

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE LORENA – EEL/USP

MATHEUS PATUSCO BASCUR

**UTILIZAÇÃO DE APACHE SPARK PARA CÁLCULO DE POTENCIAL ENERGÉTICO DE FONTE
DE ENERGIA RENOVÁVEL POR MUNICÍPIO**

LORENA

2021

MATHEUS PATUSCO BASCUR

**UTILIZAÇÃO DE APACHE SPARK PARA CÁLCULO DE POTENCIAL ENERGÉTICO DE FONTE
DE ENERGIA RENOVÁVEL POR MUNICÍPIO**

Projeto de Trabalho de Conclusão de Curso
apresentado na Escola de Engenharia de Lorena
EEL-USP como requisito obrigatório para a
conclusão do curso de Engenharia Física

Orientador: Prof^a. Dr^a. Rosa Ana Conte

LORENA

2021

NÃO AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, SERÁ DISPONIBILIZADO AUTOMATICAMENTE APÓS 2 ANOS DA PUBLICAÇÃO

Ficha catalográfica elaborada pelo Sistema Automatizado
da Escola de Engenharia de Lorena,
com os dados fornecidos pelo(a) autor(a)

Bascur, Matheus

Utilização de apache spark para cálculo de
potencial energético de fonte de energia renovável
por município / Matheus Bascur; orientadora Rosa
Conte. - Lorena, 2021.

50 p.

Monografia apresentada como requisito parcial
para a conclusão de Graduação do Curso de Engenharia
Física - Escola de Engenharia de Lorena da
Universidade de São Paulo. 2021

1. Energia limpa. 2. Energia eólica. 3. Energia
solar. 4. Apache spark. 5. Cloud pública. I. Título.
II. Conte, Rosa, orient.

A todos os meus professores, colegas e família que contribuíram para a realização deste trabalho

Agradecimentos

Agradeço a minha família, por ter me ajudado muito em todo o percurso de estudo e trabalho, e sempre me incentivar a ser melhor do que no dia anterior.

Aos meus amigos da faculdade, que me ajudaram e me apoiaram em toda essa grande batalha, em especial ao Douglas Sardisco e Matheus Bianco, que estávamos juntos desde o primeiro dia do curso.

A todos os professores que me ensinaram com garra e muito conhecimento, me direcionando a desenvolver da melhor maneira possível um conhecimento lógico e de resolução de problemas que hoje é fundamental na minha carreira.

Em especial ao Prof. Dr. Carlos Shigue, que desde o primeiro ano incentivou a curiosidade, pesquisa e que ensinou que o tempo é um grande aliado. Também a Prof. Dra. Rosa Ana Conte, que sempre cobrou excelência nas atividades a serem realizadas, que hoje é um ponto forte no ambiente de trabalho.

Agradeço aos meus colegas de profissão, em especial ao Rafael Teru, Rodolfo Simões e Sergio Patrocínio, que me apoiaram e me direcionaram para a carreira que estou percorrendo hoje, e que foi um dos motivos do tema desenvolvido dentro desse trabalho.

Agradeço também a todos que de alguma forma contribuíram para a realização deste estudo.

*“Com organização e tempo,
acha-se o segredo de fazer tudo
e bem-feito.”*

Pitágoras

RESUMO

PATUSCO, M. UTILIZAÇÃO DE APACHE SPARK PARA CÁLCULO DE POTENCIAL ENERGÉTICO DE FONTE DE ENERGIA RENOVÁVEL POR MUNICÍPIO. 2021. Número de folhas 50f. Monografia (Trabalho de Conclusão de Curso) – Escola de Engenharia de Lorena, Universidade de São Paulo, Lorena, 2021

A partir de bases disponibilizadas pelo Sistema Nacional de Organização de Dados Ambientais (SONDA), onde contém dados sobre irradiação solar e velocidade dos ventos em diferentes regiões do Brasil será fornecido um sistema de fácil utilização para mostrar ao usuário o potencial energético do seu município e a área necessária para construção de uma usina de energia eólica/solar para atender todos os munícipes. Para isso é utilizada a ferramenta Apache Spark, que é amplamente utilizada no mundo por conta de sua alta performance com big data e uso de recursos de cloud pública para manter a disponibilidade. Ao final é comparado o resultado de três diferentes regiões com os mapas fornecidos pelo SONDA e avaliamos se para estes casos recursos escolhidos foram os ideais.

Palavras-Chave: Energia limpa; Energia eólica; Energia solar, Apache Spark, Cloud Pública

ABSTRACT

PATUSCO, M. USE OF APACHE SPARK TO CALCULATE THE ENERGY POTENTIAL OF A RENEWABLE ENERGY SOURCE PER MUNICIPALITY. 2021. Number of sheets 50s. Monograph (Course Conclusion Paper) - School of Engineering of Lorena, University of São Paulo, Lorena, 2021

Based on databases provided by the National System for the Organization of Environmental Data (SONDA), which contain data on solar irradiation and wind speed in different regions of Brazil, an easy-to-use system will be provided to show the user the energy potential of their municipality and area needed for the construction of a wind/solar power plant to serve all residents. For this, an Apache Spark tool is used, which is widely used in the world because of its high performance with big data and use of public cloud resources to maintain availability. At the end, the results of three different regions are compared with the maps provided by SONDA and we evaluate whether for these chosen resources they were ideal.

Keywords: Clean energy; wind energy; Solar Power, Apache Spark, Public Cloud

Lista de figuras

Figura 2.1: Evolução da população mundial por continente no próximo século	13
Figura 2.2: Redução de CO ₂ emitido na produção de energia a partir das políticas apresentadas.	14
Figura 2.3: Matriz elétrica brasileira em maio de 2017 de acordo com dados disponíveis no Banco de Informações de Geração da Agência Nacional de Energia Elétrica.....	15
Figura 2.4: Representação das estações do ano em decorrência da inclinação da Terra em relação a posição do Sol.....	16
Figura 2.5: Representação das estações do ano em decorrência da inclinação da Terra em relação a posição do Sol.....	16
Figura 2.6: Espectro eletromagnético da radiação solar. A área entre as curvas representa a atenuação da radiação solar incidente durante sua passagem através da atmosfera. Para comparação é apresentada uma curva de emissão de um corpo negro a 6000 K.....	17
Figura 2.7: Irradiância global horizontal (G): é a taxa de energia total por unidade de área incidente numa superfície horizontal.	18
Figura 2.8: Irradiância direta normal (G _n): também conhecida como DNI, é a taxa de energia por unidade de área proveniente diretamente do Sol que incide perpendicularmente à superfície....	19
Figura 2.9: Sazonalidade do potencial de geração solar fotovoltaica para os 12 meses do ano, em termos de rendimento energético anual para todo o Brasil (medido em kWh/kWp.mês no perfil de cores), admitindo uma taxa de desempenho de 80% para geradores fotovoltaicos fixos.....	20
Figura 2.10: Edifício Sede da Eletrosul em Florianópolis-SC, com geradores solares fotovoltaicos integrados na cobertura do prédio e nos estacionamentos, somando 1 MWp de potência.....	21
Figura 2.11: mapa do total anual de irradiação solar direta normal com destaque para regiões com maior potencial de aproveitamento de energia solar concentrada.....	22
Figura 2.12: Coletores solares parabólicos em Marrocos.....	23
Figura 2.13: Usina Gemasolar na Espanha.....	23
Figura 2.14: Perfil vertical da velocidade do vento desde a superfície até a altura do vento geotrópico. O comprimento da rugosidade (z ₀) é a altura onde a velocidade é nula.....	24
Figura 2.15: Modelo de rugosidade, sobreposto ao relevo sombreado.....	25
Figura 2.16: Velocidade média trimestral do vento a 50m de altura.....	26
Figura 2.17: Representação do distanciamento recomendado entre as turbinas instaladas, sendo D o diâmetro do rotor.....	28
Figura 2.18: Curva de energia gerada em um ano em condições de velocidade do vento constante em 100% do tempo e sem perdas.....	29
Figura 3.1: Representação simples do funcionamento do Apache Spark.....	30
Figura 3.2: Exemplificação de uma API.....	31
Figura 3.3: Logo do framework Streamlit.....	32

Figura 3.4: dados presentes no banco de dados do SONDA.....	34
Figura 3.5: Metadata dos dados presentes no banco de dados do SONDA.....	34
Figura 3.6: Estações pertencentes a rede de coleta de dados SONDA	35
Figura 3.7: Pastas alocadas dentro do bucket S3, na pasta solar.....	36
Figura 3.8: Arquivo criado com os dados sobre população e área dos municípios.....	36
Figura 3.9: Configuração EMR utiliza para executar os códigos de tratamento do banco de dados.....	38
Figura 3.10: Alta arquitetura do projeto desenvolvido.	39
Figura 4.1: Base de dados disponibilizado está com valores fixos na posição onde deveria informar velocidade do vento a 50m em Brasília.	39
Figura 4.2: Imagem da tela de consulta do usuário.....	40
Figura 4.3: Valores obtidos em tabela Excel para comparação.....	40

Lista de equações

Equação 2.1.....	27
Equação 2.2.....	27
Equação 2.3.....	27
Equação 3.1.....	37

Sumário

1 Introdução.....	12
2 Revisão da literatura.....	13
2.1 Consumo de energia elétrica.....	13
2.2 Energias renováveis.....	14
2.2.1 Energia Solar.....	15
2.2.1.1 Energia Solar Fotovoltaica.....	20
2.2.1.2 Energia Solar Térmica.....	22
2.2.2 Energia Eólica.....	24
3 Metodologia.....	29
3.1 Recursos.....	29
3.1.1 Apache Spark.....	29
3.1.2 API Rest.....	31
3.1.3 Streamlit.....	31
3.1.4 Cloud Pública AWS.....	32
3.1.4.1 Instância EC2.....	32
3.1.4.2 Amazon S3.....	33
3.1.4.3 EMR.....	33
3.2 Banco de dados SONDA.....	33
Desenvolvimento.....	35
4 Resultados e discussão.....	39
5 Conclusão.....	41
6 Referencias.....	42
7 Apêndices.....	44
Apêndice I.....	44
Apêndice II.....	47
Apêndice III.....	50
Apêndice IV.....	50

1 INTRODUÇÃO

A partir do dia 3 de novembro de 2020, grande parte da população do estado brasileiro Amapá ficou sem acesso à energia elétrica regular por 22 dias, com alto impacto socioeconômico. Esse problema ocorreu em decorrência de uma explosão seguida de incêndio que ocorreu nos três transformadores da principal subestação de energia do estado, ou seja, o problema em uma única subestação impactou 90% da população do estado.

Em maio de 2021, por conta de um período de estiagem e volumes de chuva insuficientes o Brasil está próximo de um período de racionamento do fornecimento de energia e alta nos preços das contas de energia. Além disso existe a perspectiva de um período mais severo de estiagem nos próximos meses na região sudeste e centro-oeste do país, piorando ainda mais a situação e as expectativas.

Caso a população tivesse acesso a outras fontes de energia, esse problema teria um impacto menor. Mas para essas regiões, qual outra fonte de energia seria capaz de atender a demanda aproveitando ao máximo o potencial energético disponibilizado? Sabendo das necessidades do país, essa pergunta é de grande importância.

Além dessa preocupação, também precisamos nos preocupar com o meio ambiente e como essa geração de energia pode o impactar, então é necessário oferecer opções de energias renováveis e consideradas limpas. Dentro do território brasileiro essas opções contemplam principalmente a energia solar e a eólica.

Para responder as perguntas apresentadas acima, vamos utilizar dados disponibilizados de maneira pública pelo INPE sobre radiação e velocidade dos ventos, aplicando as fórmulas de eficiência energética dos geradores de energia solar e eólica disponibilizados no mercado, permitindo o acesso a essas informações de maneira simples em um site.

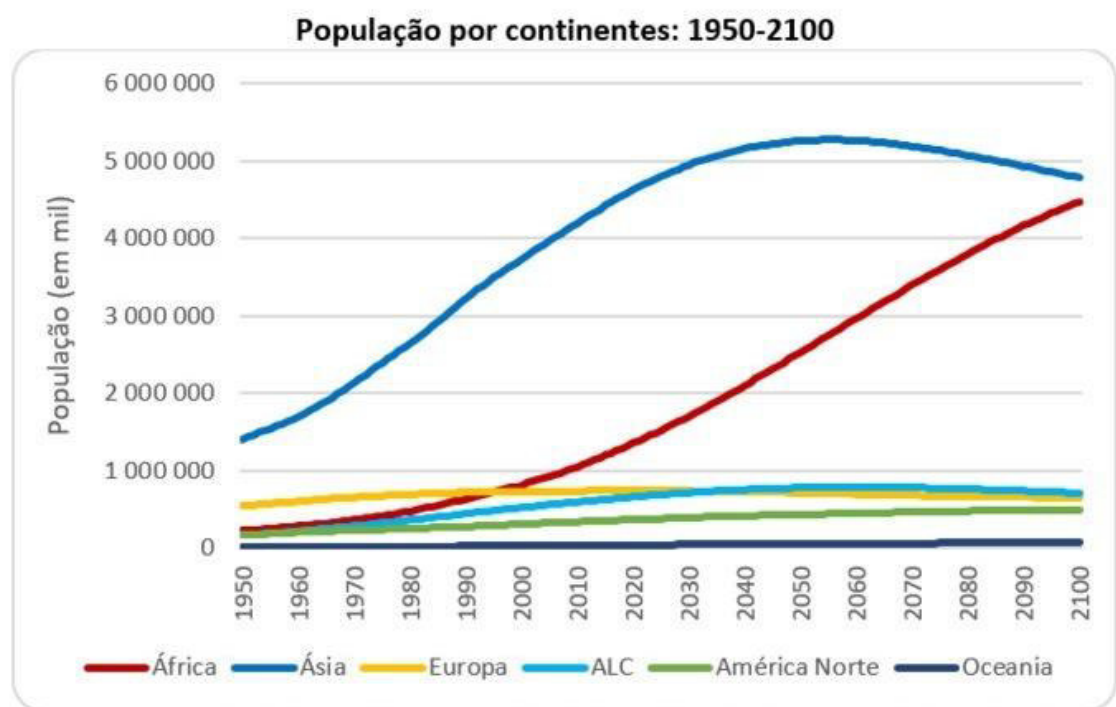
Para armazenar esses dados e trabalhar sobre eles, será utilizado uma das principais tecnologias relacionadas a big data disponibilizada atualmente que é o Apache Spark, com as informações armazenadas dentro do cloud publica AWS, com isso, além de disponibilizar as informações de maneira fácil, iremos otimizar o tratamento dos dados e aplicação das fórmulas.

