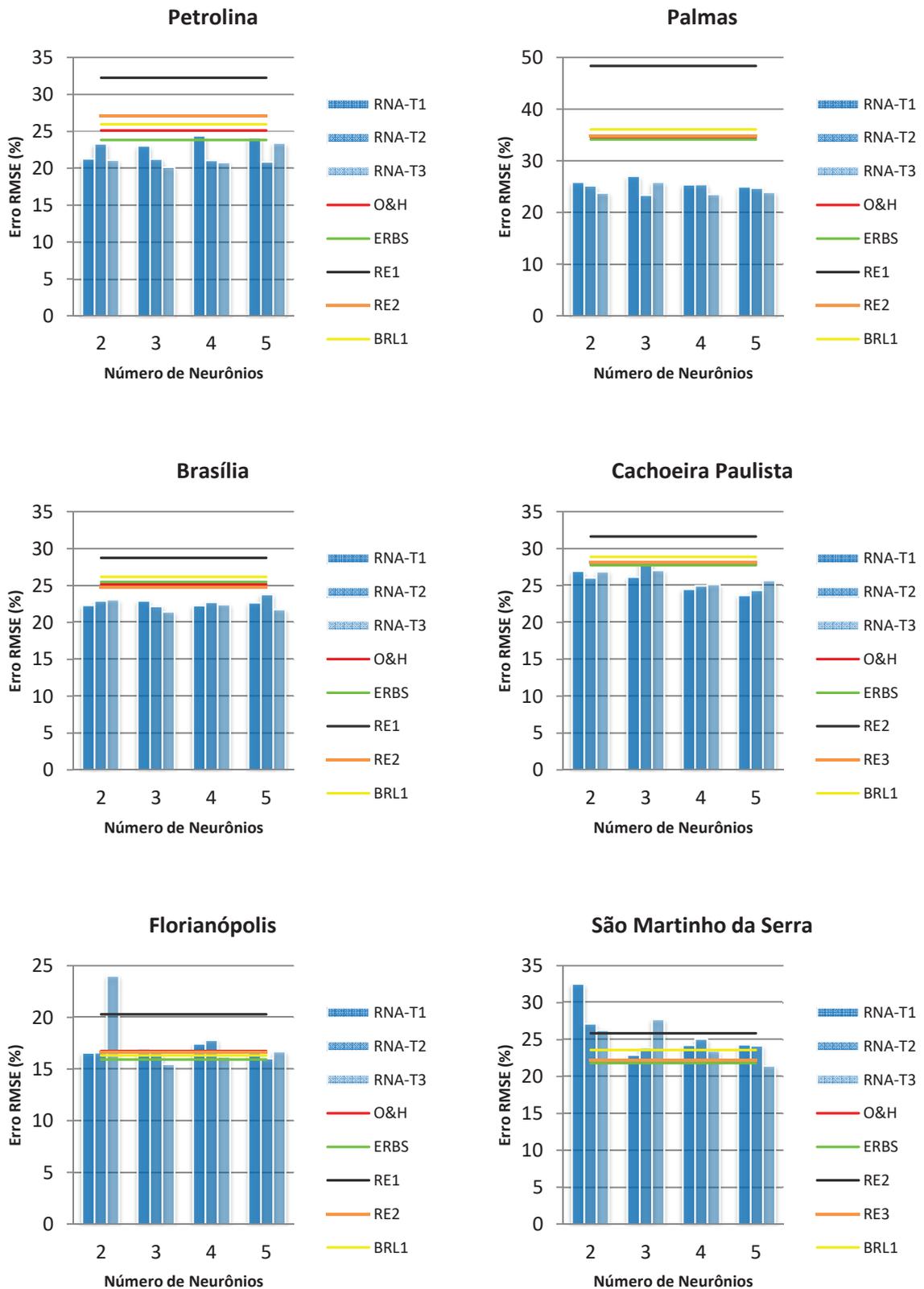


Figura 5.6 – Comparação dos erros  $RMSEr$  obtidos pelas RNAs com função de ativação sigmoide nos testes de precisão local em relação aos erros obtidos por meio dos modelos empíricos.

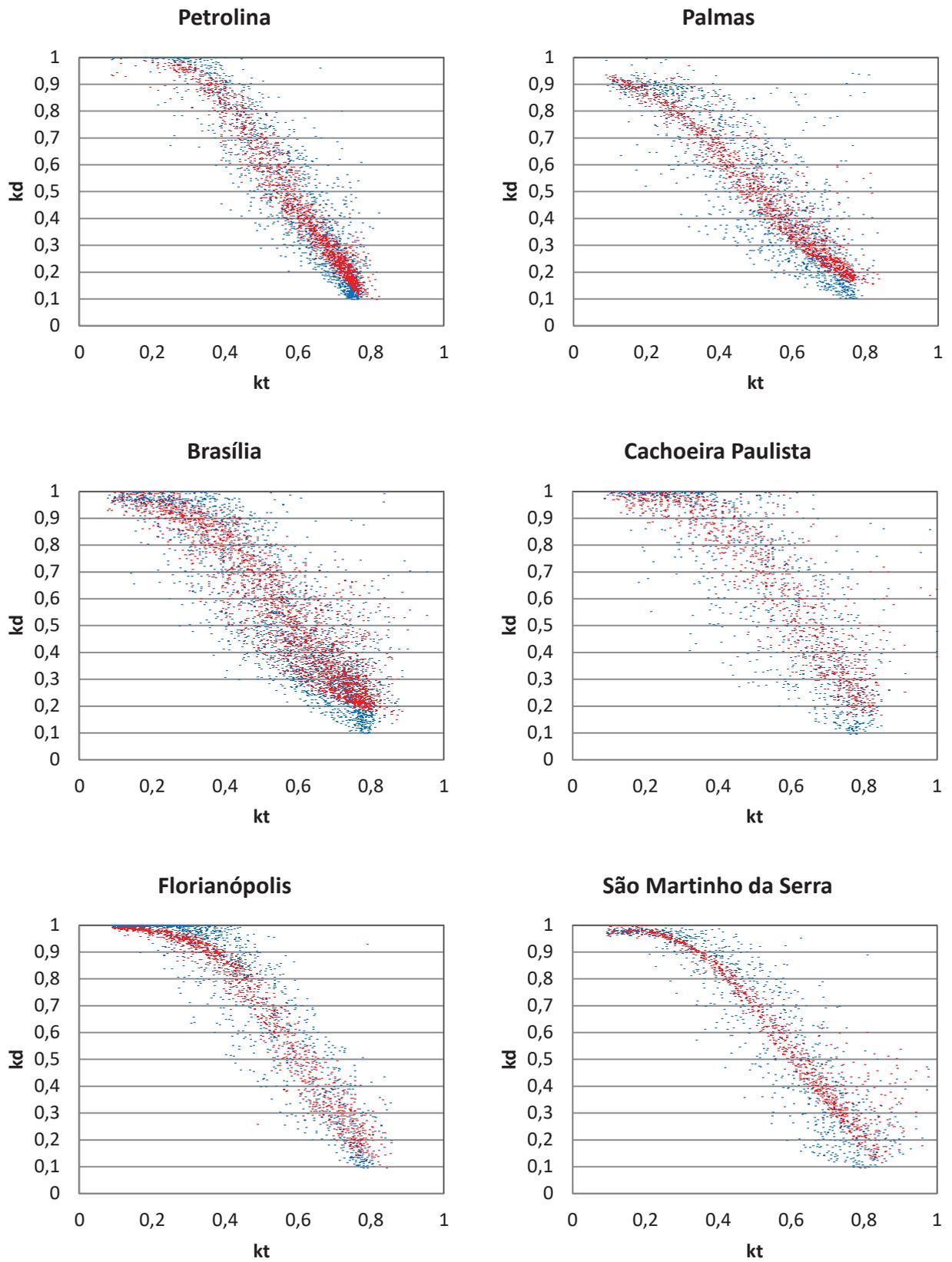


Figura 5.7 – Correlação entre dados medidos [azul] em relação aos dados estimados com menor erro [vermelho] pela RNA de quatro neurônios com função de ativação sigmoide nos testes de precisão local.

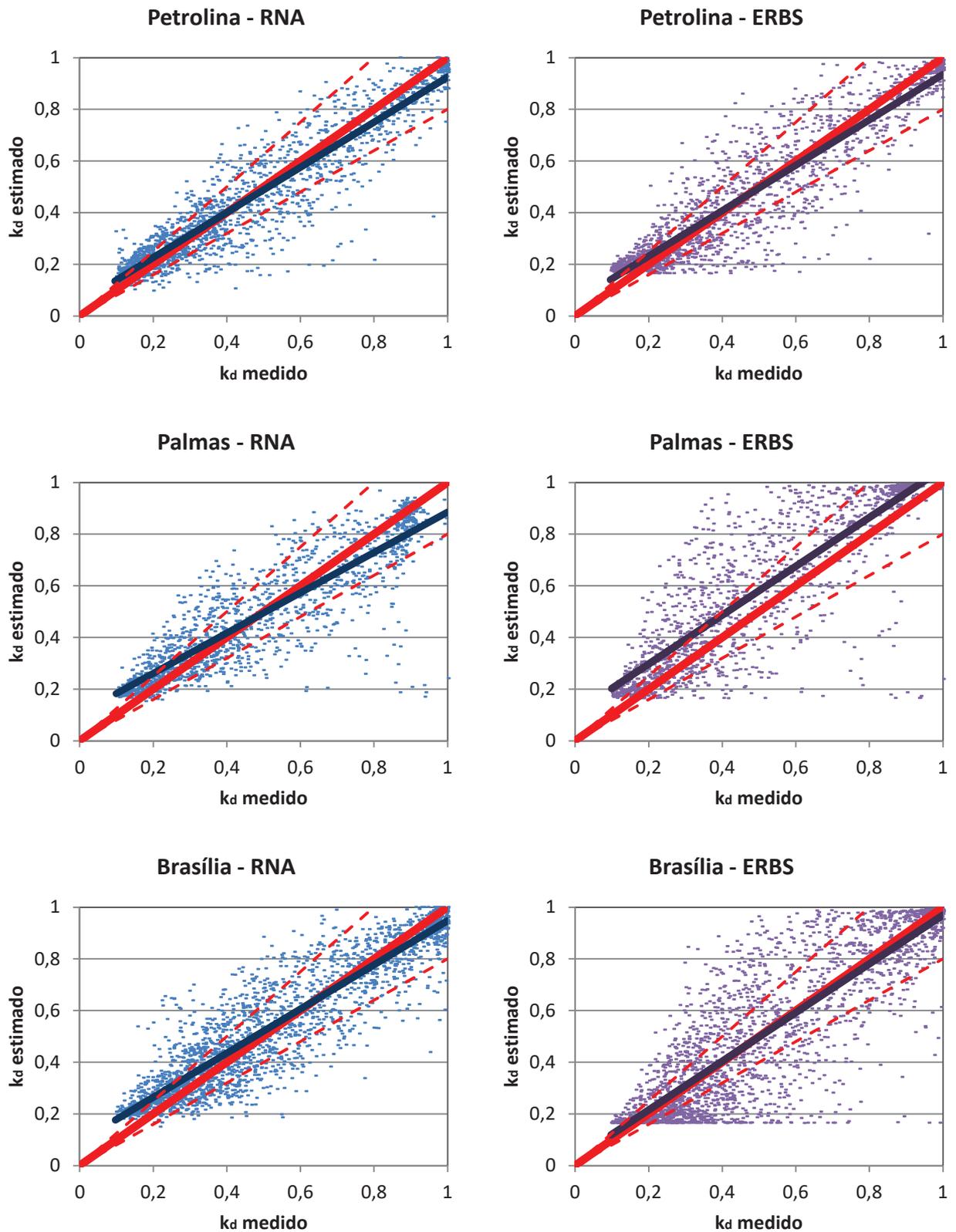


Figura 5.8 – Comparação entre estimativas realizadas pela RNA de quatro neurônios em testes de precisão local com função de ativação sigmoide e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA.

