

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

**Avaliação das mudanças no armazenamento de água subterrânea na
bacia do rio Amazonas a partir do downscaling de dados
GRACE/GRACE-FO com modelos Machine Learning**

DIEGO ALEJANDRO SATIZÁBAL ALARCÓN

Dissertação apresentada ao Programa de Pós-
Graduação em Recursos Minerais de
Hidrogeologia para obtenção do título de
Mestrado

Área de concentração: Hidrogeologia e Meio
Ambiente

Orientadora: Profa. Dra. Alexandra Vieira
Suhogusoff

São Paulo
2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Serviço de Biblioteca e Documentação do IGc/USP

Ficha catalográfica gerada automaticamente com dados fornecidos pelo(a) autor(a)
via programa desenvolvido pela Seção Técnica de Informática do ICMC/USP

Bibliotecários responsáveis pela estrutura de catalogação da publicação:

Sonia Regina Yole Guerra - CRB-8/4208 | Anderson de Santana - CRB-8/6658

Satizábal-Alarcón, Diego Avaliação das mudanças no armazenamento de água subterrânea na bacia do rio Amazonas a partir do downscaling de dados GRACE/GRACE-FO com modelos Machine Learning / Diego Satizábal-Alarcón; orientador Alexandra Suhogusoff. -- São Paulo, 2023. 159 p.

Tese (Doutorado - Programa de Pós-Graduação em Recursos Minerais e Hidrogeologia) -- Instituto de Geociências, Universidade de São Paulo, 2023.

1. Mudanças no armazenamento de água subterrânea na bacia do rio Amazonas. 2. Downscaling de dados GRACE/GRACE-FO com modelos de Machine learning. I. Suhogusoff, Alexandra, orient. II. Título.

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

**Avaliação das mudanças no armazenamento de água subterrânea na
bacia do rio Amazonas a partir do downscaling de dados
GRACE/GRACE-FO com modelos Machine Learning**

DIEGO ALEJANDRO SATIZÁBAL ALARCÓN

Orientadora: Profa. Dra. Alexandra Vieira Suhogusoff

Dissertação de Mestrado

Nº 919

COMISSÃO JULGADORA

Dr^a. Alexandra Vieira Suhogusoff

Dr. Ingo Daniel Wahnfried

Dr^a. Alice César Fassoni de Andrade

SÃO PAULO

2023

*À minha família, amigos, professores
e mestres que me guiaram até aqui.*

AGRADECIMENTOS

Após pouco mais de três anos, incluindo uma pandemia, expresso minha sincera gratidão à minha família, orientadores e amigos pelo constante apoio e companhia, pois sem eles, o desenvolvimento deste projeto não teria sido possível. Aos meus pais, Cristina e Ricardo, dedico um agradecimento especial pelo carinho, apoio, cuidado e valiosos conselhos oferecidos ao longo de todo esse processo. Foram eles que me forneceram a coragem necessária nos momentos mais desafiadores. À minha irmã, Adriana, por sua disposição em me ouvir e manter meus pés no chão. À Marian, por ser um apoio e uma confidente. Ao Andrés, agradeço sinceramente por seus conselhos e amizade incondicional, foram de grande valor. Aos amigos e colegas colombianos que compartilharam comigo a experiência no Brasil, quero expressar minha gratidão pelas conversas e momentos de descontração que proporcionaram alívio em momentos desafiadores. Um agradecimento especial a Sebas H e Verónica, cuja cumplicidade, conversas profundas e conselhos foram fundamentais para o meu progresso.

Este projeto, que representa um sonho e uma meta acadêmica, não teria sido possível sem a confiança e paciência da minha orientadora, a Professora Alexandra V. Suhogusoff. Desde o início de nossa colaboração, seu apoio, confiança, motivação e supervisão foram pilares essenciais para meu sucesso acadêmico. Gostaria também de expressar minha gratidão ao Professor Luiz Ferrari por suas valiosas observações, conselhos e contribuições que enriqueceram este trabalho.

Por fim, quero expressar minha profunda gratidão ao ensino público e, em particular, ao Instituto de Geociências da Universidade de São Paulo e ao Centro de Pesquisas de Águas Subterrâneas (CEPAS) por proporcionarem a mim e a outros alunos acesso a conhecimentos e pesquisas de excelente qualidade acadêmica. Agradeço aos professores que, com rigor e dedicação, compartilharam seu conhecimento.

É importante mencionar que este projeto foi financiado pela bolsa da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) com o processo e número de projeto FAPESP no 2014/00844-7.

“Não há ensino sem pesquisa e pesquisa sem ensino”

Paulo Freire

RESUMO

Satizábal-Alarcón, D.A., 2023. **Avaliação das mudanças no armazenamento de água subterrânea na bacia do rio Amazonas a partir do *downscaling* de dados GRACE/GRACE-FO com modelos Machine Learning**. [Dissertação de Mestrado], São Paulo, Instituto de Geociências, Universidade de São Paulo, 155p.

As alterações no armazenamento de águas subterrâneas (ΔGWS) na bacia do rio Amazonas (ARB) desempenham um papel importante no comportamento hidrológico da região, com uma influência significativa sobre os ecossistemas da floresta tropical, bem como no abastecimento de água potável para as comunidades locais. As missões satelitais GRACE e GRACE-FO fornecem dados gravimétricos a partir dos quais é possível monitorar as mudanças no armazenamento de água terrestre (ΔTWS), embora com baixa resolução espacial (1° ou 3°). Este estudo realizou uma redução de escala dos dados do GRACE/GRACE-FO a partir de modelos de aprendizado de máquina de 1° (aproximadamente 110 km) para 0.25° (aproximadamente 27.5 km). Paralelamente, foram comparados e analisados os algoritmos Random Forest e AdaBoost, obtendo magnitudes de erro menores e maior precisão nas previsões do AdaBoost. Foi fornecida uma estimativa da variabilidade espaço-temporal das anomalias de armazenamento de água terrestre e subterrânea a 0.25° entre 2002 e 2021 para a bacia do rio Amazonas. Foram identificados padrões espaciais e temporais associados aos efeitos do clima e das atividades antrópicas, para a bacia em geral, com ênfase no Sistema Aquífero Amazonas (SAA) e em alguns aquíferos menores (Parecis, Ronuro e Boa Vista). As séries temporais mostraram uma alta susceptibilidade aos fenômenos El Niño e La Niña nas ΔGWS dos aquíferos confinados ou livres. Além disso, a expansão da fronteira agrícola, o desmatamento e queimadas prolongadas ao longo do tempo têm levado à redução das ΔGWS com tendências lineares de aproximadamente 7mm/ano em áreas de 0.25° nos aquíferos de Parecis e Ronuro. O comportamento regional da ARB apresenta uma tendência positiva de longo prazo no armazenamento terrestre e subterrâneo de $14.26 \pm 1.18 \text{ km}^3/\text{ano}$ e $+22.24 \pm 1.18 \text{ km}^3/\text{ano}$, respectivamente. A validação entre as séries temporais das ΔGWS e os níveis de água em poços de monitoramento obteve coeficientes de correlação máximos de 0,85 com níveis de confiança de 0.01. Esses resultados são promissores para o uso de informações satelitais na gestão da água, especialmente no monitoramento regional de aquíferos não confinados.

Palavras-chave: Sensoriamento remoto; AdaBoost; Random Forest; Inteligência Artificial; Séries temporais; Sistema Aquífero Amazônico; Impacto antropogênico

ABSTRACT

Satizábal-Alarcón, D.A., 2023. **Assessment of changes in groundwater storage in the Amazon River basin through downscaling of GRACE/GRACE-FO data with Machine Learning models.** [Master's thesis], São Paulo, Institute of Geosciences, University of São Paulo, 155p.

Changes in groundwater storage (Δ GWS) in the Amazon River Basin (ARB) play an important role in the hydrological behavior of the region, with a significant influence on rainforest ecosystems as well as on the supply of drinking water to local communities. GRACE and GRACE-FO satellite missions provide gravimetric anomalies from which it is possible to monitor changes in terrestrial water storage (Δ TWS), albeit at a low spatial resolution (1° or 3°). This study downscaled GRACE/GRACE-FO data from 1° machine learning models (approx. 110 km) to 0.25° (approx. 27.5 km). Concurrently, the Random Forest and AdaBoost algorithms were compared and analyzed, with AdaBoost demonstrating lower error magnitudes and superior prediction accuracy. An estimation of the space-time variability of terrestrial and groundwater storage anomalies at 0.25° between 2002 and 2021 for the Amazon River basin is provided. Spatial and temporal patterns associated with climate effects and anthropogenic activities were identified, focusing on the Amazon Aquifer System (SAA) and some smaller aquifers (Parecis, Ronuro, and Boa Vista). Time series data show a high susceptibility to El Niño and La Niña phenomena in the Δ GWS of unconfined or free aquifers. Additionally, the expansion of the agricultural frontier, deforestation, and prolonged burning have led to a decline in Δ GWS, with linear trends of approximately 7mm/year in the Parecis and Ronuro aquifers. The regional behavior of the ARB presents a long-term positive trend in terrestrial and groundwater storage of $14.26 \pm 1.18 \text{ km}^3/\text{year}$ and $+22.24 \pm 1.18 \text{ km}^3/\text{year}$, respectively. Validation between Δ GWS time series and water levels in monitoring wells achieved maximum correlation coefficients of 0.85 with confidence levels of 0.01. These results hold promise for the use of satellite information in water management, particularly in the regional monitoring of unconfined aquifers.

Keywords: Remote sensing; AdaBoost; Random Forest; Artificial intelligence; Time series; Amazon Aquifer System; Anthropogenic impact.

RESUMEN

Satizábal-Alarcón, D.A., 2023. **Evaluación de los cambios en el almacenamiento de agua subterránea en la cuenca del río Amazonas a través de la reducción de escala de datos GRACE/GRACE-FO con modelos de Aprendizaje Automático.** [Tesis de Maestría], São Paulo, Instituto de Geociencias, Universidad de São Paulo, 155p.

Los cambios en el almacenamiento de aguas subterráneas (ΔGWS) en la cuenca del río Amazonas (ARB) desempeñan un papel importante en el comportamiento hidrológico de la región, con una influencia significativa sobre los ecosistemas de la selva tropical, así como en el abastecimiento de agua potable para las comunidades locales. Las misiones satelitales GRACE y GRACE-FO proporcionan datos gravimétricos a partir de los cuales es posible monitorizar los cambios en el almacenamiento de agua terrestre (ΔTWS), aunque a baja resolución espacial (1° o 3°). Este estudio realizó una reducción de escala de los datos de GRACE/GRACE-FO a partir de modelos de *machine learning* de 1° (110 km aprox.) a $0,25^\circ$ (27,5 km aprox.). Paralelamente, se compararon y analizaron los algoritmos *Random Forest* y *AdaBoost*, obteniendo magnitudes de error menores y una precisión superior en las predicciones de *AdaBoost*. Se proporciona una estimación de la variabilidad espacio-temporal de las anomalías de almacenamiento de agua terrestre y subterránea a $0,25^\circ$ entre 2002 y 2021 para la cuenca del río Amazonas. Se identificaron patrones espaciales y temporales asociados a los efectos del clima y las actividades antrópicas, para la cuenca en general con énfasis en el Sistema Acuífero Amazonas (SAA) y en algunos acuíferos de menor tamaño (Parecis, Ronuro y Boa Vista). Las series temporales muestran una alta susceptibilidad a los fenómenos del Niño y la Niña en lo ΔGWS de los acuíferos inconfiados o libres. Así mismo, la expansión de la frontera agrícola, la deforestación y quemadas prolongadas en el tiempo, han generado el descenso de los ΔGWS con tendencias lineales para áreas de $0,25^\circ$ del orden de 7mm/año en los acuíferos de Parecis y Ronuro. El comportamiento regional de la ARB presenta una tendencia positiva a largo plazo en el almacenamiento terrestre y subterráneo de $14,26 \pm 1,18 \text{ km}^3/\text{año}$ y $+22,24 \pm 1,18 \text{ km}^3/\text{año}$, respectivamente. La validación entre las series temporales de los ΔGWS con los niveles de agua en los pozos de monitoreo obtuvo coeficientes de correlación máximos de 0.85 con niveles de confianza de 0.01. Estos resultados son prometedores para el uso de información satelital en la gestión del agua, especialmente en el monitoreo regional de acuíferos no confinados.

Palabras clave: Sensores remotos; AdaBoost; Random Forest; Inteligencia artificial; Series temporales; Sistema Acuífero Amazónico; Impacto antropogénico.

Sumário

RESUMO	i
ABSTRACT	ii
Sumário	iv
Lista de Figuras	vi
Lista de Tabelas	ix
Lista de Siglas	ix
1. INTRODUÇÃO	1
1.1. Objetivos	4
1.1.1. Objetivo geral	4
1.1.2. Objetivos específicos	4
1.2. Área de Estudo	4
2. REVISÃO DE LITERATURA	9
2.1. Geologia	9
2.2. Sistema Aquífero Amazonas (AAS)	15
2.2.1. Aquífero Alter do Chão	16
2.2.2. Aquífero Içá	17
2.2.3. Acuitardo Solimoes-Repouso	18
2.2.4. Aquífero Aluvial	18
2.3. Sensoriamento remoto em Hidrogeologia	19
2.3.1. Medições gravimétricas da água subterrânea	19
2.3.2. Medições da deformação da superfície	21
2.3.3. Sistemas eletromagnéticos aéreos (AEM)	23
2.3.4. Umidade do solo e evapotranspiração	24
2.4. Downscaling GRACE/GRACE-FO	24
2.4.1. Random Forest	26
2.4.2. Adaptative Boosting	26
3. MATERIAIS E MÉTODOS	27
3.1. Materiais	27
3.1.1. Sensoriamento remoto	27
3.1.2. Medições In Situ	29
3.2. Métodos	30
3.3. Compilação de informação	31
3.3.1. Aquisição de dados	32
3.3.2. Informações complementares	32
3.4. Pré-processamento	33
3.5. Geoprocessamento	34
3.5.1. Downscaling	34
3.5.2. Estimativa das mudanças no armazenamento de água subterrânea (Δ GWS)	35
3.6. Validação da informação	36
3.6.1. Comparação de informações e análise de erro	36
3.6.1. Teste de significância	37

3.7. Análise e processamento dos resultados	37
3.7.1. Decomposição das séries temporais.....	37
3.7.2. Análise de Componentes Principais (PCA).....	38
3.7.3. Índice de umidade topográfica ou Topographic Wetness Index (TWI).....	39
3.7.4. Zoneamento das variações no Δ GWS.....	40
3.8. Programas Utilizados	41
4. RESULTADOS	43
4.1. ARTIGO CIENTÍFICO: Characterization of groundwater storage changes in the Amazon River Basin based on downscaling of GRACE/GRACE-FO data with machine learning models	43
4.2. Efeitos do clima e das atividades antropogênicas nos ΔGWS da ARB	86
4.2.1. Decomposição das séries temporais.....	86
4.2.2. Variabilidade hidrológica.....	93
4.2.3. Relação das componentes sazonal e total entre Δ GWS e outras variáveis.....	101
4.2.4. TWI.....	106
4.2.5. Atividade antrópica.....	109
4.2.6. Correlações espaço-temporais e com Δ GWS.....	111
4.2.7. PCA Pearson.....	115
4.2.8. Zoneamento de Δ GWS.....	124
5. CONCLUSÕES	127
REFERÊNCIAS BIBLIOGRÁFICAS	129
ANEXO A	144

Lista de Figuras

- Figura 1. Área de estudo onde A) Topografia da ARB, B) Bacias sedimentares e localização dos poços do RIMAS, C) Sub-bacias hidrográficas, D) Subdivisão da ARB, e E) toda a AAS. 5
- Figura 2. Área de estudo com a) localização dos poços de monitoramento e b) principais unidades chronostratigráficas extraídas do mapa geológico da América do Sul. 7
- Figura 3. Seção geológica esquemática do Arcode Gurupá. Modificado de Wanderley Filho & Costa (1991). 9
- Figura 4. Seções longitudinais e transversais da Bacia do Amazonas. Tomado de Caputo (1984). 10
- Figura 5. Seções longitudinais e transversais da Bacia do Solimões. Tomado de Caputo (1991). 11
- Figura 6. Seção geológica esquemática da Bacia de Acre. Tomada y modificada de Cunha (2006). 12
- Figura 7. Seção geológica esquemática da Bacia do Marañon. Tomado de Perupetro (2002). 13
- Figura 8. Seção geológica esquemática da Bacia de Ucayali. 13
- Figura 9. Seção geológica esquemática da bacia Madre de Dios. Tomado de Hermoza (2004). 14
- Figura 10. Proposta de coluna geológica sintética do sistema aquífero amazônico. Modificado e recentemente definido por Encarnacion, 2008; Rakhit, 2002; Cunha, 2007 14
- Figura 11. Seção geológica esquemática da bacia de Parecis. 15
- Figura 12. Descrição de como o GRACE/GRACE-FO mede a gravidade. Extraído da NASA. 20
- Figura 13. Fluxograma da metodologia. 31
- Figura 14. Componentes temporais do sinal TWS e GWS para o ARB, onde A) série temporal original, total ou derivada do GRACE/GRACE-FO, B) componente sazonal, C) ajuste sazonal, D) componente de tendência, E) componente residual e F) valor médio do boxpot da série temporal para o ARB. 87
- Figura 15. Série temporal média das principais variáveis hidrológicas da ARB. Onde A) Componente de tendência e B) Componente sazonal. Em vermelho, eventos de El Niño, em azul, eventos de La Niña de acordo com Moura et al., 2019 e em amarelo 88
- Figura 16. Média multianual das mudanças no armazenamento terrestre (ΔTWS) e de água subterrânea (ΔGWS) do ARB. 90
- Figura 17. Média multianual de: A) umidade do solo (SM), B) precipitação (Prec) e C) evaporação (Evap) da ARB. 91
- Figura 18. Distribuição espacial dos componentes sazonais de ΔTWS e ΔGWS para ARB. 92
- Figura 19. Representação espaço-temporal da variabilidade de ΔTWS para a ARB. Onde: A) SD de ΔTWS entre 2002-2021, B) tendência linear de ΔTWS entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente de tempo da sazonalidade entre 2002 e 2021. No eixo temporal dos bloxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha. 94
- Figura 20. Representação espaço-temporal da variabilidade de ΔGWS para a ARB. Onde: A) SD de ΔGWS entre 2002-2021, B) tendência linear de ΔGWS entre 2002-2021, C,D,E,F,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente temporal da sazonalidade entre 2002 e 2021. No eixo temporal dos bloxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha. 96
- Figura 21. Representação espaço-temporal da variabilidade da umidade do solo (SM) para o ARB. Onde: A) SD da SM entre 2002-2021, B) tendência linear da SM entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente temporal da sazonalidade entre 2002 e 2021. No eixo temporal dos bloxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha. 98
- Figura 22. Representação espaço-temporal da variabilidade da água da neve (SW) para o ARB. Onde: A) SD de SW entre 2002-2021, B) tendência linear de SW entre 2002-2021, C,D,E,F,G,H)

componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente temporal da sazonalidade entre 2002 e 2021. No eixo temporal dos bloxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha. 99

