

MODELAGEM COMPUTACIONAL DE CONTROLE DE QUALIDADE DE DADOS DE IRRADIÂNCIA SOLAR

Marcus Vinicius Contes Calca - mcontes@outlook.com
Matheus Rodrigues Raniero - matheus.raniero@hotmail.com
Faculdade de Ciências Agrônômicas - UNESP de Botucatu.
Carlos Roberto Pereira Padovani - crppadovani@gmail.com
Faculdade de Botucatu UNIBR.
Domingos Mario Zeca Fernando - dmzeca@gmail.com
Faculdade de Ciências Agrônômicas - UNESP de Botucatu
e Universidade Pedagógica de Moçambique - UPM.
Alexandre Dal Pai - adalpai@fca.unesp.br
Faculdade de Ciências Agrônômicas - UNESP de Botucatu.

Resumo. *O objetivo deste estudo foi desenvolver uma modelagem computacional de controle de qualidade de dados medidos de irradiância solar global, difusa e direta obtidos a partir da Estação de Radiometria Solar de Botucatu da Faculdade de Ciências Agrônômicas - UNESP de Botucatu. O correto dimensionamento de fontes de dados permite que órgãos públicos, entidades acadêmicas e organizações não-governamentais planejem e gerenciem, de forma mais adequada, recursos energéticos renováveis. Além de permitir que instituições de ensino superior e pesquisa correlacionem informações de energia solar em estudos de conversão de energia térmica, fotovoltaica e de biomassa, e em estudos sobre a influência dos raios solares em cultivos de campo. A fase inicial de dimensionamento de uma fonte de dados é o procedimento de controle de qualidade, que neste caso, foi desenvolvido se baseando em uma série de filtros computacionais, onde foram aplicados testes físicos e lógicos, a partir da linguagem de consulta estruturada SQL. A execução dos filtros se deu por meio do sistema de gerenciamento de banco de dados - SGBD - MySQL. Pôde-se observar que do total de dados brutos, somente 40% restaram após o controle de qualidade, originando, desta forma, um conjunto de dados de qualidade. A modelagem computacional em seu estado final forneceu métodos que puderam ser adaptados a outros sistemas de gerenciamento de banco de dados ou linguagens de programação, de acordo com a necessidade da pesquisa. Os filtros foram construídos sob algoritmos que permitiram a seleção de valores medidos considerados adequados.*

Palavras-chave: *Irradiância solar, Radiação solar medida, Filtragem de dados.*

1. INTRODUÇÃO

O crescimento da demanda, carência de oferta, e restrições socioeconômicas e ambientais, são indicativos de que recursos energéticos alternativos sejam mais explorados no Brasil. A evolução da tecnologia neste setor resultou em diversos avanços no decorrer dos anos, atuando em estudos sobre a geração, transmissão e distribuição final da energia. Onde conjuntos de dados devidamente estruturados podem ser interpretados por órgãos públicos, entidades acadêmicas e organizações não-governamentais, no sentido de apoiar o planejamento e a gestão de recursos energéticos renováveis, nas diferentes regiões brasileiras, de acordo com a ANEEL (2005). Além de possuir importância fundamental em áreas como arquitetura, agricultura, engenharia e biologia, reforça Muneer *et al.* (2007).

A perspectiva levantada coloca em evidência a importância de se ter bons conjuntos de dados medidos, pois servem de subsídio em diversas aplicações da energia solar. No entanto o caminho percorrido, até a obtenção de um conjunto de dados de qualidade, contém diversos impasses. Muneer *et al.* (2007) citam que existem duas grandes categorias de problemas em dados medidos de irradiância solar, os erros oriundos dos equipamentos e os erros oriundos das atividades de manutenção. Pashiards e Kalogirou (2016) apontam algumas causas de erros de equipamentos, como o desalinhamento do anel de sombreamento da radiação solar difusa, nivelamento incorreto dos sensores, excremento de pássaros sobre a cúpula, campos elétricos na proximidade dos cabos, e o mais recorrente, apagões no sistema de fornecimento de energia. Já os erros causados por atividades de manutenção são ocasionados por processos repetitivos de processamento de dados e por pouco domínio de conhecimento do perfil de irradiância incidente na localidade.

O controle de qualidade em dados de irradiância solar é uma atividade de grande importância no ramo meteorológico, responsável, principalmente, por identificar erros e corrigi-los. Qualificando-se portanto, como um excelente procedimento para filtragem de dados espúrios e obtenção de conjuntos de dados de qualidade. Muitos dos procedimentos encontrados na literatura apresentam filtros de controle de qualidade da irradiância solar global, difusa e direta, que se baseiam na relação física entre elas, delimitando valores máximos e mínimos possíveis, de acordo com a localidade da estação de coleta de dados. Outro procedimento bastante utilizado é a análise gráfica de uma determinada série temporal de irradiâncias, que na maioria dos casos é feita após a aplicação dos filtros físicos (Muneer *et al.*, 2007; Moreno-Tejera *et al.*, 2015; Pashiard e Kalogirou 2016; Il Sung-Zo *et al.*, 2017). Neste sentido, uma das maneiras de

contornar a problemática apresentada neste estudo é o desenvolvimento de filtros de controle de qualidade dos dados, contribuindo, fundamentalmente, no desenvolvimento de procedimentos computadorizados, que podem auxiliar na automatização do processo e eliminar aspectos subjetivos do processamento por meio da interação humana.

