

IMPLEMENTAÇÃO DE UM MODELO DE CÉU CLARO EM ALGORITMOS DE CLASSIFICAÇÃO

**RELATÓRIO FINAL DE PROJETO DE INICIAÇÃO CIENTÍFICA
(PIBIC/INPE/CNPq)**

**Alberto Carvalho Sardinha (Unifesp, Bolsista PIBIC/CNPq)
E-mail: alberto.sardinha@unifesp.br**

**Ênio Bueno Pereira (CCST/INPE, Orientador)
Email: enio.pereira@inpe.br**

Alberto Carvalho Sardinha (Unifesp, Bolsista PIBIC/CNPq)
E-mail: alberto.sardinha@unifesp.br

Ênio Bueno Pereira (CCST/INPE, Orientador)
Email: enio.pereira@inpe.br

IMPLEMENTAÇÃO DE UM MODELO DE CÉU CLARO EM ALGORITMOS DE CLASSIFICAÇÃO

**RELATÓRIO FINAL DE PROJETO DE INICIAÇÃO CIENTÍFICA
(PIBIC/INPE/CNPq)**

Relatório final de projeto de iniciação científica apresentado ao Instituto Nacional de Pesquisas Espaciais – INPE, como parte das atividades de bolsista do Programa Institucional de Bolsas de Iniciação Científica.

Instituto Nacional de Pesquisas Espaciais
Centro de Ciência do Sistema Terrestre
Programa Institucional de Bolsas de Iniciação Científica

São José dos Campos, SP
Julho de 2019

Resumo

O acesso a novas fontes de energia, principalmente renováveis, requer níveis específicos de dados para sua classificação correta. As estações de coleta capturam informações em um espaço curto de tempo com uma precisão estável, executando o tratamento das informações coletadas. O controle de qualidade é um aspecto principal em qualquer sistema de tratamento de dados, para garantir confiabilidade e permitir a comparação com outras amostras. O Instituto Nacional de Pesquisas Espaciais (INPE), com base nesses conceitos, implementou a rede SONDA para prover dados meteorológicos e solarimétricos com a finalidade de suporte ao setor de energia no Brasil. Este trabalho consiste na continuação da implementação do algoritmo de validação das informações anemométricas e solarimétricas coletadas na rede SONDA, consistindo em adaptar, em seu último estágio, um modelo de céu claro proposto por Muhammad Iqbal. Para dar continuidade a este projeto de Iniciação Científica estão programadas as atividades: Revisão bibliográfica e escolha dos principais modelos de céu claro; reestruturação do algoritmo implementando novos testes de qualidade para o aumento de confiabilidade das análises; Validação de modelos computacionais utilizados no levantamento destes recursos; implementação do modelo de céu claro baseado no proposto por Muhammad Iqbal, e agregá-lo ao conjunto de validação de dados solarimétricos na rede SONDA.

Lista de ilustrações

Figura 1 – Espectro de radiação solar	8
Figura 2 – Processos de interação da radiação solar com a atmosfera	8
Figura 3 – Ângulos notáveis em solametria.	9
Figura 4 – Componentes da irradiância solar.	11
Figura 5 – Modelo Brasil-SR.	13
Figura 6 – Algoritmo de classificação de dados da rede SONDA.	17
Figura 7 – Bloco de código na linguagem Java.	20
Figura 8 – Bloco de código na linguagem Python.	20

Sumário

1	Introdução	5
2	Objetivos	6
3	Revisão Bibliográfica	7
3.1	Conceitos de Irradiância Solar	7
3.2	Modelagem numérica aplicada à Energia Solar	11
3.3	Modelo Brasil-SR	12
3.4	Modelo de Céu Claro	13
4	Materiais e Métodos Utilizados	16
5	Análises e Resultados	19
6	Conclusão	21
	Referências	22

1 Introdução

No cenário de pesquisas de ciência terrestre existe grande relevância do mapeamento do potencial solar e eólico, sendo precisamente necessário quantificar a disponibilidade e variabilidade deste, tanto no quesito espacial, quanto no temporal. O acesso a novas fontes de energia, principalmente renováveis, requer níveis específicos de dados para sua classificação correta, a pesquisa científica em meteorologia aplicada ao setor de energia e modelagem computacional é de extrema importância para a avaliação de dados coletados por estações.

As estações de coleta capturam informações em um espaço curto de tempo com uma precisão estável, executando o tratamento das informações coletadas. O controle de qualidade é um aspecto principal em qualquer sistema de tratamento de dados, para garantir confiabilidade e permitir a comparação com outras amostras. Para o tratamento desse enorme volume de dados é necessário estabelecer rotinas computacionais de coleta, extrair as informações necessárias de dados brutos, verificar e sinalizar possíveis inconsistências e corrigi-las se necessário. Gerando assim, dados com confiabilidade necessária para permitir uma base de análise.

O Instituto Nacional de Pesquisas Espaciais (INPE), com base nesses conceitos, implementou a rede SONDA para prover dados meteorológicos e solarimétricos com a finalidade de suporte ao setor de energia no Brasil. O Laboratório de Modelagem e Estudos de Recursos Renováveis de Energia (LABREN) realiza atividades de pesquisa e ensino na área de meteorologia aplicada ao setor de energia, com ênfase nas relações entre energias e sistema climático, mantendo a rede SONDA de coleta de dados, composta por estações solarimétricas medindo radiação solar incidente global, difusa, direta normal, PAR, iluminância, espectrofotometria solar e dados meteorológicos complementares, como também torres anemométricas com 50 metros de altura dotadas de anemômetros e termômetros a duas alturas, 25 e 50 metros.