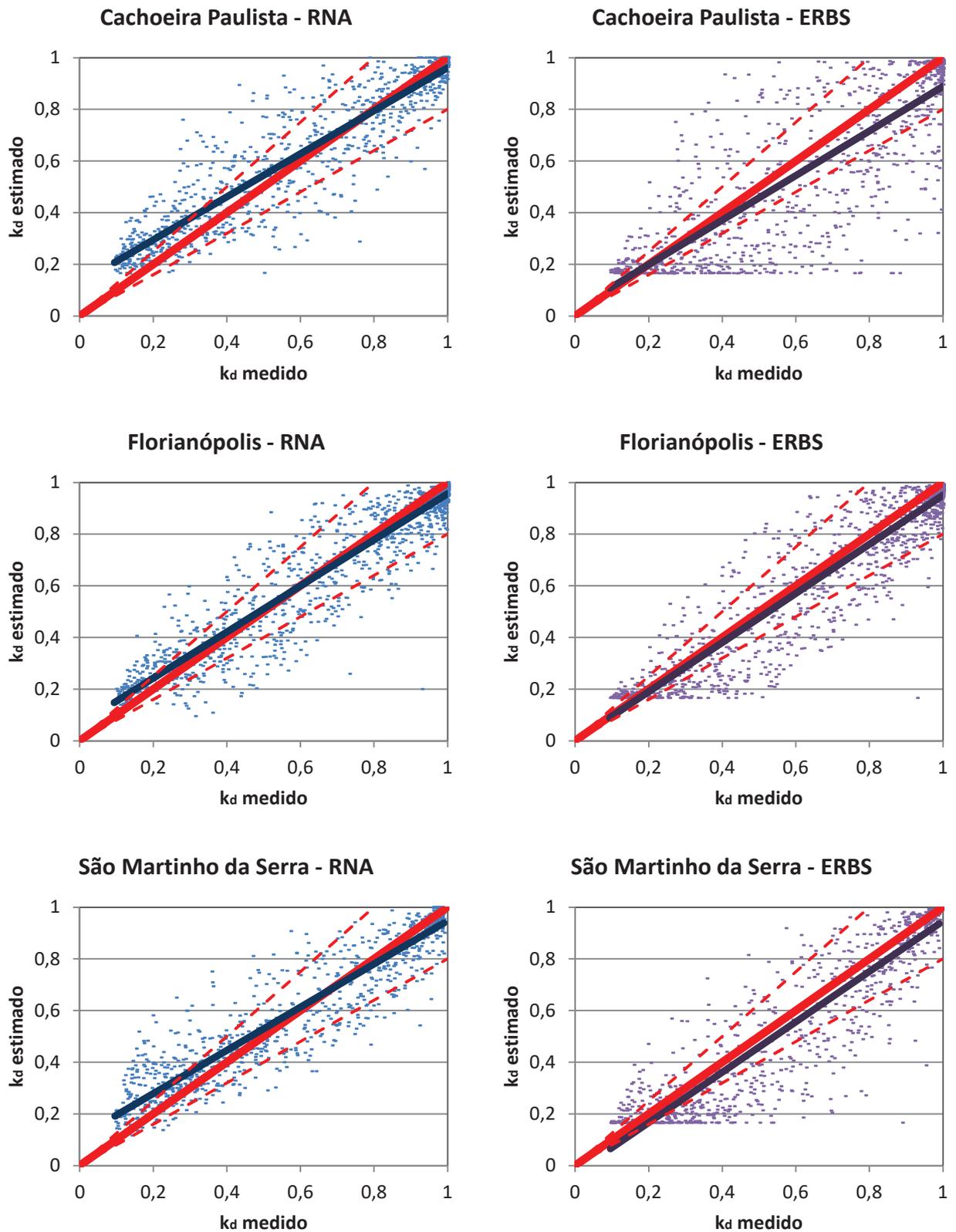


Figura 5.9 – Comparação entre estimativas realizadas pela RNA de quatro neurônios em testes de precisão local com função de ativação sigmoide e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA.

Tabela 5.1 – Coeficientes de correlação ‘ $r$ ’ para as diferentes variáveis na configuração de generalização geográfica com 7 neurônios na camada escondida e função de ativação sigmoide.

CIDADES	$k_t$	$\psi$	$\alpha$	$T_{amb}$	UR
Petrolina	-0.9683	-0.6994	-0.3575	-0.5398	0.6011
	-0.9665	-0.7014	-0.3396	-0.5399	0.6042
	-0.9640	-0.7077	-0.3194	-0.5428	0.6082
Palmas	-0.8917	-0.3179	-0.1423	-0.0117	0.5843
	-0.9019	-0.3341	-0.1107	0.0123	0.5807
	-0.9223	-0.3299	-0.1337	0.0088	0.5218
Brasília	-0.9492	-0.3635	-0.1889	-0.3284	0.5902
	-0.9530	-0.3474	-0.1606	-0.3152	0.5795
	-0.9541	-0.3550	-0.1585	-0.3181	0.5854
C. Paulista	-0.9507	-0.8293	-0.1788	-0.3757	0.4689
	-0.9540	-0.8067	-0.1902	-0.3832	0.4664
	-0.9437	-0.8198	-0.1953	-0.3847	0.4718
Florianópolis	-0.9696	-0.8331	-0.2309	-0.3546	0.6148
	-0.9721	-0.8447	-0.2089	-0.3436	0.6111
	-0.9688	-0.8471	-0.2047	-0.3430	0.6107
S. M. Serra	-0.9111	-0.8061	-0.1702	0.0684	0.5142
	-0.8825	-0.7853	-0.1899	0.1725	0.6034
	-0.8796	-0.7832	-0.1893	0.2482	0.6104

Tabela 5.2 – Coeficientes de correlação ‘ $r$ ’ para as diferentes variáveis na configuração de precisão local com 4 neurônios na camada escondida e função de ativação sigmoide.

CIDADES	$k_t$	$\psi$	$\alpha$	$T_{amb}$	UR
Petrolina	-0.9684	-0.8403	-0.1920	-0.4435	0.5703
	-0.9771	-0.8469	-0.1872	-0.4562	0.5839
	-0.9715	-0.8432	-0.2167	-0.4551	0.5830
Palmas	-0.9753	-0.7618	-0.3083	0.0784	0.4198
	-0.9716	-0.7682	-0.3207	0.0803	0.4173
	-0.9642	-0.8105	-0.3191	0.0854	0.4224
Brasília	-0.9218	-0.3075	-0.2124	-0.3321	0.6439
	-0.8712	-0.3169	-0.2195	-0.5120	0.5762
	-0.8985	-0.3240	-0.2120	-0.4307	0.5917
C. Paulista	-0.8982	-0.7651	-0.2036	-0.5446	0.5718
	-0.8852	-0.7479	-0.2198	-0.5562	0.5827
	-0.8948	-0.7605	-0.1961	-0.5315	0.5643
Florianópolis	-0.9598	-0.8659	-0.0610	-0.4757	0.5967
	-0.9551	-0.8641	-0.0357	-0.4664	0.5841
	-0.9551	-0.8687	-0.0541	-0.4772	0.5920
S. M. Serra	-0.9632	-0.8936	-0.1212	-0.5569	0.7961
	-0.9606	-0.8917	-0.1259	-0.5609	0.8113
	-0.9578	-0.8894	-0.0988	-0.5447	0.7922

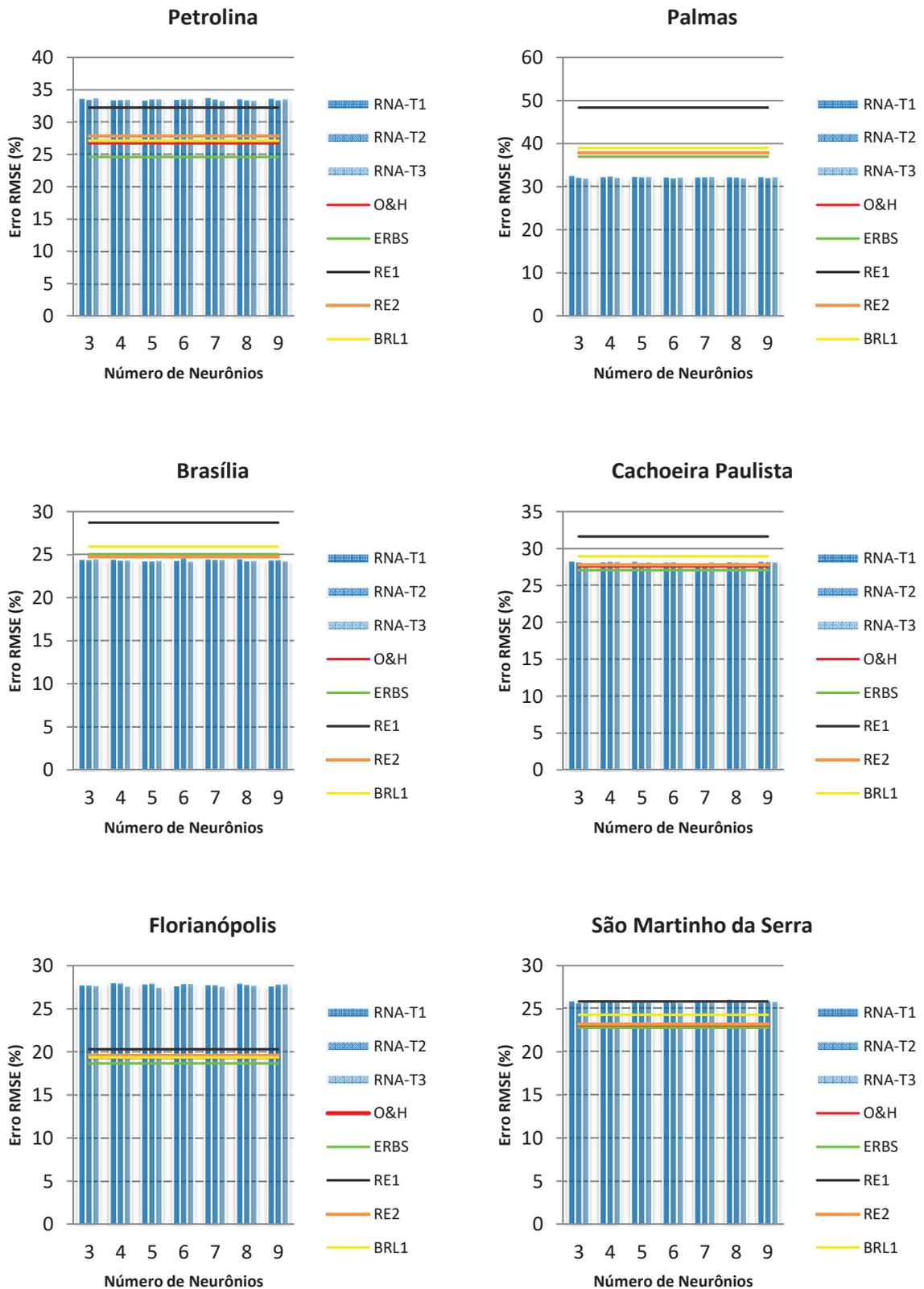


Figura 5.10 – Comparação dos erros  $RMSE_r$  obtidos pelas RNAs com função de ativação linear nos testes de generalização geográfica em relação aos erros obtidos por meio dos modelos empíricos.

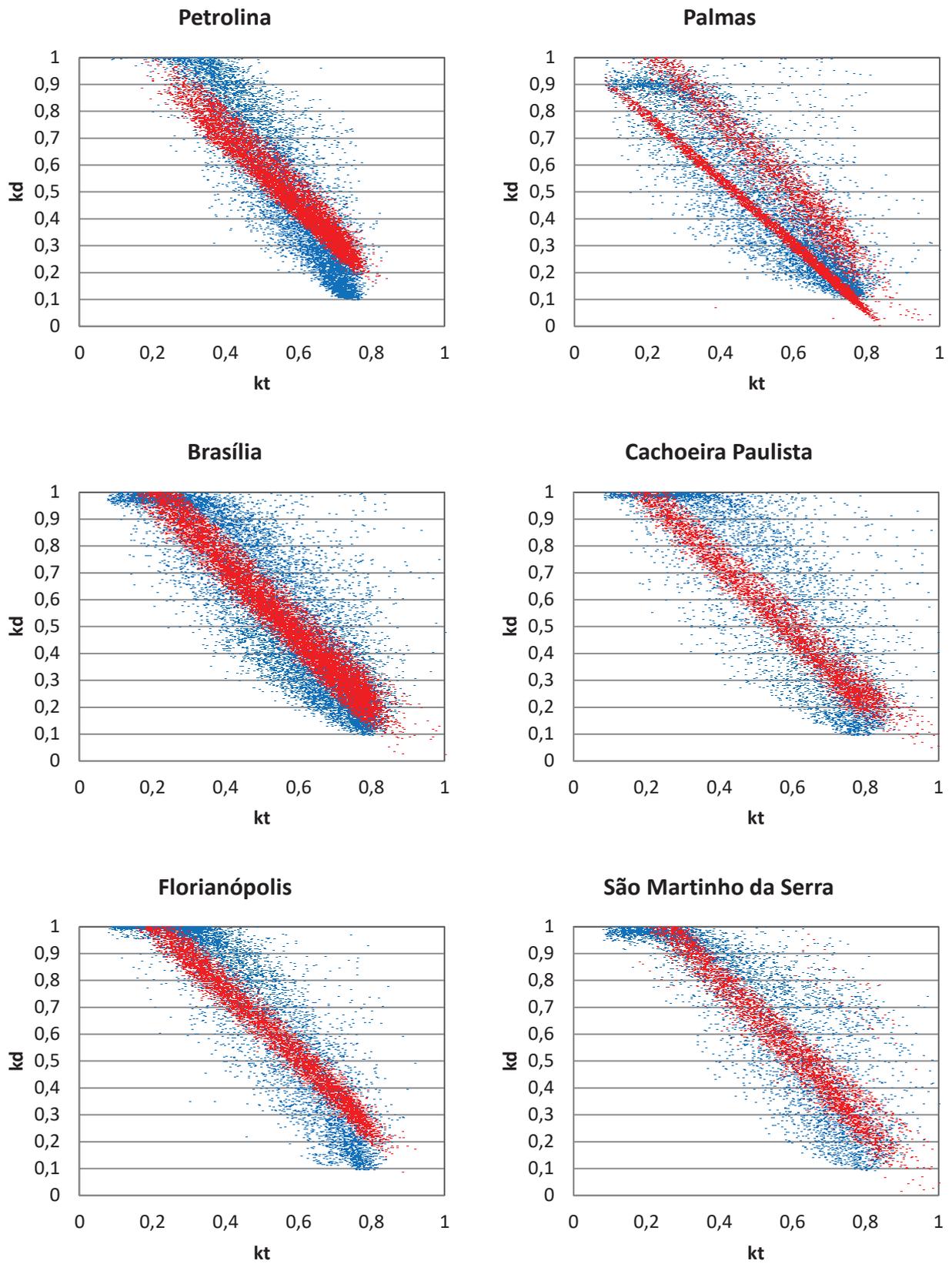


Figura 5.11 – Correlação entre dados medidos [azul] em relação aos dados estimados com menor erro [vermelho] pela RNA de cinco neurônios com função de ativação linear nos testes de generalização geográfica.