Figura 23. Representação espaço-temporal da variabilidade de precipitação do CHIRPS para o ARB. Onde: A) SD de TWS entre 2002-2021, B) tendência linear de TWS entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente de tempo de sazonalidade entre 2002 e 2021. No eixo temporal dos bloxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha.100

Figura 24. Representação espaço-temporal da variabilidade da evaporação para o ARB. Onde: A) SD da Evap entre 2002-2021, B) tendência linear da Evap entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente temporal da sazonalidade entre 2002 e 2021.....101

Figura 25. Variações na correlação sazonal entre o GWS e as principais variáveis hidrológicas para os pixels de análise no ARB.102

Figura 26. Mapas de correlação entre GWS e TWS, SM, Prec e Evap para a ARB. Onde A, B, C e D) correlação entre TS total, E, F, G e H) correlação sazonal máxima (fase) e I, J, K e L) defasagem para atingir a correlação sazonal máxima.104

Figura 27. Parâmetros topográficos derivados do MERIT Hydro-DEM para a ARB em uma resolução espacial de 0,25°. Onde A) Flow width (FW), B) Slope em grados, C) Specific catchment área (SCA).....106

Figura 28. Índice de umidade topográfica (TWI) do ARB. Onde a resolução espacial do TWI é: A) 90m e B) 0,25° (27500m aproximadamente).108

Figura 29. Localizam-se as hidrelétricas, áreas protegidas e zonas de incêndio em relação aos principais aquíferos da ARB. Dados da RAISG, 2022110

Figura 30. Vegetação da ARB. Dados do MapBiomass, 2022.111

Figura 31. Matriz de correlação espaço-temporal para ARB. Variáveis constantes ou fixas em fonte vermelha.....113

Figura 32. Mapas de correlação entre o GWS TS e as 38 variáveis espaço-temporais para o ARB....114

Figura 33. Boxplot dos valores de correlação entre o GWS e as 38 variáveis espaço-temporais para o ARB.....115

Figura 34. Proporção da variância cumulativa explicada.....116

Figura 35. Gráfico de dispersão dos loadings de PC1, PC2 e PC3. A magnitude extrema das cargas de PC1 é destacada na linha preta, a de PC2 na linha azul e a de PC3 na linha vermelha. Onde: A) PC1 com PC2, B) PC2 com PC3 e C) PC1 com PC3.117

Figura 36. Mapa dos dados transformados para PC1 e gráfico de barras com as magnitudes dos loadings. As cores das fontes correspondem ao azul nas magnitudes negativas e ao vermelho nas magnitudes.118

Figura 37. Mapa dos dados transformados para PC2 e gráfico de barras com as magnitudes dos loadings. As cores das fontes correspondem ao azul para magnitudes negativas e ao vermelho para magnitudes positivas.119

Figura 38. Mapa dos dados transformados para PC3 e gráfico de barras com as magnitudes dos loadings. As cores das fontes correspondem ao azul para magnitudes negativas e ao vermelho para magnitudes positivas.120

Figura 39. Composição RGB de PC1, 2 e 3 obtida a partir da correlação entre a TS total do GWS e as variáveis hidrometeorológicas. No diagrama ternário, a cor da fonte indica a direção do loading: azul é negativo e vermelho é positivo.121

Figura 40. Diagrama resumido da magnitude das CPs, correlações entre GWS, variáveis hidrometeorológicas e WL dos poços de monitoramento do RIMAS, para pixels de análise a 0,25° da ARB.....	123
Figura 41. Diagrama ternário de PC1, 2 e 3, sobreposto às correlações entre os STs totais dos WLs dos poços RIMAS e o GWS a 0,25°.....	124
Figura 42. Zoneamento do GWS de acordo com SD e TL para o ARB.....	125

Lista de Tabelas

Tabela 1. Resumo dos parâmetros hidráulicos do Aquífero Alter do Chão.....	17
Tabela 2. Resumo dos parâmetros hidráulicos do Aquífero Içá.....	17
Tabela 3. Resumo dos parâmetros hidráulicos do Acuitardo Solimoes-Repouso.....	18
Tabela 4. Resumo dos parâmetros hidráulicos do Aquífero Aluvial.....	19
Tabela 5. Resumo do conjunto de dados e variáveis usadas na pesquisa. Onde ground-based measurements (GBM), remote sensing products (RS), e land surface models (LSM).	32

Lista de Siglas

Δ	Mudanças
AB	<i>AdaBoost</i>
AI	<i>Artificial Intelligence</i> ou Inteligência Artificial
AAS	<i>Amazon Aquifer System</i> ou Sistema Aquífero Amazônico
ARB	<i>Amazon River Basin</i> ou Bacia Amazônica ou Bacia do Rio Amazonas
CW	<i>Canopy Water</i> ou Água do Dosel
CWS	<i>Canopy Water Storage</i> ou Armazenamento de Água do Dosel
DL	<i>Deep Learning</i> ou Aprendizado Profundo
E	Leste
GWS	<i>Groundwater Storage</i> ou Armazenamento de Água Subterrânea
HR	<i>High Resolution</i> ou Alta Resolução
LR	<i>Low Resolution</i> ou Baixa Resolução
ML	<i>Machine Learning</i> ou Aprendizado de Máquina
N	Norte
O	Oeste
PC	Componentes Principais
Qs	Escoamento Superficial
RF	<i>Random Forest</i>
SD	<i>Standard Deviation</i> ou Desvio Padrão
SM	<i>Soil moisture</i> ou Umidade do Solo
SMS	<i>Soil moisture storage</i> ou Armazenamento de Água na Umidade do Solo
SR	Sensores Remotos
ST	Séries Temporais
S	Sul
SuWS	<i>Superficial Water Storage</i> ou Armazenamento de Água Superficial
SWS	<i>Snow Water Storage</i> ou Armazenamento de Água da Neve
TL	Tendência linear
TWS	<i>Terrestrial Water Storage</i> ou Armazenamento de Água Terrestre
TS	<i>Time series</i> ou Séries Temporais
WL	Níveis de Água Subterrânea

1. INTRODUÇÃO

O entendimento da interação entre os componentes do balanço hídrico de uma bacia hidrográfica é base fundamental para o monitoramento e gestão sustentável das águas subterrâneas. A bacia hidrográfica do rio Amazonas (ARB) e sua floresta tropical são um dos maiores e mais importantes ecossistemas de água doce do mundo, pois os fluxos de água que ela contribui para o oceano e a atmosfera afetam o ciclo global da água (Coe et al., 2016). Nos últimos anos, os esforços para entender o comportamento hidrodinâmico da superfície e subsuperfície do ARB aumentaram. Com base nas contribuições de Ferreira(2006), UNESCO (2007), Rosário (2011) e Abreu (2013), tem sido definido o Sistema Aquífero Amazonas (AAS), uma unidade hidrodinâmica regional, multibacia, multicamadas, livre e aflorante com uma mesma superfície de água subterrânea e fluxo heterogêneo, composta por nove unidades hidrogeológicas (Alter do Chão, Aquíferos Aluviais, Barreiras, Cobertura Cenozóica de Porto Velho, Içá, Pirabas, Rio Branco, Tucunaré e Aquitarde Solimões- Repouso), que representam um armazenamento permanente de 123,838 km³, apenas no território do Brasil (ANA, 2015). Somente o Sistema Aquífero Alter do Chão (ACAS) ocupa uma área de 290,500 km² e representa armazenamento de importantes reservas hídricas subterrâneas, preliminarmente estimadas em 33,000 km³ (ANA, 2015), abastecendo cidades como Manaus e Santarém.

A atividade antrópica pode afetar o ciclo natural da água, gerando mudanças no balanço hídrico e conseqüentemente na sua circulação. Atividades como o desmatamento aumentaram 22% em 2021 em relação a 2020 com 13,235 km² no total da Amazônia legal (INPE, 2021), causando alterações não só físicas, mas hidroclimáticas, hidrogeológicas e ecológicas em diferentes escalas (Neill, 2011; Nobre, 2009; Costa et al., 2003; Dias et al., 2015; Coe et al., 2017; Latrubesse, 2017) . Em média e pequena escala, a implantação de pastagens no lugar de florestas aumenta o escoamento superficial em até 18 vezes em relação à infiltração, diminuindo proporcionalmente a recarga dos aquíferos (Neill, 2011). Em larga escala, os modelos globais espectrais *Atmospheric General Circulation Model* (AGCM) e o *Ocean General Circulation Model* (OGCM) mostram que a remoção de florestas afeta o balanço hídrico na bacia, com reduções na precipitação de até 60%, e até a temperatura do Oceano Pacífico, cujo aumento acaba por modificar a interação oceano-atmosfera e o clima global (Nobre, 2009). O desmatamento também reduz as taxas de evapotranspiração e aumenta a temperatura da superfície (Costa et al., 2003; Dias et al., 2015; Coe et al., 2017). Outras atividades, como a instalação de barragens para geração de energia elétrica, afetam o clima localmente devido à alteração da evaporação e do grau de infiltração e saturação dos solos, além de reduzir a capacidade de transporte de sedimentos. Áreas de preservação ambiental e reservas indígenas podem atuar como elementos de mitigação para proteger as áreas mais importantes e sensíveis do sistema, assumindo um papel relevante na redução de tais impactos (Latrubesse, 2017).

Estudos sobre os efeitos da evapotranspiração e outras variáveis hidrológicas sobre as mudanças nos recursos hídricos têm sido realizados para a ARB (Costa et al., 2003; Dias et al., 2015; Coe et al., 2017; Latrubesse, 2017), porém, sobre as mudanças no armazenamento de água subterrânea e seus efeitos têm sido poucos. A água subterrânea da ARB regula as taxas de fluxo na drenagem e evapotranspiração (Miguez-Macho e Fan, 2012). Pesquisas como a de Panday (2015), onde através da integração do GRACE e do modelo de superfície terrestre IBIS, foi evidenciada a influência do armazenamento de água subterrânea associada a eventos de seca na vazão interanual do rio Xingu. Reduções locais a regionais têm sido evidenciadas nos níveis piezométricos do aquífero Alter do Chão em Manaus, produto da superexploração das águas subterrâneas, com cerca de 3,200 poços cadastrados e mais de 6.000 poços clandestinos (Saraiva, 2015). Além disso, as contribuições das águas subterrâneas nas mudanças totais no armazenamento de água na ARB foram investigadas por De Paiva (2013) e Pokhrel (2013). A primeira por meio do modelo MGB-IPH, mostrando que 56% são contribuídos pelas águas superficiais, a água contida no solo com 27% e subterrânea com 8%. A segunda, utilizando o modelo LEAF-Hydro-Flood (LHF), estimou a água contida no solo em 71%, nas planícies de inundação em 24% e os rios em 5%. Essa discordância entre os dois métodos de simulação destaca a necessidade de considerar as águas subterrâneas nos estudos da hidrologia da Amazônia (Heerspink, 2020).

No território brasileiro, o Serviço Geológico do Brasil (CPRM) implantou a Rede Integral de Monitoramento de Águas Subterrâneas (RIMAS) e o Sistema de Informação de Águas Subterrâneas (SIAGAS). A primeira, correspondente a uma rede de piezômetros de natureza quantitativa, que mantém um registro contínuo das variações dos níveis de água (WL), bem como de alguns parâmetros químicos. O segundo consiste em um banco de dados de poços permanentemente atualizado, composto por informações construtivas, geológicas, hidrodinâmicas e hidroquímicas (CPRM, 2020). O monitoramento local e regional das mudanças de quantidade ou armazenamento de água, tanto superficial quanto subterrânea, apresenta limitações espaciais e temporais. Impõe-se, dessa forma, um desafio para a medição das variáveis hidroclimatológicas e hidrogeológicas em grandes escalas devido aos altos custos de obtê-las em atividades de campo e às dificuldades para se representar a heterogeneidade espacial das variáveis de interesse, além dos erros inerentes das amostragens (Gao et al., 2010).

Outras ferramentas para manter um registro contínuo do comportamento das variáveis hidroclimáticas e hidrogeológicas são os Sensores Remotos (SR). A implementação de metodologias baseadas em informações de SR permite monitorar diversas variáveis, em larga escala e em tempo real, facilitando a aquisição de informações em áreas de difícil acesso e com custos baixíssimos. O lançamento dos satélites GRACE (*Gravity Recovery and Climate Experiment*) e GRACE FO (*Follow One*), em 2002 e 2018 respectivamente, permitiram monitorar as mudanças no armazenamento de água terrestre (ΔTWS) e, conseqüentemente, nas águas subterrâneas. Sua operação consiste em medir as mudanças espaço-temporais do campo gravitacional terrestre, representadas em concentrações de massa, as quais são

causadas pelo movimento da água em todos os seus estados ao longo da superfície da terra (Tapley, 2019). Durante os últimos 20 anos, GRACE e GRACE-FO tornaram possível monitorar o ciclo hidrológico terrestre, apoiando pesquisas sobre balanços de massa em geleiras (Sasgen, 2013), mudanças no nível do mar (Llovel, 2014), mudanças no armazenamento de água terrestre (Scanlon, 2018), mudanças no armazenamento de água subterrânea (Rodell, 2018) e potencial de inundação (Reager, 2015) entre muitos outros.

No campo da hidrogeologia, GRACE e GRACE-FO subsidiaram magnitudes do armazenamento de água terrestre com as quais são estabelecidas as mudanças no armazenamento de água subterrânea em diferentes partes do mundo, como África (Ahmed, 2014), norte da China (Feng, 2013), Índia (Rodell, 2009), vale central da Califórnia (Scanlon, 2012) e centro do México (Castellazzi, 2018). No entanto, a baixa resolução espacial dos satélites ($1^{\circ}\times 1^{\circ}$ ou $3^{\circ}\times 3^{\circ}$) restringe sua aplicação, obtendo-se erros de menos de 1 cm de coluna de água equivalente, em bacias maiores que 400.000 km² (Swenson et al., 2003). Muito embora a precisão de 1 cm seja aparentemente baixa, isso equivale a uma diferença de 121ML de água numa área de 400.000 km². No caso de aquíferos confinados, grandes mudanças potenciométricas traduzem-se em pequenas mudanças volumétricas e conseqüentemente na massa, onde GRACE não seria muito útil (Alley & Konikow, 2015). Apesar de suas limitações, ainda é uma ferramenta muito valiosa em áreas sem acesso e onde a informação é escassa, como é o caso do ARB e do AAS.

Esta pesquisa busca implementar uma metodologia baseada em SR e medições *in situ* para caracterizar mudanças no armazenamento de água subterrânea na ARB e especialmente no AAS, determinando os componentes do armazenamento total de água terrestre a partir de variáveis hidrometeorológicas (balanço hídrico terrestre). A resolução espacial dos dados de GRACE/GRACE-FO foi aumentada implementando algoritmos de machine learning em variáveis hidrometeorológicas de maior resolução, permitindo assim extrair características espaciais e temporais dos dados, replicando-as nas previsões de maior resolução. Esse aumento na resolução espacial proporciona uma maior clareza na identificação de novos padrões espaciais, bem como a capacidade de realizar análises mais detalhadas sobre os efeitos da atividade antrópica e dos fenômenos locais. Os algoritmos Adaboost e Random Forest foram utilizados, contribuindo para os processos de pesquisa junto com outros algoritmos e em outras áreas, conforme demonstrado em trabalhos anteriores (Jyolsna et al., 2021; Milewski et al., 2019; Miro & Famiglietti, 2018; Sahour et al., 2020; Seyoum et al., 2019; Verma & Katpatal, 2020; Vishwakarma et al., 2021; Zhong et al., 2021; Zuo et al., 2021). Esta pesquisa propõe identificar os efeitos e alterações que tanto a atividade antrópica quanto as mudanças climáticas têm sobre o armazenamento de água subterrânea na ARB e particularmente no AAS, identificando o alcance e as limitações do uso de ML e da metodologia proposta como ferramenta de monitoramento de água subterrânea.

1.1. Objetivos

1.1.1. Objetivo geral

Essa pesquisa tem por objetivo caracterizar e avaliar as mudanças no armazenamento de água subterrânea no Sistema Aquífero Amazônico (AAS) e a Bacia do rio Amazonas (ARB), com base na integração das anomalias gravitacionais medidas pelo satélite GRACE / GRACE FO e sua validação com os registros de poços de monitoramento (RIMAS) da região Norte e Centro-Oeste do Brasil durante o período entre 2002 e 2021.

1.1.2. Objetivos específicos

1. Implementar e validar um modelo de *downscaling* (de 1° a 0,25°) para as mudanças de armazenamento de água terrestre (Δ TWS) do GRACE/GRACE-FO a partir das variáveis hidrometeorológicas e morfológicas;
2. Obter os valores espaço-temporais das mudanças de armazenamento de água subterrânea (Δ GWS) a 0,25° e validar suas séries temporais com os níveis de água (WL) registrados pelos piezômetros da RIMAS;
3. Caracterizar os padrões espaciais e temporais de Δ GWS de curto e longo prazo em uma resolução de 0,25°;
4. Identificar as principais relações espaciais e temporais de Δ GWS a 0,25° com as variáveis hidrometeorológicas e morfológicas;
5. Determinar os possíveis efeitos das atividades antropogênicas sobre a Δ GWS a 0,25° na ARB;
6. Propor um zoneamento hidrogeológico derivado dos comportamentos espaço-temporais do Δ GWS e sua relação com as variáveis hidrometeorológicas e morfológicas.