2 REVISÃO DA LITERATURA

2.1 Consumo de energia elétrica

O consumo de energia elétrica tem a tendência de aumentar de maneira significativa nas próximas décadas, em contrapartida, existe a expectativa que o crescimento da população mundial alcance um platô (com exceção do continente africano), como é possível visualizar na figura 2.1.

Figura 2.1: Evolução da população mundial por continente no próximo século



Fonte: WORLD Population Prospects 2019

Esses dados parecem contraditórios se analisados sem a perspectiva da mudança de cultura global. Com o avanço cada vez maior dos equipamentos tecnológicos, cada vez mais estamos substituindo equipamentos analógicos, equipamentos mecânicos e até as relações pessoais por equipamentos eletroeletrônicos e atividades virtuais. Como um grande exemplo, podemos citar os bancos, que estão investindo cada vez mais para que os clientes não precisem sequer pisar nas agências para resolver problemas, e em grandes centros, não há mais a necessidade de se utilizar dinheiro físico no dia a dia.

Nesse estudo iremos utilizar a média de consumo energético de uma família de tamanho médio (3 pessoas) para realizarmos cálculos. Além disso será considerado que alcançamos o platô populacional, ou seja, não será considerado a variável de crescimento populacional.

Será considerado os seguintes equipamentos elétricos utilizados nas residências para estimar a necessidade de consumo de energia elétrica para uma residência com três pessoas:

- Chuveiro elétrico: 72 kWh/mês;
- Geladeira de uma porta: 25,2 kWh/mês;
- Micro-ondas: 13,98 kWh/mês;

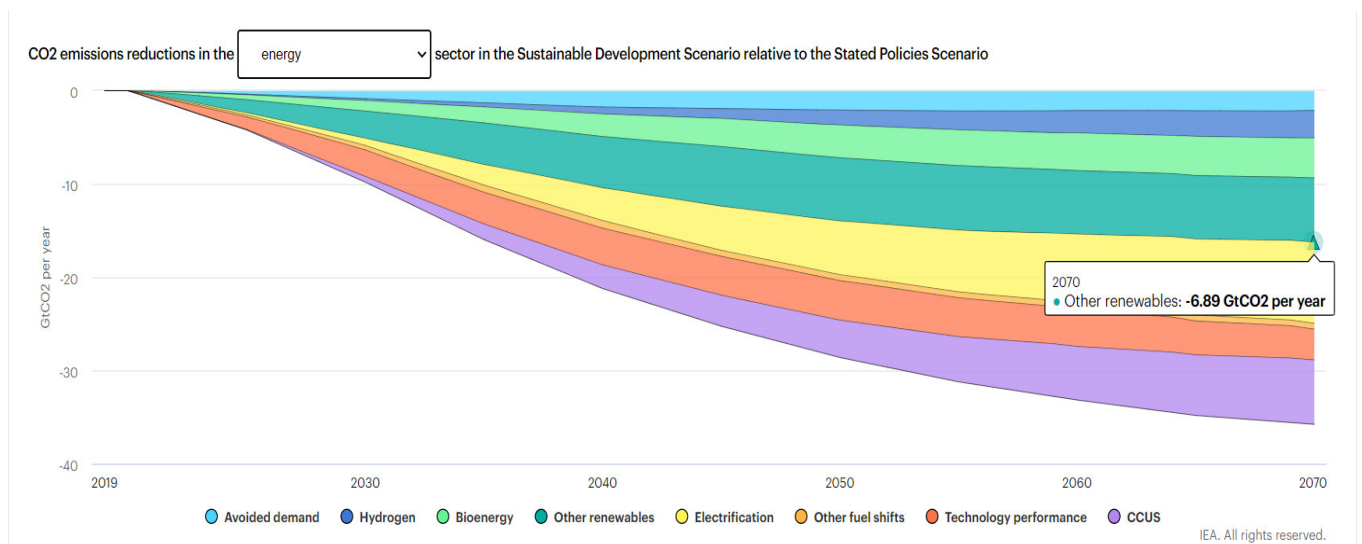
- Ferro elétrico a vapor: 7,2 kWh/mês;
- Lavadora de roupas: 3 kWh/mês.
- Televisão 40 polegadas: 8,6 kWh/mês

Com essas informações, consideramos o consumo médio mensal de energia é de 130kWh/mês

2.2 Energias renováveis

As fontes de energias renováveis são assim chamadas por gerarem energia a partir de recursos naturais que são repostos com uma taxa igual ou superior ao consumo durante a produção de energia. Além disso a opção por fontes de energias renováveis está diretamente ligada a questões de preservação ambiental. O mundo hoje tenta reduzir ao máximo as fontes de energia que geram principalmente o gás carbônico CO₂, e como é possível ver no gráfico 2.2 o uso de energias renováveis poderá representar 15% de toda redução de gás carbônico gerado durante a produção de energia.

Figura 2.2: Redução de CO₂ emitido na produção de energia a partir das políticas apresentadas

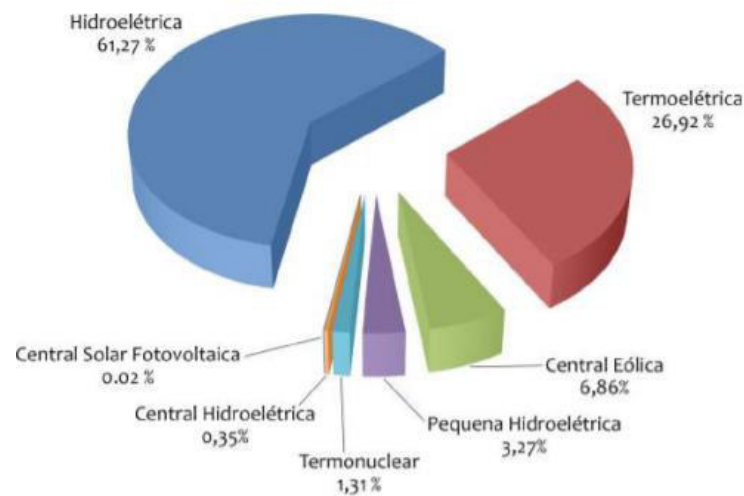


Fonte: ENERGY Technology Perspectives 2020

Também é possível obter a informação a partir da figura 2.2, até 2070 as políticas ambientais preveem a redução de 40GtCO₂ emitidos por ano, sendo a utilização de fontes de energia renováveis responsáveis pela redução de quase 6.9GtCO₂ emitido.

Mas, no sentido contrário ao que está sendo buscado, o Brasil está atendendo a demanda crescente de energia elétrica junto com os recentes problemas de secas nas grandes usinas elétricas que atendem o território nacional com fontes térmicas não renováveis, e hoje elas já representam mais de um quarto de todas as fontes de energia instaladas no Brasil, como é possível visualizar na figura 2.3.

Figura 2.3: Matriz elétrica brasileira em maio de 2017 de acordo com dados disponíveis no Banco de Informações de Geração da Agência Nacional de Energia Elétrica.



Fonte: ATLAS Brasileiro de Energia SOLAR

Mas além da energia termoelétrica, podemos destacar as energias provindas de central Solar Fotovoltaica, Central Hidroelétrica, Pequena Hidroelétrica e central eólica, que juntas representam hoje um pouco mais de 10% das fontes de energia no Brasil.

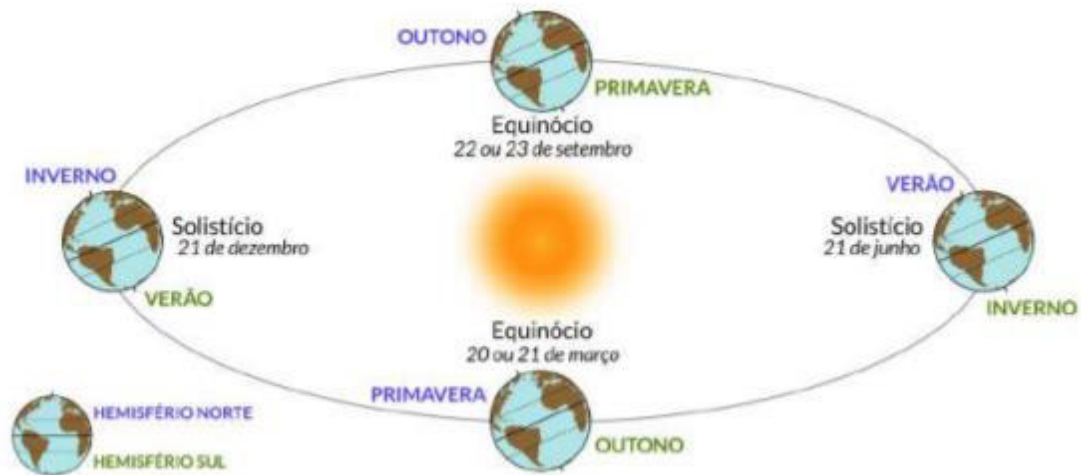
2.2.1 Energia Solar

Considerando a definição de energia renovável, a energia solar não é uma opção, mas pode ser considerada uma fonte de energia inesgotável, uma vez que é estimado que o Sol continuará a fornecer sua energia para a Terra por milhares de anos, além de emitir uma altíssima potência de energia constante a bilhões de anos na ordem de $3,86 \cdot 10^{26}$ W.

Mas a maneira como essa energia solar alcança a Terra para poder ser aproveitada em processos naturais e tecnológicos dependem intrinsecamente de conceitos astronômicos, como por exemplo a distância entre a órbita da Terra e o Sol, que varia entre 1,47.108 km e 1,52.108 km e o valor médio da irradiância solar igual a 1.366 W/m^2 é definido como a constante solar (NATIONAL Renewable Energy Laboratory, 2017).

Outro conceito muito importante para entender o aproveitamento energético da energia solar é a variabilidade temporal, que depende de dois ciclos: o ciclo anual e o diário. A variabilidade durante a variabilidade do ciclo anual é decorrente da inclinação de $23,45^\circ$ do eixo da Terra, e como consequência apresentam-se as chamadas estações do ano que são descritas na figura 2.4.

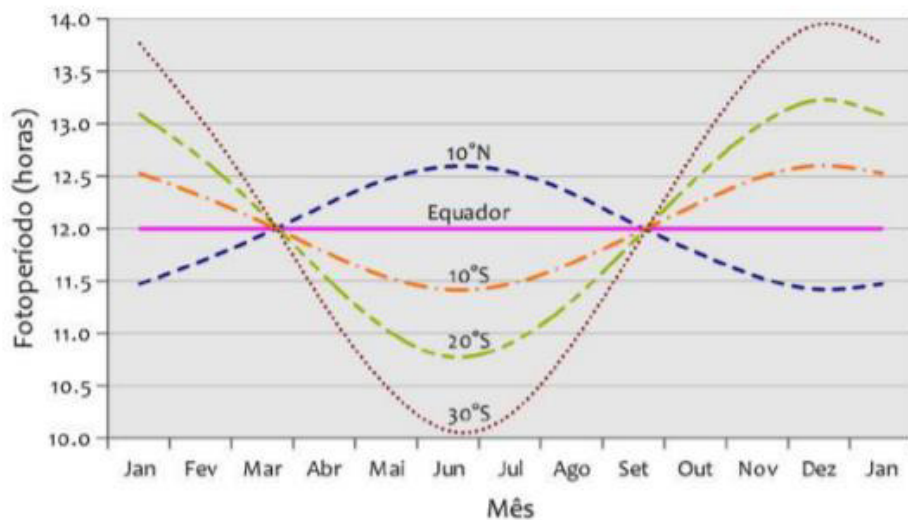
Figura 2.4: Representação das estações do ano em decorrência da inclinação da Terra em relação a posição do Sol



Fonte: ATLAS Brasileiro de Energia SOLAR

A duração do dia além de depender do movimento de translação, também está relacionada com as estações do ano e a posição latitudinal do observador. Isso fica evidente na figura 2.5.

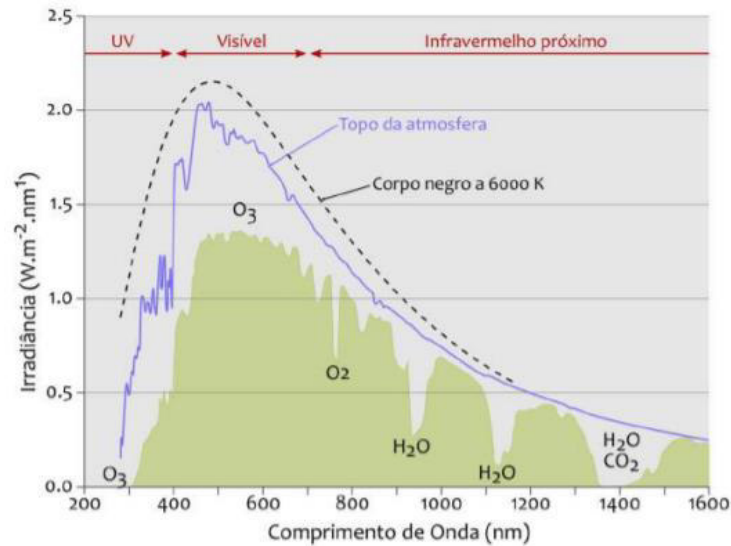
Figura 2.5: Representação das estações do ano em decorrência da inclinação da Terra em relação a posição do Sol



Fonte: ATLAS Brasileiro de Energia SOLAR

Esses aspectos explicados se limitam a estabelecer uma explicação da variabilidade de radiação solar até a camada mais externa da atmosfera. A partir do momento que a radiação solar inicia o percurso do início da atmosfera até o solo da Terra, temos uma série de processos de absorção e espalhamento de acordo com reação das diferentes amplitudes de onda com os compostos químicos presentes que atenuam a radiação. É possível entender o impacto desses compostos químicos a partir do gráfico na figura 2.5.

Figura 2.6: Espectro eletromagnético da radiação solar. A área entre as curvas representa a atenuação da radiação solar incidente durante sua passagem através da atmosfera. Para comparação é apresentada uma curva de emissão de um corpo negro a 6000 K (linha tracejada).