Este trabalho consiste na continuação da implementação do algoritmo de validação das informações anemométricas e solarimétricas coletadas na rede SONDA, consistindo em adaptar, em seu último estágio, um modelo de céu claro proposto por Muhammad Iqbal (IQBAL, 1983). No capítulo 2 serão mostrados os objetivos gerais e específicos relacionados ao trabalho, o capítulo 3 irá tratar da fundamentação teórica e o modelo de céu claro proposto e os capítulos 4, 5 e 6 serão, respectivamente, os métodos utilizados, resultados e análises e a conclusão.

2 Objetivos

Entre os objetivos principais do projeto de implementação do modelo estão:

- Estudo e melhoria dos métodos de tratamento e qualificação aplicados aos dados meteorológicos e espaciais da rede SONDA, que são utilizados na quantificação dos recursos solar e eólico.
- Revisar algoritmos, implementando novos testes de qualidade para o aumento de confiabilidade das análises.
- Validação de modelos computacionais utilizados no levantamento dos recursos.
- Implementar o modelo de céu claro baseado no modelo proposto por Muhammad Iqbal ([IQBAL, 1983](#)), e agregá-lo ao conjunto de validação de dados solarimétricos na rede SONDA.

3 Revisão Bibliográfica

Para iniciar as instruções em qualificação de dados, foi necessário inicialmente um período de estudo e pesquisa para absorver os conceitos básicos sobre radiação solar. Entre os conteúdos principais estão a Irradiância Solar e como funcionam os fenômenos referentes à ela (seção 3.1), modelagem numérica aplicada à energia solar (seção 3.2), o entendimento básico sobre o modelo Brasil-SR (seção 3.3) e o modelo de céu claro de Iqbal (1983) para a última etapa de validação da rede SONDA (seção 3.4).

3.1 Conceitos de Irradiância Solar

A energia solar é uma das alternativas energéticas mais promissoras atualmente, sendo sua característica principal não ser renovável, porém tendo uma fonte inesgotável. O Sol é uma estrela de tamanho médio que irradia energia como consequência de reações de fusão nuclear dos átomos de hidrogênio, sua temperatura de superfície é de aproximadamente 5778 Kelvin (5505 graus Celsius) gerando uma taxa de energia constante com potência de ordem $3.86 \cdot 10^{26}$ Watts. A radiação solar é definida como o fluxo de energia emitida pelo Sol e transmitida sob a forma de radiação eletromagnética, que consiste em duas medidas de radiação principais, a Irradiância Solar Total (TSI) e a Irradiância Solar Espectral (SSI), a Irradiância Solar Total é a potência total de energia proveniente do Sol por unidade de área a uma Unidade Astronômica (UA), que é a distância entre o Sol e a Terra.

As medidas realizadas sobre a radiação solar ocorrem em função de comprimento de onda, a Lei de Wien relaciona comprimento de onda referente a máxima emissão de radiação eletromagnética de corpo negro e sua temperatura:

$$\lambda_{max} = \frac{\alpha}{T}$$

Onde λ_{max} é o comprimento de onda, α constante de Wien ($2898K * \mu m$) e T a Temperatura em Kelvin. A lei de Stefan-Boltzmann diz que a energia total irradiada por unidade de área superficial de um corpo negro na unidade de tempo (potencial emissora) é diretamente proporcional à quarta potência de sua temperatura. A energia irradiada pelo Sol cobre uma ampla faixa do espectro eletromagnético, sendo que 81 por cento da energia que chega está em faixa de comprimento de onda indo do visível ao infravermelho próximo, a radiação solar recebida na superfície da Terra é de onda curta (0.1 a $0.4\mu m$) e a radiação emitida pela Terra é de onda longa (3.0 a $30.0\mu m$). Segundo Pereira et al. (2017), 44 por cento de toda radiação incidente se concentra entre 0.4 e $0.7\mu m$, como mostrado na figura 1.

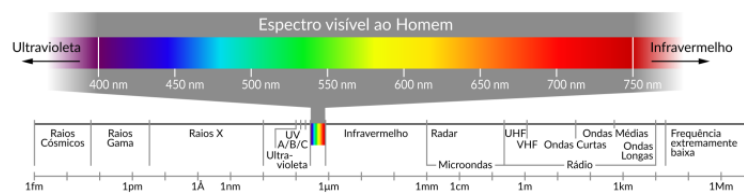


Figura 1: Espectro de radiação solar (PEREIRA et al., 2017)

A atmosfera terrestre atua sobre o espectro de emissão do Sol através de fenômenos em determinados comprimentos de onda, no fenômeno de Espalhamento a fração do feixe de radiação direta incidente é re-irradiada em várias direções pelos componentes atmosféricos, separando-se em Espalhamento de Rayleigh (constituintes gasosos) e espalhamento de Mie (partículas em suspensão), no fenômeno de Atenuação a fração de energia é subtraída do fluxo incidente e incorporada à energia interna do meio atravessado, já a Absorção consiste em fenômenos quânticos de interação da radiação a nível atômico e molecular, componentes que mais contribuem são O₃, CO₂, e vapor d’água, a figura 2 demonstra esses fenômenos.