Este estudo propõe uma modelagem computacional de controle de qualidade de dados medidos de irradiância solar baseada em uma série de testes físicos e lógicos de filtragem de valores. Os filtros foram adaptados para serem aplicados de forma automatizada por meio da linguagem de consulta estruturada SQL, na plataforma de gerenciamento de banco de dados MySQL. A proposta de controle de qualidade visa ser aplicada nas irradiâncias global, difusa e direta, visto que são as mais comuns em estações solarimétricas e meteorológicas.

2. MATERIAL E MÉTODOS

Este estudo apresenta uma modelagem computacional para controle de qualidade de dados de irradiância solar. Portanto, contempla algoritmos com filtros físicos e lógicos de classificação das medidas. Necessitando, então, que em sua recriação, adaptações sejam feitas, principalmente com relação ao conjunto de dados. Uma vez que, neste caso, compreende um padrão lógico de distribuição das colunas, possibilitando a execução dos algoritmos na melhor condição possível.

2.1 Estação e Instrumentação

A Estação Experimental de Radiometria Solar (latitude 22,85° sul, longitude 48,45° oeste, 786 metros) está inserida no ambiente rural da Faculdade de Ciências Agrônomicas - UNESP de Botucatu, e exerce atividades de monitoramento da radiação solar desde 1996. As informações obtidas, nas atividades de monitoramento, são enviadas para o Laboratório de Radiometria Solar, localizado na mesma instituição da estação.

O Laboratório de Radiometria Solar contempla um banco de dados de mais de 20 anos de medidas em diferentes repositórios de armazenamento. As medidas são armazenadas em arquivos de texto no padrão ASCII e em uma tecnologia computacional de gerenciamento de banco de dados, conhecida por SGBD MySQL, ou Sistema de Gerenciamento de Banco de Dados. O banco de dados, em ambos os repositórios citados, compreende as seguintes componentes solares: radiação global, radiação direta na incidência normal, radiação difusa (métodos da diferença, disco e anel de sombreamento), radiação refletida, radiação global e difusa em superfície inclinada, radiação global em superfície vertical, radiação espectral ultravioleta, fotossinteticamente ativa e infravermelha curta, além das radiações de ondas longas terrestre e atmosférica.

A modelagem computacional apresentada por este estudo utilizou apenas os dados medidos das irradiâncias global, difusa, e direta, pois são as medidas mais encontradas em estações meteorológicas e de radiometria solar ao redor do mundo. A irradiância global (I_g) foi medida por um piranômetro Eppley - PSP, a irradiância direta em incidência normal (I_b) foi medida por um pireliômetro Eppley - NIP equipado com um dispositivo Eppley - ST-3 de rastreamento solar, enquanto a irradiância solar difusa (I_d) foi medida por um piranômetro Eppley - PSP equipado com um anel de sombreamento MEO, e corrigida a partir dos fatores apresentados por Dal Pai *et al.* (2016). As medidas foram gerenciadas por meio de diferentes sistemas de aquisição de dados ao longo dos anos, como o CR23X, CR1000 e CR3000 da Campbell Scientific.

2.2 Banco de Dados

Conforme descrito na subseção anterior (subseção 2.1) o banco de dados de radiometria solar abrange dois tipos de repositórios de dados e informações, os arquivos em formato texto no padrão ASCII e os arquivos incorporados a tecnologia de gerenciamento de banco de dados MySQL.

O arquivo ASCII é constituído por letras, números e símbolos, além de ser considerado como um padrão universalmente aceito. Este arquivo é geralmente utilizado para ser inserido em programas computacionais matemáticos como o Microsoft Excel®, Microcal Origin® e MatLab®.

A tecnologia de gerenciamento de banco de dados MySQL é um programa computacional do tipo SGBD, ou seja, um Sistema de Gerenciamento de Banco de Dados. O SGBD é uma ferramenta de computador responsável por permitir o acesso, organização e manipulação de dados a partir de tabelas e relacionamentos. As tabelas de um SGBD são estruturadas por linhas e colunas, de forma similar as planilhas eletrônicas do Microsoft Excel®, e são interconectadas a partir de relacionamentos, conhecidos também por associações, ou seja, uma tabela pode estar associada a outra. A maior vantagem ao se utilizar de um SGBD como um repositório de dados é a existência de uma linguagem interna, de alto nível, de gerenciamento do conteúdo, conhecida como SQL (*structured query language*), ou linguagem de consulta estruturada (Takai *et al.*, 2005).