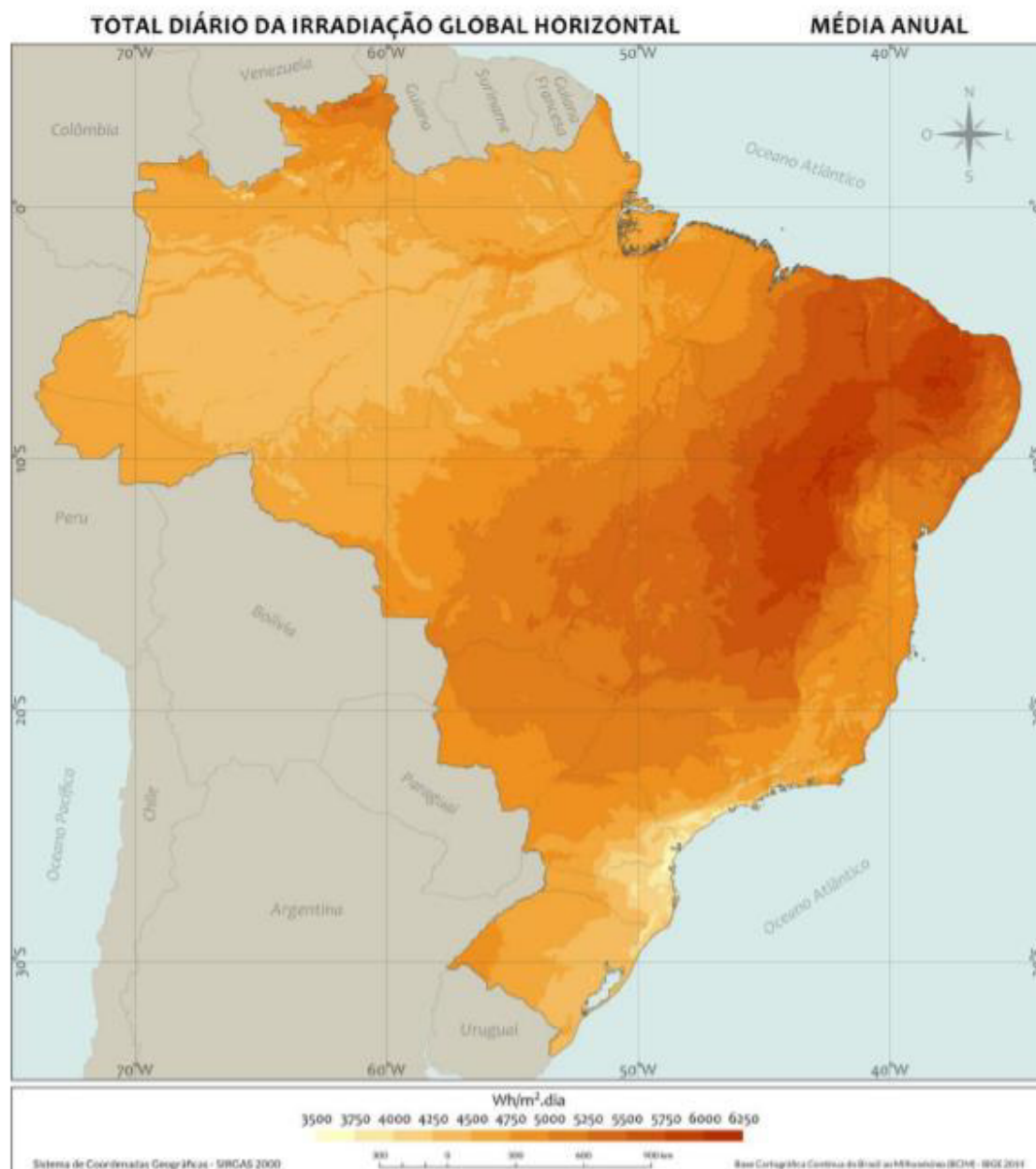


Fonte: ATLAS Brasileiro de Energia SOLAR

Um composto muito importante a ser observado é a água (H_2O). Ela tem alto impacto de atenuação da irradiação solar principalmente nas faixas de amplitude de onda de infravermelho. Além disso, sua composição na atmosfera tem grande variabilidade dependendo da região, como por exemplo, em áreas desérticas ou semiáridas, a água compõe aproximadamente 1% da atmosfera, enquanto em áreas de floresta tropical compõem 4%. E dentro do território brasileiro temos essas regiões.

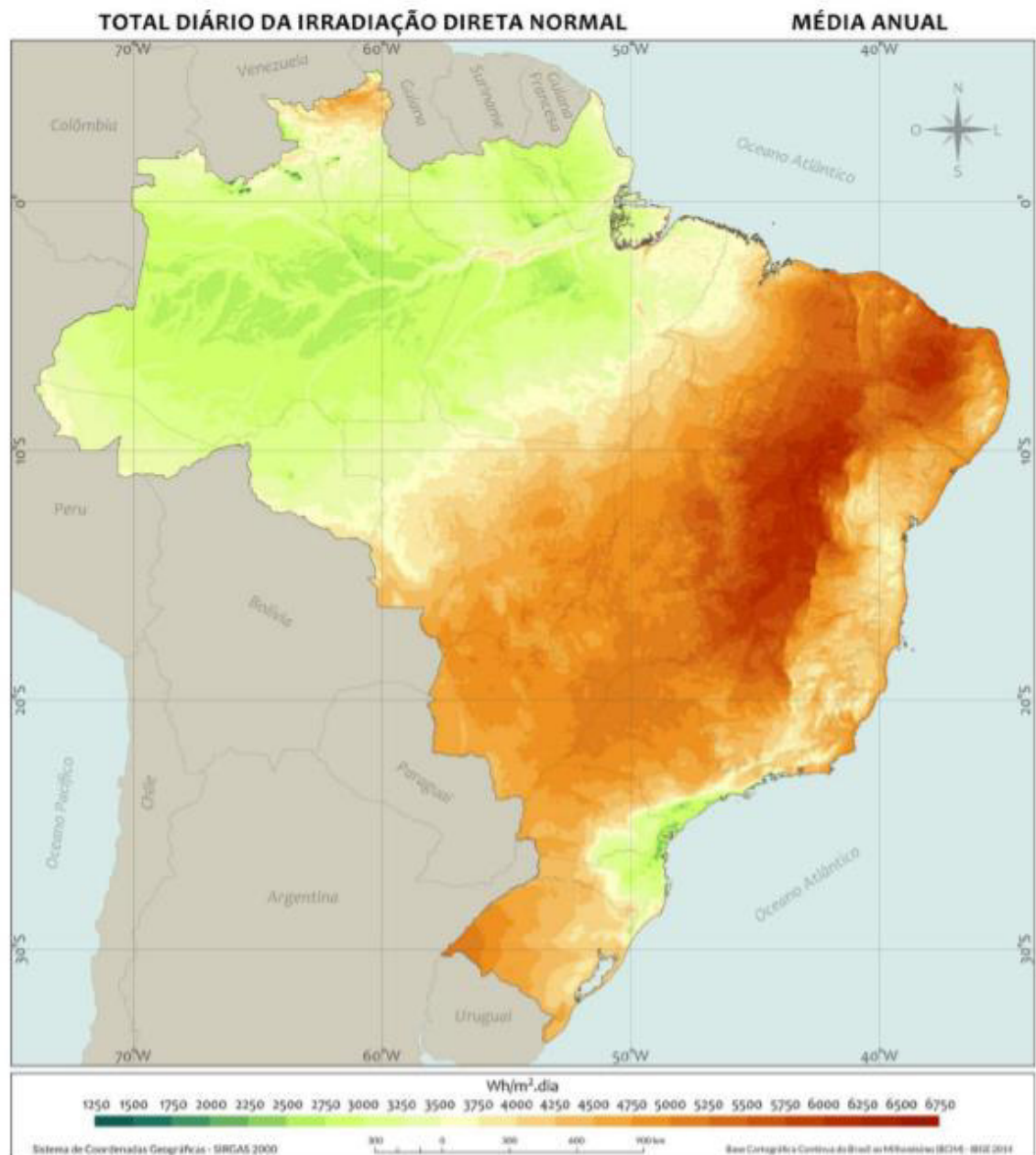
A irradiância solar que incide em uma superfície é composta por suas componentes direta e difusa. A irradiância solar direta apresenta direção de incidência na linha imaginária entre a superfície e o Sol e representa a parcela que não sofreu os processos radiativos de absorção e espalhamento que ocorrem na atmosfera. A componente difusa engloba a radiação proveniente de todas as demais direções que são decorrentes dos processos de espalhamento. A seguir teremos a representação gráfica de dois tipos de irradiação solar em território brasileiro. São essas representações: Irradiação Solar Global Horizontal (figura 2.7) e Irradiação Solar Direta Normal (figura 2.8).

Figura 2.7: Irradiância global horizontal (G): é a taxa de energia total por unidade de área incidente numa superfície horizontal.



Fonte: ATLAS Brasileiro de Energia SOLAR

Figura 2.8: Irradiância direta normal (Gn): também conhecida como DNI, é a taxa de energia por unidade de área proveniente diretamente do Sol que incide perpendicularmente à superfície.



Fonte: ATLAS Brasileiro de Energia SOLAR

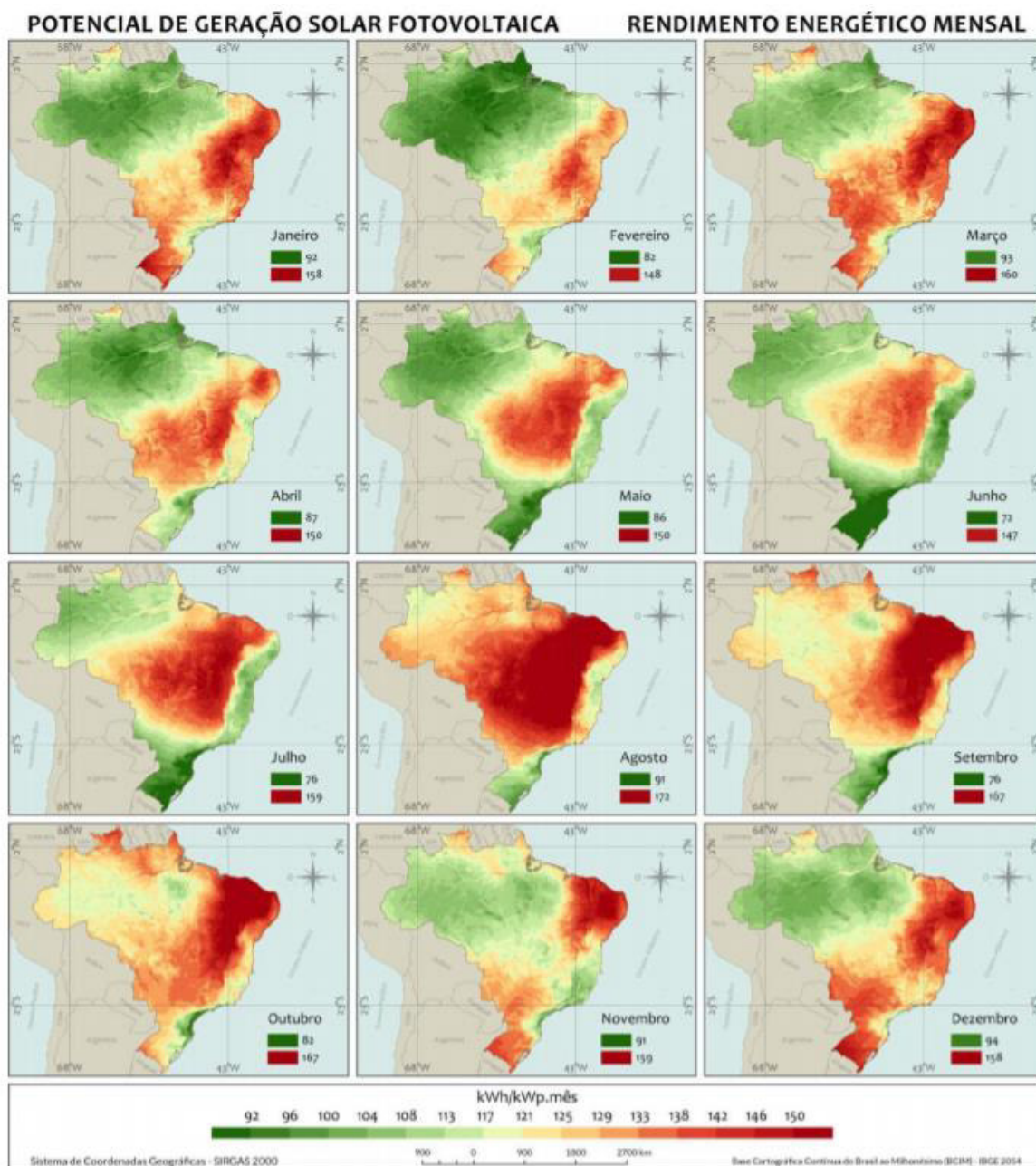
Com essas informações conseguimos identificar pontos a serem considerados para identificar a capacidade de aproveitamento energético da radiação Solar em determinada região, que são a posição latitudinal e as condições atmosféricas. Nos próximos tópicos vamos especificar melhor os dois tipos de tecnologia que se utilizam da radiação solar para obtenção de energia, a energia solar fotovoltaica e a energia solar térmica.

Nesse trabalho, iremos discutir a geração de energia solar a partir de usinas de médio e grande porte, para facilitar a comparação com as outras fontes de energia renovável aqui discutidas.

2.2.1.1 Energia Solar Fotovoltaica

Antes de entendermos como é calculado a eficiência e geração de energia solar fotovoltaica, vamos entender o potencial da utilização dessa energia em território brasileiro. Essa análise será baseada na figura 2.9 onde é mostrado o rendimento energético mensal máximo (medido em kWh de energia elétrica gerada por ano para cada kWp de potência fotovoltaica instalada e considerando uma eficiência energética de 80% para facilitar os cálculos)

Figura 2.9: Sazonalidade do potencial de geração solar fotovoltaica para os 12 meses do ano, em termos de rendimento energético anual para todo o Brasil (medido em kWh/kWp.mês no perfil de cores), admitindo uma taxa de desempenho de 80% para geradores fotovoltaicos fixos.



Fonte: ATLAS Brasileiro de Energia SOLAR

Esse valor de 80% de capacidade foi escolhido porque é a capacidade média real dos painéis solares instalados nos primeiros 25 anos de funcionamento desde que a instalação seja feita corretamente. Essa diferença é decorrente da definição de kWp, que é a produção de kW em um ambiente com temperatura de 25° e massa de ar de 1.5, fatores que dificilmente representam as condições reais de funcionamento.

É importante notar que nos meses de verão, principalmente de dezembro a março, a geração é máxima nos estados do extremo Sul e Sudeste do Brasil e coincide com os máximos de demanda registrados pelo Operador Nacional do Sistema – ONS para essas regiões.