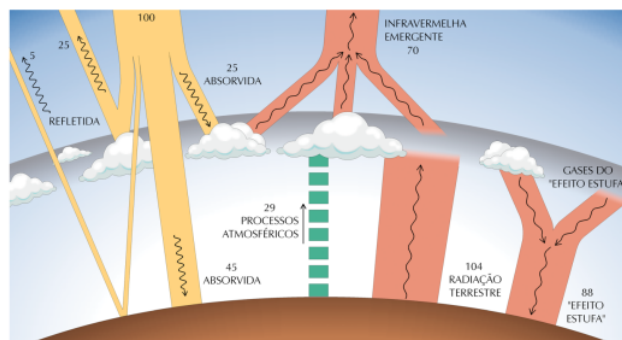


Figura 2: Processos de interação da radiação solar com a atmosfera (PEREIRA et al., 2017)

A Terra contém órbita elíptica com distância média de 149598106Km do Sol (1UA), sendo a variação da distância: *entre* $1.47 \cdot 10^8\text{Km}$ e $1.52 \cdot 10^8\text{Km}$. O fluxo de radiação solar oscila entre $1325\text{W}/\text{m}^2$ e $1412\text{W}/\text{m}^2$, com valor médio de $1366\text{W}/\text{m}^2$ definido como constante solar (I_{sc}). A quantidade de energia solar incidente em um ponto de superfície apresenta variabilidade de dois tipos, ciclo anual e ciclo diário. Entre os conceitos geométricos principais da relação Terra-Sol os seguintes mostrados pela figura 3:

- Declinação Solar (δ): ângulo formado pela inclinação do plano equatorial da Terra e a linha de direção Sol-Terra, com variação entre $-23^{\circ}27'$ e $+23^{\circ}27'$ durante um ano.

$$\delta = 23.45 \cdot \sin\left(\frac{360}{365} \cdot (ndia + 284)\right)$$

- Ângulo Solar (ω): deslocamento angular do movimento aparente do Sol devido à rotação da Terra, varia entre -180° e $+180^{\circ}$, variando 15° por hora (meio dia valor 0°).

$$\omega = (12 - hora_{solar}) \cdot 15$$

- Ângulo zenital solar (θ_z): formado entre a vertical no ponto de observação e a direção da linha que liga o mesmo ponto da superfície da Terra ao Sol, pode ser calculado conhecendo valores da latitude local (ϕ), declinação solar (δ) e ângulo horário solar (ω).

$$\cos \theta_z = \cos \phi \cdot \cos \delta \cdot \cos \omega + \sin \phi \cdot \sin \delta$$

- Elevação ou altura solar (δ): ângulo complementar ao ângulo zenital, apresenta variação diurna de $0^\circ < \delta < 90^\circ$.
- Ângulo azimutal de uma superfície (γ): projeção em um plano horizontal à superfície e o meridiano local (-180° a $+180^\circ$).
- Ângulo azimutal solar (γ_s): formado entre a projeção da direção do Sol no plano horizontal com o meridiano do observador (-180° a $+180^\circ$).
- Ângulo de incidência (θ): formado entre a normal a superfície coletora e a reta determinada pela direção da radiação solar direta, varia entre 0° e 90° e pode ser calculado pela declinação solar, a inclinação da superfície, o ângulo azimutal da superfície coletora, o ângulo horário solar e a latitude do local.

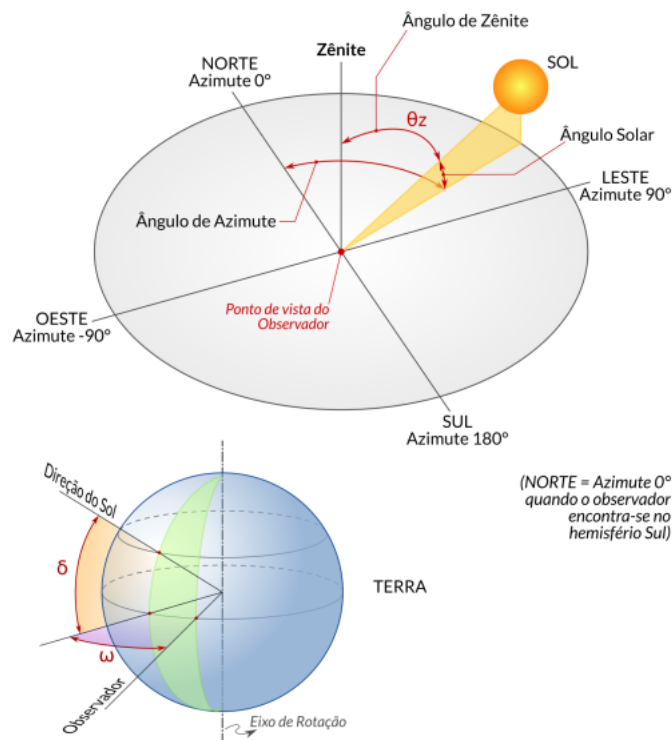


Figura 3: Ângulos notáveis em solimetria.
(PEREIRA et al., 2017)

Devido a órbita terrestre elíptica a taxa de energia solar (E_0) incidente sobre uma superfície horizontal unitária, situada no topo da atmosfera terrestre, sofre uma variação de 6,7% ao longo de um ano.

