

UNIVERSIDADE FEDERAL DE MINAS GERAIS - UFMG
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
NÍVEL MESTRADO

ANDRE MOURA GOMES DA COSTA

**ANÁLISE DA DINÂMICA DE LONGO PRAZO DO RECURSO
SOLAR NO TERRITÓRIO BRASILEIRO UTILIZANDO
DECOMPOSIÇÃO DE SÉRIES TEMPORAIS E TÉCNICAS
DE ESTATÍSTICA MULTIVARIADA**

BELO HORIZONTE

2019

DISSERTAÇÃO DE MESTRADO Nº 1110

**ANÁLISE DA DINÂMICA DE LONGO PRAZO DO RECURSO SOLAR NO
TERRITÓRIO BRASILEIRO UTILIZANDO DECOMPOSIÇÃO DE SÉRIES
TEMPORAIS E TÉCNICAS DE ESTATÍSTICA MULTIVARIADA**

André Moura Gomes da Costa

DATA DA DEFESA: 20/02/2019

Universidade Federal de Minas Gerais

Escola de Engenharia

Programa de Pós-Graduação em Engenharia Elétrica

**ANÁLISE DA DINÂMICA DE LONGO PRAZO DO RECURSO
SOLAR NO TERRITÓRIO BRASILEIRO UTILIZANDO
DECOMPOSIÇÃO DE SÉRIES TEMPORAIS E TÉCNICAS DE
ESTATÍSTICA MULTIVARIADA**

André Moura Gomes da Costa

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do Título de Mestre em Engenharia Elétrica.

Orientadora: Profa. Wadaed Uturbey da Costa

Belo Horizonte - MG

Fevereiro de 2019

C837a	<p>Costa, Andre Moura Gomes da. Análise da dinâmica de longo prazo do recurso solar no território brasileiro utilizando decomposição de séries temporais e técnicas de estatística multivariada [recurso eletrônico] / Andre Moura Gomes da Costa. – 2019. 1 recurso online (89 f.: il., color.) : pdf..</p> <p>Orientador: Wadaed Uturbey da Costa.</p> <p>Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.</p> <p>Apêndice: f. 89.</p> <p>Bibliografia: f. 81-87. Exigências do sistema: Adobe Acrobat Reader.</p> <p>1. Engenharia elétrica - Teses. 2. Energia solar - Teses. 3. Planejamento energético - Teses. I. Costa, Wadaed Uturbey da. II. Universidade Federal de Minas Gerais. Escola de Engenharia. III. Título.</p>
	CDU: 621.3(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
Programa de Pós-Graduação em Engenharia Elétrica

**ATA DA 1110ª DEFESA DE DISSERTAÇÃO DE MESTRADO
DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

ATA DE DEFESA DE DISSERTAÇÃO DE MESTRADO do aluno **André Moura Gomes da Costa** - registro de matrícula de número 2017664752. Às 14:30 horas do dia 20 do mês de fevereiro de 2019, reuniu-se na Escola de Engenharia da UFMG a Comissão Examinadora da DISSERTAÇÃO DE MESTRADO para julgar, em exame final, o trabalho intitulado "Análise da Dinâmica de Longo Prazo do Recurso Solar no Território Brasileiro Utilizando Decomposição de Séries Temporais e Técnicas de Estatística Multivariada" da Área de Concentração em Engenharia de Potência, Linha de Pesquisa Compatibilidade Eletromagnética e Qualidade de Energia. A Profa. Wadaed Uturbey da Costa, orientadora do aluno, abriu a sessão apresentando os membros da Comissão e, dando continuidade aos trabalhos, informou aos presentes que, de acordo com o Regulamento do Programa no seu Art. 8.16, será considerado APROVADO na defesa da Dissertação de Mestrado o candidato que obtiver a aprovação unânime dos membros da Comissão Examinadora. Em seguida deu início à apresentação do trabalho pelo Candidato. Ao final da apresentação seguiu-se a arguição do candidato pelos examinadores. Logo após o término da arguição a Comissão Examinadora se reuniu, sem a presença do Candidato e do público, e elegeu o Prof Wallace do Couto Boaventura para presidir a fase de avaliação do trabalho, constituída de deliberação individual de APROVAÇÃO ou de REPROVAÇÃO e expedição do resultado final. As deliberações individuais de cada membro da Comissão Examinadora foram as seguintes:

Membro da Comissão Examinadora	Instituição de Origem	Deliberação	Assinatura
Profa. Dra. Wadaed Uturbey da Costa - Orientador	DEE (UFMG)	Aprovado	
Profa. Dra. Elizabeth Duarte Pereira	UNA (Centro Universitário UNA)	Aprovado	
Prof. Dr. Matheus Pereira Porto	DEMEC (UFMG)	Aprovado	
Prof. Dr. Wallace do Couto Boaventura	DEE (UFMG)	Aprovado	

Tendo como base as deliberações dos membros da Comissão Examinadora a Dissertação de Mestrado foi APROVADA..... O resultado final de Aprovado..... foi comunicado publicamente ao Candidato pelo Presidente da Comissão, ressaltando que a obtenção do Grau de Mestre em ENGENHARIA ELÉTRICA fica condicionada à entrega do TEXTO FINAL da Dissertação de Mestrado. O Candidato terá um prazo máximo de 30 (trinta) dias, a partir desta data, para fazer as CORREÇÕES DE FORMA e entregar o texto final da Dissertação de Mestrado na secretaria do PPGE/UFMG. As correções de forma exigidas pelos membros da Comissão Examinadora deverão ser registradas em um exemplar do texto da Dissertação de Mestrado, cuja verificação ficará sob a responsabilidade do Presidente da Banca Examinadora. Nada mais havendo a tratar o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada pelo Presidente da Comissão Examinadora. Belo Horizonte, 20 de fevereiro de 2019.

ASSINATURA DO PRESIDENTE DA COMISSÃO EXAMINADORA

AGRADECIMENTOS

Deixo meus agradecimentos a todos que estiveram presentes durante a minha trajetória no mestrado.

Agradeço à minha família, por sempre me dar suporte em todo o meu processo de educação, valorizando o crescimento pessoal e compreendendo os momentos mais estressantes desta trajetória.

Agradeço à minha namorada, Daniela, pelo carinho, pelos momentos de trabalho juntos, pelo suporte mútuo em fases difíceis de trabalho e pelo crescimento obtido em conversas e discussões.

Agradeço a todos os meus amigos, em especial, ao Rafael, ao Bráulio, ao Juliano, ao Pedro, ao Lucas, ao Eduardo e ao Arthur, pela antiga amizade.

Agradeço à minha orientadora, Wadaed, pela confiança e apoio na execução deste trabalho.

Agradeço aos meus colegas de laboratório, em especial ao Bruno Marciano, pelos trabalhos desenvolvido juntos, pelo conhecimento repassado, pelas boas discussões e pela amizade. Também agradeço aos colegas Luciana e Hendrigo pela companhia e conversas nos almoços no bandeirão.

Agradeço à UFMG e a todos os seus integrantes, estudantes, professores e servidores, por tornarem possível o funcionamento desta instituição e a formação de profissionais de diversas áreas para o país.

Agradeço à comunidade que se esforça em desenvolver programas gratuitos, que podem trazer diversos benefícios à sociedade. Em especial, à comunidade do R, pelo desenvolvimento desta ferramenta e seus diversos pacotes utilizados neste trabalho.

Agradeço aos usuários do site *stack exchange*, onde pude aprofundar conhecimentos sobre a programação em R e sobre estatística.

Agradeço à comunidade acadêmica em geral pelo esforço de produzir mais conhecimento sobre o funcionamento do mundo para a sociedade.

RESUMO

A utilização de sistemas de geração de energia elétrica que têm como fonte direta a radiação solar está em rápido crescimento no Brasil e no mundo, sobretudo por meio de sistemas fotovoltaicos, principalmente devido à constante redução de custos tecnológicos relacionados. Como esta tecnologia não possui sistema de armazenamento intrínseco, como é opção em hidrelétricas, e sistemas de armazenamento externos ainda têm alto custo, é importante compreender o comportamento da variabilidade relacionada à disponibilidade de sua fonte de energia. Assim, esta dissertação explora aspectos espaçotemporais do comportamento do recurso solar ao longo do território brasileiro, a partir de bancos de dados de satélite, disponibilizados pelo sistema POWER (*Prediction Of Worldwide Energy Resources*), da NASA, e de dados de estações meteorológicas utilizadas no projeto internacional SWERA (*Solar and Wind Energy Resource Assessment*), ambos obtidos gratuitamente nos *sites* dos respectivos projetos. A análise, focada em variabilidades de longo prazo, é feita a partir de uma metodologia proposta pelo autor, baseada na decomposição das séries históricas de irradiância em componentes de tendência, de oscilações interanuais, de sazonalidade e de oscilações mensais. Cada componente da série temporal é avaliada separadamente por meio de técnicas de estatística multivariada, como a análise de componentes principais e a análise de clusters. Ainda, é proposto um novo modelo de índice para avaliação da dinâmica do recurso solar em pares de locais, cujo resultado possui maior valor prático do que o do coeficiente de correlação, usado em vários estudos. É verificado que a metodologia proposta é útil para a compreensão do comportamento do recurso solar, especialmente levando em conta o planejamento energético. Todavia, a análise de algumas componentes, em especial a de oscilações interanuais, mostra a necessidade de um maior contingente de dados para um estudo mais consistente. Entre os resultados encontrados está a sugestão de que o desenvolvimento da utilização da energia solar no estado do Mato Grosso do Sul pode ser interessante, visando diminuir a incerteza mensal da geração solar, principalmente quando esta fonte for mais predominante na matriz energética nacional. Ademais, os bancos de dados utilizados são comparados, e são verificadas consideráveis divergências entre estes, o que põe sob questionamento a qualidade dos dados utilizados.

Palavras-chaves: Variabilidade do Recurso Solar. Complementaridade Energética. Decomposição de Séries Temporais. Análise de Cluster.

ABSTRACT

The use of electrical energy generation systems based on the direct use of solar radiation is showing a rapid growth both on Brazil and worldwide, mainly through photovoltaic systems, due to the continuous reduction of the prices associated with the components fabrication. As this technology does not present storage systems intrinsically associated, and external storage technologies still present high costs, it is important to understand the behaviour of the variability of the source. This dissertation explores spatiotemporal aspects of solar resource behaviour throughout the Brazilian territory, using satellite based data, from the NASA's project POWER (Prediction Of Worldwide Energy Resource), and meteorological stations data, from the international project SWERA (Solar and Wind Energy Resource Assessment), both freely obtained from their's respective websites. The analyses, focused on long-term variability, is performed using a methodology proposed by the author, based on the decomposition of historical irradiance time series into trend, interannual, seasonal, and monthly remainder components. Each component is separately evaluated, using multivariate statistic techniques, such as principal component analysis and cluster analysis. Moreover, a new model of index for pairwise evaluation of the solar resource dynamics is proposed, which presents greater practical meaning than the correlation index, commonly used. It is shown that the proposed methodology is useful for the comprehension of the solar resource dynamics, specially considering its use regarding energy planning. However, the analysis of some components, mainly the interannual one, shows that a larger contingency of data might be needed for obtaining more consistent results. Among the results found, there is the suggestion that the development of the solar energy on the state of Mato Grosso do Sul can be interesting, aiming at lessening global solar energy monthly variability, specially when this source of energy presents a larger participation on the national energy mix. Furthermore, the two datasets used on the study are compared, and considerable dissimilarities are verified between them. This result brings some questions regarding the quality of the satellite based data and its use for understanding of local variability.

Key-words: Solar Resource Variability. Energy complementarity. Time Series Decomposition. Cluster analysis.

LISTA DE FIGURAS

Figura 1 – Crescimento do Total de Unidades de GDFV no Brasil	14
Figura 2 – Crescimento da Potência Instalada Total em Unidades de GDFV no Brasil	15
Figura 3 – Potência Média e Mediana dos Sistemas GDFV no Brasil ao Longo do Tempo	15
Figura 4 – Potência de GDFV Instalada por Subsistema e Estados Mais Destacados	16
Figura 5 – Localização das Usinas Solares Vencedoras de Leilões	17
Figura 6 – Irradiação Global Horizontal Média Verificada para os <i>Pixels</i> Analisados Para os Anos de 1984 a 1987	28
Figura 7 – Irradiação Global Horizontal Média Verificada para os <i>Pixels</i> Analisados Para o Período de 1988 a 2007	29
Figura 8 – Esquemático da Metodologia Proposta	39
Figura 9 – Obtenção das Componentes de Série Temporal para Patos - PB	41
Figura 10 – Obtenção das Componentes de Série Temporal para Belo Horizonte - MG	42
Figura 11 – Componente Sazonal para Um Ano	43
Figura 12 – Variância Explicada pelo Número k de partições	44
Figura 13 – Partições Obtidas	44
Figura 14 – Comparação do Comportamento dos Diversos Índices para a Avaliação de Pares: Componente de Resíduo Mensal	45
Figura 15 – Correlação Versus Distância Considerando a Componente de Resíduo Mensal	46
Figura 16 – Gráfico de Calor das Correlações para a Componente de Resíduo Mensal	47
Figura 17 – Irradiância Diária Média versus Coeficiente de Correlação Considerando a Componente de Resíduo Mensal	48
Figura 18 – Área de Interesse do Gráfico da Figura 17	48
Figura 19 – OPCV versus Distância Considerando a Componente de Resíduo Mensal	49
Figura 20 – Gráfico de Calor dos OPCVs para a Componente de Resíduo Mensal . .	50
Figura 21 – Coeficientes de Variabilidade da Componente de Resíduo Mensal Para o Brasil, com Identificação dos Medoides	51
Figura 22 – Irradiância Diária Média versus OPCV Considerando a Componente de Resíduo Mensal	51
Figura 23 – Áreas de Interesse do Gráfico da Figura 22	52
Figura 24 – Mapas de OPCVs e de Correlações em Análises Baseada em Pontos de Referência	54
Figura 25 – Dinâmica ao Longo dos Anos do Comportamento da correlação e do OPCV para a Componente de Resíduo Mensal considerando o Par de Medoides 32-16	56

Figura 26 – Correlação Versus Distância Considerando a Componente Interanual . . .	57
Figura 27 – Gráfico de Calor das Correlações para a Componente Interanual	57
Figura 28 – Correlação versus Irradiância Média Diária Considerando a Componente Interanual	58
Figura 29 – Área de Interesse do Gráfico da Figura 28	58
Figura 30 – OPCV versus Distância Considerando a Componente Interanual	59
Figura 31 – Gráfico de Calor dos OPCV para a Componente Interanual	60
Figura 32 – Coeficientes de Variabilidade da Componente Internaul Para o Brasil, com Identificação dos Medoides	60
Figura 33 – Irradiância Diária Média versus OPCV Considerando a Componente Interanual	61
Figura 34 – Áreas de Interesse do Gráfico da Figura 33	62
Figura 35 – Mapas de OPCV e de Correlações em Análises Baseada em Pontos de Referência para a componente interanual	64
Figura 36 – Componentes Principais Analisadas Relativas à Componente Sazonal	65
Figura 37 – Componentes Principais Analisadas Relativas à Componente de Tendência	67
Figura 38 – Mapa com a Localização das Estações Utilizadas do SWERA e Comparadas com os Dados do Projeto POWER	69
Figura 39 – Dados Diários do GEWEX versus Dados Diários do SWERA.	69
Figura 40 – Componentes de Tendência nas Localidades das Estações de Medição utilizados no SWERA, Para os Dois Bancos de Dados Usados	71
Figura 41 – Componentes Sazonal nas Localidades das Estações de Medição utilizados no SWERA, Para os Dois Bancos de Dados Usados	72
Figura 42 – Dados mensais do GEWEX versus Dados Mensais do SWERA	73
Figura 43 – Comparação das Correlações Obtidas em Cada Banco de Dados, para o período de 1983 a 2002 - Componente de Resíduo Mensal	74
Figura 44 – Comparação dos OPCV Obtidas em Cada Banco de Dados, para o período de 1983 a 2002 - Componente de Resíduo Mensal	75
Figura 45 – Comparação de Correlações e OPCV Obtidos em Cada Banco de Dados, para o período de 1983 a 2002 - Componente de Resíduo Mensal	75
Figura 46 – Comparação de Correlações e OPCV Obtidos em Cada Banco de Dados, para o período de 1983 a 2002. Componente Interanual.	76
Figura 47 – Partições Obtidas com a Metodologia Apresentada	89

LISTA DE ABREVIATURAS E SIGLAS

ANEEL	Agência Nacional de Energia Elétrica
CV	Coeficiente de Variação
CCEE	Câmara Comercializadora de Energia Elétrica
DNI	<i>Direct Normal Irradiance</i> Irradiância Direta Normal
EPE	Empresa de Pesquisa Energética
FV	Fotovoltaico
GDFV	Geração Distribuída Fotovoltaica
GHI	<i>Global Horizontal Irradiance</i> Irradiância Global no plano Horizontal
loess	<i>locally estimated scatterplot smoothing</i> (Suavização local estimada do gráfico de dispersão)
ONS	Operador Nacional do Sistema
OPCV	<i>Optimal Pair Coefficient of Variation</i> (Coeficiente de Variação Ótima do Par)
PCV	<i>Pair Coefficient of Variation</i> (Coeficiente de Variação do Par)
REN	Resolução Normativa
SIN	Sistema Interligado Nacional

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Objetivos	19
1.1.1	Objetivos específicos	19
1.2	Estrutura do Trabalho	19
2	VARIABILIDADE DO RECURSO SOLAR	21
2.1	Uso de Técnicas de Estatística Multivariada na Caracterização da Variabilidade do Recurso Solar	23
2.2	Decomposição de Séries temporais	24
2.3	Estudos da Variabilidade do Recurso Solar no Brasil	25
3	DADOS UTILIZADOS E METODOLOGIA PROPOSTA	27
3.1	Dados Utilizados	27
3.2	Decomposição das Séries Temporais	30
3.3	Particionamento em Torno de Medoides	32
3.3.1	Definição do Número de Partições	33
3.3.1.1	<i>Análise Baseada em Pontos de referência</i>	35
3.4	Estatísticas dos pares	35
3.5	Análise de Componentes Principais	37
3.6	Resumo	38
4	RESULTADOS E DISCUSSÕES	40
4.1	Decomposição das Séries Temporais	40
4.2	Análise das Componentes de Oscilações Aleatórias	43
4.2.1	Coeficientes Utilizados para Avaliação de Pares	43
4.2.2	Análise da Componente de Resíduo Mensal	46
4.2.2.1	<i>Análise Baseada nas Correlações</i>	46
4.2.2.2	<i>Análise Baseada no OPCV</i>	49
4.2.2.3	<i>Análise Baseada em Pontos de Referência</i>	53
4.2.3	Comportamento ao Longo do Tempo	55
4.2.4	Análise da Componente Interanual	55
4.2.4.1	<i>Análise Baseada em Pontos de Referência</i>	61
4.3	Análise das Componentes Sazonal e de Tendência	63
4.3.1	Componente Sazonal	63
4.3.2	Componente de Tendência	66
4.4	Comparações entre Banco de Dados	68

4.5	Discussões Finais	74
5	CONCLUSÕES	77
5.1	Trabalhos Futuros	79
5.2	Produtos do Trabalho de Mestrado	80
	REFERÊNCIAS BIBLIOGRÁFICAS	81
	APÊNDICES	88
	APÊNDICE A – MAPA DOS AGRUPAMENTOS	89

1 INTRODUÇÃO

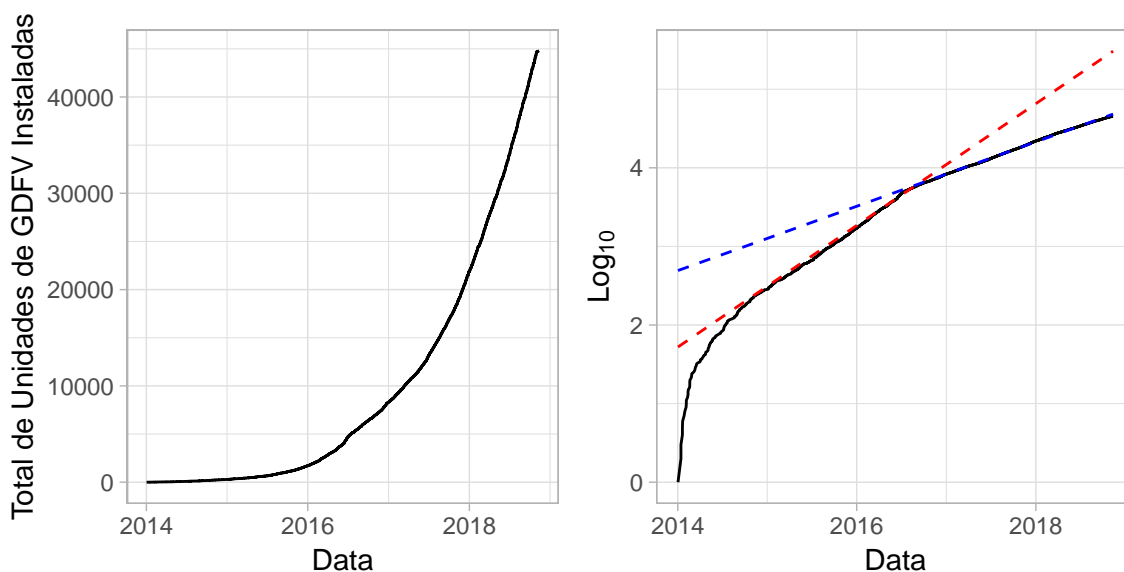
A utilização de sistemas de conversão de energia fotovoltaicos (FV) está em acelerado crescimento em todo o mundo. A potência instalada mundial acumulada de tais sistemas, que era de 15 GWp em 2008, chegou à faixa de 100 GWp em 2012 e à de 300 GWp em 2016. Ainda, é previsto que esse valor possa chegar ao patamar de 900 GWp em 2021 (SCHMELA et al., 2017).

Esta tecnologia, que antes era utilizada principalmente em situações específicas, como em sistemas isolados e em satélites e espaçonaves, começou a ser mais difundida a partir da década de 2000, por se tratar de uma forma de geração de energia elétrica relativamente limpa, que não emite poluentes durante a geração e que, por isso, começou a receber diversos incentivos (JONES; BOUAMANE, 2012).

Antes, seus custos associados ainda eram muito elevados, e países como Alemanha e Japão adotaram políticas de incentivo à utilização deste tipo de geração, que envolviam diversos tipos de subsídios. Devido a questões de economia de escala e de curva de aprendizado, a tecnologia foi se desenvolvendo e aumentando seu desempenho e sua eficiência, enquanto seus custos relacionados foram diminuindo (JONES; BOUAMANE, 2012). Assim, a geração de energia elétrica por meio de sistemas fotovoltaicos foi se tornando cada vez mais comum e apresenta, hoje, custos bastantes competitivos com outras formas de geração. Isto, somado às questões ambientais, justifica o iminente crescimento da utilização da geração fotovoltaica. A redução dos custos de geração pode ser verificada em documentos como o *US solar photovoltaic system cost benchmark* (FU et al., 2017) e na constante redução nos valores de lances em leilões de energia gerada por tal tecnologia, batendo recordes de preço anualmente (SCHMELA et al., 2017). É destacável que, em outubro de 2017, um leilão feito na Arábia Saudita para venda de energia solar, registrou lances de até 17,80US\$/MWh (BELLINI, 2017).

No Brasil, a utilização de sistemas de conversão de energia solar em elétrica começou mais tardiamente. A geração distribuída (GD), modalidade de geração que representa uma parcela significativa da utilização de sistemas fotovoltaicos, foi regularizada apenas em 2012, com a ANEEL REN 482 (ANEEL, 2012). A partir de 2014, devido ao ganho de competitividade com a redução de custos, o uso de sistemas de geração distribuída fotovoltaica (GDFV) começou a ganhar popularidade e o número de instalações de sistemas instalados passou a apresentar um rápido crescimento, como mostram os dados do registro de micro e minigeradores da ANEEL (ANEEL, 2018b), apresentados na figura 1. Desde essa data, pode-se verificar um crescimento que segue padrões exponenciais. Regressões logarítmicas por partes mostram um crescimento de unidades instaladas equivalente a 5,77 vezes ao ano entre início de 2015 e meados de 2016 e, para o período de meados de 2016

Figura 1 – Crescimento do Total de Unidades de GDFV no Brasil



Regressões logarítmicas para os períodos de janeiro de 2015 até junho de 2016 (vermelho) e de julho de 2016 até novembro de 2018 (azul) mostradas em pontilhado.

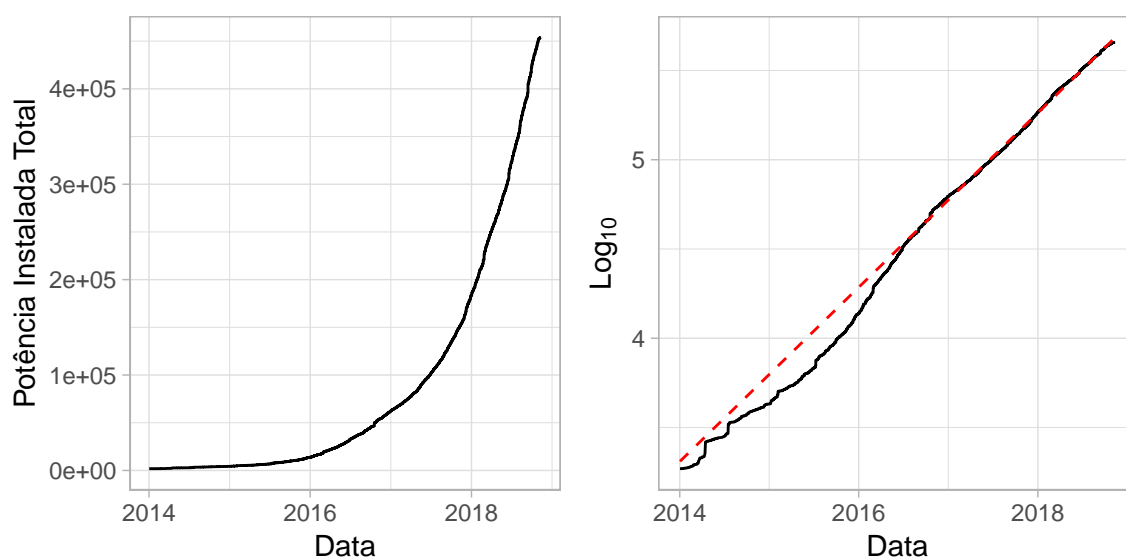
Elaboração própria

Dados: ANEEL (2018b)

até novembro de 2018, um crescimento de 2,55 vezes ao ano. É interessante notar que em 2016 entrou em vigor a ANEEL REN 687 (ANEEL, 2015), que modifica a resolução normativa 247 e questões relacionadas à regulamentação de sistemas de GD.

Em relação à potência instalada, também pode ser verificado um crescimento exponencial, seguindo padrões um pouco diferentes, como pode ser visto na figura 2. Uma regressão logarítmica para o período de 2016 até final de 2018 mostra um crescimento de 3,08 vezes ao ano. Para os períodos anteriores, um crescimento médio similar também pode ser observado, mas um pouco menos regular. Diferenças nas dinâmicas das taxas de crescimento da quantidade de sistemas instalados e da potência total instalada podem ser justificadas verificando a dinâmica do dimensionamento dos sistemas. A figura 3 mostra que o dimensionamento dos sistemas FV médios começou a crescer depois de 2016, fazendo com que a taxa de crescimento da potência instalada para este período pudesse ser maior do que a taxa relativa ao número de sistemas instalados. É importante notar que uma das alterações da resolução normativa 687 foi a possibilidade de instalações de sistemas em novas modalidades, que fazem mais viável a utilização de sistemas maiores. A figura 4 mostra a distribuição atual dos empreendimentos de geração distribuída por subsistema, destacando os estados com maiores potências instaladas. É interessante observar que as regiões de maior desenvolvimento econômico são as que apresentam maiores potências instaladas, sem necessariamente serem regiões de maiores níveis relativos de irradiação média.

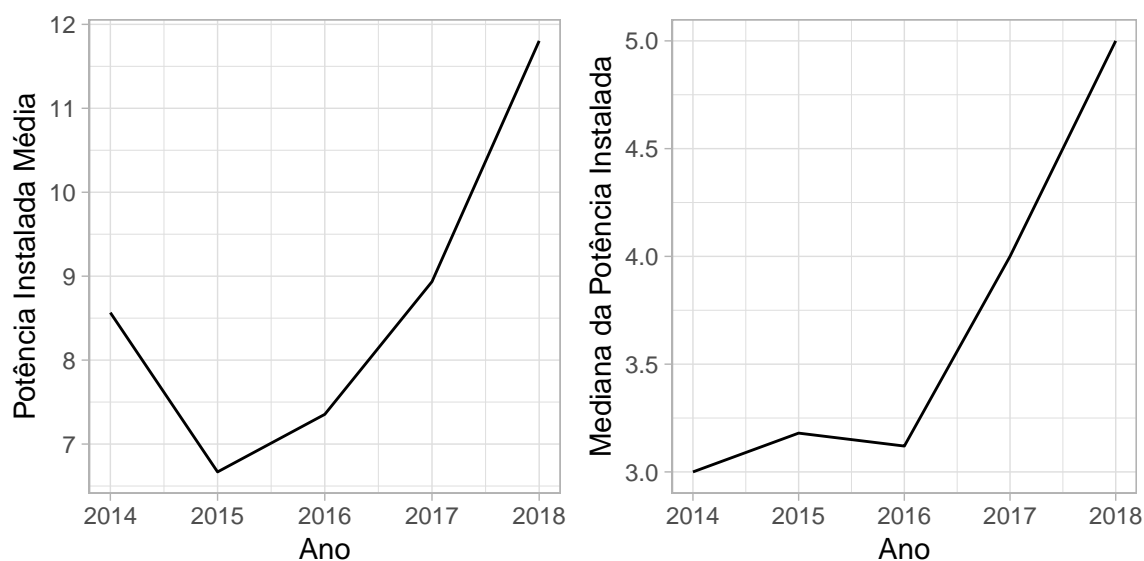
Figura 2 – Crescimento da Potência Instalada Total em Unidades de GDFV no Brasil



Regressão logarítmica para os período de julho de 2016 até novembro de 2018 mostrada em pontilhado.

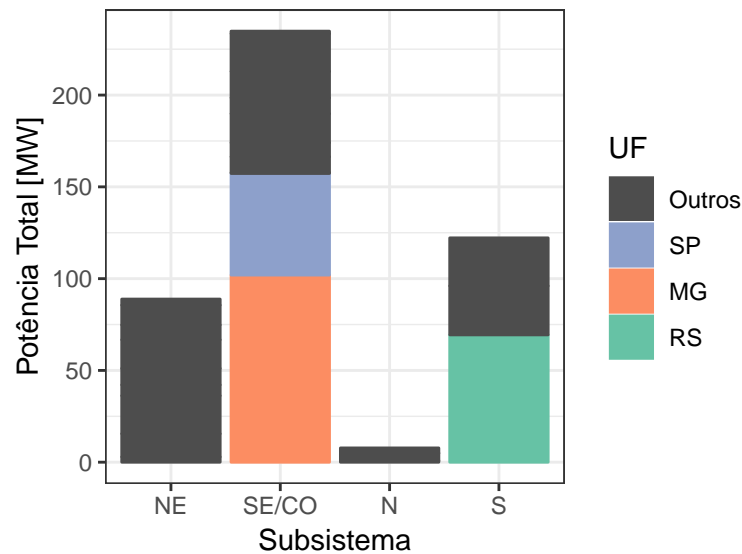
Elaboração própria
 Dados: ANEEL (2018b)

Figura 3 – Potência Média e Mediana dos Sistemas GDFV no Brasil ao Longo do Tempo



Elaboração própria
 Dados: ANEEL (2018b)

Figura 4 – Potência de GDFV Instalada por Subsistema e Estados Mais Destacados



Elaboração própria
Dados: ANEEL (2018b)

No contexto de grandes unidades geradoras centralizadas, o principal ambiente de participação de empreendimentos é o ambiente regulado, devido a seus baixos riscos. Neste contexto, a energia é contratada por meio de leilões. A fonte solar participou pela primeira vez em um leilão como produto exclusivo ao fim de 2014 (ANEEL, 2018a). Antes disso, nenhum empreendimento de energia solar havia ganhado um leilão de geração, devido a seus elevados custos e falta de competitividade. Deste então, mais quatro leilões foram realizados para esta modalidade de geração (ANEEL, 2018a), com destaque para o mais recente à data de escrita deste documento, o leilão 01/2018, realizado em abril do referido ano. Os preços médios de venda de energia através da fonte solar (118,07R\$/MWh) foram inferiores aos verificados para as fontes de biomassa (198,94R\$/MWh) e hidráulica (198,12R\$/MWh), no mesmo leilão, com o mesmo início de provimento. Neste leilão, empreendimentos relacionados à fonte solar apresentaram valores médios superiores apenas aos relacionados à fonte eólica (67,602R\$/MWh) (EPE, 2018). A redução nos preços da energia solar no Brasil pode ser observada com as informações relativas aos leilões mostrados na tabela 1, baseada em dados dos respectivos editais e nas taxas de câmbio do real em relação ao dólar americano nas datas de cada leilão. O preço em dólares americanos é exposto devido ao fato de que grande parte da tecnologia utilizada é proveniente de outros países, o que faz com que o preço em real apresente maiores flutuações em relação à sua cotação. Nestes dados, não são consideradas as variações do valor das moedas com o tempo, que ainda poderiam aumentar as diferenças nos preços de venda ao logo do período analisado.