No mundo, mais de 90% das placas solares são produzidas com a mesma composição, sendo o principal componente a célula fotovoltaica. Essa célula é feita a partir do semicondutor Silício, e quando recebe a radiação solar, seus elétrons se agitam e recebem energia o suficiente para se desprenderem do átomo, com isso eles iniciam uma movimentação dentro da célula fotovoltaica. Quando essa movimentação apresenta uma direção majoritária é gerado uma diferença de potencial e com isso é gerado a energia elétrica.

Nesse processo, grande parte da energia absorvida é utilizada para desprender o elétron, e a energia liberada pelo elétron nesse processo não é equivalente a energia recebida, com isso a eficiência energética das placas solares não ultrapassam 20%, sendo que essa informação foi considerada para a geração da figura 2.9, ou seja, está dentro dos 80% de capacidade.

Hoje as placas solares fotovoltaicas têm uma previsão de funcionamento com o máximo de capacidade de aproximadamente 25 anos, mas esse tempo pode ser reduzido em decorrência da exposição à radiação ultravioleta, porque a resistência interna das células pode aumentar devido à infiltração de contaminantes por rachaduras no material que encapsula sob ação dessa radiação (Jordan e Kurtz, 2012). Com isso, problemas relacionados a camada de ozônio podem impactar diretamente o tempo de vida das placas solares fotovoltaicas.

Figura 2.10: Edifício Sede da Eletrosul em Florianópolis-SC, com geradores solares fotovoltaicos integrados na cobertura do prédio e nos estacionamentos, somando 1 MWp de potência.



Fonte: ATLAS Brasileiro de Energia SOLAR

As usinas fotovoltaicas centralizadas de grande porte apresentam duas configurações típicas: sistemas fixos e sistemas com seguimento da trajetória aparente do Sol, sendo essa segunda sendo a mais utilizada em território brasileiro. Esse sistema pode ser estruturado em um ou dois

eixos, e para utilização do trabalho, vamos considerar a estrutura com apenas um eixo, que otimiza a absorção da energia solar durante o dia, mas não tem sua posição alterada para acompanhar as mudanças em decorrência da angulação da Terra.

Para os cálculos iremos considerar a radiação solar direta normal por m^2 disponibilizada na base de dados

Em 2019 a Usina Solar de São Gonçalo realizou a ampliação do seu potencial energético em 133MW, com um custo aproximado de 422 milhões de reais. Realizando a divisão temos um valor aproximado de mais de R\$3000/W.

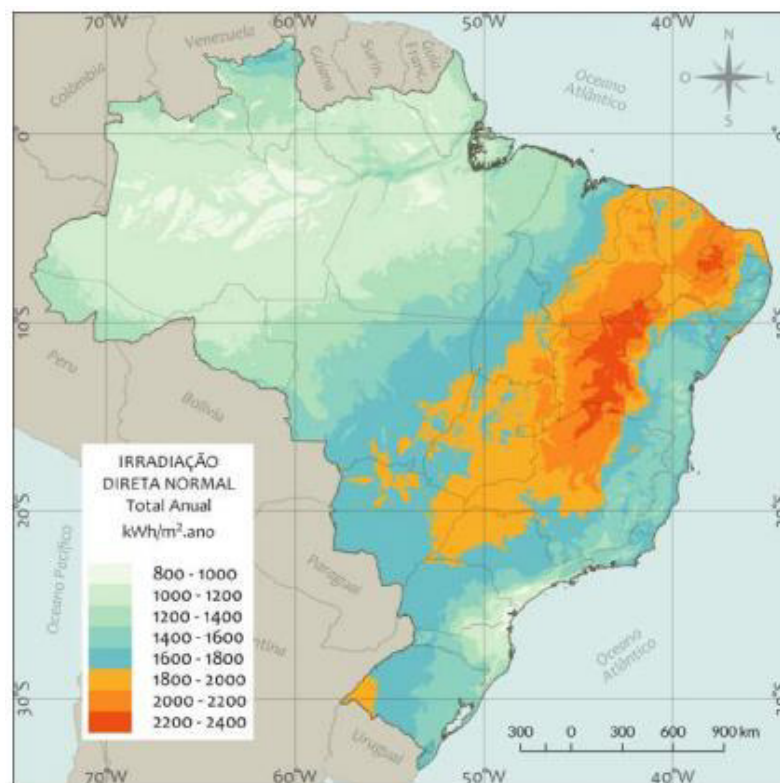
2.2.1.2 Energia Solar Térmica

Hoje quando falamos sobre a utilização de energia solar térmica, normalmente relacionamos com o aquecimento de água para temperatura de até 100°C para uso doméstico. Mas esse não é o foco desse estudo. Nosso objetivo é avaliar a utilização da energia solar térmica para geração de energia elétrica, mundialmente conhecido como “Concetrated Solar Power” (CSP) e como vamos nos referenciar a partir desse ponto.

Diferente da energia solar fotovoltaica, a CSP só é eficiente quando existe grande irradiação solar direta, em geral, regiões com disponibilidade de irradiação solar anual acima de 2000 kWh/m^2 e baixa nebulosidade. (Schlecht e Meyer; 2012).

No Brasil, essas características ambientais se apresentam quase que exclusivamente na região do semiárido nordestino, como é possível visualizar na figura 2.11.

Figura 2.11: mapa do total anual de irradiação solar direta normal com destaque para regiões com maior potencial de aproveitamento de energia solar concentrada.



Fonte: ATLAS Brasileiro de Energia SOLAR

Sobre as instalações desse sistema, hoje existem dois principais grupos: Concentração linear e concentração bidimensional.

As instalações de concentração linear aumentam a densidade de fluxo da irradiância sobre uma linha onde localiza-se o dispositivo absorvedor de energia solar (ATLAS Brasileiro de Energia SOLAR). A tecnologia mais utilizada nas usinas termoelétricas que se utilizam da energia solar são os coletores parabólicos, que é possível observar uma instalação deles em Marrocos na figura 2.12.

Figura 2.12: Coletores solares parabólicos em Marrocos



Fonte: ATLAS Brasileiro de Energia SOLAR

As instalações de concentração bidimensional concentram os raios solares sobre um ponto aumentando a densidade de fluxo e irradiância e permitindo temperaturas de trabalho mais altas e consequentemente, melhor eficiência de conversão do ciclo termodinâmico para geração de potência (ATLAS Brasileiro de Energia SOLAR).

Esse segundo tipo de instalação é comprovadamente eficiente, e para esse estudo, vamos nos basear nos números referentes a instalação da usina Gemasolar na Espanha, a primeira usina desse tipo a conseguir gerar energia por 24h e está representada na figura 2.12.

Figura 2.13: Usina Gemasolar na Espanha



Fonte: USINA ESPANHOLA

A usina Gemasolar ocupa uma área aproximada de 185 hectares, aproximadamente 185 campos de futebol. O custo de instalação foi de US\$ 325 milhões, com uma geração média anual de 110 GWh e fornece energia atualmente para 27500 residências na Espanha (USINA SOLAR DA ESPANHA). Convertendo para reais em uma cotação de 5,60, temos que a construção custou aproximadamente R\$25000/Wh produzido.

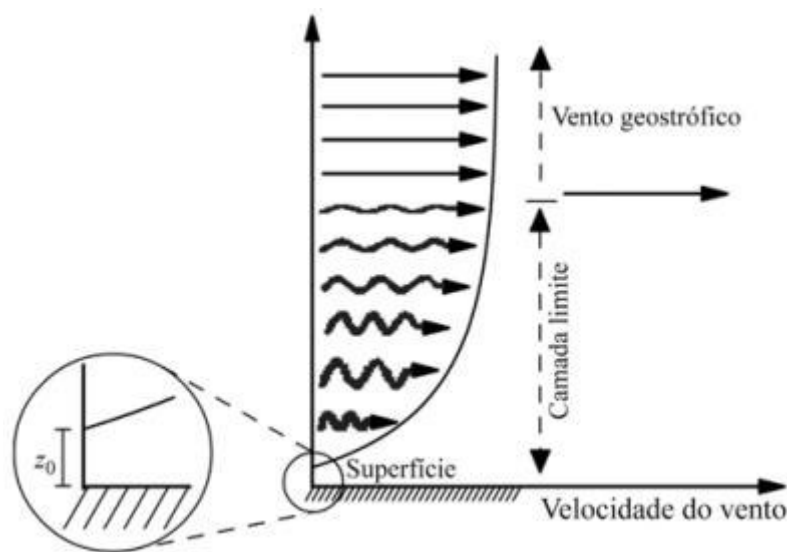
Diferente da energia solar fotovoltaica, não encontramos valores para instalação individual, por módulos, então por fator comparativo, quando formos comparar o potencial de energia disponibilizado, vamos levar em conta a área do município e sua população, considerando assim a energia gerada por uma usina e não residencial.

2.2.2 Energia Eólica

Energia eólica é a utilização da energia cinética do vento para geração de energia. Por definição o vento é o movimento das parcelas do ar, sendo comum ser considerado apenas o deslocamento horizontal e paralelo a superfície. E para a geração de energia eólica essa definição é suficiente.

A velocidade do vento nas proximidades da superfície está diretamente relacionada a rugosidade da superfície, uma propriedade física que molda o escoamento do ar e a turbulência do fluxo. O impacto da distância da superfície pode ser mais bem visualizado na representação gráfica da figura 2.14.

Figura 2.14: Perfil vertical da velocidade do vento desde a superfície até a altura do vento geotrófico. O comprimento da rugosidade (z_0) é a altura onde a velocidade é nula.



Fonte: O APROVEITAMENTO DA ENERGIA EÓLICA

Na figura 2.15 percebemos que o vento geostrófico tem a maior velocidade possível dentro da troposfera. Esse vento é um vento horizontal, não acelerado, que sopra ao longo de trajetórias retilíneas e apresenta um equilíbrio entre a força gradiente de pressão e a força de Coriolis. (O VENTO GEOSTRÓFICO) Sendo força de Coriolis é uma força aparente decorrente da consideração da segunda lei de Newton em um sistema de coordenadas não-inercial.

Sabendo dessa relação entre a energia cinética do vento com a rugosidade da superfície, observando o mapa do Brasil com um modelo de rugosidade da figura 2.16 já podemos ter uma visão das regiões com maior potencial para geração de energia eólica.

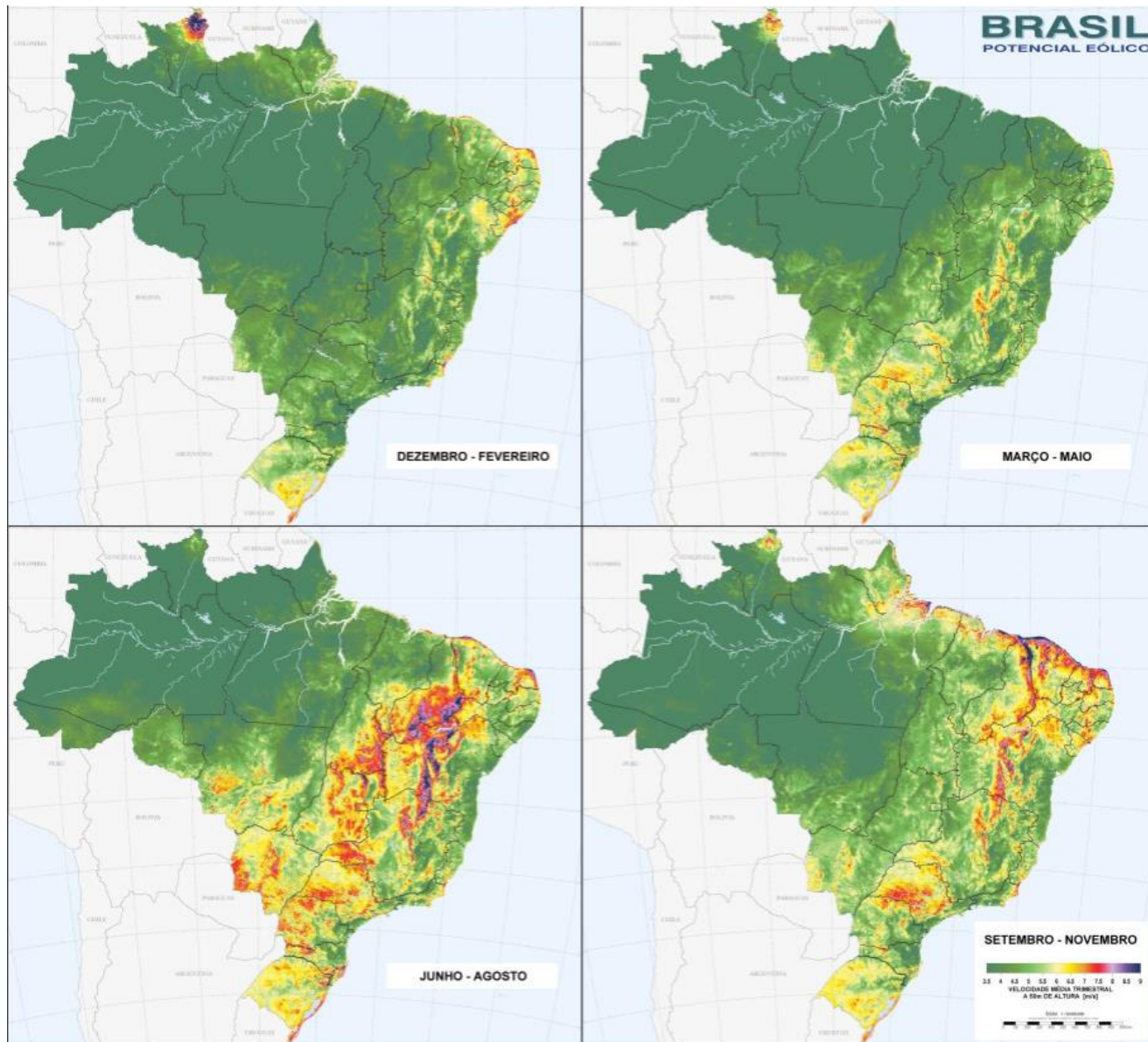
Figura 2.15: Modelo de rugosidade, sobreposto ao relevo sombreado



Fonte: Atlas do Potencial Eólico Brasileiro

Na figura 2.16 temos a distribuição da velocidade média dos ventos por trimestre, e comparando com a figura 2.15, podemos relacionar diretamente a rugosidade da superfície com a velocidade dos ventos.