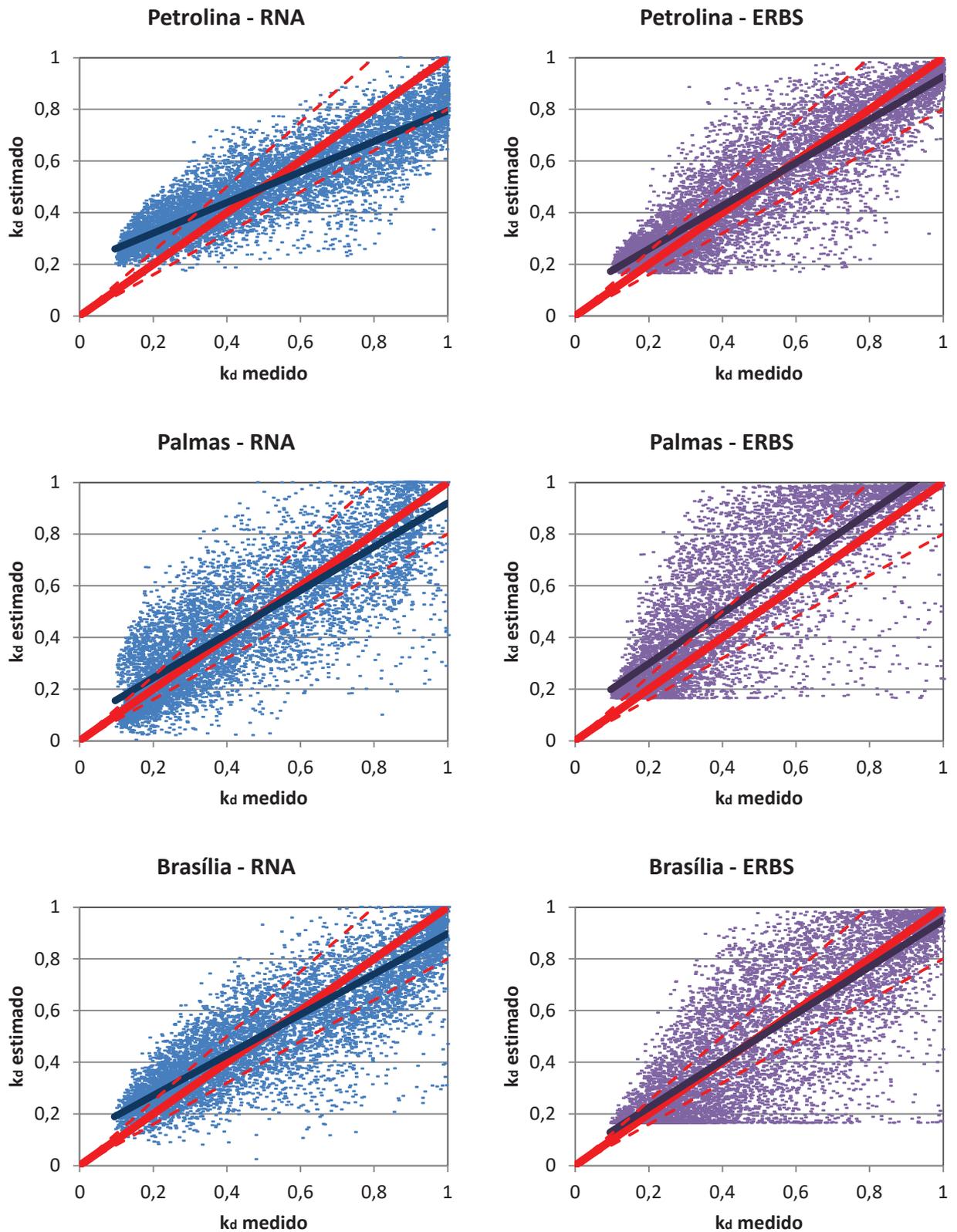


Figura 5.12 – Comparação entre estimativas realizadas pela RNA de cinco neurônios em testes de generalização geográfica com função de ativação linear e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA.

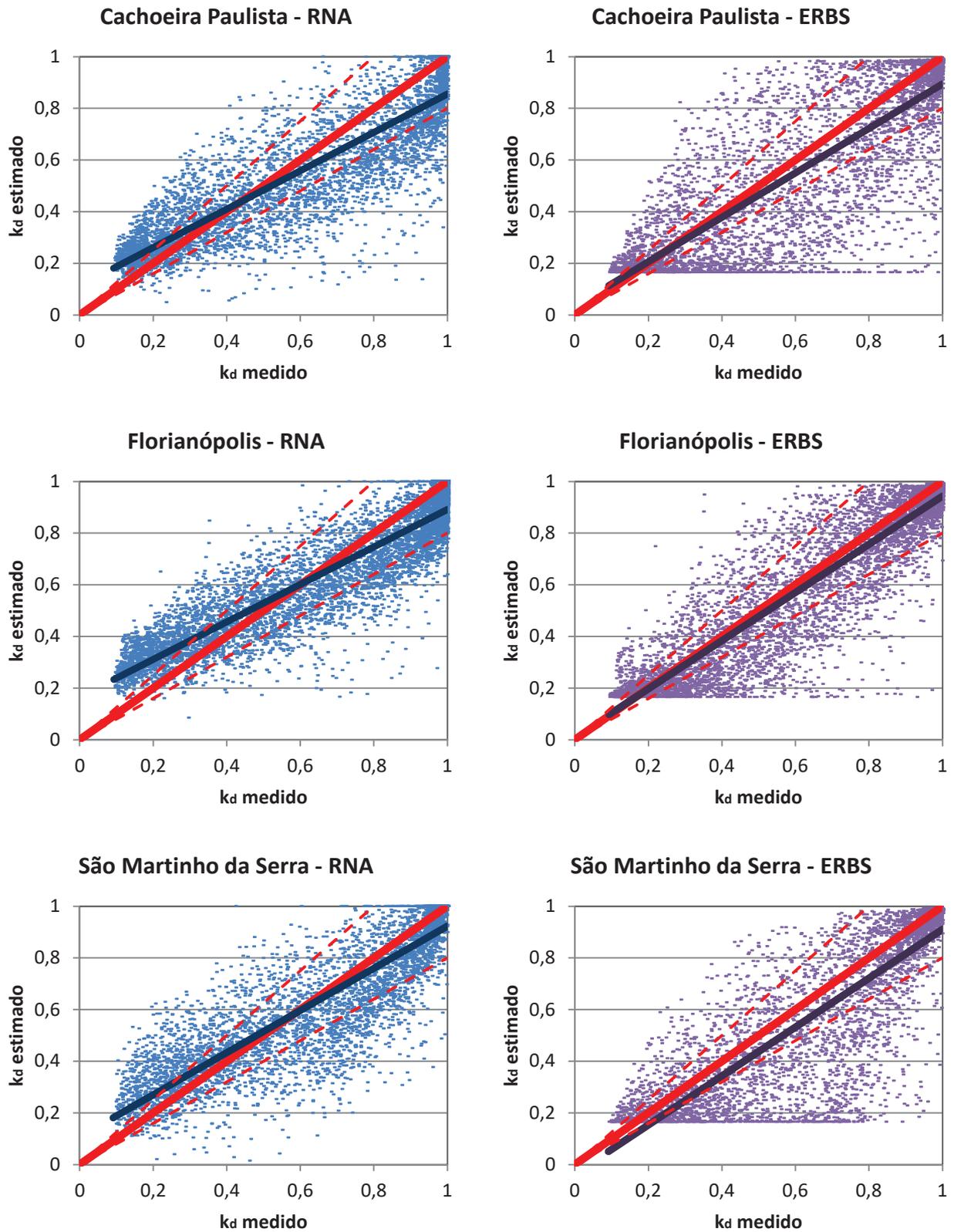


Figura 5.13 – Comparação entre estimativas realizadas pela RNA de cinco neurônios em testes de generalização geográfica com função de ativação linear e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA.

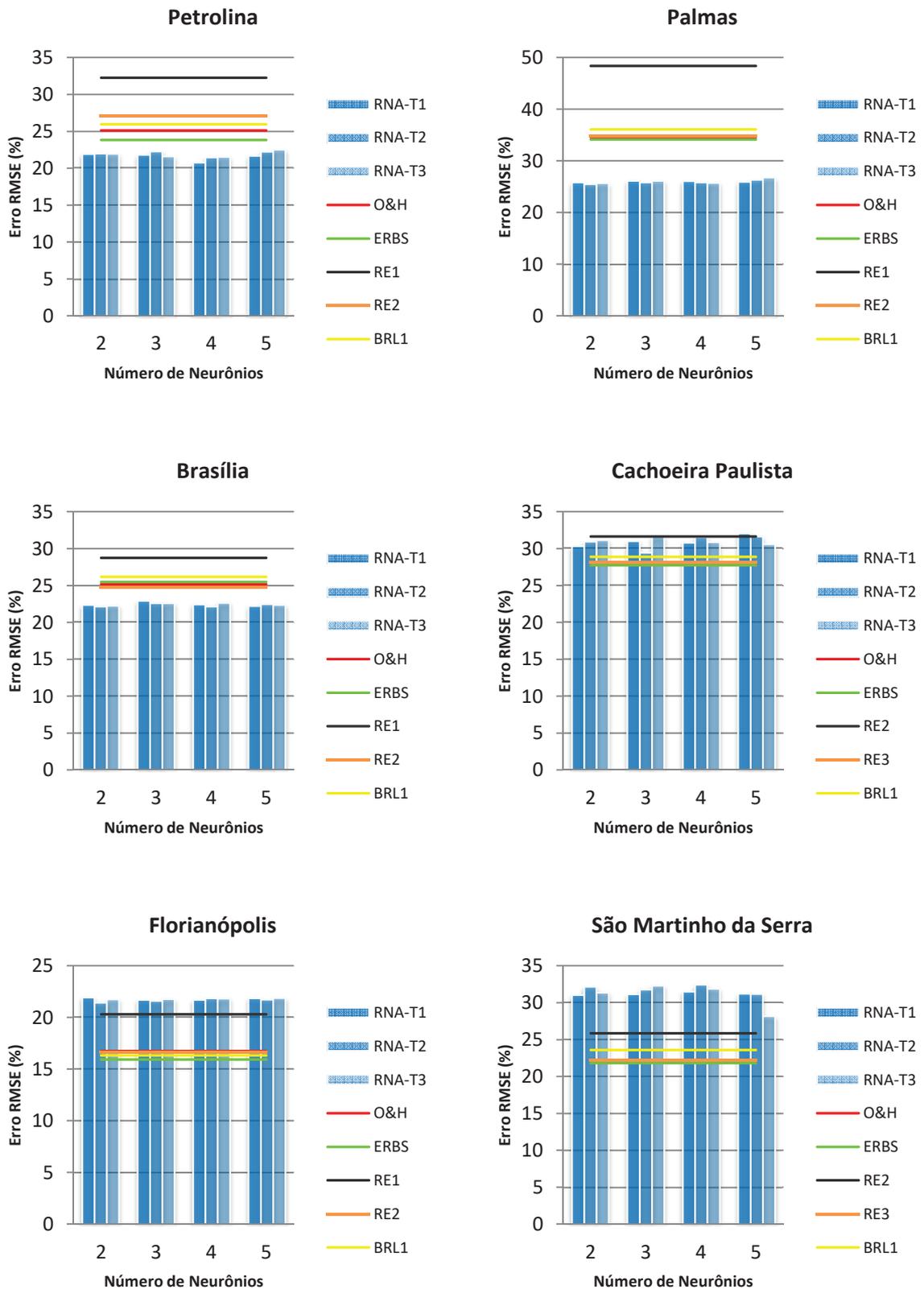


Figura 5.14 – Comparação dos erros  $RMSE_r$  obtidos pelas RNAs com função de ativação linear nos testes de precisão local em relação aos erros obtidos por meio dos modelos empíricos.

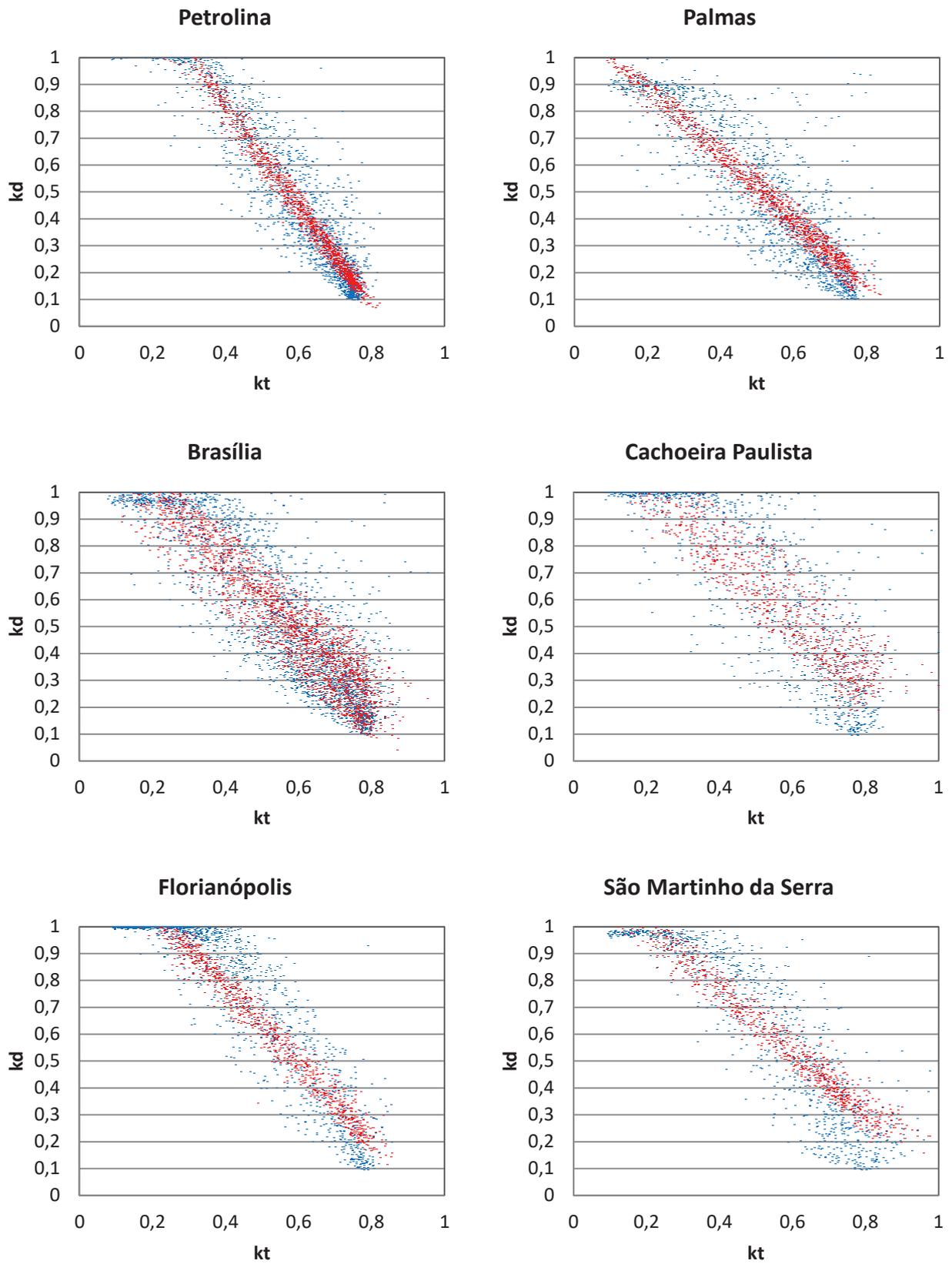


Figura 5.15 – Correlação entre dados medidos [azul] em relação aos dados estimados com menor erro [vermelho] pela RNA de dois neurônios com função de ativação linear nos testes de precisão local.