$$E_0 = I_{sc} \left(\frac{\text{distancia}_{media}}{\text{distancia}} \right)^2$$

A radiação extraterrestre (E_h) que incide perpendicularmente a uma superfície unitária imaginária situada no topo da atmosfera é função da constante solar multiplicada pela correção devido à órbita de terra e o cosseno do ângulo zenital

$$E_h = E_0 \cdot \cos(\theta z) = I_{sc} \left(\frac{\text{distancia}_{media}}{\text{distancia}} \right)^2 \cdot \cos(\theta z)$$

Dentre os principais componentes da radiação, temos os demonstrados pela figura 4:

- Radiação Solar: energia oriunda do Sol por meio de ondas eletromagnética (J).
- Irradiância Solar (G): taxa de energia solar ou a potência que incide sobre uma superfície em um determinado momento por unidade de área (W/m^2).
- Irradiação Solar (H): quantidade de energia que incide sobre uma superfície durante um certo intervalo de tempo, por unidade de área (J/m^2).
- Irradiância extraterrestre (G_0): taxa de energia incidente em um plano horizontal unitário situado no topo da atmosfera (irradiância no topo da atmosfera).
- Irradiância difusa (G_{dif}): taxa de energia incidente por unidade de área, decorrente do espalhamento do feixe solar direto pelos constituintes atmosféricos sobre uma superfície horizontal.
- Irradiância direta normal (G_n): taxa de energia por unidade de área proveniente diretamente do Sol que incide perpendicularmente à uma superfície.
- Irradiância direta horizontal (G_{dir}): taxa de energia por unidade de área do feixe solar direto numa superfície horizontal, produto entre a irradiância direta normal e o cosseno do ângulo zenital solar.
- Irradiância global solar (G): taxa de energia total por unidade de área incidente numa superfície horizontal.

$$G = G_{dif} + G_{dir} = G_{dif} + G_n \cdot \cos(\theta z)$$

- Irradiância no plano inclinado (G_i): taxa de energia total por unidade de área incidente sobre um plano inclinado em relação à superfície

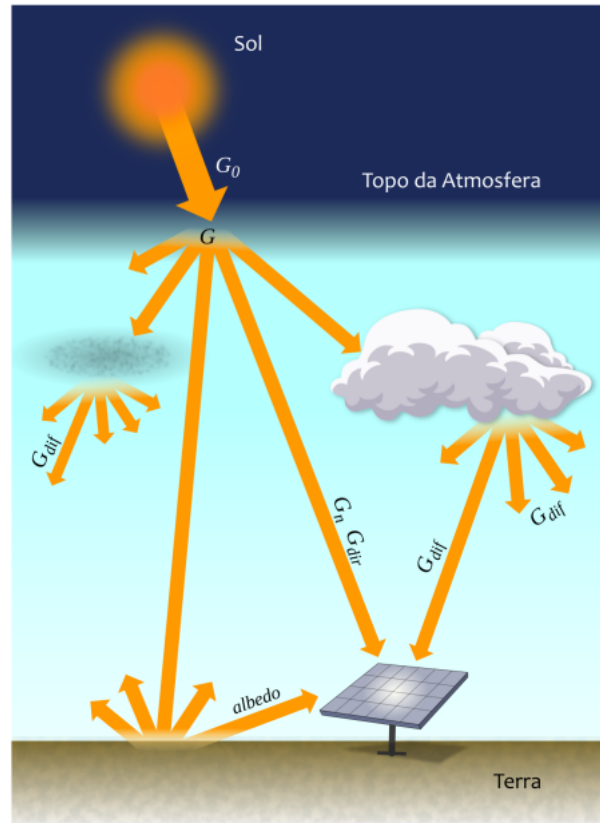


Figura 4: Componentes da irradiância solar.
(PEREIRA et al., 2017)

3.2 Modelagem numérica aplicada à Energia Solar

A modelagem numérica e computacional trata a radiação solar como fonte de energia que regula o equilíbrio termodinâmico e climático do planeta, é necessário conhecimento preciso da radiação solar incidente na superfície do planeta e da radiação solar que interage com constituintes atmosféricos, sendo uma condição básica para a meteorologia e climatologia. O desenvolvimento de modelos numéricos baseado no conhecimento dos processos radiativos que ocorrem na atmosfera é de uma importância, ao atravessar a atmosfera, radiação sofre complexas interações de espalhamento e absorção com os constituintes atmosféricos e a superfície do planeta.

O topo da atmosfera a radiação recebida são o ponto de partida para modelos numéricos que simulam processos radiativos que ocorrem na atmosfera e modulam a incidência da radiação solar na superfície, tendo foco no total diário (H_0) e total horário (I_0).

$$H_0 = \frac{24 \cdot 3600}{\pi} \cdot I_{sc} \left(\frac{dm}{D}\right)^2 \cdot \left(\cos \phi \cos \delta \cos \omega_s + \frac{\pi \cdot \omega_s}{180} \cdot \sin \phi \sin \delta\right)$$

$$I_0 = \frac{12 \cdot 3600}{\pi} \cdot I_{sc} \left(\frac{dm}{D}\right)^2 \cdot \left(\cos \phi \cos \delta \cdot (\sin \omega_2 - \sin \omega_1) + \frac{\pi \cdot (\omega_2 - \omega_1)}{180} \cdot \sin \phi \sin \delta\right)$$

ω_s = ângulo horário do por do Sol = $\arccos(-\tan \phi \tan \delta)$.