1.2. Área de Estudo

A bacia do rio Amazonas (ARB) é reconhecida como a maior do mundo, abrangendo aproximadamente $6,3 \times 10^6$ km² distribuídos entre seis países: Brasil (67,9%), Peru (13,0%), Bolívia (11,2%), Colômbia (5,5%), Equador (1,7%) e Venezuela (0,7%). Essa bacia é caracterizada por sua complexidade e diversidade, composta por vários rios, extensas planícies aluviais e áreas úmidas (Reis et al., 2019). Ela se estende no centro e leste da América do Sul, com faixas de latitude e longitude entre 5°N-20°S e 80°W-50°W, respectivamente. Apresenta altitudes máximas na cordilheira dos Andes, na montanha Huayhuash, no Peru, com 6.450 metros acima do nível do mar, até a foz do rio Amazonas no Oceano Atlântico, no nordeste do Brasil. O rio Amazonas percorre aproximadamente 7.100 km com um fluxo médio de $2,10 \times 10^5$ m³/s (Braga et al., 2011). Seus principais afluentes incluem os rios Madeira, Xingu, Negro, Tapajós e Japurá. A ARB destaca-se por sua grande extensão e complexidade hidrológica, com altas taxas de precipitação e evapotranspiração, e grandes variações no armazenamento e descarga de água, o que a torna um ator chave no sistema climático global (Gash et al., 2013; Gatti et al., 2021; Nagy et al., 2016). A precipitação anual média na maior parte da bacia varia entre 1.000 e 3.000 mm,

1.2. Área de Estudo

alcançando valores máximos de 4.000 mm no Noroeste e mínimos de 100 mm na região dos Andes (Haghtalab et al., 2020; Maeda et al., 2017). A floresta tropical contribui significativamente para a quantidade de água atmosférica, com uma média anual de 1.000 a 1.500 mm (Maeda et al., 2017). As mudanças sazonais na água são fundamentais para o desenvolvimento e a sustentação das florestas tropicais (Leite et al., 2012) e a biodiversidade (Junk, 1977; Junk et al., 2010). Neste estudo, a ARB é dividida em três unidades de acordo com sua elevação: bacia alta (>1000 m), bacia média (170 m-1000 m) e bacia inferior (<170 m) (Figura 1).

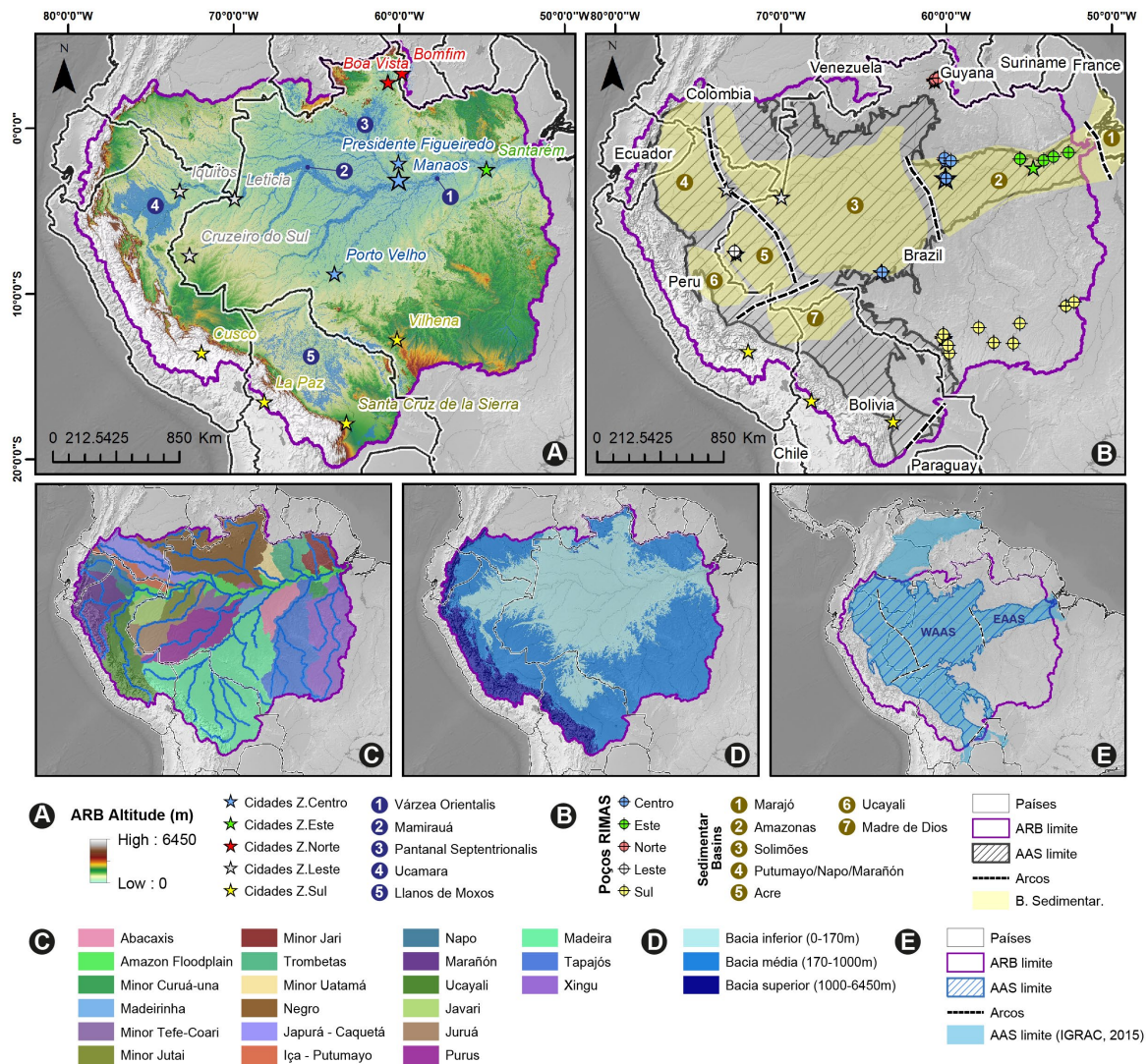


Figura 1. Área de estudo onde A) Topografia da ARB, B) Bacias sedimentares e localização dos poços do RIMAS, C) Sub-bacias hidrográficas, D) Subdivisão da ARB, e E) toda a AAS.

No subsolo da ARB encontra-se o Sistema Aquífero Amazônico (AAS), um dos maiores aquíferos transfronteiriços do mundo, abrangendo as províncias hidrogeológicas do Orinoco e da Amazônia, definido em 2005 pela UNESCO/OEA. A extensão real do AAS ainda apresenta alguma incerteza, conforme evidenciado nos mapas propostos por diferentes autores (ISARM e UNESCO, 2009; Margat, 2007; Anjos et al., 2007 e Rosário et al., 2016). Neste estudo, o AAS é definido dentro da província hidrogeológica amazônica com base nas unidades sedimentares do mapa geológico da

América do Sul, cobrindo uma área total de $3,1 \times 10^6$ km². O AAS é delimitado a oeste pelo sistema de falhas subandino, ao norte e ao sul pelos escudos de rochas cristalinas da Guiana e do Brasil (Figura 2). O AAS é um sistema de fluxo regional, heterogêneo, multicamadas e multibacias, com uma superfície potenciométrica livre e conectada a rios, disposta sobre rochas paleozoicas (afloramentos de diabásio e/ou formações sedimentares) e embasamento cristalino (Agência Nacional de Águas - ANA, 2015). É formada pelas bacias sedimentares do Amazonas, Solimões, Acre, Marañon (Oriente e Putumayo), Ucayali e Madre de Dios, que são preenchidas por sedimentos consolidados e não consolidados com semelhanças estratigráficas do Cretáceo Inicial ao Terciário Final (Rosário et al., 2016). Entre os principais aquíferos não confinados que compõem o AAS estão os aquíferos Alter do Chão, Içá, Tucunaré, Aluvial e Solimões-Repouso. Os materiais sedimentares no AAS variam em granulometria, desde argilas, siltes, areias até cascalhos, com espessuras que variam de 500 a 1.000 m na bacia Amazônica, 4.000 m na bacia do Marajó e de 150 a 1.000 m nas sequências arenosas das bacias do Solimões e do Acre. A ANA assume a hipótese de uma única superfície potenciométrica entre as unidades cenozoicas, cretáceas e paleozoicas, independentemente da natureza dos materiais, e elas estão inter-relacionadas por variações laterais ou curtos-circuitos de fluxo através de fraturas dentro da porção brasileira da AAS.

As principais estruturas geológicas que delimitam as bacias sedimentares são os arcos Iquitos, Purus, Gurupá, Fitzcarrald, Vaupes e Michicola e o cinturão de falhas subandino (Figura 2). A AAS é dividida em duas regiões: o Sistema Aquífero da Amazônia Ocidental (WAAS) e o Sistema Aquífero da Amazônia Oriental (EAAS), separados pelo Arco do Purus (Figura 2). Apesar das diferenças de nomenclatura entre os países, subunidades como o Sistema Aquífero Tikuna (Rosário et al., 2016) ou o Sistema Aquífero Grande Amazônia (SAGA) (Matos de Abreu et al., 2013) foram propostas para descrever determinados comportamentos hidráulicos regionais no AAS. Os aquíferos de Parecis e Ronuro ao SE, e Boa Vista ao N, foram considerados durante a análise devido à sua localização dentro

da ARB, apesar de não fazerem parte do AAS.

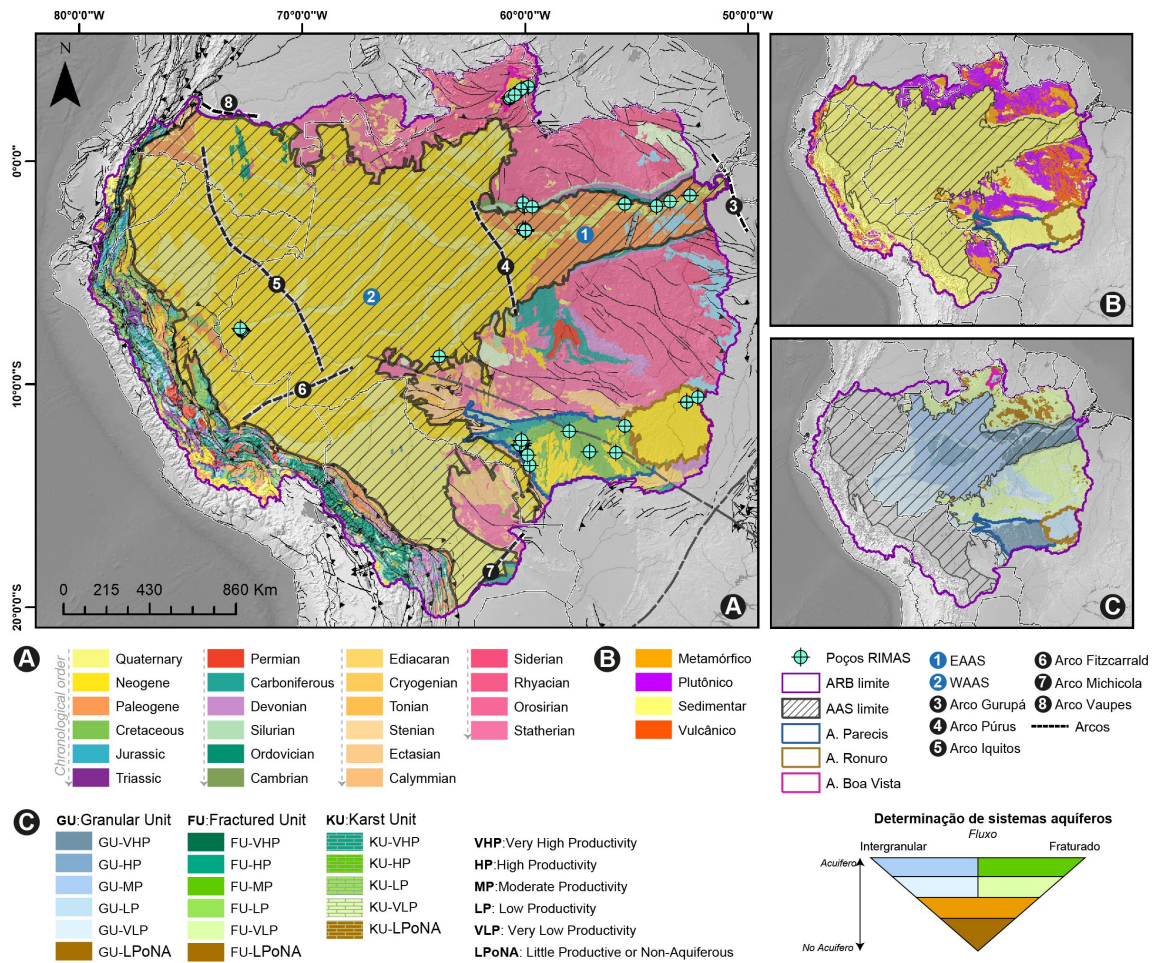


Figura 2. Área de estudo com a) localização dos poços de monitoramento e b) principais unidades chronostratigráficas extraídas do mapa geológico da América do Sul.

Na porção brasileira da ARB, há 69 poços de monitoramento pertencentes à Rede Integrada de Monitoramento de Águas Subterrâneas (RIMAS). Esses poços estão localizados em 26 cidades classificadas nas zonas Norte, Leste, Sul, Oeste e Central. Desses 69 poços totais, 41 foram considerados na etapa de validação devido a lacunas nas informações dos registros, tais como falta de continuidade na série temporal ou a valores anômalos nos registros. Os furos de sondagem atravessam diferentes materiais geológicos e foram instalados com profundidades que variam entre 9 e 260 metros. Algumas das cidades com o maior número de poços são Manaus, Vilhena, Cruzeiro do Sul, Bonfim, Boa Vista e Presidente Figueiredo, sendo os locais ideais para a análise de validação dos resultados. A localização dos poços nos pixels analisados corresponde a um número de pixels individuais de acordo com as soluções, sendo 19 em CSR a 1°, 21 em GFZ a 1°, 21 em JPL a 1° e 28 em uma resolução de 0,25°. Além disso, 18 pixels individuais distribuídos dentro da ARB nos domínios dos Andes, da AAS e dos aquíferos de Parecis, Ronuro e Boa Vista foram definidos para a interpretação espacial das séries temporais de diferentes variáveis e da análise de componentes principais (PCA). A distribuição desses pixels obedece à análise espacial da ARB, na qual foram incluídos diferentes domínios, características

geológicas, cobertura vegetal, características climáticas e aspectos geográficos. Alguns desses pixels coincidem com cidades ou locais com poços de monitoramento do RIMAS.

2. REVISÃO DE LITERATURA

2.1. Geologia

A Bacia do Rio Amazonas (ARB) consiste em três domínios principais compostos pelas unidades sedimentares da bacia inferior, os escudos brasileiro e guianense, bem como a Cordilheira dos Andes. As principais estruturas geológicas regionais são os arcos Gurupá, Purus, Vaupes, Iquitos, Fitzcarrald e Michicola, bem como o sistema de falhas subandino. No cráton amazônico, as unidades sedimentares da bacia inferior e média constituem as bacias sedimentares de Amazonas, Solimões, Acre, Madre de Dios, Beni, Ucayali, Marañon, Oriente e Putumayo, preenchidas por sedimentos do início do Cretáceo ao final do Terciário, consolidados e não consolidados com semelhanças estratigráficas (Rosario et al., 2016).

A leste da ARB, em território brasileiro e na foz do rio Amazonas, está a bacia sedimentar do Marajó (Figura 3), também chamada de bacia de rifte, com uma área aproximada de 55000 km² na zona insular e 177000 km² na área composta por cobertura cenozoica. É uma bacia distensional intracontinental neocomiana do tipo aulacogênica, ou seja, formada em uma zona de junção tectônica tripla. É composta por sedimentos predominantemente arenosos pré, sin e pós-rifte. Algumas das principais formações que o compõem são Limoeiro, Marajó, Barreiras e Tucunaré. A primeira, definida por Schaller et al., 1971, consiste em arenitos friáveis, finos a grossos, com intercalações de conglomerados e argilas cinza-escuras, com espessura de cerca de 3000 m e idade variada entre o Turoniano e o Eoceno (Zalán & Matsuda, 2007). A Formação Marajó, definida por Schaller et al., 1971, é composta por arenitos, argilas, xistos e calcários com espessura máxima de 2000 m, com idade entre o Eoceno e o início do Eomioceno. A Formação Barreiras tem uma espessura de 700 m, composta de conglomerados, arenitos e argilas, com idade entre o Neo-Oligoceno-Eomioceno e o Mesomioceno. A Formação Tucunaré, composta principalmente por arenitos de idade Plio-Pleistocênica e com espessura de aproximadamente 1000 m.

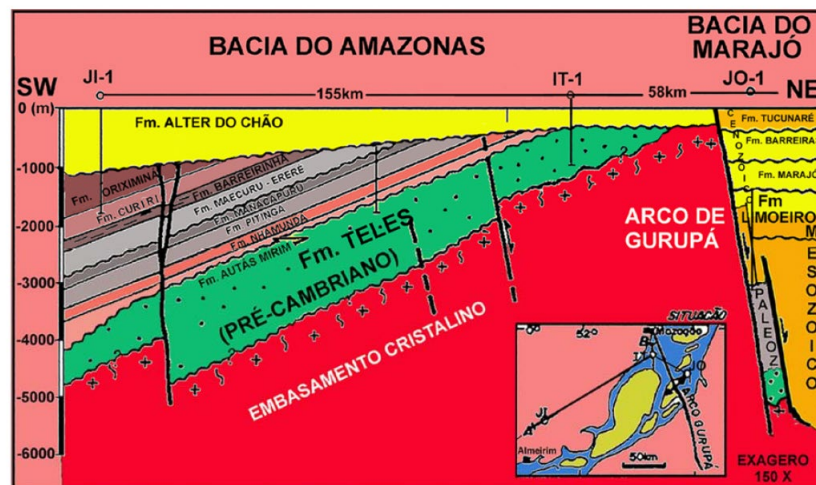


Figura 3. Seção geológica esquemática do Arco de Gurupá. Modificado de Wanderley Filho & Costa (1991).

Sobre a região subequatorial ao norte do Brasil, encontra-se a bacia sedimentar do Amazonas (Figura 4), a qual tem uma forma alongada nas direções WSE-ENE. Trata-se de uma bacia paleozoica intracratônica com cerca de 620000 km², 1300 km de espessura e 380 km de comprimento médio. Limitada pelos Arcos de Gurupá ao leste e de Purus ao oeste, ao norte pelo Escudo das Guianas e ao sul pelo Escudo Brasileiro. O preenchimento sedimentar e ígneo da bacia tem espessura de 5000 a 6000 m e é composto por rochas paleozoicas do Ordoviciano ao Permiano. Algumas das principais formações da bacia são Prosperança, Trombetas, Cajari, Iracema, Urupadi, Curuá, Curiri, Faro, Oriximiná, Urucuricaia, Itaituba, Nova Olinda, Monte Alegre, Arari, Andirá, Jazida da Fazendinha e Alter do Chão. Esta última é a de maior interesse no desenvolvimento deste projeto, composta por intercalações de arenitos, argilas, limos e alguns conglomerados, predominantemente vermelhos. Com uma espessura registrada em perfuração de 1266 m (ANA, 2016), embora com uma espessura média entre 200 a 400 m. Em direção às margens norte e sul, cobre várias unidades paleozoicas com discordâncias erosivas, na margem leste cobre parcialmente o arco de Gurupá. Em direção à porção central da bacia, a unidade repousa discordantemente sobre as formações Jazida da Fazendinha, Arari e parte da formação Nova Olinda.

