

XXIV SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS

DADOS GLOBAIS DE QUALIDADE DA ÁGUA, IMAGENS DE SATÉLITE E GOOGLE EARTH ENGINE PARA PRODUÇÃO DE MODELOS EMPÍRICOS DE QUALIDADE DA ÁGUA

Philippe Maillard ¹; Alexandre H. Dias Pereira ² & Arthur Lanna Neves ³

Resumo – Um trio de ferramentas de geotecnologias é apresentado como alternativa eficiente para a geração de modelos de regressão linear múltipla de parâmetros de qualidade de água. Aliando a base de dados globais GEMStat ao Google Earth Engine, um aplicativo de fonte aberta foi criado para gerar modelos de regressão linear múltipla – SIMOA. O uso do Earth Engine permitiu de buscar dados espectrais do satélite Sentinel-2 sem necessidade de armazenar as imagens. Com suas 12 bandas espectrais, o Sentinel-2 abre novas possibilidades no monitoramento remoto da água. O SIMOA é descrito no artigo, assim como sua relação com o GEMStat e o Earth Engine. Modelos de turbidez e de clorofila-a são apresentados e aplicados à represa de Três Marias a fim de ser comparados a modelos gerados a partir de dados in situ locais. Apesar de os modelos tenderem a exagerar os valores dos dois parâmetros, o SIMOA apresenta-se como uma ferramenta ágil e como plataforma de teste.

Palavras-Chave – Modelos empíricos de qualidade da água; Imagens Sentinel-2; Google Earth Engine

GLOBAL WATER QUALITY DATA, SATELLITE IMAGES AND GOOGLE EARTH ENGINE TO PRODUCE EMPIRICAL WATER QUALITY MODELS

Abstract – A trio of geotechnology tools is presented as an efficient alternative for the generation of multiple linear regression models of water quality parameters. Combining the GEMStat global database with Google Earth Engine, an open source application was created to generate multiple linear regression models – SIMOA. The use of Earth Engine made it possible to acquire spectral data from the Sentinel-2 satellite without the need to store the images. With its 12 spectral bands, Sentinel-2 opens up new possibilities in remote water monitoring. SIMOA is described as well as its relationship with GEMStat and Earth Engine. Turbidity and chlorophyll-a models are presented and applied to the Três Marias reservoir in order to be compared to models generated from local in situ data. Although the models tend to exaggerate levels of the two parameters, SIMOA presents itself as an agile tool and as a efficient test platform.

Keywords – Empirical water quality models, Sentinel-2 images Google Earth Engine.

1) Afiliação: Departamento de Geografia - IGC, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31.270-901, Brazil
+55 (31) 99600-8465, philippermaillard@yahoo.com.br

2) Afiliação: Departamento de Engenharia de Controle de Automação - ENG, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31.270-901, Brazil +55 (31) 98235-8181, alexandrehdias@gmail.com

3) Afiliação: Departamento de Geografia - IGC, Universidade Federal de Minas Gerais, Belo Horizonte, MG 31.270-901, Brazil
+55 (31) 99697-2869, arthurlannan245@gmail.com

INTRODUÇÃO

Os modelos empíricos de parâmetros de qualidade da água (QA) baseados na relação entre dados in situ e dados espectrais de imagens de satélite tem sido relativamente comuns nos últimos 20-30 anos (Liu et al., 2003). Entre os parâmetros que podem ser monitorados com sensoriamento remoto, a turbidez, a clorofila-a, a transparência e os sedimentos em suspensão são os mais comuns (Yigit-Avdan et al., 2019). O atual projeto sobre a QA da represa de Três Marias – MG está sendo desenvolvido em parte baseado em regressão linear múltipla (RLM) aplicada às bandas espectrais das imagens dos satélites Sentinel-2 e Landsat-8. A RLM é também aliada às técnicas de seleção de variáveis explicativas (bandas espectrais) que possuem o maior impacto nos modelos. Alguns desses modelos podem ser considerados semi-empíricos quando empregam variáveis pré-selecionadas na base do conhecimento de relações entre as características ópticas da água e a resposta espectral de certas faixas espectrais inseridas nas bandas das imagens. A inclusão de índices como o “Normalized Difference Turbidity Index” (NDTI) ou o “Maximum Chlorophyll Index” (MCI) também constituem uma forma de modelo semi-empíricos (Pinardi et al., 2015).

O elemento crítico para a construção desses modelos reside no acesso a dados medidos in situ para servir de treinamento e para dar validade estatística aos resultados. Em dois anos de projeto, duas dificuldades foram enfrentadas para poder produzir modelos robustos. A primeira vem da pequena amostra de dados in situ que conseguimos adquirir principalmente causada pela pandemia de Covid-19, mas também pela necessidade de sincronizar os levantamentos in situ com a passagem de satélite, que deve também coincidir com a quase ausência de nuvens. Três campanhas de campo produziram uma amostra de apenas 78 pontos in situ. Adicionalmente, a relativa homogeneidade dos parâmetros de QA da represa, que tem uma água de boa qualidade (o que pode dificultar a medição de certos parâmetros em vista dos limites de sensibilidade das sondas) é um fato que tende a enfraquecer os modelos de regressão múltipla.