ω_i = ângulo horário da i-ésima hora.

$$H_0 = J/(m^2 \cdot dia)$$
$$I_0 = J/(m^2 \cdot hora)$$

Como um princípio geral dos modelos satelitais, onde os dados são coletados por satélites artificiais, deve existir um balanço de energia na atmosfera de forma que a energia incidente no topo da atmosfera deve ser igual a soma de energia que retorna ao espaço e a energia que é absorvida na atmosfera e na superfície do planeta.

$$E_{TOA} \downarrow = E_{TOA} \uparrow + E_{atm} + E_{solo}$$

3.3 Modelo Brasil-SR

O modelo Brasil-SR é um modelo físico para obtenção de estimativas da radiação solar incidente na superfície que combina aproximação em dois-fluxos na solução da equação de transferência radiativa, desenvolvido na Alemanha, adaptado e aperfeiçoado no Brasil (INPE e UFSC). O uso dos parâmetros é determinado de forma estatística a partir de imagens de satélite, a cobertura de nuvens é considerada como principal fator de modificação da transmitância atmosférica, e as demais propriedades ópticas da atmosfera são parametrizadas a partir de seus valores climatológicos.

A obtenção de uma estimativa da radiação solar incidente na superfície pode ser dividida em 3 etapas:

- Tratamento dos dados climatológicos e imagens de satélite.
- Aplicação do método de dois-fluxos.
- Cálculo da radiação global.

Para a determinação do Coeficiente de Cobertura de Nuvens (Ceff) é necessário composição de imagens de céu claro e de céu completamente encoberto por meio de análise estatística de imagens do período de interesse, valores de Ceff descrevem dois efeitos causados pela presença de nuvens:

- Cobertura de nuvens.
- Variações espaciais da profundidade óptica das nuvens.

O modelo Brasil-SR assume a existência de uma relação linear entre a irradiância global na superfície e o fluxo de radiação refletida no topo da atmosfera. A figura 5 demonstra o modelo Brasil-SR.

$$\phi_g = \phi_0 \{ (\tau_{clear} - \tau_{cloud}) \cdot (1 - C_{eff}) + \tau_{cloud} \}$$

ϕ_g = fluxo de radiação solar incidente na superfície.
 ϕ_0 = radiação incidente no topo da atmosfera.

$\tau_{clear}, \tau_{cloud}$ = transmitância atmosférica em condições de céu claro e completamente nublado.

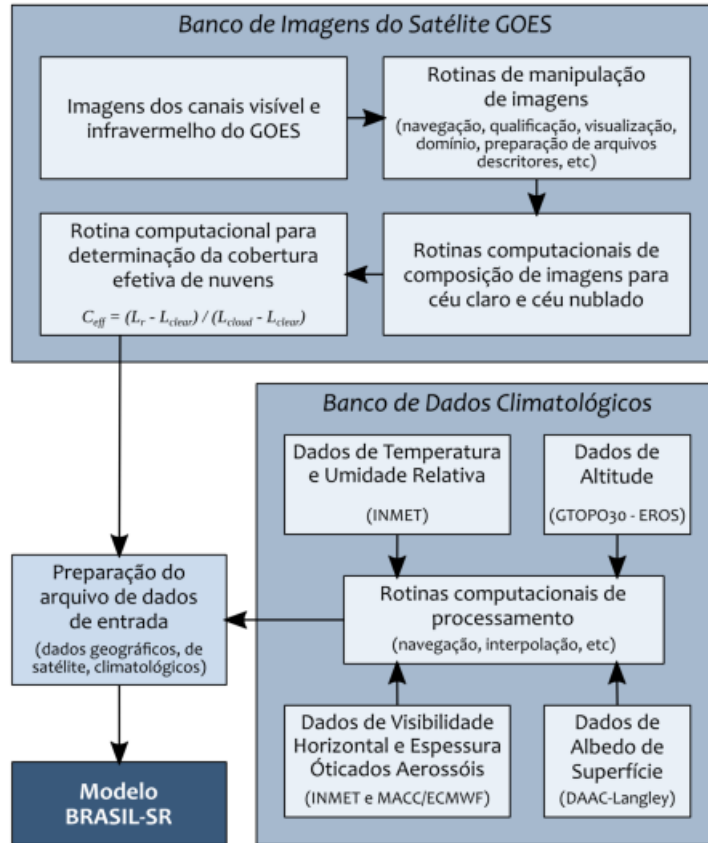


Figura 5: Modelo Brasil-SR. (PEREIRA et al., 2017)

3.4 Modelo de Céu Claro

O modelo de céu claro escolhido para algoritmo de validação é o Modelo C (IQBAL, 1983) baseado nos modelos de Bird e Hulstrom, por meio de parametrizações é possível determinar as irradiâncias normal, difusa por espalhamento de Rayleigh, difusa por espalhamento de aerossóis e consequentemente, irradiância global solar.