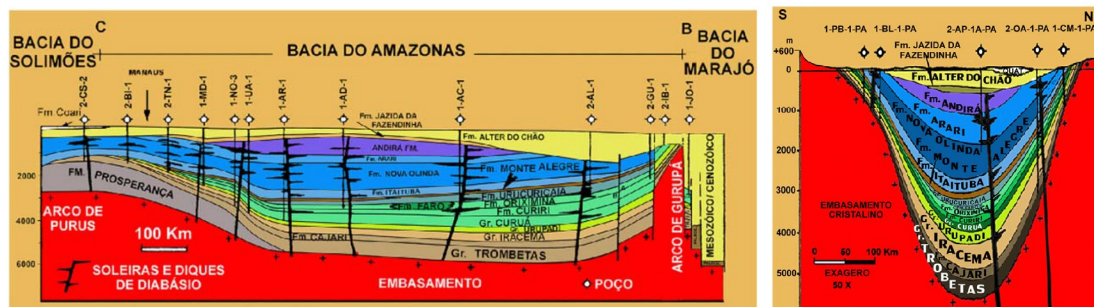


Figura 4. Seções longitudinais e transversais da Bacia do Amazonas. Tomado de Caputo (1984).

A bacia sedimentar de Solimões (Figura 5) está localizada entre Brasil, Colômbia e Peru, com uma orientação E-W, limitada pelos escudos das Guianas ao norte e Brasileiro ao sul, pelo arco de Purus a leste e o arco de Iquitos a oeste. É composta por uma área ocidental (Jandiatuba) com sedimentação de facies marinhas profundas e uma área oriental (Juruá) com sedimentação rasa. Em termos gerais, a bacia de Solimões comporta-se como uma rampa, cuja parte mais profunda corresponde à Sub-bacia de Jandiatuba e a parte mais rasa à sub-bacia de Juruá, que constituía uma plataforma. As espessuras máximas das colunas ígneas-sedimentares das sub-bacias de Jandiatuba e Juruá são de 3500 e 3800 m, respectivamente. Não se estabeleceu uma origem concreta para a bacia sedimentar, no entanto, Eiras & Wanderley Filho (2006) propõem a hipótese de uma subsidência de origem flexural, progressiva de oeste a leste, onde a bacia poderia ser uma rampa vinculada à tectônica paleozoica da região ocidental da América do Sul. Algumas das principais formações presentes na bacia são Benjamin Constant, Grupo Marimari, Jutai, Itamarati, Uerê, Grupo Jandiatuba, Jaraqui, Sernambi, Grupo Tefé, Juruá, Carauari, Taititu, Fonte Boa, Javari, Repouso, Solimões e Ica. Esta última, formada durante o pleistoceno e proposta por Maia et al. (1977), é constituída principalmente por arenitos amarelos avermelhados, de

granulometria fina a conglomerática, às vezes com limolitas argilosas e conglomerados, limolitas e lutitas acinzentadas. A formação provém dos Andes, de ambiente de deposição fluvial, possivelmente formando a terraça mais antiga depositada sobre a formação Solimões (Maia et al., 1977). A formação Iça tem espessura de algumas dezenas de metros até 140m e abrange uma área de mais de 600000 km² no Brasil. Limita-se a leste, um pouco além do Arco de Purus, dentro da bacia do Amazonas, em contato discordante com a Formação Alter do Chão. Quanto à Formação Solimões, é composta principalmente por argilas, limolitas, arenitos argilosos finos a médios, brechas com fragmentos subangulares, com concreções carbonáticas, gipsíferas e ferruginosas, com intercalações de alguns lignitos (Radambrasil, 1977; Hoorn, 1993). Tem espessura variável entre 200 a 600m, dependendo da paleotopografia do embasamento, afinando-se para leste, perto do arco de Purus. Trata-se de uma molassa distal do orógeno andino, que recobre discordantemente a Formação Repouso e parte do embasamento cristalino nas margens da bacia. Distribui-se sobre as bacias sedimentares de Solimões e Acre, de ambiente continental paludal, lacustre, flúvio-lacustre, fluvial e flúvio deltaico (Carlotto et al., 2008). Tem idade neomiocênica-eopliocênica, sendo mais recente que a formação Alter do Chão (Cruz, 1984; Nogueira et al., 2013).

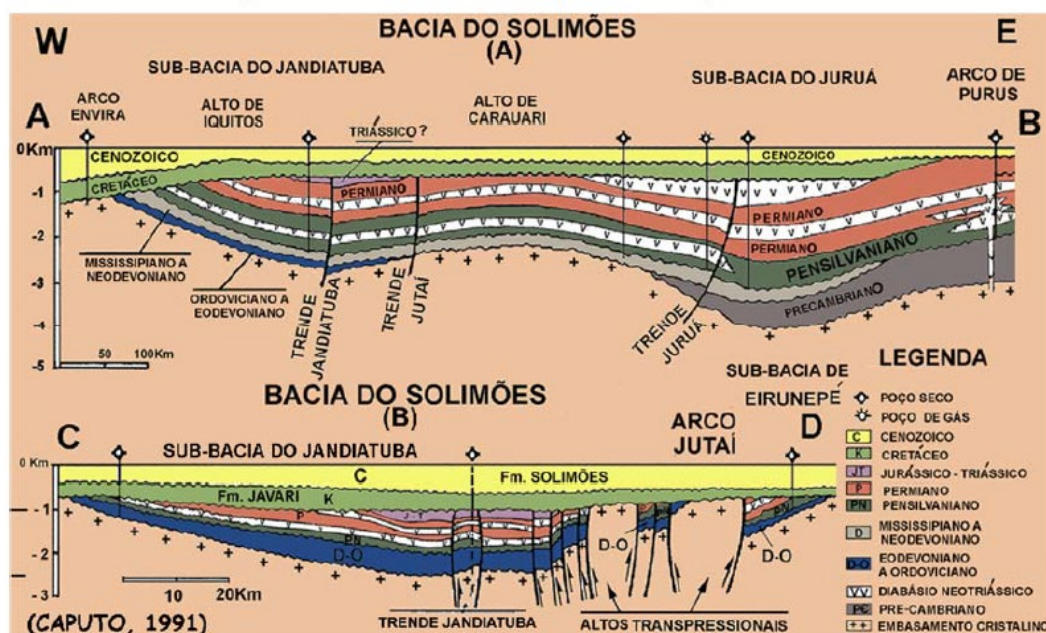


Figura 5. Seções longitudinais e transversais da Bacia do Solimões. Tomado de Caputo (1991).

A bacia sedimentar do Acre (Figura 6) localiza-se a oeste do Brasil e ocupa uma área total de 23000 km². Ela separa-se da bacia de Solimões a leste pelo arco de Iquitos. Ao nordeste, tem continuidade parcial com a bacia de Marañón, no Peru. A oeste, separa-se da bacia de Ucayali pela divisória de águas que forma a Serra do Divisor, no Peru. E ao sudeste, limita-se com a Bacia Madre de Dios, na Bolívia, pelo arco de Fitzcarrald. Trata-se de uma bacia de antepaís com alta atividade tectônica, devido à sua proximidade com o Orógeno Andino. O preenchimento sedimentar pode alcançar entre 6000 a 5000 metros, de acordo com dados sísmicos (ANA, 2016), abrigando as formações Formosa,

Apuí, Cruzeiro do Sul, Serra do Divisor, Rio do Moura, Juruá Mirim, Capanaua, Moa, Rio Azul, Divisor, Ramon, Repouso e Solimões.

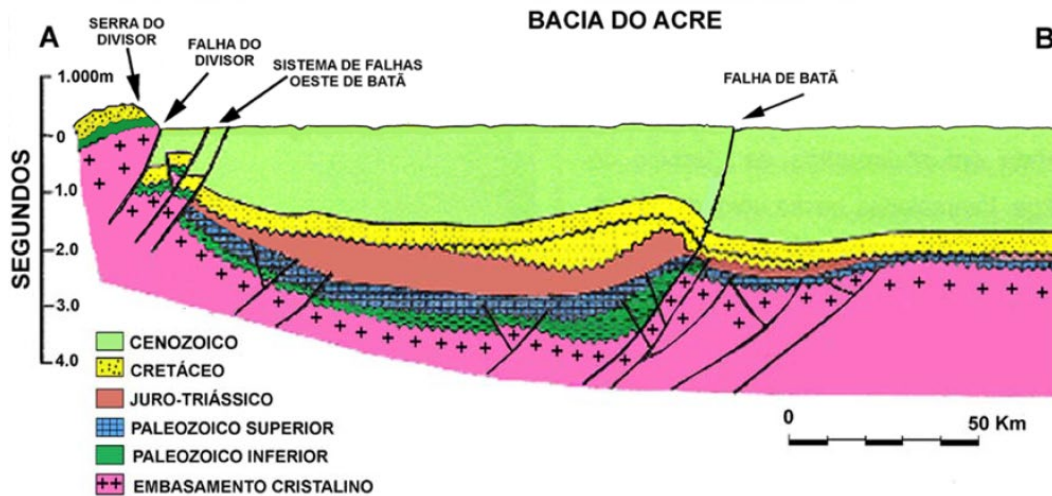


Figura 6. Seção geológica esquemática da Bacia do Acre. Tomada y modificada de Cunha (2006).

Ao noroeste da ARB, entre o arco de Iquitos e o sistema de falhas subandino, encontra-se uma cadeia de bacias sedimentares de antepaís, denominadas como Putumayo na Colômbia, Oriente no Equador e Marañón no Peru (Figura 7). Essa série de bacias sedimentares, limitada entre a cordilheira dos Andes e o cráton Guyano-brasileiro, é fortemente assimétrica, com eixo N-S e subsidências máximas em direção ao oeste. A bacia de Putumayo possui uma área total de 30000 km², cujo preenchimento sedimentar é perfeitamente correlacionável com o da Bacia Oriente (Higley, 2001; Baby et al., 2004; Gonçalves et al., 2002). É composta pelas formações Saldaña, Caballos, Villeta, Rumiyaco, Pepino, Orteguaza, Orito-Belen, Ospina e Caiman. Por outro lado, a bacia Oriente abrange uma área de aproximadamente 100000 km², com um depocentro localizado no Peru, com espessuras que atingem 5000m em sedimentos terciários e Cretáceos (Young R. e Harás J., 1996). O preenchimento sedimentar da bacia Oriente é composto pelas formações Pumbuiza, Macuma, Santiago, Chapiza, Hollín, Napo, Tena, Tiyuyacu, Orteguaza, Chalcana, Arajuno, Chambira, Mesa e Mera. A bacia de Marañón possui uma área de 450000 km² e constitui uma grande bacia subandina que se estende até as bacias de Oriente e Putumayo, limitando-se com o arco de Iquitos a leste, com a zona subandina a oeste (bacias de Huallafa e Santiago) e ao sul com o alto de Cushabatay-Contaya, que a separa da bacia de Ucayali. Algumas das formações que compõem o preenchimento sedimentar da bacia são Cabanillas, Ambo, Tarma, Copacaba, Mitu, Pucara, Sarayaquillo, Cushabatay, Raya, Agua Caliente, Chonta, Vivian, Cachiyacu, Yahuarango, Pozo, Chambira, Pebas, Marañón e Corrientes.

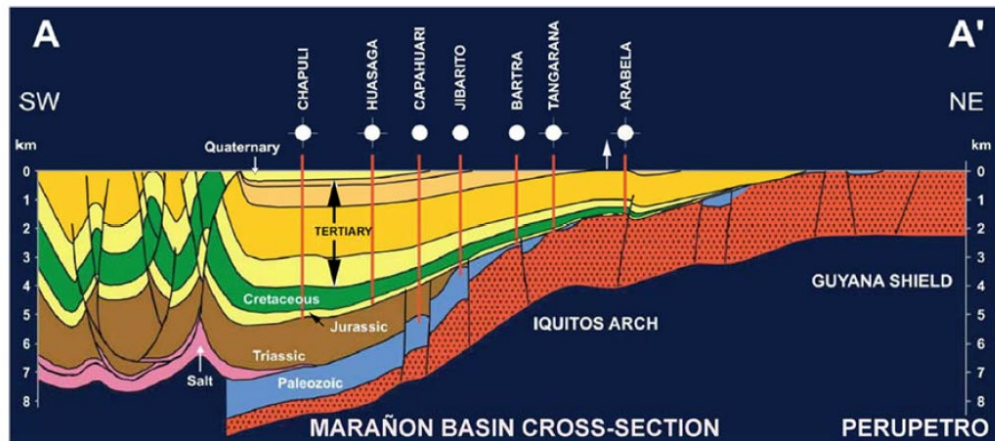


Figura 7. Seção geológica esquemática da Bacia do Marañon. Tomado de Perupetro (2002).

A bacia sedimentar de Ucayali (Figura 8) limita-se a leste com a Serra do Divisor (também chamada de Divisor Moa), ao nordeste com o Arco de Contaya que a separa da bacia de Marañón, a oeste com a zona subandina e ao sul com o Arco de Fitzcarrald. É constituída pelas formações Mitu, Pucará, Sarayaquillo, Cushabatay, Raya, Agua Caliente, Chonta, Vivian, Cachiyacu, Huchpayacu, Casa Blanca, Yahunango, Chamira e Ipururo.

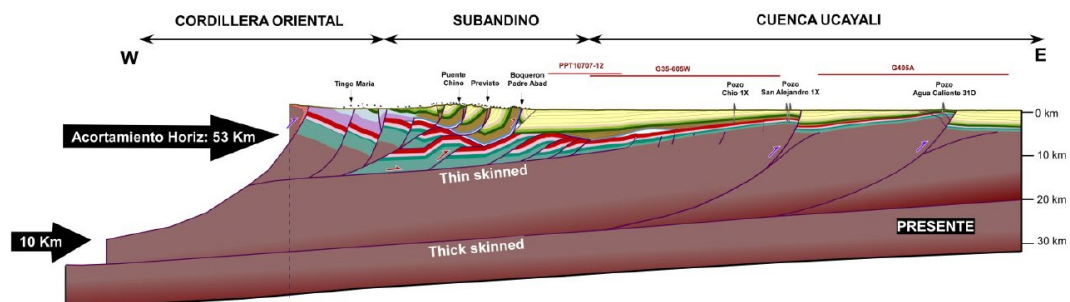


Figura 8. Seção geológica esquemática da Bacia de Ucayali. Tomado de Hurtado et al., 2018.

Ao Sul, localiza-se a bacia sedimentar de Madre de Dios, que possui uma área aproximada de 43000 km² e um preenchimento sedimentar entre 5000 a 8000 m (Hurtado et al., 2022). Localizada entre Peru, Brasil e Bolívia, é limitada pelos arcos de Fitzcarrald ao norte e o sistema de falhas subandino a oeste. Trata-se de uma bacia de antepaís, que começou a se desenvolver durante o Cretáceo tardio (Baby et al., 2018; Louterbach et al., 2018). Algumas das formações que compõem o seu preenchimento sedimentar são Nanea, Cabanillas, Ambo, Tarma, Copacabana, Nia, Cushabatay, Chonta, Vivian, Huchpayacu, Casa Blanca, Yahurango e Ipururo.

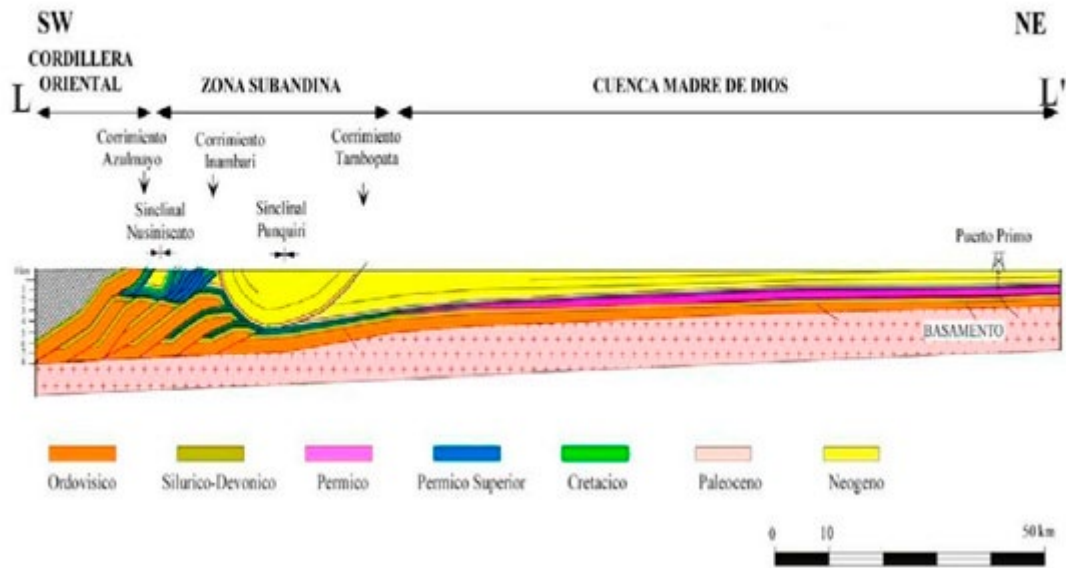


Figura 9. Seção geológica esquemática da bacia Madre de Dios. Tomado de Hermoza (2004).

De acordo com Rosario et al., 2016, uma comparação entre os preenchimentos sedimentares das diferentes bacias evidencia semelhanças entre os depósitos cretáceos e terciários (períodos Paleógeno e Neógeno dentro da era Cenozoica), postulando a hipótese de uma mesma história deposicional durante esses períodos, permitindo desenvolver a coluna geológica sintética apresentada e adaptada na Figura 10.