É para remediar a esses problemas que resolvemos buscar dados de outros lagos e represas para ter uma maior quantidade de dados que permita uma melhor relação estatística e de investigar abordagens de aprendizagem de máquina como “Random Forest” (RF) e “Support Vector Machine” (SVM). Em segundo lugar, a inclusão de dados de outros lugares permite uma maior variabilidade de dados de QA (e consequentemente de resposta espectral) a fim de criar modelos mais genéricos (aplicáveis em outros contextos) e robustos.

MATERIAL E MÉTODO

A abordagem adotada usa uma série de soluções informatizadas disponíveis gratuitamente. O programa intitulado “SIMOA” (Sistema Inteligente de Monitoramento Operacional da Água) foi escrito em Python 3 e incorpora bibliotecas de gerenciamento de dados e de processamento estatístico (ver seção abaixo). Além do SIMOA, a estratégia é focada na utilização da base global de dados de QA GEMStat (<https://gemstat.org/>), uma iniciativa de um grupo alemão apoiada pelo programa ambiental das Nações Unidas e da plataforma de processamento geoespacial Google Earth Engine (GEE – <https://earthengine.google.com/>). É a combinação desses três elementos que permite uma maior agilidade na criação de modelos. Os três itens dessa colaboração são descritos abaixo.

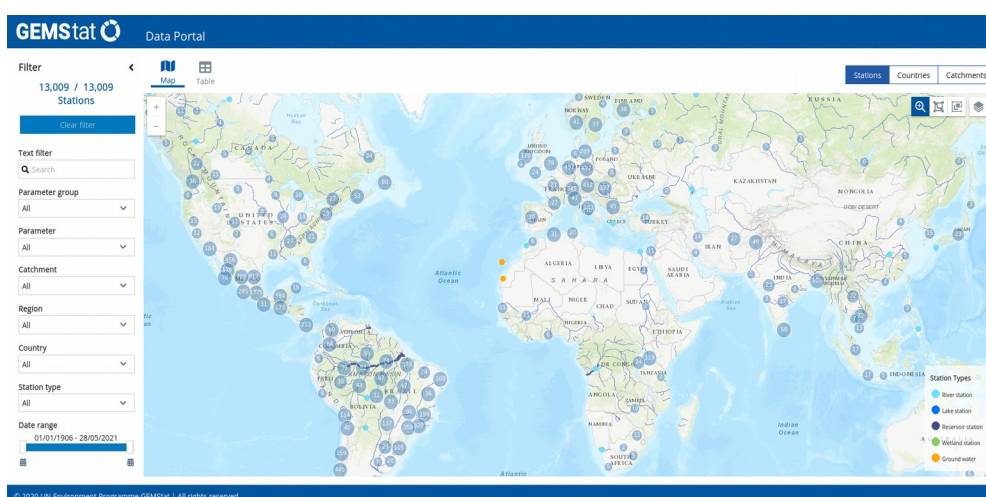
A base de dados GEMStat

A base de dados GEMStat (Global Freshwater Quality Database) é um repositório de dados globais de QA baseada na Alemanha e disponível gratuitamente (Färber, 2018). O texto a seguir é uma tradução.

“O Banco de Dados Global de Qualidade de Água Doce GEMStat fornece dados e informações cientificamente confiáveis sobre o estado e as tendências da QA continental global. Como parte operacional do GEMS / Programa de Água do Programa das Nações Unidas para o Meio Ambiente (UNEP), o GEMStat é hospedado pelo GEMS / Water Data Center (GWDC) dentro do Centro Internacional de Recursos Hídricos e Mudanças Globais (ICWRGC) em Koblenz, Alemanha.” <https://gemstat.org/>

Os dados do GEMStat são organizados e classificados de várias maneiras, o que permite uma seleção focada em tipos de corpos de água (rios, lagos, represas), regiões (países, continentes) e parâmetros de QA específicos (Figura 1). Assim, o recorte apenas em lagos e represas foi utilizado para atender aos estudos da represa de Três Marias. A plataforma permite a aquisição de dados diretamente através de uma seleção no mapa. Entretanto, a seleção de grandes quantidade de dados é feito por requerimento, via e-mail à equipe do GEMStat. Apesar de existirem outras plataformas de disponibilização de dados de QA (União Europeia, Estados Unidos, etc.), nem todas fornecem os dados in situ originais ou apresentam a mesma facilidade de seleção de parâmetros de região de interesse da plataforma GEMStat.