A irradiância normal é determinada a partir da constante solar (I_{sc}), transmitância por espalhamento de Rayleigh (τ_r), transmitância por ozônio (τ_o), transmitância por gases uniformes (τ_g), transmitância por vapor de água (τ_w), transmitância por aerossóis (τ_a) e uma constante inclusa devido ao intervalo espectral considerado no modelo.

$$I_n = 0.9751 \cdot I_{sc} \cdot \tau_r \tau_o \tau_g \tau_w \tau_a$$

Cada transmitância é calculada separadamente, a transmitância por espalhamento de Rayleigh (τ_r) é calculada por meio da massa óptica relativa e absoluta (m_r e m_a), que podem ser estipuladas por valores de pressão local, assim como a transmitância por gases uniformes.

$$\begin{aligned}\tau_r &= \exp[-0.0903m_a^{0.84}(1.0 + m_a - m_a^{1.01})] \\ \tau_g &= \exp(-0.0127m_a^{0.26})\end{aligned}$$

A transmitância por ozônio pode ser calculada como:

$$\begin{aligned}\tau_o &= 1 - \alpha_o \text{ onde} \\ \alpha_o &= [0.1611U_3(1.0 + 139.48U_3)^{-0.3035} - 0.002715U_3(1.0 + 0.044U_3 + 0.0003U_3^2)^{-1}]\end{aligned}$$

Sendo U_3 o comprimento óptico relativo que a radiação incidente passa pela cama de ozônio, levando em conta argumentos como a grossura da mesma em centímetros por meio de aproximação e a massa óptica relativa.

A transmitância por vapor de água pode ser calculada como:

$$\begin{aligned}\tau_w &= 1 - \alpha_w \text{ onde} \\ \alpha_w &= 2.4959U_1[(1.0 + 79.034U_1)0.6828 + 6.385U_1]^{-1}\end{aligned}$$

Sendo U_1 o comprimento de caminho óptico relativo corrigido por pressão de precipitação de água, podendo ser estipulado a partir da quantidade de água precipitável em uma região e a massa óptica relativa.

A transmitância por aerossóis leva em conta o valor k_a que é uma atenuação espectral de dois tamanhos de onda usados comumente em modelos parecidos ao proposto, sendo estes tamanhos 0.38 e $0.5\mu m$

$$\begin{aligned}\tau_a &= \exp[-k_a^{0.873}(1.0 + k_a - k_a^{0.7088})m_a^{0.9108}] \text{ onde} \\ k_a &= 0.2758k_{a\lambda|0.38\mu m} + 0.35k_{a\lambda|0.5\mu m}\end{aligned}$$

A irradiância difusa por espalhamento de Rayleigh é baseada além das transmitâncias anteriormente citadas, a transmitância de radiação direta por meio de absorção de aerossóis.

$$\begin{aligned}I_{dr} &= 0.79 \cdot I_{sc} \cdot \cos \theta_z \cdot \tau_o \tau_g \tau_w \tau_{aa} \cdot 0.5(1 - \tau_r)/(1 - m_a + m_a^{1.02}) \text{ onde} \\ t_{aa} &= 1 - (1 - \omega_o)(1 - m_a + m_a^{1.06})(1 - \tau_a)\end{aligned}$$

Iqbal (1983) recomenda que, caso não se possua uma retratação mais precisa de ω_o , o valor que deve ser utilizado é de 0.9.

A irradiância difusa por espalhamento de aerossóis é dada pela seguinte fórmula:

$$I_{da} = 0.79 \cdot I_{sc} \cdot \cos \theta_z \cdot \tau_o \tau_o \tau_g \tau_w \tau_{aa} \cdot F_c (1 - \tau_{as}) / (1 - m_a + m_a^{1.02}) \text{ onde}$$

$$t_{as} = \frac{\tau_a}{\tau_{aa}}$$

Iqbal (1983) recomenda que, para o modelo, o valor 0.84 deve ser usado para F_c caso não exista melhor informação sobre os aerossóis.

Por final, a irradiância solar global é dada pela soma da irradiância normal multiplicada com o ângulo zenital com as irradiâncias difusas, Iqbal (1983) descreve que caso seja necessário uma maior precisão na modelagem de céu claro, os valores de albedo (reflexão) devem ser calculados e multiplicados à global posteriormente.

$$I = I_n \cos \theta_z + I_d = (I_n \cos \theta_z + I_{dr} + I_{da}) \left(\frac{1}{1 - p_g p_a} \right)$$

onde p_g e p_a respectivamente valores de albedo do chão e atmosférico.

4 Materiais e Métodos Utilizados

Para o desenvolvimento da última fase do algoritmo de validação da rede SONDA, que valida os dados obtidos comparados ao modelo de céu claro, foi necessário um período de estudo de toda fundamentação teórica com base em [Iqbal \(1983\)](#) e na apostila de [Pereira et al. \(2014\)](#). Um entendimento maior do próprio sistema de classificação, assim como suas fases subdivididas de validação foi necessário. As estações da rede SONDA empregam piranômetros e utilizam radiômetros padrão secundário onde as estimativas de incertezas nas medições de irradiação são as constantes da ISO 9060 para instrumentos de padrão secundário. Os procedimentos de operação e manutenção seguem as recomendações estabelecidas pela WMO (*World Meteorological Organization*) para a BRSN (*Baseline Surface Radiation Network*) ([DRIEMEL et al., 2018](#)), que definem critérios para as estações de coleta. A aquisição de dados deve ser feita em intervalos de um segundo e o armazenamento de valores médios a cada minuto. O sistema de classificação de dados da rede SONDA possui atualmente quatro fases de validação para variáveis anemométricas e três fases para variáveis radiométricas e meteorológicas, seguindo a seguinte ordem:

Anemométricas

1. Dado suspeito quando fisicamente impossível.
2. Dado suspeito quando o evento é extremamente raro.
3. Dado suspeito quando apresenta uma evolução temporal não condizente com o esperado para a variável.
4. Dado suspeito quando inconsistente com medidas apresentadas por outras variáveis da mesma estação.