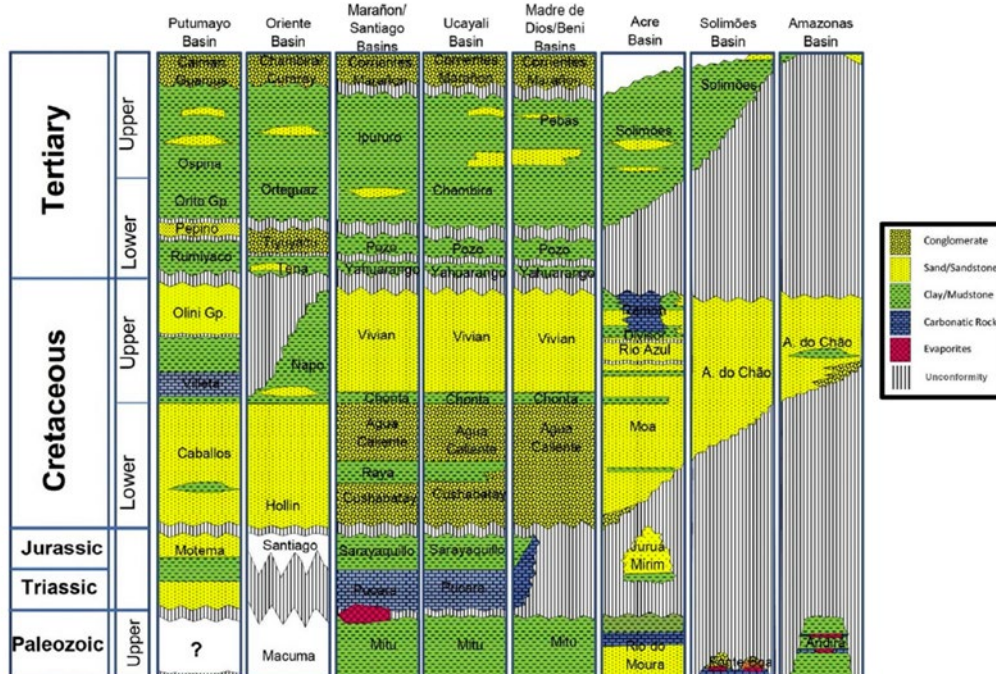


Figura 10. Proposta de coluna geológica sintética do sistema aquífero amazônico. Modificado e recentemente definido por Encarnacion, 2008; Rakhit, 2002; Cunha, 2007

No Brasil, nos estados do Amazonas, Mato Grosso e Pará, localiza-se a bacia sedimentar do Alto Tapajós, uma bacia intracratônica que repousa sobre o cráton amazônico na direção NO-SE, com uma

área de 97000 km², articulando-se ao NW com as bacias sedimentares de Amazonas e Solimões (Toczeck et al., 2019). Seu preenchimento é composto por seqüências sedimentares de idade mesozoica e paleozoica com espessura inferior a 1000 m. De acordo com Santiago et al. (1980), o preenchimento sedimentar corresponde a formações proterozoicas e paleozoicas com mais de 1700m de espessura, constituídas pelas formações Borrachudo, Capoeiras, São Bendito, Ipixuna e Navalha.

Ao norte da ARB, sobre o estado de Roraima, no Brasil, e parte da república da Guiana, encontra-se a bacia sedimentar de Tacutu. Formada por um sistema de grabens com direções NE-SW, com uma extensão de aproximadamente 300 km de comprimento e entre 30 a 40 km de largura (Eiras et al., 1990), desenvolvida sobre a área central do Escudo das Guianas. Possui uma área total de cerca de 12.500 km², dos quais dentro da porção brasileira conta com uma área de 4500 km². Seu preenchimento sedimentar e vulcânico tem mais de 7000 m de espessura (Eiras et al., 1990) e é composto pelas formações Apoteri, Manari, Rupununi, Pirara, Tacutu, Serra do Tucano e Boa Vista.

No centro-oeste do Brasil encontra-se a bacia sedimentar dos Parecis (Figura 11), que é uma bacia do tipo IF (Fratura interior continental) evoluindo para uma IS (Sineclise interior continental) (CPRM, 2006). Ela é subdividida em três domínios tectonossedimentares separados pelas elevações de Vilhena e Serra Formosa (Siqueira e Teixeira, 1993), os quais são a depressão tectônica que forma a sub-bacia de Rondônia, a região central composta pela Sub-bacia de Jurueña e a bacia tipo sineclise que forma a sub-bacia do Alto Xingu. Seus preenchimentos sedimentares são definidos pelas formações Cacoal, Furnas, Ponta Grossa, Pimenta Bueno, Fazenda da Casa Branca, Botucatu, Anari & Tapirapuã, Parecis e Araguaia.

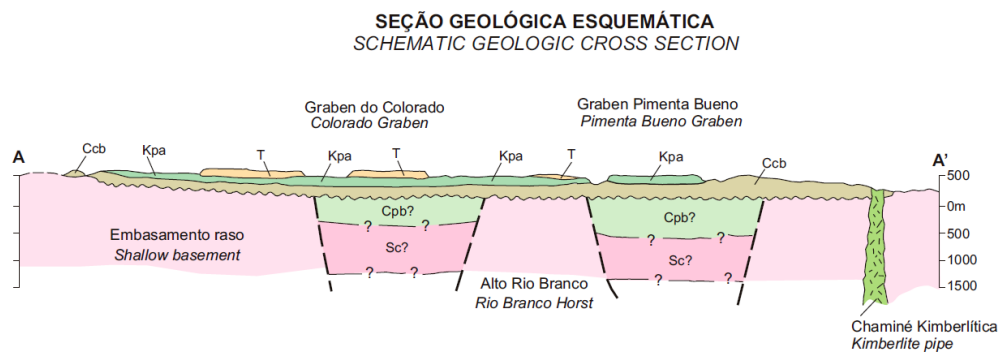


Figura 11. Seção geológica esquemática da bacia de Parecis. Tomado de CPRM, 2006.

2.2. Sistema Aquífero Amazonas (AAS)

O Sistema Aquífero Amazônico (AAS) é um aquífero transfronteiriço catalogado como um dos maiores do mundo. Foi definido em 2005 pela UNESCO/OEA e compreende as províncias hidrogeológicas do Orinoco e do Amazonas. Existe considerável incerteza sobre sua extensão, como evidenciado pelos mapas propostos por diferentes autores, como Puri e Aureli, 2009; Margat, 2007; Unesco, 2007; Rosario et al., 2016. Para este estudo, o AAS é definido com base nas unidades

sedimentares do Cenozoico ao Cretáceo do mapa geológico da América do Sul (CPRM, 2019), limitado a oeste pelo sistema de falhas subandinas, ao norte e ao sul pelas divisórias de água da AHB e pelos escudos das Guianas e do Brasil, abrangendo uma área total de $3.1 \times 10^6 \text{ km}^2$. Considera-se o AAS como um sistema de fluxo regional, heterogêneo, multicamada e multibacia, com uma única superfície potenciométrica livre conectada aos rios, assentada sobre rochas paleozoicas (derrames diabásicos ou formações sedimentares) e o embasamento cristalino (ANA, 2015). Desenvolve-se sobre o cráton amazônico e é composto pelas bacias sedimentares de Solimões, Acre, Madre de Dios, Beni, Ucayali, Marañón, Oriente e Putumayo, preenchidas com sedimentos consolidados e não consolidados com semelhanças estratigráficas do Cretáceo Inferior ao Terciário Tardio (Rosario et al., 2016). Apesar das diferenças na nomenclatura entre países, alguns dos principais aquíferos, principalmente do tipo não confinado, que compõem o AAS são o Alter do Chão, Içá, Tucunaré, Aluvial, bem como o aquífero Solimões-Repouso. A composição dos materiais varia entre argilas, limos, areias e cascalhos, com espessuras entre 500 a 1000 metros na bacia do Amazonas, 4000 metros na bacia de Marajó e entre 150 a 1000 metros nas sequências arenosas das bacias de Solimões e Acre. A ANA assume a hipótese de uma mesma superfície potenciométrica entre as unidades cenozoicas, cretáceas e paleozoicas, independentemente da natureza dos materiais, inter-relacionadas por variações laterais e/ou curtos circuitos de fluxo através de fraturas dentro da porção brasileira do AAS. As principais estruturas geológicas que limitam as bacias sedimentares são os arcos de Iquitos, Purus, Gurupá, Fitzcarrald, Vaupés e Michicola, além do cinturão de falhas subandinas. O AAS é dividido em duas regiões, o Western Amazon Aquifer System (WAAS) e o Eastern Amazon Aquifer System (EAAS), a partir do Arco de Purus. Ainda existem muitas incertezas em relação ao funcionamento regional do AAS; no entanto, a partir do agrupamento de diferentes unidades geológicas, surgiram a definição de subsistemas como o Sistema Aquífero Tikuna (Rosario, 2016) ou o Sistema Aquífero Grande Amazônia (SAGA) (Abreu, 2013).

2.2.1. Aquífero Alter do Chão

Aquífero de tipo livre ou não confinado, localizado nos estados do Pará, Amazonas e Amapá do Brasil, com uma área de aproximadamente 290.500 km^2 . De idade cenozoica e constituído por intercalações de arenitos, argilas, limolitas e alguns conglomerados avermelhados. Foram registradas espessuras de 1266m em perfurações, mas sua espessura média varia de 200 a 400m (ANA, 2015). Aquífero de grande importância para o abastecimento de cidades da região, como Manaus ou Santarém.

Investigações de Tancredi (1996), a partir de 16 poços tubulares localizados em Santarém, registraram parâmetros hidrodinâmicos com valores médios de Transmissividade de $3.33 \times 10^{-2} \text{ m}^2/\text{s}$, Condutividade hidráulica de $9.78 \times 10^{-5} \text{ m/s}$ e Coeficiente de armazenamento de 3.63×10^{-4} . Obtendo faixas de transmissividades entre 132 a 790 m^2/dia e condutividades hidráulicas entre 4.6 a 18.8 m/dia para a região, variações que são explicadas pelos diferentes espessores dos materiais geológicos nos locais dos testes de bombeamento. Foram identificadas porções confinadas do aquífero com espessuras

de 430m, intercaladas com aquícludes e aquíardos. Foi realizada uma estimativa das reservas de água subterrânea de $226 \times 10^6 \text{ m}^3$ para uso sazonal e de $86550 \times 10^6 \text{ m}^3$ para uso permanente.

De acordo com Aguiar & Mourão (2012), na cidade de Manaus, o aquífero apresenta uma espessura média de aproximadamente 160m, sobrepondo-se às evaporitas e calcários da Formação Nova Olinda. Uma análise da transmissividade mostra que os maiores valores estão nos setores SO, N e E da cidade, e os menores na zona central. A média dos principais parâmetros hidrodinâmicos é de transmissividade de $4.7 \text{ m}^2/\text{h}$, condutividade hidráulica de $2.9 \times 10^{-2} \text{ m/h}$, porosidade efetiva de 18%, um gradiente hidráulico da ordem de 0.002 e uma velocidade em torno de 2cm/dia. O estudo de França et al., 2005, por meio de testes de aquífero realizados em 36 poços, obteve valores médios de transmissividade de $1.48 \times 10^{-3} \text{ m}^2/\text{s}$ e condutividade hidráulica de $1.54 \times 10^{-5} \text{ m/s}$.

Tabela 1. Resumo dos parâmetros hidráulicos do Aquífero Alter do Chão.

Cidade	Autor	Transmissividade (m^2/s)	Condutividade Hidráulica (m/s)	Coefficiente de Armazenamento
Santarem	Tancredi, 1996	3.33×10^{-2}	9.78×10^{-5}	3.63×10^{-4}
Manaos	Aguiar & Mourão, 2012	1.30×10^{-3}	8.05×10^{-6}	-
Manaos	França et al 2005	1.48×10^{-3}	1.54×10^{-5}	-

2.2.2. Aquífero Içá

Aquífero de tipo livre a semi-confinado presente nos estados do Amazonas, Roraima, Rondônia e Acre no Brasil, ocupando uma área de aproximadamente 932700 km^2 . De idade pleistocênica, composta principalmente por camadas de arenito fino a médio, algumas limo-argilosas, pouco consolidadas, com estratificação cruzada de ambientes fluviais. Presença de alguns níveis ricos em matéria orgânica e/ou ferro em direção aos níveis basais. Foram registradas espessuras de 143m (Maia & Marmos, 2010), no entanto, sua espessura média é inferior a 100m. Fonte principal de abastecimento de várias cidades na região, como Coari, Fonte Boa e Barcelos.

Estimativas de parâmetros hidrodinâmicos obtidas por Maia & Marmos, 2010, na cidade de Tefé, obtiveram valores médios de transmissividade de $6 \times 10^{-4} \text{ m}^2/\text{s}$ e condutividade hidráulica de $1.7 \times 10^{-5} \text{ m/s}$. Testes de bombeamento realizados em cidades vizinhas a Caracarái pela CPRM, 2002, obtiveram valores de transmissividade da ordem de $4.5 \times 10^{-2} \text{ m}^2/\text{s}$ e condutividade hidráulica média de $3 \times 10^{-3} \text{ m/s}$.

Tabela 2. Resumo dos parâmetros hidráulicos do Aquífero Içá

Cidade	Autor	Transmissividade (m^2/s)	Condutividade Hidráulica (m/s)	Coefficiente de Armazenamento
Tefé	Maia & Marmos, 2010	6×10^{-4}	1.7×10^{-5}	-
Cerca de Caracarái	CPRM,2002	4.5×10^{-2}	3×10^{-3}	-

2.2.3. Acuitardo Solimoes-Repouso

Acuitardo aflorante nos estados de Acre, Amazonas e parte de Rondônia, com uma área de aproximadamente 417.000 km². Essa unidade geológica tem continuidade lateral nas bacias sedimentares de Marañon e Madre de Dios, onde são denominadas como Formação Pebas, Ipunuro e Nauta, e na Colômbia na bacia do Putumayo como Terciario Amazonico e/ou capas de La Tagua (Hoorn, 1993). As formações Solimões e Repouso são compostas principalmente por argilas macias, laminadas ou massivas, com intercalações de lentes finas de lignito. Na parte superior ocorrem intercalações de arenitos finos a grossos brancos. De idade cenozoica e com espessuras máximas de 760m (Caputo et al., 1972), mas com espessura média de 300m (Galvão, 2011).

Testes de aquífero em um poço em Urucu realizados por Souza, 2009, obtiveram transmissividade de 3.7×10^{-3} m²/s, condutividade hidráulica de 7.4×10^{-5} m/s e coeficientes de armazenamento de 8.0×10^{-4} .

Tabela 3. Resumo dos parâmetros hidráulicos do Acuitardo Solimoes-Repouso

Cidade	Autor	Transmisividade (m ² /s)	Conductividade Hidráulica (m/s)	Coefficiente de Armazenamento
Urucu	Souza, 2009	3.7×10^{-3}	7.4×10^{-5}	8.00×10^{-4}

2.2.4. Aquífero Aluvial

Aquífero formado pela presença de depósitos de planícies aluviais. Segundo a CPRM, 2006, os principais rios que integram as bacias Solimões e Amazonas apresentam áreas expressivas com planícies aluviais, chegando a alcançar dezenas de quilômetros de comprimento e largura. As bacias dos rios Madeira, Purus, Solimões, Japurá, Juruá, Jutai, Amazonas, Uatumã e Jatapu possuem planícies sedimentares com canais fluviais meândricos, lagos e meandros abandonados que caracterizam esses depósitos. As planícies são compostas por sedimentos arenosos a argilosos inconsolidados a semi-consolidados, associados a níveis de cascalhos e matéria orgânica. Quanto aos materiais inconsolidados, eles são compostos principalmente de areia com granulometria variável, geralmente mostrando diminuição ascendente dos grãos. Os materiais semi-consolidados consistem em camadas argilosas, maciças, com alguns níveis arenosos de espessura variável (ANA, 2015). De acordo com Pereira & Cajazeiras, 2012, este é um aquífero contínuo, poroso e de baixa permeabilidade, com espessuras inferiores a 40m.

Testes realizados pela CPRM, 1998, obtiveram em Rondônia transmissividades de 6.6×10^{-6} m²/s a 4.27×10^{-3} m²/s, com um valor médio de 6.03×10^{-4} m²/s.

Tabela 4. Resumo dos parâmetros hidráulicos do Aquífero Aluvial.

Cidade	Autor	Transmissividade (m ² /s)	Condutividade Hidráulica (m/s)	Coefficiente de Armazenamento
Rondônia	CPRM, 1998	6.03 x 10 ⁻⁴	-	-

2.3. Sensoriamento remoto em Hidrogeologia

A água subterrânea é difícil de quantificar com precisão, pois está abaixo da superfície terrestre. Com a evolução dos sensores remotos (SR), foram desenvolvidas múltiplas técnicas para identificar diretamente ou indiretamente as mudanças nas águas subterrâneas. Toda técnica baseada em sensores remotos deve ser analisada em conjunto com observações terrestres, como medições de carga hidráulica em poços de monitoramento. Além disso, são integrados monitoramento do Sistema Satelital de Navegação Global e modelagem numérica para complementar as abordagens das técnicas. Algumas das técnicas com melhores resultados são: 1) inferência de mudanças na água subsuperficial a partir de mudanças de massa usando medições gravitacionais, 2) análise de mudanças na altura da superfície terrestre usando Radar interferométrico de abertura sintética, 3) sistemas eletromagnéticos aerotransportados e 4) altimetria satelital.

As observações satelitais e aéreas revolucionaram a compreensão da hidrologia e da disponibilidade de água em escala regional e global de maneiras que não seriam possíveis com observações in situ relativamente escassas. Os satélites e sistemas aerotransportados podem fornecer tanto a cobertura espacial quanto as variações temporais em grandes regiões de variáveis, como mudanças no armazenamento de água subterrânea (Δ GWS). No entanto, ainda existem desafios no uso integrado dos dados de SR em relação às observações hidrológicas, redução de escala (espacial e temporal) dos conjuntos de dados e sua integração com a modelagem numérica de águas subterrâneas.

2.3.1. Medições gravimétricas da água subterrânea

As missões GRACE/GRACE-FO fornecem estimativas das variações temporais do campo gravitacional terrestre. Seu funcionamento consiste em um par de satélites idênticos voando em formação líder-seguidor, separados por cerca de 200 km. Equipados com um instrumento de alcance de micro-ondas de banda K/Ka que mede a distância entre os satélites com precisão micrométrica enquanto orbitam a Terra. As mudanças na aceleração devido à atração das anomalias gravitacionais locais resultam em alterações na distância entre os satélites, rastreando continuamente o alcance relativo entre eles (Figura 12). O GRACE-FO, além de dar continuidade à missão GRACE, adiciona um interferômetro a laser que mede as mudanças de alcance aproximadamente 100 vezes mais preciso que o GRACE. As missões fornecem estimativas mensais do campo geopotencial variável no tempo da Terra, que é então convertido em altura equivalente de água usando os métodos de Wahr et al., 1998 ou Rowlands et al., 2005. Estima-se uma precisão de aproximadamente 2 cm para uma região de 300 km

(Tapley et al., 2019). É importante notar que as medições podem ser interferidas por variações de massa da Terra sólida, e, portanto, correções podem ser aplicadas usando modelos (Han et al., 2008; Peltier et al., 2018).