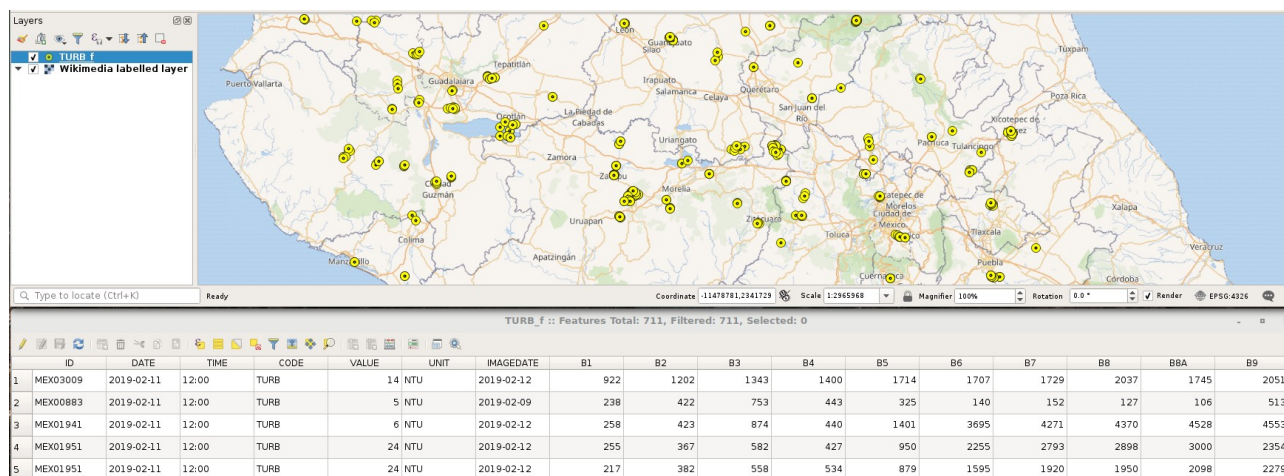
Figura 2. Aspecto gráfico do Portal do GEMStat usado para selecionar os dados a serem baixados.



Inicialmente apenas os parâmetros opticamente ativos foram escolhidas para serem analisados e processados pelo aplicativo SIMOA: temperatura (T°), clorofila (Chl-a), turbidez (TURB), transparência (Secchi), sólidos em suspensão (TSS) e carbono orgânico total (TOC).

Os dados do GEMStat são fornecidos em duas planilhas, uma com os dados de QA e a outra com as coordenadas geográficas associadas. Os códigos de ponto de coleta in situ estão compartilhados nas duas planilhas. Essas duas planilhas foram agrupadas em uma planilha única (.csv) e depois transformadas (usando um programa desenvolvido em Python) em arquivos do tipo *shapefile*, sendo um formato geo-codificado universalmente reconhecido. O formato *shapefile* permite que os dados sejam processados, tanto como planilhas (em formato .dbf) quanto como dados geo-localizados através de um Sistema de Informação Geográfica (SIG - Quantum GIS no caso do projeto). A Figura 2 ilustra no QGIS uma porção dos dados de turbidez localizados no México, assim como a tabela de atributos que contém os dados. A tabela é organizada por linhas, que contém as seguintes informações: identificador do ponto, data da coleta, hora, código do parâmetro, valor, unidade de medição e os dados espectrais do satélite adquiridos pela plataforma GEE.

Figura 2. Representação de dados de turbidex do GEMStat na região do México: (topo) na interface geográfica do QGIS e (em baixo) na interface de planilha que contem os dados.



A Plataforma de Dados e Processamentos Google Earth Engine (GEE)