Meteorológicas

1. Dado suspeito quando fisicamente impossível.
2. Dado suspeito quando o evento é extremamente raro.
3. Dado suspeito quando apresenta uma evolução temporal não condizente com o esperado para a variável.
4. (Ainda não aplicado) Dado suspeito caso a medida esteja inconsistente quando comparada com estimativas de modelos computacionais.

Radiométricas

1. Dado suspeito quando fisicamente impossível.
2. Dado suspeito quando o evento é extremamente raro.
3. Dado suspeito quando inconsistente com medidas apresentadas por outras variáveis da mesma estação.
4. (Ainda não aplicado) Dado suspeito caso a medida esteja inconsistente quando comparada com estimativas de modelos computacionais.

A aprovação em cada etapa é requisito para a continuidade do processo. Assim, somente quando um dado for considerado aprovado numa etapa, a etapa seguinte será iniciada. Se não houver aprovação, o processo será interrompido e o dado receberá o código equivalente a suspeito. Se houver aprovação, o dado receberá o código de aprovado e a etapa seguinte será iniciada, como demonstrado pela figura 6.

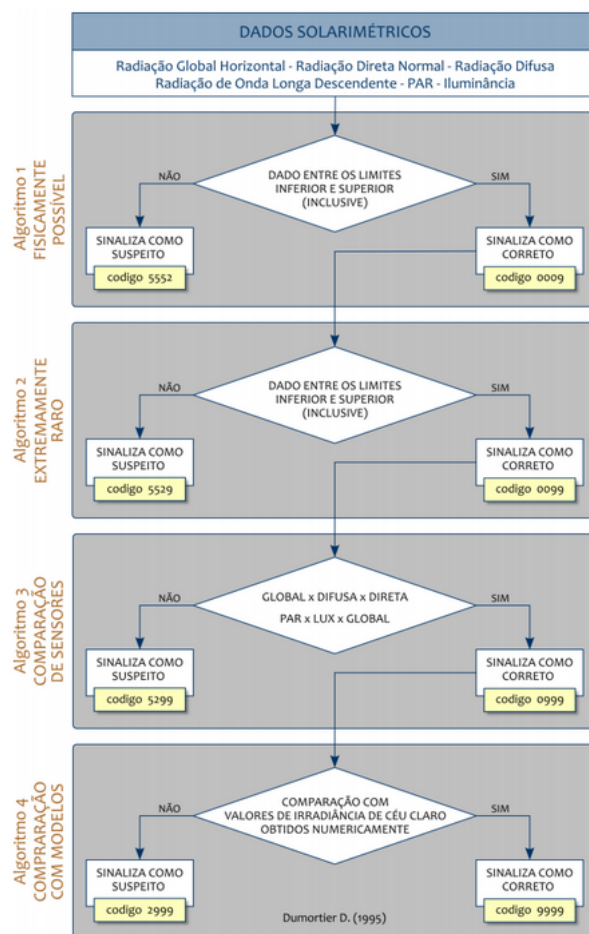


Figura 6: Algoritmo de classificação de dados da rede SONDA. (PEREIRA et al., 2017)

A versão mais atual do algoritmo de classificação é a de número 3.5, onde a maior porcentagem do código está em linguagem Java, com códigos extras criados em Fortran. Como uma primeira tarefa principal foi necessário converter e reestruturar o código para uma única linguagem como um *software* único, facilitando a execução e portabilidade em ambientes variados, a linguagem escolhida após algumas reuniões foi Python, pela sua facilidade de uso e importação de bibliotecas.

5 Análises e Resultados

Após a escolha da linguagem, a reestruturação do código foi feita observando-se falhas ou resultados imprecisos, sendo necessário reformulação do código local com os ajustes necessários. Novamente o código foi revisado buscando melhorias com foco no tempo de execução, garantindo além de melhor confiabilidade uma disponibilidade maior no uso.

O início da quarta fase de validação foi feito, preparando as variáveis necessárias, assim como fórmulas auxiliares do modelo de céu claro, sendo somente necessário desenvolver as fórmulas principais, aplicar mudanças necessárias ao código e incorporá-lo ao *software* previamente reestruturado, finalizando assim o algoritmo de classificação.

Após a reestruturação do código principal e sua determinada revisão, notou-se uma melhoria na precisão dos dados calculados, dados sinalizados anteriormente como suspeitos passaram a ser checados novamente, passando pela etapa de validação. A versão anterior também apresentava etapas de validação ignoradas dependendo de certas variáveis na inserção, o que foi corrigido e regularizado melhorando conceitos de integridade e confiabilidade dos dados coletados e verificados. Melhorias em passagem de dados, leitura de arquivos e exclusão de etapas redundantes ou duplicadas diminuíram o tempo de execução do algoritmo melhorando sua disponibilidade e resultados rápidos para verificação.