As anomalias de armazenamento total ou terrestre de água (TWS) contêm informações integradas das mudanças de armazenamento de água de uma região, incluindo mudanças na água superficial (lagos e rios), a água contida na vegetação (água do dossel), umidade do solo, neve e água subterrânea. Ao estimar as mudanças de armazenamento de água subterrânea (ΔGWS), os erros dos modelos de superfície de onde provêm as outras variáveis hidrológicas se acumulam na água subterrânea residual calculada, de modo que uma contabilidade cuidadosa de erros deve ser realizada (Rodell et al., 2007). A incerteza nos resíduos de água subterrânea derivados do ΔTWS de mês a mês costumam ser substanciais, mas a abordagem pode ser útil na identificação de tendências de longo prazo (Rodell et al., 2009).

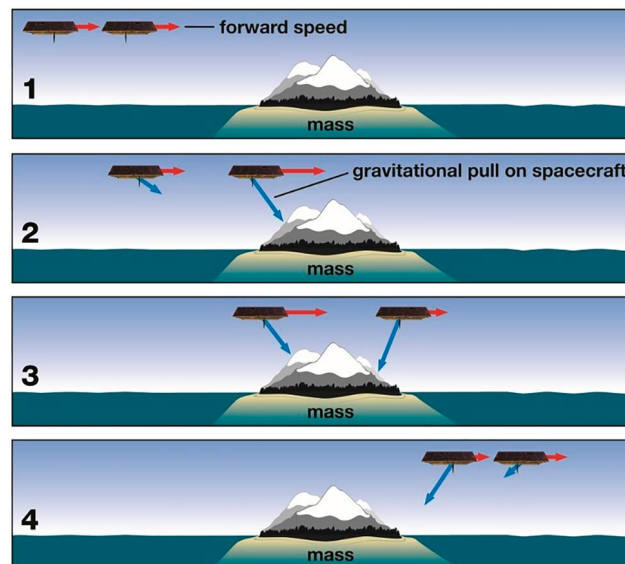


Figura 12. Descrição de como o GRACE/GRACE-FO mede a gravidade. Extraído da NASA.

Algumas de suas principais limitações estão relacionadas a erros de processamento de dados, limitações na altitude orbital e na geometria de observação. Durante o processamento dos dados para resolver as variações temporais da gravidade, as medições das mudanças terrestres são acumuladas mensalmente, resultando na criação de medições sintéticas de alcance entre os satélites. Esse sistema terrestre dinâmico criado inclui modelos de variações de massa de alta frequência, como as marés oceânicas, variações atmosféricas de alta frequência e variações de massa oceânica. Uma vez que as medições são acumuladas ao longo de 30 dias para obter uma solução estável, o GRACE/GRACE-FO não consegue captar variações de massa de alta frequência. Portanto, é necessário aplicar um modelo para considerar o efeito dessas variações na órbita dos satélites. Os erros associados a esses modelos têm um impacto maior na recuperação da gravidade do que o previsto pelos sistemas de medição a bordo dos satélites

A altitude dos satélites também é uma limitação, pois eles foram colocados em uma órbita com altitude aproximada de 500 km, que naturalmente decai devido ao arrasto atmosférico. Esse decaimento aumenta a sensibilidade às perturbações do campo de gravidade e permitiria uma melhor resolução espacial, porém, aumenta a resistência atmosférica e limita a vida útil da missão. A última limitação é a geometria de observação. Como os satélites estão em órbita polar para obter cobertura global, eles fazem medições na direção de sua trajetória e apresentam limitações nas variações E-W. Isso se manifesta em erros ao longo das faixas longitudinais nas soluções de gravidade. Uma ampla variedade de filtros de processamento foi desenvolvida (Luthcke et al., 2013; Save et al., 2016; Swenson & Wahr, 2002; Watkins et al., 2015), no entanto, qualquer abordagem de remoção das faixas corre o risco de também eliminar o sinal geofísico real. Recomenda-se aplicar conjuntamente os chamados "fatores de escala" para restaurar o sinal geofísico perdido durante o filtrado (Landerer e Swenson, 2012; Wiese et al., 2016).

2.3.2. Medições da deformação da superfície

A análise de séries temporais de subsidência e elevação da superfície em regiões de aquíferos explorados com sedimentos compressíveis tem sido utilizada para estimar as mudanças no armazenamento de águas subterrâneas (tanto devido à recarga quanto à extração) (Adams, et al., 2022). Em áreas de geologia suscetível, a extração de águas subterrâneas pode levar à compactação de certas camadas, resultando em afundamentos elásticos ou inelásticos na superfície (Amelung et al., 1999; Argus et al., 2005; Bawden et al., 2001; Chaussard et al., 2013, 2014; Smith et al., 2017). A deformação é causada por uma mudança na pressão dos poros nos sedimentos do subsolo, o que provoca uma mudança na tensão efetiva de sentido oposto (Terzaghi, 1925). Mudanças na tensão efetiva fazem com que a matriz do aquífero se consolide ou se expanda, respectivamente. O aumento dos níveis de água subterrânea gera uma diminuição da tensão efetiva e a expansão do material do aquífero, enquanto a diminuição dos níveis de água subterrânea provoca o oposto.

A deformabilidade por unidade de mudança do nível de água subterrânea depende da compressibilidade e história de tensões da camada. Materiais não consolidados de grãos finos, como as argilas, são mais compressíveis e geralmente deformam mais do que outros materiais. Os fatores-chave na geração de subsidências devido à extração de águas subterrâneas são: a) uma queda historicamente alta nos níveis de água subterrânea e b) camadas amplas, não consolidadas e compressíveis de materiais argilosos. Devido as variações espaciais significativas nas deformações, é difícil relacionar observações geodésicas com as mudanças no armazenamento de água subterrânea. O Sistema Global de Navegação por Satélite pode fornecer medições contínuas ao longo do tempo nas localizações das estações (Bürgmann et al., 2006; Dehghan-Soraki et al., 2015; Ferretti et al., 2007). A interferometria radar de abertura sintética e a altimetria radar podem fornecer um mapa contínuo da deformação. Portanto, a deformação do terreno pode servir como um indicador indireto da mudança volumétrica das águas subterrâneas. No entanto, como a resposta da superfície à mudança das águas subterrâneas pode ser não

linear e espacialmente heterogênea, essas abordagens são melhor utilizadas quando complementadas com o conhecimento das propriedades geológicas e hidrológicas gerais do aquífero (Adams, et al., 2022).

Interferometria radar de abertura sintética

O interferômetro de abertura sintética por radar (InSAR) é uma ferramenta geodésica que possibilita a medição de deformações e mudanças na superfície terrestre, sendo particularmente útil em áreas onde as medições locais *in situ* não são viáveis, difíceis ou impraticáveis. Um satélite InSAR que passa sobre um local antes e depois de um movimento na superfície, como subsidência devido à retirada de fluidos do subsolo, exatamente no mesmo ponto do espaço inercial, mede como o terreno se desloca entre passagens, através de um interferograma de radar. As imagens SAR são compostas por pixels que representam amplitude e fase de retroespalhamento coerente. O interferograma consiste no padrão de interferência de fase da onda entre duas imagens SAR, sendo sensível ao movimento do terreno ao longo da linha de visão do radar. Uma imagem InSAR da diferença de fase ponto a ponto da onda na superfície é utilizada para criar um mapa do movimento da superfície ao longo do tempo. As imagens de radar podem realizar observações atravessando nuvens e sem a necessidade de luz solar, proporcionando resultados de mudanças subcentimétricas. Para a estimativa das águas subterrâneas, onde as mudanças ocorrem lentamente, as medições de séries temporais podem ser essenciais, especialmente para sistemas com comprimentos de onda curtos, como a banda X (3 cm de comprimento de onda), onde mudanças em outras propriedades da superfície, como vegetação ou umidade do solo, reduzem a coerência e introduzem ruído, tornando necessário empilhar muitas observações ao longo do tempo para rastrear o sinal (Adams et al., 2022).

O padrão espacial e o alcance das subsidências superficiais dependem da heterogeneidade geológica, elasticidade da matriz, tensão efetiva e limite de carga crítica (Castellazzi et al., 2016; Hoffmann et al., 2003). Coberturas vegetais ou mudanças na superfície devido à atividade antrópica, como agricultura, podem causar descorrelação do sinal (Castellazzi et al., 2016), um aspecto que pode ser resolvido com imagens de radar com comprimento de onda menos sensível a tais mudanças. Outros fatores, como vapor d'água na atmosfera, podem ser limitantes, pois geram atrasos na trajetória do sinal, causando erros. É necessário usar diversas amostras ao longo do tempo para médias fontes de ruído não correlacionadas e utilizar apenas aquelas relacionadas às deformações do aquífero mais altamente correlacionadas com tendências seculares ou sazonais (Li et al., 2009, 2012).

As principais limitações no uso do InSAR em problemas de águas subterrâneas residem na disponibilidade de dados, repetição de cobertura e amostragem. As mudanças nos aquíferos ocorrem em gradientes espaciais baixos e escalas temporais longas; portanto, produtos InSAR obtidos com imagens de resolução espacial de centenas de metros, amostradas semanalmente por um longo período, podem observar deformações na superfície relacionadas a alterações no armazenamento de águas subterrâneas.

Nesses casos, a missão Sentinel-1 da Agência Espacial Europeia (ESA) é uma boa opção, pois fornece cobertura espacial e amostragem adequadas (Adams et al., 2022).

Altimetria (radar)

Ao contrário do InSAR, que utiliza interferometria entre as formas de onda emitidas e refletidas, a altimetria obtida por radar utiliza o eco nadir dos pulsos de radar para medir o retardo do tempo de viagem bidirecional e estimar a elevação da superfície (Hwang et al., 2016; Kuo et al., 2015). A altimetria de radar não é tão restrita pela necessidade de que as características do terreno estejam altamente correlacionadas temporalmente, como é o caso do InSAR. Isso permite oferecer registros de subsidência a longo prazo com uma maior flexibilidade espacial, pois não está tão limitada pelas condições específicas do terreno. Atualmente, a altimetria é capaz de obter subsidências com escalas verticais maiores que 1 cm/ano. Sua utilidade na medição de subsidências é restrita a terrenos planos, com vegetação baixa e com um mínimo de edifícios. São especialmente úteis missões com longo período de repetição e pequeno espaçamento entre trajetórias (Adams, et al., 2022).

Outra aplicação proposta por Pfeffer, et al (2014), na região central da Amazônia, consistiu em implementar pontos ou estações altimétricas nos principais cursos dos rios alimentados intrinsecamente por águas subterrâneas de aquíferos não confinados. Com essas alturas, foram gerados mapas de níveis freáticos interpolados com resoluções de 50 a 100 km, que foram coerentes com dados de poços de águas subterrâneas, permitindo obter informações espaço-temporais nas mudanças de armazenamento de água superficial e subterrânea.

LIIDAR

O ICESat (2003-2006) e o ICESat-2 (2018-presente) da NASA são espaçonaves de observação da Terra com altímetros a laser para medir a elevação da camada de gelo, a espessura do gelo marinho, bem como a altura da vegetação, batimetria e topografia terrestre (Abdalati et al., 2010). Dado que ambos os ICESats coletam amostras de pontos ao longo de suas trajetórias na Terra, em vez de varrer seus feixes sobre um campo de visão, o mapa da superfície é sintetizado a partir de múltiplas passagens da mesma área, com trajetórias na Terra ligeiramente diferentes (Adams, et al., 2022). Em 2015, foi obtida uma aplicação bem-sucedida do ICESat na planície norte da China para quantificar as subsidências do terreno devido à exploração das águas subterrâneas e medidas obtidas a partir do GRACE (An, 2015). Os efeitos da vegetação sobre o sinal do ICESat foram avaliados usando o Índice de Vegetação por Diferença Normalizada (NDVI). Assim como outras pesquisas, é comum relacionar as medições de deformações superficiais junto com os dados do GRACE para obter uma visão geral das águas subterrâneas e a resposta dos aquíferos.

2.3.3. Sistemas eletromagnéticos aéreos (AEM)

Os sistemas de AEM são métodos rápidos e de baixo custo para explorar sistemas aquíferos em grande escala, desde rasos (1-3 m) até profundos (300-400 m), utilizando a resposta medida de uma

fonte eletromagnética ativa para estimar a condutividade elétrica do subsolo (Paine e Minty, 2005; Siemon et al., 2009). Devido à sua dependência da condutividade elétrica, os AEM são sensíveis às variações no nível freático, às propriedades geológicas do aquífero (especialmente as interfaces entre materiais resistentes, como leitos rochosos, materiais moderadamente resistentes, como areias, e materiais condutivos, como argilas), e às distribuições de salinidade (Abraham & Cannia, 2011; Dewar & Knight, 2020; Kirsch, 2006; Knight et al., 2018; O'Connell et al., 2020).

Esses sistemas funcionam gerando correntes elétricas por meio de bobinas; à medida que essas correntes mudam, seja desligando rapidamente ou alternando, campos magnéticos são produzidos e penetram no solo, gerando correntes parasitas no subsolo. A partir dessas correntes parasitas, é gerado um campo eletromagnético secundário, cuja intensidade depende da distribuição da condutividade do subsolo e que é detectado pelas bobinas receptoras (Palacky, 1993). A partir das medições, o sinal passa por um processamento padrão, uma série de correções (primárias, de altura, geométricas e eliminação de interferências elétricas) para realizar um processo de inversão e estimar a condutividade do subsolo em altas resoluções espaciais e moderadas resoluções verticais. A partir das estimativas de condutividade do subsolo, é possível gerar modelos 3D, que, integrados aos modelos litológicos obtidos por sondagens *in situ*, fornecem informações sobre as propriedades geológicas e hidrogeológicas.

2.3.4. Umidade do solo e evapotranspiração

A avaliação da umidade do solo, da evaporação ou da vegetação na superfície também pode fornecer informações sobre o uso das águas subterrâneas (Chen e Hu, 2004; Jackson, 2002; Maxwell et al., 2007), uma vez que as águas subterrâneas próximas à superfície afetarão direta ou indiretamente a vegetação, especialmente em regiões áridas, de alta infiltração (por exemplo, sistemas cársticos) e de irrigação (Hartmann et al., 2020). As bandas visíveis, infravermelhas próximas e térmicas dos satélites Landsat e Sentinel-2 permitiram a implementação de metodologias para inferir a evapotranspiração e a umidade do solo a partir do espaço. Geralmente, a comparação de regiões agrícolas irrigadas com regiões adjacentes não irrigadas possibilitou inferências sobre o aquecimento/resfriamento devido à umidade e associado à evapotranspiração (ET). Com relação à inferência sobre as águas subterrâneas, seria necessário um conhecimento avançado sobre o tipo de vegetação e seus respectivos índices de ET, bem como uma medição e contabilização adequadas de todas as fontes de água.

2.4. Downscaling GRACE/GRACE-FO

Em sensoriamento remoto, *downscaling* refere-se ao aumento da resolução espacial ou diminuição do tamanho dos pixels de uma imagem, e *upscaling* refere-se à diminuição ou engrossamento

da resolução espacial (Atkinson e Tate, 2000). Nos últimos anos, têm aumentado os métodos e estratégias para o *downscaling* do TWS de GRACE e GRACE-FO, bem como, respectivamente, do Δ GWS (Ali et al., 2021; Agarwal et al., 2023; Chen et al., 2019; Khorrami et al., 2023; Seyoum et al.,

2019; Pulla et al., 2023 and Wang et al., 2021;). As abordagens de *downscaling* podem ser divididas em duas categorias: 1) *downscaling* baseado em modelos ou "*downscaling* dinâmico" e *downscaling* baseado em dados ou "*downscaling* estatístico" (Pascal, et al., 2022). A primeira abordagem consiste em assimilar os dados de ΔTWS de GRACE/GRACE-FO em modelos hidrológicos ou de superfície terrestre baseados fisicamente para obter ΔGWS nas resoluções temporais e espaciais do modelo, que geralmente são superiores às de GRACE (Giroto et al., 2016; Houborg et al., 2012; Nie et al., 2019; Schumacher et al., 2018; Tian et al., 2017; Zaitchik et al., 2008). As limitações desta abordagem são: (a) a discrepância entre GRACE/GRACE-FO e as resoluções dos dados de entrada do modelo e (b) as limitações inerentes aos modelos: hipóteses e parametrização do modelo, a incerteza do forçamento meteorológico e, em particular, a falta de representação de processos antropogênicos como a irrigação de cultivos (Long et al., 2013). A abordagem baseada em dados consiste em (a) derivar um modelo estatístico de ΔTWS a partir de dados auxiliares disponíveis em alta resolução (HR), (b) calibrá-lo em baixa resolução (LR), (c) aplicá-lo em HR e (d) eliminar a contribuição das reservas de água da superfície e da umidade do solo para isolar o ΔGWS (Pascal, et al., 2022). Essa abordagem é baseada na hipótese de que os processos hidrológicos e físicos que vinculam essas variáveis são idênticos em todas as resoluções (Ali et al., 2021; Jyolsna et al., 2021; Karunakalage et al., 2021; Sahour et al., 2020; Seyoum e Milewski, 2017; Vishwakarma et al., 2021; J. Zhang et al., 2021). Nesta abordagem, ganharam popularidade a implementação de algoritmos de inteligência artificial (IA) dentro dos campos do Machine Learning (ML) e do Deep Learning (DL). Estes métodos aproveitam o poder dos algoritmos de aprendizagem automática e das redes neurais para aprender relações não lineares entre os dados de gravidade em grande escala e outras variáveis de entrada, com o objetivo de fornecer estimativas mais detalhadas em escalas menores. Alguns dos algoritmos de machine learning mais frequentes são Gradient Tree Boosting (GTB), Random Forest (RF), Extra Tree (ET), Adaptive Boosting (AB), K-Nearest Neighbors regressor (KNN), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost) e Gradient Boosting (GB) (Seyoum, et al., 2019; Chen, et al., 2021; Agarwal, et al., 2022; Yin, et al., 2022; Pulla, et al., 2023; Tao, et al., 2023; Zhang, et al., 2023) e de deep learning como Multi-layer perceptron (ANN) (Foroumandi, et al., 2022; Pulla, et al., 2023). Em ambas as abordagens, a validação do ΔGWS escalado consiste em comparar as séries temporais com medições *in situ* dos níveis de água subterrânea (GWL) convertidos em anomalias, às vezes com o uso do rendimento específico (Sy) representativo da região. Essa comparação entre as séries temporais de ΔGWS e ΔWL é realizada com métricas como o coeficiente de determinação (R^2) ou o coeficiente de eficiência de Nash-Sutcliffe (NSE), o coeficiente de correlação de Pearson (R), o erro quadrático médio (RMSE) ou o erro absoluto médio (MAE) (Ali et al., 2021; Jyolsna et al., 2021; Karunakalage et al., 2021; Sahour et al., 2020; Yin et al., 2018; J. Zhang et al., 2021; G. Zhang et al., 2021; Zuo et al., 2021).