A figura 5 mostra a distribuição por estado e subsistema da potência esperada instalada em 2022, quando os ganhadores dos leilões realizados até 2018 deverão estar

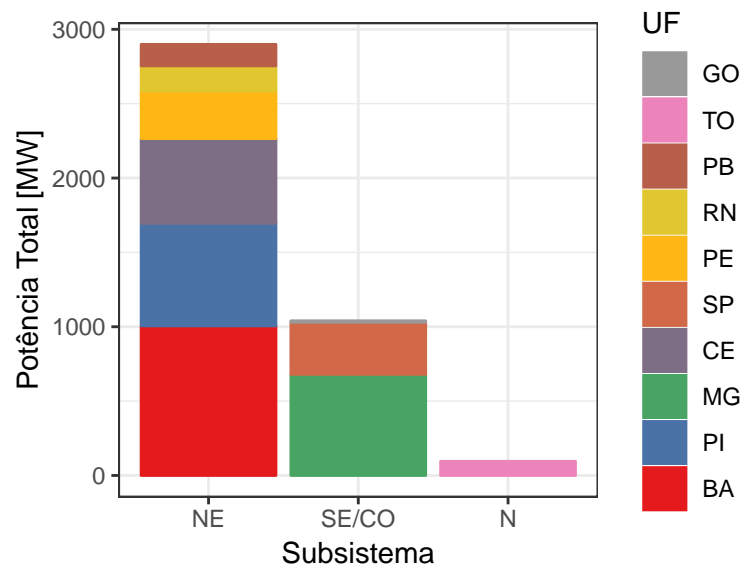
Tabela 1 – Resumo dos Leilões de Energia Solar

Leilão	Data do Leilão	Início do Provimento	Preço Médio [R\$]	Preço Médio [US\$]	Potência [MW]
08/2014	31/10/2014	01/10/2017	215,17	89,47	889,660
08/2015	28/08/2015	01/08/2017	301,79	84,88	833,802
09/2015	13/11/2015	01/11/2018	297,75	78,98	929,340
04/2017	18/12/2017	01/01/2021	145,71	44,22	574,000
01/2018	04/04/2018	01/01/2022	118,07	35,38	806,640

Elaboração própria

Dados: ANEEL (2018a)

Figura 5 – Localização das Usinas Solares Vencedoras de Leilões



Elaboração própria

Dados: ANEEL (2018a)

funcionando. Para esta modalidade, é interessante observar a grande predominância de sistemas na região Nordeste, seguida pelo estado de Minas Gerais, que apresenta diversos empreendimentos no seu noroeste, como Paracatu, Pirapora e Vazante. Esta área, do nordeste até o noroeste de MG é a área que apresenta maiores índices de radiação. Comparando os empreendimentos de geração centralizada com os de geração distribuída, pode-se verificar distribuições espaciais bastante distintas para as duas modalidades, que pode ser explicada considerando as diferentes regulações relativas a cada uma.

Levando em consideração um sistema eletroenergético que está passando por fortes transformações técnicas e estruturais, como destacado na Nota Técnica N° 5 de 2017 do Ministério de Minas e Energia (MME, 2017), junto a uma crescente utilização de sistemas de energia solar, o conhecimento sobre a dinâmica relacionada à disponibilidade de geração desta fonte de energia é importante para a tomada de decisão relativa às suas regras.

Como a produção de energia elétrica via sistemas fotovoltaicos ocorre apenas quando há radiação incidente em seus módulos, uma das características mais marcantes de sua produção energética é a da variabilidade. Esta variabilidade se dá devido a fatores previsíveis, relacionados à rotação e translação da terra, que determinam a inclinação relativa entre o sol e a superfície, e também por fatores altamente estocásticos, como as condições atmosféricas e de nebulosidade do céu (PEREZ; FTHENAKIS, 2015). Devido a estes fenômenos, não é possível dizer com certeza quando e o quanto de energia pode ser produzido por esta fonte. As dificuldades para se armazenar energia fazem com que tais variabilidades possam afetar o funcionamento dos sistemas eletroenergéticos, sendo, assim, importante compreender estes comportamentos variáveis (ENGELAND et al., 2017).

Variabilidades na produção energética solar podem ser analisadas em diferentes escalas de tempo e áreas de abrangência. Variabilidades de curtos intervalos de tempo, como de escalas de até uma hora, causadas por passagens de sistemas de nuvens menores, afetam o sistema em uma escala mais local e seus maiores impactos são relativos a questões de qualidade de energia elétrica, destacando-se a de flutuação de tensão em redes de distribuição de média e baixa tensão com alta densidade de GDFV (LAVE; RENO; BRODERICK, 2015; ELA et al., 2013; MILLS; WISER, 2010). Variabilidades relativas a análises de maiores intervalos de tempo persistem em áreas mais abrangentes, como nacionais ou regionais e, dentre seus impactos, tem-se questões tais como de disponibilidade de energia, mas que ainda foram pouco exploradas (HINKELMAN; SCHAEFFER, 2017). Bryce et al. (2018) destacam a importância da consideração de variabilidades a longo prazo, interanual, em análises de investimentos. Há, no entanto, muito poucos estudos na literatura sobre a dinâmica de longo prazo do recurso solar em áreas de tamanho continentais, como é o Brasil.

Para a análise de variabilidades espaçotemporais do recurso solar, diversos estudos utilizam técnicas de estatística multivariada, como análise de clusters e análise de correlações (ZAGOURAS; PEDRO; COIMBRA, 2014; GUTIÉRREZ et al., 2017; YANG et al., 2017). Essas análises são adequadas a este contexto uma vez que tem-se diversos elementos (locais de medição ou de estimação da irradiação incidente) com diversas variáveis cada (irradiações ao longo de um período). Assim, essas técnicas permitem a manipulação mais efetiva dos dados e a geração de conhecimento por meio destes. Entretanto, Engeland et al. (2017) observaram que não há uma estrutura de trabalho comum para estas análises. Além disso, algumas das práticas comumente utilizadas em escalas espaciais e temporais menores não são totalmente adequadas para a utilização em escalas maiores, como será discutido em seções posteriores. Uma das questões levantadas é relativa às variáveis a serem consideradas na análise. Em relação a isto, é proposta a utilização neste trabalho de decomposição de séries temporais, técnica comum nos estudos de econometria, que consiste em separar componentes (que, se somadas ou multiplicadas, geram a série de dados original), que são causadas por fenômenos diferentes e implicam em efeitos diferentes nos sistemas

(BEVERIDGE; NELSON, 1981; WEST, 1997).

1.1 Objetivos

Este trabalho busca contribuir para a compreensão e caracterização do recurso solar e de suas variabilidades e complementariedades espaçotemporais ao longo do território nacional, para servir de ferramenta que possa auxiliar no planejamento energético e definições de políticas relacionadas, visando a redução dos custos relativos à geração de energia elétrica e à operação eletroenergética, que, em última instância, recaem à população.

Considerando o fato de que a dinâmica do recurso solar é pouco explorada para escalas espaciais e temporais maiores, e a baixa disponibilidade de fonte de dados com resoluções menores, foi decidido focar este trabalho em efeitos de longas escalas, que afetam, principalmente, questões relacionadas ao planejamento. Assim, o objetivo deste trabalho é a proposição de uma metodologia para a compreensão do comportamento variável do recurso solar em grandes áreas, e a sua aplicação na área do Brasil.

1.1.1 Objetivos específicos

Os objetivos específicos deste trabalho são:

- Revisar estudos que fazem a caracterização da variabilidade do recurso solar em largas áreas.
- Obter e tratar dados relacionados ao recurso solar, provenientes de fontes gratuitas para a área do Brasil, para gerar um banco de dados que possa ser utilizado com maior facilidade.
- Propor e implementar uma metodologia de análise do recurso solar baseada na decomposição de séries temporais e na estatística multivariada, que adicionam às metodologias de exploração do recurso solar da literatura.
- Verificar a aplicabilidade da metodologia proposta.
- Extrair resultados relativos ao recurso solar no território brasileiro que possam ser úteis no contexto do planejamento eletroenergético, identificando locais que apresentem complementariedades entre si do recurso em questão.
- Comparar dados de bancos diferentes, verificando a consistência entre eles.

1.2 Estrutura do Trabalho

Este trabalho está estruturado da seguinte maneira: neste capítulo, foi apresentado a questão abordada na presente dissertação e os seus objetivos. No capítulo 2 é feita

uma revisão bibliográfica relativa à variabilidade espaçotemporal do recurso solar, com foco maior nos estudos que consideram escalas mais amplas de tempo e espaço. O capítulo 3 apresenta os dados e discute os métodos utilizados na análise deste trabalho. O Capítulo 4 apresenta os resultados obtidos e discussões em relação a estes. O capítulo 5 conclui o trabalho e apresenta uma discussão final sobre os métodos utilizados e resultados obtidos.

2 VARIABILIDADE DO RECURSO SOLAR

Este capítulo apresenta uma revisão da literatura relacionada à investigação da variabilidade do recurso solar e às estratégias para se lidar com seus efeitos relacionados nos sistemas eletroenergéticos. Dado que o crescimento da utilização da geração distribuída fotovoltaica ocorre de maneira desequilibrada, alguns alimentadores, geralmente em regiões com casas de proprietários com alto poder financeiro, passam a apresentar penetrações de GDFV muito maiores do que a de sua região. Assim, impactos elétricos locais são, geralmente, os primeiros limitadores da utilização da energia solar. Portanto, há, na literatura, um foco maior em estudos relacionadas às dinâmicas espaçotemporais do recurso solar que afetam áreas pequenas, de redes elétricas locais, que são relacionados a escalas de tempo intradiárias e com resoluções, em alguns casos, inferiores à de minutos (ENGELAND et al., 2017). Apesar da literatura mais ampla sobre variabilidades relativas a baixas escalas de tempo, neste trabalho há um foco maior nos estudos que tratam efeitos em escalas maiores, com resoluções diárias ou superiores, utilizando dados interanuais, uma vez que estes também são importantes e se assemelham mais aos objetivos deste trabalho, como discutido no capítulo 1.

Mills e Wiser (2010), Hoff e Perez (2010), Lave e Kleissl (2010), apresentam estudos pioneiros em avaliar como a dispersão geográfica afeta a variabilidade do recurso solar, focados na variação de passo da produção. Em Mills e Wiser (2010), é observado que a correlação das variabilidades do recurso, definidas como as variações do índice de céu claro entre dois intervalos, diminui com a distância, sendo que esta diminuição ocorre mais rapidamente para passos de escalas de tempo mais curtas. Isso gera o efeito de suavização, que faz com que a variabilidade relativa total de um portfólio de sistemas dispersos seja consideravelmente menor do que aquela apresentada por um único sistema. Assim, a dispersão de sistemas geradores é uma das estratégias mais interessantes para se lidar com fenômenos relacionados à variabilidade do recurso solar, dado os altos custos de outras soluções, como a de sistemas armazenadores. Diversos artigos aprofundaram no assunto, ajustando a modelagem do efeito de suavização e das correlações, ou estudando efeitos em locais específicos (HOFF; PEREZ, 2012; PEREZ et al., 2012; HINKELMAN, 2013; DAVID; ANDRIAMASOMANANA; LIANDRAT, 2014; PEREZ; FTHENAKIS, 2015; TRIPATHY et al., 2017). Perez e Hoff (2013), Perez et al. (2016) resumem os principais estudos que visam compreender a variabilidade espaçotemporal do recurso solar considerando o efeito de suavização das variabilidades no índice de claridade. É destacado que essas correlações são bem modeladas utilizando informações sobre as distâncias entre os locais, a dimensão temporal do passo, e a velocidade do vento, levando em conta sua direção.

Há, também, estudos que focam em outras questões que levam em conta a variabili-

dade solar espaçotemporal, como relacionados à previsão do comportamento do recurso energético (RAZA; NADARAJAH; EKANAYAKE, 2016; MURATA; OHTAKE; OOZEKI, 2018; AGUIAR et al., 2016). Outros estudos otimizam a qualidade de portfólios de sistemas geradores ou de redes de estações solarimétricas, visando minimizar impactos relacionados à variabilidade (DAVY; TROCCOLI, 2014; DEETJEN et al., 2018; JEREZ et al., 2013; URQUHART; SENGUPTA; KELLER, 2013) .

Estudos que avaliam variabilidades em escalas espaciais e temporais mais longas, que são úteis aos planejadores dos sistemas elétricos e energéticos, entretanto, ainda não são tão comuns (PEREZ; FTHENAKIS, 2015; ENGELAND et al., 2017). Boa parte dos artigos sobre o recurso solar em grandes áreas, focam apenas em médias de longo prazo, sejam elas diárias, mensais ou anuais. (NUÑO et al., 2018; KAPLANI; KAPLANIS; MONDAL, 2018). Os principais estudos da variabilidade solar em largas áreas, considerando extensões de países de médio porte até tamanhos continentais, ou em elevadas escalas temporais, considerando dinâmicas entre dias ou mais lentas, são apresentados a seguir.

Alguns artigos estudam a variabilidade espacial e temporal separadamente, como em Gueymard e Wilcox (2011), que utiliza dados de 37 estações solarimétricas e dados modelados a partir de imagens de satélite para avaliar o comportamento do recurso solar a longo prazo, em todo os Estados Unidos da América. Para isso, usa o coeficiente de variação e faz estimativas do número de anos necessários para a média das radiações anuais convergir à média de longo prazo considerando o pior cenário. Neste artigo, também, é feita uma análise da variabilidade de médias anuais e mensais entre localidades próximas. Esta análise, entretanto, apenas quantifica as diferenças entre as médias de longo prazo das localidades, sem levar em consideração, por exemplo, a correlação entre as variações nas localidades.

Kariuki e Sato (2018) utilizam a mesma metodologia para avaliar a variabilidade do recurso solar no Quênia. Gil et al. (2015) também estudam a variabilidade interanual para diversas localizações da península ibérica de maneira similar. Perez e Fthenakis (2012) avaliam, a partir de dados de satélite, variabilidades de longas escalas do recurso solar em torno de um ponto específico. Para isso, verificam o comportamento da variabilidade da radiação e do coeficiente de claridade, considerando diferentes escalas de tempo e de espaço de integração separadamente. É verificado que, em geral, quanto maior a escala de integração de tempo e de área, menor a variabilidade.

Em Perez e Fthenakis (2015), os autores utilizam dados de satélite referentes a todo o continente americano para realizar estudos de correlação entre séries temporais de variações do índice de claridade para pares de localizações, visando compreender a dinâmica do recurso solar. As escalas de tempo utilizadas apresentam janelas variando entre 1 e 30 dias. Assim, é verificado um comportamento médio das correlações entre pares que apresenta decaimento exponencial com a distância, que ocorre mais lentamente quanto maior a escala

de tempo considerada, já verificado para escalas de tempo intradiárias (HOFF; PEREZ, 2012). Apesar de o estudo apresentar esta relação para a tendência média, pode ser observado que há uma dispersão considerável em torno deste comportamento, especialmente quando considerados intervalos de tempo mais elevados. Isto sugere que uma análise generalista pode não ser suficiente para se ter um bom entendimento de como ocorrem as correlações entre duas sequências temporais de locais específicos. Um entendimento de localizações que desviam deste comportamento médio pode ser estrategicamente importante para o planejamento dos sistemas elétricos e energéticos.

Outra abordagem para analisar e explorar variabilidades espaçotemporais do recurso energético solar se dá com a verificação de como a dinâmica da radiação solar ao longo do tempo interage com dinâmicas de modos climáticos, que são padrões de variabilidade climática de baixa-frequência, que podem ser expressos por um índice (SANTOS-ALAMILLOS et al., 2012; JEREZ; TRIGO, 2013; KRAKAUER; COHAN, 2017; MOHAMMADI; GOUDARZI, 2018). Por exemplo, Krakauer e Cohan (2017) apresentam mapas mundiais de correlação interanual da radiação solar e da velocidade do vento com diversos modos climáticos e (MOHAMMADI; GOUDARZI, 2018) verificam como o ENSO (*El Niño Southern Oscillation*, ou, em português, Oscilação Sul-El Niño) afeta a dinâmica de recursos renováveis na Califórnia. Esta abordagem permite uma boa compreensão de alguns aspectos da dinâmica solar e permite um entendimento de como ela ocorre junto a outros fenômenos. Entretanto, verifica o comportamento do recurso solar apenas indiretamente, pois nem sempre os modos climáticos podem ser utilizados com um bom poder de explicação da dinâmica solar.

2.1 Uso de Técnicas de Estatística Multivariada na Caracterização da Variabilidade do Recurso Solar

Uma estratégia utilizada para se estudar a dinâmica do recurso solar em determinada área é a utilização de análise de agrupamentos ou análise de *cluster*, que consiste em agrupar elementos de um banco de dados a partir de um determinado algoritmo. No contexto da dinâmica solar, Zagouras et al. (2013) inicialmente propôs a utilização desta técnica para se determinar regiões coerentes e utilizá-las como base na definição de locais para se instalar sensores de radiação. O mesmo autor, depois utilizou a análise de agrupamentos para auxiliar no entendimento do comportamento do recurso solar em certas pequenas áreas, como uma ilha do Havaí e uma parte da Califórnia (ZAGOURAS; INMAN; COIMBRA, 2014; ZAGOURAS; PEDRO; COIMBRA, 2014). Em relação a áreas maiores, como regionais ou de países de porte médio, a análise de *cluster* foi utilizada para avaliar a variabilidade do recurso solar na Península Ibérica (GUTIÉRREZ et al., 2017), no Vietnã (POLO et al., 2015) e no Japão (WATANABE; TAKAMATSU; NAKAJIMA, 2016). Há, também, estudos que definem regiões para a análise manualmente, sem a utilização de algoritmos de agrupamento, como Bett e Thornton (2016), que dividem as ilhas britânicas em seis regiões, visando avaliar a

dinâmica dos recursos solares e eólicos. Uma vez que regiões coerentes são definidas por meio de agrupamentos, pode-se estudá-las separadamente ou avaliar suas interações com as outras. Por exemplo, Gutiérrez et al. (2017) utilizam, após a estratificação, análises de correlação, para identificar complementariedades locais relacionadas à energia solar. Na seção 3.4 esta abordagem será discutida e uma proposta alternativa, apresentada.

Neste trabalho, análise de agrupamentos é proposta para se estudar a dinâmica do comportamento do recurso solar em uma área maior, de tamanho continental. O algoritmo utilizado para realizar a análise de agrupamentos é o PAM (sigla inglês para Particionamento em Torno de Medóides) (KAUFMAN; ROUSSEEUW, 1987), que apresenta similaridades com o k-means, utilizado nos trabalhos supracitados, mas com algumas vantagens e que ainda não foi usado na literatura para análise do recurso solar. Mais detalhes sobre a o algoritmo e sua escolha podem ser vistos na seção 3.3. Na mesma seção, também é discutida a escolha do número de partições utilizadas na análise, que difere das utilizadas nos trabalhos que utilizam análise de *cluster* no contexto da dinâmica do recurso solar. Os grupos obtidos são então utilizados para avaliar dinâmicas e complementariedades entre pares.

Outra técnica muito usual no contexto de estatística multivariada é a de análise de componentes principais (ou PCA, do inglês *principal components analysis*), um método de redução de dimensionalidades. Na literatura sobre o recurso solar, a análise de componentes principais foi utilizada principalmente visando fazer com que algoritmos, como de clusterização, fossem executados mais rapidamente (ZAGOURAS et al., 2013; ZAGOURAS; INMAN; COIMBRA, 2014; ZAGOURAS; PEDRO; COIMBRA, 2014; GUTIÉRREZ et al., 2017). Yang et al. (2017) utilizam tal técnica para caracterizar localizações, usando como base características, ou *features*, relacionadas a sistemas de energia solar nestas, como ângulos de azimute e de inclinação ótimos. Entretanto, não utiliza a técnica na compreensão do comportamento dinâmico do recurso energético, como é proposto neste trabalho.

2.2 Decomposição de Séries temporais

Uma técnica empregada comumente em análise de séries temporais é a da decomposição destas séries, dado que diferentes componentes de uma série podem ter causas e consequências em diferentes âmbitos. O estudo de cada uma dessas componentes separadamente permite a criação de melhores modelos e análises mais interpretáveis, como se dá no caso da econometria, em que efeitos de tendência, sazonais e de ciclos de mercado são, muitas vezes, separados (BEVERIDGE; NELSON, 1981; WEST, 1997).

Em relação ao recurso solar, este conceito já foi utilizado, principalmente, em estudos de previsão de seu comportamento (YANG et al., 2015; YANG; DONG, 2018). Entretanto, não é utilizado diretamente em artigos focados na caracterização e entendimento das variabilidades do recurso solar. Nestes, muitas vezes, índices, como o índice de claridade,

o índice de céu claro, ou o índice de claridade corrigido, são utilizados visando remover efeitos nas séries temporais causados pela posição e inclinação relativa do sol com os locais analisados (PEREZ et al., 2016; LUIZ et al., 2018). A utilização destes índices, entretanto, diminui o significado prático da análise, especialmente quando são consideradas áreas mais abrangentes, uma vez que o mesmo valor do índice pode representar valores muito diferentes de irradiação. Outro problema da utilização destes é que eles podem não remover todos os efeitos de sazonalidade, pois aqueles devidos a efeitos climáticos, e não apenas da posição relativa do sol, persistem.

Como discutido neste capítulo, diversos estudos utilizam de séries de variações do índice de claridade, ou similar, entre dois intervalos de tempo consecutivos, considerando diferentes resoluções temporais para caracterizar a variabilidade solar para diferentes escalas de tempo (PEREZ et al., 2016). Esta é uma estratégia interessante, uma vez que consegue filtrar efeitos de outras escalas de tempo, como tendências, sendo também útil para fazer a série estacionária. Por essas características a diferenciação permite boas modelagens do sistema, sendo por isso, usado em métodos como o ARIMA (CONTRERAS et al., 2003). Também, um dos maiores problemas da utilização de sistemas fotovoltaicos, em relações a escalas de tempo menores, é relativo às dinâmicas de rampa de geração, que estão associados a essas variações entre um período e outro, e à consequente necessária alocação de outros sistemas geradores com alta taxa de rampa. Considerando escalas de tempo mais longas, entretanto, a variabilidade solar afeta principalmente questões relacionadas ao planejamento energético, de modo em que é mais interessante compreender o comportamento da irradiação do que o comportamento de suas variações entre dois períodos. Assim, neste trabalho, que busca um entendimento prático da dinâmica do recurso solar, a estratégia da avaliação das variações entre intervalos consecutivos de medição não é utilizada e é proposta a utilização da decomposição de séries temporais para o estudo da variabilidade solar a longo prazo considerando suas diferentes dinâmicas.

2.3 Estudos da Variabilidade do Recurso Solar no Brasil

No contexto brasileiro, poucos estudos relativos à variabilidade do recurso solar são encontrados. Um destaque se dá para as duas edições do Atlas Brasileiro de Energia Solar (PEREIRA et al., 2006; FERREIRA et al., 2017). Estes documentos caracterizam, a partir de modelos calibrados de dados de satélite, a radiação solar no território brasileiro. Em relação à variabilidade, nestes documentos são feitos estudos sobre a tendência e a variabilidade interanual (em um período de 10 anos cada). A tendência é, na primeira edição, estimada por meio de uma regressão linear simples utilizando as séries de dados de irradiação anuais médios de cada uma das cinco regiões do Brasil. Já na segunda, estas séries (para um período posterior) são avaliadas com o teste de Mann-Kendall, para verificar se há tendência de crescimento ou decrescimento para a irradiação média ao longo dos anos analisados.

Uma questão relacionada a essa abordagem é a utilização das regiões geopolíticas do Brasil, sem considerar a estratificação por clima ou pela própria observação do comportamento do recurso solar.

Em relação a variabilidades interanuais, a primeira edição do atlas as avalia por meio índices, calculados para cada pixel analisado, que estima a variabilidade do local. Nesta edição, variabilidades intra-sazonais também são avaliadas com índices similares. Além disso, diagramas de caixa são utilizados para mostrar a dispersão do comportamento das médias anuais e das médias sazonais para cada ano por região. Na segunda edição, apenas esta abordagem é utilizada. Essas duas metodologias permitem a visualização da variabilidade a longo prazo no território brasileiro e a identificação de regiões com menos variações. Entretanto, não analisa variabilidades espaçotemporais e possíveis complementariedades do recurso solar ao longo da área nacional.

Dentre outros estudos que avaliam o comportamento variável do recurso solar no Brasil, destacam-se Anjos et al. (2015), que observam e correlacionam o comportamento a longo prazo do recurso solar e do recurso eólico, Luiz et al. (2018), que investigam o comportamento da variabilidade de curto prazo em 3 locais no país, em regiões climáticas distintas, Souza, Nicácio e Moura (2005), que investigam este comportamento para Maceió, e Silva et al. (2010), que avaliam o comportamento de tendência do recurso solar na região nordeste do país. É importante observar que nenhum estudo encontrado avalia variabilidades espaçotemporais de maneira conjunta, como é feito neste trabalho.

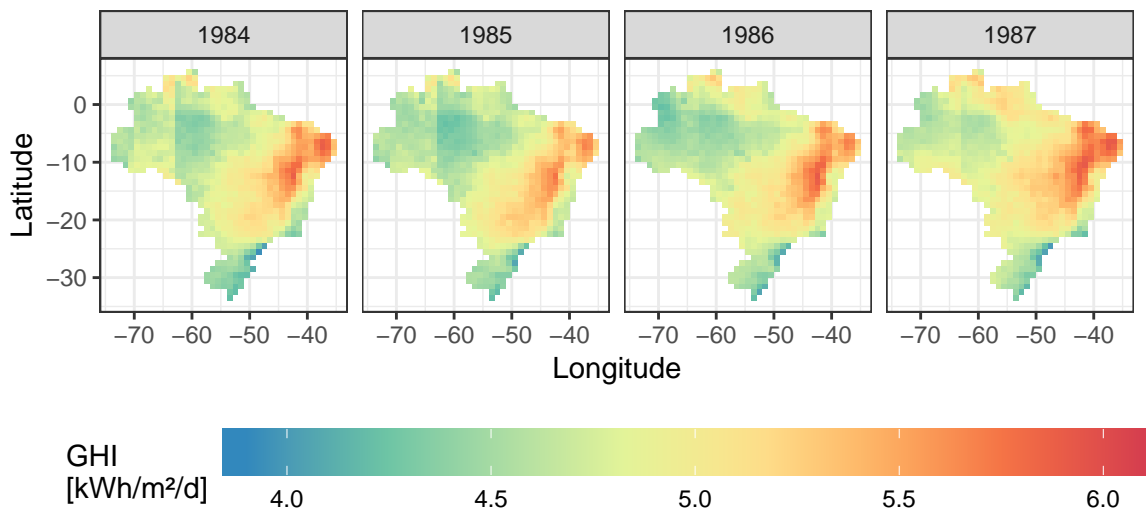
3 DADOS UTILIZADOS E METODOLOGIA PROPOSTA

Este capítulo apresenta os dados utilizados, as ferramentas computacionais e os métodos propostos e implementados para a compreensão da variabilidade a longo prazo do recurso solar. Alguns pontos de metodologias referenciadas no capítulo 2 são aprofundados visando destacar diferenças com abordagens propostas.

3.1 Dados Utilizados

Este trabalho utiliza, principalmente, dados provenientes do projeto POWER (*Prediction Of Worldwide Energy Resources*, em tradução livre para o português, Predição Mundial de Fontes de Energia) da agência espacial americana, NASA. Este projeto disponibiliza, gratuitamente, dados climáticos relacionados à geração de energia (como velocidade de vento em diferentes alturas, temperatura, precipitação e irradiação), provenientes de processamento de imagens satelitais. São disponibilizados dados diários de irradiação obtidos com o modelo SRB (*Solar Resource Budget*, ou, em tradução livre para o português, Orçamento do Recurso Solar) v3.0 do projeto GEWEX (*Global Energy and Water Exchanges*, ou, em tradução livre para o português, Trocas Globais de Energia e Água) para o período de Julho de 1983 até Dezembro de 2007. A partir desta data, até poucos dias antes do dia de consulta, também são disponibilizados dados obtidos com diferentes versões do modelo FLASHFlux. Neste trabalho foram utilizados apenas dados provenientes do modelo do GEWEX, uma vez que contem a mais longa série consistente em um só modelo. Em Stackhouse et al. (2018), podem ser encontradas as modelagens utilizadas no projeto POWER, junto às suas validações.

Os bancos de dados de irradiação são disponíveis em uma base de resolução diária para uma grade mundial com dimensões de 1° por 1° . Cada elemento com dimensões de 1° de longitude por 1° de latitude, ou *pixel*, contem séries de dados médios ao longo da área que representa. Para cada *pixel* no Brasil foi baixada a respectiva série histórica de irradiação diária. Para isso, foi desenvolvido pelo autor, um programa em R (R Core Team, 2018) que acessa e extrai dados do sistema POWER, por meio de sua API (*Application Programming Interface*, ou, em tradução para o português, Interface de Programação de Aplicativos). *Pixels* em fronteiras com o mar cujas áreas são predominantemente ocupadas por águas, foram retirados da análise por apresentarem comportamentos singulares e não esperados, como valores muito altos de irradiação. Tais comportamentos são apresentados possivelmente devido a dinâmicas climáticas diferentes em áreas marítimas que compõem o *pixel*, ou questões referentes ao processamento de imagens de satélites nestas áreas. Dados relativos ao período entre 1988 e 2007 foram utilizados neste trabalho, uma vez

Figura 6 – Irradiação Global Horizontal Média Verificada para os *Pixels* Analisados Para os Anos de 1984 a 1987

Elaboração própria

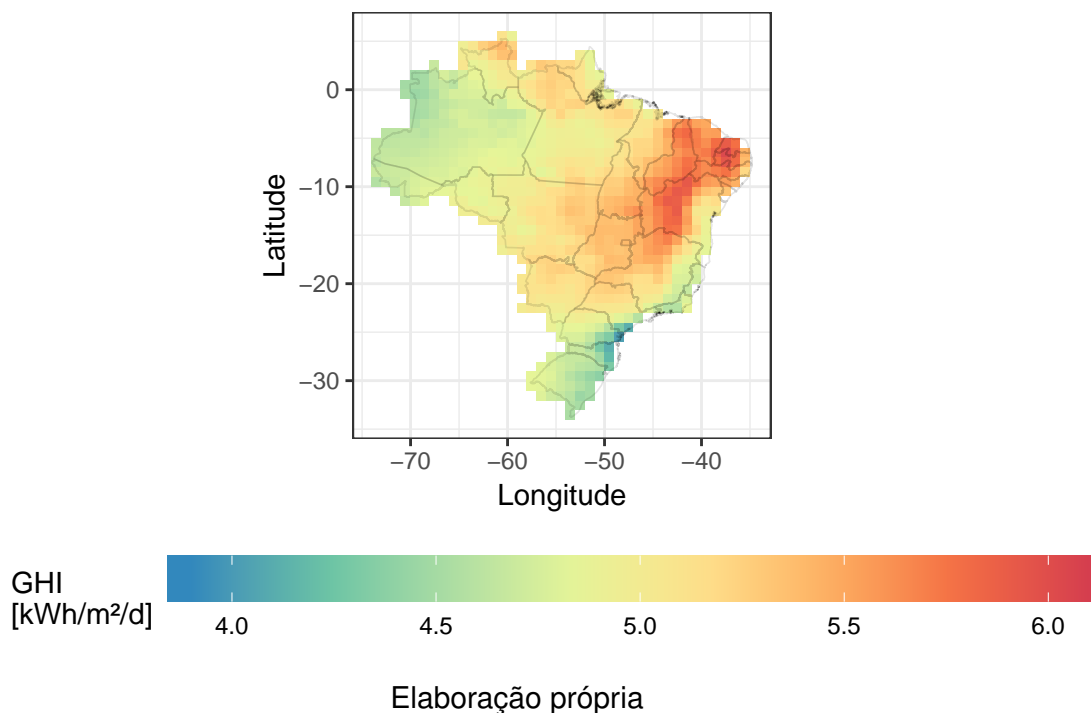
que são dados obtidos por um mesmo modelo, e que não apresentaram comportamentos anormais. Foi verificado que dados anteriores a 1988 apresentavam uma descontinuidade espacial em médias anuais próximo à longitude -63° , como é mostrado na figura 6, por isso foram desconsiderados. A figura 7 mostra as médias diárias para o período considerado da irradiação global no plano horizontal para os *pixels* utilizados na análise.