O procedimento de controle de qualidade se utiliza do arquivo de dados brutos em partição instantânea. Os valores foram lidos pelo sistema de aquisição de dados (*data logger*) em um intervalo de 5 segundos, e os valores médios, em um intervalo de 5 minutos, foram calculados e armazenados. Neste sentido, o arquivo contempla 288 valores registrados para um dia, uma vez que, mesmo no período da noite, o *data logger* persiste em execução.

2.3 MySQL e HeidiSQL

O MySQL é um SGBD que atua como servidor, enquanto o HeidiSQL é uma interface de gerenciamento do SGBD MySQL que atua como cliente. O conceito de cliente e servidor é definido da seguinte forma por Takai *et al.* (2005): o servidor é o elemento computacional responsável por prover serviços, por consequência, o cliente é o elemento computacional responsável por requisitar e utilizar tais serviços. Desta forma, o MySQL (servidor) e o HeidiSQL (cliente) são elementos computacionais em forma de software livre e forneceram as funcionalidades necessárias para executar as operações computacionais deste estudo, desde a alocação do banco de dados, até o desenvolvimento do algoritmo de filtragem das irradiâncias solares. Por proporcionarem um ambiente genérico, os procedimentos podem ser operados por outros tipos de SGBD, ou até mesmo por outras linguagens de programação ou programas de computadores, neste estudo o SGBD MySQL e o cliente HeidiSQL foram aplicados devido a simplicidade de operação, resposta em tempo ágil, devido ao grande volume de dados utilizados, e com foco na integração desta modelagem computacional a outras modelagens computacionais do Laboratório de Radiometria Solar de Botucatu, que se utilizam das mesmas ferramentas.

2.4 Procedimentos de Controle de Qualidade

Os procedimentos de controle de qualidade foram baseados na aplicação de filtros de seleção de dados. Os filtros aplicaram uma série de testes lógicos nos valores, retornando uma resposta *booleana*, onde são possíveis apenas dois estados (sim ou não). Neste sentido, foi implementado um algoritmo, pela linguagem de consulta estruturada SQL, diretamente na plataforma MySQL, que aplicou o processamento de forma sequencial.

O processamento foi responsável por classificar se um determinado valor de irradiância solar estava adequado ou não para compor o arquivo final de dados, desta forma, os valores foram submetidos a três filtros, caso estivessem de acordo com as condições aplicadas pelos filtros, eram classificados como adequados, caso não correspondessem a nenhuma das condições, eram classificados como inadequados. O algoritmo de filtragem não eliminou registros do banco de dados, apenas os classificou de acordo com suas respostas. Os valores classificados como adequados foram selecionados e realocados para outras tabelas, já os valores classificados como inadequados foram mantidos armazenados na tabela original de dados na plataforma MySQL.

A primeira fase de filtragem foi responsável por marcar como nulos todos os valores das irradiâncias global (I_g), difusa (I_d) e direta na horizontal (I_b) que fossem negativos. Posteriormente a eliminação dos valores negativos, esta primeira etapa realizou a seleção de valores registrados entre as horas de fotoperíodo, onde foram utilizadas informações das efemérides astronômicas. Neste caso, o filtro é responsável por selecionar apenas os valores compreendidos entre o horário do nascer ($12-H/15$) e pôr do sol ($12+H/15$), onde H representa o ângulo horário do nascer e pôr do sol (Eq. 1), juntamente a valores que se encontram entre o ângulo zenital (Eq. 3) no valor de 80° , que representa o sol no horizonte.

$$H = \arccos(-tg \phi tg \delta) \quad (1)$$

onde ϕ é a latitude do local e δ (Eq. 2) é a declinação solar.

$$\delta = 23.45 \operatorname{sen}\left(\frac{360}{365}(DJ - 80)\right) \quad (2)$$

onde DJ é o dia juliano, que representa os dias decorridos do ano.

$$\cos Z = \operatorname{sen} \phi \operatorname{sen} \delta + \cos \phi \cos \delta \operatorname{sen} h \quad (3)$$

onde ϕ é a latitude do local, δ é a declinação solar e h é o ângulo horário.