Qualquer método de *downscaling* pode melhorar ou degradar a precisão dos dados de GRACE/GRACE-FO para a resolução desejada, dependendo de: (a) a variabilidade espacial sub-pixel de $\Delta TWS/\Delta GWS$ e (b) as incertezas nos parâmetros de entrada do modelo e do forçamento (Pascal, et al., 2022). O *downscaling* é um problema de regressão, para o qual é necessário ter um bom número de dados de boa qualidade para treinar os modelos.

2.4.1. Random Forest

Em 2001, Breiman propôs o modelo Random Forest (RF) como um novo algoritmo de Machine Learning (ML) integrado, consistente em múltiplas combinações de Árvores de Classificação e Regressão (CART) (Breiman, 2001). Este algoritmo pode ser utilizado tanto para problemas de classificação como de regressão. Dentro do ML, é um algoritmo de aprendizado supervisionado. É um método ensemble, ou seja, combina a saída de várias árvores de decisão para obter um resultado médio. O algoritmo gera subconjuntos aleatórios de características, o que reduz a correlação entre as árvores de decisão e melhora a precisão do modelo. Três de seus principais hiperparâmetros são: número de estimadores, características máximas e divisão mínima de amostras. Este algoritmo aplica o Bootstrapping, uma técnica de reamostragem iterativa utilizada para extrair aleatoriamente várias amostras para o treinamento, aplicada nos dados de treinamento, e as amostras restantes são agrupadas como dados fora da amostra e são usadas para testar a precisão do modelo. É um algoritmo amplamente utilizado na redução de escala de GRACE/GRACE FO (Agarwal et al., 2023; Chen et al., 2019; Seyoum et al., 2019). Algumas de suas vantagens são a baixa probabilidade de overfitting, detecção de interações e importância das variáveis, incorporação de métodos para estimar valores ausentes e estabilidade ao usar múltiplas amostras, prevalecendo a média da votação das árvores de decisão. Algumas de suas desvantagens são os elevados tempos de treinamento e o baixo desempenho com conjuntos de dados pequenos.

2.4.2. Adaptive Boosting

Em 1995, Freund e Schapire propuseram o modelo Adaptive Boosting ou AdaBoost (AB) como um método ensemble, cujo funcionamento consiste em criar vários preditores sequenciais simples a partir de um algoritmo comum (Freund & Schapire, 1995). Esses preditores são treinados com o mesmo conjunto de dados, recebendo pesos derivados dos erros cometidos, melhorando sequencialmente. No final, quando todos os preditores foram treinados, o AdaBoost realiza uma predição final, estimando um voto ponderado ou soma para cada preditor de acordo com um peso relativo (Elith et al., 2008). Este algoritmo pode ser utilizado tanto para problemas de classificação como de regressão. Alguns de seus principais hiperparâmetros são: *number of estimators*, *learning rate* e *the regression loss function*. É um algoritmo simples, fácil de programar, flexível, com alta capacidade de generalização, baixa taxa de erro e alta precisão. No entanto, a influência de alguns preditores fracos nos valores atípicos pode levar ao overfitting, tornando-o vulnerável ao ruído.

3. MATERIAIS E MÉTODOS

3.1. Materiais

3.1.1. Sensoramento remoto

GRACE e GRACE FO

O *Gravity Recovery and Climate Experiment* (GRACE) e seu sucessor *Follow-On* (GRACE FO) são um produto da missão de satélite da cooperação entre as agências espaciais dos Estados Unidos o National Aeronautics and Space Administration (NASA) e Alemanha o German Center for Aerospace Travel (DLR). As operações do segmento terrestre do GRACE são atualmente cofinanciadas pelo Centro Alemão de Pesquisa em Geociências (GFZ) e pela Agência Espacial Europeia (ESA). O jogo de satélite gêmeos GRACE foi lançado em 17 de março de 2002 e funcionou até 12 de outubro de 2017, sucedido em 22 de maio de 2018 pelo GRACE-FO. A missão mede as variações da gravidade na superfície da Terra, produzindo um novo mapa do campo gravitacional a cada 30 dias e mostrando como a gravidade do planeta difere de um lugar para outro de um período de tempo para outro. Existem três instituições que fornecem dados GRACE / GRACE-FO processados através de três algoritmos, que são CSR (U. Texas / Center for Space Research), GFZ (GeoForschungsZentrum Potsdam) e JPL (Jet Propulsion Laboratory) (Landerer, FW e Cooley, 2019). A diferença entre os algoritmos se deve às opções na combinação de parâmetros e estratégias na solução para derivar as variações do campo gravitacional, de acordo com o equipamento de processamento.

As informações de GRACE / GRACE-FO estão dividida em cinco níveis diferentes: Nível-0, Nível-1A, Nível-1B, Nível-2 e Nível-3. Os dados do Nível-0 correspondem aos dados RAW (brutos) gerados durante um estágio de recebimento e armazenamento. O Nível-1A corresponde à transformação de dados binários do Nível-0 para unidades de engenharia. Os dados produzidos para o nível 1B são o resultado de um processo possivelmente destrutivo dos dados do nível 1A, obtendo rotulagem e redução temporária. Os dados de nível 2 incluem campo gravitacional estático e de curto prazo (mensal e semanal). O processamento para atingir o Nível 2 foi desenvolvido de forma independente por cada instituto. Por fim, os dados do Nível 3 passaram por uma série de correções e tratamentos, como a Substituição de C20 e C30, Correções de 1 grau e Correção GIA (ajuste isostático glacial) (Save et al., 2016). Nessa investigação, são utilizados os produtos de nível 3 do CSR e JPL, nos quais uma série de processos foi aplicada, como alisamento gaussiano, filtro de eliminação, ajuste isostático glacial (GIA) e as anomalias deduzidas da linha de base média de 2004 a 2009.

GLDAS

O *Global Land Data Assimilation System* (GLDAS) é uma iniciativa colaborativa fornecida pelo *Goddard Space Flight Center* (GSFC) da NASA. Ele combina observações da Terra de diversas fontes, como satélites e estações meteorológicas, com modelos de computador. O objetivo é fornecer

estimativas globais da umidade do solo, temperatura da superfície, precipitação e outras variáveis ambientais. O GLDAS inclui três modelos globais de processos de superfície terrestre (CLM, Mosaic, NOAH) e um modelo hidrológico (VIC), dividido em duas versões, GLDAS1 e GLDAS-2. O GLDAS-2 tem três componentes GLDAS 2.0, 2.1 e 2.2: o primeiro é completamente forçado com os dados de entrada do forçamento meteorológico de Princeton entre 1948 e 2014; o GLDAS-2.1 é forçado com uma combinação de dados de modelagem e observação de 2000 até o presente; e o GLDAS-2.2 usa assimilação de dados (DA), incluindo assimilação de dados do GRACE/GRACE-FO de 2003 até o presente (Rui & Beaudoin, 2020). O GLDAS fornece várias variáveis hidroclimatológicas em resoluções espaciais de $0.25^\circ \times 0.25^\circ$ e $1^\circ \times 1^\circ$ e temporais de 3 horas, diárias e mensais. Nessa investigação, é utilizado o produto GLDAS-2.1 NOAH com resolução espacial de $0.25^\circ \times 0.25^\circ$ e temporal mensal.

CHIRPS

O *Climate Hazards Group InfraRed Precipitation with Station data* (CHIRPS) é um produto da colaboração entre o *United States Geological Survey* (USGS) e o *Earth Resources Observation and Science* (EROS). O CHIRPS fornece dados diários, mensais (1, 2 ou 3 meses) e anuais de precipitação quase globais, com resolução espacial de 0.05° e 0.25° . O conjunto de dados é baseado em abordagens anteriores de técnicas de interpolação e estimativas de precipitação de alta resolução e longo período de registro baseadas em observações infravermelhas de duração de nuvens frias (CCD) (C. Funk et al., 2015). As principais fontes de dados utilizadas em sua criação são: 1) climatologia de precipitação mensal, CHPClim; 2) observações infravermelhas térmicas (IR) geoestacionárias quase globais; 3) o produto 3B42 da Missão de Medição de Chuvas Tropicais (TRMM) da NASA (Huffman et al., 2007; Scheel et al., 2011); 4) campos de precipitação de modelos atmosféricos do Climate Forecast System version 2 (CFSv2) da NOAA (Saha et al., 2014; Yuan et al., 2011); e 5) observações de precipitação in situ (C. C. Funk et al., 2014).

Neste estudo, foram utilizados dados de precipitação com periodicidade mensal e resolução espacial de 0.05° , obtidos do *Climate Hazards Center* da Universidade da Califórnia. Esse conjunto de dados assimila diversas fontes de informações, incluindo dados de satélite e estações de campo, com alta resolução espacial e temporal, além de apresentar boas correlações para a Região Amazônica (ARB) conforme Haghtalab et al (2020).

GLEAM

O *Global Land Evaporation Amsterdam Model* (GLEAM) foi desenvolvido e fornecido pela *School of Geographical Sciences da University of Bristol*. É um conjunto de algoritmos que estimam separadamente os diferentes componentes da evaporação terrestre (ET), incluindo 10 componentes, tais como: Evaporação real (E), Evaporação do solo (Eb), Perda por interceptação (Ei), Evaporação potencial (Ep), Sublimação da neve (Es), Transpiração (Et), Evaporação de água aberta (Ew), Estresse evaporativo

(S), Umidade do solo da zona radicular (SMroot) e Umidade superficial do solo (SMsurf). Tem uma resolução espacial de 0.25° X 0.25° e uma resolução temporal dividida em três tipos: dia, mês e ano, durante um período de tempo entre 1980 e 2020. Nessa investigação, é utilizado o produto de dados v3.5a, com base em dados de satélite e reanálise (radiação líquida ERA5 e temperatura do ar) (Martens et al., 2017; Miralles et al., 2011).

MERIT Hydro DEM

O MERIT DEM é um Modelo Digital de Elevação (DEM) desenvolvido através da remoção de diferentes componentes de erro, como viés absoluto, ruído de faixa, ruído de ponto e viés de altura de árvores, dos DEMs espaciais existentes (SRTM3 v2.1 e AW3D-30m v1) (Yamazaki et al., 2017). O MERIT Hydro é um conjunto de dados de hidrografia global derivado do MERIT DEM e conjuntos de dados de corpos d'água como G1WBM, GSWO e OpenStreetMap, que inclui: Direção do fluxo, Elevações ajustadas hidrologicamente, Área de drenagem a montante, Número de pixels de drenagem a montante, Largura do canal do rio e Altura sobre a drenagem mais próxima (HAND) (Yamazaki et al., 2019). Os dados de elevação do terreno das Elevações Ajustadas Hidrologicamente têm uma resolução espacial de 90 m (aproximadamente 0.0008°). Os valores de declive, área total de captação, largura de fluxo e aspecto foram estimados com o suporte de sistemas de informação geográfica (SIG) e, em seguida, foram transformados para 0.25° aplicando uma técnica de agregação com a média.

3.1.2. Medições In Situ

RIMAS y SIAGAS

A Rede Integrada de Monitoramento de Águas Subterrâneas (RIMAS) é um projeto do Serviço Geológico do Brasil (CPRM), de acordo com suas atribuições estabelecidas na Lei nº 8.970, de 28/12/1994. Essa rede de monitoramento é fundamentalmente quantitativa, mantendo registros consecutivos das variações dos Níveis de Água (NA) ao longo do tempo, bem como medições anuais de condutividade elétrica, pH, potencial redox, além de atender parcialmente aos parâmetros mínimos estabelecidos pelo CONAMA Resolução nº 396 (CPRM, 2020). A Rede é composta por poços existentes (cedidos) e poços construídos, buscando obter uma distribuição e densidade suficientes para obter valores representativos. Esses poços de monitoramento ou piezômetros consistem em medidores automáticos de nível de água ou *dataloggers*, alguns dos quais são do tipo bóia-contrapeso e borbulhamento. Em 14/02/2022 existem 409 poços cadastrados, os quais podem ser consultados na página web do projeto (<http://rimasweb.cprm.gov.br>), encontrando dados gerais, construtivos, hidráulicos e geológicos, bem como monitoramento de NA e parâmetros hidroquímicos.

O Sistema de Informação de Água Subterrânea (SIAGAS), também desenvolvido pela CPRM, é uma ferramenta complementar que dá consistência às informações cadastradas. Consiste em um banco de dados constantemente atualizado com informações sobre as fontes hídricas, mais especificamente sobre os poços de exploração. Por meio de sua plataforma web (<http://siagasweb.cprm.gov.br>) é possível

realizar consultas, pesquisas, além de extrair e gerar relatórios. Em 14 /02/2022 conta com um total de 347.327 registros, incluindo poços e fontes naturais. É possível encontrar dados gerais, construtivas, geológicas, hidráulicas e análises químicas.

Nessa pesquisa, são utilizados os registros de variação do nível de água de todos os poços de monitoramento localizados no ARB e no AAS. Também são empregados os registros dos níveis estáticos registrados nos testes de bombeamento, bem como os valores das vazões de exploração dos poços cadastrados no sistema RIMAS.

3.2. Métodos

Esta pesquisa é de natureza observacional e tem como objetivo caracterizar as mudanças no armazenamento de água subterrânea (Δ GWS) para o monitoramento regional e local do AAS e da ARB, com base em informações de acesso livre obtidas a partir de sensoriamento remoto e estações de medição. Seu desenvolvimento envolve a implementação de metodologias de geoprocessamento em Sistemas de Informação Geográfica (SIG), geoestatística, sensoriamento remoto (SR) e *Machine Learning* (ML). Os parâmetros com os quais será realizado o tratamento dos dados de base dessa investigação correspondem principalmente a dados quantitativos, como coluna armazenamento de água terrestre (TWS), precipitação (P), evapotranspiração (ET), escoamento superficial (Qs), nível de água subterrânea (WL), armazenamento de água na umidade do solo (SMS), entre outros. No entanto, durante a etapa de análise e conclusões, serão utilizadas informações qualitativas, configurando-se também como uma investigação abrangente ou mista.

A metodologia usada para atingir os objetivos deste projeto está dividida em cinco etapas (Figura 13). Diversas fontes de informação foram integradas para a obtenção das variáveis hidrometeorológicas, morfológicas, geológicas, de cobertura vegetal e de atividades antrópicas. Durante a validação dos resultados do modelo e da série temporal do Δ GWS, foram realizados testes de hipóteses com relação à média e aos coeficientes de correlação. Foram utilizados os algoritmos de ML, *Adaptive Boosting* e *Random Forest*. O primeiro corresponde a um algoritmo de boosting, o que significa que seu funcionamento é sequencial para reduzir o viés, e o segundo é um algoritmo de bagging, cujo funcionamento tem como objetivo minimizar a variância.

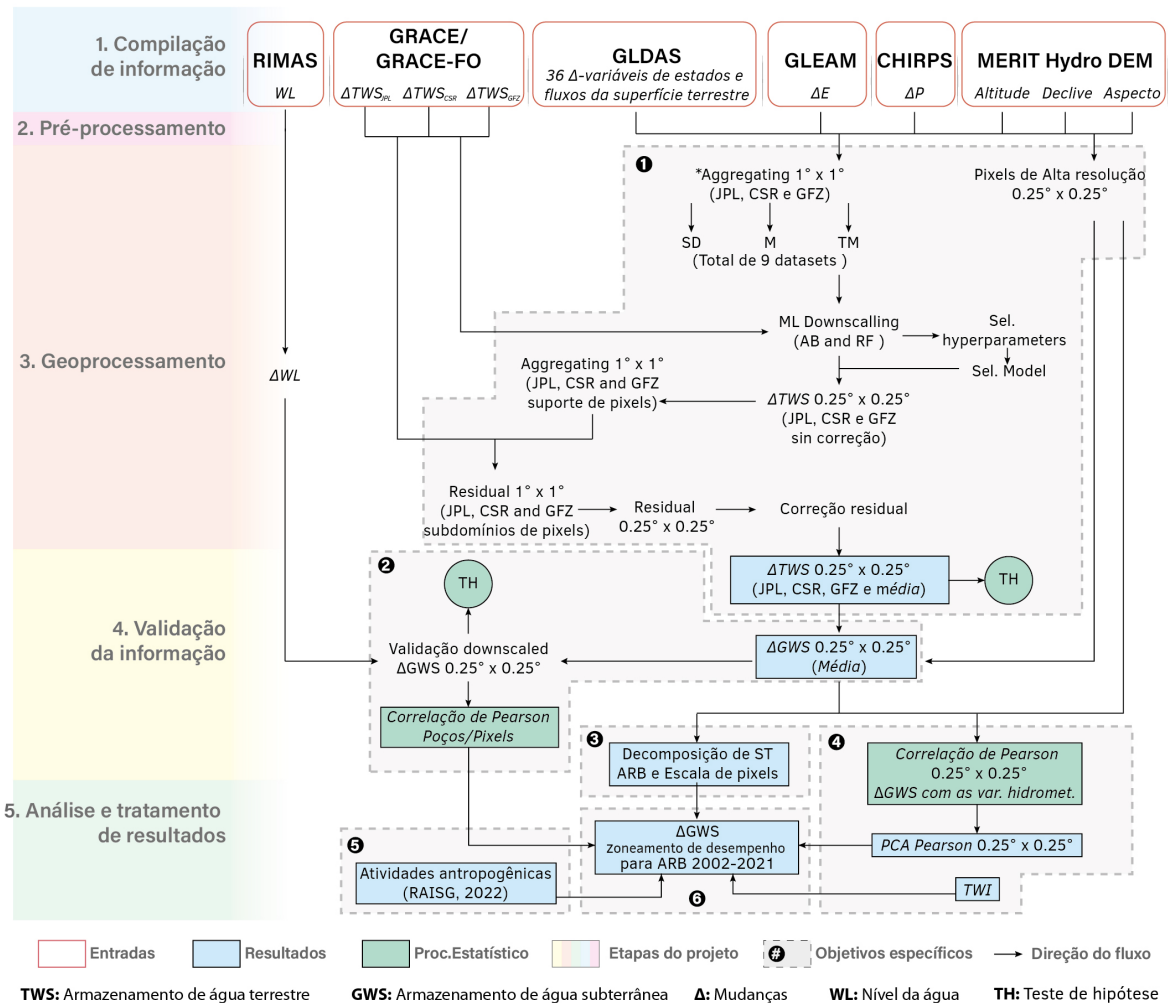


Figura 13. Fluxograma da metodologia.