Além destes dados, outro banco de dados foi utilizado, com intuito de comparar e validar resultados. Este banco é o que foi utilizado para geração dos anos típicos meteorológicos no projeto SWERA (RENNÉ et al., 2005) (*Solar and Wind Energy Resource Assessment*, ou, em tradução livre para português, *Avaliação dos Recursos Energéticos Solar e Eólico*), que foi um projeto de parceria internacional com missão de fornecer informações sobre energia renovável pelo mundo, com foco em países em desenvolvimento. No Brasil, o projeto começou em 2001 e teve como principal produto a primeira edição do Atlas Brasileiro de Energia Solar.

O banco de dados utilizado possui séries temporais horárias para 20 estações meteorológicas, para períodos de até 30 anos. Os dados de cada estação são disponíveis em um arquivo compactado, contendo arquivos com dados anuais que podem ser lidos por meio de uma macro de Excel disponibilizada. Para poder extrair os dados de todos os anos e para todas as estações meteorológicas, foi desenvolvido um programa em R (R Core Team, 2018) para ler cada arquivo anual contido nos arquivos compactados e obter um único arquivo contendo as séries temporais completas para todas as estações meteorológicas.

Considerando o objetivo deste trabalho de observar dinâmicas de escalas de tempo

Figura 7 – Irradiação Global Horizontal Média Verificada para os *Pixels* Analisados Para o Período de 1988 a 2007



maiores, os dados de ambas as fontes foram utilizados para gerar séries de resolução mensal, que apresentam tamanhos reduzidos e, conseqüentemente, maior manuseabilidade.

Como discutido no Capítulo 2, na maioria dos estudos referentes à variabilidade do recurso solar, a análise é baseada em valores de índices, como o índice de céu claro K_t^* , definido como a razão da irradiação no plano horizontal terrestre pela irradiação incidente para a mesma localização considerando um dia de céu claro (PEREZ et al., 2016). Tais índices são usados visando remover efeitos determinísticos da dinâmica solar. Neste trabalho, entretanto, são utilizados valores de irradiação diretamente. Esta escolha se dá por alguns motivos, sendo o primeiro, a consideração de que a irradiação tem um significado muito mais prático, dado que a produção solar é altamente dependente da irradiação incidente para qualquer tempo, enquanto que, para avaliar a produção energética utilizando o K_t^* , é necessário um maior processamento de informações. Ainda, a utilização de tais índices em áreas grandes pode falhar em observar dinâmicas do recurso energético em questão, dado que áreas distantes, com radiação extraterrestre muito diferentes, ao serem analisadas ao mesmo tempo, podem apresentar dinâmicas muito distintas entre os valores de K_t^* e os do recurso energético.

Considerando a utilização de dados de irradiação, é desejável que componentes das séries históricas sejam estudadas separadamente, uma vez que efeitos com características diferentes causam impactos nos sistemas elétricos e energéticos distintos e que podem ser tratados por mecanismos também distintos. Na literatura técnica, a maneira mais comum

de se analisar efeitos em escalas distintas é a avaliação das séries de diferenças das observações entre dois intervalos, ou *step-changes*, considerando janelas de integração variáveis. A utilização desta estratégia faz muito sentido quando resoluções temporais menores são analisadas, devido a que nessa situação trata-se de evitar, principalmente, variações rápidas nas potências de saídas dos geradores. Entretanto, ao analisar intervalos maiores, esta estratégia não apresenta muito significado prático. Para tal análise, é proposto, neste trabalho, o uso de decomposição das séries históricas em séries temporais que representam impactos diferentes no planejamento, como descrito na sequência.

3.2 Decomposição das Séries Temporais

Para cada *pixel*, é proposto decompor as séries temporais históricas de irradiação mensal em quatro componentes: de tendência, sazonal, interanual e de resíduo mensal. A seguir é discutido o motivo e o método de obtenção de cada componente.

A componente de tendência é obtida visando remover possíveis vieses persistentes dos dados originais, que podem mudar lentamente com o tempo. Desta maneira, mudanças relativas a alterações de satélites, calibrações e metodologias podem ser suavizadas com a remoção desta componente. Investigações sobre essa componente podem, também, fornecer informações sobre uma dinâmica lenta do comportamento do recurso solar. Entretanto, esta análise deve ser feita com cuidado, pois as variações analisadas podem ser apenas em erros das medições e estimativas. Levando em conta o objetivo de obtenção desta componente, neste trabalho, ela é estimada utilizando *loess* (*locally estimated scatter-plot smoothing*, ou, em tradução livre, suavização estimada localmente em gráfico de pontos.), um método não paramétrico comum na econometria para estimação de tendências dinâmicas, que inclusive serve de base para o método de decomposição STL (*Seasonal Trend decomposition using Loess*, ou, em tradução livre para o português, Decomposição Sazonal e de Tendência usando *loess*), proposto por Cleveland et al. (1990), e amplamente utilizado. O *loess* é utilizado devido às suas qualidades de não diminuir o horizonte da análise, como, por exemplo, faria um filtro de médias móveis, e conseguir capturar dinâmicas locais que podem variar ao longo do tempo, algo que pode falhar em regressões lineares ou polinomiais de baixa ordem, por exemplo. A sua popularidade, devido a seu considerado bom desempenho em estudos de séries temporais, também foi levada em conta nesta escolha.

O funcionamento do *loess* consiste em, para cada ponto, estimar uma regressão polinomial de baixa ordem em torno de sua região. Esta regressão é estimada usando mínimos quadrados ponderados, em que pontos próximos ao local da estimativa possuem pesos maiores. Detalhes sobre algoritmo podem ser verificados em ampla bibliografia, como no artigo que ajusta o método para o utilizado atualmente (CLEVELAND; DEVLIN, 1988). Neste trabalho, é utilizado um *loess* com regressão de segunda ordem e considerando 15 anos de dados para cada regressão, que foi estimado usando a função `loess()` presente

no R (R Core Team, 2018). Os parâmetros foram escolhidos verificando visualmente as componentes de tendência obtidas.

A componente interanual, como o nome diz, busca capturar fenômenos interanuais. Esta componente é obtida para separar fenômenos mais lentos dos efeitos intra-anuais, sazonais ou não. Essa escala de tempo foi definida levando em conta, também, questões de funcionamento do sistema energético, como o fato de que o saldo de energia relativo a empreendimentos solares contratados nos leilões que já ocorreram no Brasil é contabilizado em um período de um ano (ANEEL, 2018a), de forma que, para um investimento, oscilações com escalas de tempo mais curtas não fazem tanta diferença. Esta componente é obtida utilizando um filtro de médias móveis, de largura igual a 12 meses, nas séries temporais com suas componentes de tendência retiradas. Assim, pode ser considerada como a agregação de 12 séries temporais de resolução anual, cada uma centrada em um mês. Desta maneira, esta componente é útil para a avaliação de questões relativas ao planejamento no contexto interanual. Esta componente, juntamente com as componentes sazonal e de resíduo mensal, descritas no parágrafo subsequente, são extraídas simultaneamente com a função `decompose()`, que opera a decomposição baseada em médias móveis.

Retiradas as componentes interanuais e de tendência, a série restante retrata dinâmicas intra-anuais, com resolução mensal. Estas dinâmicas podem ser geradas por fenômenos que aqui são divididos em duas categorias. A primeira é a de efeitos que ocorrem recorrentemente nos anos, devido a padrões climáticos típicos e a questões da inclinação relativa da Terra em relação ao Sol. Note que apesar de K_t^* e suas variações removerem as questões relativas à dinâmica espacial da Terra com o Sol, eles não levam em conta os padrões climáticos recorrentes. A segunda, se refere aos desvios desses padrões de comportamento. Estas duas questões afetam o planejamento de maneira diferente, sendo que a primeira pode ser usada para buscar um portfólio que coincida mais com o consumo típico de energia e que seja complementar a efeitos sazonais da geração por meio de outras fontes renováveis. A segunda categoria, por sua vez, gera incertezas, que são indesejadas e busca-se minimizá-las. Desta maneira, são geradas componentes para analisar cada um desses pontos separadamente. Para cada mês, a componente sazonal é definida como o valor médio de irradiação diária (retiradas as componentes de tendência e interanual) para aquele mês considerando todo o período analisado. A série restante, retiradas todas as componentes descritas anteriormente, é definida como a série de resíduo mensal, que representa os desvios das sazonalidades da dinâmica do recurso solar.

Pode-se observar que o método de decomposição de séries temporais utilizado consiste na aplicação de um *loess* seguido por uma decomposição clássica com médias móveis. Ressalta-se aqui os motivos de aplicação do *loess* para obtenção da componente de tendência, visando a remoção de possíveis vieses persistentes para a análise interanual. Se não fosse aplicado este filtro, a decomposição por médias móveis iria obter uma componente

chamada na econometria de componente de tendência mais ciclo. Assim, o *loess* serve para dividir essas componentes, e a componente de ciclo é chamada, neste trabalho, de componente interanual.

3.3 Particionamento em Torno de Medoides

Para o estudo das componentes interanual e de resíduo mensal, são utilizados dados de *pixels* representativos dos dados, que são extraídos visando reduzir as dimensões dos dados e assim gerar análises mais interpretáveis. A obtenção destes pontos é feita por meio do algoritmo PAM (*Partitioning Around Medoids*, ou, em tradução livre, Particionamento em Torno de Medoides). Este é uma implementação do método de agrupamento chamado de *k-medoides*, que é baseado na ideia do *k-médias*, sendo, também, um método de agrupamento baseado em partições. No método *k-medoides*, cada grupo é representado por um medoide, que é o ponto mais representativo do respectivo grupo, enquanto que, no *k-médias*, os grupos são representados por centroides, que são compostos pelas médias entre os elementos do grupo para cada variável da análise. Esse tratamento faz com que o método escolhido seja mais robusto. No contexto deste trabalho, utilizar uma implementação do *k-medoides* trás, também, a vantagem sobre o *k-médias* de definir um pixel representativo para cada grupo obtido. Com isso, é possível, por exemplo, ter uma distância geográfica exata entre dois grupos, representados por seus medoides, que pode ser avaliada junto a outras informações dos pares, como a correlação de suas séries temporais. A implementação do *k-medoides* mais comumente utilizada é o algoritmo PAM, que possui a vantagem, sobre o *k-médias*, de ser determinístico para bancos de dados não simétricos, como é na maior parte dos casos.

O algoritmo PAM é estruturado em duas fases, a de construção e a de permutação. A fase de construção consiste em selecionar um número definido k de pontos para comporem o grupo inicial de candidatos a serem os medoides. A fase de permutação busca refinar este grupo de medoides, em uma busca local.

A fase de construção opera da seguinte maneira: inicialmente considera-se o grupo dos medoides iniciais vazio. Então, é encontrado o elemento cuja soma das dissimilaridades com todos os outros elementos seja a menor possível. Este elemento é adicionado ao grupo dos medoides iniciais. Então, iterativamente é encontrado o elemento que, quando adicionado ao grupo de medoides iniciais, minimiza a soma das dissimilaridades entre os elementos não pertencentes ao grupo de medoides iniciais e os respectivos elementos mais similar presentes no grupo. Este procedimento é repetido até que este grupo de medoides iniciais seja composto por k elementos.

A fase de permutação consiste em verificar se qualquer operação de permutação entre elementos no grupo de medoides e elementos fora deste grupo poderia diminuir a soma das dissimilaridades entre cada ponto e seu medoide mais próximo. Se forem encontradas

operações de permutação que reduzam a soma de dissimilaridades, é executada a operação que a reduz em maior valor. Esta operação é executada iterativamente, até que não haja mais operação de permuta que possa reduzir a soma de dissimilaridades. Mais detalhes deste algoritmo podem ser encontrados na referência, (KAUFMAN; ROUSSEEUW, 1987; REYNOLDS et al., 2006). É interessante ressaltar que este método é, muitas vezes, descrito de maneira errada em sites da internet e que a consulta à bibliografia referenciada apresenta o algoritmo original e implementado no R. Vale notar que, devido a sua estrutura, este algoritmo pode executar de maneira lenta para bancos de dados maiores. Neste caso, variações deste como o CLARA (*Clustering for Large Applications* ou, em tradução livre, Agrupamento para Grandes Aplicações) podem ser bons substitutos (KAUFMAN; ROUSEEUW, 1990).

3.3.1 Definição do Número de Partições

Uma das questões mais importantes na análise de agrupamento é a da seleção do número k de partições utilizados na análise. No contexto da análise do recurso solar utilizando técnicas de agrupamento, alguns estudos escolhem o número ótimo de partições por tentativas sucessivas (POLO et al., 2015), enquanto outros utilizam critérios matemáticos (ZAGOURAS et al., 2013; ZAGOURAS; INMAN; COIMBRA, 2014; ZAGOURAS; PEDRO; COIMBRA, 2014; GUTIÉRREZ et al., 2017). Entretanto, as estratégias utilizadas nestes artigos consistem em combinações não convencionais feitas utilizando índices de validação interna do particionamento e o método do cotovelo (ou método-L), que pode não apresentar muito significado prático, como discutido a seguir.

Índices de validação interna são índices com valores que tentam definir a qualidade de um agrupamento sem requerer informação externa sobre os dados. Na formulação da maioria destes índices, geralmente são consideradas duas características dos conjuntos. A primeira é relativa a quão similares são os elementos em um mesmo grupo e a segunda, a quão diferentes são os grupos entre si. Essas definições variam para cada índice, mas essencialmente representam essas características (ARBELAITZ et al., 2013). Na maioria dos casos, o analista calcula o índice escolhido para um conjunto de valores interessantes do número partições, e escolhe aquele que maximiza (ou minimiza, dependendo de como é definido o índice) o valor obtido.

Considerando o agrupamento de uma área em diversas regiões para a dinâmica do recurso solar, especialmente baseado em satélites, o uso destes índices pode não ser a melhor estratégia, uma vez que, devido à persistência espacial, elementos nas fronteiras entre partições tendem a ser muito similares. Isso pode fazer com que vários dos índices apresentem comportamentos não usuais, como piorar quase monotonicamente seus resultados com o aumento de k , uma vez que o aumento de partições também aumenta as regiões de fronteira. Em Zagouras, Inman e Coimbra (2014), pode-se observar que os índices de Calinski-Harabasz e da Silhueta, dois índices de validação interna comumente utilizados,

apresentam o comportamento quase monotonicamente decrescente com o aumento de k . Observando estes valores e considerando o comportamento dos dados utilizados, é bastante plausível considerar que este tipo de conjunto de dados não seja intrinsecamente agrupável, o que significa que não há uma estrutura de grupos nos dados. Assim, a análise de partições deve ser usada neste tipo de trabalho não para identificar grupos bem separados e definidos nos dados (o que é buscado com a utilização de índices de validação interna), mas sim para reduzir o número de elementos a serem analisados mas mantendo informações o suficiente dos dados originais.

Devido a este comportamento em relação aos índices de validação interna, em estudos analisados, estes índices foram usados junto ao método do cotovelo (ZAGOURAS; INMAN; COIMBRA, 2014; GUTIÉRREZ et al., 2017). Este método consiste em encontrar um ponto de cotovelo em um gráfico de um índice em função do número de partições, k . Usualmente, é utilizado com índices como a soma de quadrados internos ou a percentagem da variância explicada. Desta maneira, encontra-se um ponto para o qual o aumento do número de partições não aumenta o poder explicatório do particionamento de maneira significativa, fazendo-se não valer a pena aumentar a complexidade da análise que ocorre com o aumento de k . Uma questão que se coloca em relação à utilização de tal método é o fato de que há múltiplas maneiras de se definir este ponto de cotovelo, o que o faz bastante subjetivo e assim, faz com que seja possível determinar um valor de k de acordo com o interesse do analista, com uma justificativa aparentemente técnica, mas que pode ser manipulada. Uma questão identificada é que, ao utilizar o método do cotovelo e índices de validação interna como o da Silhueta, o que se verifica é que o ponto de cotovelo não é um ponto onde pouca informação adicional pode ser obtida com o aumento de k , como se dá nos casos usuais. Um leitor pode verificar, em Zagouras, Inman e Coimbra (2014), Gutiérrez et al. (2017) que o ponto de cotovelo nestes casos é um ponto em que não é possível **piorar** o agrupamento com o aumento do k . Assim, não faz muito sentido a justificativa para a escolha desse valor.

Levando em conta o objetivo deste trabalho, que é o de manter informações o suficiente do banco de dados, mantendo apenas pontos representativos de grupos de localizações, obter clusteres bem definidos e distintos de outros, não é de máxima importância. Portanto, não são usados índices de validação interna, e sim uma medida relativa aos objetivos desejados. Assim, a seleção de k proposta para esta análise e outras similares é a de se usar o número mínimo de clusters que possuem poder de explicação de no mínimo um valor determinado da variância dos dados. Em comparação com o método de escolha de k baseado na união do método do cotovelo aplicado a índices de validação interna, este método apresenta maior simplicidade de implementação e maior interpretabilidade, dado que possui significado prático muito maior. Também, esta estratégia pode ser facilmente adaptada, podendo-se alterar o valor da variância explicada, ou utilizar outros índices que avaliam a quantidade de informação mantida pelo particionamento. Neste caso, foi escolhido, arbitrariamente, o

valor limiar de variância explicada de 75%. Este valor pode ser considerado não muito alto, se comparado a valores utilizados para outras técnicas de estatística multivariada (como análise de componentes principais), mas foi definido considerando as diferenças entre as técnicas, dado que um grupo de componentes principais possui poder explicativo muito maior do que um grupo de partições. Ainda, *pixels* longe dos medoides são explicados de maneira pior, e, por sua vez, reduzem a variância explicada total. Estes locais, muitas vezes se comportam de maneira similar à partição vizinha, e a observação do comportamento de sua partição junta à da partição vizinha pode ser utilizada para estimar o comportamento destes locais de fronteira. Também, no caso de interesse em algum ponto específico em fronteiras de partições, pode-se explorá-lo por meio de um programa desenvolvido para se executar análises baseadas em pontos de referência, discutida a seguir.

3.3.1.1 Análise Baseada em Pontos de referência

Para complementar a análise baseada em partições foi desenvolvido um programa para realizar análises baseada em pontos de referência, de escolha do usuário. Neste programa, o usuário pode selecionar, em um mapa principal, um *pixel* de interesse, clicando nele. Então, são geradas figuras com mapas de estatísticas de pares (que serão discutidas na seção 3.4) em que um dos elementos do par considerado é o elemento selecionado, enquanto o outro, o do *pixel* do mapa. No capítulo 4 pode-se ver exemplos de figuras obtidas com o programa.

3.4 Estatísticas dos pares

Para avaliar o comportamento da energia solar nos grupos obtidos, Gutiérrez et al. (2017) buscam identificar partições que apresentam complementariedade com outras, por meio de uma análise de correlações. Assim, estima, para cada par de clusters, o coeficiente de correlação de Pearson, definido na equação 1, que é estimado por meio da equação 2.

$$\rho(X_c, Y_c) = \frac{Cov(X_c, Y_c)}{\sqrt{Var(X_c)Var(Y_c)}} \quad (1)$$

$$r(X_c, Y_c) = \frac{\sum_{i=1}^n (x_{c,i} - \bar{x}_c)(y_{c,i} - \bar{y}_c)}{\sqrt{[\sum_{i=1}^n (x_{c,i} - \bar{x}_c)] [\sum_{i=1}^n (y_{c,i} - \bar{y}_c)]}} \quad (2)$$

Nas equações acima, $\rho(X, Y)$ representa a correlação de Pearson entre as variáveis genéricas X e Y , $COV(X, Y)$ a covariância entre X e Y e $VAR(X)$ a variância da variável genérica, X . $r(X, Y)$ representa a correlação de Pearson amostral entre X e Y e \bar{x} a média amostral da variável genérica X . O subscrito c representa a componente da variável analisada, assim, X_m e X_a seriam as variáveis de resíduo mensal e interanual obtidas na decomposição da variável X .

A utilização do coeficiente de correlação de *pearson* é interessante, uma vez que ele é adimensional e apresenta informações sobre se desvios das médias em pares de locais tendem a ocorrer de maneira similar ou distinta (GUTIÉRREZ et al., 2017). Entretanto, este índice não mostra, por exemplo, quão grandes são estes desvios, informação que pode ser útil no planejamento, dado que normaliza a covariância. Assim, neste trabalho, também é explorada a dinâmica da covariância sem ser normalizada.

Ainda, estes índices podem falhar em ter um significado prático e de expressar características dos pares que podem ser úteis ao planejamento, como qual seria a variabilidade total se houvesse sistemas instalados nesses pares. Levando isso em conta, são propostos dois índices para avaliação dos pares, baseados no coeficiente de variação (CV). Deve-se ressaltar que, neste trabalho, o CV é definido de maneira um pouco diferente da situação padrão. Isto é feito uma vez que as variáveis analisadas possuem média 0, já que são componentes da série temporal em que a média foi extraída. Assim, no denominador da fórmula do coeficiente de variação, a média considerada não é a média da componente em questão, mas sim, a média a longo prazo da irradiância que ao ser decomposta, gera a componente analisada. A equação 3 mostra a fórmula de obtenção do CV considerada, em que $\widehat{CV}(X_c)$ representa o coeficiente de variação estimado da componente c da variável X , $\hat{\sigma}(X_c)$, o desvio padrão estimado para esta componente e X_{lp} , a média de longo prazo desta variável antes da decomposição.

$$\widehat{CV}(X_c) = \frac{\hat{\sigma}(X_c)}{X_{lp}} \quad (3)$$

Com base nisso, é definido o PCV (*Pair Coeficient of Variation*, ou, em português, Coeficiente de Variação de Par) que representa qual o coeficiente de variação da soma das séries da componente analisada das duas localidades, definido na equação 4. Este índice, sendo um coeficiente de variação, é, assim como a correlação, adimensional, e pode ser usado para análise de outros recursos sem perder o poder de comparação.

$$\widehat{PCV}(X_c, Y_c) = \widehat{CV}(X_c + Y_c) = \frac{\hat{\sigma}(X_c + Y_c)}{X_{lp} + Y_{lp}} \quad (4)$$

Este índice, então, contabiliza questões relacionadas à variabilidade da irradiação para pares de locais, levando em conta a variação em cada ponto integrante de um dado par e a complementariedade entre os pontos do par. Ele pode sugerir, por exemplo, um par que, quando considerada a produção nos dois locais apresenta baixa variação. Essa informação pode ser mais interessante do que a indicação de correlação negativa, já que possui um significado mais prático. Uma outra versão deste índice também é proposta, o OPCV (*Optimal Pairwise Coeficient of Variation*, ou, em português, Coeficiente de Variação de Par Ótima), definido na equação 5, que representa o menor coeficiente de correlação que a série da média ponderada das séries dos elementos de um certo par pode apresentar,

considerando todas as ponderações possíveis. Assim, ele contém informações que podem ser mais úteis do que a do PCV simples (que seria equivalente ao caso de uma média simples, ou com ponderações iguais).

$$\widehat{OPCV}(X_c, Y_c) = \min_{a \in (0,1)} \widehat{CV}(aX_c + (1-a)Y_c) = \min_{a \in (0,1)} \frac{\hat{\sigma}(aX_c + (1-a)Y_c)}{aX_{lp} + (1-a)Y_{lp}} \quad (5)$$

Considerando a irradiação como um bom estimador da produção energética, este índice representa o mínimo coeficiente de variação que um portfólio que contenha sistemas apenas nos dois locais do par pode apresentar. Assim, também é interessante obter o valor da irradiação média a longo prazo que este portfólio apresentaria, visando comparar os portfólios de pares no quesito de variação total e de irradiação média, relacionada aos custos da energia. Este valor é o denominador da função a ser otimizada, quando o a é o valor que otimiza a função. Assim, pode ser definido como $a_{opt}X_{lp} + (1 - a_{opt})Y_{lp}$.

Uma vez verificada graficamente que a função a ser minimizada é unimodal, a minimização é feita utilizando a função `optim()` presente no R, que se baseia na combinação do método da seção áurea com sucessivas interpolações parabólicas, apresentado em Brent (1973).

3.5 Análise de Componentes Principais

As componentes sazonal e de tendência podem ser analisadas visualmente para cada local. A componente sazonal pode ser caracterizada por um padrão de 12 meses, que se repete a cada ano, enquanto a componente de tendência apresenta poucas oscilações. Levando isso em conta, é proposta a utilização de PCA (*principal component analysis*, ou análise de componentes principais) para auxiliar na análise de tais componentes, uma vez que, por serem visualmente interpretáveis, as PCs (*principal component*, ou componentes principais) obtidas com tal método, também podem ser. A ideia é de se obter para as componentes de tendência e sazonal, alguns poucos padrões temporais (descritos por suas componentes principais), que, ao serem analisados, conseguem explicar a dinâmica de seus comportamentos ao longo da área nacional.

PCA é um método de redução de dimensionalidade bastante conhecido e popular em análise de dados. Baseia-se em transformar variáveis originais do problema em variáveis linearmente independentes, chamadas de componentes, por meio de transformação ortogonal. A transformação é feita de maneira em que a primeira componente explica o máximo possível da variância dos dados que uma componente poderia explicar. As componentes em sequência também explicam o máximo de variância possível dos dados, sendo sempre ortogonais a todas as componentes anteriores.

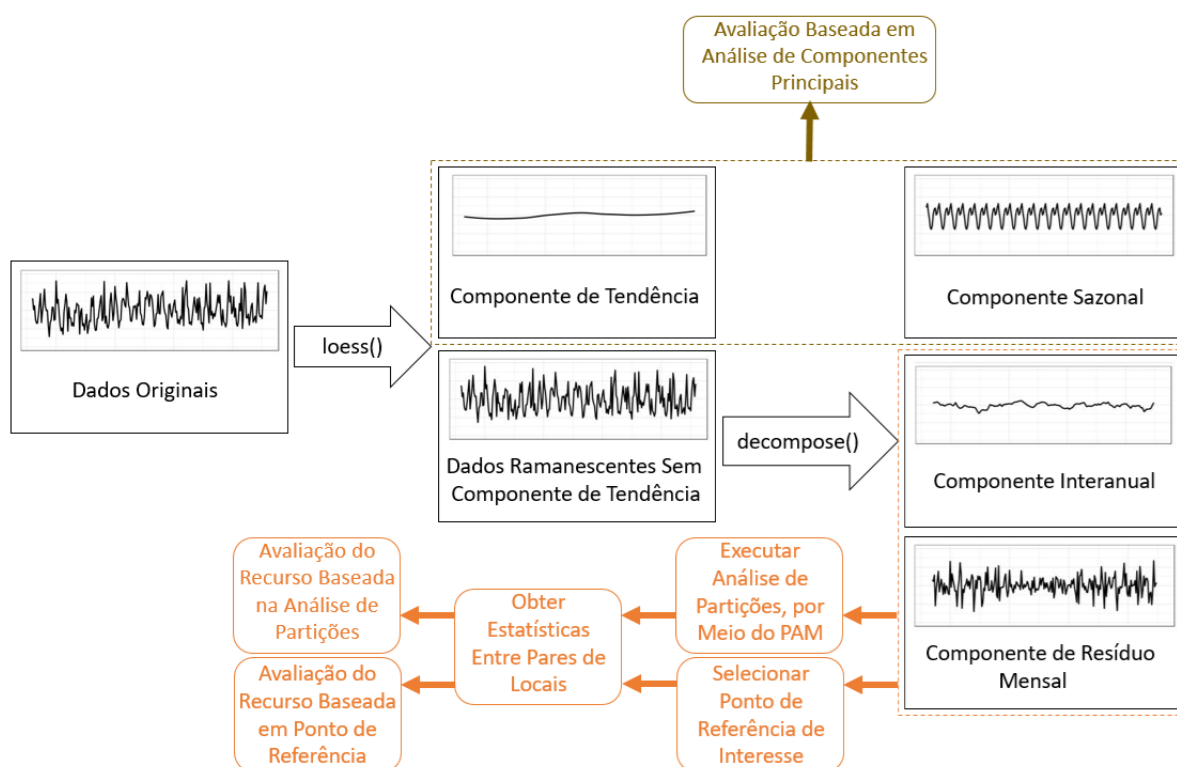
O objetivo deste método é manter o máximo possível da variabilidade dos dados mantendo o mínimo necessário de componentes. Desta maneira, para compreender o comportamento de um elemento, ao invés de ter que se observar cada variável, pode-se observar apenas as componentes principais, que conseguem expor as maiores características destas variáveis. Uma análise de componentes principais desejável produz PCs com significados interpretáveis e práticos. Assim, essas PCs obtidas permitem o entendimento do comportamento das variáveis. Há ampla bibliografia sobre o assunto. O leitor interessado pode se referir a Hotelling (1933), Wall, Rechtsteiner e Rocha (2003), Mingoti (2005).

Neste trabalho, PCA é utilizada separadamente para as séries temporais sazonal e de tendência. Em ambos os casos, as variáveis são as séries históricas destas componentes (no caso da sazonal, de apenas um ano), e os elementos são cada *pixel*. As componentes obtidas devem apresentar padrões temporais típicos destas séries. Mapas são gerados mostrando quanto cada componente é presente em cada *pixel*. São avaliadas as componentes principais que juntas explicam no mínimo 90% da variância dos dados. Os cálculos dos autovalores e a obtenção das componentes principais são feitos no R via a função `prcomp()`.

3.6 Resumo

Nesta seção, a metodologia desenvolvida é sintetizada. Os dados utilizados são extraídos no site do sistema POWER da NASA, por meio de um programa desenvolvido pelo autor em R. Também, são utilizados dados do Projeto SWERA, das estações consideradas para geração dos anos típicos meteorológicos neste projeto, que são lidos e interpretados também no R. Os dados são integralizados em escala mensal e as séries com esta escala, para todas as localidades, são decompostas em componentes de tendência, interanual, sazonal e de resíduo mensal. A decomposição é feita utilizando a função `loess()`, que aplica um filtro *loess* para obter a componente de tendência, e a função `decompose()` para obter as outras componentes utilizando a decomposição baseada em médias móveis. As componentes interanual e de resíduo mensal são avaliadas por meio de análise de partições e análises baseadas em pontos de referência. Estas análises são feitas verificando o comportamento de pares de localidades, por meio de estatísticas descritas na seção 3.4. As componentes de tendência e sazonal são avaliadas visualmente utilizando de análise de componentes principais para tal. A figura 8 apresenta um esquemático da metodologia de análise utilizada.

Figura 8 – Esquemático da Metodologia Proposta



Componente de tendência obtida com `loess()`. A série remanescente é utilizada para obter as outras componentes, com `decompose()`.

As cores são relativas a grupos de componentes que são analisados com métodos distintos

Elaboração própria

4 RESULTADOS E DISCUSSÕES

Neste capítulo, são expostos resultados obtidos com a metodologia apresentada anteriormente. Também são feitas discussões baseadas nestes.

4.1 Decomposição das Séries Temporais

A metodologia para obtenção das componentes das séries temporais apresentada na seção 3.2 foi utilizada, de modo que para cada pixel as quatro componentes de séries temporais (isto é: de tendência, interanual, de sazonalidade e de resíduo mensal) foram obtidas.