Figura 2.16: Velocidade média trimestral do vento a 50m de altura



Fonte: Atlas do Potencial Eólico Brasileiro

Uma vez entendido como é distribuído o vento em território brasileiro, vamos entender melhor como podemos utilizar disso para geração de energia.

A utilização da conversão de energia cinética dos ventos para outra energia vem de milhares de anos atrás, quando foram construídos os primeiros redemoinhos para moagem de grãos e bombeamento de água. Já a utilização dessa energia para a conversão de energia elétrica é bem mais recente, sendo os primeiros estudos publicados a cerca de 150 anos. (G.M. Joselin Herbert, S. Iniyar, E. Sreevalsan and S. Rajapadian, Renew. Sustain. Energ. Rev. 11, 1117 (2007).

Mas a fabricação e implantação das primeiras turbinas eólicas é ainda mais recente, a partir de 1980, na Dinamarca, hoje o maior fabricante de turbinas eólicas no mundo. As primeiras turbinas construídas tinham uma capacidade máxima de geração de 55kW, bem abaixo do potencial máximo de geração de energia eólica das turbinas atuais que é de 2500kW. (O APROVEITAMENTO DA ENERGIA EÓLICA)

A potência das turbinas atuais em relação a velocidade do vento perpendicular que se choca com suas pás é descrita pela equação (2.1)

$$2.1 \quad P = \frac{1}{2} \rho A v^3$$

Onde P é a potência do vento, A é a área das pás onde o vento entra em contato, v é a velocidade do vento e ρ é a densidade do vento, que pode ser descrita pela equação (2.2)

$$2.2 \quad \rho(z) = \frac{p_0}{RT} e^{\frac{gz}{RT}}$$

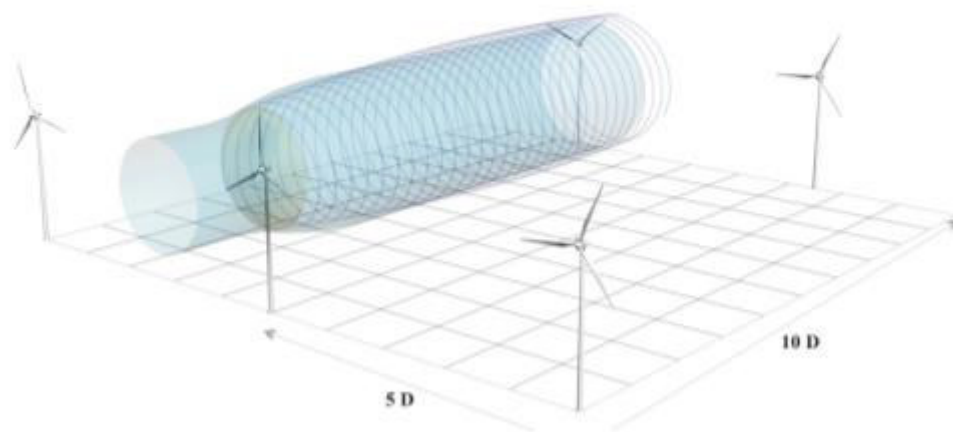
onde $\rho(z)$ é a densidade do ar (em kg/m³) em função da altitude z, p_0 é a pressão atmosférica ao nível do mar (em kg/m³), T é a temperatura do ar (em K), g é a aceleração da gravidade (em m/s²) e R é a constante específica do ar (em J/K mol).

Porém essas equações representam um valor teórico, o real valor de potencial máximo chega a ser de 42% do valor calculado (R. Gasch and J. Tvele (eds), Wind Power Plants: Fundamentals, Design, Construction and Operation (Solarpraxis AG, Alemanha, 2002)) por conta da eficiência do conjunto gerador /transmissão η e o coeficiente aerodinâmico de potência do rotor C_p , que adicionados a equação 2.1 temos a equação 2.3.

$$2.3 \quad P = \frac{1}{2} \rho A v^3 \eta C_p$$

O contato com as pás da turbina faz com que o vento perca parte de sua energia cinética, porém logo é recuperada quando se encontra com as massas que não se chocaram com as pás, com o movimento circular das pás forma-se uma esteira helicoidal de vórtices atrás, o que poderia impactar diretamente e negativamente uma turbina que estivesse logo atrás, por isso existe uma recomendação de distanciamento entre as turbinas que é da ordem de 10 vezes o diâmetro D do rotor, se instalada a jusante, e 5 vezes D, se instalada ao lado, em relação ao vento predominante (Atlas do Potencial Eólico Brasileiro) Na figura 2.17 temos uma representação sobre esse distanciamento.

Figura 2.17: Representação do distanciamento recomendado entre as turbinas instaladas, sendo D o diâmetro do rotor

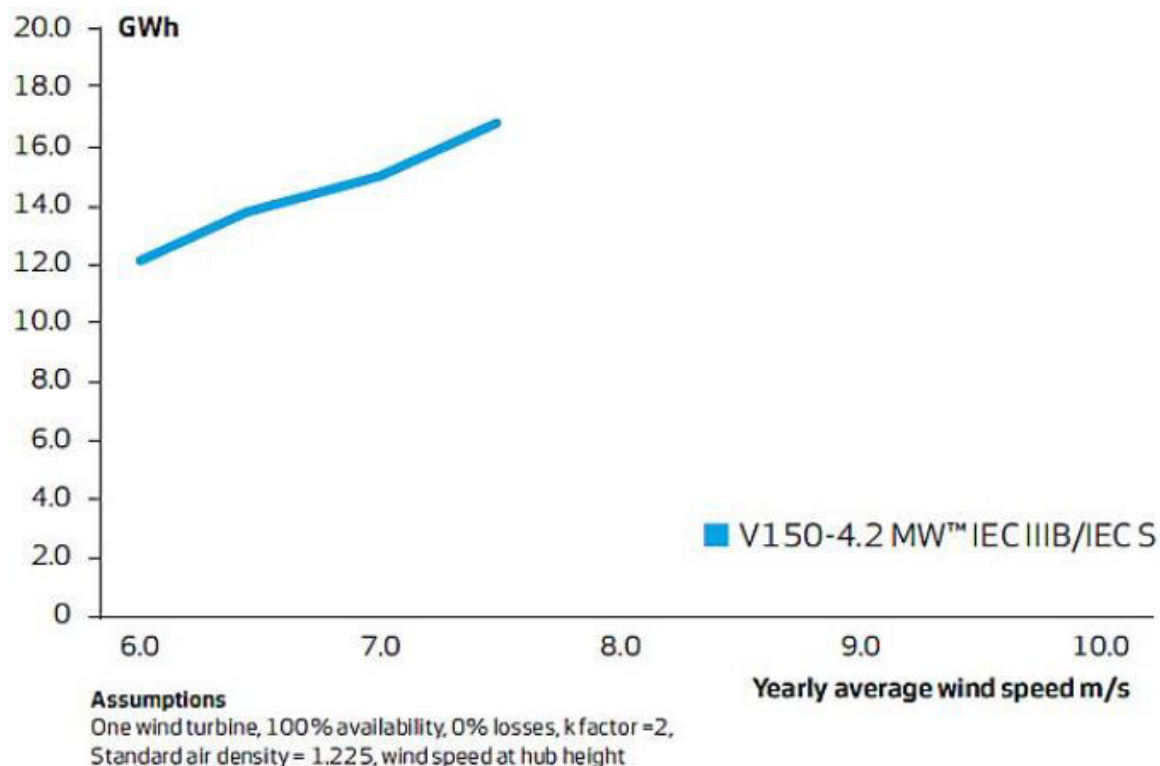


Fonte: Atlas do Potencial Eólico Brasileiro

No Brasil, uma das principais fontes de energia eólica é o Conjunto Eólico Umburanas, que finalizou sua primeira fase de implantação em 2019, com 360MW de capacidade instalada a partir de 144 geradores da Vestas, modelo V150, com 120 metros de altura e 4,2 MW de potência instalados em 18 parques eólicos. De acordo com Boletim Mensal de Geração Eólica de junho de 2021 (BOLETIM MENSAL DE GERAÇÃO EÓLICA), a média de geração mensal em 2020 foi de 185,4 MW, ou seja, foi gerado mais de 2,2GW no ano de 2020 (51,5% da capacidade total). O custo de sua instalação foi de R\$1,4 bilhões. Realizando a divisão, encontramos que o custo é de R\$5.436.893,00/Wh.

Para esse projeto será utilizado como base de cálculos o mesmo modelo de gerador da usina citada acima. Para saber a relação da velocidade do vento com a energia gerada, iremos utilizar a relação vista na figura 2.16, onde uma velocidade média anual do vento de 6m/s geram 12GWh no máximo e 7.5m/s geram 16GWh. Outra informação presente no site da fabricante é que a velocidade máxima suportada é de 20 m/s, e a mínima é de 3m/s.

Figura 2.18: Curva de energia gerada em um ano em condições de velocidade do vento constante em 100% do tempo e sem perdas



Fonte: VESTAS: V150-4.2 MW™

Como todos os dados teóricos, vamos aplicar como referência a produção da usina mostrada na figura 2.18, colocando como potencial 51% do valor calculado, fazendo uma relação linear de 3 m/s até 7,5m/s, e acima disso será fixado a potência.

3 Metodologia

3.1 Recursos

Para disponibilizar o valor do potencial de energia solar ou eólica disponível no município selecionado de maneira simples e garantir o correto funcionamento do projeto, serão necessários a utilizados de alguns recursos tecnológicos, que são o framework Apache Spark para comunicação com o banco de dados, o desenvolvimento de uma API rest para relacionar a solicitação do usuário com as informações desejadas, o desenvolvimento de uma tela para utilização do usuário com Angular, sendo todos essas aplicações gerenciada por hardwares dentro da cloud pública AWS. As informações que utilizaremos estão todas presentes dentro do banco de dados públicos do SONDA. Todos esses recursos estão mais bem explicados nos subtópicos a seguir.

3.1.1 Apache Spark

O Apache Spark é uma ferramenta para processamento de BigData amplamente utilizada por engenheiros de dados, cientistas de dados, engenheiros de Machine Learning. Isso acontece por algumas características que não são encontradas em nenhuma outra ferramenta disponibilizada no mercado.

A primeira característica é sua velocidade de execução, sendo até 100x mais rápido que o principal concorrente no mercado, o Hadoop MapReduce. Outra característica importante é sua generalidade de uso, por ter todos os componentes integrados, sendo eles SQL e DataFrames, MLlib (machine learning), GraphX (geração de gráficos) e Spark Streaming (processamento de dados em tempo real), com isso, cobre qualquer caso de uso de dados.

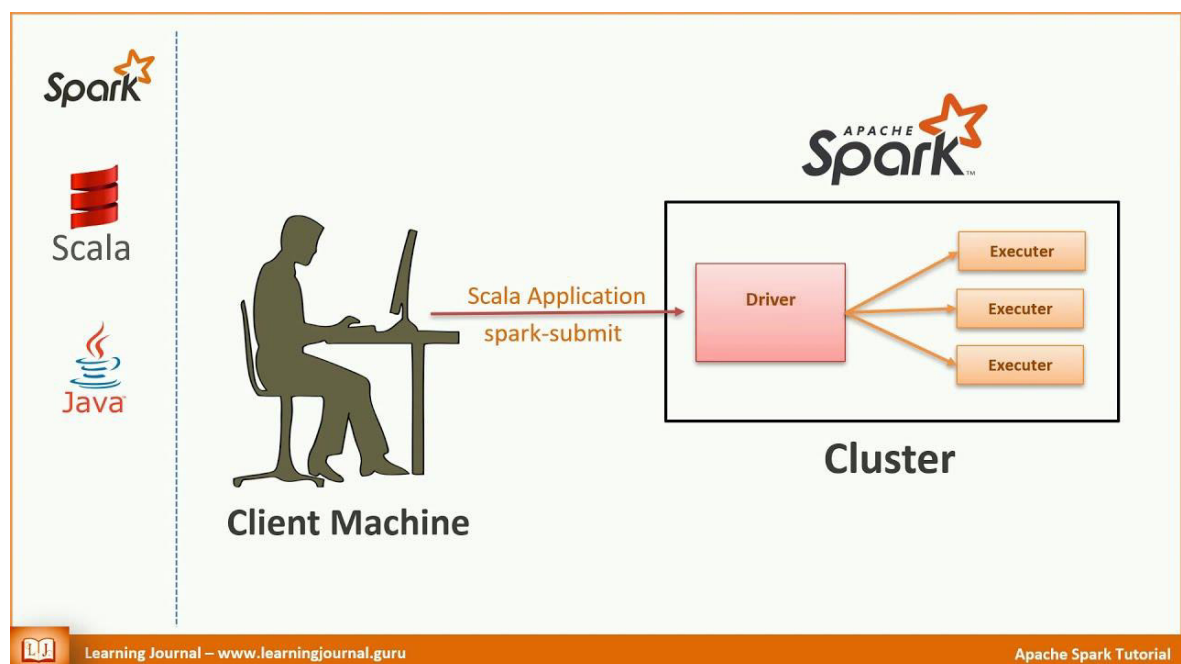
Outro que impulsiona sua utilização é poder programar com diferentes linguagens, que são Python, Java, R, SQL e Scala. Todas as linguagens abstraem o RDD (Resilient Distributed Datasets), que por si só abstrai os objetos dentro do cluster. Esses objetos são os responsáveis por executar o processamento de dados.

Para se obter o programa mais performático possível, precisamos utilizar a linguagem com menor camada de abstrações. Como dito anteriormente, o RDD é a primeira camada de abstração, porém sua lógica é muito complexa, uma vez que está atuando diretamente sobre os objetos de execução. A próxima camada é a Scala, tornando-a mais performática possível. A linguagem Python abstrai a linguagem Java que abstrai o Scala, logo o Python é o menos performático

Na figura 1.3 podemos ver um desenho esquemático simplificado do funcionamento do Apache Spark. Um usuário inicia a execução de uma aplicação Scala, que por sua vez envia a requisição ao Spark, onde o Driver recebe a requisição e distribui aos Executors. É importante deixar claro que o driver é responsável por administrar e distribuir as requisições de acordo com o que foi programado dentro da aplicação, atuando por exemplo como os dados serão distribuídos entre os executors e quantos serão executados por vez.

Os executors colocam os dados que serão utilizados em cache e começa a executar as tarefas sobre esse cache, e como dito anteriormente, os comandos são realizados por RDD, como por exemplo agrupar, somar, contar entre outros.