A segunda fase de filtragem se baseou na seleção de valores que compreendiam as frações instantâneas das irradiâncias global ($K_t = I_g/I_o$), difusa ($K_d = I_d/I_g$) e direta ($K_b = I_b/I_g$), onde I_o representa a irradiância no topo da atmosfera (Eq. 4). As frações foram calculadas e gravadas em colunas na plataforma MySQL, onde seu resultado deveria compreender um valor entre zero e um, para este filtro, descartando da seleção os valores fora do padrão. Os resultados entre o intervalo de zero e um representam para K_t a quantidade de irradiância global, com relação a irradiância no topo da atmosfera (I_o), que chegou a superfície, representam também, para K_d a porção de irradiância difusa (I_d) compreendida na irradiância global (I_g), e para K_b a porção de irradiância direta na horizontal (I_b) compreendida na radiação global (I_g).

$$I_o = I_{sc} E_o \cos Z \quad (4)$$

onde I_{sc} é a constante solar, que representa 1367 W/m^2 , e E_o é a correção da excentricidade da órbita terrestre (Eq. (5)).

$$E_o = 1 - 0,0009464 \operatorname{sen}(F) - 0,01671 \operatorname{cos}(F) - 0,0001489 \operatorname{cos}(2F) - 0,00002917 \operatorname{sen}(3F) - 0,0003438 \operatorname{cos}(4F) \quad (5)$$

onde F é $360^\circ DJ/365$.

A terceira fase se baseou na seleção de valores que apresentassem resposta positiva a uma série de comparações lógicas propostas por Dal Pai *et al.* (2016), e implementadas nos estudos do Laboratório de Radiometria Solar de Botucatu, sede deste trabalho. Foram selecionados valores onde a irradiância global fosse menor do que a irradiância no topo da atmosfera, onde a irradiância direta na horizontal fosse menor do que a constante solar, e onde a irradiância difusa fosse maior do que 10% da irradiância global ($0,1 I_g$) e menor do que o valor da irradiância global. Para melhor visualização, pode-se observar na Tab. 1 as condições lógicas utilizadas nas comparações da terceira fase de filtragem.

Tabela 1 - Condições Lógicas da Terceira Fase de Filtragem.

Irradiâncias	Condições lógicas
Irradiância Global	$I_g < I_o$
Irradiância Direta	$I_b < I_{sc}$
Irradiância Difusa	$0,1 I_g \leq I_d \leq I_g$

As condições das comparações lógicas da terceira fase de filtragem (Tab. 1) foram desenvolvidas com base nas pesquisas de Kudish e Evseev (2008), além dos resultados encontrados na pesquisa de Dal Pai *et al.* (2016), onde a irradiância difusa não chegou a valores menores do que 10% da global de acordo com a análise feita, e para este caso, foi corrigida a partir dos fatores de correção do anel de sombreamento, uma vez que, este método de medida, utilizado no Laboratório de Radiometria Solar, causa uma certa subjetividade nos valores medidos da irradiância difusa.

Em termos computacionais os dados brutos em partição instantânea foram armazenados em tabelas na plataforma MySQL, os valores classificados como adequados, a partir dos testes descritos acima, foram armazenados em elementos auxiliares das tabelas, conhecidos como *view*. A *view* representa uma estrutura lógica da plataforma MySQL que é resultado de uma consulta de dados, desta forma, os valores armazenados em uma *view* necessitam, antes, passar por uma instrução de seleção na linguagem de consulta estruturada SQL, de acordo com o teste lógico.

A Fig. 1 apresenta um fluxograma das etapas do procedimento de controle de qualidade. Iniciando pela leitura do arquivo ASCII pela plataforma MySQL, percorrendo as três fases de filtragem, armazenando os valores filtrados e os não filtrados em tabelas distintas, e por fim, gerando o arquivo final, com os dados processados, no formato ASCII.

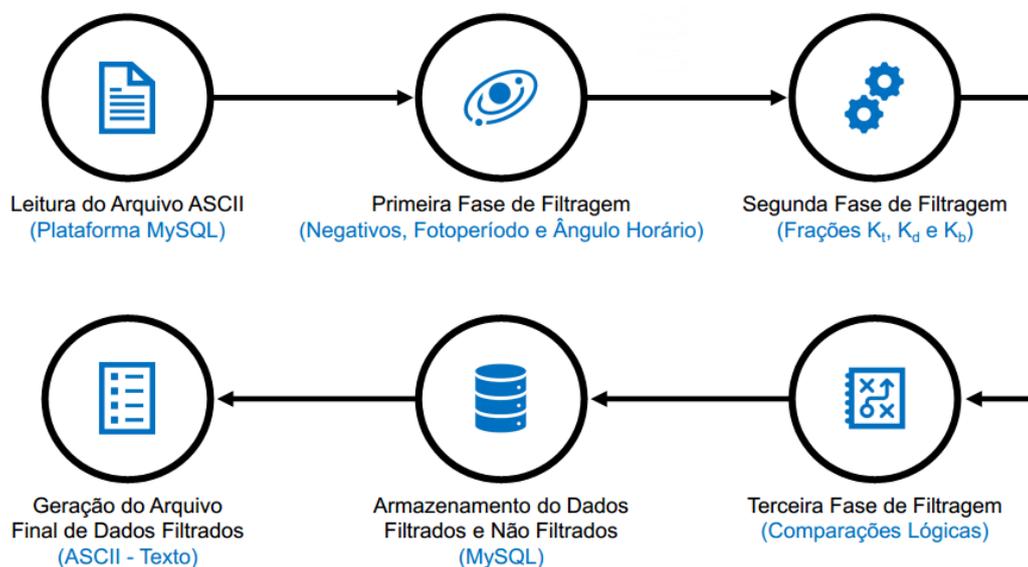


Figura 1 - Fluxograma do Procedimento de Controle de Qualidade.