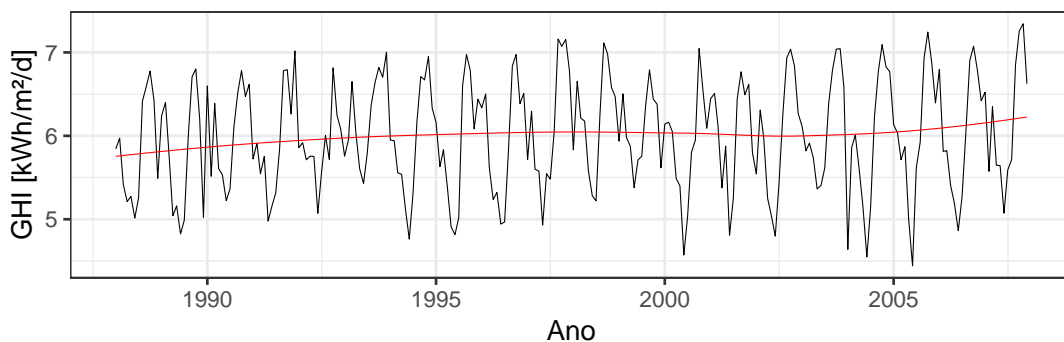
Para exemplificar este processo, a figura 9 demonstra a obtenção das séries de resolução mensal para a localidade centrada na Latitude -6.5° e Longitude $-37,5^\circ$. Esta localidade foi escolhida por ter apresentado maior nível de irradiação diária média no banco de dados utilizado, de $6,00 \text{ kWh/m}^2$. Este *pixel* contém a cidade de Patos, na Paraíba, e é referido com o nome da cidade.

Na figura 9a, é verificável que a componente de tendência, para a localidade em questão, identificou um comportamento de crescimento da medida da irradiação diária média ao longo do tempo, com certa oscilação. A Figura 9b mostra que a componente interanual apresenta pouca variabilidade para essa localidade, o que também é verificado na figura 9d para a componente de resíduo mensal. De fato, se tratando de uma localização com altos índices de irradiação, é de se esperar mais baixas variabilidades devido ao clima mais seco. Finalmente, é visível na figura 9c, um padrão sazonal bem definido ao longo do tempo, com um vale em Maio e um pico em Novembro, ressaltado na figura 11a.

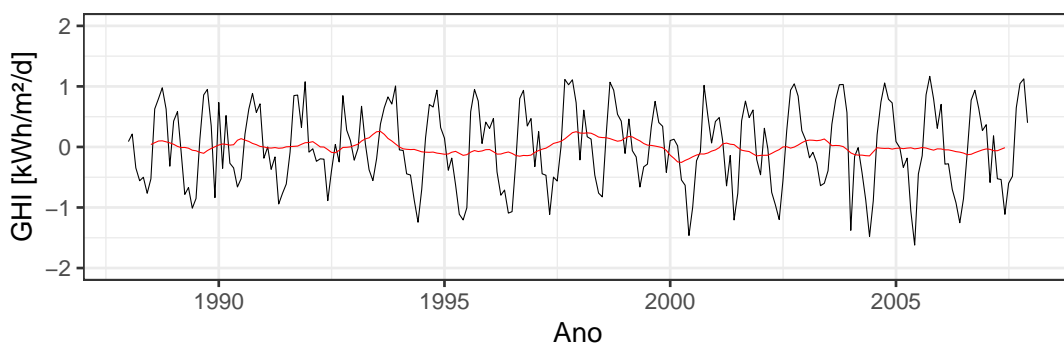
A Figura 10 mostra as séries temporais obtidas para o pixel contendo a cidade de Belo Horizonte, que foi escolhida devido à sua localidade central e por ser a cidade em que se encontra a Escola de Engenharia da Universidade Federal de Minas Gerais. Esta cidade também é uma das dez cidades com mais potência instalada de GDFV no Brasil. Comparando com a localidade anterior, pode-se ver que esta localidade também apresenta uma componente de tendência predominantemente crescente, especialmente entre 1992 e 1997 (figura 10a). Comparando as componentes de oscilações aleatórias, percebe-se que em Belo Horizonte, elas apresentam maiores variabilidades (figuras 10b e 10c), enquanto a componente sazonal apresenta menor amplitude e uma dinâmica consideravelmente distintas, como dois picos, em Fevereiro e em Outubro (figura 11b).

Figura 9 – Obtenção das Componentes de Série Temporal para Patos - PB

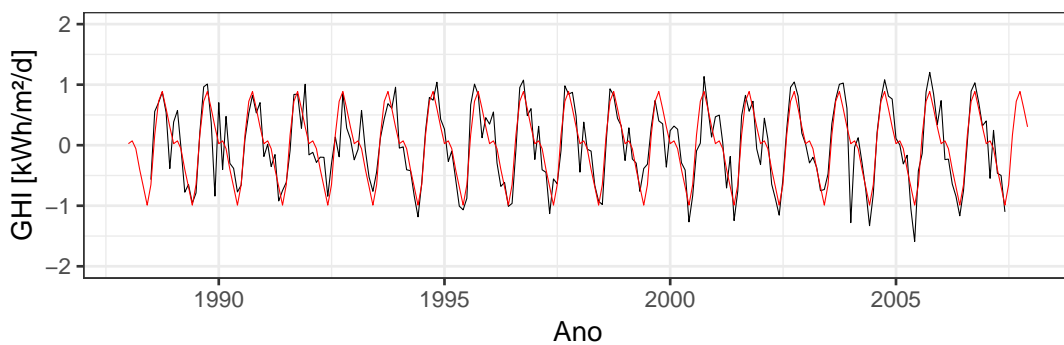
(a) Série Histórica (Preto) e Componente de Tendência (Vermelho)



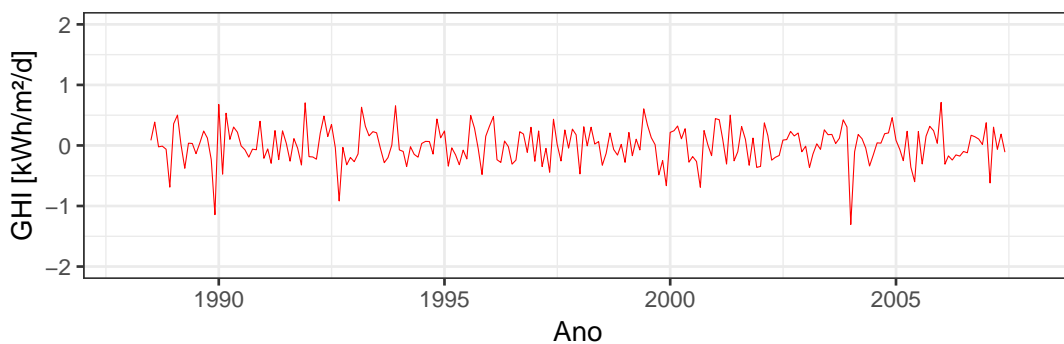
(b) Série Histórica com Componente de Tendência Retirada (Preto) e Componente Interanual (Vermelho)



(c) Série Histórica Mensal (Preto) e Componente Sazonal (Vermelho)



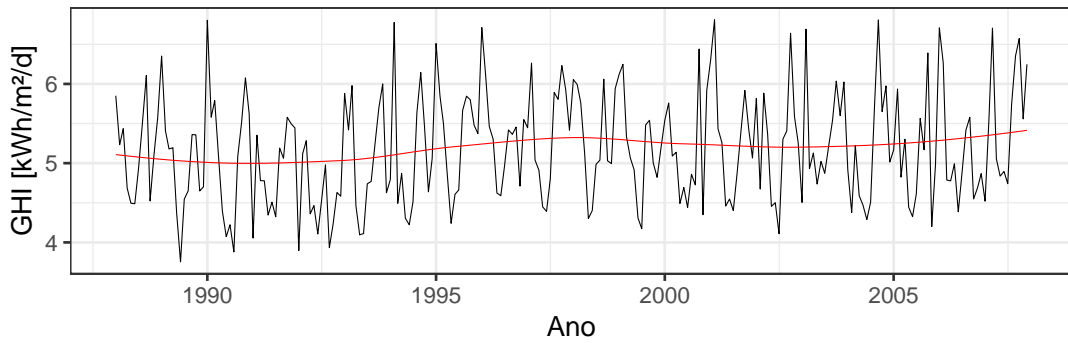
(d) Série Histórica do Resíduo Mensal



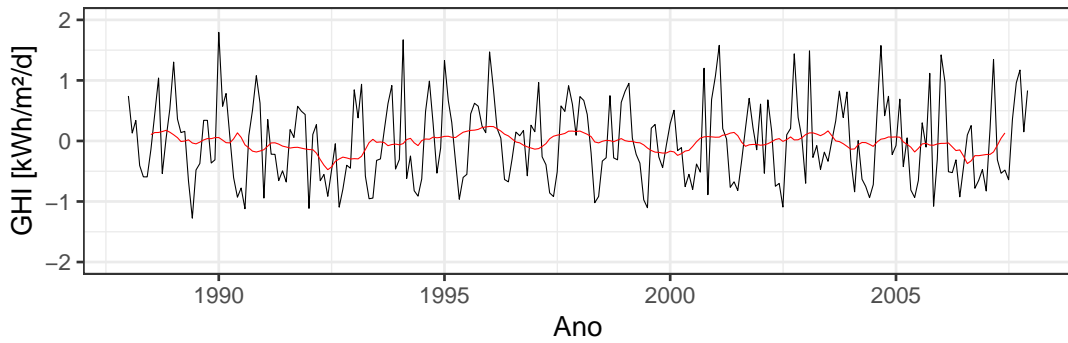
Elaboração própria

Figura 10 – Obtenção das Componentes de Série Temporal para Belo Horizonte - MG

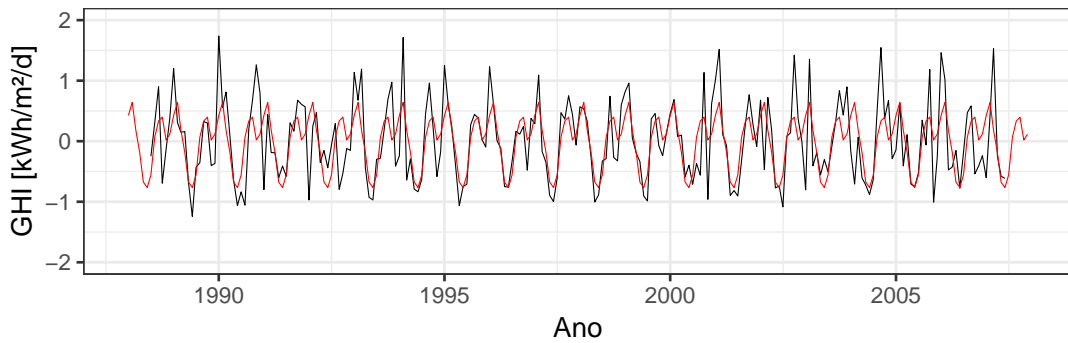
(a) Série Histórica (Preto) e Componente de Tendência (Vermelho)



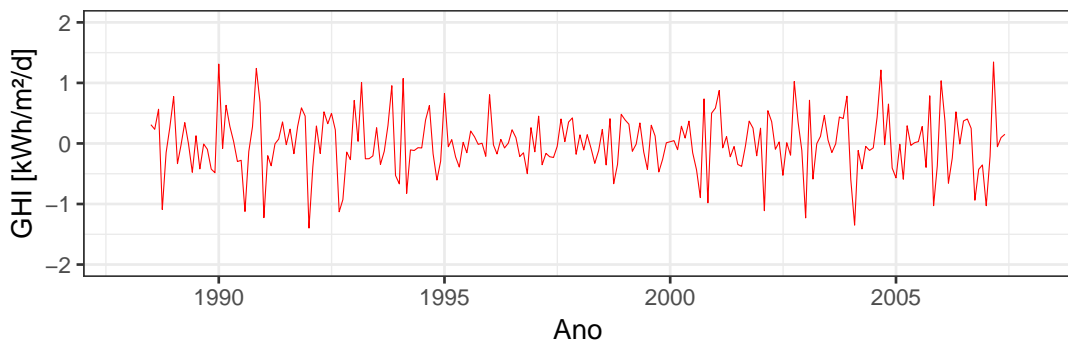
(b) Série Histórica com Componente de Tendência Retirada (Preto) e Componente Interanual (Vermelho)



(c) Série Histórica Mensal (Preto) e Componente Sazonal (Vermelho)

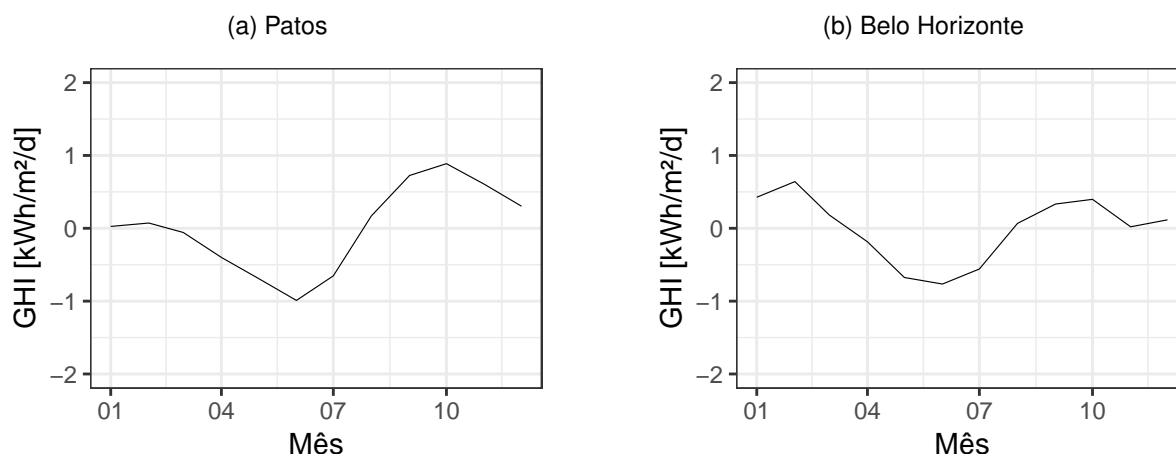


(d) Série Histórica do Resíduo Mensal



Elaboração própria

Figura 11 – Componente Sazonal para Um Ano



Elaboração Própria

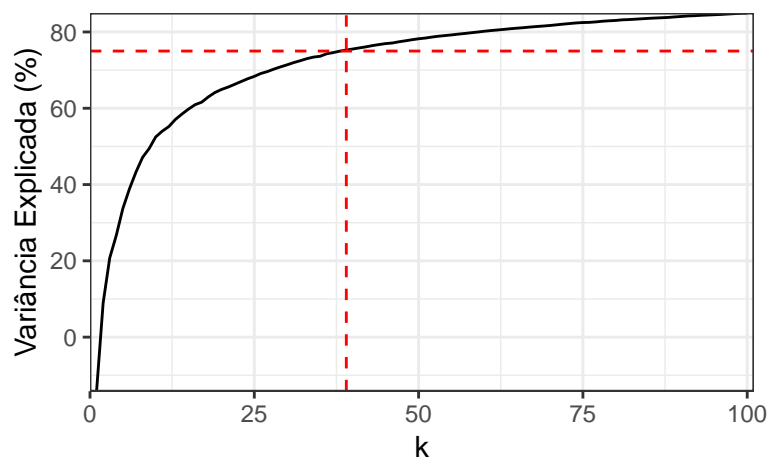
4.2 Análise das Componentes de Oscilações Aleatórias

As componentes de oscilações aleatórias, isto é, componentes interanual e de resíduo mensal, foram avaliadas considerando pontos representativos do banco de dados, para identificar pares de regiões que juntas apresentam baixa variabilidade do recurso e alto potencial de geração. Para obter os pontos representativos, foi utilizada a metodologia descrita na seção 3.3. O algoritmo PAM foi usado de forma em que cada elemento do banco de dados correspondia a um *pixel*, e continha, como variável, a série temporal correspondente à soma das componentes interanual e de resíduo mensal. É interessante ressaltar que o particionamento poderia ter sido executado separadamente para as duas componentes, de modo a obter conjuntos de partições diferentes para cada caso, mas foi feito considerando-as juntas visando uma maior coerência e interpretabilidade do resultado em contexto da aplicação. O algoritmo foi rodado para k variando entre 1 e 100. A figura 12 mostra os valores da variância explicada para cada valor de k considerado. Pode ser visto que, para alcançar o valor desejado de 75%, são necessárias 39 partições. A figura 13 mostra as partições obtidas e os seus medoides, que são os pontos representativos para a análise. No Anexo A, há uma versão maior do mapa, que pode ser utilizado para acompanhar este capítulo.

4.2.1 Coeficientes Utilizados para Avaliação de Pares

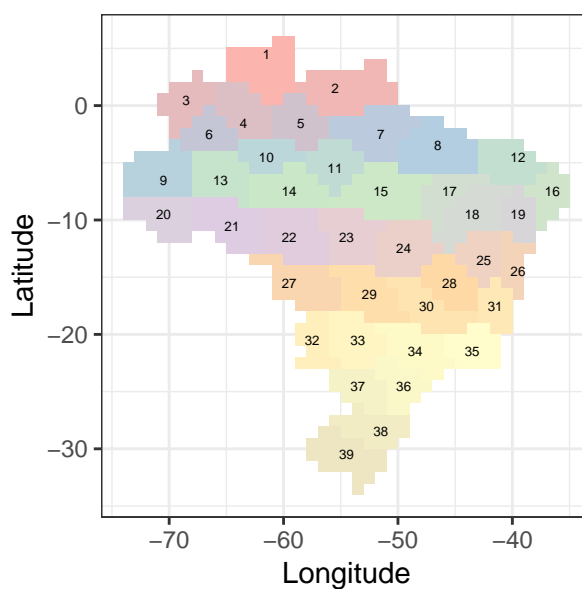
Foram calculados, para cada par de medoides obtido, os coeficientes para avaliação de pares apresentados na seção 3.4, nominalmente, o coeficiente de correlação de Pearson, a covariância e os coeficientes de variação de pares simples (PCV) e ótimos (OPCV). Para compará-los, a figura 14 mostra os gráficos de dispersão para cada possível par de índices calculados e as respectivas correlações, considerando a componente de resíduo mensal. Também, mostra uma curva de densidade de probabilidade para cada índice. Pode ser

Figura 12 – Variância Explicada pelo Número k de partições



Elaboração própria

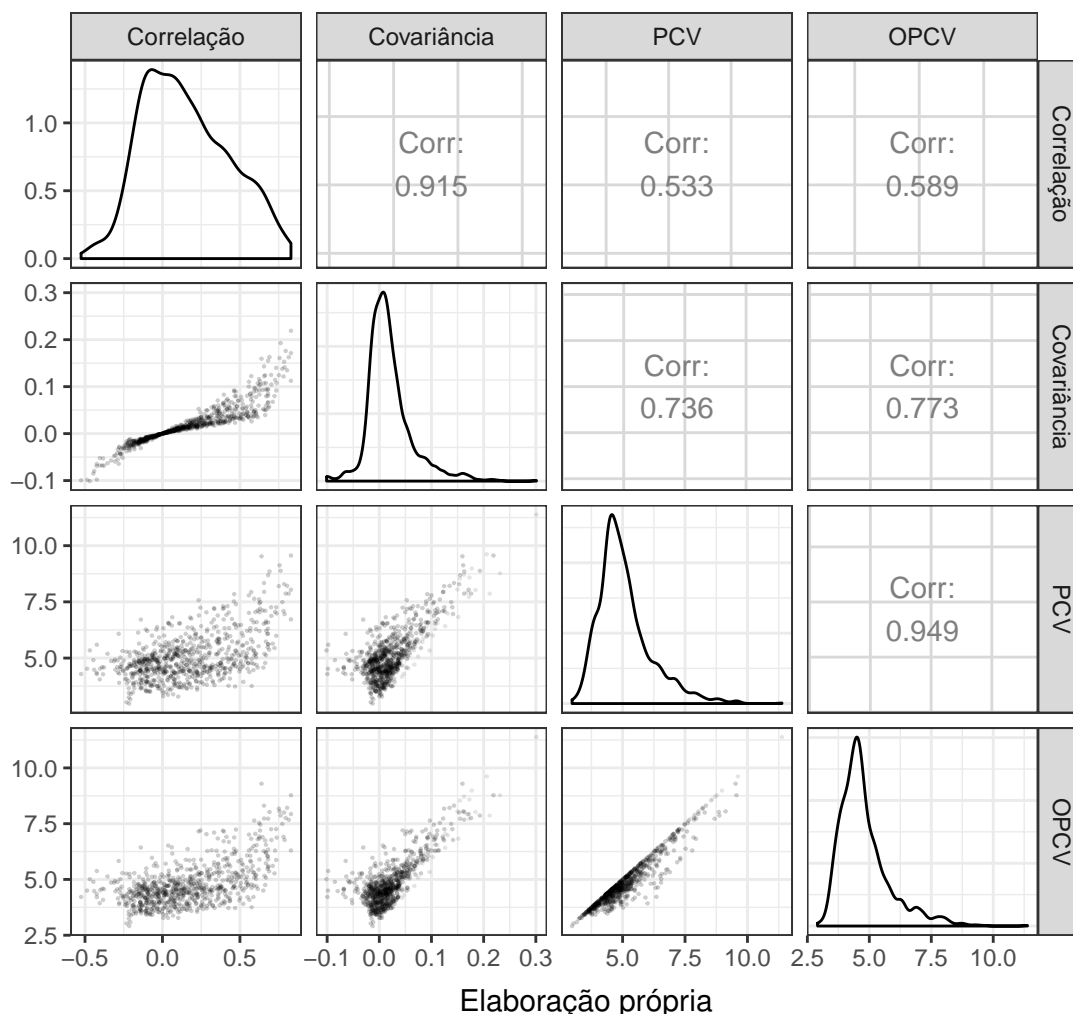
Figura 13 – Partições Obtidas



Cada cor representa uma partição. Partições são numeradas de norte a sul.

Elaboração própria

Figura 14 – Comparação do Comportamento dos Diversos índices para a Avaliação de Pares: Componente de Resíduo Mensal

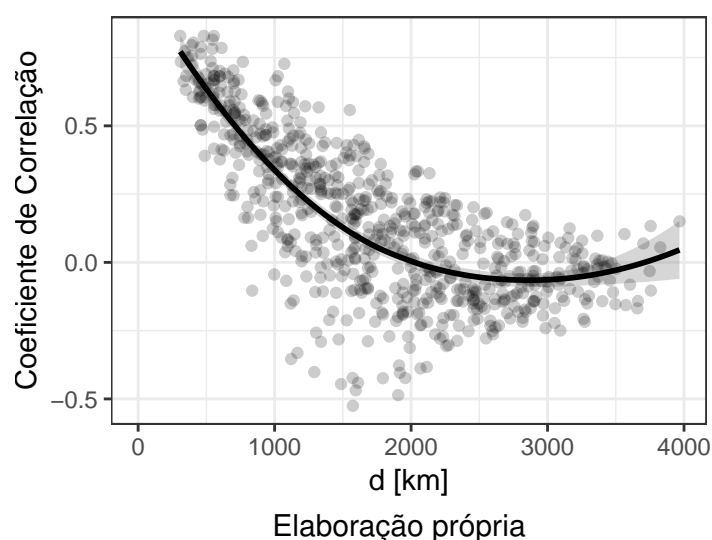


observado um comportamento similar entre a correlação e a covariância, especialmente para valores próximos de 0. Isso faz sentido, dado que a correlação de Pearson é obtida por meio da multiplicação da covariância por um fator normalizador. Os coeficientes de variação de pares, por sua vez, apresentaram comportamentos consideravelmente diferentes e correlações mais baixas com os outros índices, especialmente, com a correlação. Essas diferenças sugerem que a análise utilizando tais índices pode chegar a resultados diferentes do que o obtido utilizando somente a correlação. Estes índices são similares entre si para a maior parte dos pares e exceções a esta similaridade ocorrem principalmente para pares em que um dos elementos do par possui coeficiente de variação muito maior do que o outro. É interessante observar que o valor do OPCV é sempre menor ou igual ao do PCV, uma vez que o OPCV representa a solução ótima enquanto o PCV é uma das soluções de uma combinação arbitrária do espaço de busca. Neste trabalho são utilizados nas análises posteriores os valores obtidos para as correlações de Pearson, devido ao seu uso em outros estudos e para os OPCV, devido ao seu significado prático.

4.2.2 Análise da Componente de Resíduo Mensal

Nesta seção, a componente de resíduo mensal é avaliada. Inicialmente, a relação entre as distâncias entre os pontos representativos e as correlações dos pares é investigada. A figura 15 mostra o gráfico destas duas medidas, junto com a tendência média para este gráfico, estimada com um *loess*. Pode ser observado que os pontos são consideravelmente dispersos, o que sugere que a utilização da distância como estimador para a correlação desta componente pode não capturar todos os resultados de interesse. Como mencionado no Capítulo 1, um resultado similar pode ser verificado em Perez e Fthenakis (2015), em que são avaliadas correlações de variações do índice de claridade com resolução de até 30 dias.

Figura 15 – Correlação Versus Distância Considerando a Componente de Resíduo Mensal

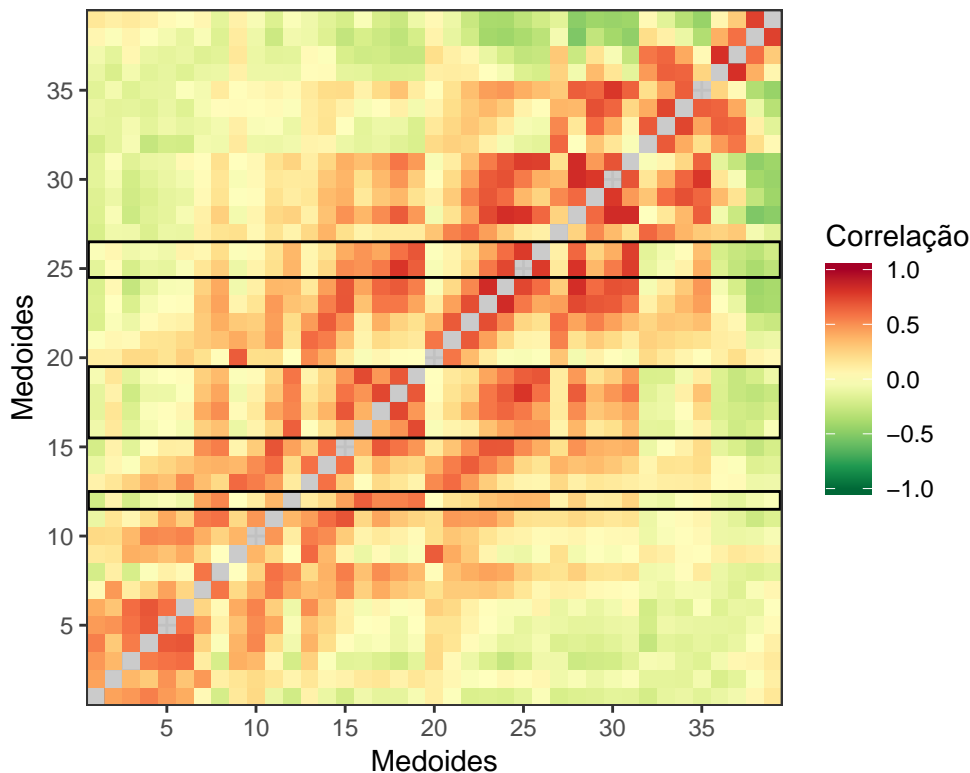


4.2.2.1 Análise Baseada nas Correlações

Para avaliar as correlações entre os pares de localidades, e compreender melhor as dinâmicas do recurso solar no território brasileiro, são utilizadas duas maneiras de visualização destas correlações. A primeira, similar às usadas em artigos como Zagouras, Inman e Coimbra (2014), Gutiérrez et al. (2017), consiste em apresentar os valores do coeficiente de correlação para cada par por meio de um gráfico de calor (figura 16). Este tipo de gráfico pode ser de difícil interpretação. Considerando isto, é proposto o seguinte procedimento para sua avaliação: 1. Escolher uma região de interesse; 2. Encontrar na figura 13 ou no apêndice A os medoides relativos à região escolhida; 3. Observar as linhas relativas a esses medoides, identificando, nas colunas, localidades que apresentem maiores complementariedades com a de interesse, verificando valores de correlação mais negativos, ou, se positivos, mais próximos de zero.

A região de interesse que é avaliada com maior foco neste trabalho é a nordeste, devido a seu alto potencial de geração. Na figura 16, linhas referentes aos *pixels* 12, 16-19,

Figura 16 – Gráfico de Calor das Correlações para a Componente de Resíduo Mensal



Eixos representam os medoides, que podem ser identificados na figura 13 ou no apêndice A. Retângulos ressaltam linhas relativas a medoides no Nordeste, aos quais é dado um foco maior na discussão.

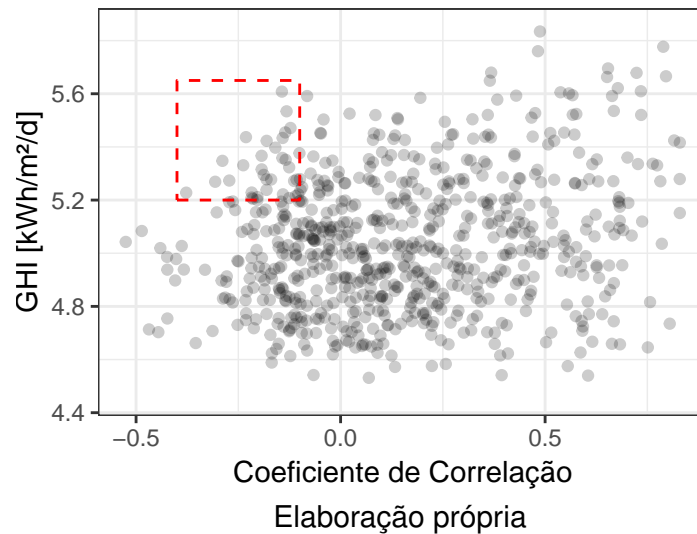
Elaboração própria

25-26, relativos a esta região, são demarcadas com retângulos. É verificável que correlações baixas ou negativas em relação a tal região ocorrem, principalmente, no extremo norte do país (medoides 1, 3-6, 9), na região próxima ao Mato Grosso do Sul (32,33) e na região sul do país (36-39).

As limitações do uso desta figura, devido à dificuldade de interpretação e ao fato de conter apenas dados de correlação, não expressando outras variáveis que podem ser úteis, como o potencial solar, fazem com que também sejam interessantes outras maneiras de explorar os dados. Visualizando valores de irradiação juntamente com informações de correlações, pode-se obter melhor entendimento do comportamento do recurso. Considerando esta questão, é proposto o uso de um gráfico da média dos valores de irradiação direta horizontal (GHI) a longo prazo entre os dois elementos de cada par versus a respectiva correlação (figura 17).

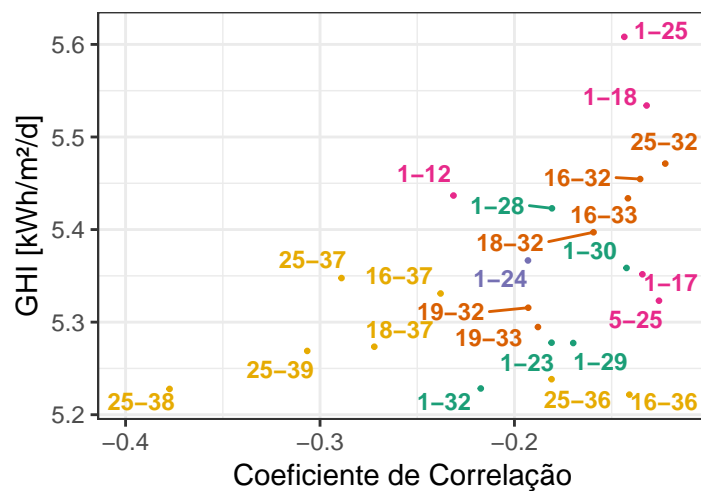
Utilizando ferramentas, como a de ampliação, para explorar os dados da figura 17, é possível a identificação de pares de interesse, com elevados valores de irradiação, e que apresentam complementariedade. A região no gráfico, visualmente identificada, com pares

Figura 17 – Irradiância Diária Média versus Coeficiente de Correlação Considerando a Componente de Resíduo Mensal



que apresentam maiores valores de irradiação e correlações negativas é demarcada com um retângulo e plotada separadamente na figura 18, onde cada ponto é rotulado com os números dos medoides que fazem parte de seu par relativo e as cores representam em quais subsistemas estão cada ponto, para auxiliar na interpretação.