Figura 3.1: Representação simples do funcionamento do Apache Spark



Fonte: YouTube – Learning Journal

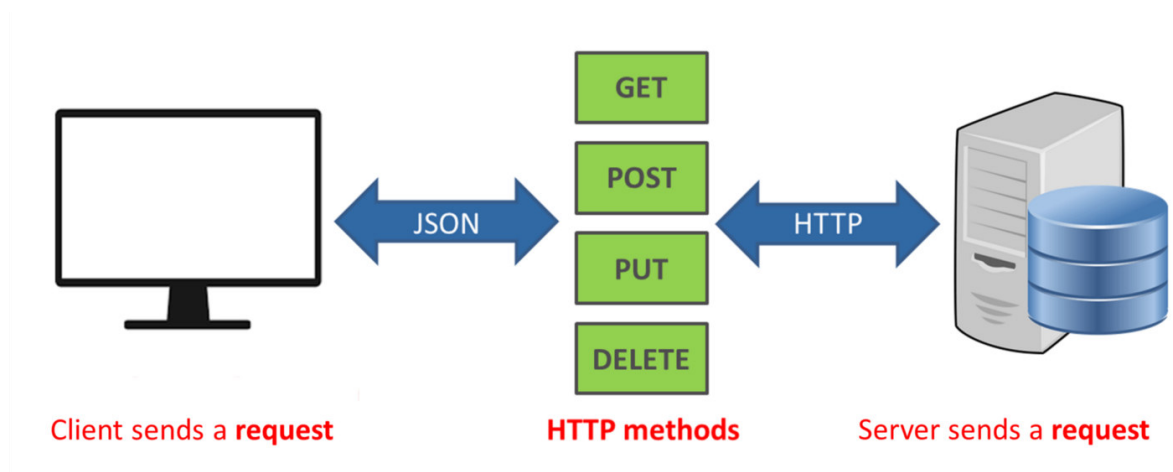
3.1.2 API Rest

Uma definição de API é a troca de informações entre um usuário e um provedor. A partir de uma requisição o usuário passa as informações necessárias para a API que envia ao servidor, essas informações são processadas e executadas, então o servidor responde a API que passa a informação ao usuário.

Quando falamos de API Rest, falamos de um conjunto de regras a serem seguidas para garantir estabilidade e seguranças nessas requisições. Algumas dessas regras é a utilização de requisições HTTP, representadas na figura 3.2, sendo que cada verbo utilizado tem uma função específica. Outro ponto importante é que a requisição precisa estar em formato JSON.

Na figura 3.2 temos uma representação gráfica de uma API Rest. Observa-se que apesar das regras descritas acima, é uma solução muito simples, além de ser extremamente performática. Por isso é amplamente utilizadas em aplicativos, site e IoT(internet das coisas).

Figura 3.2: Exemplificação de uma API



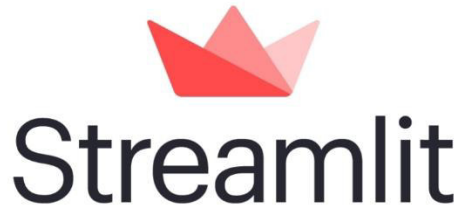
Fonte: REST no Protheus

3.1.3 Streamlit

Hoje existem muitas linguagens que permitem a criação de um front end ao usuário de maneira simples, sendo uma dessas opções o framework Streamlit. Esse framework é desenvolvido em Python com uma forma nativa HTML e Javascript, ou seja, conseguimos desenvolver uma aplicação web conhecendo apenas a linguagem Python.

É uma forma simples de disponibilizar ao usuário informações de banco de dados e de previsões de modelos de machine learning.

Figura 3.3: Logo do framework Streamlit



Fonte: streamlit

3.1.4 Cloud Pública AWS

Cloud pública são servidores fornecidos por empresas a outros usuários a partir de uma rede de internet com um valor pré-estabelecido de acordo com os recursos utilizados. Esses servidores estão cada vez mais comuns dentro do mercado de tecnologia por fornecer maior praticidade e baixo custo para as empresas em relação a ambientes de servidores on-premisse, os quais a própria empresa compra e administra os recursos.

Isso acontece porque quando a empresa administra seus próprios recursos, caso necessite de algum upgrade dos recursos disponibilizados, precisaria comprar um novo hardware, ou também pode acontecer de comprar um hardware superdimensionado, com alto custo de manutenção de maneira desnecessária.

Quando se contrata um recurso de cloud pública, você paga exatamente os recursos que são utilizados, se precisar de mais, com poucos clique e minutos você atende a nova demanda de recursos.

Hoje, uma das empresas líderes do mercado é a Amazon, que fornece o produto com o nome AWS (Amazon Web Service). Nela, podemos contratar serviços de banco de dados, máquinas similares que temos em casa, pipeline de Machine Learning entre vários outros que permitem empresas de grande porte ter todo seu ecossistema dentro da cloud pública.

Nos próximos pontos explicaremos melhor alguns dos recursos específicos que estão sendo utilizados.

3.1.4.1 Instância EC2

Quando falamos de instância EC2 falamos de algo muito similar com os computadores que temos em casa dentro da cloud. Podemos criar e acessar pastas, arquivos e aplicativos.

Para obter essa instância é simples, no próprio site da AWS é selecionado o sistema operacional, quantidade de memória e capacidade de processamento, além de podermos selecionar o tipo de instância de acordo com a nossa necessidade. Existe instâncias com todos os recursos equilibrados, outra com a memória RAM dedicada, ou com uma GPU, selecionando essas opções você terá acesso a sua instância em poucos minutos.

3.1.4.2 Amazon S3

É um serviço de armazenamento de objetos. Podemos comparar esse recurso ao HD dos nossos computadores, onde armazenamos arquivos, imagens, documentos, mas em uma escala maior que exabyte (1.048.576 Terabytes), uma vez que temos empresas multinacionais armazenando todas suas informações dentro desse recurso da Amazon, como por exemplo o Itaú e o Mercado Livre.

Uma vantagem de usar esse tipo de recurso é a possibilidade de escolher com qual frequência você vai utilizar as informações contidas nele. Quanto maior a frequência de uso, maior o custo, e vai diminuindo gradualmente.

3.1.4.3 EMR

É uma plataforma de cluster que simplifica a execução do Apache Spark descrito anteriormente. Como podemos ver na figura 3.1, um driver administra os executors, mas ao executar nas máquinas locais, temos um limite para a criação e tamanho dos executors, podendo inclusive tornar inviável a execução do Apache Spark. Quando utilizamos o Amazon EMR esse risco é mais baixo, porque podemos colocar o recurso necessário para executar todo o tratamento necessário nos dados, e salvar onde precisar.

3.1.4.4 VPC

VPC significa Virtual Private Cloud, como se fosse uma rede interna dentro da cloud AWS. Todos os componentes que estiverem dentro de uma mesma VPC consegue trocar dados e requisições de maneira livre, sem impedimentos. Porém, se estiver em VPCs diferentes, é necessário criar regras de acesso controlando também as atividades permitidas.

Um exemplo seria uma instância EC2 em uma VPC1 precisa acessar um banco de dados S3 na VPC2, como a instância só precisa ler as informações, eu só irei dar o acesso de leitura do banco de dados.

3.1.5 Banco de dados SONDA

A rede SONDA tem sua base de dados dividida em dois grupos: dados Ambientais e dados Anemométricos. Na figura 3.4 podemos ver quais são as informações em cada grupo de dados. Essa divisão ocorre porque existem dois tipos de estações que pertencem ao SONDA, as estações Radiométricas (dados ambientais) e as Aneométricas (dados anemométricos), que são as responsáveis por alimentar essas bases de dados. Também temos a metadata de ambas as tabelas na figura 3.5.

Figura 3.4: Dados presentes no banco de dados do SONDA



Fonte: Sistema de Organização Nacional de Dados Ambientais SONDA

Figura 3.5: Metadata dos dados presentes no banco de dados do SONDA

Variáveis	Descrição
VARIÁVEIS INFORMATIVAS	
Identificador (Id)	Código identificador da estação
Ano (year)	Ano corrente da coleta dos dados
Dia (day)	Dia no calendário Juliano
Data e Hora (datetm)	Data no calendário Gregoriano e hora:minuto:segundo
Minuto (min)	Variáveis Ambientais:
	o minuto especificado (de 0 a 1439) indica a média dos 60 segundos posteriores
	Variáveis Anemométricas:
	o minuto especificado (de 0 a 1430) indica a média dos 10 minutos posteriores
VARIÁVEIS ANEMOMÉTRICAS Periodicidade: 10 minutos	
Velocidade do Vento a 25m (ws_25m)	Velocidade média do vento a 25m em ms-1
Direção do Vento a 25m (wd_25m)	Direção média do vento a 25m de 0° (Norte) a 360° (sentido horário)
Temperatura a 25m (tp_25m)	Temperatura do ar a 25m em °C
Velocidade do Vento a 50m (ws_50m)	Velocidade média do vento a 50m em ms-1
Direção do Vento a 50m (wd_50m)	Direção média do vento a 50m de 0° (Norte) a 360° (sentido horário)
Temperatura a 50m (tp_50m)	Temperatura do ar a 50m em °C
VARIÁVEIS METEOROLÓGICAS Periodicidade: 1 minuto	
Temperatura do Ar na Superfície (tp_sfc)	Temperatura do ar superficial em °C
Umidade do Ar (humid)	Umidade relativa do ar em %
Pressão Atmosférica (press)	Pressão atmosférica em milibares
Precipitação de Chuva (rain)	Precipitação de chuva em milímetros
Velocidade do Vento a 10m (wd_10m)	Velocidade média do vento a 10m em ms-1
Direção do Vento a 10m (wd_10m)	Direção média do vento a 10m de 0° a 360° (horário)
VARIÁVEIS RADIOMÉTRICAS Periodicidade: 1 minuto	
Radiação Global Horizontal (glo_avg)	Média da radiação global horizontal em Wm-2
Radiação Direta Normal (dir_avg)	Média da radiação direta normal em Wm-2
Radiação Difusa (diff_avg)	Média da radiação difusa em Wm-2
Radiação de Onda Longa (lw_avg)	Média da radiação de Onda Longa descendente em Wm-2 com correção
Radiação PAR (par_avg)	Média da radiação fotossinteticamente ativa em μmols s-1m-2
Iluminância (lux_avg)	Média da iluminância em kLux

Fonte: Sistema de Organização Nacional de Dados Ambientais SONDA

Na figura 3.6 temos todas as estações existentes no Brasil, onde na coluna tipo conseguimos identificar quais informações as estações registram, sendo A as anemométricas, S as solarimétricas e AS que possuem sensores que fornecem ambas as informações.

Figura 3.6: Estações pertencentes a rede de coleta de dados SONDA

<i>Estação (SIGLA)</i>	<i>Tipo</i>	<i>UF</i>	<i>ID</i>	<i>Latitude</i>	<i>Longitude</i>	<i>Alt.(m)</i>
Belo Jardim (BJD)	A	PE	22	08° 22' 03" S	36° 25' 46" O	718
Brasília (BRB)	SA	DF	10	15° 36' 03" S	47° 42' 47" O	1023
Cachoeira Paulista (CPA)	S	SP	13	22° 41' 22,65" S	45° 00' 22,28" O	574
Caicó (CAI)	S	RN	20	06° 28' 01" S	37° 05' 05" O	176
Campo Grande (CGR)	S	MS	12	20° 26' 18" S	54° 32' 18" O	677
Cuiabá (CBA)	S	MT	21	15° 33' 19" S	56° 04' 12" O	185
Ourinhos (ORN)	SA	SP	28	22° 56' 55" S	49° 53' 39" O	446
Palmas (PMA)	S	TO	19	10° 10' 40" S	48° 21' 43" O	216
Petrolina (PTR)	SA	PE	11	09° 04' 08" S	40° 19' 11" O	387
Rolim de Moura (RLM)	S	RO	27	11° 34' 54" S	61° 46' 25" O	252
São Luiz (SLZ)	S	MA	16	02° 35' 36" S	44° 12' 44" O	40
São João do Cariri (SCR)	A	PB	25	07° 22' 54" S	36° 31' 38" O	486
São Martinho da Serra (SMS)	SA	RS	08	29° 26' 34" S	53° 49' 23" O	489
Triunfo (TRI)	A	PE	23	07° 49' 38" S	38° 07' 20" O	1123

Fonte: Sistema de Organização Nacional de Dados Ambientais SONDA

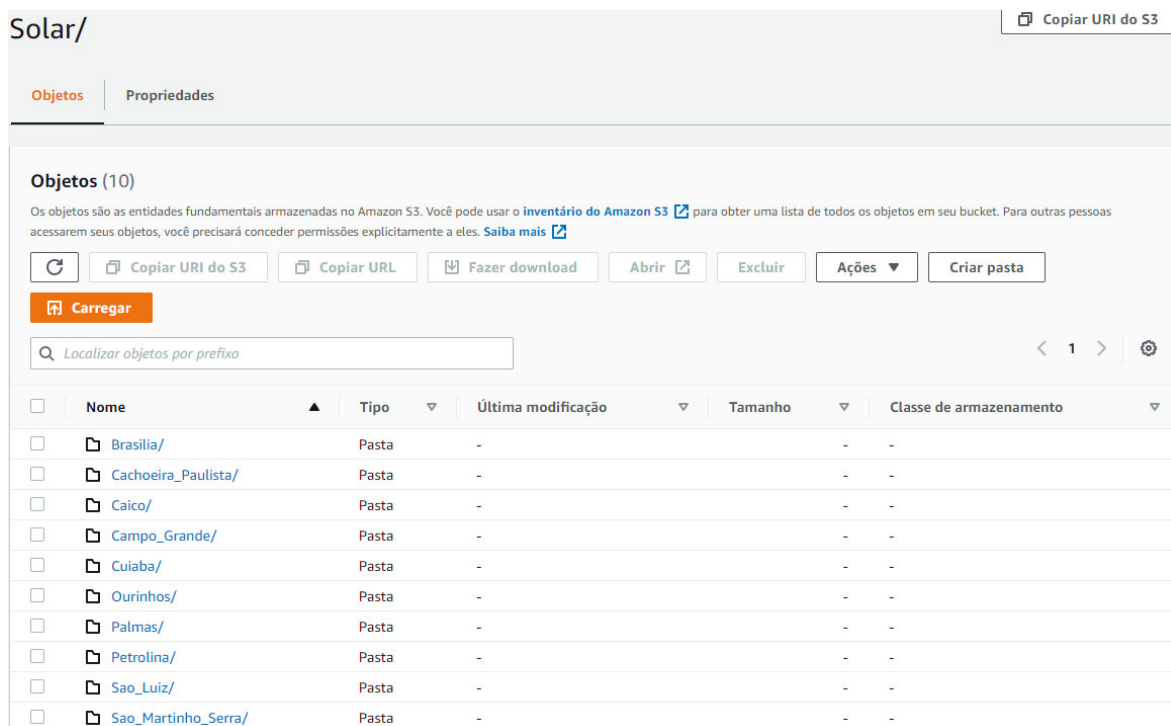
Todos os dados extraídos dessas estações estão disponibilizados no próprio site do SONDA de maneira pública em arquivos separados por mês de coleta. Um problema são as estações exclusivamente anemométricas que possuem dados de mais de 10 anos.