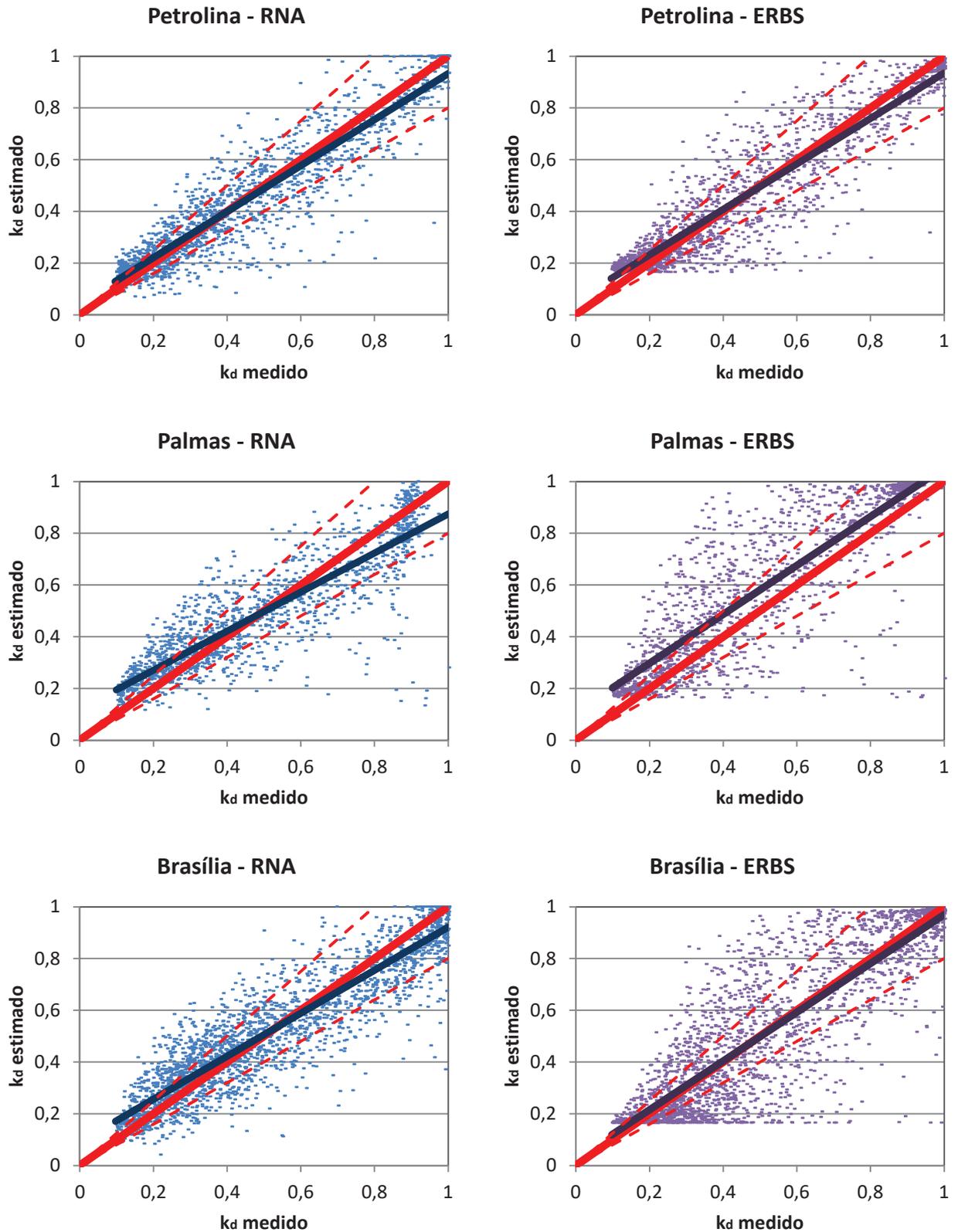


Figura 5.16 – Comparação entre estimativas realizadas pela RNA de dois neurônios em testes de precisão local com função de ativação linear e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA.

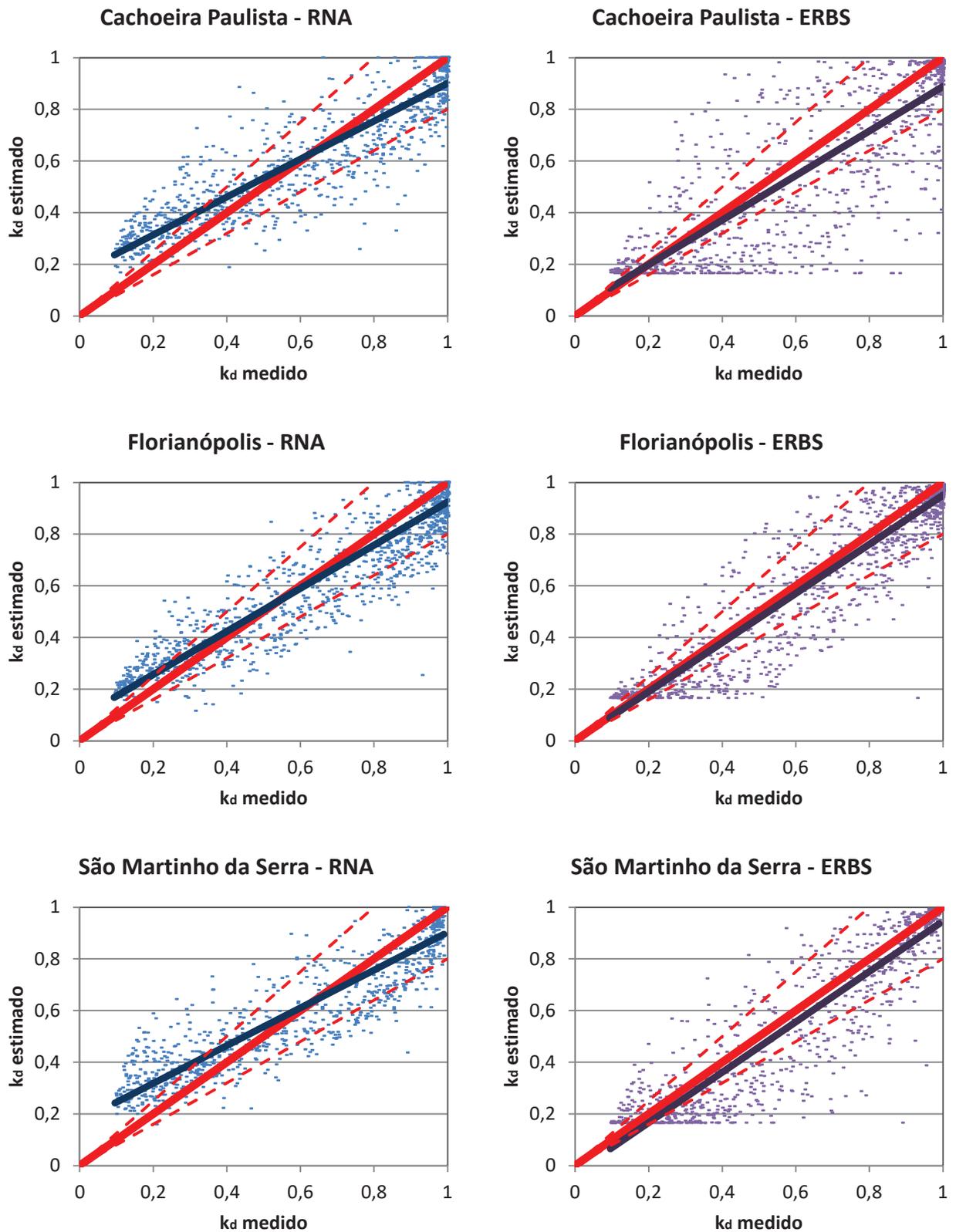


Figura 5.17 – Comparação entre estimativas realizadas pela RNA de dois neurônios em testes de precisão local com função de ativação linear e estimativas realizadas pelo modelo de Erbs et al. (1982) em relação aos coeficientes difusos medidos pelas estações SONDA.

Tabela 5.3 – Coeficientes de correlação ‘ $r$ ’ para as diferentes variáveis na configuração de generalização geográfica com 5 neurônios na camada escondida e função de ativação linear.

CIDADES	$k_t$	$\psi$	$\alpha$	$T_{amb}$	UR
Petrolina	-0.9698	-0.6998	-0.4213	-0.5520	0.6103
	-0.9697	-0.6992	-0.4194	-0.5531	0.6118
	-0.9696	-0.6994	-0.4215	-0.5526	0.6111
Palmas	-0.9181	-0.3600	-0.1679	-0.0432	0.5830
	-0.9190	-0.3692	-0.1622	-0.0370	0.5816
	-0.9196	-0.3629	-0.1626	-0.0408	0.5805
Brasília	-0.9816	-0.3742	-0.1842	-0.3208	0.5750
	-0.9820	-0.3567	-0.1843	-0.3212	0.5755
	-0.9819	-0.3640	-0.1843	-0.3210	0.5752
C. Paulista	-0.9852	-0.8396	-0.1533	-0.3571	0.4313
	-0.9850	-0.8406	-0.1562	-0.3581	0.4317
	-0.9852	-0.8401	-0.1555	-0.3572	0.4305
Florianópolis	-0.9889	-0.8512	-0.2254	-0.3597	0.6371
	-0.9888	-0.8509	-0.2252	-0.3597	0.6381
	-0.9894	-0.8507	-0.2221	-0.3586	0.6361
S. M. Serra	-0.9667	-0.8676	-0.0885	0.0888	0.5504
	-0.9661	-0.8664	-0.0860	0.0946	0.5528
	-0.9637	-0.8639	-0.0875	0.1155	0.5576

Tabela 5.4 – Coeficientes de correlação ‘ $r$ ’ para as diferentes variáveis na configuração de precisão local com 2 neurônios na camada escondida e função de ativação linear.

CIDADES	$k_t$	$\psi$	$\alpha$	$T_{amb}$	UR
Petrolina	-0.9926	-0.8288	-0.2017	-0.4652	0.5939
	-0.9928	-0.8347	-0.2011	-0.4568	0.5858
	-0.9933	-0.8295	-0.1987	-0.4576	0.5850
Palmas	-0.9877	-0.7863	-0.3232	0.0836	0.4283
	-0.9892	-0.7803	-0.3114	0.0842	0.4275
	-0.9878	-0.7831	-0.3159	0.0848	0.4351
Brasília	-0.9425	-0.3331	-0.2082	-0.3013	0.6661
	-0.9450	-0.3274	-0.2037	-0.2990	0.6613
	-0.9409	-0.3318	-0.2140	-0.2917	0.6742
C. Paulista	-0.9438	-0.8072	-0.2561	-0.5709	0.6013
	-0.9407	-0.7977	-0.2613	-0.5801	0.6097
	-0.9489	-0.8121	-0.2427	-0.5596	0.5893
Florianópolis	-0.9883	-0.9044	-0.0284	-0.5109	0.6300
	-0.9902	-0.9028	-0.0451	-0.5032	0.6273
	-0.9904	-0.8993	-0.0327	-0.5047	0.6276
S. M. Serra	-0.9849	-0.9220	-0.1218	-0.5653	0.8002
	-0.9888	-0.9206	-0.0989	-0.5626	0.7974
	-0.9852	-0.9214	-0.1191	-0.5660	0.8013