Figura 18 – Área de Interesse do Gráfico da Figura 17

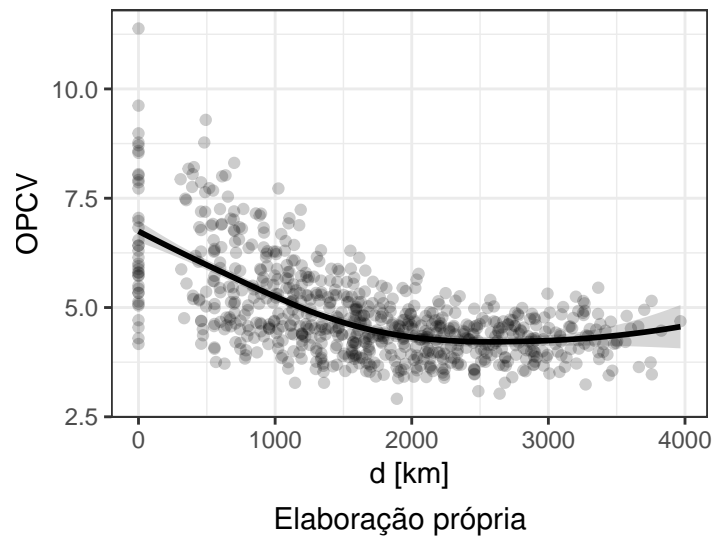


Rubricas mostram os medoides dos pares, que podem ser identificados no apêndice A
 Cores são relativas aos subsistemas: N-N=Roxo; N-SE/CO=Azul Esverdeado; N-NE=Rosa;
 NE-S=Amarelo; NE-SE/CO=Laranja

Elaboração própria

As figuras 18 e 17 indicam que coeficientes de correlação negativos associados a altos índices de irradiação ocorrem principalmente em pares em que um elemento encontra-se no subsistema Nordeste, enquanto tem-se o outro elemento no subsistema Sul (pares em

Figura 19 – OPCV versus Distância Considerando a Componente de Resíduo Mensal



amarelo), no sudoeste do subsistema SE-CO (pares em laranja) e na parte do subsistema Norte próximos a Roraima (pares em rosa). Note que estes pares coincidem com os já analisados na figura 16. A região próxima a Roraima também apresenta bons resultados quando associada a elementos do SE-CO (pares em azul esverdeado), e a um elemento da região Norte próximo à região Nordeste (par em lilás). Pode-se verificar nesta área destacada que a região Sul apresenta as correlações mais negativas, mas pares com menor potencial. Vale ressaltar que mesmo esse menor potencial ainda é significativo, especialmente quando comparados com os valores na Europa.

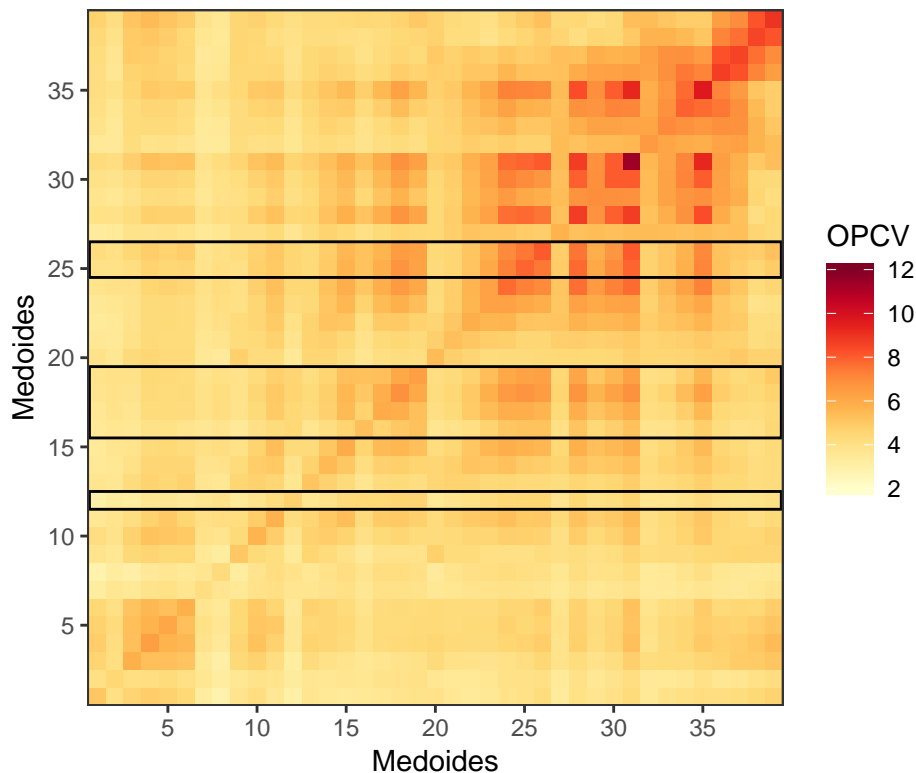
4.2.2.2 Análise Baseada no OPCV

Como já discutido, o índice de correlação apresenta pouco significado prático, e uma análise baseada apenas nele é incompleta, dado que não avalia algumas questões importantes. Assim, uma análise similar à feita com este coeficiente também é realizada com o proposto OPCV. Primeiramente é avaliado o gráfico do OPCV pela distância (figura 19). É interessante observar que para os dados analisados, o valor máximo do OPCV apresenta uma forte relação com a distância até valores de 2000km . Entretanto, o valor mínimo apresenta menor dependência da distância.

A figura 20 mostra o mapa de calor para este índice. Novamente, é dado um foco maior a linhas relativas a medoides no Nordeste (12; 16-19; 25-26), que são destacadas com retângulos, devido ao interesse na região, relacionado ao mais elevado potencial solar. Neste caso, semelhanças são verificadas quando é feita uma comparação com a figura 16, como a complementariedade entre Nordeste e as regiões Norte, Sul e do Mato Grosso do Sul.

Também, é interessante observar na figura 16 que alguns medoides apresentam valores baixos quando associados a quase todos os outros elementos. Verificando a figura

Figura 20 – Gráfico de Calor dos OPCVs para a Componente de Resíduo Mensal



Eixos representam os medoides, que podem ser identificados na figura 13 ou no apêndice A. Retângulos ressaltam linhas relativas a medoides no Nordeste, aos quais é dado um foco maior na discussão.

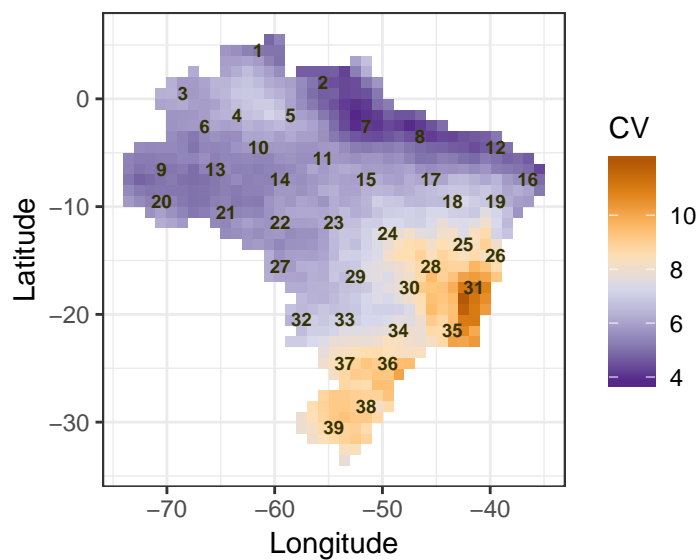
Elaboração própria

21, que mostra os valores dos coeficientes de variação mensal para o território nacional, pode-se verificar que tais colunas são relativas a elementos com coeficiente de variação mensal mais baixos, como o 7, o 8 e o 12.

Para melhor visualização, a figura 22 mostra os valores do OPCV mensal em função do GHI da combinação ótima que minimiza o CV mensal do par. As regiões da figura com maiores valores de GHI e menores valores de OPCV, demarcadas com retângulos neste gráfico, são destacadas na figura 23. Nesta figura, pode-se observar que pares compostos por elementos nas regiões N-NE (Rosa), NE-NE (Verde) e NE-SE/CE (Laranja) são os mais destacados.

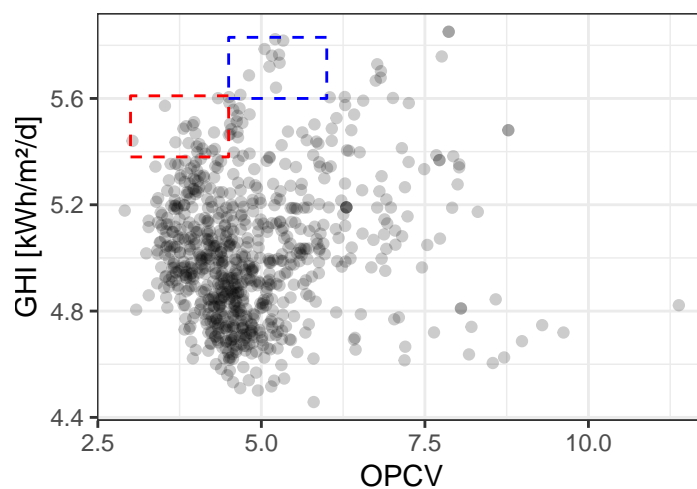
É interessante comparar estes resultados com os obtidos utilizando os coeficientes de correlação. Inicialmente pode-se observar que pares NE-S (Amarelo) ficam em evidência muito menor quando se utiliza o OPCV. Note que quando avaliados os coeficientes de correlação, pares NE-S apresentavam os menores valores, mas que estes não corresponderam com valores mínimos de OPCV. Essas diferenças nos resultados podem ser explicadas verificando a figura 21, já que altos valores do CV mensal para a região Sul afetam os CVs dos

Figura 21 – Coeficientes de Variabilidade da Componente de Resíduo Mensal Para o Brasil, com Identificação dos Medoides



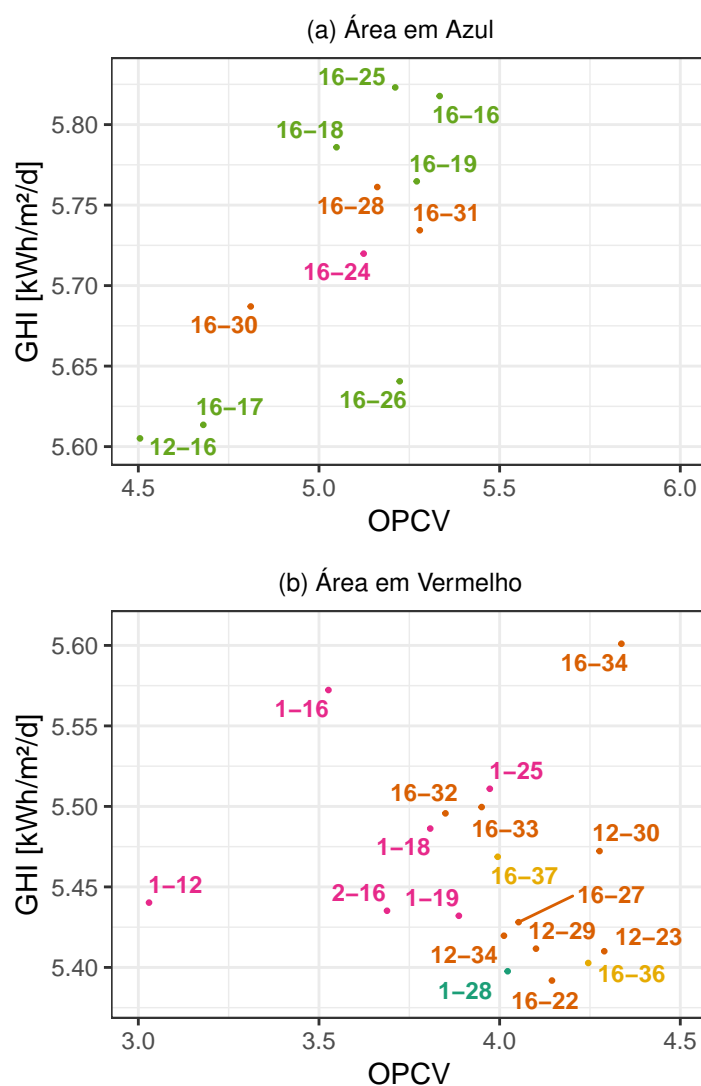
Elaboração própria

Figura 22 – Irradiância Diária Média versus OPCV Considerando a Componente de Resíduo Mensal



Elaboração própria

Figura 23 – Áreas de Interesse do Gráfico da Figura 22



Rubricas mostram os medoides dos pares, que podem ser identificados no apêndice A

Cores são relativas aos subsistemas: N-SE/CO=Azul Esverdeado; N-NE=Rosa; NE-NE=Verde; NE-S=Amarelo; NE-SE/CO=Laranja.

Elaboração própria

pares com elementos nela. Também, pode-se observar que o ponto com maior destaque do subsistema Nordeste, passou a ser o ponto 16, ao invés do 25, verificado na análise baseada em correlação. É possível observar, que o primeiro apresenta CV consideravelmente inferior, o que afeta os valores de seus OPCV.

A utilização do OPCV trás, também, a possibilidade de diretamente comparar o benefício de se ter mais potencial de geração em um par com a redução da sua variabilidade. Assim, tal índice possibilita o entendimento mais direto de questões como a do aproveitamento de regiões de menor potencial que, embora possuam maiores custos de geração, podem contribuir para reduzir a variabilidade da geração do ponto de vista sistêmico. A utilização da correlação não trata esta redução da variabilidade diretamente.

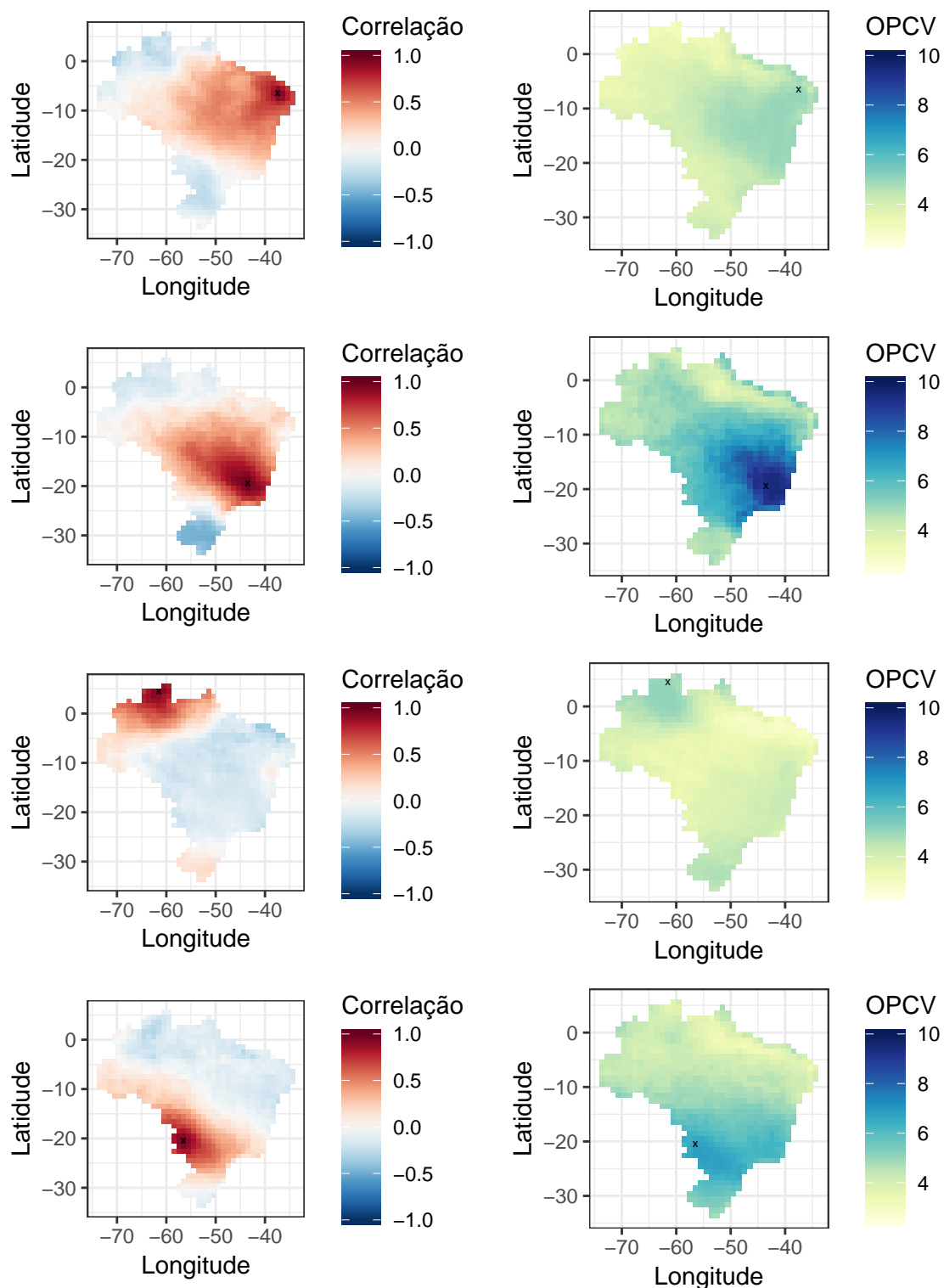
Considerando a tendência de um maior crescimento da utilização da energia solar no Nordeste, especialmente na modalidade centralizada, devido a seus altos valores de irradiância, os resultados da figura 23 mostram que seriam interessantes incentivos da utilização da energia solar em regiões que não apresentam níveis tão elevados de irradiação incidente ou de penetração de FV, como a do Mato Grosso do Sul (medoides 32 e 33), que já possui uma rede interligada mais forte, visando diminuir a variabilidade mensal aleatória do recurso solar. Outras áreas do subsistema SE/CO e a região Sul, que também apresentam complementariedade com o Nordeste, naturalmente já devem apresentar um desenvolvimento da utilização de energia solar na modalidade da geração distribuída. A região de Roraima (medoide 1) também se destaca com alta complementariedade com grande parte do país. Entretanto, este estado ainda não está conectado ao SIN, sendo prevista a conexão em 2027 junto à instalação da usina hidrelétrica de Bem Querer, no estado. Mesmo no caso de conexão, para planejar o desenvolvimento da energia solar neste estado deve-se prestar mais atenção às limitações nas capacidades de escoamento de energia entre este local e os centros de carga.

4.2.2.3 *Análise Baseada em Pontos de Referência*

Para servir de ferramenta complementar, também foi desenvolvido, em R, um programa para explorar os dados. Este programa permite a exploração baseada em pontos de referência, em que o usuário pode clicar em um *pixel* e valores dos índices de complementariedade do pixel selecionado com todos os outros são expostos em mapas. A figura 24 mostra mapas que podem ser facilmente obtidos com a ferramenta e que complementam a análise prévia, considerando os pontos de referência em Patos e Belo Horizonte, além de pontos em Roraima (medoide 1) e no Mato Grosso do Sul (medoide 32).

As análises baseadas nas localidades de interesse mostram, novamente, os bons resultados dos estados MS e RR. O noroeste da região Nordeste também é ressaltado observando os valores de OPCV. Ademais, é verificável que a região Sul apresenta correlações negativas com o ponto do Nordeste analisado e com Belo Horizonte. O uso destas figuras é

Figura 24 – Mapas de OPCVs e de Correlações em Análises Baseada em Pontos de Referência



Correlações na esquerda e OPCV na Direita.

Pontos de referência: de cima para baixo: Patos; Belo Horizonte; Medoide 1 (em Roraima); Medoide 32 (no Mato Grosso do Sul)

Elaboração própria

um bom complemento à análise de partições.

4.2.3 Comportamento ao Longo do Tempo

Nesta seção é explorada a dinâmica ao longo do tempo para um par de local de interesse para a componente de resíduo mensal. A figura 25a mostra o comportamento da correlação calculada para janelas anuais (com 12 dados mensais) para o par de medoide 32, no Mato Grosso do Sul e 16, na Paraíba. Pode-se observar que a dinâmica interanual do comportamento da correlação apresenta um comportamento consideravelmente variável. A figura 25b mostra o comportamento do CV, medido em intervalos de 12 meses, da combinação entre os elementos do par que minimiza o CV a longo prazo (ou seja, a combinação relativa ao OPCV no longo prazo, mas que não necessariamente otimiza o CV para cada ano) calculado anualmente para este par. A figura também mostra o CV dos elementos do par. Pode-se observar que também há uma considerável variabilidade entre os anos para estes valores, que é um pouco menor para o relativo ao CV do par.

4.2.4 Análise da Componente Interanual

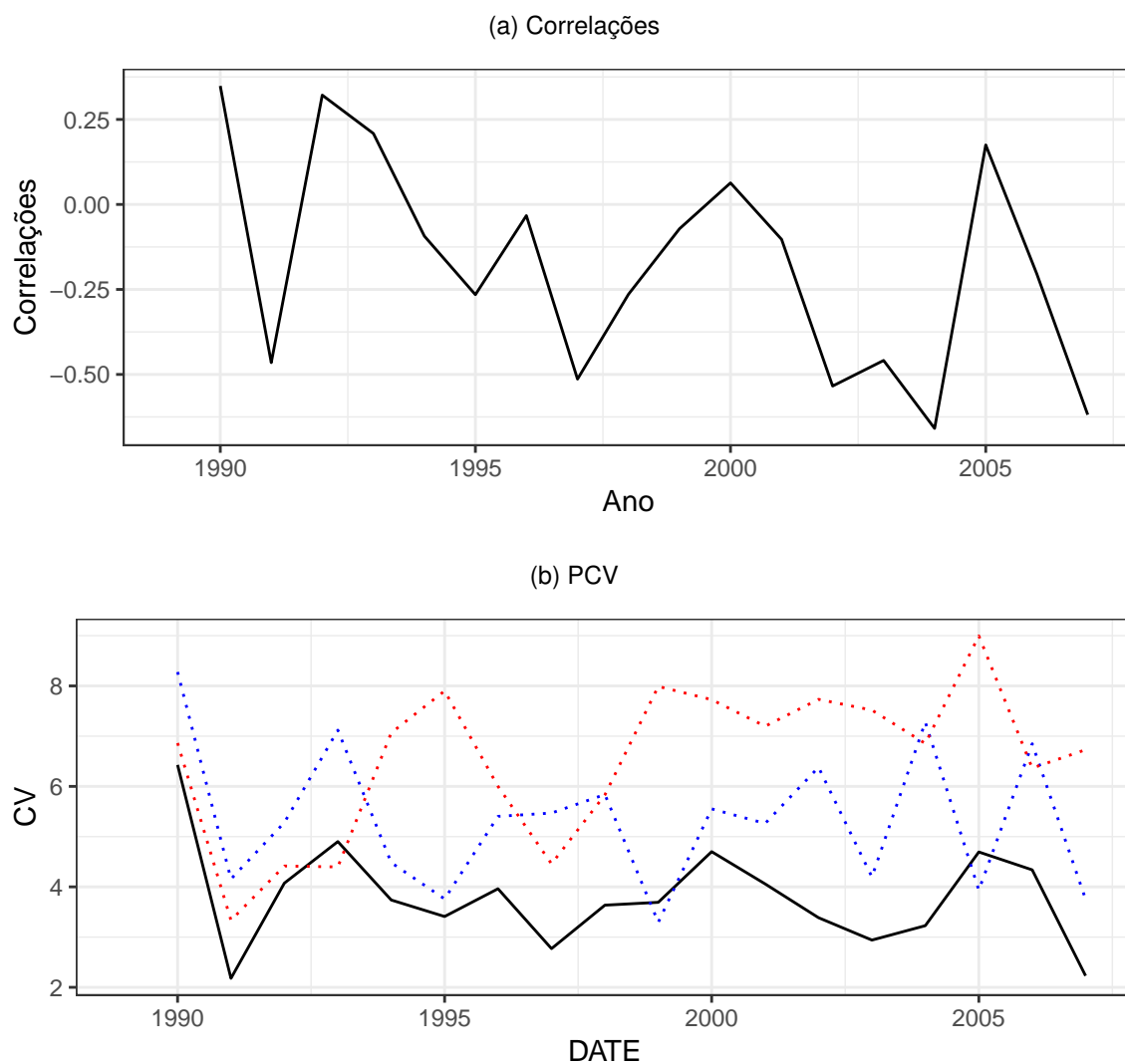
A componente interanual é analisada de maneira similar à componente de resíduo mensal. A figura 26 mostra a correlação em função da distância entre os elementos de cada par para esta componente. É possível observar que, em relação à componente mensal, a interanual decai levemente mais devagar. Também, pode ser observado um espalhamento maior em torno do comportamento médio obtido com o *loess*.

O mapa de calor das correlações para esta componente é mostrado na figura 27, em que linhas referentes a elementos no Nordeste são destacadas com retângulos. Comparando esta figura com a figura da componente mensal (figura 16), pode-se verificar que há muito mais elementos com valores mais altos de correlação, o que era esperado devido à maior persistência espacial de fenômenos mais lentos. Em relação à região Nordeste, é verificado que as menores correlações se dão com elementos no sul do subsistema SE/CO e no subsistema sul (32-34 e 36-39).

A Figura 28 mostra as médias dos valores de irradiação a longo prazo entre os dois elementos de cada par, versus a respectiva correlação. A área de maior interesse para alta geração e com correlações negativas, marcada com um retângulo na figura, é destacada na figura 29, em que os pares são explicitados via texto e são coloridos de acordo com os subsistemas em que cada elemento do par se encontra. A figura destaca o bom desempenho de pares N-SE/CO (Azul esverdeado), com o elemento do Norte encontrando-se em Roraima, pares NE-SE/CO (laranja), com os elementos em SE/CO encontram-se próximos ao Mato Grosso do Sul, e pares NE-S (amarelo).

A figura 30 mostra os valores do OPCV interanual para cada par, junto com a distância entre seus respectivos elementos. É interessante observar que há uma dependência com a

Figura 25 – Dinâmica ao Longo dos Anos do Comportamento da correlação e do OPCV para a Componente de Resíduo Mensal considerando o Par de Medoides 32-16



CV da componente de resíduo mensal calculado anualmente relativo às localidades do medoide 16 (azul) e 32 (vermelho). Em linha contínua, o CV da combinação ótima, a longo prazo, para o par analisado.

Elaboração própria

Figura 26 – Correlação Versus Distância Considerando a Componente Interanual

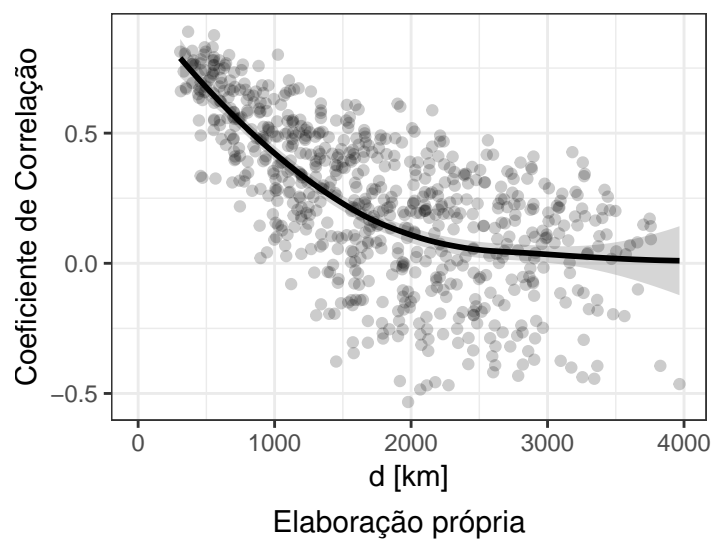
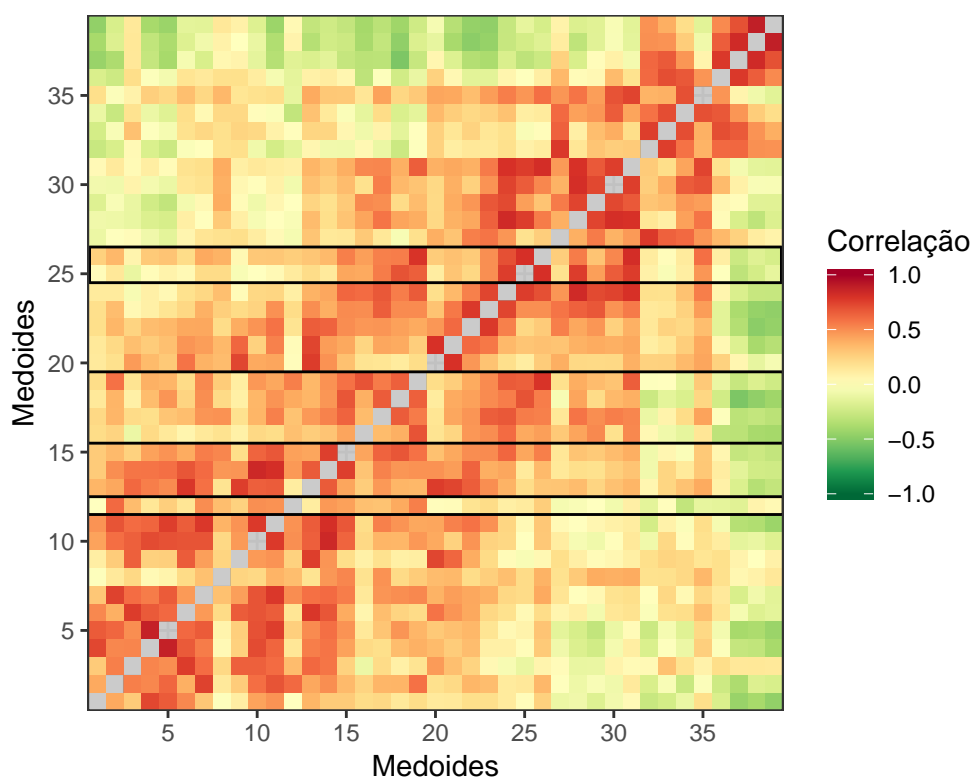


Figura 27 – Gráfico de Calor das Correlações para a Componente Interanual



Eixos representam os medoides, que podem ser identificados na figura 13 ou no apêndice A. Retângulos ressaltam linhas relativas a medoides no Nordeste, aos quais é dado um foco maior na discussão.