3. RESULTADOS E DISCUSSÕES

O conjunto de dados utilizado como amostra para a criação da modelagem computacional pertence ao ano de 1996, e representa a partição de dados instantâneos, composta pela média das medidas em um intervalo de 5 minutos, registrando para cada dia 288 valores de irradiância solar. O arquivo original continha no total 101.972 registros das irradiâncias global, difusa e direta na horizontal em W/m^2 . No sentido de oferecer suporte aos processos de filtragem dos valores foram agregados aos dados, colunas com informações de efemérides como o ângulo horário do nascer e pôr do sol (Eq. 1), declinação solar (Eq. 2), ângulo zenital (Eq. 3) e irradiância no topo da atmosfera (Eq. 4).

Conforme informado nas seções anteriores, o controle de qualidade dos dados foi feito a partir de três fases de filtragem, de maneira estruturada, ou seja, uma etapa é totalmente dependente dos resultados obtidos na etapa anterior. Conseqüentemente, o fluxograma do procedimento de controle de qualidade (Fig. 1) ilustra de uma forma simplificada, a ordem lógica em que o algoritmo de controle de qualidade executa os procedimentos computacionais na plataforma MySQL.

Os registros das irradiâncias solares foram segmentados a partir dos meses do ano de 1996, para que desta forma, o resultado das filtragens pudesse apresentar maiores informações a comunidade científica, uma vez que, o procedimento realizado de forma computacional, a partir de algoritmos, pode ser replicado para outros universos de dados que apresentem valores medidos das irradiâncias solares. É importante ressaltar que o algoritmo foi desenvolvido a partir da linguagem SQL que é uma linguagem fundamentalmente empregada na manipulação de dados em bancos de dados relacionais, como a plataforma MySQL, portanto, as diretrizes foram programadas para fazer uma seleção dos valores que apresentassem resposta positiva as condições aplicadas pelos filtros, não eliminando registro algum do banco de dados. Neste contexto, a Tab. 2, a seguir, retrata a quantidade de valores medidos que continha o arquivo ASCII original e a quantidade que foi restando a partir das três fases de filtragem dos dados.

Tabela 2 - Quantidade de Valores Medidos.

Mês	Quantidade de Registros			
	Original	Primeira Fase	Segunda Fase	Terceira Fase
Janeiro	8.744	4.213	4.065	3.759
Fevereiro	8.217	3.853	3.820	3.702
Março	8.470	3.666	3.636	3.472
Abril	8.627	3.571	3.566	3.022
Maiο	8.609	3.325	3.324	2.879
Junho	8.331	3.118	1.929	1.372
Julho	8.783	3.402	3.402	2.792
Agosto	8.803	3.620	3.615	2.775
Setembro	8.222	3.635	3.616	3.471
Outubro	8.917	4.131	4.109	3.979
Novembro	8.495	4.156	4.141	3.715
Dezembro	7.754	3.804	3.775	3.750
Total	101.972	44.494	42.998	38.688

Nota-se a partir dos valores apresentados na Tab. 2 que o mês que possui maior quantidade de registros, com relação ao arquivo original de dados é o mês de Outubro com 8.917 valores, e o que possui menor quantidade é o mês de Fevereiro com 8.217 valores. O ano de 1996 possui uma média de 8.498 registros de valores medidos das irradiâncias solares para cada mês do ano.

Na aplicação da primeira fase de filtragem foi observado que, o mês de Julho apresentou maior diminuição na quantidade de registros das irradiâncias, reduzindo de 8.783 para 3.402 registros (5.381 valores ignorados) e o mês de Dezembro apresentou menor diminuição, reduzindo de 7.754 para 3.804 registros (3.804 valores ignorados). Esta é a etapa de filtragem que possui os maiores índices de redução dos valores, uma vez que, mantém apenas os valores das irradiâncias monitorados no período do dia, quando realmente se obtém as radiações de ondas curtas. Ignorando, portanto, os valores registrados no período noturno. Os dados apresentados na primeira fase de filtragem, mostram que aproximadamente 56% dos valores medidos, foram filtrados com relação ao arquivo original, reduzindo de 101.972 para 44.494 registros (57.478 valores ignorados).