Como os modelos são essencialmente baseados na resposta multi-espectral dos parâmetros, é necessário associar dados de bandas espectrais aos pontos de coleta, sejam aqueles adquiridos por nossa equipe, sejam de terceiros. O GEMStat oferece milhares de pontos de coleta, mas sem ter necessariamente uma relação com dados de satélite. Também, nem todos os pontos registram todos os parâmetros de QA necessários para nosso estudo. A diversidade de localizações e de datas de levantamento ocasionaria uma tarefa árdua e gigantesca de aquisição das imagens de satélites correspondentes, sem mencionar o espaço de disco que seria necessário. Para resolver esse impasse, o Google Earth Engine (GEE) veio como a solução natural mais apropriada. O GEE contém uma base de milhares de imagens de satélite (Landsat, Sentinel, MODIS, etc.), assim como muitas ferramentas de processamento. Um código em Python foi escrito que permite buscar dados de bandas espectrais a partir de uma coordenada e de uma data. Além disso, uma ferramenta do GEE permite calcular a média dos pixels num raio informado pelo usuário; uma precaução necessária para contornar os problemas de imprecisão de GNSS e de deriva dos barcos durante as coletas. Também foi permitido uma tolerância na correspondência entre as datas de coleta in situ e da imagem de satélite. Inicialmente, essa tolerância foi fixada a dois dias antes e depois da coleta in situ. Como a data dos dados da imagem é também registrada, uma filtragem dos dados baseada na diferença entre data de coleta e data de imagem, é sempre possível. O código Python abaixo mostra como essa busca é realizada.

```
buffer = 60 # raio de busca para o GEE retornar a média dos pixels da imagem
```

```
lst = ee.ImageCollection('COPERNICUS/S2_SR').filterDate(i_date, f_date).  
filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 20))
```

```
# escolha da coleção de imagens. i_date e f_date são as datas de início e fim da busca
```

```
# o parâmetro CLOUDY_PIXEL_PERCENTAGE permite excluir imagen com mais de 20% de  
nuvens
```

```
p_poi = ee.Geometry.Point(p_lon, p_lat) # onde p_lon e p_lat são listas das coordenadas desejadas
```

```
lst_p_poi = lst.getRegion(p_poi, buffer).getInfo() # retorna as informações espectrais em uma lista
```

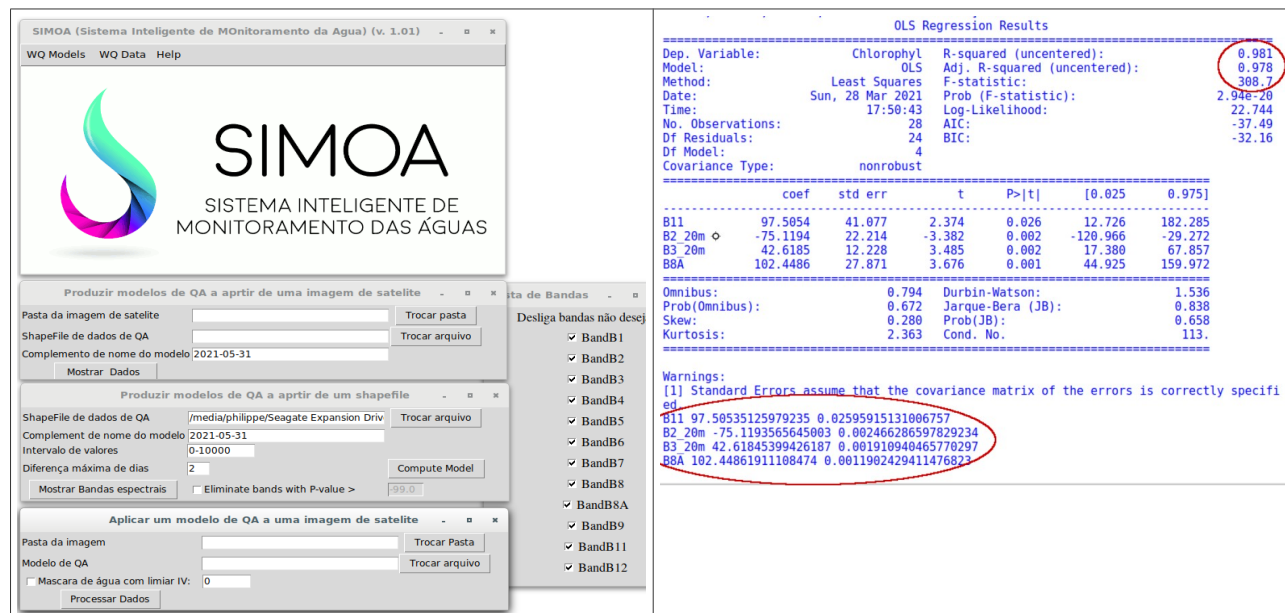
O resto do código utiliza a lista de valores espectrais para preencher a tabela de atributos do *shapefile*, que contém os dados de QA. A escolha, no código acima da coleção

COPERNICUS/S2_SR, significa que são dados de refletância da superfície (SR de *surface reflectance*) das imagens dos satélites Sentinel-2 (S2) do programa Europeu Copernicus a serem utilizadas. No caso, um limiar de 20% de nuvens foi aceito para incluir as imagens.