Figura 28 – Correlação versus Irradiância Média Diária Considerando a Componente Interanual

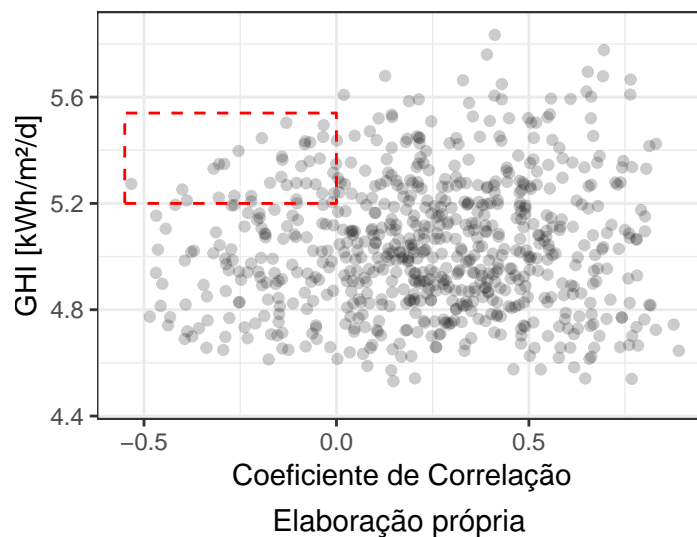
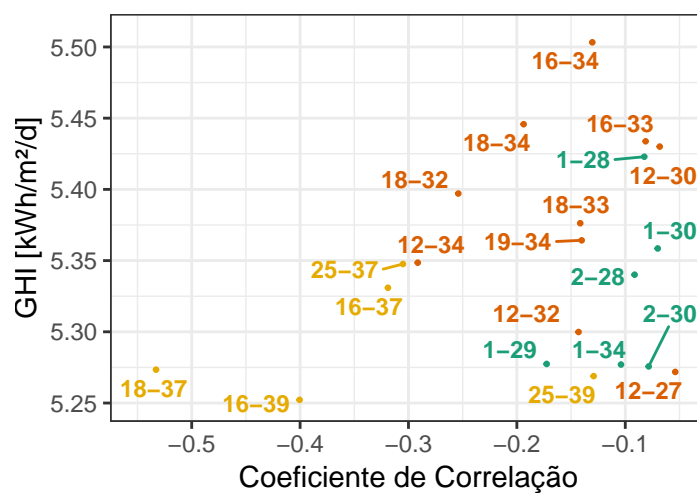


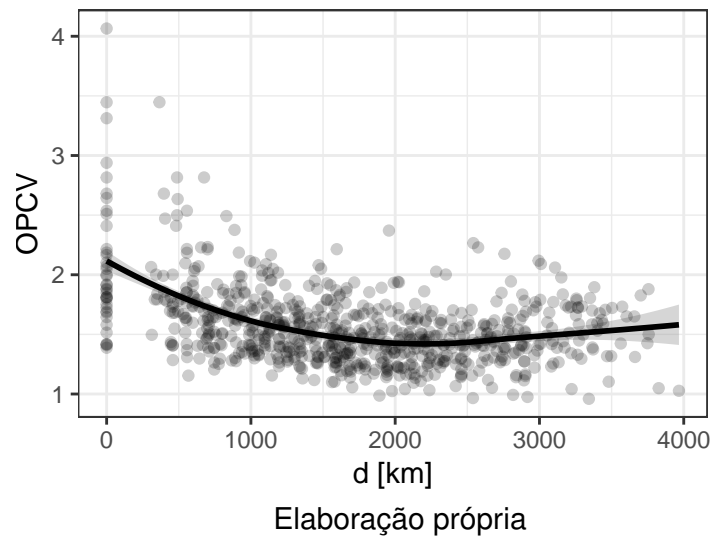
Figura 29 – Área de Interesse do Gráfico da Figura 28



Rubricas mostram os medoides dos pares, que podem ser identificados no apêndice A são relativas aos subsistemas: N-SE/CO=Azul Esverdeado; N-NE=Rosa; NE-S=Amarelo; NE-SE/CO=Laranja

Elaboração própria

Figura 30 – OPCV versus Distância Considerando a Componente Interanual



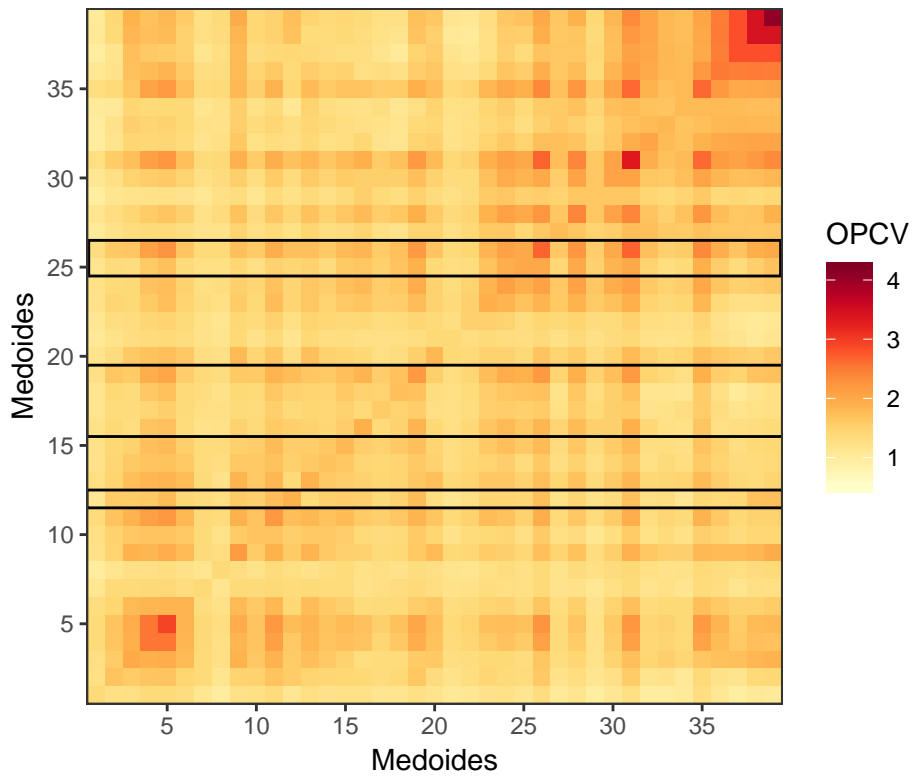
distância, especialmente quando se olha para o limite superior, mas que esta dependência só é forte para distâncias de até 1500 km. A partir deste valor, parece não haver muita relação entre o OPCV e a distância.

A figura 31 mostra o mapa de calor para o OPCV considerando a componente interanual. Pode-se observar que este gráfico parece mais difícil de interpretar que os outros mapas de calor. São observáveis baixos valores de OPCV para algumas colunas, de forma similar ao da figura 20, mas em maior evidência. Em relação às linhas destacadas com um retângulo, referentes à região de mais alta geração, no Nordeste, pode-se observar valores mais baixos de OPCV relativos a pares com o outro elemento no sul do subsistema SE/CO ou na a região Sul (32-34, 36-39). As regiões do Acre e de Rondônia (20-22) e do norte do Pará (7 e 8), também se destacam. Pode ser observado na figura 32, que tais regiões apresentam baixos valores de CV anuais.

Um ponto interessante em relação à figura 32 é que ela apresenta transições muito menos suaves, e aparentemente mais ruidosas, do que a figura 21. Isto sugere a necessidade de um contingente maior de dados maiores para a avaliação adequada desta componente, justificável pelas baixas amplitudes de suas oscilações e pelo fato de apresentar oscilações mais lentas. A análise é prosseguida para esta componente, mas ressalta-se que o resultado deve ser utilizado com cuidado.

Para identificar pares com alto potencial de geração e baixa variabilidade global, é explorada a figura 33, que mostra o valor da irradiância relativa à combinação de cada OPCV pelo respectivo valor do OPCV. Duas regiões de interesse são demarcadas e plotadas separadamente na figura 34. Em 34b é ressaltada a região do gráfico que apresenta os melhores valores de OPCV junto a altos valores do respectivo GHI diário. Dentre os pares em evidência, verificam-se combinações do subsistema Norte, no estado de Roraima, com

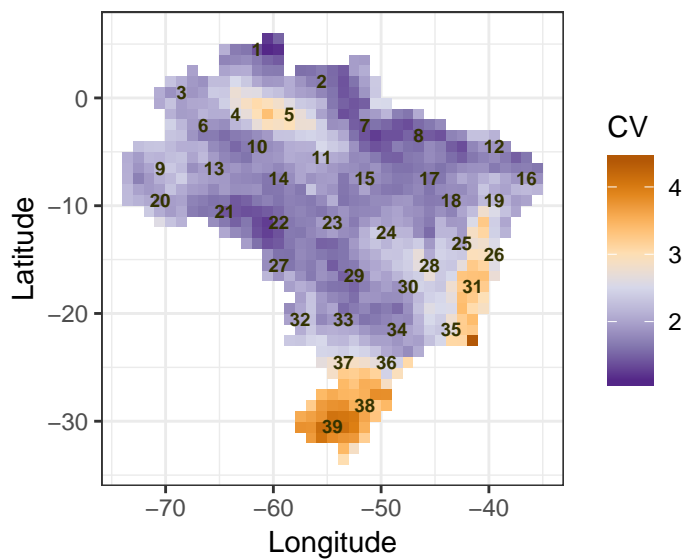
Figura 31 – Gráfico de Calor dos OPCV para a Componente Interanual



Eixos representam os medoides, que podem ser identificados na figura 13 ou no apêndice A. Retângulos ressaltam linhas relativas a medoides no Nordeste, aos quais é dado um foco maior na discussão.

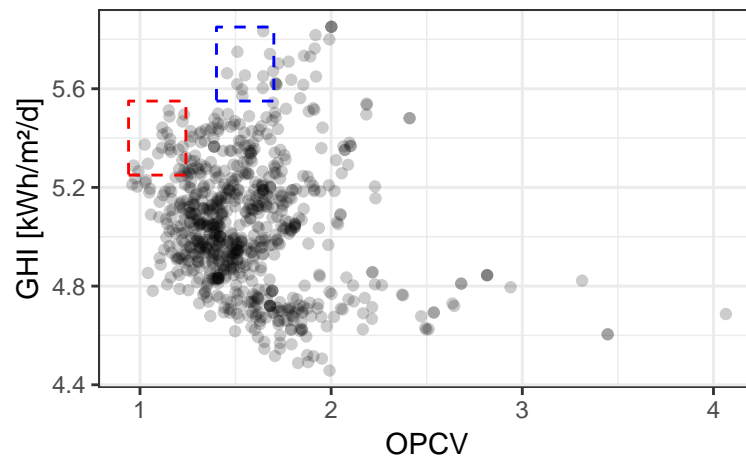
Elaboração própria

Figura 32 – Coeficientes de Variabilidade da Componente Internaual Para o Brasil, com Identificação dos Medoides



Elaboração própria

Figura 33 – Irradiância Diária Média versus OPCV Considerando a Componente Interanual



Elaboração própria

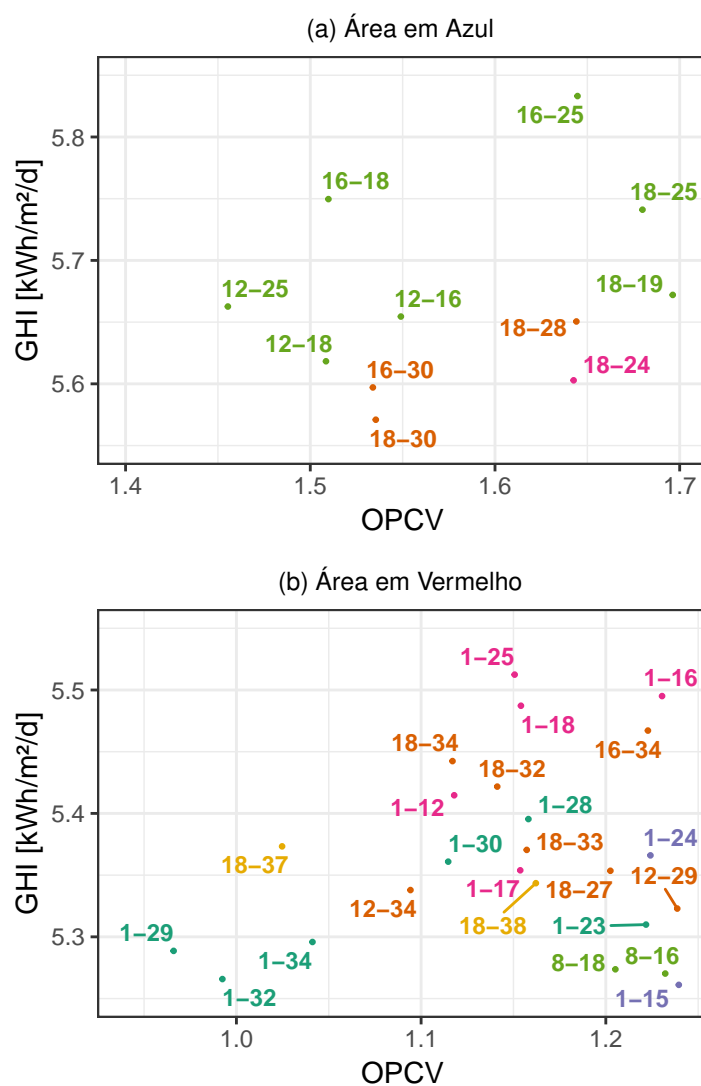
elementos no mesmo subsistema, mais distantes (lilás), com elementos do subsistema Nordeste (rosa) e com diversas partes do subsistema SE/CO (azul esverdeado). Os pares de regiões em evidência na análise baseada nas correlações (NE-S, amarelo; NE-SE/CO, laranja), apesar de ainda se mostrarem com bons resultados, perdem a importância quando avaliado o OPCV, devido aos resultados apresentados na região de Roraima, dado seu baixo valor de CV interanual, como pode ser visto na figura 32. A figura 34a mostra uma segunda região de interesse, que apresenta valores máximos de irradiância junto aos menores valores de OPCV para esta faixa de potencial. Nesta região da figura, são observados, principalmente, pares com os dois elementos no subsistema Nordeste (verde).

A mesma discussão feita quando analisada a componente de resíduo mensal sobre a utilização do OPCV e seu resultado prático pode ser feita: o uso do OPCV permite a visualização diretamente de um *trade-off* entre potencial de geração por metro quadrado e, conseqüentemente, menor custo de geração e a variabilidade total. Também, ressaltam-se as diferenças dos resultados entre os estudos baseados no OPCV e na correlação. As análises nessa seção mostraram similaridades com as análises da componente de resíduo mensal. Ambas mostraram as regiões de Roraima e do Mato Grosso do Sul como interessantes para reduzir a variabilidade global da geração solar.

4.2.4.1 Análise Baseada em Pontos de Referência

A ferramenta para analisar pontos de referência também foi utilizada para avaliar a componente interanual. Os gráficos com os mesmos pontos de interesse avaliados para a componente de resíduo mensal são apresentados para esta componente na figura 35. Um

Figura 34 – Áreas de Interesse do Gráfico da Figura 33



Rubricas mostram os medoides dos pares, que podem ser identificados no apêndice A
 Cores são relativas aos subsistemas: N-N=Roxo; N-NE=Rosa; N-SE/CO=Azul Esverdeado;
 NE-NE=Verde; NE-S=Amarelo; NE-SE/CO=Laranja.

Elaboração própria

ponto interessante de se observar nestes gráficos é o comportamento mais ruidoso relativo a esta componente.

4.3 Análise das Componentes Sazonal e de Tendência

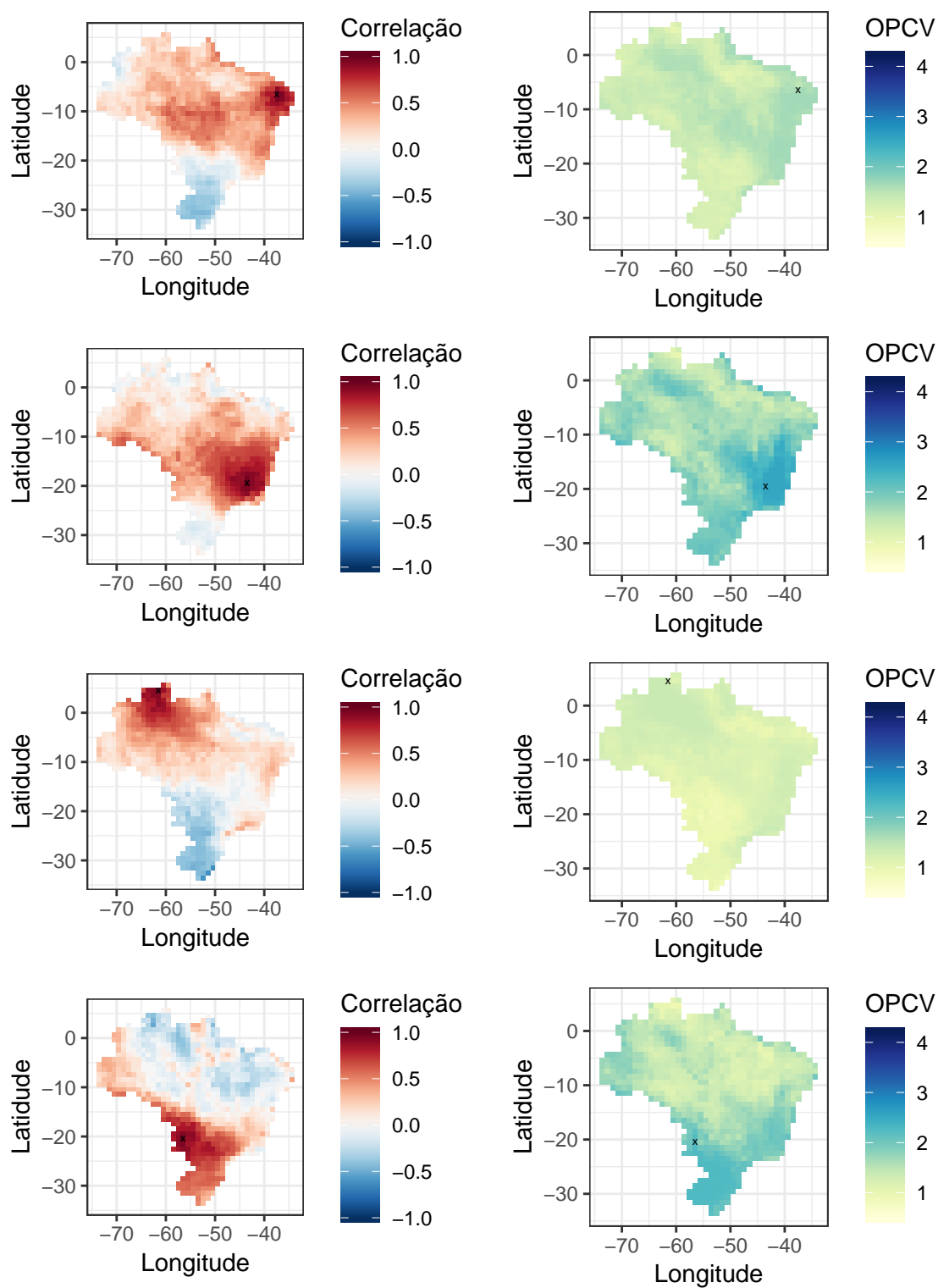
Nesta seção, as componentes sazonal e de tendência são avaliadas. Essas componentes podem ser avaliadas visualmente para cada ponto, pois não apresentam oscilações aleatórias de alta frequência. Assim, a maneira proposta para as avaliar é baseada na utilização de análise de componentes principais, como já discutido na seção 3.5, para auxiliar a visualização global do comportamento destas componentes com a utilização de poucos mapas. Relembrando, são avaliadas, para cada componente de série temporal, as primeiras CPs que juntas explicam até 90% da variância total. Essas componentes são avaliadas com gráficos que apresentam suas composições e seus comportamentos em relação ao tempo, junto a mapas que mostram os fatores, ou multiplicadores, de cada *pixel* para cada componente principal avaliada. O comportamento da série temporal para cada *pixel* é, então, aproximadamente igual ao comportamento da soma de cada componente principal vezes o respectivo multiplicador na localidade.

4.3.1 Componente Sazonal

Para a componente sazonal, foi utilizada a técnica de análise de componentes principais, considerando um período de 12 meses, de janeiro a dezembro, já que os valores mensais desta componente se repetem em todos os anos, por definição. Com duas componentes principais, explica-se 94% da variância dos dados, assim, as duas primeiras PCs, conseguem explicar mais de 90% da variância dos dados e são analisadas. A figura 36a mostra o comportamento dessas componentes principais ao longo de um ano.

As figuras 36b e 36c mostram mapas com os fatores das duas CPs para cada pixel analisado. Estes fatores representam o quanto de uma componente há em certa localidade, de modo em que a série sazonal em um dado ponto seja aproximadamente igual à primeira componente (PC1) vezes o seu fator nesta localidade mais a segunda componente (PC2) vezes seu respectivo valor no local. Pelas figuras, é possível observar, por exemplo, uma diversidade de padrões sazonais na região Nordeste. Com os dados usados para gerar a figura, pode-se verificar que a área com irradiâncias diárias com valores superiores a $5.5kWh/m^2/dia$ apresenta fatores da primeira componente principal que vão de valores negativos ($-0,97$) a elevados valores positivos ($2,75$). Esta região também apresenta um alcance considerável da segunda componente, de ($0,53$ a $1,88$). É interessante ressaltar que a utilização de componentes principais permite a visualização de forma mais compacta deste comportamento do que se 12 figuras mensais fossem utilizadas. Isto auxilia na visualização e compreensão da diversidade do comportamento sazonal.

Figura 35 – Mapas de OPCV e de Correlações em Análises Baseada em Pontos de Referência para a componente interanual

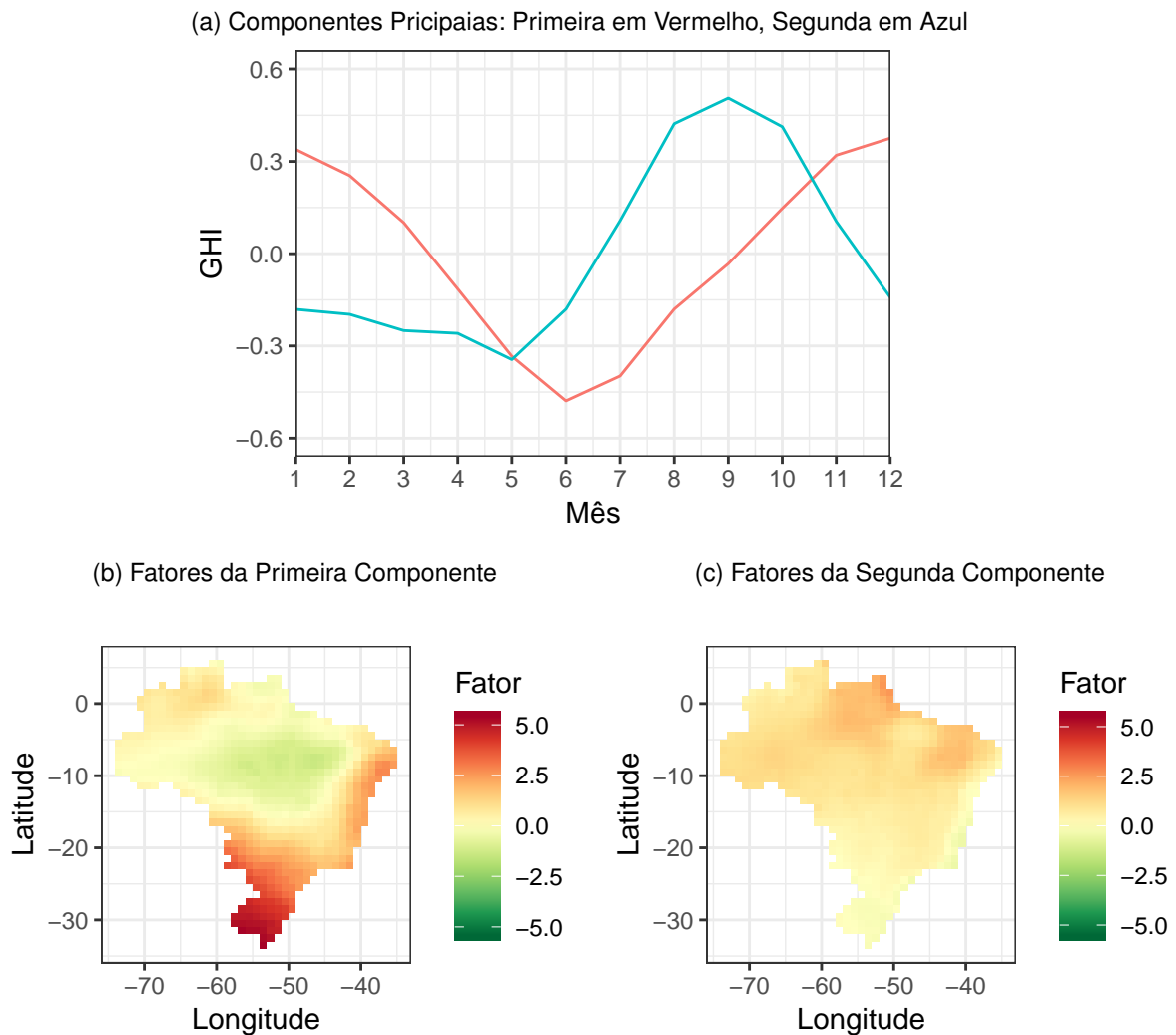


Correlações na esquerda e OPCV na Direita.

Pontos de referência: de cima para baixo: Patos; Belo Horizonte; Medoide 1; Medoide 32

Elaboração própria

Figura 36 – Componentes Principais Analisadas Relativas à Componente Sazonal



Elaboração Própria

Interpretando as componentes obtidas, pode ser visto que a primeira componente, que explica 66% da variância total dos dados, apresenta um pico no solstício de verão do hemisfério sul e um vale no solstício de inverno. Esta componente, conforme esperado, apresenta fatores mais elevados na região temperada, no sul, onde sua predominância ocorre principalmente devido a efeitos da inclinação relativa da terra em relação ao sol. Entretanto, é importante observar que tal componente também é presente em uma longa área do país, seguindo em sua costa até latitudes tropicais, de até -10° . Nestas localizações, os principais fatores que causam esse padrão são climáticos, sendo interessante observar que, mais ao centro do país, os fatores desta componente na mesma latitude, são negativos. É interessante observar que o uso do K_t poderia ofuscar o fato de que o litoral nordestino e o sul do país apresentam padrões sazonais semelhantes, dado que são causados por efeitos distintos.

A segunda componente principal explica 28% da variância dos dados. Esta componente possui um pico em setembro, mês em que ocorre o equinócio da primavera, e que, para grande parte do país, é mais seco do que o equinócio de outono. Esta componente é mais forte para regiões mais equatoriais, como a região do Amapá e Norte do Pará e a região do Piauí, Maranhão e Ceará. Com menos intensidade, também é presente em quase todo o país, com exceção da região Sul e da região litorânea do Sudeste e da Bahia. Quando associado a valores positivos da componente principal 1, representa um desvio do vale e do pico daquela componente para ocorrerem um pouco antes no ano.

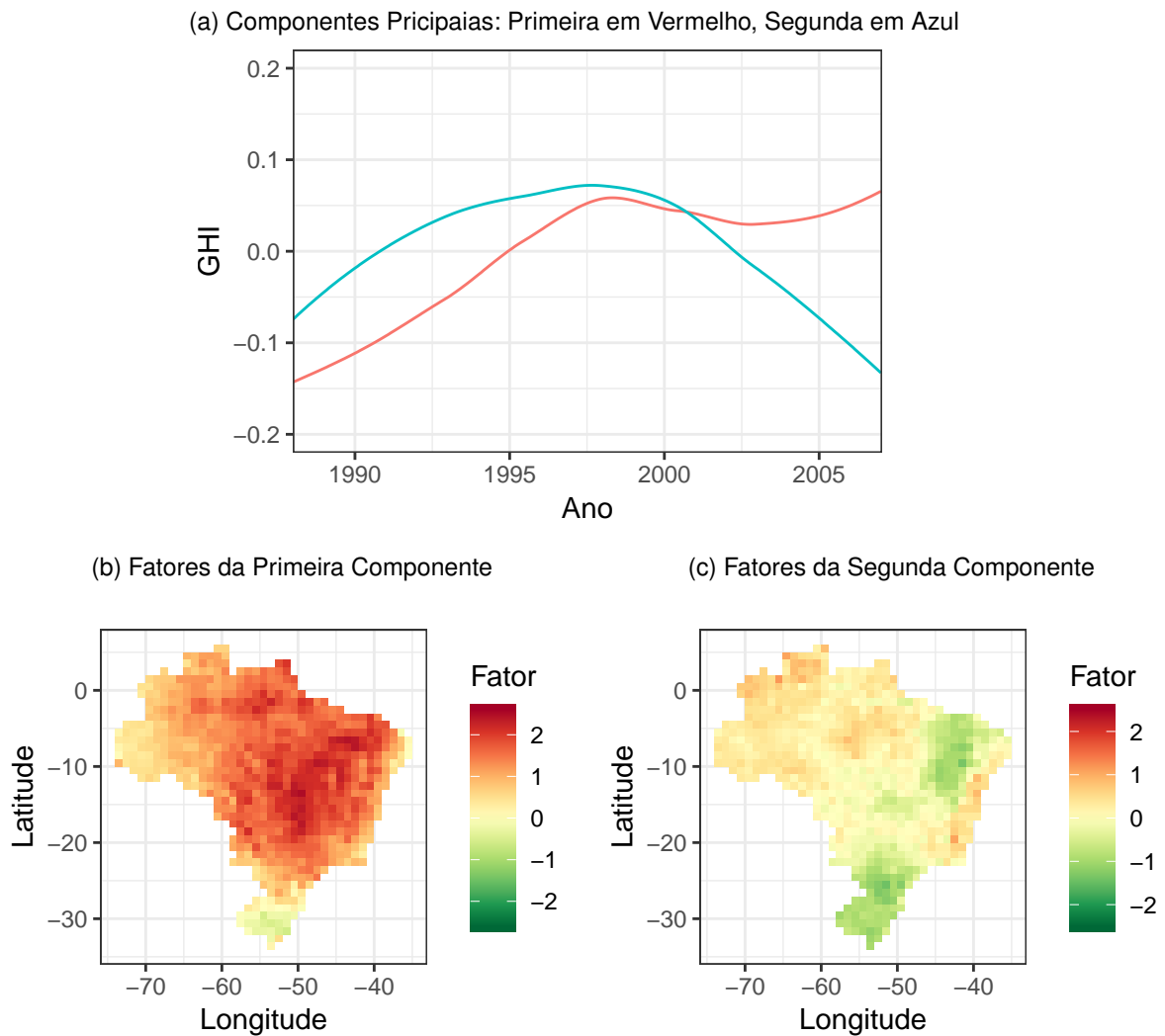
Considerando a dinâmica de consumo de energia no Sistema Interligado Nacional (SIN), em que os meses com maiores níveis de consumo energético e picos de energia diários são entre janeiro e março, é interessante ter mais energia solar instalada em locais que apresentam valores mais altos da primeira componente principal e mais baixos da segunda componente principal. Analisando os mapas, pode-se observar que este padrão é apresentado pela região mais litorânea do Nordeste e pela região mais ao sul do país, incluindo o estado do Mato Grosso do Sul.

4.3.2 Componente de Tendência

Como mencionado anteriormente, a principal razão para se considerar separadamente uma componente para a tendência é a de se remover possíveis vieses persistentes da análise. Entretanto, é interessante observar que dados obtidos com o algoritmo GEWEX SRB, o mesmo utilizado neste estudo, foram usados por Pinker, Zhang e Dutton (2005), em um estudo para verificar se os bancos de dados de satélite identificavam tendências. Neste artigo, publicado na revista *Science*, tendências foram modeladas como lineares ou de segunda ordem ao longo do período analisado, e os resultados foram estimados para agregação de longas áreas no globo, como o cinto tropical e a região ártica. Não há análises mostrando a dinâmica ao longo dessas áreas analisadas, como mapas, mas apenas um estudo da média da tendência nas regiões. Neste trabalho, é proposta a utilização da técnica PCA, como discutida na seção 3.5 e de maneira similar a como utilizado com a componente sazonal.

Para a avaliação da dinâmica da componente de tendência, foi retirada a média de longo prazo para analisá-la, uma vez que o comportamento das médias já pode ser avaliado com a figura 7. Duas PCs foram necessárias para se explicar 90% da variabilidade dos dados. De fato, estas duas componentes explicam 83% e 9% da variabilidade cada. A figura 37a mostra as dinâmicas dessas componentes. A primeira componente principal é caracterizada por um considerável aumento nos níveis de irradiação entre os anos de 1987 até 1998, seguida de uma leve oscilação na década seguinte. É interessante observar seus elevados valores positivos ao longo de quase todo o Brasil, especialmente na região do Nordeste e no estado de Goiás. Apenas no Rio Grande do Sul são verificáveis valores negativos, ainda que

Figura 37 – Componentes Principais Analisadas Relativas à Componente de Tendência



Elaboração Própria

bem levemente. Esta componente sugere que houve um aumento nos níveis de irradiância no Brasil ao longo do tempo (fatores próximos a 2, por exemplo, representam um aumento significativo, de aproximadamente $0,4 \text{ kWh/m}^2/\text{dia}$), principalmente entre 1988 e 1998, ou que houve alguma mudança sistemática que fez com que as medições relatassem este comportamento. Ressalta-se, novamente que estes efeitos verificados nos dados podem ser devidos a questões de vieses persistentes que afetam os dados.

A segunda componente explica apenas 9% da variabilidade dos dados. Ela representa um aumento dos níveis de irradiância nos primeiros dez anos da análise, seguida por um decréscimo nestes níveis na década seguinte. Está mais destacavelmente presente com sinal negativo, no Sul e no Nordeste. Quando com sinal negativo e associada a valores positivos da CP1, como no Nordeste, representa um aumento dos níveis solares menos concentrada nos dez primeiros anos e mais distribuídos ao longo do tempo da análise.

Um ponto interessante de se observar nos mapas é que os padrões se apresentam bastante ruidosos. Estes fenômenos, mais lentos, apresentam, em geral, maior persistência espacial, de forma que este comportamento não é esperado. Entretanto, como as oscilações são consideravelmente suaves, esta componente fica mais susceptível a ruídos.

É interessante se observar que, caso o período de análise fosse maior, poderiam ser necessárias mais componentes principais para a compreensão do comportamento da componente tendência, o que poderia deixar sua interpretação mais difícil. Uma opção, neste caso, poderia ser a análise desta componente considerando diferentes janelas de tempo separadamente.