Na aplicação da segunda fase de filtragem foi observado que, o mês de Junho apresentou maior diminuição na quantidade de registros, reduzindo de 3.118 para 1.929 registros (1.189 valores ignorados) e o mês de Julho apresentou menor resposta, não reduzindo valor algum. De acordo com esta filtragem, o mês de Junho mostrou algum tipo de erro crítico no monitoramento das irradiâncias, uma vez que, se comparado aos demais valores foi o que sofreu maior redução, sendo esta, oito vezes maior do que a segunda maior redução do filtro. Os dados apresentados na segunda fase de filtragem, mostram que aproximadamente 3% dos valores medidos foram filtrados com relação a fase anterior de filtragem, reduzindo de 44.494 para 42.998 registros (1.496 valores ignorados).

Na aplicação da terceira fase de filtragem foi observado que, o mês de Agosto apresentou maior diminuição na quantidade de registros, reduzindo de 3.615 para 2.775 registros (840 valores ignorados), enquanto o mês de Dezembro apresentou o menor índice de diminuições, reduzindo de 3.775 para 3.750 registros (25 valores ignorados). Esta é a filtragem final do procedimento de controle de qualidade das irradiâncias, portanto, os registros que restaram são os que apresentaram resposta positiva em todos as filtragens. É importante ressaltar que o procedimento de filtragem empregado não eliminou registro algum, apenas os ignorou na seleção dos valores. Por fim, os dados apresentados na terceira fase de filtragem, mostram que aproximadamente 10% dos valores medidos foram filtrados com relação a fase anterior de filtragem, reduzindo de 42.998 para 38.688 registros (4.310 valores ignorados).

Após a realização das etapas de filtragem, os dados selecionados que apresentaram resposta positiva as condições impostas pelos filtros, foram armazenados em uma outra tabela da plataforma MySQL. Enquanto que os dados que o

arquivo original continha, ainda se encontram disponíveis na tabela para onde foram importados na plataforma MySQL. Essa medida foi planejada para que caso outras investigações e/ou procedimentos sejam realizados, os dados originais possam estar disponíveis.

O algoritmo na primeira fase de filtragem selecionou 43,63% (44.494 valores) do total de dados do arquivo original. Já na segunda fase de filtragem a seleção reduziu para 42,16% (42.998 valores) do total, com relação aos dados originais, e a última fase de filtragem selecionou 37,93% (38.688 valores) dos dados totais, também com relação ao arquivo original. A primeira fase de filtragem é a que apresenta maior rigor na filtragem das irradiâncias, uma vez que, é responsável por eliminar todos os valores negativos e delimitar os valores medidos de acordo com a hora do nascer e ocaso do sol. Isto significa que aproximadamente 1/3 dos dados medidos foram selecionados para compor o conjunto de dados filtrados, permitindo que este conjunto possa ser utilizado como fonte de investigações científicas em diversos segmentos de pesquisa.

Com relação ao conjunto de dados que foram utilizados neste trabalho, pôde-se observar que o mês de Junho foi o que mais obteve os dados filtrados no processo de controle de qualidade, uma vez que, comparado aos demais meses, foi o que apresentou maior redução de valores medidos, principalmente se levada em consideração a segunda fase de filtragem, referente as frações das irradiâncias, onde a quantidade de dados ignorados na seleção foi de aproximadamente 40% dos dados que haviam sido selecionados na primeira fase de filtragem.

A partir dos resultados descritos foi iniciado um processo de investigação dos dados para obtenção da resposta do motivo pelo qual o mês de Junho apresentou a maior redução de valores medidos por meio da terceira fase de filtragem. Desta forma foram criados gráficos com os dados completos antes da filtragem e depois da filtragem (Fig. 2).