Existem mudanças significativas no código final com a passagem de estruturação da linguagem Java para a linguagem Python, tanto no nível de sintático de código quanto no de execução e comportamento de variáveis. Como um exemplo simples destaca-se a diferença na decisão de arredondamento de casas decimais em variáveis de ponto flutuante quando o valor específico está em seu intermédio, enquanto na linguagem Python a decisão é inclusiva de incrementar a casa decimal, na linguagem Java a decisão é de truncamento. Outra diferença se refere ao número de linhas necessárias para funções de mesmo comportamento semântico, na linguagem Python se conseguiu restringir o número de linhas com média de 300 por arquivo de código em comparação na linguagem Java, permitindo uma melhor leitura e compreensão do código como um todo, as figuras 7 e 8 demonstram a distinção de um mesmo bloco de código nas diferentes linguagens.

```

    }
    j++;
}
}
if (contWdirValid >= 40) {
    if (variation_wd1018h > 10) {
        loader.code[i][25] = 999;
        lastWd10Valid = loader.code[i][25];
    } else {
        loader.code[i][25] = 299;
    }
} else {
    int l = 0;
    while (l <= wd1018h) {
        if ((loader.code[i + l][25] != 3333) && (loader.code[i + l][25] != 552) && (loader.code[i + l][25] != 529)) {
            if (i == 0) {
                loader.code[i][25] = 559;
            } else {
                loader.code[i][25] = lastWd10Valid;
            }
        }
        l++;
    }
    contWdirValid = 0;
    wd10_max = 0;
    wd10_min = 100;
}
}
if (i >= totalWd1018h 2) {
    loader.code[i][25] = loader.code[totalWd1018h_1][25];
}
}
}

```

Figura 7: Bloco de código na linguagem Java.

```

if self.contWdirValid >= 40:
    if self.variation_wd1018h > 10:
        self.loader.code[i][25] = 999
        self.lastWd10Valid = self.loader.code[i][25]
    else:
        self.loader.code[i][25] = 299
else:
    l = 0
    while l <= self.wd1018h:
        if self.loader.code[i+l][25] != 3333 and self.loader.code[i+l][25] != 552 and self.loader.code[i+l][25] != 529:
            if i == 0:
                self.loader.code[i][25] = 559
            else:
                self.loader.code[i][25] = self.lastWd10Valid
        l += 1
    self.contWdirValid = 0
    self.wd10_max = 0
    self.wd10_min = 999

if i >= totalWd1018h 2:
    self.loader.code[i][25] = self.loader.code[totalWd1018h_1][25]

# End of the routine validation: Wind Direction 10m (*) level 3
# -----

```

Figura 8: Bloco de código na linguagem Python.

Um dos grandes fatores do *software* é o acoplamento de último estágio de classificação utilizando o modelo de céu claro, como dito no capítulo 4 após toda a reestruturação do código para Python as equações secundárias e todas as variáveis necessárias foram passadas ao modelo do código, necessitando somente modelar e portar as equações principais do modelo de Iqbal (1983), facilitando o trabalho posterior e a implementação ao *software* mais complexo.

6 Conclusão

O algoritmo de validação de dados da rede Sonda é uma importante fase da modelagem computacional com foco no cenário de pesquisas de ciência terrestre, a necessidade de quantificar a disponibilidade e variabilidade do potencial eólico e solar é extremamente relevante no cenário atual de energia. Os dados coletados nas estações da rede SONDA precisam ter garantia de confiabilidade e integridade em todas suas fases de validação e classificação, podendo ser uma base transparente para usos futuros. O último processo de validação se baseia em comparações com o modelo de céu claro estabelecido por [Iqbal \(1983\)](#) e sua aplicação ao algoritmo de classificação de dados é de suma importância para a finalização do mesmo. A reestruturação do algoritmo para a linguagem Python permitiu sua maior portabilidade, além de estruturar todos os passos de validação em um único *software* executável, e a preparação das equações e variáveis para a finalização da última etapa de validação permite concluir com facilidade o *software* proposto.

Referências

- DRIEMEL, A. et al. Baseline surface radiation network (bsrn): structure and data description (1992–2017). *Earth System Science Data*, v. 10, n. 3, p. 1491–1501, 2018. Disponível em: <<https://www.earth-syst-sci-data.net/10/1491/2018/>>. Citado na página 16.
- IQBAL, M. Chapter 6 - solar spectral radiation under cloudless skies. In: IQBAL, M. (Ed.). *An Introduction to Solar Radiation*. Academic Press, 1983. p. 107 – 168. ISBN 978-0-12-373750-2. Disponível em: <<http://www.sciencedirect.com/science/article/pii/B9780123737502500112>>. Citado 9 vezes nas páginas 5, 6, 7, 13, 14, 15, 16, 20 e 21.
- PEREIRA, E. et al. *Atlas Brasileiro de Energia Solar – 2ª Edição*. [S.l.: s.n.], 2017. ISBN 978-85-17-00089-8. Citado 6 vezes nas páginas 7, 8, 9, 11, 13 e 17.
- PEREIRA, E. et al. *TREINAMENTO TEÓRICO-PRÁTICO EM AVALIAÇÃO DE POTENCIAL DE RECURSO ENERGÉTICO SOLAR*. [S.l.: s.n.], 2014. Citado na página 16.