O Aplicativo SIMOA

O programa SIMOA foi desenvolvido em Python 3 como ferramenta de trabalho para gerar e testar modelos de regressão múltipla unindo dados in situ e dados de imagens de satélite (ou de drone no futuro). Inicialmente SIMOA aceita um *shapefile* com os dados de QA (sendo os atributos dos pontos geo-codificados) e os dados espectrais associados ou ainda uma imagem de satélite de nível 2 (correção atmosférica) do Sentinel-2 ou do Landsat-8 para produzir modelos de regressão. Um parâmetro de intervalos de valores (mínimo e máximo) e um de tolerância de diferença de dias (entre data da coleta e da imagem) permitem de testar como os modelos reagem às alterações desses dois parâmetros. O programa permite selecionar manualmente as bandas ou deixar um processo *stepwise* eliminar as bandas espectrais de baixo impacto no modelo, baseado na significância (P) da variável espectral. SIMOA tem sua própria interface gráfica constituída de uma janela principal e algumas janelas (filhas) dependentes (Figura 3). SIMOA utiliza os pacotes “Shapefile”, “Statsmodel” e “Pandas” para ler os dados e calcular e armazenar os modelos de regressão. Os resultados são apresentados na forma de um arquivo de modelo (pesos) e de um relatório estatístico (Figura 3) ambos armazenados. Um módulo permite aplicar qualquer modelo em uma imagem desde que ela seja do mesmo tipo e tenha o mesmo nível de correção (L2). Uma vez aplicado o resultado é armazenado em um arquivo do tipo GeoTiff além de ser ilustrado em uma figura.

Figura 3. Interface gráfica do SIMOA mostrando a janela principal (esquerda) assim como as três principais janelas filha e a janela de escolha de bandas e exemplo de relatório estatístico produzido pelo pacote Statsmodel no SIMOA (direita). Áreas cercadas mostram o coeficiente de determinação e os pesos calculados para as bandas retidas pela regressão (aqui de a clorofila-a).



RESULTADOS

O trio de ferramentas permite uma agilidade muito grande para a produção de modelos empíricos que devem ser posteriormente testados e validados. A ideia de “melhor modelo” é relativa. Por um lado pode ser o modelo que mais se aproxima da situação real de um local

específico (represa, lago, rio) ou pode ser um modelo que consegue ser mais robusto na maior quantidade de situações diferentes (exportabilidade do modelo). Para os fins deste artigo, modelos de clorofila-a e de turbidez são apresentados e discutidos.

Turbidez

A turbidez é, com a clorofila e a transparência, um dos parâmetros de QA mais estimado através do sensoriamento remoto (Yigit-Avdan et al., 2019). A turbidez é associada à presença de partículas inorgânicas e orgânicas que provocam o espalhamento e a absorção da luz. É considerado que o sensoriamento remoto pode ser usado para medir a turbidez superficial e subsuperficial em função da concentração dessas partículas. Apesar disso, medir a turbidez em ambientes complexos representa ainda um desafio (Robert et al. 2017). Vários autores utilizam a turbidez justamente para caracterizar os tipos de água como clara, turva e muito turva (Uudeberg et al., 2020).

Doze modelos de turbidez foram construídos com o SIMOA, seis utilizando todas as bandas espectrais do Sentinel-2 (Tabela 1) e seis usando o processo *stepwise* de eliminação de bandas. Dois níveis máximos de turbidez foram criados: ≤ 500 NTU e ≤ 100 NTU (essa subdivisão foi criada em função das necessidades do projeto da represa Três Marias que apresenta níveis baixos de turbidez). Finalmente, foram estabelecidos três níveis de tolerância em função da diferença de data entre a coleta e da aquisição da imagem: 0, 1 e 2 dias (Tabela 2).

Tabela 1. Descrição das 13 bandas espectrais do Sentinel-2. Legenda: NIR (*near infrared*) = infravermelho próximo; RE (*red edge*) = limite vermelho-infravermelho próximo.

Banda	1	2	3	4	5	6	7	8	8a	9	10	11	12
Comprimento de onda (nm)	443	490	560	665	705	740	783	842	865	945	1375	1610	2190
Largura da banda (nm)	20	128	128	108	15	15	20	115	20	20	20	90	180
Nomenclatura	Costal	Azul	Verde	Verm.	RE	RE	RE	NIR	NIR	Vapor	Cirrus	SWIR	SWIR
Resolução espacial (m)	60	10	10	10	20	20	20	10	20	60	60	20	20

A Tabela 2 mostra o resumo dos resultados obtidos pelas 12 regressões (R^2 e significância) assim com a relação das bandas espectrais que mais se encontram nos modelos (última linha da tabela). Aparece claro pelos valores de coeficiente de determinação (R^2) que o fato de limitar o valor máximo de turbidez afeta menos o resultado que a tolerância de dias de diferença entre imagem e coleta. Entretanto, a redução da tolerância afeta muito o número de observações (N) que tem um efeito na performance do modelo. O processo *stepwise* tem a vantagem de reduzir o número de variáveis e simplificar o modelo, mas tende a reduzir também o coeficiente de determinação. Por outro lado ele informa sobre a importância de cada variável, especialmente no contexto de vários modelos. No caso, são especialmente as bandas 4 (vermelho) e 6 (*Red Edge*) que são privilegiadas (6 em 6). Em segundo lugar, as bandas 3 (verde) e 8 (IVP) tem uma recorrência de 4 em 6). Essa observação corrobora o que vários autores apontaram, especialmente sobre a banda do vermelho. No caso da turbidez, a diferença de data entre a coleta e a imagem não parece ter uma importância tão grande, pois em 3 dos 4 casos (≤ 500 e ≤ 100 com e sem *stepwise*) os modelos tiveram performance melhor com a maior tolerância.