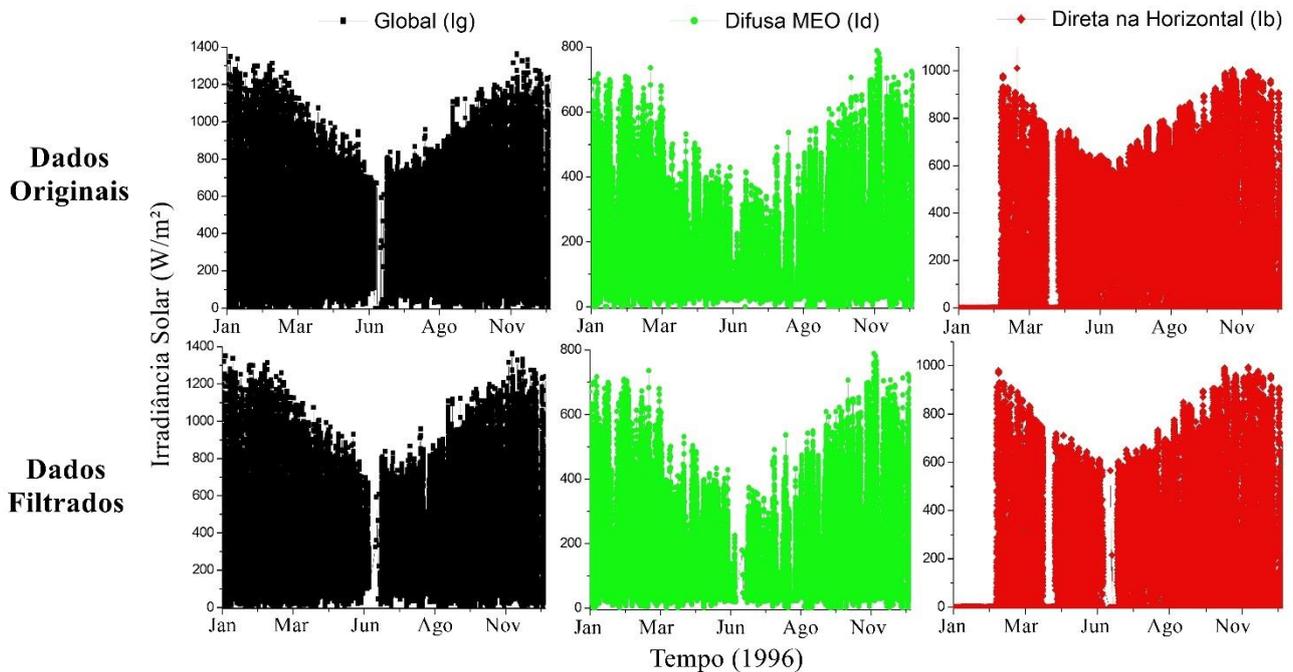


Figura 2 - Gráficos dos Dados Totais.

A plotagem de gráficos de séries temporais é de extrema importância, uma vez que, expõe ao pesquisador lacunas de valores e até mesmo possíveis erros de instrumentos ou de operação, nas atividades de monitoramento da radiação solar. Neste caso, os resultados expostos nos gráficos das irradiâncias (Fig. 2) mostram que antes do procedimento de filtragem a irradiância difusa apresentava valores espúrios em alguns dias. Já com relação a irradiância global e direta, antes da filtragem, pôde-se observar que existia uma lacuna de dados no mês de Junho, e nos primeiros meses do ano, respectivamente. A ausência dos valores medidos pode ter sido causada pela falta de operação do instrumento de medida. A análise dos gráficos mostrou também que diversos valores espúrios foram filtrados com relação aos dados originais, bem como, que o ano de 1996 apresenta diversos períodos de ausência de dados e que a lacuna de dados da irradiância global no mês de Junho resultou na não seleção dos valores, neste mesmo período, para as irradiâncias difusa e direta, por isso este mês apresentou maiores indícios de redução de seus valores na segunda e terceira fase de filtragem. Isto quer dizer que, devido ao fato de um conjunto de dados ter de possuir períodos iguais para todas as componentes solares, os algoritmos desenvolvidos foram programados para eliminar qualquer discrepância no conjunto, ou seja, no caso da global não apresentar dados em um período específico, este mesmo período será descartado da difusa e da direta, para que no final se tenham valores de períodos iguais para todas as radiações.

A ausência de valores, em muitos casos, pode ser um transtorno. Uma forma lógica de se contornar esta situação é realizar a soma dos valores medidos da irradiância difusa e direta (Eq. 6) para a global, uma vez que é compreendida por dois tipos de medidas, a difusa, que é resultado da ação de espalhamento da atmosfera, e a direta, que é a irradiância provinda diretamente do disco solar. Bem como, uma forma de contornar a ausência dos valores da irradiância direta e

difusa, é a partir do método da diferença. Que é uma técnica que requer que, ao menos, se tenha registro de duas radiações. No caso da irradiância direta, usa-se a diferença entre a global e a difusa (Eq. 7) e no caso da irradiância difusa, usa-se a diferença entre a global e a direta (Eq. 8).

$$I_g = I_d + I_b \quad (6)$$

$$I_b = I_g - I_d \quad (7)$$

$$I_d = I_g - I_b \quad (8)$$

É importante ressaltar que as análises e propostas de soluções apresentadas a respeito do preenchimento de valores das irradiâncias nos períodos de ausência podem ser realizadas para qualquer universo de dados, desde que, tais propostas sejam implementadas antes do procedimento de controle de qualidade, já que, o algoritmo desenvolvido pode ignorar tais dados se apresentarem resposta negativa as condições lógicas das comparações, principalmente, na terceira fase de filtragem.