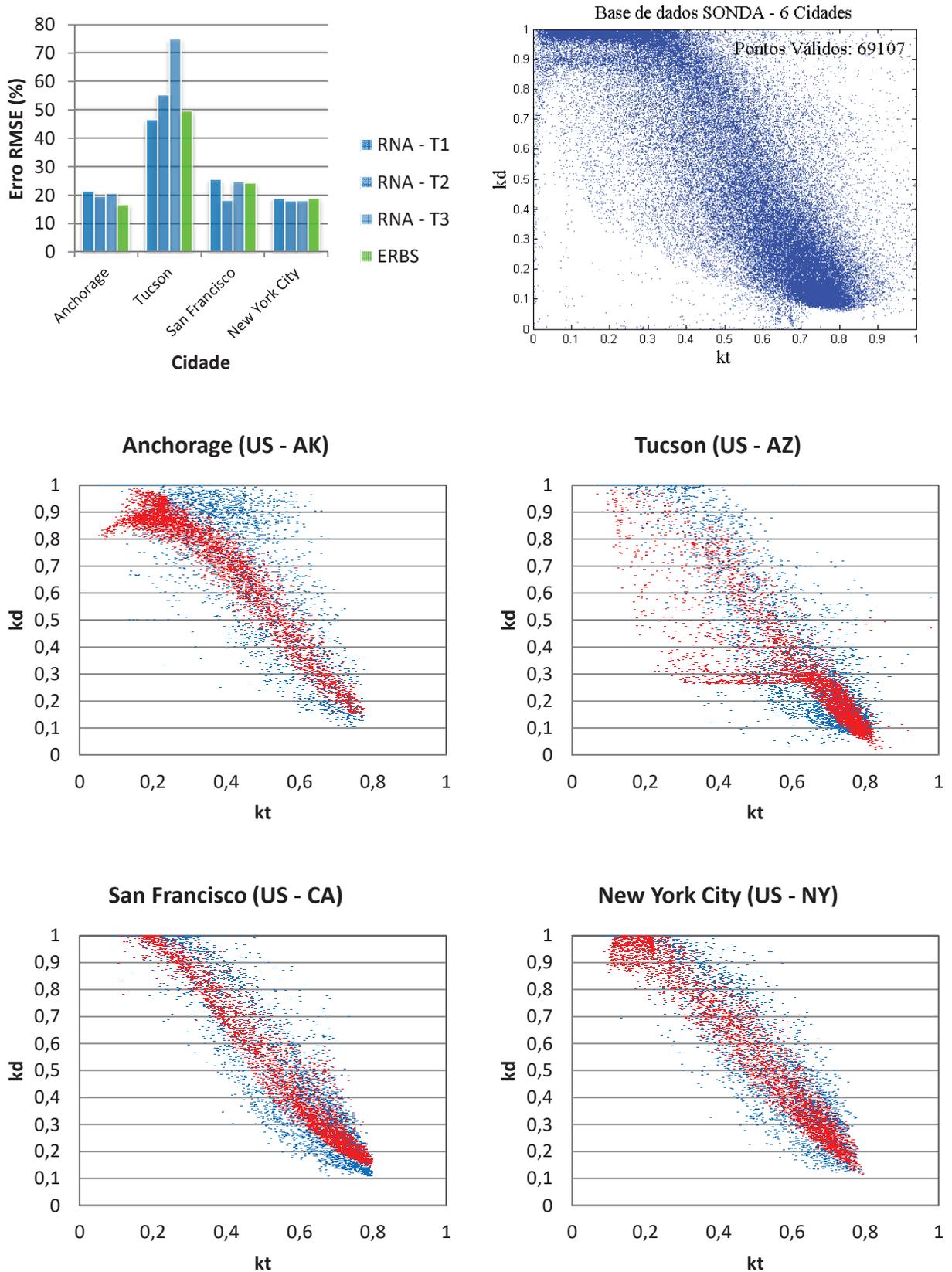


Figura 5.18 – Testes extremos de robustez: Comparação de erros  $RMSE_r$  da RNA em relação à erros do modelo de Erbs et al. (1982), base de dados brasileiros utilizada para treinamento das RNAs e correlações de coeficientes de claridade e coeficientes difusos estimados [vermelho] em comparação com dados TMY3 medidos [azuis].

Os resultados *MBEr* e *RMSEr* para os testes do modelo neural demonstraram pouca variação em relação à quantidade de treinos realizados. Além disso, seu comportamento de generalização geográfica, mesmo em casos onde as características climáticas eram muito diferentes da cidade teste, foram equivalentes ou melhores do que os modelos empíricos. Na precisão local, todas as cidades com mais de cinco mil dados medidos obtiveram melhores resultados em pelo menos um dos testes. Em relação às variáveis utilizadas, o coeficiente de correlação linear demonstrou que a elevação solar em graus, tida como uma variável importante pelo estudo de Ridley et al. (2010), possui pouca ou quase nenhuma correlação com os valores do coeficiente difuso obtidos pelas RNAs treinadas em todos os testes e configurações de camada escondida. A temperatura ambiente (temperatura de bulbo seco) também demonstrou uma correlação variada em relação ao coeficiente difuso, pois em algumas cidades houve correlação moderada e em outras, houve correlação muito fraca. Os resultados dos testes extremos de robustez demonstraram que o presente modelo possui alta generalização e grande precisão/confiabilidade, mesmo quando comparado com o modelo de Erbs et al. (1982).

## 6 ANÁLISE DOS RESULTADOS

O modelo de Erbs et al. (1982) demonstrou ser o mais generalizável dentre todos os modelos empíricos, corroborando assim a afirmação de Ricieri et al. (1995). No entanto, houve uma considerável variação de cidade para cidade em relação às tendências de erro ( $MBEr$  e  $RMSEr$ ) tanto para o modelo com o melhor desvio, quanto para os modelos com o pior e maior desvio.

Uma característica muito interessante dos resultados obtidos com modelos empíricos é o comportamento dos modelos de Reindl et al. (1990). Esta característica de erros mais altos do que modelos simplificados, como o de Erbs e Orgill & Hollands, já foi demonstrado em estudos como o de Ridley et al. (2010). Neste artigo, os erros  $RMSEr$  chegavam a diferir em quase 10% a mais do que outros modelos em localidades como Maputo e Darwin, e em quase 5% na cidade de Lisboa. Macagnan (1993) também demonstrou em seus resultados que os erros do modelo de Reindl et al. (1990), para a cidade de Madri, foram quase 7% mais altos do que os obtidos pelo modelo de Erbs et al. (1982). Levando em conta as características climáticas e latitudes dessas localidades, a relação de erro do modelo RE2 para estimativas realizadas no Brasil e Austrália é bastante relevante, e o alto erro para a cidade de Lisboa é também corroborado pela proximidade com *Madrid* por meio dos resultados de Macagnan (1993). Ao que esses estudos indicam, o referido modelo não possui um bom grau de generalização e, em conjunto com o fato de não ter o segundo “*bin*” representado por um polinômio de quarto grau, a dispersão dos dados estimados - que aparentemente acabaria por beneficiar este modelo em relação a outros totalmente lineares/polinomiais - acaba por aumentar seu erro dependendo da localidade em que é utilizado. Um comportamento bastante curioso, que pode ser observado na Fig. 6.1.

Ridley et al. (2010) utilizou o ângulo de elevação solar em graus, enquanto Reindl et al. (1990) utilizou o seno do ângulo de elevação solar. Assim, ao analisar o coeficiente de correlação linear do ângulo de elevação solar em graus, pode-se admitir que o mesmo não possui correlação e há a necessidade de realização de mais testes, como coeficientes de correlação não linear.

A temperatura ambiente possuiu correlações irregulares com o coeficiente difuso. Nos testes de generalização geográfica, essa variável possuiu uma correlação fraca com as cidades de Brasília, Cachoeira Paulista e Florianópolis além de correlação moderada com a cidade de Petrolina. Nos testes de precisão local, a correlação fraca ocorreu nas cidades de Petrolina e Brasília, e a correlação moderada, nas cidades de Cachoeira Paulista, Florianópolis e São

Martinho da Serra. Esta variável foi convertida em Kelvin para a RNA possuir a maior robustez possível e não divergir com temperaturas extremas. Isso pode reduzir a sensibilidade da rede nos testes de precisão local e minimizar seu coeficiente de correlação em cidades com menor variação de temperatura ao longo do ano. Já em testes de generalização geográfica, pode-se levantar a hipótese de que há mais dados de cidades com temperaturas relativamente constantes ao longo do ano e isso pode influenciar o comportamento da RNA quando dados de regiões subtropicais úmidas são inseridos em suas entradas. Mesmo assim, ainda há a necessidade de um estudo mais aprofundado em relação a essa variável de entrada.

Com os dados brasileiros, o modelo BRL2 de Ridley et al. (2010) demonstrou um comportamento não ideal e totalmente fora do padrão de sua representação, conforme o artigo publicado por seus autores. Em tentativas de alteração desse comportamento, até mesmo a variação manual ou até a exclusão de seus parâmetros, não foram suficientes para que o modelo se comportasse como demonstrado a partir de dados australianos. Com isso, é possível concluir que o modelo RE2 se comportou adequadamente, mas o modelo BRL2 foi bastante errático. Ao que tudo indica, o aumento da complexidade e dependência de variáveis físicas locais em modelos empíricos especializa demais os mesmos e/ou os torna pouco generalizáveis geograficamente. Esse comportamento pode ser observado na Fig. 6.1, onde há quatro sobreposições de resultados obtidos em simulações com dados da cidade de Brasília.

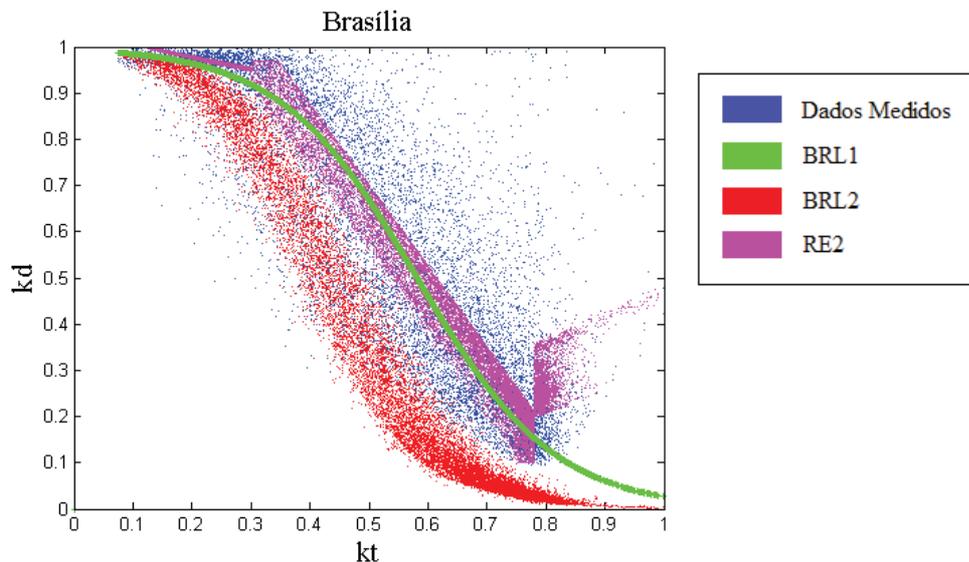


Figura 6.1 - Sobreposição de dados estimados pelos modelos BRL1, BRL2 e RE2, para a cidade de Brasília.

Pela análise dos modelos difusos horários citados, percebe-se que há duas correntes dominantes, as quais podem ser definidas como paradigmas: modelos empíricos e modelos

com aprendizagem de máquina. Esses dois paradigmas de modelagem utilizam como base definições bastante rígidas e podem ser ainda subdivididos em três: modelos empíricos utilizando somente a correlação de Orgill e Hollands (1977) ou Erbs et al. (1982); modelos utilizando variações do estudo de Reindl et al. (1990); e modelos que usam RNAs para previsão e/ou classificação – a definição clássica de utilização de RNAs.

Modelagens analíticas derivadas da análise empírica da influência de parâmetros físicos, no entanto, estão cada vez mais complexas para melhorar sua precisão ou generalização, como no caso do estudo de Starkveit et al. (1998). Com isso, modelos lineares de correlação simples, como o de Erbs et al. (1982), foram considerados por muito tempo como os melhores e mais confiáveis para utilização, visto que eles representam o comportamento médio da irradiância difusa, possuem uma generalização geográfica muito boa em detrimento da precisão local e ainda contam com uma metodologia de adaptação de parâmetros e utilização bastante simples. Os melhores resultados obtidos nos testes realizados com cada modelo/paradigma podem ser vistos na Tab. 6.1. As siglas entre parênteses representam o nome dos modelos empíricos, enquanto os números entre parênteses representam o número de neurônios das redes com função de ativação sigmoide (redes que apresentaram menores erros absolutos dentre todos os testes realizados).

Tabela 6.1 – Menores erros  $RMSE(\%)$  obtidos por meio dos modelos clássicos, modelos neurais generalistas e modelos neurais locais.

CIDADE	Modelos empíricos Gen. Geográfica	RNAs - Gen. Geográfica	Modelos empíricos Prec. Local	RNAs - Prec. Local
Petrolina	24,63 ( <b>ERBS</b> )	24,58 (7)	23,85 ( <b>ERBS</b> )	20,08 (3)
Palmas	36,98 ( <b>ERBS</b> )	29,57 (7)	34,17 ( <b>ERBS</b> )	23,31 (3)
Brasília	24,72 ( <b>RE2</b> )	20,10 (7)	24,73 ( <b>RE2</b> )	21,67 (5)
C. Paulista	27,07 ( <b>ERBS</b> )	22,01 (7)	27,72 ( <b>ERBS</b> )	24,28 (5)
Florianópolis	16,66 ( <b>ERBS</b> )	22,78 (7)	15,94 ( <b>ERBS</b> )	16,00 (5)
S. M. da Serra	22,82 ( <b>ERBS</b> )	21,25 (6)	21,80 ( <b>ERBS</b> )	21,39 (5)

Três cidades apresentaram resultados de precisão local maiores que os de generalização geográfica, mas duas delas possuíam menos dados para treinamento/teste (menos de 5000 dados válidos para treinamento) do que as outras. Uma RNA com variáveis físicas em suas entradas não apenas possui a mesma simplicidade e robustez para obtenção de resultados gerais como também é muito simples para adaptação a localidades específicas com possibilidade de utilização confiável em localidades próximas, sem adaptação de parâmetros

em equações ou variáveis físicas. A dispersão dos dados apresentados pelas RNAs também demonstra um comportamento da irradiância muito mais natural em relação ao comportamento dos resultados obtidos com modelos empíricos.