Tabela 2. Resultado das regressões lineares múltiplas entre turbidez e dados espectrais (Sentinel-2) realizada com o SIMOA usando os dados da base de dados GEMStat.

Intervalo de valores	Dias de diferença	N	R ² ajustado	Signif. a 95%	Bandas Sentinel-2
≤ 500	2	709	0.618	Sim	Todas (12)
≤ 500	1	540	0.623	Sim	Todas (12)
≤ 500	0	182	0.561	Sim	Todas (12)
≤ 100	2	686	0.633	Sim	Todas (12)
≤ 100	1	509	0.619	Sim	Todas (12)
≤ 100	0	180	0.745	Sim	Todas (12)
≤ 500	2	709	0.578	Sim	Stepwise (3,4,6,8,11)
≤ 500	1	540	0.573	Sim	Stepwise (3,4,6,8,11)
≤ 500	0	182	0.377	Sim	Stepwise (4,6)
≤ 100	2	686	0.602	Sim	Stepwise (1,3,4,6,7,8,8a,11)
≤ 100	1	509	0.574	Sim	Stepwise (1,3,4,6,7,8,8a,9,11)
≤ 100	0	180	0.473	Sim	Stepwise (1,4,6)

Recorrencia das bandas: B4 → 6; B6 → 6; B3 → 4; B8 → 4; B11 → 4

Clorofila-a

A Tabela 3 mostra os resultados da performance dos modelos de regressões lineares múltiplas para os dados de clorofila-a. Neste caso, além de manter a subdivisão “sem seleção” e “com seleção” (*stepwise*) e pela tolerância de dias, três critérios de intervalos de valores foram usados: ≤ 1000 ug/l, ≤ 500 ug/l e ≤ 100 ug/l a fim de refletir melhor a situação da represa de Três Marias que apresenta valor relativamente baixo de clorofila. Os coeficientes de determinação são notavelmente mais baixo que para a turbidez. A clorofila é associada aos organismos clorofilianos, que para a fotossíntese absorvem partes específicas da luz (azul e vermelho) e costumem refletir mais no infravermelho próximo (IVP). Como a água tem um poder de absorção muito forte no IVP, essa porção é mais dificilmente observável no meio aquático.

A seleção *stepwise* de bandas tem um efeito forte nos resultados que aparecem melhores com todas as bandas. Entretanto isso não significa que os modelos com todas as bandas sejam mais robustos. Neste sentido, o processo *stepwise* informa sobre as bandas com maior impacto nos modelos. São as bandas 6 e 7, ambas da região do limite Vermelho-IVP (*red edge*) que são as mais frequentes seguidas das bandas 8a e 9 do IVP que indicam que essa relação entre clorofila e maior reflexão no IVP foi detectada pelos modelos. Finalmente a presença da banda 4 (vermelho) deve ser reflexo da maior absorção desta faixa pela clorofila.

A diferença de dias entre a coleta e a data da imagem parece ter um efeito mais importante para a clorofila que para a turbidez. Em todos os casos, os dados sem diferença de data tiveram um coeficiente de determinação maior que as tolerâncias de 1 e 2 dias.

Tabela 3. Resultado das regressões lineares múltiplas entre clorofila-a e dados espectrais (Sentinel-2) realizada com o SIMOA usando os dados da base de dados GEMStat.

Intervalo de valores	Dias de diferença	N	R ² ajustado	F (* signif.)	Bandas Sentinel-2
≤ 1000	2	615	0.407	Sim	Todas (12)
≤ 1000	1	456	0.428	Sim	Todas (12)
≤ 1000	0	149	0.547	Sim	Todas (12)
≤ 500	2	596	0.476	Sim	Todas (12)
≤ 500	1	440	0.472	Sim	Todas (12)
≤ 500	0	143	0.639	Sim	Todas (12)
≤ 100	2	455	0.435	Sim	Todas (12)
≤ 100	1	329	0.419	Sim	Todas (12)
≤ 100	0	114	0.633	Sim	Todas (12)
≤ 1000	2	615	0.277	Sim	Stepwise (4,6,7,8,8a,9)
≤ 1000	1	456	0.278	Sim	Stepwise (4,6,7,8,8a,9)
≤ 1000	0	149	0.348	Sim	Stepwise (4,6,7,9,12)
≤ 500	2	596	0.339	Sim	Stepwise (4,6,7,8,8a,9)
≤ 500	1	440	0.317	Sim	Stepwise (6,7,8,8a,9)
≤ 500	0	143	0.465	Sim	Stepwise (6,7,9,11)
≤ 100	2	455	0.308	Sim	Stepwise (5,6,7,8a,9)
≤ 100	1	329	0.326	Sim	Stepwise (5,6,7,8a,9)
≤ 100	0	114	0.397	Sim	Stepwise (2,6,7,9)