5. CONCLUSÃO

De acordo com os testes físicos e lógicos realizados a partir dos filtros computacionais aplicados, a modelagem de controle de qualidade de dados de irradiâncias solares mostrou-se eficiente, pois foi capaz de realizar a filtragem dos valores medidos de acordo com o planejado baseando-se na literatura da área de especialização. Foi observado que a primeira fase de filtragem, referente a eliminação de valores negativos, e de delimitação a partir do nascer e ocaso do sol e do ângulo zenital menor que 80°, mostrou o maior corte nos dados, eliminando a maior quantidade de registros no procedimento de seleção. Constatou-se, também, que a segunda fase de filtragem apresentou maiores reduções de valores para o mês de Junho do ano de 1996, que também foi o mês que, ao final do procedimento, apresentou a menor quantidade de valores devido a um período de 12 dias em que a irradiância global não foi monitorada, onde o algoritmo ignorou esses registros. Propõem-se que ao replicar este estudo uma análise seja feita no conjunto de dados e que valores com lacunas sejam preenchidos a partir de métodos de estimativa antes de aplicar o procedimento de controle de qualidade.

A partir da averiguação dos resultados, levanta-se a proposta de aplicar este método para um conjunto ainda maior e mais completo de dados, verificando sua resposta de forma segmentada, por mês e por ano, e investigando o motivo de reduções fora da normalidade, conforme foi apresentado para o mês de Junho, nesta ocasião. Portanto, a modelagem computacional para controle de qualidade de dados de irradiâncias solares se mostrou efetiva, e pode ser aplicada para o processo de filtragem de dados espúrios em estações solarimétricas e meteorológicas de forma simplificada, uma vez que, realiza a seleção de valores medidos de uma forma simples a partir de um algoritmo, possibilitando que, pesquisadores com os mais diversificados níveis de conhecimento na área possam realizar o mesmo processo em suas pesquisas.

Agradecimentos

Os autores agradecem a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), a Faculdade de Botucatu UNIBR e ao Programa de Pós-Graduação em Energia na Agricultura da Faculdade de Ciências Agrônomicas - UNESP de Botucatu pelo apoio financeiro. Bem como ao Professor Dr. Enzo Dal Pai (Faculdade de Ciências Agrônomicas - UNESP de Botucatu) pelas instruções, e apoio, com relação aos filtros computacionais aplicados neste estudo.

REFERÊNCIAS

- Agência Nacional de Energia Elétrica. Atlas de Energia Elétrica do Brasil. ANEEL, Brasília, 2 ed., 2005.
- Dal Pai, A., *et al.* MEO shadowring method for measuring diffuse solar irradiance: Corrections based on sky cover. *Renewable Energy*, n. 99, p. 754-763, 2016.
- Il Sung-Zo., *et al.* Baseline Surface Radiation Network (BSRN) Quality Control of Solar Radiation Data on the Gangneung-Wonju National University Radiation Station. *Asia-Pacific Journal of Atmospheric Sciences*, n. 53, p. 11-19, 2017.
- Kudish, A. I., Evseev, E. G. The assessment of four different correction models applied to the diffuse radiation measured with a shadow ring using global and normal beam radiation measurements for Beer Sheva, Israel. *Solar Energy*, n. 82, p. 144-156, 2008.
- Muneer, T., *et al.* Discourses on solar radiation modeling. *Renewable & Sustainable Energy Reviews*, n. 11, p. 551-602, 2007.
- Moreno-Tejera, S., *et al.* A proposed methodology for quick assessment of timestamp and quality control results of solar radiation data. *Renewable Energy*, n.78, p. 531-537, 2015.
- Pashiards, S., Kalogirou, S. A. Quality control of solar shortwave and terrestrial longwave radiation for surface radiation measurements at two sites in Cyprus. *Renewable Energy*, n. 96, p. 1015-1033, 2016.
- Takai, O. K., *et al.* Apostila: Introdução à Banco de Dados. São Paulo: DCC-IME-USP. 2005. Acesso em: 17 out. 2017. Disponível em: < <https://www.ime.usp.br/~jef/apostila.pdf> >.

World Meteorological Organization. The Global Observing System for Climate: Implementation Needs. 2016.

COMPUTATIONAL MODELING OF QUALITY CONTROL OF SOLAR IRRADIANCE DATA

Abstract. *The aim of this study was to develop a computational model of quality control of data measured of global, diffuse and direct solar irradiance obtained by Botucatu Solar Radiometry Station. The correct data sources sizing allows government agencies, academic institutions and non-governmental organizations to plan and manage, more appropriately, renewable energy resources. In addition to help higher education and research with information about solar energy into thermal energy conversion studies, photovoltaic and biomass, in studies on the influence of sunlight on field crops. The initial dimensioning process of a data source is the quality control procedure, which in this case was developed based on a series of computational filters, where physical and logical tests were applied from the SQL query language. The implementation of the filters was done by the database management system - DBMS - MySQL. It can be observed that of the total raw data only 40% remained after the quality control, thus giving rise to a set of quality data. The computational modeling in final state provided methods that could be adapted to other database management systems or programming languages, according to with the need for research. The filters were constructed under algorithms that allowed the selection of measured values considered adequate.*

Key words: *Solar irradiance, Measured solar radiation, Data filtering.*