Nos resultados dos testes de robustez extrema, mesmo quando confrontado com resultados de um modelo geral (ERBS et al., 1982) que possuía cidades próximas das cidades teste, as RNAs com variáveis físicas em suas entradas demonstraram um aumento na precisão em pelo menos um dos três testes e pouca diferença de precisão quando não superou o modelo de confronto (com exceção do terceiro teste realizado na cidade de Tucson e os testes realizados na cidade de Anchorage). Levando em conta que a função de ativação sigmoide é mais complexa e passível de uma maior liberdade na obtenção dos pesos sinápticos da rede neural, as RNAs com variáveis físicas em suas entradas demonstraram um excelente desempenho ao ser utilizado em cidades com extrema diferença de latitude e longitude.

## 7 CONCLUSÃO

Com a velocidade de computação cada vez mais alta, mais núcleos nos processadores de última geração, o fator de aumento de tempo em um projeto é o humano. Portanto, nos tempos atuais, afirmações sobre o aumento da complexidade de modelos para pouco ganho em relação a resultados (LORENZO, 2006) somente são relevantes pelo ponto de vista do tempo humano para desenvolvimento, implementação e reprodutibilidade de modelos, sejam eles de qualquer natureza em se tratando de radiação solar (na qual os modelos difusos horários encontram-se inclusos).

O presente estudo conseguiu demonstrar, portanto, por meio de seus resultados, que modelos difusos horários com RNAs possuem uma metodologia de implementação bastante simplificada, apesar de sua robustez. Modelos empíricos como os de Erbs et al. (1982) e Reindl et al. (1990) necessitam de ferramentas computacionais para obter seus polinômios a partir do método de mínimos quadrados. Já os modelos com RNAs necessitam de um software auxiliar para treinamento das redes.

Além da fácil implementação, os modelos que utilizam RNAs também possuem uma grande robustez em relação à qualidade dos dados utilizados. Com um procedimento de limpeza muito simplificado, a RNA comportou-se de maneira adequada, com grande precisão e uma alta capacidade de generalização em dados contendo regiões de pontos que poderiam ser classificados como *outliers*. As RNAs também demonstram que não há necessidade de ferramentas de limpeza que causariam ainda mais custo de tempo humano e necessidade de especialistas em áreas como matemática, algo recorrente em estudos como os de Ridley et al. (2010) e Boland e Ridley (2008).

Em relação aos resultados apresentados, às características geográficas do Brasil e aos testes específicos de robustez da RNA, essa metodologia também pode ser classificada como um paradigma para modelos difusos gerais e locais. A abordagem deste estudo difere tanto dos modelos matemáticos empíricos com ou sem a utilização de variáveis físicas quanto dos modelos de previsão por meio de aprendizado de máquina, abrindo então uma nova possibilidade para futuros estudos a partir da adaptação deste mesmo modelo a outras regiões, da inclusão de diferentes variáveis físicas ou da alteração de parâmetros/arquitetura da RNA. Modelos difusos contendo RNAs também podem ser utilizados para obtenção das parcelas da radiação do Sol necessárias para projetos de viabilidade de instalação solar com grande confiança no comportamento do “espalhamento” do coeficiente difuso, seja para geração de energia ou para outras aplicações que utilizem esta variável.

É importante ressaltar que os modelos que utilizam mutuamente uma RNA com entradas representadas por variáveis físicas demonstraram uma melhora mútua entre precisão e generalização geográfica, algo não atingido em estudos anteriores. O aumento da precisão local em modelos empíricos aumenta sua incapacidade de generalização geográfica e vice-versa. Portanto, pode se afirmar, que a partir do ponto de vista da engenharia, da estatística e da computação, esta abordagem de utilização de RNA - mesmo em estágio embrionário - demonstra potencial para melhorias no estudo da modelagem da radiação difusa. Esta abordagem pode ser ainda mais efetiva, melhorar os resultados obtidos e aumentar a confiança em relação às estimativas da irradiância difusa horizontal se aplicada, difundida e ampliada em estudos de outras regiões do globo terrestre.

## REFERÊNCIAS

- ANEEL. **Resolução normativa nº 687/2015**. Brasil, 2015.
- BASSETO, E. L.; ESCOBEDO, J. F.; PAI, A. D. **Estimativa da fração transmitida da irradiação difusa com técnicas de aprendizagem de máquinas**. VII Congresso Brasileiro de Energia Solar, Gramado, R. 2018.
- BECQUEREL, A. E. **Memoire sur les effets électriques produits sous l'influence des rayons solaires**. C.R.Acad.Sci., Paris, 9, 561 - 567. 1839.
- BENDT, P.; COLLARES-PEREIRA, M.; RABL, A. **The Frequency Distribution of Daily Radiation Values**. *Solar Energy*, 27, 1. 1981.
- BOLAND, J.; RIDLEY, B. Models of Diffuse Solar Fraction. In: BADESCU, V. **Modeling Radiation at the Earth's Surface**. Romênia: Springer, ano. p. 193-219. 2008.
- CORREA, S. M. B. B. **Probabilidade e estatística**. Puc Minas Virtual, 3a ed, Belo Horizonte, 2003.
- CRESESB. **Energia Solar - Princípios e Aplicações**. 2008. Disponível em: [http://www.cresesb.cepel.br/download/tutorial/tutorial\\_solar\\_2006.pdf](http://www.cresesb.cepel.br/download/tutorial/tutorial_solar_2006.pdf). Acessado em 22/08/2017.
- DAI-PRÁ, L. B.; DIAS, J. B.; KIELING, A. G. **Levantamento de potenciais aspectos e impactos ambientais na fabricação de células de silício para a produção de módulos fotovoltaicos**. 4º Congresso Internacional de Tecnologias para o Meio Ambiente, Bento Gonçalves, RS. 2014.
- DUFFIE, J. A.; BECKMAN, W. A. **Solar Engineering of Thermal Processes**. Ed. John Wiley & Sons. 3a ed. 2006.
- ENERDATA. **Global Energy Statistical Yearbook**. 2017. Disponível em <https://yearbook.enerdata.net/>. Acessado em 22/08/2017.
- ERBS, D. G.; KLEIN S. A.; DUFFIE, J. A. **Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation**. *Solar Energy*, 28, 293-302, 1982.

- GARRISON, J. D. **A study of solar irradiation data for six sites.** *Solar Energy*, p. 237-249. 1984.
- GARRISON, J. D. **A study of the division of global irradiance into direct and diffuse irradiance at thirty-three U.S. sites.** *Solar Energy*, p. 341-351. 1985.
- HAYKIN, S. **Neural Networks: A Comprehensive Foundation**, Second ed. Macmillan College Publishing, New York, 1998.
- HEAT, G. A.; BURKARDT III, J. J. **Meta-Analysis of Estimates of Life Cycle Greenhouse Gas Emissions from Concentrating Solar Power**, SolarPACES, Espanha. 2011.
- HERZOG, M. E. **Estimation of Hourly and Monthly Average Daily Insolation on Tilted Surfaces.** M. S. Thesis, Trinity University. 1985.
- HOLLANDS, K. G. T; HUGET R. G. **A Probability Density Function for the Clearness Index, with Applications.** *Solar Energy*, 30, 195. 1983.
- HULSTRON, R. L. **Solar resources.** The MIT Press, EUA. 1989.
- IBGE. **Mapas Temáticos.** Disponível em: <http://mapas.ibge.gov.br/tematicos/>. Acessado em 22/08/2017.
- IQBAL, M. **An introduction to solar radiation.** Kindle Edition, Academic Press, 386p, 2012.
- KALOGIROU, S. A. **Solar energy engineering: processes and systems.** 1a ed. Academic Press, Elsevier, EUA. 2009.
- LIU, B. Y. H.; JORDAN, R. C. **The interrelationship and characteristic distribution of direct, diffuse and total solar radiation.** *Solar Energy*, [S.l.], v. 4, p. 1-19, 1960.
- LORENZO, E. P. **Electricidad solar fotovoltaica, Volumen II: Radiación solar y dispositivos fotovoltaicos.** PROGENSA, 2006.
- LOVEGROVE, K.; STEIN W. **Concentrating solar power technology - Principles, developments and applications.** Woodhead Publishing Limited. 2012.

MACAGNAN, M. H. **Caracterización de la radiación solar para aplicaciones fotovoltaicas en el caso de Madrid.** Tesis Doctoral, Universidad Politécnica de Madrid, Escuela Superior de Ingenieros de Telecomunicación, Madrid. 1993.

MAXWELL, E. L. **A quasi-physical model for converting hourly global horizontal to direct normal insolation.** SERI/TR-215-3087, 1987.

MELLIT, A.; ELEUCH, H.; BENGHANEM, M.; ELAOUN, C.; PAVAN, A. M. **An adaptive model for predicting of global, direct and diffuse hourly solar irradiance.** Energy Conversion and Management 51, p. 771-782, 2010.

ORGILL, J. F.; HOLLANDS, K. G. T. **Correlation equation for hourly diffuse radiation on a horizontal surface.** Solar Energy, 19, p. 357-359, 1977.

PALMIERE, S. E. **Arquiteturas e Topologias de Redes Neurais Artificiais.** 2018. Disponível em: <https://www.embarcados.com.br/redes-neurais-artificiais/>. Acessado em 01/12/2018.

PAULESCU, E.; BLAGA, R. **Regression models for hourly diffuse solar radiation.** Solar Energy, 125, p. 111-124, 2016.

PEREIRA, E. B.; MARTINS, F. R. GONÇALVES, A. R.; COSTA, R. S.; LIMA, F. J. L.; RÜTHER, R.; ABREU, S. L.; TIEPOLO, G. M.; PEREIRA, S.V.; SOUZA J. G.; **Atlas brasileiro de energia solar.** INPE, São José dos Campos. 2a ed. 2017.

PILLOT, B.. **Planification de l'électrification rurale décentralisée en Afrique subsaharienne à l'aide de sources renouvelables d'énergie: le cas de l'énergie photovoltaïque en République de Djibouti.** Tese de Doutorado, Université De Corse - Pascal Paoli, École Doctorale Environnement Et Société, França, 2014.

PILLOT, B.; MUSELLI, M.; POGGI, P. **Retrieve monthly averages of ambient air temperature daily profile from satellite-derived land surface temperature data for sizing off-grid PV systems in Djibouti's arid climate.** 27th European Photovoltaic Solar Energy Conference and Exhibition, 2012.

REINDL, D.T.; BECKMAN, W. A.; DUFFIE, J. A. **Diffuse fraction correlations.** Solar Energy, 45, 1-7, 1990.

- RICIERI, R. P.; KOEHLER, K.; KOEHLER, A. C.; SOUZA S. N. M. S. **Modelos de estimativa da radiação difusa diária em Cascavel**. Enc. Energ. Meio Rural An. 4, 2002.
- RIDLEY, B.; BOLAND, J.; LAURET, P. **Modelling of diffuse solar fraction with multiple predictors**. *Renewable Energy*, 35, 478–483, 2010.
- SCALCO, P. **Análise óptica e térmica do receptor de um sistema de concentradores fresnel lineares**. Dissertação de Mestrado, Programa de Pós-graduação em Engenharia Mecânica. Unisinos, 2016.
- SHAMSHIRBAND, S.; MOHAMMADI, K.; KHORASANIZADEH, H.; YEE, P. L.; LEE, M.; DALIBOR PETKOVIĆ, D.; ZALNEZHAD, E. **Estimating the diffuse solar radiation using a coupled support vector machine–wavelet transform model**. *Renewable and Sustainable Energy Reviews*, 56, p. 428–435. 2016.
- SIMAS, E. **Introdução às Redes Neurais Artificiais**. 2018. Disponível em: <http://www.dee.eng.ufba.br/home/simas/introdRNA.pdf>. Acessado em 01/12/2018.
- SKARTVEIT, A.; OLSETH, J. A.; TUFT, M. E. **An hourly diffuse fraction model with correction for variability and surface albedo**. *Solar Energy*, 63, p. 173–183, 1998.
- STOFFEL, T.; RENNÉ, D.; MYERS, D.; WILCOX, S.; SENGUPTA, M.; GEORGE, R.; TURCHI, C. **Concentrating solar power - Best practices handbook for the collection and use of solar resource data**. Technical Report, NREL. Setembro, 2010.
- SONDA. **Portal SONDA**. 2018. Disponível em: <http://sonda.ccest.inpe.br/infos/index.html>. Acessado em 30/11/2018.
- TORRESOL ENERGY. **Portal Torresol Energy**. 2018. Disponível em: <http://torresolenergy.com/>. Acessado em 01/12/2018.
- WENHAM, S. R.; GREEN, M. A.; WATT, M. E.; CORKISH, R. **Applied photovoltaics**. Ed. Earthscan. 2a ed. 2007.
- WHILLIER, A. **Solar energy collection and its utilization for house heating**. Ph.D. Dissertation, Dept. of Mechanical Engineering, MIT. 1953.

VAUGHAN, A. **Solar power growth leaps by 50% worldwide thanks to US and China.** The Guardian Solar Power, 2017. Disponível em: <https://www.theguardian.com/environment/2017/mar/07/solar-power-growth-worldwide-us-china-uk-europe>. Acessado em 22/08/2017.