Recorrencia das bandas: B6 → 9; B7 → 9; B9 → 7; B8a → 6; B4 → 4; B8 → 4

Comparação de modelos

Na intenção de avaliar os modelos gerados a partir dos dados GEMStat (de diversos países, datas e intervalos de valores) eles foram aplicados na represa de Três Marias para o dia 31 de maio 2020. Dados in situ (sonda Exo) do dia 31 de maio também foram usados para gerar modelos locais a fim de ser comparados com os modelos GEMStat. Tantos os valores de turbidez como de clorofila-a da represa naquele dia foram considerados baixos. A turbidez não ultrapassou 6.27 NTU e a clorofila-a 1.0 µg/l, valores que caracterizam água limpa. Por outro lado, a grande maioria dos dados do GEMStat apresentam valores maiores nos dois parâmetros. Deve-se considerar que na maioria das vezes, corpos hídricos são monitorados justamente quando os parâmetros de qualidade excedam potencialmente os níveis aceitáveis.

A Figura 4 mostra o modelo de turbidez usando valores até 100 NTU com tolerância de até dois dias de diferença e um modelo gerado exclusivamente com os dados de campo coletados no dia 31 de maio 2020 (R² de 0.94). Fica notável que o modelo GEMStat tende a exagerar a turbidez daquele dia por um fator de até 2 ou 3. Entretanto, mesmo com esse exagero, a turbidez da represa é relativamente baixa.

O mesmo exercício foi feito com a clorofila-a e é representado na Figura 5. Neste caso o resultado ficou pior que para a turbidez. O modelo GEMStat exagera o nível de clorofila-a por um fator de até 5 (a escala de cor da figura teve de ser ajustada para poder representar melhor as diferenças). Foi observado, que quanto mais baixa a concentração de clorofila, pior o resultado do modelo GEMStat.

CONCLUSÃO

É um dos objetivos do projeto de determinar como modelos mais robustos e replicáveis podem ser obtidos a partir de uma grande base de dados diversificada. O projeto SIMOA está ainda em fase de desenvolvimento e outros dados estão sendo procurados para compor nossa base global de dados de QA. Os primeiros resultados mostram que a tarefa de construir modelos robustos e genéricos é complexa e requer um maior entendimento sobre a relação entre parâmetros de QA e entre esses parâmetros e as propriedades ópticas das águas de superfície. O trio GEMStat-Google Earth Engine-SIMOA permite uma grande agilidade em produzir e testar modelos empíricos de QA. Os modelos de turbidez e clorofila demonstraram que a dinâmica dos valores não permitem modelos lineares de serem exportáveis facilmente; uma conclusão esperada em vista de muitas outras publicações de sensoriamento remoto aplicado à QA. Modelos não lineares e não paramétricos vão ser implementados no SIMOA como próxima etapa do projeto.

AGRADECIMENTOS

Os autores gostariam de agradecer a CEMIG pelo apoio institucional e financeiro desse projeto e o grupo do GEMStat pela disposição em disponibilizar os dados utilizados neste trabalho. Enfim somos reconhecetes a toda equipe do projeto P&D 607 que permitiu desenvolver o SIMOA e adquirir os dados de campo.

REFERÊNCIAS

- FÄRBER, C., LISNIAK, D., SAILE, P., KLEBER, S.H., EHL, M., DIETRICH, S., FADER, M. AND DEMUTH, S. (2018). “*Water quality at the global scale: GEMStat database and information system*” in EGU General Assembly Conference Abstracts, p. 15984.
- LIU, Y., ISLAM, M. A., & GAO, J. (2003). “*Quantification of shallow water quality parameters by means of remote sensing*”. Progress in Physical Geography, 27(1), pp. 24-43.
- PINARDI, M., FENOCCHI, A., GIARDINO, C., SIBILLA, S., BARTOLI, M., BRESCIANI, M. (2015). “*Assessing potential algal blooms in a shallow fluvial lake by combining hydrodynamic modelling and remote-sensed images*”. Water, 7(5), pp. 1921-1942.
- QUININO, U.C.; CAMPOS, L.F.; GADELHA, C.L. (2000). “*Avaliação da qualidade das águas subterrâneas na bacia do rio Gramame no Estado da Paraíba*” in Anais do V Simpósio de Recursos Hídricos do Nordeste, Natal, Nov. 2000, 1, pp. 162-176.
- ROBERT, E., KERGOAT, L., SOUMAGUEL, N., MERLET, S., MARTINEZ, J. M., DIAWARA, M., & GRIPPA, M. (2017). “*Analysis of suspended particulate matter and its drivers in Sahelian ponds and lakes by remote sensing (Landsat and MODIS): Gourma region, Mali*”. Remote Sensing, 9(12), p.1272.
- UUDEBERG, K.; AAVASTE, A.; KÖKS, K.L.; ANSPER, A.; UUSÕUE, M.; KANGRO, K.; ANSKO, I.; LIGI, M.; TOMING, K.; REINART, A. (2020). “*Optical Water Type Guided Approach to Estimate Optical Water Quality Parameters*”. Remote Sensing, 12(6), 931.
- YIGIT AVDAN, Z., KAPLAN, G., GONCU, S., AVDAN, U. (2019). “*Monitoring the water quality of small water bodies using high-resolution remote sensing data*”. ISPRS International Journal of Geo-Information, 8(12), p. 553.