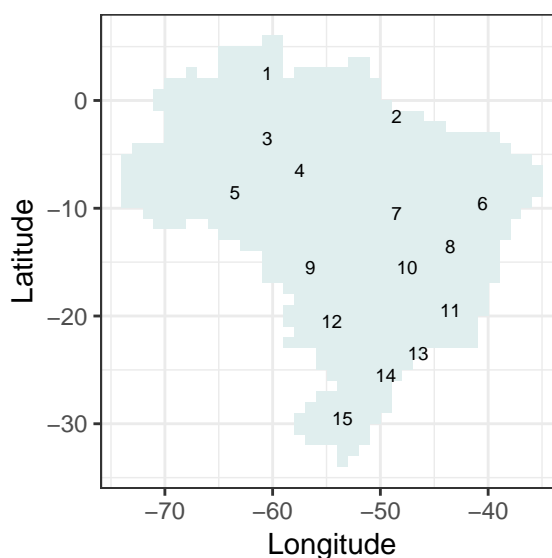
4.4 Comparações entre Banco de Dados

Nesta seção são comparados o banco de dados utilizado nas análises com aquele utilizado no projeto internacional SWERA, para determinar os anos meteorológicos típicos, baseados em dados de estações meteorológicas. A tabela 2 mostra as cidades das estações, suas coordenadas geográfica e aquelas relativas aos pontos centrais dos *pixels* nos quais tais estações se encontram e a distância entre estas localidades. Além disso, a coluna “ID” mostra um identificador, utilizado na comparação entre dados dos dois bancos para as componentes interanual e de resíduo mensal. Tais valores foram determinados em ordem de latitude (de norte para o sul) e longitude (de oeste para leste). Para essa comparação, estações em cidades litorâneas não foram consideradas, devido aos comportamentos dos píxeis nestas. A figura 38 mostra a localização das estações consideradas na comparação destas componentes.

A figura 39 mostra os dados de irradiação diários das estações não litorâneas junto aos dados do modelo SRB do GEWEX para as mesmas datas. É verificável uma dispersão muito grande dos dados, visualmente muito superior à observada no documento de validação dos dados do GEWEX SRB (STACKHOUSE et al., 2018), que os compara com dados de algumas estações certificadas, principalmente na região temperada do hemisfério norte. Um ponto interessante de se observar no gráfico é do limite inferior verificado nos dados diários do SWERA, de aproximadamente 0,9, e que não é verificado nos dados do projeto da NASA. Também, valores máximos do SWERA são um pouco menores do que os do modelo do GEWEX.

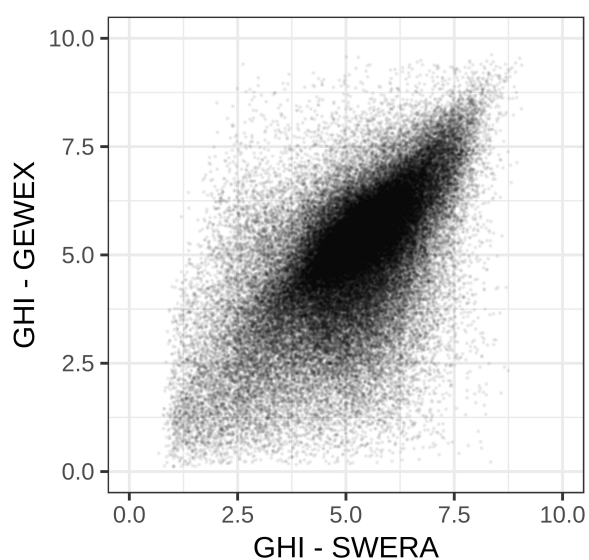
A figura 40 compara as componentes de tendência obtidas para os dados da estação do SWERA e para os dados do GEWEX SRB considerados. Chama a atenção que os comportamentos para mesmos períodos de dados sejam tão diferentes entre as fontes de dados para a maioria das estações, com poucas exceções, como em Brasília. Isso mostra uma elevada incoerência entre os bancos de dados. Algo interessante de se observar nos gráficos é que, no banco de dados do SWERA, não se verifica, nas maiorias das estações, o

Figura 38 – Mapa com a Localização das Estações Utilizadas do SWERA e Comparadas com os Dados do Projeto POWER



Elaboração Própria

Figura 39 – Dados Diários do GEWEX versus Dados Diários do SWERA.



Elaboração Própria

aumento dos níveis de radiação até o ano de 1997, que é representado pela primeira PC desta componente e perceptível em muitos dos locais do gráfico com os dados da NASA.

A figura 41 mostra as componentes sazonais dos dois bancos de dados para estas localidades. É interessante observar uma similaridade muito maior para esta componente do que para a anterior. De fato, as múltiplas medições e a maior amplitude contribuem para com que as diferenças nesta componente sejam menores. Em relação às localidades, Cuiabá chama a atenção pelos resultados bastante diferentes entre as fontes de dados, mesmo estando distante do oceano.

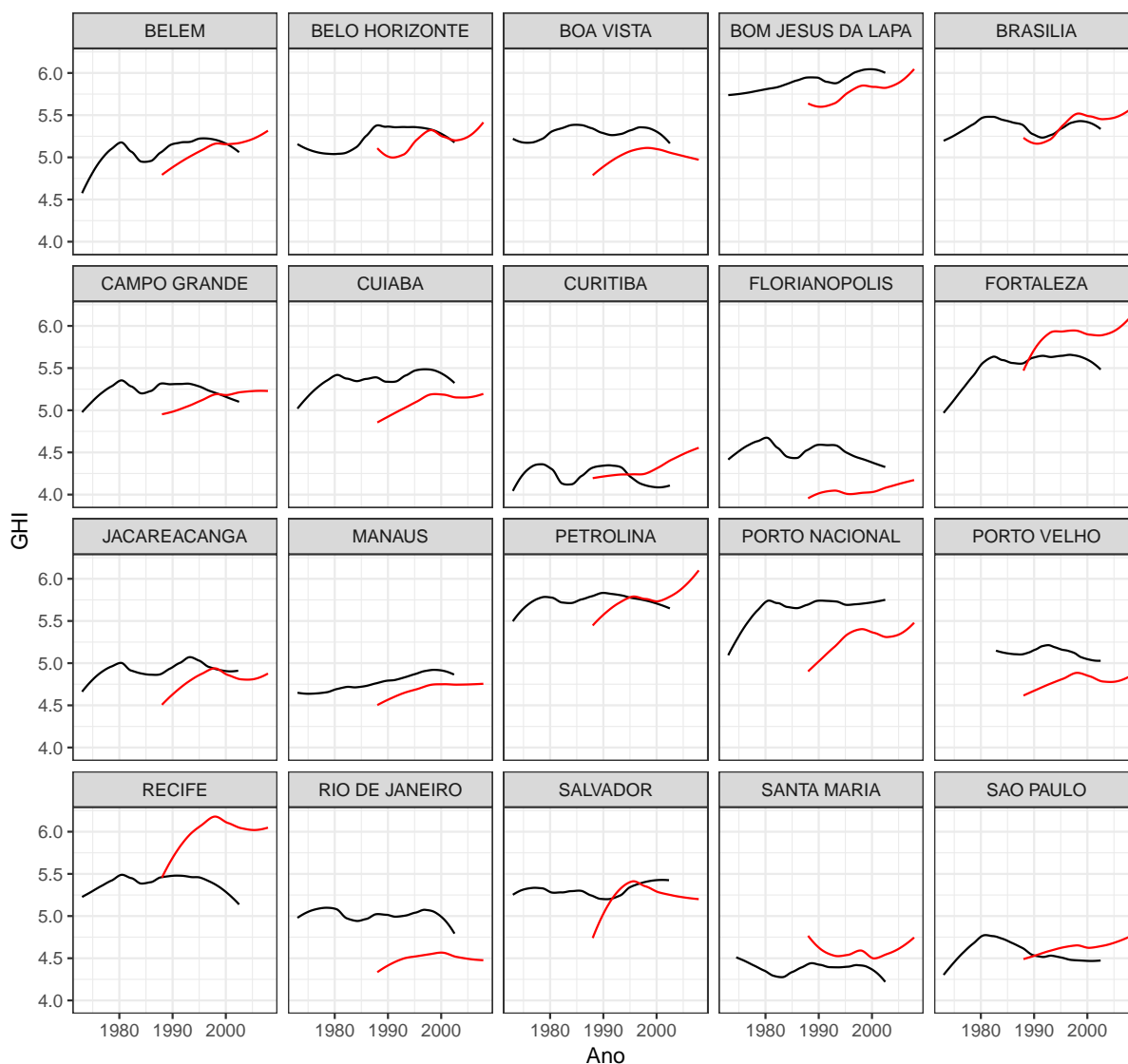
Pode-se verificar na figura 42 os gráficos dos dados mensais do modelo GEWEX pelos dados mensais do SWERA, considerando as séries antes da decomposição, além das componentes de resíduo mensal e interanual. Nesta figura, só foram consideradas as estações não litorâneas, com números na coluna ID na Tabela 2. Comparando com a figura 39, é verificável uma dispersão muito menor para os dados mensais, o que é esperado devido à diminuição da variabilidade que ocorre na agregação mensal. Entretanto, mesmo para as séries sem decomposições, a dispersão é consideravelmente elevada. Ao lado, vê-se que, quando analisadas as componentes de resíduo mensal e interanual, há baixas correlações entre as fontes de dados consideradas.

Mesmo com as correlações baixas entre os dados pontuais dos bancos de dados, a

Tabela 2 – Estações disponíveis no banco de dados do SWERA

Estação	ID	Latitude Swera	Longitude Swera	Latitude Pixel	Longitude Pixel	Distância [km]
BELEM	2	-1.38	-48.48	-1.5	-48.5	13.5
BELO HORIZONTE	11	-19.85	-43.95	-19.5	-43.5	61.1
BOA VISTA	1	-2.83	-60.70	-2.5	-60.5	42.7
BOM JESUS DA LAPA	8	-13.27	-43.42	-13.5	-43.5	26.9
BRASILIA	10	-15.87	-47.93	-15.5	-47.5	61.7
CAMPO GRANDE	12	-20.47	-54.67	-20.5	-54.5	18.0
CUIABA	9	-15.65	-56.10	-15.5	-56.5	46.0
CURITIBA	14	-25.52	-49.17	-25.5	-49.5	33.2
FLORIANOPOLIS	-	-27.67	-48.55	-27.5	-48.5	19.5
FORTALEZA	-	-3.78	-38.53	-3.5	-38.5	31.1
JACAREACANGA	4	-6.27	-57.73	-6.5	-57.5	36.0
MANAUS	3	-3.13	-60.02	-3.5	-60.5	67.2
PETROLINA	6	-9.35	-40.55	-9.5	-40.5	17.5
PORTO NACIONAL	7	-10.70	-48.40	-10.5	-48.5	24.7
PORTO VELHO	5	-8.77	-63.92	-8.5	-63.5	55.0
RECIFE	-	-8.07	-34.85	-8.5	-34.5	61.2
RIO DE JANEIRO	-	-22.90	-43.17	-22.5	-43.5	55.8
SALVADOR	-	-13.02	-38.52	-13.5	-38.5	53.1
SANTA MARIA	15	-29.70	-53.70	-29.5	-53.5	29.4
SAO PAULO	13	-23.62	-46.65	-23.5	-46.5	20.3

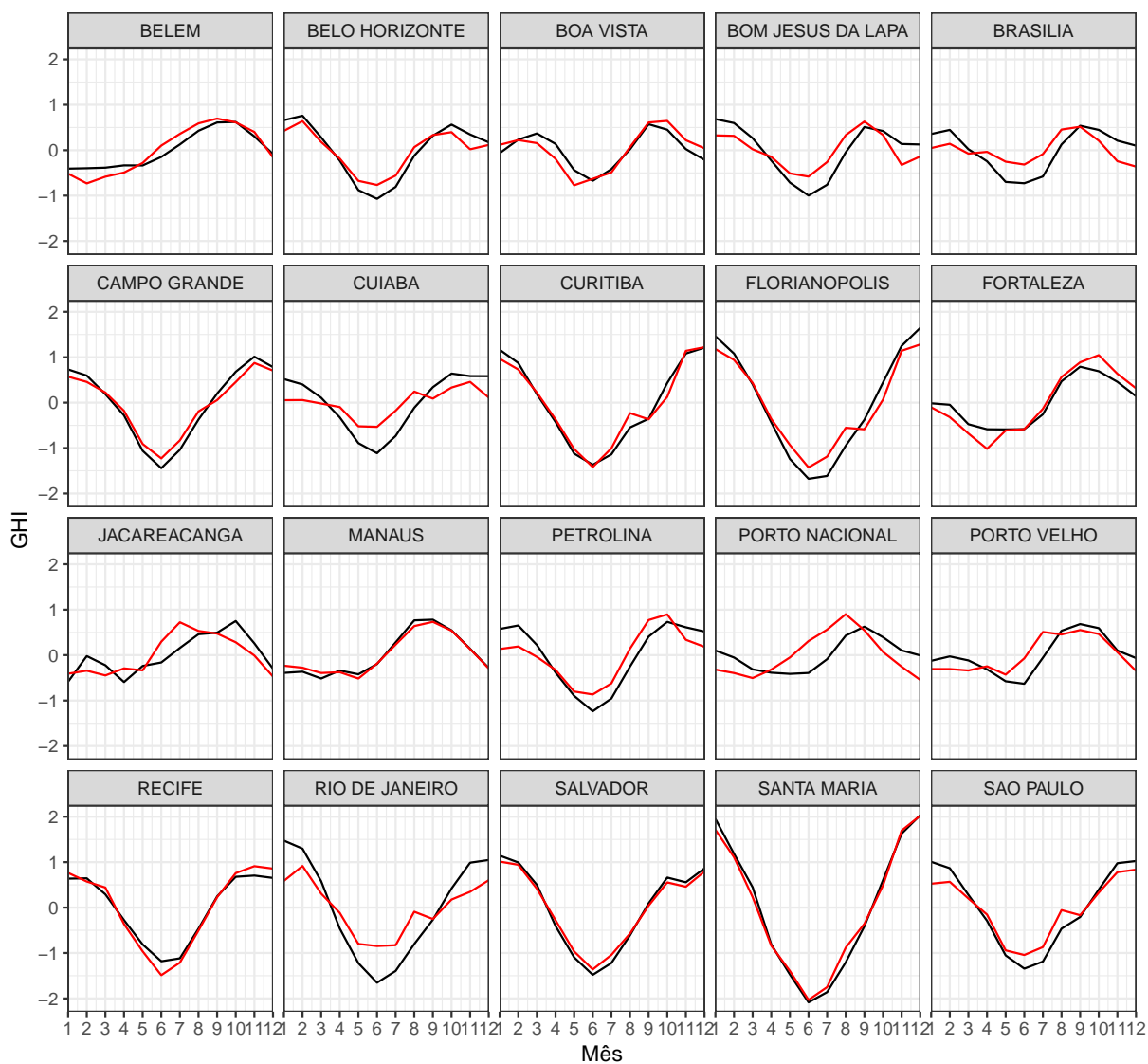
Figura 40 – Componentes de Tendência nas Localidades das Estações de Medição utilizados no SWERA, Para os Dois Bancos de Dados Usados



Em preto: Dados das estações utilizadas no SWERA. Em vermelho: Dados do sistema POWER, considerados na análise, obtidos com o modelo GEWEX SRB v3.0

Elaboração Própria

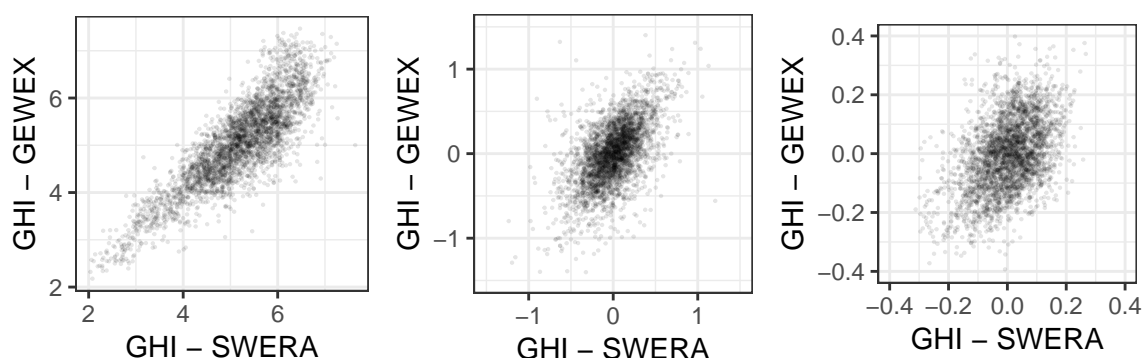
Figura 41 – Componentes Sazonal nas Localidades das Estações de Medição utilizados no SWERA, Para os Dois Bancos de Dados Usados



Em preto: Dados das estações utilizadas no SWERA. Em vermelho: Dados do sistema POWER, considerados na análise, obtidos com o modelo GEWEX SRB v3.0

Elaboração Própria

Figura 42 – Dados mensais do GEWEX versus Dados Mensais do SWERA



Da esquerda para a direita: Dados mensais sem decomposição; Componente de resíduo mensal; Componente interanual.

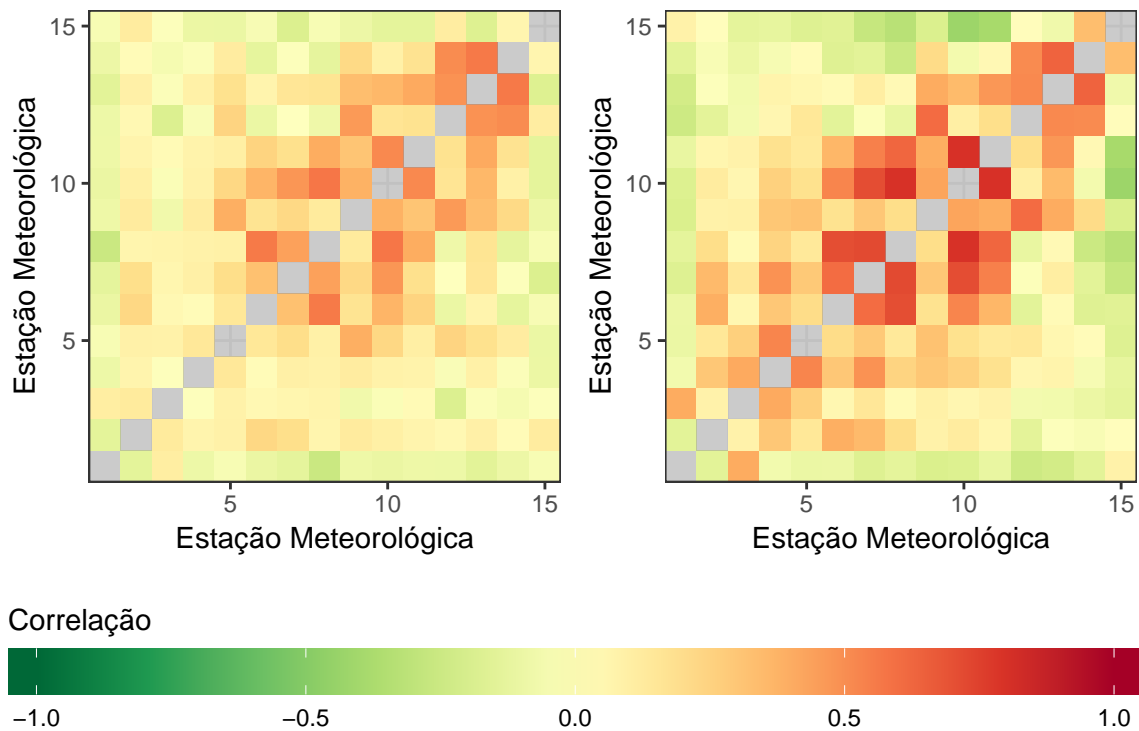
Elaboração Própria

seguir são comparados os comportamentos dos índices de correlação e do OPCV de pares de estações meteorológicas e de *pixels* nestas localizações para as duas fontes. As figuras 43 e 44 mostram, respectivamente, os mapas de calor das correlações e dos OPCV das séries de resíduo mensal para os dois bancos de dados, considerando os anos de 1983 a 2002, em que há dados disponíveis para ambos. A figura 45 mostra as correlações e os OPCV do SWERA versus os valores destes índices para os dados obtidos com o modelo GEWEX SRB no projeto POWER.

Quando comparadas as correlações, percebe-se um comportamento semelhante entre as fontes de dados para a maior parte dos pares de dados (como é o caso do par 6-12 e do par 8-12, entre o Nordeste e o MS), entretanto, em geral, as correlações das estações do SWERA, tanto negativas quanto positivas, parecem ser muito mais próximas de zero. Uma regressão linear considerando os dados do GEWEX como preditores dos dados do SWERA, também mostrada na figura 45, tem *offset* muito próximo de zero e uma inclinação de 0,6. É interessante observar que esta regressão consegue explicar 76,7% da variância dos dados, como é expresso pelo seu r^2 . Observando a figura 44, pode-se verificar padrões semelhantes entre os bancos de dados em relação ao OPCV, de uma maneira em que os valores para o banco do SWERA se apresentam inferiores e um pouco menos variáveis. A regressão linear para esta componente apresenta um r^2 um pouco inferior à apresentada pelas correlações, de 71% mas, também, razoável.

Para a componente interanual, entretanto, os resultados não coincidem tanto. Por simplicidade, é mostrado apenas os gráficos dos índices entre o banco de dados do SWERA e da NASA, e as regressões lineares considerando o segundo como preditor, na figura

Figura 43 – Comparação das Correlações Obtidas em Cada Banco de Dados, para o período de 1983 a 2002 - Componente de Resíduo Mensal



Da esquerda para a direita: Dados SWERA; Dados GEWEX SRB

Elaboração Própria

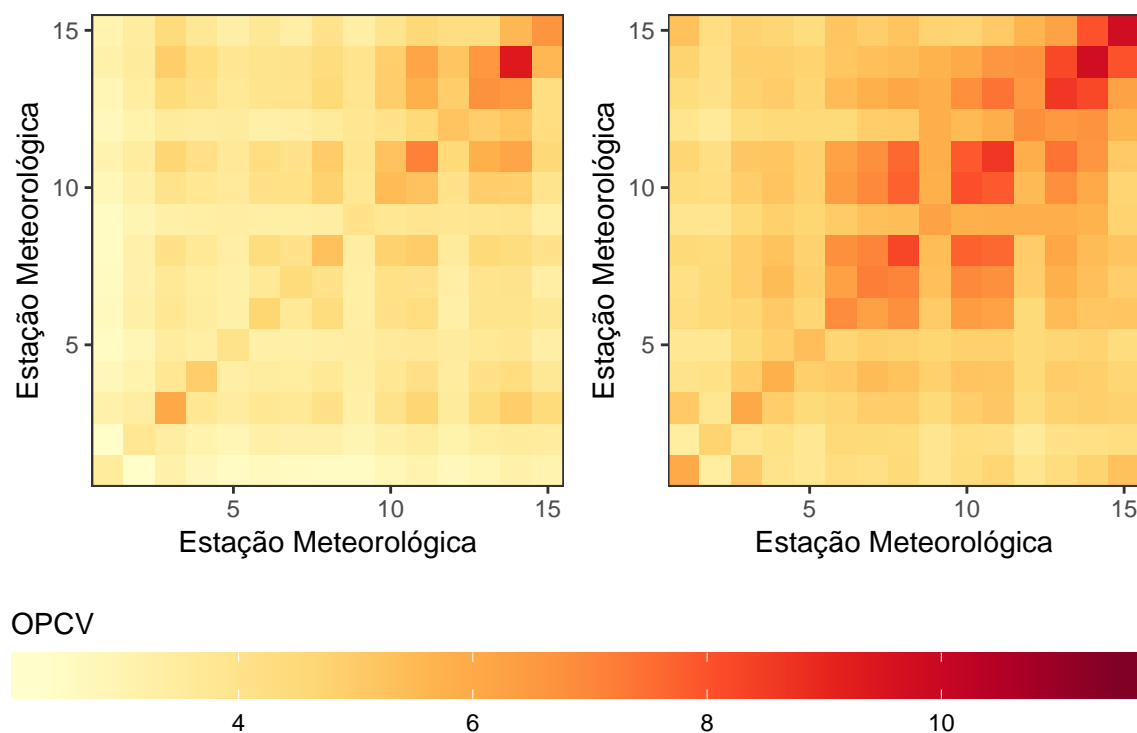
46. Em relação às regressões, pode-se verificar que os valores encontrados dos termos estimados se assemelham aos valores verificados para a componente de resíduo mensal. Entretanto, apresentam valores de r^2 muito inferiores, o que mostra uma grande divergência entre os bancos de dados.

4.5 Discussões Finais

Dentre os resultados, destacam-se os relativos ao da região do Mato Grosso do Sul, que é uma região que não apresenta uma tendência muito grande de crescimento do uso de energia solar, e está fortemente conectada ao SIN. Os resultados mostraram que este estado apresenta certa complementariedade com a dinâmica do comportamento do Nordeste para a componente de resíduo mensal; baixas variações desta componente; bons índices de irradiação e uma dinâmica sazonal com picos de produção próximos ao verão, época de picos de consumo. A região do Roraima também apresentou bons resultados, mas não está conectada ao SIN.

Os resultados apresentados ressaltaram que há dinâmicas diversas em relação ao comportamento do recurso solar no Brasil. Diferentes locais se diferem em relação a seus

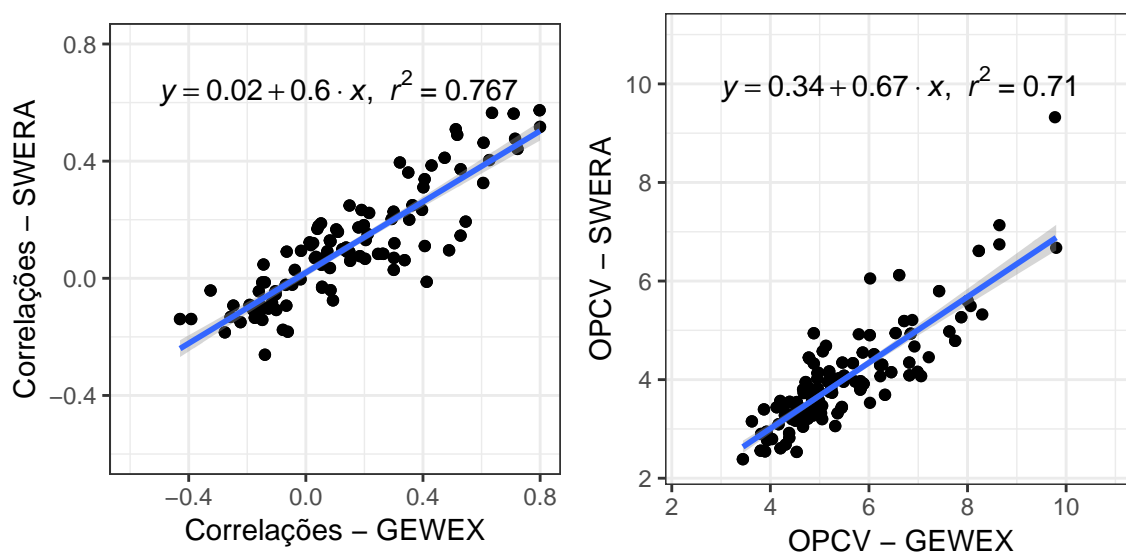
Figura 44 – Comparação dos OPCV Obtidas em Cada Banco de Dados, para o período de 1983 a 2002 - Componente de Resíduo Mensal



Da esquerda para a direita: Dados SWERA; Dados GEWEX SRB

Elaboração Própria

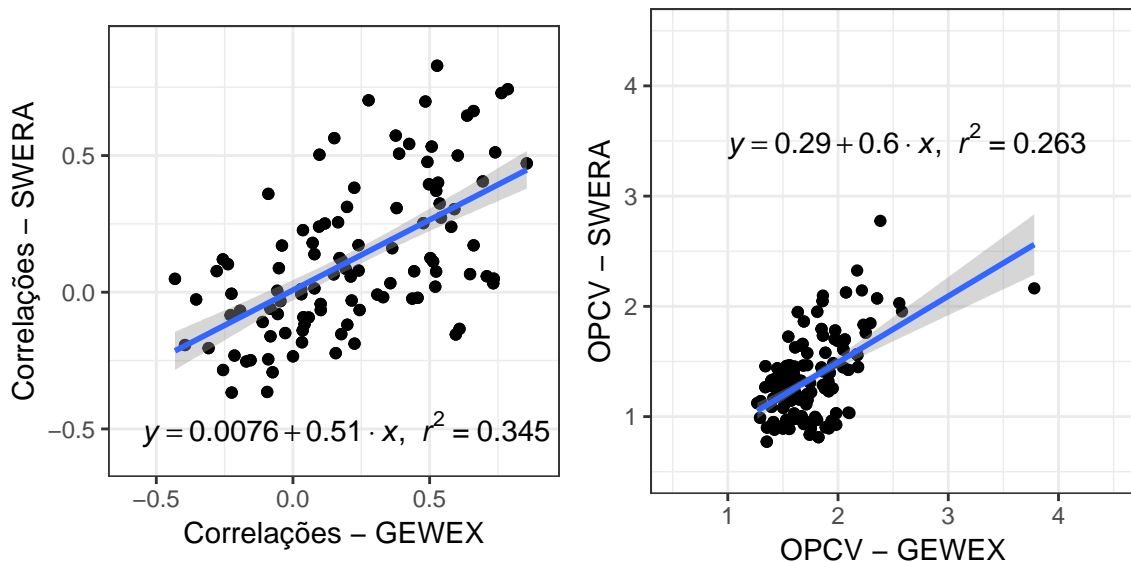
Figura 45 – Comparação de Correlações e OPCV Obtidos em Cada Banco de Dados, para o período de 1983 a 2002 - Componente de Resíduo Mensal



Da esquerda para a direita: Correlações; OPCV

Elaboração Própria

Figura 46 – Comparação de Correlações e OPCV Obtidos em Cada Banco de Dados, para o período de 1983 a 2002. Componente Interanual.



Da esquerda para a direita: Correlações; OPCV

Elaboração Própria

padrões sazonais e em relação a quão variável é o recurso em cada localidade. Também, foram observadas certas complementariedades entre as dinâmicas do recurso entre algumas regiões. Estes efeitos afetam a operação do SIN e, conseqüentemente, afetam os seus custos de funcionamento. Entretanto, eles não são considerados em leilões de energia solar, cujo produto a ser vendido é a energia. É interessante notar que, apesar de estes efeitos não serem tão importantes agora, podem ser no futuro, com a maior penetração de energia solar no sistema e que é interessante estudar como eles afetam os custos de operação do SIN.

Algumas sugestões de alteração nos processos de leilão e de regulamentação da geração distribuída são citadas. No quesito de redução de aleatoriedades, relacionado a distribuições espaciais dos geradores, leilões poderiam ser feitos para áreas de interesse específicas. Também, regiões de interesse poderiam ter incentivos maiores para a GDFV. No quesito sazonal, a energia contratada em leilões poderia, por exemplo, apresentar preços diferentes para cada mês.

É interessante notar que há certo incentivo implícito para a dispersão espacial, que afeta grupos empresariais proprietários de muitos geradores. Um grupo pode utilizar da dispersão espacial com o objetivo de reduzir flutuações em seus fluxos de caixa, considerando o saldo anual da energia solar, visando a minimização da flutuação total da produção de energia em seu portfólio, e riscos relacionados envolvidos.

5 CONCLUSÕES

Considerando a tendência de crescimento da utilização de energia solar no Brasil, este trabalho buscou avaliar a dinâmica da disponibilidade do recurso solar a longo prazo no país, utilizando dados integralizados a uma escala de tempo mensal. Esta escala de tempo foi escolhida levando em conta a limitação de dados de resoluções finas disponíveis para o Brasil, a extensão do território nacional e o fato de que, apesar de apresentar importância no funcionamento de sistemas eletroenergéticos, especialmente no contexto do planejamento, análises da variabilidade do recurso solar a longo prazo ainda não foram muito exploradas na literatura, como foram variabilidades intradiárias e intra-horárias. Assim, buscou-se obter um conjunto de ferramentas para a análise do recurso solar a longo prazo para uma área continental.

Nas análises relativas à dinâmica solar, muitas vezes são utilizados o índice de claridade ou o índice de céu claro. A utilização de tais índices, entretanto, diminui o significado prático da análise, especialmente quando são consideradas áreas mais abrangentes. Neste contexto, um mesmo valor de dado índice em pontos distantes pode representar valores muito diferentes de irradiação medida, uma vez que a irradiação extraterrestre ou a irradiação de céu claro podem ser muito distintas entre as localidades. Uma outra questão que se tem com a utilização destes índices é que eles podem não remover todos os efeitos de sazonalidade, dado que estes efeitos podem ser causados por fenômenos climáticos recorrentes, e não apenas pelas variações na inclinação relativa do sol. Assim, neste trabalho foram analisados dados de séries históricas de irradiação, que são bons estimadores da potencial produção energética por área.