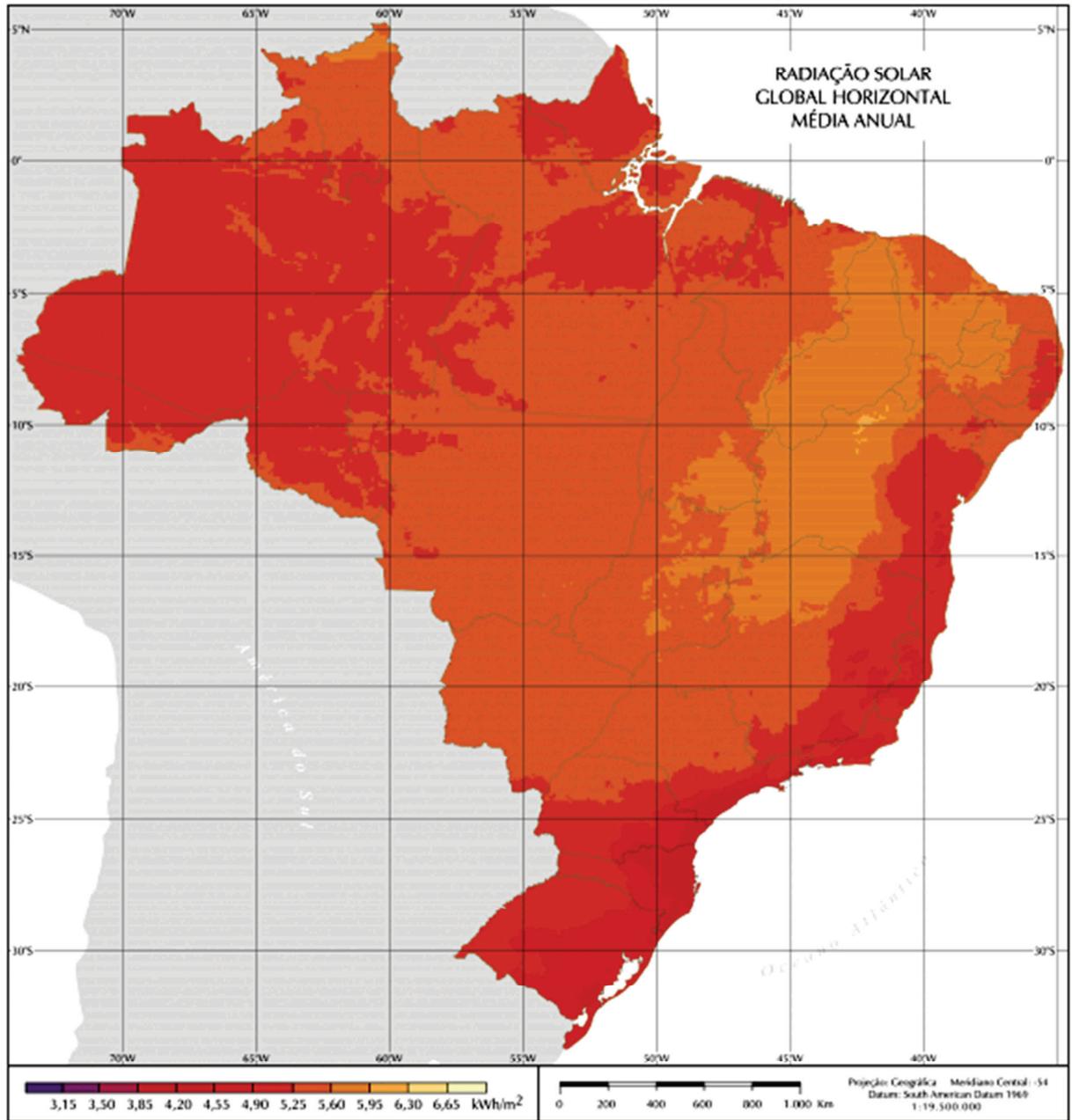
**ANEXO A – RADIAÇÃO SOLAR GLOBAL HORIZONTAL NO BRASIL**

Figura A.1 – Imagem da distribuição da irradiação global horizontal no território brasileiro.

Fonte: Pereira et. al (2017).

## ANEXO B – RADIAÇÃO SOLAR DIFUSA HORIZONTAL NO BRASIL

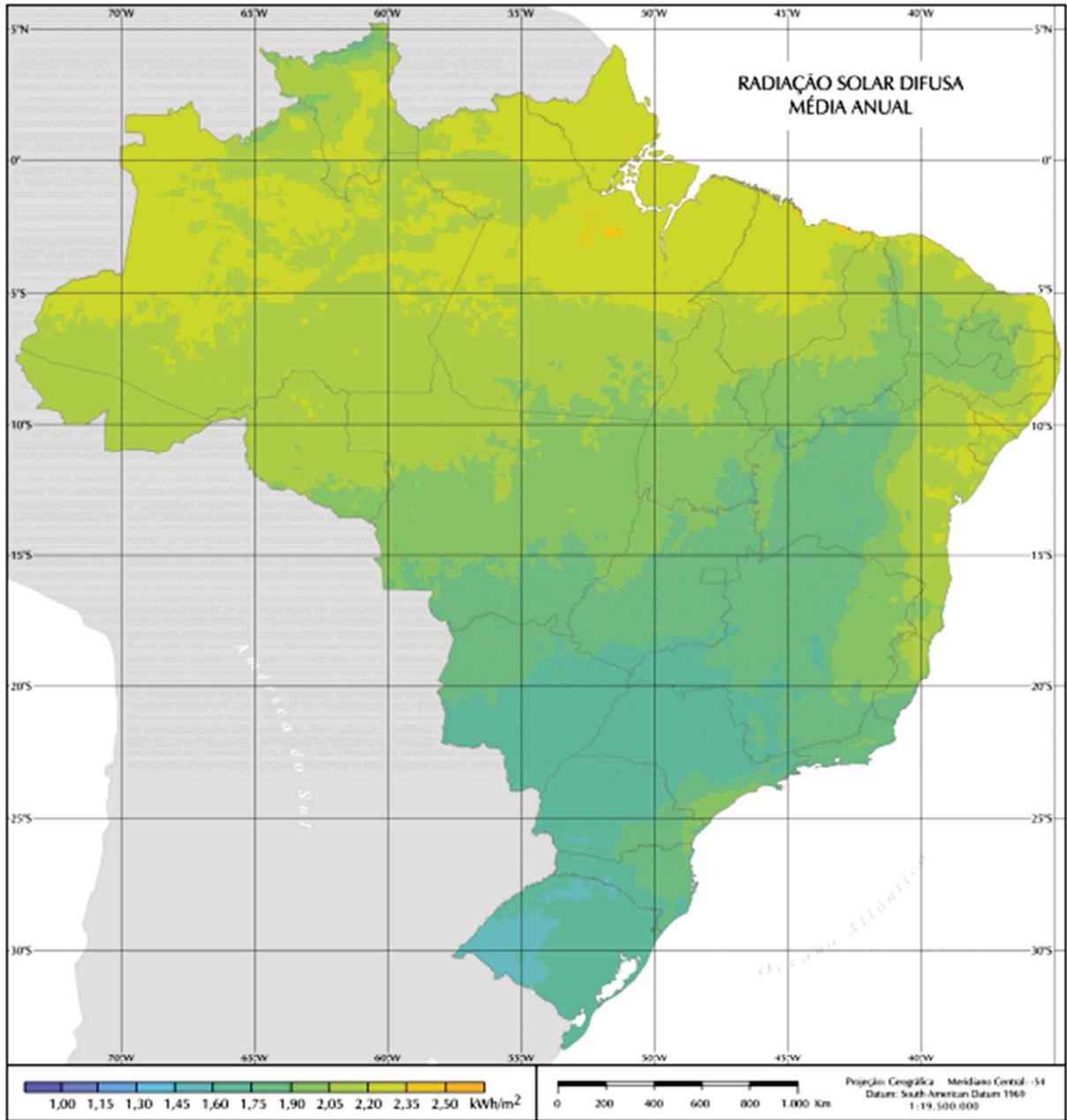
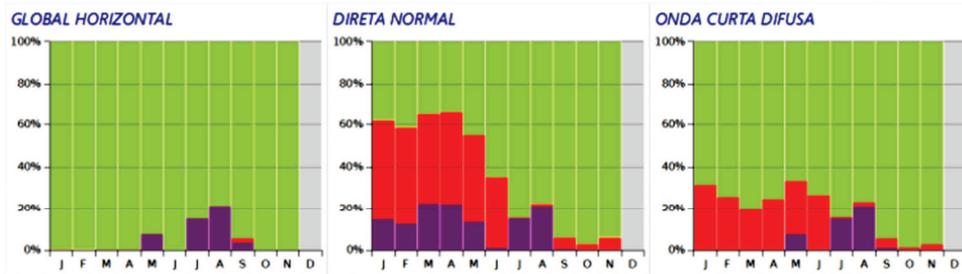
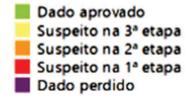


Figura B.1 – Distribuição da irradiação difusa horizontal no território brasileiro.

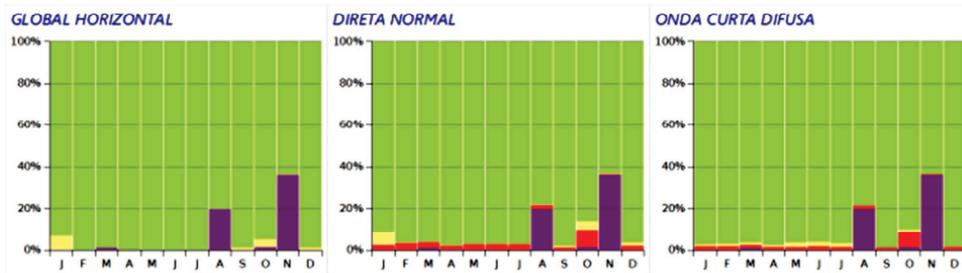
Fonte: Pereira et. al (2017).

ANEXO C – AMOSTRAS DE QUALIDADE DE DADOS (SONDA)

Brasília (BRB 10)  
2015



Petrolina (PTR 11)  
2015



Campo Grande (CGR 12)  
2015



Natal (NAT 17)  
2015



Figura C.1 – Demonstração da qualidade das medições do projeto SONDA.

Fonte: Adaptado de SONDA (2018).

**APÊNDICE A – DEMONSTRAÇÃO DA QUALIDADE DOS DADOS MEDIDOS  
PELAS ESTAÇÕES DO PROJETO SONDA**

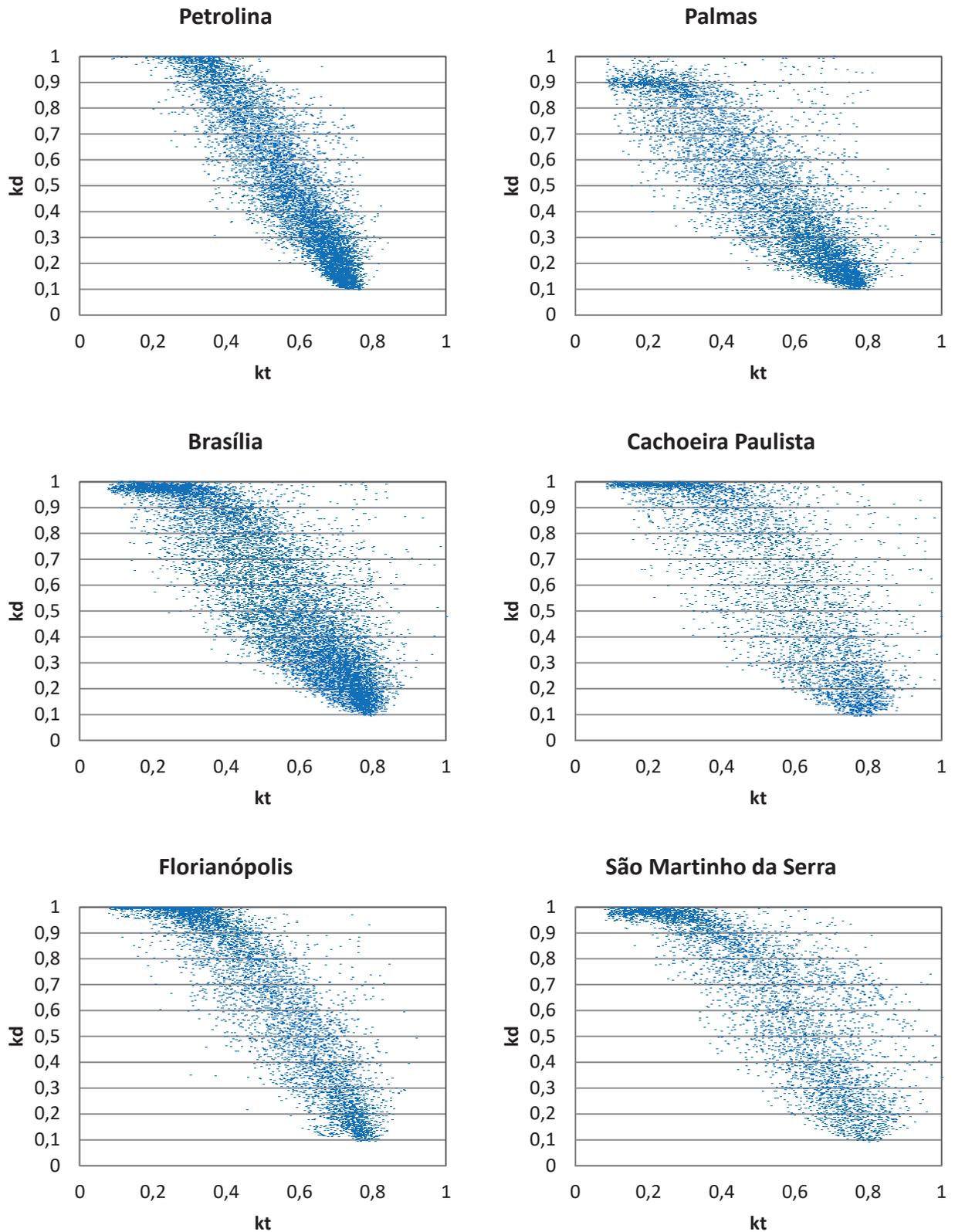


Figura A.1 – Demonstração da qualidade dos dados utilizados neste estudo.

**APÊNDICE B – ERROS RMSE E MBE OBTIDOS A PARTIR DOS MODELOS EMPÍRICOS**

Tabela B.1 – Erro *RMSEr* (%) dos modelos difusos empíricos obtidos a partir da base de testes de generalização geográfica.