3.2 Desenvolvimento

O primeiro passo, foi realizar o download dos dados de todas as estações para a memória local, que estão separados por mês e em arquivos .zip, além disso, cada estação está em um link diferente, o que torna esse processo bem oneroso. Por questão de capacidade da máquina e recursos a serem utilizados, foi limitado o número de dados em 36 meses por estação. (A estação Rolim de Moura não tem dados suficientes, a estação de São Martinho da Serra possui apenas dados ambientais).

Os arquivos em formato comprimido totalizaram 200MB, fazendo a descompressão temos 1,15GB de informação. Por questão de organização, os arquivos foram divididos em duas pastas, Solar e Eólica, onde as bases que fornecem dados suficiente para cálculo do potencial de energia solar da região estão na pasta Solar, e as que fornecem dados sobre o potencial eólico estão na pasta Eólica. E dentro dessas pastas foram subdivididas por pastas com os nomes das cidades. Essa mesma configuração foi utilizada dentro do S3, como é possível visualizar nas figuras 3.7.

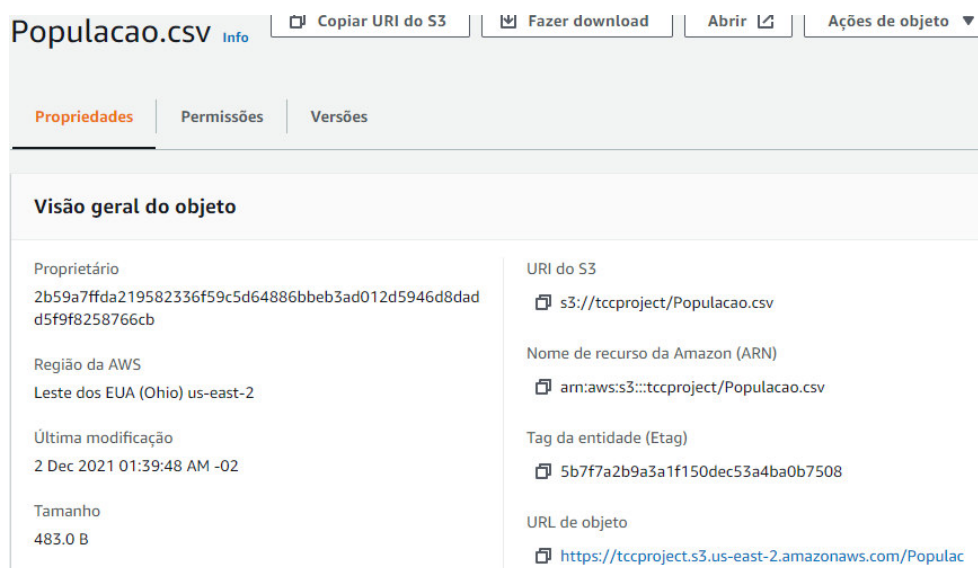
Figura 3.7: Pastas alocadas dentro do bucket S3, na pasta solar



Fonte: Autor

Além dessas informações, precisamos ter as informações de área do município e população. Para isso foi criado um arquivo csv com dados de população, área (km²) e um cálculo de consumo previsto (MW), seguindo as informações presentes na revisão da literatura onde é explicado o consumo de energia elétrica por uma família composta de 3 pessoas.

Figura 3.8: Arquivo criado com os dados sobre população e área dos municípios



Fonte: Autor

Com os dados já prontos, vamos iniciar o desenvolvimento da aplicação que será responsável por nos dizer o potencial energético dos municípios. Antes, precisamos definir a lógica de programação a ser usada.

Teremos dois objetos Scala, um para tratamento do potencial de energia eólica (apêndice I) e outro para tratamento do potencial de energia solar fotovoltaica e energia solar térmica (apêndice II).

Para o cálculo de energia eólica vamos agrupar os dados por mês, utilizando a data completa (datetm) e por cidade, acessando uma pasta de cidade por vez. Vamos obter a velocidade média a 50m de altura. Se a velocidade estiver entre 0m/s e 6m/s vamos considerar que não será gerada energia elétrica. Entre 6m/s e 7,5m/s vamos considerar uma 50% do valor representado na figura 2.16 em uma variação linear, logo, isso representa uma equação 3.1.

$$3.1 \quad P = 4V - 12$$

Onde P é GWh/mês e v a velocidade do vento. E caso a velocidade mensal do vento seja maior que 7,5 vamos considerar a geração máxima também presente na figura 2.16 que é 18 GWh. Lembrando que esse valor é por turbina.

Com isso dividimos o valor da demanda de energia do município pela geração de uma turbina e ver quantas seriam necessárias para atender a demanda mensal. Essa área será o diâmetro do rotor (150m (VESTAS: V150-4.2 MW™)) multiplicado por 50 de acordo com a figura 2.15, e dividido por 4, porque temos 4 geradores dentro dessa área, totalizando 1875m².

E com isso teremos como resultado uma tabela com as colunas de município, tipo de energia, potencial energético de geração de uma turbina e se atende ou não a demanda do município.

Para o cálculo de energia solar, vamos agrupar os dados por cidade, acessando uma pasta de cidade por vez. Vamos obter a média de radiação normal. Com esse valor vamos calcular que as placas solares têm o potencial de converter 16% de toda energia recebida em energia elétrica, utilizando a informação de eficiência máxima de 20% de toda radiação recebida, junto com a eficiência real de 80% do que estipulado em ambiente controlado. Caso a radiação seja maior que 2000 MW/m² também será considerado a energia solar térmica, considerando a mesma taxa de conversão anterior, por não ter sido encontrado um valor de referência.

Com isso já teremos o potencial gerado por m² e faremos a mesma relação citada anteriormente, adicionando informação na tabela criada anteriormente.

Com a tabela criada, agora podemos disponibilizar essas informações. Vamos usar uma aplicação streamlit, que vai buscar o banco de dados no bucket S3 e salvar na memória em cache, tornando acessível de maneira rápida. Mas como citado anteriormente existe regras a serem seguidas para que possamos realmente ter uma API segura e que segue as boas práticas de requisições HTTP, mas pela simplicidade do streamlit isso não é possível, logo não é recomendado utilização dessa aplicação para dados sensíveis e aplicações produtivas. A estrutura do streamlit é visto no apêndice III.

Figura 3.9: Configuração EMR utiliza para executar os códigos de tratamento do banco de dados

Configuração geral

Nome do cluster

Eolico

☒ Registro em log

Pasta do S3

s3://aws-logs-204792748194-us-east-2/elasticmap

Modo de execução

☒ Cluster

☐ Execução da etapa

Configuração de software

Versão

emr-5.33.1

Aplicativos

☐ Core Hadoop: Hadoop 2.10.1, Hive 2.3.7, Hue 4.9.0, Mahout 0.13.0, Pig 0.17.0, and Tez 0.9.2

☐ HBase: HBase 1.4.13, Hadoop 2.10.1, Hive 2.3.7, Hue 4.9.0, Phoenix 4.14.3, and ZooKeeper 3.4.14

☐ Presto: Presto 0.245.1 with Hadoop 2.10.1 HDFS and Hive 2.3.7 Metastore

☒ Spark: Spark 2.4.7 on Hadoop 2.10.1 YARN and Zeppelin 0.9.0

☐ Usar o catálogo de dados do AWS Glue para obter uma tabela de metadados

Configuração do hardware

Tipo de instância

m4.large

O tipo de instância selecionado adiciona um volume do EBS GP2 de 32 GiB padrão por instância. Saiba mais

Número de instâncias

3

(1 principal e 2 nós core)

Escalabilidade do cluster

☐ scale cluster nodes based on workload

Encerramento automático

☐ Habilitar encerramento automático

Saiba mais

Fonte: Autor

Agora com a lógica de programação definida, precisamos definir a estrutura a ser utilizada. O banco de dados já está construído, e o destino da tabela final será no mesmo bucket S3 da amazon. Isso para facilitar a configuração de segurança de acesso, deixando todas as informações dentro de uma mesma rede de acesso.

Os objetos Scala são executados em instâncias EMR, com versionamento SPARK 2.4.7, 3 máquinas com 32GB com configurações equilibradas em relação a velocidade de processamento e capacidade de processamento. Sendo 3 máquinas, uma driver e duas executors, ou seja, serão 64GB para processamento dos dados, que é mais que suficiente para colocar todos os dados em memória e ir realizando as funções necessárias. Aqui não precisamos passar nenhuma configuração já que a versão spark já vem instalada, passando só o local onde a aplicação está armazenada.

Por último, precisamos de uma instância EC2 para disponibilizar a aplicação para acesso do usuário, o único componente disponibilizado e acessível por qualquer usuário com a URL. Como vamos colocar a tabela final em memória e ainda precisa sobrar espaço para executar as solicitações, vamos utilizar uma instância t2. medium, que apresenta 4 GB de memória, 2 vCPUs e sistema operacional Linux, que para uso exclusivo desse projeto é o suficiente, e na criação solicitamos a instalação dos requisitos e a execução do programa (Apêndice IV).

Com a descrição da lógica de programação aplica, estrutura e banco de dados, podemos observar toda arquitetura do projeto desenvolvido na figura 3.10.

Figura 3.10: Alta arquitetura do projeto desenvolvido.



Fonte: Autor

4 Resultados e discussão

Quando executado o programa, foi identificado que os dados de 2009 relacionados a velocidade dos ventos em Brasília estavam com algum problema, porque a média ficou muito diferente dos outros anos e causou uma significativa distorção. Podemos observar na figura 4.1 que as velocidades dos ventos estavam em -12.5 m/s fixos. Como representam o único ano completo de dados e mais de 60% dos dados relacionado a Brasília, foi descartado o cálculo de energia eólica neste município.

Figura 4.1: Base de dados disponibilizado está com valores fixos na posição onde deveria informar velocidade do vento a 50m em Brasília.

```

2009;1;2009-1-1 00:00:00;0;0.584;305.2;-75;-12.5;270;25.56;
2009;1;2009-1-1 00:10:00;10;0.732;310.4;-75;-12.5;270;25.67;
2009;1;2009-1-1 00:20:00;20;0.49;321.2;-75;-12.5;270;25.79;
2009;1;2009-1-1 00:30:00;30;0.063;247.9;-75;-12.5;270;25.68;
2009;1;2009-1-1 00:40:00;40;0.622;131.6;-75;-12.5;270;25.58;
2009;1;2009-1-1 00:50:00;50;1.359;114.2;-75;-12.5;270;25.36;
2009;1;2009-1-1 01:00:00;60;1.256;128.2;-75;-12.5;270;25.06;
2009;1;2009-1-1 01:10:00;70;0.094;107.4;-75;-12.5;270;24.88;
2009;1;2009-1-1 01:20:00;80;-0.017;107.5;-75;-12.5;270;24.79;
2009;1;2009-1-1 01:30:00;90;0.06;107.6;-75;-12.5;270;24.73;
2009;1;2009-1-1 01:40:00;100;-0.016;107.6;-75;-12.5;270;24.86;
2009;1;2009-1-1 01:50:00;110;0.129;347.8;-75;-12.5;270;24.7;
2009;1;2009-1-1 02:00:00;120;0.634;110.6;-75;-12.5;270;24.61;
2009;1;2009-1-1 02:10:00;130;2.012;146.2;-75;-12.5;270;24.47;
2009;1;2009-1-1 02:20:00;140;2.124;156.2;-75;-12.5;270;24.52;
2009;1;2009-1-1 02:30:00;150;1.159;167.6;-75;-12.5;270;24.38;
2009;1;2009-1-1 02:40:00;160;1.115;180.1;-75;-12.5;270;24.38;
2009;1;2009-1-1 02:50:00;170;0.938;171.5;-75;-12.5;270;24.36;
2009;1;2009-1-1 03:00:00;180;0.135;144.9;-75;-12.5;270;24.19;
2009;1;2009-1-1 03:10:00;190;0.383;240.2;-75;-12.5;270;23.96;
2009;1;2009-1-1 03:20:00;200;0.082;239.1;-75;-12.5;270;24.05;
2009;1;2009-1-1 03:30:00;210;0.371;325.3;-75;-12.5;270;23.93;
2009;1;2009-1-1 03:40:00;220;-0.016;325.2;-75;-12.5;270;23.82;
2009;1;2009-1-1 03:50:00;230;0.594;201.9;-75;-12.5;270;24.12;
2009;1;2009-1-1 04:00:00;240;0.908;192.4;-75;-12.5;270;24.23;
2009;1;2009-1-1 04:10:00;250;0.004;112.0;-75;-12.5;270;22.47;

```

Fonte: Autor

Como é possível observar nas figuras 2.9 e 2.14, as regiões nordeste e sul tem velocidade de ventos e intensidade de energia solar altos comparados a demais regiões do Brasil, então iremos focar em resultados de três estações, a de São Martinho da Serra, a de Ourinhos e a de Petrolina,

que possuem informações tanto para cálculo de potencial de energia eólica quanto de energia Solar.

Em São Martinho da Serra é um município localizado no Rio Grande do Sul, estado onde já existe a instalação de algumas usinas eólicas, com isso os dados sobre velocidade dos ventos seriam interessantes do ponto de vista do projeto, mas não está disponibilizado.

Com isso só obtivemos os dados de energia solar. Na figura 4.2 temos uma representação de como fazer a consulta de dados pela aplicação gerada pelo streamlit.

Figura 4.2: Imagem da tela de consulta do usuário.

Potencial elétrico do município

Informações dos dados

Selecione a fonte de energia desejada

Solar

Selecione o município

Sao Martinho

Fonte: Autor

Por questão de comparação, agrupamos todos os valores em uma planilha do Excel, que está na figura 4.3

Figura 4.3: Valores obtidos em tabela Excel para comparação

Município	Energia	Potência (Eolica = W/gerador; Solar = W/m²)	Área (km²)	Possível
Ourinhos	Eolica	0	0	Não
Ourinhos	Solar	12,44252	4,84007	Sim
Petrolina	Eolica	3,28786	0,087171	Sim
Petrolina	Solar	59,17785	2,58306	Sim
São Martinho da Serra	Solar	35,2629	0,00477	Sim

Fonte: Autor

Em Ourinhos, uma cidade do interior de São Paulo, a média anual de ventos foi abaixo de 6m/s, portanto, de acordo com o valor encontrado a instalação de uma usina eólica nesta cidade não teria utilidade.