A maior parte dos artigos que buscam caracterizar a dinâmica espaçotemporal do recurso solar avaliam dinâmicas das variações entre dois intervalos de tempo consecutivos das variáveis analisadas. Entretanto, no contexto de longas escalas de tempo, relacionado ao planejamento e à disponibilidade energética, esta avaliação não contém o significado prático que apresenta no contexto de escalas de tempo intradiárias, em que as taxas de rampa da geração fotovoltaica causam problemas na operação das redes elétricas. Levando isso em conta, a estratégia adotada neste trabalho para avaliar dinâmicas do recurso solar em diferentes escalas de tempo não se deu com a consideração de séries de variações entre intervalos de tempo consecutivos para diversas resoluções de integração, como é feito em estudos para escalas de tempo intradiárias. Foi proposta, então, uma análise baseada na decomposição de séries temporais, que divide a série histórica de irradiação em séries de tendência, interanual, sazonal e de resíduo mensal.

Cada conjunto de dados referente a uma componente das séries temporais foi avaliado separadamente, por meio de técnicas de estatística multivariada. Análise de compo-

mentes principais foi utilizada para a observação visual do comportamento das componentes sazonais e de tendência ao longo do território nacional. A maneira proposta de utilização desta ferramenta permitiu a visualização do comportamento destas componentes com poucas figuras, o que faz com que o entendimento de suas dinâmicas se dê de maneira clara e de fácil interpretação. Deve ser ressaltado que os resultados para a componente de tendência apresentaram certo ruído nos mapas dos fatores das componentes principais. É interessante buscar compreender se tal ruído está associado a algum padrão nos dados ou à forma em que foi definida a componente.

As componentes interanual e de resíduo mensal foram avaliadas por meio de análise de partições. Neste contexto, foram exploradas as maneiras em que são feitas as escolhas dos números de partições na literatura técnica da área, e discutido que a estratégia adotada na maioria destes, baseada no uso de índices de validação interna do particionamento junto ao método do cotovelo, apresenta pouco significado prático. Também, é argumentado que os comportamentos destes índices sugerem que não haja estrutura de partições intrínseca aos dados em diversos casos. Assim, é proposta a utilização de um método mais simples, baseado na quantidade da informação mantida ao utilizar do particionamento.

As partições foram obtidas com o algoritmo PAM, e utilizadas avaliando-se os pares de medoides. Para isso foram estudados os comportamentos dos coeficientes de correlação de Pearson para cada par, visando identificar locais com correlações negativas, que representam complementariedades, ou próximas de zero, que representam baixas ocorrências simultâneas de desvios. Também, foram avaliadas outras estatísticas, como a covariância e os propostos coeficientes de variação de pares simples e ótimos. As diferenças de resultados das análises com a correlação e o proposto OPCV ressaltam a utilidade do proposto índice, que leva em conta tanto as complementariedades quanto a variação local. Destaca-se que correlações mais baixas (ou negativas) não necessariamente representam variabilidades globais menores, medidas pelos propostos coeficientes de variação de pares.

Na análise das componente de resíduo mensal, a metodologia deste trabalho possibilitou alcançar resultados interessantes, como o referente à boa performance da região do Mato Grosso do Sul, que apresenta níveis de irradiação razoáveis, baixas variabilidades e complementariedade com a região do nordeste. A proposta da análise baseada em pontos de referência se mostrou muito interessante e complementar à análise baseada nas partições. Ao observar a dinâmica desta componente ao longo do tempo, foi verificado que, tanto as correlações, quanto os coeficientes de variação de pares apresentam variabilidade significativas entre um ano e outro, o que faz com que os valores obtidos nas análises sejam válidos a longo prazo, mas não necessariamente para todos os anos. Em relação à componente anual, os resultados se apresentaram bastante ruidosos, e é interessante ter cuidado ao interpretá-los, pois, devido às baixas amplitudes das oscilações desta componente, e ao número limitado de anos analisados, podem apresentar maiores erros. Para uma análise

mais aprofundada, seria interessante um contingente maior de dados. Também, pode ser feita a discussão de como a definição da componente de tendência afeta o comportamento da componente interanual.

O uso da ferramenta para a análise baseada em pontos de referência se mostrou muito interessante para complementar a análise baseada no particionamento. Pode-se até mesmo questionar o uso da análise de partições para estudos com objetivo de explorar comportamentos de pares, dado que a ferramenta desenvolvida consegue gerar resultados mais simples e interpretáveis.

Em relação à metodologia proposta para a decomposição das séries temporais, pode-se chamar a atenção a algumas questões que podem ser mais investigadas, como a maneira de obtenção da componente de tendência, que utiliza parâmetros que foram determinados de uma maneira um pouco arbitrária, baseada em observações visuais da componente. Esta definição também impacta os resultados da componente interanual. Ademais, é interessante a discussão de como foi definida a componente sazonal, considerada igual para todos os anos. Algumas modelagens consideram uma lenta dinâmica de variação desta componente ao longo do tempo, e é interessante explorar se esta consideração pode aumentar a qualidade da análise.

Dois bancos de dados de irradiação foram comparados, e diferenças significativas foram verificadas entre seus comportamentos. Contudo, foi verificado que mesmo com essas diferenças, dinâmicas de pares relativas à componente de resíduo mensal apresentam padrões com similaridades, o que traz certa corroboração aos resultados obtidos para esta componente.

5.1 Trabalhos Futuros

Como trabalhos futuros a serem realizados, que dão continuidade a este são sugeridos:

- Estudos utilizando e comparando outros dados, como os das estações meteorológicas do INMET, e dados intermediários gerados para a realização da segunda edição do Atlas Brasileiro de Energia Solar (como as séries de dados de irradiação). Estes dados, apesar de não estarem disponíveis na *internet*, pertencem a órgãos públicos, que podem estar dispostos a compartilhá-los.
- O desenvolvimento e o teste de modelos de previsão do comportamento a longo prazo do recurso solar. Neste caso, pode-se verificar se as técnicas apresentadas neste trabalho, como a decomposição das séries temporais e a análise de partições podem ser úteis para melhorar o desempenho dos modelos.

- A expansão da metodologia proposta neste trabalho para o estudo de outros recursos renováveis, como o eólico e o hidráulico, considerando as interações entre as diferentes fontes e suas dinâmicas espaçotemporais. Também, cabe a exploração da dinâmica da irradiação direta e a geração de usinas solares considerando diferentes aspectos. Ainda, é interessante avaliar essas interações associadas aos comportamentos dos consumidores. Neste contexto, o estudo associado às dinâmicas de modos climáticos pode ser interessante.
- A modelagem financeira de efeitos das variabilidades do recurso solar a longo prazo, visando a quantificação dos benefícios que a redução destas podem trazer ao sistema.
- A exploração das dinâmicas das componentes das séries de irradiação no contexto de avaliação de portfólios maiores, explorando a variabilidade de cenários e comparando-os. Neste contexto, também é interessante a aplicação de métodos de otimização multiobjetivo, buscando-se portfólios ótimos.

5.2 Produtos do Trabalho de Mestrado

Ao longo do período de trabalho no mestrado foram redigidos os seguintes artigos de congresso:

COSTA, A. M. G. d.; LOPES, B. M.; UTURBEY, W. Mapeamento do Potencial de Geração Solar Fotovoltaica no Brasil - Uma Abordagem Preliminar. In: *Anais do VII Congresso Brasileiro de Energia Solar*. Gramado, RS: [s.n.], 2018. Disponível em: <<http://anaiscbens.emnuvens.com.br/cbens/article/view/699/699>>.

COSTA, A. M. G. d.; UTURBEY, W. Voltage regulation strategies for distributed PV generation. In: *2017 IEEE Manchester PowerTech*. Manchester, United Kingdom: IEEE, 2017. ISBN 978-1-5090-4237-1. Disponível em: <<http://ieeexplore.ieee.org/document/7981067/>>.

Também, está em fase final de desenvolvimento, um programa feito em R, com interface de usuário, que utiliza das metodologias apresentadas neste trabalho e permite uma exploração mais dinâmica e abrangente dos resultados apresentados neste documento. Os bancos de dados foram tratados para utilização neste trabalho também estão disponíveis e podem ser requeridos ao autor.

REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. *Resolução Normativa Número 482*. 2012. Disponível em: <<http://www2.aneel.gov.br/cedoc/ren2012482.pdf>>. Citado na página 13.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. *Resolução Normativa Número 687*. 2015. Disponível em: <<http://www2.aneel.gov.br/cedoc/ren2015687.pdf>>. Citado na página 14.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. *Editais de Geração: Leilões de Geração*. 2018. Disponível em: <<http://www.aneel.gov.br/geracao4>>. Citado 3 vezes nas páginas 16, 17 e 31.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. *Micro e Minigeração Distribuída: Outorgas e Registros de Geração*. 2018. Disponível em: <<https://bit.ly/29qs6Cy>>. Citado 4 vezes nas páginas 13, 14, 15 e 16.

AGUIAR, L. M. et al. Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting. *Renewable Energy*, v. 97, p. 599–610, nov. 2016. ISSN 09601481. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0960148116305390>>. Citado na página 22.

ANJOS, P. S. dos et al. Long-term correlations and cross-correlations in wind speed and solar radiation temporal series from Fernando de Noronha Island, Brazil. *Physica A: Statistical Mechanics and its Applications*, v. 424, p. 90–96, abr. 2015. ISSN 03784371. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0378437115000059>>. Citado na página 26.

ARBELAITZ, O. et al. An extensive comparative study of cluster validity indices. *Pattern Recognition*, v. 46, n. 1, p. 243–256, jan. 2013. ISSN 00313203. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S003132031200338X>>. Citado na página 33.

BELLINI, E. *Saudi Arabia's 300 MW solar tender may conclude with lowest bid ever*. 2017. Disponível em: <<https://www.pv-magazine.com/2017/10/04/saudi-arabias-300-mw-solar-tender-may-conclude-with-lowest-bid-ever/>>. Citado na página 13.

BETT, P. E.; THORNTON, H. E. The climatological relationships between wind and solar energy supply in Britain. *Renewable Energy*, v. 87, p. 96–110, mar. 2016. ISSN 09601481. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0960148115303591>>. Citado na página 23.

BEVERIDGE, S.; NELSON, C. R. A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'. p. 24, 1981. Citado 2 vezes nas páginas 19 e 24.

BRENT, R. P. *Algorithms for Minimization without Derivatives*. Englewood Cliffs, New Jersey: Courier Corporation, 1973. ISBN 978-0-486-41998-5. Citado na página 37.

BRYCE, R. et al. Consequences of neglecting the interannual variability of the solar resource: A case study of photovoltaic power among the Hawaiian Islands. *Solar Energy*, v. 167, p. 61–75, jun. 2018. ISSN 0038092X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X18303414>>. Citado na página 18.

- CLEVELAND, R. B. et al. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, v. 6, n. 1, p. 3–73, 1990. Disponível em: <<http://www.nniem.ru/file/news/2016/stl-statistical-model.pdf>>. Citado na página 30.
- CLEVELAND, W. S.; DEVLIN, S. J. Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American statistical association*, Taylor & Francis, v. 83, n. 403, p. 596–610, 1988. Citado na página 30.
- CONTRERAS, J. et al. Arima models to predict next-day electricity prices. *IEEE transactions on power systems*, IEEE, v. 18, n. 3, p. 1014–1020, 2003. Citado na página 25.
- COSTA, A. M. G. d.; LOPES, B. M.; UTURBEY, W. Mapeamento do Potencial de Geração Solar Fotovoltaica no Brasil - Uma Abordagem Preliminar. In: *Anais do VII Congresso Brasileiro de Energia Solar*. Gramado, RS: [s.n.], 2018. Disponível em: <<http://anaiscbens.emnuvens.com.br/cbens/article/view/699/699>>. Citado na página 80.
- COSTA, A. M. G. d.; UTURBEY, W. Voltage regulation strategies for distributed PV generation. In: *2017 IEEE Manchester PowerTech*. Manchester, United Kingdom: IEEE, 2017. ISBN 978-1-5090-4237-1. Disponível em: <<http://ieeexplore.ieee.org/document/7981067/>>. Citado na página 80.
- DAVID, M.; ANDRIAMASOMANANA, F. H. R.; LIANDRAT, O. Spatial and Temporal Variability of PV Output in an Insular Grid: Case of Reunion Island. *Energy Procedia*, v. 57, p. 1275–1282, 2014. ISSN 18766102. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1876610214014842>>. Citado na página 21.
- DAVY, R. J.; TROCCOLI, A. Continental-scale spatial optimisation of a solar irradiance monitoring network. *Solar Energy*, v. 109, p. 36–44, nov. 2014. ISSN 0038092X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0038092X14004034>>. Citado na página 22.
- DEETJEN, T. A. et al. Modeling the optimal mix and location of wind and solar with transmission and carbon pricing considerations. *Renewable Energy*, v. 120, p. 35–50, maio 2018. ISSN 09601481. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0960148117312636>>. Citado na página 22.
- ELA, E. et al. *Impacts of variability and uncertainty in solar photovoltaic generation at multiple timescales*. [S.l.], 2013. Citado na página 18.
- EMPRESA DE PESQUISA ENERGÉTICA. *Informe - Leilões de Energia Nova de 2018*. Rio de Janeiro, 2018. Disponível em: <http://www.epe.gov.br/sites-pt/sala-de-imprensa/noticias/Documents/Informe%20Leil%c3%b5es%202018_Final.pdf>. Citado na página 16.
- ENGELAND, K. et al. Space-time variability of climate variables and intermittent renewable electricity production – A review. *Renewable and Sustainable Energy Reviews*, v. 79, p. 600–617, nov. 2017. ISSN 13640321. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1364032117306822>>. Citado 3 vezes nas páginas 18, 21 e 22.
- FERREIRA, E. B. et al. *Atlas Brasileiro Energia Solar*. Segunda edição. São José Dos Campos - SP: Instituto Nacional de Pesquisas Espaciais, 2017. ISBN 978-85-17-00089-8. Disponível em: <http://ftp.cptec.inpe.br/labren/publ/livros/Atlas_Brasileiro_Energia_Solar_2a_Edicao.pdf>. Citado na página 25.

FU, R. et al. *US solar photovoltaic system cost benchmark: Q1 2017*. [S.l.], 2017. Citado na página 13.

GIL, V. et al. Analysis of interannual variability of sunshine hours and precipitation over Peninsular Spain. *Renewable Energy*, v. 83, p. 680–689, nov. 2015. ISSN 09601481. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0960148115003626>>. Citado na página 22.

GUEYMARD, C. A.; WILCOX, S. M. Assessment of spatial and temporal variability in the US solar resource from radiometric measurements and predictions from models using ground-based or satellite data. *Solar Energy*, v. 85, n. 5, p. 1068–1084, maio 2011. ISSN 0038092X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0038092X11000855>>. Citado na página 22.

GUTIÉRREZ, C. et al. A multi-step scheme for spatial analysis of solar and photovoltaic production variability and complementarity. *Solar Energy*, v. 158, p. 100–116, dez. 2017. ISSN 0038092X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X17308241>>. Citado 8 vezes nas páginas 18, 23, 24, 33, 34, 35, 36 e 46.

HINKELMAN, L. M. Differences between along-wind and cross-wind solar irradiance variability on small spatial scales. *Solar Energy*, v. 88, p. 192–203, fev. 2013. ISSN 0038092X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X12004021>>. Citado na página 21.

HINKELMAN, L. M.; SCHAEFFER, N. Relating solar resource and its variability to weather and climate across the northwestern United States. *Solar Energy*, v. 157, p. 966–978, nov. 2017. ISSN 0038092X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0038092X1730645X>>. Citado na página 18.

HOFF, T. E.; PEREZ, R. Quantifying PV power Output Variability. *Solar Energy*, v. 84, n. 10, p. 1782–1793, out. 2010. ISSN 0038092X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0038092X10002380>>. Citado na página 21.

HOFF, T. E.; PEREZ, R. Modeling PV fleet output variability. *Solar Energy*, v. 86, n. 8, p. 2177–2189, ago. 2012. ISSN 0038092X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0038092X11004154>>. Citado 2 vezes nas páginas 21 e 23.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, Warwick & York, v. 24, n. 6, p. 417, 1933. Citado na página 38.

JEREZ, S. et al. Spatio-temporal Complementarity between Solar and Wind Power in the Iberian Peninsula. *Energy Procedia*, v. 40, p. 48–57, 2013. ISSN 18766102. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1876610213016019>>. Citado na página 22.

JEREZ, S.; TRIGO, R. M. Time-scale and extent at which large-scale circulation modes determine the wind and solar potential in the Iberian Peninsula. *Environmental Research Letters*, v. 8, n. 4, p. 044035, dez. 2013. ISSN 1748-9326. Disponível em: <<http://stacks.iop.org/1748-9326/8/i=4/a=044035?key=crossref.bd78f653acd03b2f65bd9245c21f41f2>>. Citado na página 23.

JONES, G.; BOUAMANE, L. “Power from Sunshine”: A Business History of Solar Energy. *Harvard Business School Working Paper*, v. 12-105, p. 89, maio 2012. Citado na página 13.

- KAPLANI, E.; KAPLANIS, S.; MONDAL, S. A spatiotemporal universal model for the prediction of the global solar radiation based on Fourier series and the site altitude. *Renewable Energy*, v. 126, p. 933–942, out. 2018. ISSN 09601481. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0960148118304129>>. Citado na página 22.
- KARIUKI, B. W.; SATO, T. Interannual and spatial variability of solar radiation energy potential in Kenya using Meteosat satellite. *Renewable Energy*, v. 116, p. 88–96, fev. 2018. ISSN 09601481. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0960148117309321>>. Citado na página 22.
- KAUFMAN, L.; ROUSSEEUW, P. J. *Finding Groups in Data: An Introduction to Cluster Analysis*. first edition. Hoboken, New Jersey: John Wiley and Sons, Inc, 1990. ISBN 0-471-73578-7. Citado na página 33.
- KAUFMAN, L.; ROUSSEEUW, P. *Clustering by means of medoids*. North-Holland, 1987. Disponível em: <https://www.researchgate.net/publication/243777819_Clustering_by_Means_of_Medoids>. Citado 2 vezes nas páginas 24 e 33.
- KRAKAUER, N.; COHAN, D. Interannual Variability and Seasonal Predictability of Wind and Solar Resources. *Resources*, v. 6, n. 3, p. 29, jul. 2017. ISSN 2079-9276. Disponível em: <<http://www.mdpi.com/2079-9276/6/3/29>>. Citado na página 23.
- LAVE, M.; KLEISSL, J. Solar variability of four sites across the state of Colorado. *Renewable Energy*, v. 35, n. 12, p. 2867–2873, dez. 2010. ISSN 09601481. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0960148110002314>>. Citado na página 21.
- LAVE, M.; RENO, M. J.; BRODERICK, R. J. Characterizing local high-frequency solar variability and its impact to distribution studies. *Solar Energy*, v. 118, p. 327–337, ago. 2015. ISSN 0038092X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X15002881>>. Citado na página 18.
- LUIZ, E. W. et al. Analysis of intra-day solar irradiance variability in different Brazilian climate zones. *Solar Energy*, v. 167, p. 210–219, jun. 2018. ISSN 0038092X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X18303505>>. Citado 2 vezes nas páginas 25 e 26.
- MILLS, A.; WISER, R. Implications of wide-area geographic diversity for short-term variability of solar power. 2010. Disponível em: <<http://eetd.lbl.gov/EA/EMP>>. Citado 2 vezes nas páginas 18 e 21.
- MINGOTI, S. A. *Análise de Dados Através de Métodos de Estatística Multivariada: Uma Amordagem Aplicada*. Belo Horizonte, Brasil: Editora UFMG, 2005. ISBN 85-7041-451-X. Citado na página 38.
- MINISTÉRIO DE MINAS E ENERGIA. *Nota Técnica Número 5/2017 - Aprimoramento do marco legal do setor elétrico*. 2017. Disponível em: <<https://bit.ly/2vgn4i4>>. Citado na página 17.
- MOHAMMADI, K.; GOUDARZI, N. Study of inter-correlations of solar radiation, wind speed and precipitation under the influence of El Niño Southern Oscillation (ENSO) in California. *Renewable Energy*, v. 120, p. 190–200, maio 2018. ISSN 09601481. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0960148117312739>>. Citado na página 23.

MURATA, A.; OHTAKE, H.; OOZEKI, T. Modeling of uncertainty of solar irradiance forecasts on numerical weather predictions with the estimation of multiple confidence intervals. *Renewable Energy*, v. 117, p. 193–201, mar. 2018. ISSN 09601481. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0960148117309813>>. Citado na página 22.

NUÑO, E. et al. Simulation of transcontinental wind and solar PV generation time series. *Renewable Energy*, v. 118, p. 425–436, abr. 2018. ISSN 09601481. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0960148117311400>>. Citado na página 22.

PEREIRA, E. B. et al. *Atlas Brasileiro de Energia Solar*. Primeira edição. São José Dos Campos - SP: Instituto Nacional de Pesquisas Espaciais, 2006. ISBN 978-85-17-00030-0. Disponível em: <http://ftp.cptec.inpe.br/labren/publ/livros/brazil_solar_atlas_R1.pdf>. Citado na página 25.

PEREZ, M. J.; FTHENAKIS, V. Impacts of long-timescale variability in solar resources at high PV penetrations: Quantification. In: *Photovoltaic Specialists Conference (PVSC), 2012 38th IEEE*. [S.l.]: IEEE, 2012. p. 002481–002486. Citado na página 22.

PEREZ, M. J.; FTHENAKIS, V. M. On the spatial decorrelation of stochastic solar resource variability at long timescales. *Solar Energy*, v. 117, p. 46–58, jul. 2015. ISSN 0038092X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0038092X1500208X>>. Citado 4 vezes nas páginas 18, 21, 22 e 46.

PEREZ, R.; HOFF, T. E. Solar Resource Variability. In: KLEISSL, J. (Ed.). *Solar energy forecasting and resource assessment*. First edition. Amsterdam ; Boston: Elsevier, AP, Academic Press in an imprint of Elsevier, 2013. p. 133 – 148. ISBN 978-0-12-397177-7. OCLC: ocn856889163. Citado na página 21.

PEREZ, R. et al. Short-term irradiance variability: Preliminary estimation of station pair correlation as a function of distance. *Solar Energy*, v. 86, n. 8, p. 2170–2176, ago. 2012. ISSN 0038092X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X12000928>>. Citado na página 21.

PEREZ, R. et al. Spatial and Temporal Variability of Solar Energy. *Foundations and Trends® in Renewable Energy*, v. 1, n. 1, p. 1–44, 2016. ISSN 2328-8892, 2328-8906. Disponível em: <<http://www.nowpublishers.com/article/Details/REN-006>>. Citado 3 vezes nas páginas 21, 25 e 29.

PINKER, R. T.; ZHANG, B.; DUTTON, E. G. Do Satellites Detect Trends in Surface Solar Radiation? *Science*, v. 308, 2005. Disponível em: <<http://science.sciencemag.org/content/308/5723/850>>. Citado na página 66.

POLO, J. et al. Spatial variability and clustering of global solar irradiation in Vietnam from sunshine duration measurements. *Renewable and Sustainable Energy Reviews*, v. 42, p. 1326–1334, fev. 2015. ISSN 13640321. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1364032114009459>>. Citado 2 vezes nas páginas 23 e 33.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2018. Disponível em: <<https://www.R-project.org/>>. Citado 3 vezes nas páginas 27, 28 e 31.

RAZA, M. Q.; NADARAJAH, M.; EKANAYAKE, C. On recent advances in PV output power forecast. *Solar Energy*, v. 136, p. 125–144, out. 2016. ISSN 0038-092X. Citado na página 22.

RENNÉ, D. et al. Results of solar resource assessments in the unep/swera project. In: CITESEER. *Proceedings of the Solar World Congress, Orlando, FL, International Solar Energy Soc.* [S.l.], 2005. Citado na página 28.

REYNOLDS, A. P. et al. Clustering rules: a comparison of partitioning and hierarchical clustering algorithms. *Journal of Mathematical Modelling and Algorithms*, Springer, v. 5, n. 4, p. 475–504, 2006. Citado na página 33.

SANTOS-ALAMILLOS, F. J. et al. Analysis of Spatiotemporal Balancing between Wind and Solar Energy Resources in the Southern Iberian Peninsula. *Journal of Applied Meteorology and Climatology*, v. 51, n. 11, p. 2005–2024, nov. 2012. ISSN 1558-8424, 1558-8432. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-11-0189.1>>. Citado na página 23.

SCHMELA, M. et al. *SolarPower Europe's Global Market Outlook 2017 - 2021*. Bruxelas, Bélgica, 2017. Citado na página 13.

SILVA, V. d. P. R. d. et al. Trends in solar radiation in NCEP/NCAR database and measurements in northeastern Brazil. *Solar Energy*, v. 84, n. 10, p. 1852–1862, out. 2010. ISSN 0038092X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X10002471>>. Citado na página 26.

SOUZA, J. L. d.; NICÁCIO, R. M.; MOURA, M. A. L. Global solar radiation measurements in Maceió, Brazil. *Renewable Energy*, v. 30, n. 8, p. 1203–1220, jul. 2005. ISSN 09601481. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0960148104003842>>. Citado na página 26.

STACKHOUSE, P. W. et al. POWER Release 8 (with GIS Applications) Methodology (Data Parameters, Sources, & Validation) Documentation. p. 89, jul. 2018. Citado 2 vezes nas páginas 27 e 68.

TRIPATHY, S. K. et al. Impact assessment of short-term variability of solar radiation in Rajasthan using SRR data. *Renewable and Sustainable Energy Reviews*, v. 78, p. 798–806, out. 2017. ISSN 13640321. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1364032117306500>>. Citado na página 21.

URQUHART, B.; SENGUPTA, M.; KELLER, J. Optimizing geographic allotment of photovoltaic capacity in a distributed generation setting: Optimizing geographic allotment of PV capacity. *Progress in Photovoltaics: Research and Applications*, v. 21, n. 6, p. 1276–1285, set. 2013. ISSN 10627995. Disponível em: <<http://doi.wiley.com/10.1002/pip.2334>>. Citado na página 22.

WALL, M. E.; RECHTSTEINER, A.; ROCHA, L. M. Singular value decomposition and principal component analysis. In: *A practical approach to microarray data analysis*. [S.l.]: Springer, 2003. p. 91–109. Citado na página 38.

WATANABE, T.; TAKAMATSU, T.; NAKAJIMA, T. Y. Evaluation of Variation in Surface Solar Irradiance and Clustering of Observation Stations in Japan. *Journal of Applied Meteorology and Climatology*, v. 55, n. 10, p. 2165–2180, out. 2016. ISSN 1558-8424, 1558-8432.

Disponível em: <<http://journals.ametsoc.org/doi/10.1175/JAMC-D-15-0227.1>>. Citado na página 23.

WEST, M. Time series decomposition. *Biometrika*, v. 84, n. 2, p. 489–494, 1997. Disponível em: <<https://academic.oup.com/biomet/article-abstract/84/2/489/23412>>. Citado 2 vezes nas páginas 19 e 24.

YANG, D.; DONG, Z. Operational photovoltaics power forecasting using seasonal time series ensemble. *Solar Energy*, v. 166, p. 529–541, maio 2018. ISSN 0038092X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X18301282>>. Citado na página 24.

YANG, D. et al. Analyzing big time series data in solar engineering using features and PCA. *Solar Energy*, v. 153, p. 317–328, set. 2017. ISSN 0038092X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0038092X17304796>>. Citado 2 vezes nas páginas 18 e 24.

YANG, D. et al. Forecasting of global horizontal irradiance by exponential smoothing, using decompositions. *Energy*, v. 81, p. 111–119, mar. 2015. ISSN 03605442. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0360544214013528>>. Citado na página 24.

ZAGOURAS, A.; INMAN, R. H.; COIMBRA, C. F. On the determination of coherent solar microclimates for utility planning and operations. *Solar Energy*, v. 102, p. 173–188, abr. 2014. ISSN 0038092X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0038092X14000395>>. Citado 5 vezes nas páginas 23, 24, 33, 34 e 46.

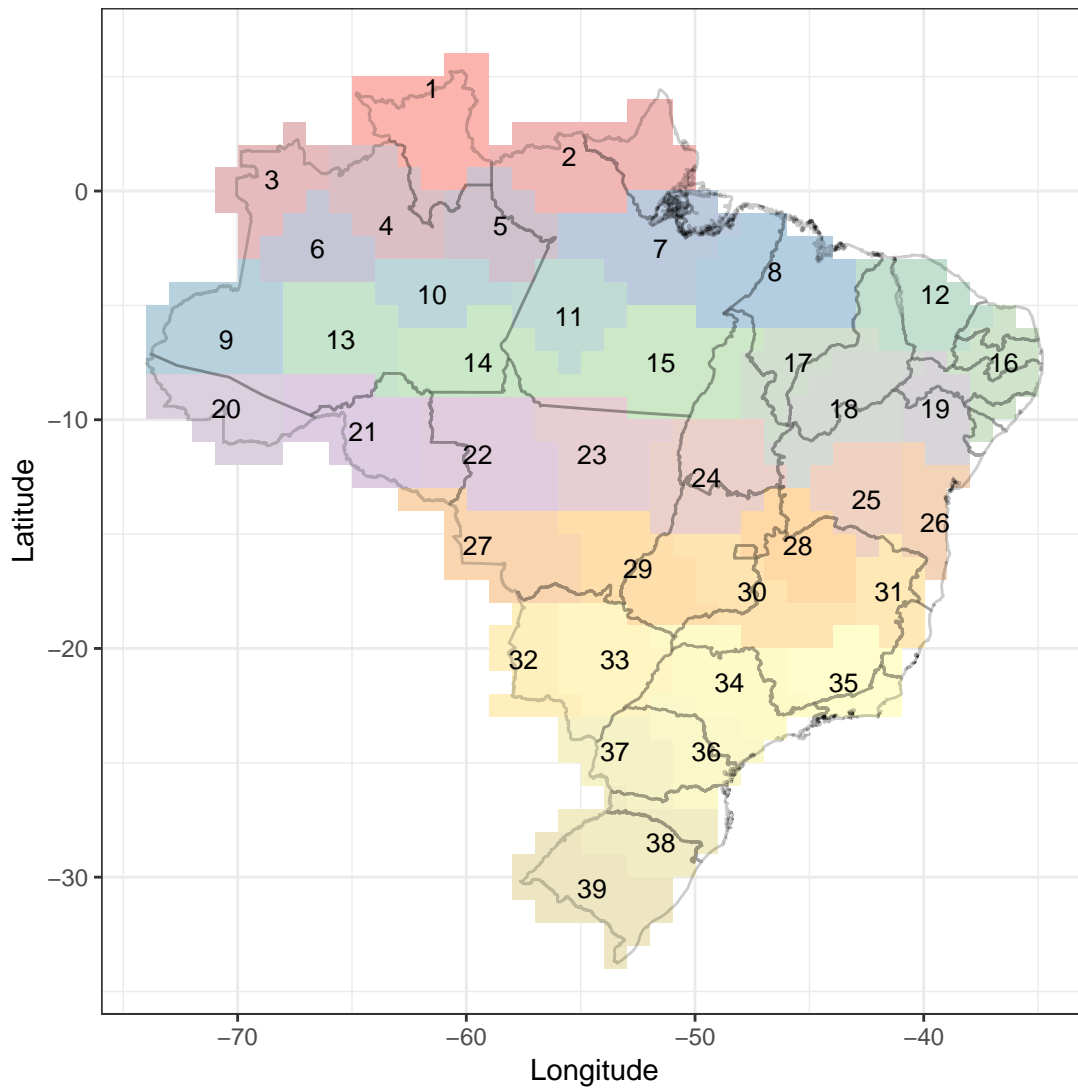
ZAGOURAS, A. et al. Determination of measuring sites for solar irradiance, based on cluster analysis of satellite-derived cloud estimations. *Solar Energy*, v. 97, p. 1–11, nov. 2013. ISSN 0038092X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0038092X13003150>>. Citado 3 vezes nas páginas 23, 24 e 33.

ZAGOURAS, A.; PEDRO, H. T.; COIMBRA, C. F. Clustering the solar resource for grid management in island mode. *Solar Energy*, v. 110, p. 507–518, dez. 2014. ISSN 0038092X. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0038092X14004836>>. Citado 4 vezes nas páginas 18, 23, 24 e 33.

Apêndices

APÊNDICE A – MAPA DOS AGRUPAMENTOS

Figura 47 – Partições Obtidas com a Metodologia Apresentada



Elaboração própria