CIDADE	O&H	ERBS	RE1*	RE2	BRL1
Petrolina	26,71	24,63	32,25	27,92	27,12
Palmas	37,86	36,98	48,38	37,82	39,00
Brasília	25,01	25,06	28,73	24,72	25,94
Cachoeira Paulista	27,58	27,07	31,64	27,79	28,94
Florianópolis	19,41	18,66	20,31	19,61	19,35
São Martinho da Serra	22,91	22,82	25,85	23,28	24,30

Tabela B.2 – Erro *MBEr* (%) dos modelos difusos empíricos obtidos a partir da base de testes de generalização geográfica.

CIDADE	O&H	ERBS	RE1*	RE2	BRL1
Petrolina	13,68	10,63	22,13	13,97	14,79
Palmas	30,45	28,50	42,92	30,10	30,91
Brasília	3,79	2,50	13,03	2,42	3,08
Cachoeira Paulista	-1,82	-3,09	9,29	-3,84	-4,68
Florianópolis	-1,39	-2,05	4,92	-2,12	-1,57
São Martinho da Serra	-8,53	-9,41	2,69	-10,06	-10,82

Tabela B.3 – Erro *RMSEr* (%) dos modelos difusos empíricos obtidos a partir da base de testes de precisão local.

CIDADE	O&H	ERBS	RE1*	RE2	BRL1
Petrolina	25,13	23,85	32,25	27,09	25,97
Palmas	34,60	34,17	48,38	34,86	36,07
Brasília	25,17	25,49	28,73	24,73	26,17
Cachoeira Paulista	27,95	27,72	31,64	28,13	28,87
Florianópolis	16,74	15,94	20,31	16,63	16,33
São Martinho da Serra	22,10	21,80	25,85	22,22	23,58

Tabela B.4 – Erro *MBEr* (%) dos modelos difusos empíricos obtidos a partir da base de testes de precisão local.

CIDADE	O&H	ERBS	RE1*	RE2	BRL1
Petrolina	11,28	8,58	22,13	12,93	12,64
Palmas	26,69	25,11	42,92	26,76	28,03
Brasília	3,44	1,96	13,03	1,99	2,57
Cachoeira Paulista	-3,72	-4,77	9,29	-5,30	-5,44
Florianópolis	-3,10	-3,10	4,92	-3,85	-3,18
São Martinho da Serra	-3,84	-5,40	2,69	-5,73	-7,66

\* Os testes do modelo RE1 foram realizados com todos os dados disponíveis para cada cidade. Assim, os resultados se repetem tanto para generalização geográfica quanto para precisão local.

**APÊNDICE C – ERROS RMSE E MBE OBTIDOS A PARTIR DA RNA EM TESTES DE GENERALIZAÇÃO GEOGRÁFICA I**

Tabela C.1 – Erro *RMSEr* (%) das RNAs por número de neurônios da camada escondida para testes de generalização geográfica com função de ativação sigmoide.

CIDADE	3	4	5	6	7	8	9
Petrolina	28,26	28,07	28,43	25,91	29,32	26,92	27,66
	29,27	27,05	28,60	28,00	24,51	27,61	27,27
	27,39	27,15	26,81	28,15	24,87	27,60	27,33
Palmas	32,59	30,16	29,88	29,84	30,58	31,21	32,85
	29,95	31,16	32,05	29,77	31,54	36,13	32,60
	31,02	30,35	30,31	32,75	29,57	32,85	29,60
Brasília	22,28	20,35	21,75	22,16	20,92	21,39	19,90
	22,22	22,97	20,49	21,58	20,10	20,86	20,33
	22,51	21,94	22,12	20,92	20,44	21,24	21,57
C. Paulista	23,63	24,75	25,58	24,21	22,42	24,75	24,44
	23,84	23,84	25,77	24,39	24,83	24,61	24,78
	24,21	25,95	25,68	24,38	22,01	26,82	23,96
Florianópolis	29,11	23,14	23,48	24,36	24,45	23,47	23,29
	23,81	23,12	23,30	24,11	23,02	24,08	22,78
	23,84	24,84	23,25	23,06	22,78	22,86	23,70
S. M. Serra	22,67	22,34	22,01	23,15	22,54	22,19	22,28
	23,67	21,07	22,65	21,25	21,95	22,85	22,11
	22,53	23,51	23,79	22,02	23,48	23,39	23,44

Tabela C.2 – Erro *MBEr* (%) das RNAs por número de neurônios da camada escondida para testes de generalização geográfica com função de ativação sigmoide.

CIDADE	3	4	5	6	7	8	9
Petrolina	13,75	13,19	13,57	9,34	15,12	12,15	12,58
	14,14	12,94	12,90	13,54	5,73	12,50	12,64
	11,93	12,89	12,16	13,61	6,16	12,76	11,87
Palmas	14,27	9,15	10,40	9,03	14,43	9,34	10,16
	11,72	10,25	6,75	7,84	11,13	26,59	7,87
	11,24	7,96	8,40	15,14	15,50	21,07	12,30
Brasília	7,53	5,39	7,11	8,09	5,36	7,68	3,79
	8,59	8,51	5,16	7,37	4,24	5,94	4,59
	7,67	7,23	8,01	6,56	5,15	6,03	7,53
C. Paulista	3,63	4,72	5,65	4,47	-3,33	3,56	4,43
	4,05	3,48	5,46	5,38	5,45	4,63	5,23
	4,28	5,47	2,66	5,01	-1,32	9,20	4,11
Florianópolis	15,93	8,97	9,61	9,40	11,78	8,21	9,43
	9,77	9,98	9,37	10,12	9,37	10,73	8,16
	10,12	12,39	9,22	9,34	9,04	9,04	8,93
S. M. Serra	10,43	8,18	7,60	11,01	9,23	8,41	9,85
	6,76	5,90	6,16	4,66	8,79	10,25	8,11
	9,61	9,94	11,90	7,96	11,01	9,21	11,08

**APÊNDICE D – ERROS RMSE E MBE OBTIDOS A PARTIR DA RNA EM TESTES DE PRECISÃO LOCAL I**

Tabela D.1 – Erro *RMSEr* (%) das RNAs por número de neurônios da camada escondida para testes de precisão local com função de ativação sigmoide.

CIDADES	2	3	4	5
Petrolina	21,21	22,97	24,32	24,05
	23,24	21,20	21,03	20,82
	21,05	20,08	20,75	23,36
Palmas	25,77	26,96	25,32	24,89
	25,08	23,31	25,37	24,67
	23,70	25,78	23,46	23,86
Brasília	22,23	22,83	22,21	22,57
	22,82	22,11	22,66	23,71
	22,99	21,37	22,34	21,67
C. Paulista	26,86	26,05	24,44	23,60
	25,96	28,06	24,89	24,28
	26,79	26,97	25,07	25,62
Florianópolis	16,55	16,96	17,43	16,64
	16,57	16,61	17,76	16,00
	23,96	15,46	16,17	16,67
S. M. Serra	32,44	22,87	24,20	24,24
	27,09	23,84	25,02	24,13
	26,22	27,67	23,41	21,39

Tabela D.2 – Erro *MBEr* (%) das RNAs por número de neurônios da camada escondida para testes de precisão local com função de ativação sigmoide.

CIDADES	2	3	4	5
Petrolina	7,70	8,56	13,25	12,06
	10,80	7,63	5,99	7,25
	6,49	5,68	5,90	11,91
Palmas	12,53	12,99	11,41	12,15
	10,30	8,97	11,10	10,98
	10,10	12,88	9,07	10,03
Brasília	9,21	10,47	11,09	10,45
	9,89	8,81	10,85	12,64
	11,13	9,34	10,63	9,65
C. Paulista	16,75	15,63	15,31	11,23
	15,22	17,52	15,46	14,70
	16,73	17,90	16,29	17,31
Florianópolis	3,86	4,44	5,59	4,16
	4,23	3,97	6,82	3,68
	12,26	3,58	3,54	3,80
S. M. Serra	27,57	12,78	11,09	14,80
	16,70	13,20	14,49	13,45
	17,00	17,92	13,16	5,99

**APÊNDICE E – ERROS RMSE E MBE OBTIDOS A PARTIR DA RNA EM TESTES DE GENERALIZAÇÃO GEOGRÁFICA II**

Tabela E.1 – Erro *RMSEr* (%) das RNAs, por número de neurônios da camada escondida, para testes de generalização geográfica com função de ativação linear.

CIDADE	3	4	5	6	7	8	9
Petrolina	33,66	33,42	33,36	33,48	33,77	33,58	33,64
	33,49	33,47	33,57	33,55	33,55	33,40	33,40
	33,72	33,51	33,58	33,57	33,33	33,37	33,57
Palmas	32,54	32,29	32,33	32,22	32,22	32,26	32,24
	32,12	32,44	32,27	32,05	32,25	32,21	32,07
	31,98	32,12	32,31	32,20	32,31	32,03	32,29
Brasília	24,45	24,50	24,29	24,33	24,52	24,55	24,39
	24,42	24,37	24,27	24,65	24,45	24,30	24,42
	24,54	24,39	24,33	24,25	24,45	24,34	24,28
C. Paulista	28,30	28,20	28,25	28,17	28,08	28,20	28,27
	28,19	28,26	28,11	28,20	28,00	28,15	28,24
	28,06	28,25	28,19	28,02	28,20	28,12	28,19
Florianópolis	27,80	28,08	27,91	27,70	27,82	28,02	27,69
	27,80	28,05	28,02	27,98	27,84	27,87	27,90
	27,73	27,69	27,53	27,98	27,64	27,76	27,95
S. M. Serra	25,89	25,91	25,79	25,79	25,98	26,06	25,73
	25,69	25,88	25,79	25,85	25,88	25,77	25,84
	25,84	25,88	25,92	25,69	25,83	25,86	25,82

Tabela E.2 – Erro *MBEr* (%) das RNAs por número de neurônios da camada escondida para testes de generalização geográfica com função de ativação linear.

CIDADE	3	4	5	6	7	8	9
Petrolina	18,43	18,15	18,01	18,18	18,63	18,30	18,44
	18,21	18,21	18,37	18,36	18,26	18,04	18,07
	18,60	18,25	18,28	18,23	17,87	18,03	18,22
Palmas	8,26	8,97	8,52	8,98	9,12	9,16	9,60
	9,32	8,31	9,08	9,21	9,14	8,78	9,26
	9,78	9,18	8,00	9,05	8,79	9,76	8,69
Brasília	9,09	9,29	8,85	8,98	9,41	9,34	8,99
	9,15	8,87	8,74	9,55	9,19	8,75	9,13
	9,30	9,06	8,93	8,77	9,18	8,92	8,91
C. Paulista	5,86	5,38	5,61	5,49	5,25	5,41	5,51
	5,56	5,65	5,30	5,38	5,02	5,44	5,64
	5,28	5,47	5,43	5,18	5,61	5,23	5,39
Florianópolis	13,06	13,40	13,20	12,73	12,85	13,33	12,74
	13,08	13,56	13,35	13,21	13,07	13,05	13,13
	12,80	12,97	12,45	13,19	12,65	12,91	13,17
S. M. Serra	8,82	8,72	8,61	8,55	8,82	9,29	8,03
	7,91	8,83	8,45	8,64	8,65	8,52	8,57
	8,45	8,66	8,85	8,00	8,62	8,72	8,57

**APÊNDICE F – ERROS RMSE E MBE OBTIDOS A PARTIR DA RNA EM TESTES DE PRECISÃO LOCAL II**

Tabela F.1 – Erro *RMSEr* (%) das RNAs por número de neurônios da camada escondida para testes de precisão local com função de ativação linear.

CIDADES	2	3	4	5
Petrolina	21,82	21,72	20,66	21,58
	21,89	22,18	21,38	22,13
	21,85	21,51	21,45	22,42
Palmas	25,78	26,08	26,05	25,90
	25,43	25,80	25,79	26,27
	25,63	26,11	25,69	26,72
Brasília	22,30	22,85	22,39	22,15
	22,08	22,54	22,08	22,40
	22,20	22,56	22,59	22,32
C. Paulista	30,23	30,95	30,74	31,99
	30,87	29,37	31,49	31,57
	31,09	31,57	30,80	30,52
Florianópolis	21,91	21,67	21,67	21,83
	21,40	21,58	21,82	21,70
	21,73	21,76	21,81	21,86
S. M. Serra	31,00	31,10	31,47	31,19
	32,13	31,74	32,40	31,15
	31,33	32,28	31,85	28,16

Tabela F.2 – Erro *MBEr* (%) das RNAs por número de neurônios da camada escondida para testes de precisão local com função de ativação linear.

CIDADES	2	3	4	5
Petrolina	5,47	5,84	3,63	5,11
	5,82	5,91	4,91	7,06
	5,45	5,53	5,11	7,17
Palmas	11,61	13,19	12,14	12,08
	10,87	11,77	11,49	13,07
	11,58	12,49	11,52	13,77
Brasília	8,14	8,98	8,14	7,70
	7,56	8,51	8,27	8,05
	7,95	8,23	8,34	7,97
C. Paulista	17,18	18,54	18,20	20,46
	18,12	15,99	19,72	19,78
	18,35	19,36	18,35	18,04
Florianópolis	7,24	6,73	6,69	7,05
	6,30	6,44	6,62	6,81
	6,62	6,71	6,86	7,17
S. M. Serra	19,26	19,39	19,79	19,57
	20,90	20,34	21,22	19,47
	19,95	21,07	20,49	7,38

**APÊNDICE G – ERROS RMSE E MBE OBTIDOS A PARTIR DE TESTES  
EXTREMOS DE ROBUSTEZ**

Tabela G.1 – Erros RMSE (%) dos testes extremos de robustez do modelo desenvolvido.

MODELO	Anchorage AK	Tucson AZ	San Francisco CA	New York City NY
RNA - T1	21,49	46,57	25,61	18,91
RNA - T2	19,59	55,25	18,26	18,03
RNA - T3	20,65	74,80	24,86	18,11
ERBS	16,72	49,60	24,31	19,03