Em relação a energia solar, os dados a partir de janeiro de 2018 estão todos vazios, resumindo a média de radiação para o período. de 2016 até 2017. Mesmo assim, chegamos em uma área necessária de quase 5 km² para uma usina fotovoltaica atender toda demanda da população.

Em Petrolina, uma cidade de Pernambuco, a velocidade média dos ventos anuais foi acima de 6m/s, e aplicando os cálculos é possível construir uma usina eólica que atenda as necessidades da população ocupando menos de 0,1Km². E para atender a demanda de energia com energia solar precisamos de cerca de 2,5 km² de placas solares fotovoltaicas.

Nenhum desses municípios apresentou uma média de radiação solar suficiente para tornar viável a utilização da energia solar térmica.

Deve-se ressaltar que os dados usados são a média anual, ou seja, também foi considerado o período noturno de irradiação solar, então muito provavelmente no pico de irradiação solar durante o dia, as placas solares gerem muito mais energia do que o necessário para atender a população do município se construída uma usina com a área encontrada.

Além desses resultados, pudemos acompanhar a velocidade de execução, recursos consumidos e custos da arquitetura implantada dentro da AWS.

O custo das instâncias dedicadas ao EMR, que foi utilizada apenas para uma única execução e depois desativadas, e tem seu valor referente ao tempo de uso, foi de U\$0,09, sendo cobrada a hora inteira de 3 instâncias, sendo que foi utilizada por apenas 10 minutos. A instância EC2, que deve estar disponível todo tempo durante o mês, terá um custo aproximado de U\$20,95. E o armazenamento do bucket S3, que tem seu valor relacionado ao armazenamento e quantidade de vezes que acessamos a informação, terá um custo aproximado de U\$0,07 por mês, somando tudo e colocando uma conversão próxima de U\$1 = R\$5,60, vamos ter um gasto aproximado de R\$118,22/mês.

5 Conclusão

Primeiro, em relação as escolhas de ferramentas nesse projeto, foram ferramentas extremamente robustas para uma situação bem simples. O gasto estimado por mês poderia ser nulo e teríamos um resultado igual ou muito parecido utilizando recursos como planilhas eletrônicas ou banco de dados locais. Outro ponto foi a utilização do Apache Spark com Scala, ideal para big data em que o tempo de execução em milissegundos é de extrema importância, diferente da situação atual onde não haverá atualização dos dados, e o total ocupado por ele é de apenas 1,5GB. Em resumo todo o processo de desenvolvimento poderia ter sido realizado de maneira mais simples e rápida com o mesmo resultado. Porém, em caso de trabalhos futuros em que essas condições estejam presentes, esse trabalho pode servir como base para desenvolvimento da aplicação Apache Spark.

Em relação aos dados utilizados, muitos deles estão muito defasados e desatualizados, muito provavelmente não representando a realidade de hoje, uma vez que o clima tem sofrido constantes mudanças.

Sobre os dados encontrados nos municípios de São Martinho da Serra, Ourinhos e Petrolina, ao compararmos com os dados disponibilizados nas figuras 2.9 e 2.16 os dados encontrados fazem sentido, onde o município de Petrolina, localizado no nordeste tem possibilidade de geração de energia eólica e solar, a região próxima ao município de Ourinhos, por apresentar uma rugosidade da superfície grande não permite a geração de energia eólica, mas, como quase todo território brasileiro, é possível obter energia solar fotovoltaica.

Por fim, os cálculos realizados, quando utilizamos uma média anual, desconsideramos variações de estação e horário, pontos de extrema importância quando consideramos a utilização de recursos naturais como vento e a radiação solar. Com isso esses dados podem conter grande imprecisão sobre a real área necessária para construção de usinas nesses municípios.

6 Referencias

WORLD Population Prospects 2019. [S. l.], 2 jan. 2020. Disponível em:

<https://population.un.org/wpp/Graphs/DemographicProfiles/>. Acesso em: 4 set. 2021.

ENERGY Technology Perspectives 2020. [S. l.], 1 set. 2020. Disponível em:

<https://www.iea.org/reports/energy-technology-perspectives-2020>. Acesso em: 7 dez. 2021.

ATLAS Brasileiro de Energia SOLAR. 2. ed. [S. l.: s. n.], 2017. Atlas. Disponível em: <http://mtc-m21b.sid.inpe.br/rep/8JMKD3MGP3W34P/3PERDJE>. Acesso em: 7 dez. 2021.

NATIONAL Renewable Energy Laboratory. [S. l.], 7 dez. 2021. Disponível em:

<https://www.nrel.gov/>. Acesso em: 7 dez. 2021.

USINA ESPANHOLA utiliza energia solar até mesmo durante a noite. [S. l.], 2 out. 2012.

Disponível em: <https://canaltech.com.br/curiosidades/Usina-espanhola-utiliza-energia-solar-ate-mesmo-durante-a-noite/>. Acesso em: 7 dez. 2021.

USINA SOLAR DA ESPANHA usa sal fundido para manter turbinas funcionando e gerar energia 24 horas, 9 mai. 2012. Disponível em: <https://petronoticias.com.br/usina-solar-da-espanha-usa-sal-fundido-para-manter-turbinas-funcionando-e-gerar-energia-24-horas/>. Acesso em: 7 dez. 2021.

MARTINS, F.R.; GUARNIERI, R.A.; PEREIRA, E.B. **O APROVEITAMENTO DA ENERGIA EÓLICA.**

Instituto Nacional de Pesquisas Espaciais, Centro de Previsão do Tempo e Estudos Climáticos, São José dos Campos, SP, Brasil, [s. l.], 17 out. 2007. Disponível em:

<https://www.scielo.br/j/rbef/a/nL6x7dJv9gJv7HnkTSJRLfs/?format=pdf&lang=pt>. Acesso em: 7 dez. 2021.

G.M. Joselin Herbert, S. Iniyar, E. Sreevalsan and S. Rajapadian, *Renew. Sustain. Energ. Rev.* **11**, 1117 (2007).

R. Gasch and J. Tewe (eds), **Wind Power Plants: Fundamentals, Design, Construction and Operation** (Solarpraxis AG, Alemanha, 2002).

SCHLECHT, M., MEYER, R. **Site selection and feasibility analysis for concentrating solar power (CSP) systems.** LOVEGROVE, K.; STEIN, W. Concentrating solar power technology. Cambridge, UK: Woodhead, 2012. Cap. 4.

O VENTO GEOSTRÓFICO. [S. l.], 1 jun. 2017. Disponível em:

<http://fisica.ufpr.br/grimm/aposmeteo/cap7/cap7-4.html>. Acesso em: 7 dez. 2021.

ATLAS do Potencial EÓLICO Brasileiro. [S. l.: s. n.], 2001. Atlas. Disponível em:

http://www.cresesb.cepel.br/publicacoes/download/atlas_eolico/Atlas%20do%20Potencial%20Eolico%20Brasileiro.pdf. Acesso em: 7 dez. 2021.

BOLETIM MENSAL DE GERAÇÃO EÓLICA. [S. l.: s. n.], 2021. Mensal. Disponível em:

<http://www.ons.org.br/AcervoDigitalDocumentosEPublicacoes/Boletim%20Mensal%20de%20Geracao%20A7%C3%A3o%20E%C3%B3lica%202021-06.pdf>. Acesso em: 7 dez. 2021.

REST no Protheus. [S. l.], 7 dez. 2021. Disponível em: <https://medium.com/totvsdevelopers/rest-no-protheus-b%C3%A1sico-do-b%C3%A1sico-a4b4431a4e3e>. Acesso em: 7 dez. 2021.

STREAMLIT. [S. l.]. Disponível em: <https://streamlit.io/>. Acesso em: 7 dez. 2021.

AWS AMAZON. [S. l.]. Disponível em: <https://aws.amazon.com/pt/>. Acesso em: 7 dez. 2021.

Sistema de Organização Nacional de Dados Ambientais SONDA. [S. l.]. Disponível em: <http://sonda.ccst.inpe.br/>. Acesso em: 7 dez. 2021.

Portal Solar. [S. l.]. Disponível em: <https://www.portalsolar.com.br/>. Acesso em: 7 dez. 2021.

What is a REST API. [S. l.]. Disponível em: <https://www.redhat.com/pt-br/topics/api/what-is-a-rest-api>. Acesso em: 7 dez. 2021.

VESTAS: V150-4.2 MW™. [S. l.], 2018. Disponível em: <https://www.vestas.com/en/products/4-mw-platform/V150-4-2-MW>. Acesso em: 7 dez. 2021.

APACHE SPARK: introdução. [S. l.], 2016. Disponível em: <https://www.devmedia.com.br/introducao-ao-apache-spark/34178>. Acesso em: 7 dez. 2021.

7 Apêndices

I – Código desenvolvido em scala para tratamento de dados da energia eólica

```
import org.apache.log4j.Logger

import org.apache.spark.sql.Session
import org.apache.spark.sql.functions._

object AggDemo extends Serializable {

  @transient lazy val logger: Logger = Logger.getLogger(getClass.getName)

  def main(args: Array[String]): Unit = {

    case class Resultado(municipio: String, energia: String, potencial: Float, area: Float, atende:
    Bool)

    val spark = Session.builder()

      .appName("Eolico")

      .master("local[3]")

      .getOrCreate()

    val cidades: List[String] = List("Belo_Jardim", "Brasilia", "Ourinhos", "Petrópolis",
    "São_João_Cairi", "Triunfo")

    val apoioDF = spark.read

      .format("csv")

      .option("header", "true")

      .load("S3://tccproject/Populacao.csv")

    for (cidade <- cidades)

    {

      var path = "S3://tccproject/" + cidade

      val invoiceDF = spark.read

        .format("csv")
```

```

.option("header", "true")

.load(path: _)

val vel_media = invoiceDF.select(
  avg("ws_50").as("media_velocidade")
)

if (vel_media < 6) {
  var potencia = 0
} else if (vel_media < 7.5) {
  var potencia = (vel_media*4) - 12
} else {
  var potencia = 1.5
}

var consumo_municipio = apoioDF
  .select("Consumo de energia estimada(MW)")
  .where("Municipio" = cidade)
)

var area_municipio = apoioDF
  .select("Área(km2)")
  .where("Municipio" = cidade)
)

if (potencia > 0){
  var turbinas = (consumo_municipio*12/1000)/potencia
  var area = turbinas * 1875/1000000
  if (area < area_municipio) {
    var possivel = true
  } else {
    var possivel = false
  }
}

```

```
    }  
  } else {  
    var possivel = false  
  }  
  val resultado_final = new resultado(cidade, "Eolica", potencia, area, possivel)  
  var df = resultado_final.toDF()  
}  
  
val df_final = df_final.union(df)  
dataFrame.write.format("com.databricks.spark.csv").save("S3://tccproject/energia_eolica.csv")  
}  
  
}
```

II – Código desenvolvido em scala para tratamento dos dados de energia solar

```
import org.apache.log4j.Logger

import org.apache.spark.sql.Session
import org.apache.spark.sql.functions._

object AggDemo extends Serializable {

  @transient lazy val logger: Logger = Logger.getLogger(getClass.getName)

  def main(args: Array[String]): Unit = {

    case class Resultado(municipio: String, energia: String, potencial: Float, area: Float, atende:
    Bool)

    val spark = SparkSession.builder()

      .appName("Solar")

      .master("local[3]")

      .getOrCreate()

    val cidades: List[String] = List("Brasilia", "Cachoeira Paulista", "Caico", "Campo Grande",
    "Cuiaba", "Ourinhos", "Palma", "Petrolina", "Sao Luiz", "Sao Martinho")

    val apoioDF = spark.read

      .format("csv")

      .option("header", "true")

      .load("S3://tccproject/Populacao.csv")

    for (cidade <- cidades)

    {

      var path = "S3://tccproject/" + cidade

      val invoiceDF = spark.read

        .format("csv")
```

```

.option("header", "true")

.load(path: _)

val rad_media = invoiceDF.select(
    avg("dir_avg").as("mediaraidacao")
)

val potencia = rad_media*0.16

var consumo_municipio = apoioDF
    .select("Consumo de energia estimada(MW)")
    .where("Municipio" = cidade)
)

var area_municipio = apoioDF
    .select("Área(km2)")
    .where("Municipio" = cidade)

)

var placas = (consumo_municipio*12/1000)/potencia
var area = placas/1000000
if (area < area_municipio) {
    var possivel = true
} else {
    var possivel = false
}

val resultado_final = new resultado(cidade, "Solar", potencia, area, possivel)
var df = resultado_final.toDF()

if (rad_media > 2000000) {

```

```
        val resultado_final_termico = new resultado(cidade, "Solar_termico", potencia, area,
possivel)

        var df_termico = esutado_final_termico.toDF()

        df = df_termico.union(df)

    }

}

val df_final = df_final.union(df)

dataFrame.write.format("com.databricks.spark.csv").save("S3://tccproject/energia_solar.csv")

}

}
```

III – Código Streamlit para execução do front end ao usuário

```
import pandas as pd

import streamlit as st

#título
st.write("Potencial elétrico do município")

#cabeçalho
st.subheader("Informações dos dados")

#Dando as opções ao usuário
opcao = st.selectbox( 'Selecione a fonte de energia desejada', ('Solar', 'Eolica'))

#Agora damos as opções de municípios
if opcao == 'Solar':

    cidade = st.selectbox( 'Selecione o município', ('Brasilia', 'Cachoeira Paulista', 'Caico', 'Campo Grande', 'Cuiaba', 'Ourinhos',

                                                    'Palma', 'Petrolina', 'Sao Luiz', 'Sao Martinho'))

if opcao == 'Eolica':

    cidade = st.selectbox( 'Selecione o município', ('Belo Jardim', 'Brasilia', 'Ourinhos', 'Petrolina', 'Sao Joao Cairi', 'Triunfo'))

#Busca no S3 qual cidade foi escolhida
string = 's3://tccproject/Eolica/' + cidade + '/resultado'
dataset = pd.read_csv(string)

#Agora iremos mostrar os resultados
st.write(base)
```

Apendice IV – Comando para execução do streamlit

```
pip install pandas
pip install streamlit
streamlit run S3://tccproject/code/front.py
```