Figura 4. Comparação entre o modelo de turbidez gerado a partir dos dados do GEMStat e do modelo de turbidez gerado a partir dos dados in situ, ambos para o dia 31 de maio 2020.

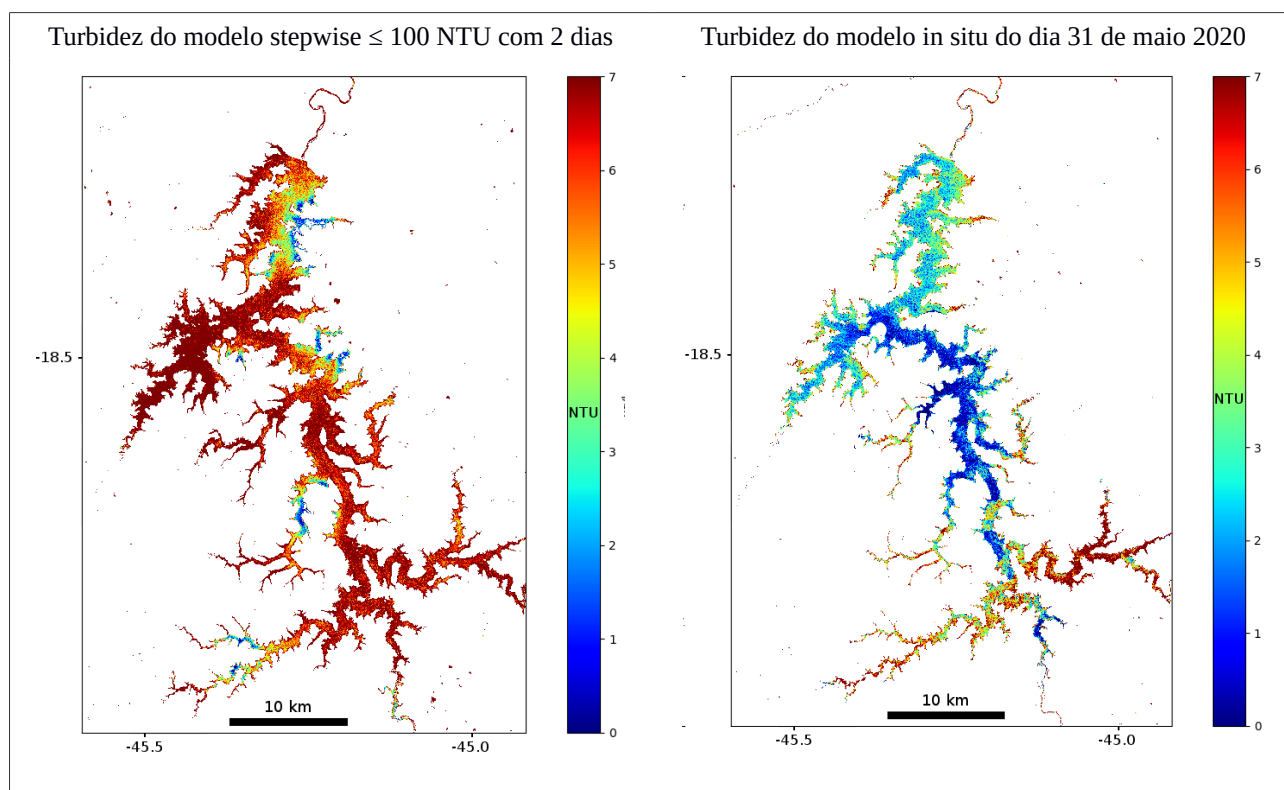


Figura 5. Comparação entre o modelo de clorofila-a gerado a partir dos dados do GEMStat e do modelo de clorofila-a gerado a partir dos dados in situ, ambos para o dia 31 de maio 2020.