3.3. Compilação de informação

Os dados espaço-temporais das variáveis hidrometeorológicas foram acessados das páginas web de cada um dos programas ou modelos correspondentes, sendo os principais RIMAS, GRACE/GRACE-FO, GLDAS, GLEAM, CHIRPS, MERIT Hydro-DEM, SIAGAS, MapBiomias e RAISG, obtendo-se os valores das Mudanças Totais de Armazenamento de Água Terrestre (ΔTWS), Umidade do solo (SM), Altura equivalente da neve (SWE), Água do dossel (CWS), Escoamento superficial (Q), Evaporação (E) e Precipitação (P) por meio de imagens multitemporais em formato raster com resoluções espaciais de 1° , 0.25° e 0.05° , e temporal de um mês. Modelo Digital de Elevação (DEM) com resolução espacial de 90 m, bem como o mapeamento geológico da América do Sul da CPRM, foram extraídos dentro dos limites da ARB. Em relação às medições *in situ*, foram compiladas as séries temporais dos níveis de água subterrâneas (WL) registrados pelos piezômetros do projeto RIMAS, assim como as informações contidas em seus perfis construtivos (localização, profundidade, distribuição das seções filtrantes, data de instalação) e lito-estratigráficos.

3.3.1. Aquisição de dados

Os dados relativos às variáveis hidrometeorológicas para a ARB e o AAS são determinados a partir de Sensores Remotos (SR) e modelos derivados. As medições *in situ*, que serão empregadas para a validação dos resultados obtidos com os SR, são obtidas por medições em terra (GBM) e disponibilizadas pela RIMAS. As características dos diferentes conjuntos de dados usados neste trabalho encontram-se resumidas na Tabela 5.

Tabela 5. Resumo do conjunto de dados e variáveis usadas na pesquisa. Onde ground-based measurements (GBM), remote sensing products (RS), and land surface models (LSM).

Variable	Missão	Produto	Resolução	Fonte
Armazenamento de água terrestre (TWS) (CSR, JPL, and GFZ)	GRACE/GRACE-FO	RS	1° - Monthly	http://www2.csr.utexas.edu/grace https://podaac.jpl.nasa.gov/ http://grace.jpl.nasa.gov
Albedo Temperatura média superficial (AvgSurfT) Água do dossel da planta (CanopInt) Evaporação da água do dossel (ECanop) Evaporação direta do solo descoberto (ESoil) Evapotranspiração (Evap) Fluxo de radiação de onda longa para baixo (LWdown) Fluxo de radiação de ondas longas (Lwnet) Taxa de evaporação potencial (PotEvap) Pressão da superfície (Psurf) Umidade específica (Qair) Fluxo de calor do solo (Qg) Fluxo de calor sensível (Qh) Fluxo de calor latente (Qle) Escoamento superficial (Qs) Fluxo de base - escoamento de água subterrânea (Qsb) Derretimento da neve (Qsm) Taxa de precipitação de chuva (Rainf) Taxa de precipitação total (RainfT) Umidade do solo na zona da raiz (RootMoist) Profundidade da neve (SnowDepth) Taxa de precipitação de neve (Snowf) Umidade do solo (0-10cm) (SM0-10cm) Umidade do solo (10-40cm) (SM10-40cm) Umidade do solo (40-100cm) (SM40-100cm) Umidade do solo (100-200cm) (SM100-200cm) Temperatura do solo (0-10 cm) (ST0_10cm) Temperatura do solo (10-40 cm) (ST10_40cm) Temperatura do solo (40-100cm) (ST40_100cm) Temperatura do solo (100-200cm) (ST100_200cm) Fluxo de radiação de ondas curtas para baixo (SWdown) Equivalente hídrico da profundidade da neve (SWS) Fluxo de radiação de ondas curtas (Swnet) Temperatura do ar (Tair) Transpiração (Tveg) Velocidade do vento (Wind)	GLDAS	LSM	0.25° - Monthly	https://disc.gsfc.nasa.gov/
Precipitação (Prec_CHIRPS)	CHIRPS	RS and GBM	0.05° - Monthly	https://www.chc.ucsb.edu/data/chirps
Evaporação (Evap_GLEAM)	GLEAM	RS and LSM	0.25° - Monthly	https://www.gleam.eu/
Declive Altitude Aspecto	MERIT Hydro DEM	DEM	90m	http://hydro.iis.u-tokyo.ac.jp/
Nível de Água (WL)	RIMAS	GB	Point - Daily	http://rimasweb.cprm.gov.br/

3.3.2. Informações complementares

Para interpretar minuciosamente os resultados, é necessário considerar aspectos como a geologia, as mudanças nas coberturas vegetais, a distribuição de áreas inundáveis, a localização de áreas protegidas (territórios indígenas e áreas naturais protegidas) e a distribuição de áreas com principais

atividades humanas que possam afetar o ciclo hidrológico. Utilizamos o conjunto de dados de geologia da América do Sul em escala de 1:5000000 (Gómez, J et al., 2019), desenvolvido sob a coordenação dos serviços geológicos do Brasil (CPRM) e da Colômbia (SGC). Este pode ser consultado e baixado em vários formatos em: <https://www2.sgc.gov.co/MGC/Paginas/gmsa5M2019.aspx>.

As mudanças nas coberturas vegetais foram obtidas através do Projeto MapBiomas Amazônia, uma iniciativa multi-institucional da Rede Amazônica de Informação Socioambiental Georreferenciada (RAISG) para gerar mapas anuais de cobertura e uso da terra a partir de processos de classificação automatizada de imagens de satélite. A descrição completa do projeto pode ser encontrada em <http://amazonia.mapbiomas.org>.

A extensão das áreas úmidas foi obtida de Hess, et al., 2015. Este conjunto de dados fornece um mapa da extensão das áreas úmidas, tipo de vegetação e estado de inundação de duas estações em toda a bacia baixa do Amazonas, obtido a partir de imagens do Radar de Abertura Sintética (SAR) do Satélite Japonês de Recursos Terrestres (JERS-1) para o período de outubro-novembro de 1995 e maio-julho de 1996. Pode ser consultado em https://daac.ornl.gov/LBA/guides/LC07_Amazon_Wetlands.html. As áreas protegidas e principais atividades humanas foram obtidas da RAISG em <https://www.raisg.org/es/mapas/>.

3.4. Pré-processamento

Devido às características espaço-temporais das variáveis obtidas a partir de SR, estas são normalmente fornecidas em *Datacubes* ou Cubos de Dados, matrizes de dados multidimensionais onde os dados e as séries temporais (sequência de imagens raster) são armazenados. Os formatos mais comuns são NetCDF ou NetCDF4, cuja manipulação difere da de um formato raster comum. Os dados obtidos foram transformados em *dataframes* (tabelas com linhas e colunas) e armazenados em bancos de dados. Para seu tratamento espacial dentro de SIG, foram definidos como Shapes tipo Ponto, os quais posteriormente foram transformados em um valor de pixel, para fins de análise.

Sobre os conjuntos de dados de todas as variáveis obtidas, é aplicada uma análise exploratória de dados para conhecer as características dos mesmos, identificando possíveis lacunas de informação, anomalias e padrões. Nos dados de SR foram aplicados os seguintes processos:

- **Conversão de formato:** os arquivos *.netcdf* (*Network Common Data Form*) originais foram transformados em dados tabulados no formato *.xlsx* ou *.csv* para processamento no *Microsoft Excel*, por meio de comandos como "*raster pixel to point*" das plataformas GIS.
- **Padronização:** foram extraídos os valores dos *pixels* contidos nos limites geográficos e temporais previstos no projeto, ou seja, a ARB e o período de tempo entre março de 2002 e dezembro de 2021, para realizar a transformação das unidades de todas as variáveis para milímetros (mm) de coluna de água equivalente.

- **Estimativa da linha de base:** de acordo com Landerer e Cooley (2019), o valor médio de cada *pixel* deve ser estimado levando-se em conta o maior período de dados contínuos de registro do GRACE, o ocorre para o período entre janeiro de 2004 a dezembro 2009. A linha de base é o valor resultante da subtração do valor original de cada mês a cada *pixel*, do valor médio de 2004 a 2009, obtendo-se as series temporais variações ou diferenciais (Δ) de cada *pixel* ajustados a um valor médio.

Ao final dessa etapa, foi obtido um conjunto de dados para o geoprocessamento da informação em um formato de fácil tratamento, em que os atributos se encontram com unidades, intervalos espaciais e temporais padronizados, ajustados a uma coerente linha de base para as necessidades de processamento GRACE/GRACE-FO.

Existem lacunas de informação temporal e espacial, como ocorre com alguns *pixels* no caso de dados GLDAS e GLEAM, ou períodos de tempo incompletos como ocorre com GRACE/ GRACE FO. Para isso, as lacunas de informações temporais foram obtidas após a aplicação dos modelos de *downscaling*, a partir dos valores das outras variáveis hidrometeorológicas. As lacunas de informações espaciais foram estimadas a partir da média dos 4 pixels mais próximos.

3.5. Geoprocessamento

Com os conjuntos de dados agora transformados em diferenciais ou mudanças (Δ), padronizados e tabulados, é possível iniciar a etapa de geoprocessamento. Consiste na aplicação de ferramentas para o processamento de dados geográficos por meio da modelagem e análise. Durante esta etapa, realizou-se a ampliação dos dados de treinamento e correções residuais. Os modelos de *downscaling* foram treinados e os hiperparâmetros dos algoritmos foram calibrados.

3.5.1. Downscaling

Foram estabelecidos os algoritmos de AdaBoost (AB) e Random Forest (RF) para gerar os modelos estatísticos não lineares de *downscaling* dos dados de TWS. A escolha dos algoritmos baseou-se no uso predominante em modelos de *downscaling* e nas diferenças de abordagem (*boosting* e *bagging*). O desenho do modelo de *downscaling* foi baseado nas propostas metodológicas de Chen et al. (2019) e Ali et al. (2021), ampliando os conjuntos de dados de treinamento, realizando uma análise de sensibilidade entre variáveis e modelos, variando as estratégias de agregação e explorando os hiperparâmetros. Os resultados de ambos os modelos/algoritmos foram comparados. Inicialmente, foram definidas as relações estatísticas entre as variáveis hidrometeorológicas e topográficas em relação ao Δ TWS de cada uma das soluções (CSR, GFZ e JPL) dentro da região do ARB em uma resolução de $1^\circ \times 1^\circ$. Definindo o Δ TWS como variável dependente e o restante das variáveis como independentes, suas relações foram estabelecidas em $1^\circ \times 1^\circ$ a partir dos modelos de ML e posteriormente aplicadas às variáveis de $0.25^\circ \times 0.25^\circ$, obtendo um TWS de alta resolução. Metodologicamente, o modelo de *downscaling* seguiu os seguintes passos:

- I. Todas as variáveis independentes (total de 41, sendo 36 do GLDAS, 1 do GLEAM, 1 do CHIRPS e 3 do DEM) dentro da região do ARB entre março de 2002 e dezembro de 2021 foram reamostradas de suas resoluções originais para 1°, de acordo com as diferentes soluções de pixels CSR, GFZ e JPL. A reamostragem utilizou três estratégias de agregação: Desvio Padrão (SD), Média (M) e Média Truncada (TM), obtendo um total de 3 reamostragens para cada solução ΔTWS , com um total de 9 conjuntos de dados de treinamento. Cabe destacar que o processo de reamostragem consistiu em obter um valor médio dos 16 pixels de 0.25° (que compõem um pixel de 1°) da resolução original, de acordo com a configuração de pixels de 1° dos dados de GRACE.
- II. Os modelos foram treinados e avaliados utilizando os algoritmos AB e RF com cada um dos nove conjuntos de dados com hiperparâmetros padrão, usando 70% dos dados para treinamento e 30% para validação. O conjunto de dados de melhor desempenho de cada solução (CSR, GFZ e JPL) foi submetido à calibração dos hiperparâmetros tanto para AB quanto para RF, selecionando os melhores modelos por solução. A partir de uma análise multiplicativa das variáveis, com base nos coeficientes de correlação e na Técnica de Importância de Recursos por Permutação (*Permutation Feature Importance technique* ou PFIT), as variáveis mais influentes no modelo foram definidas e com elas um modelo final foi treinado, usando a totalidade dos dados de cada conjunto de dados.
- III. O ΔTWS derivado dos modelos AB e RF com resolução de 0.25° foi reamostrado novamente para 1° com seus respectivos soluções. Os resíduos entre o ΔTWS original e o ΔTWS derivado a 1° foram estimados. Por fim, os resíduos foram adicionados ao ΔTWS derivado de 0.25° (correção residual). Essa correção residual garantiu que o ΔTWS reduzido correspondesse aos dados originais (derivados de GRACE/GRACE-FO) e corrigiu o viés de previsão que poderia resultar de variáveis omitidas. A partir dos três conjuntos de dados resultantes (CSR, GFZ e JPL), foi obtido um ΔTWS MÉDIO.

3.5.2. Estimativa das mudanças no armazenamento de água subterrânea (ΔGWS)

O sinal registrado pelo GRACE / GRACE FO corresponde às mudanças no campo gravitacional da Terra medidas a partir do sistema de alcance de microondas de alta precisão entre as duas aeronaves. Essas medições são convertidas em concentrações de massa e suas variações são interpretadas como Mudanças no Armazenamento de Água Terrestre ou Total (ΔTWS). Essas medições são usadas para derivar a distribuição global das águas subterrâneas em grande escala em conjunto com informações hidrológicas adicionais, como um produto do balanço hídrico terrestre:

$$P - E - Q = \Delta TWS \text{ [balanço hídrico terrestre]} \quad (1)$$

$$\Delta TWS = \Delta GW + \Delta SM + \Delta SWE + \Delta SWS \quad (2)$$

$$\Delta GW = \Delta TWS - \Delta SM - \Delta SWE - \Delta SWS \quad (3)$$

Onde:

P	= Precipitação (CHIRPS)
E	= Evaporação (GLEAM)
Qs	= Descarga do rio (GLDAS)
ΔTWS	= Mudança no armazenamento de água terrestre (GRACE/GRACE-FO)
ΔGWS	= Mudança no armazenamento de água subterrânea
ΔSMS	= Mudança na umidade do solo (GLDAS)
ΔSWE	= Mudança no equivalente em água da neve (GLDAS)
ΔCWS	= Mudança no armazenamento de água de superfície (GRACE)

O ΔGWS foi obtido após o downscaling dos dados para 0.25° (Tabela 5), para validar e comparar com as estimativas em 1° . Os componentes de armazenamento foram expressos em volume com unidades de km^3 e altura de água equivalente em mm.

3.6. Validação da informação

Os dois elementos principais a serem validados foram os modelos de *downscaling* para TWS e as séries temporais (ST) de ΔGWS . Em relação aos modelos de aprendizado de máquina (ML), foram comparados os valores originais (1°) de GRACE/GRACE-FO com os valores preditos (0.25°), estimando o desvio ou erro e analisando-os por meio de métricas como RMSE (Equação 4), MAE (Equação 5) e R^2 . As séries temporais de ΔGWS e ΔWL foram validadas pelo coeficiente de correlação de Pearson. Foram realizadas correlações nas ST totais, bem como em suas componentes de tendência, sazonalidade e residual. Também foram aplicados dois testes de hipóteses, um em relação à média entre os dados originais (1°) e escalados (0.25°), e o outro a partir de valores críticos nos coeficientes de correlação das ST de ΔWL e ΔGWS .

3.6.1. Comparação de informações e análise de erro

As métricas usadas para avaliar os resultados foram:

Raiz do Erro Quadrático Médio ou Root-Mean-Square Error (RMSE)

Também chamado de erro padrão, corresponde à raiz quadrada da MSE. Quanto menor o valor, mais próximos os valores previstos estão dos reais, ou seja, quanto menor o RMSE, maior a precisão do modelo.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}} \quad (4)$$

Onde X_i é o valor original e Y_i o estimado pelo modelo e n o número total de amostras.

Erro médio absoluto o Mean absolute error (MAE)

Ele representa a média do valor absoluto do desvio de cada um dos valores simulados pela média aritmética, o que pode evitar problemas de cancelamento entre os valores de erro, podendo assim refletir com precisão a previsão do erro atual.

$$MAE = \frac{1}{n} \left(\sum_{i=1}^n |X_i - Y_i| \right) \quad (5)$$

Coefficiente de correlação Pearson (R)

Teste estatístico que permite analisar a relação linear entre duas variáveis. Seu valor pode variar entre -1 e 1, onde o sinal indica a direção da correlação e o valor numérico, sua magnitude.

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (6)$$

Coefficiente de determinação (R²)

Estatísticas usadas com o propósito de prever resultados futuros ou testar uma hipótese. Ajuda a determinar a qualidade de um modelo para replicar os resultados. É o quadrado do coeficiente de Pearson, indicando o percentual de variação de uma variável em relação a outra. Ou seja, a proporção de uma variável explicada pela outra. Pode assumir valores entre 0 e 1.

3.6.1. Teste de significância

Foram aplicados dois testes de significância utilizando o método do teste t de Student. O primeiro foi usado para comparar as médias das séries temporais entre os pixels escalados e os originais nas diferentes soluções (CSR, GFZ e JPL), identificando pixels com diferenças significativas na média como resultado do *downscaling*. Um segundo teste de correlação entre as séries temporais dos níveis de água (WL) nos poços de monitoramento RIMAS e os Δ GWS derivados e preditos (1° e 0.25°) foi realizado, estimando os coeficientes de correlação de Pearson (R) e os valores críticos. Os níveis de significância definidos foram 0.05 (5%) e 0.01 (1%).

3.7. Análise e processamento dos resultados**3.7.1. Decomposição das séries temporais**

As séries temporais foram decompostas em seus componentes de tendência [T(t)], sazonais [S(t)] e residuais [R(t)] aplicando a abordagem não paramétrica STL (*Seasonal Trend decomposition using Loess*) de Cleveland et al., (1990), conforme descrito na Equação (7)

$$h(t) = T(t) + S(t) + R(t) \quad (7)$$

Com relação ao sinal total $[h(t)]$ e ao componente de tendência $[T(t)]$, é possível obter a tendência linear $[TL(t)]$, para descrever o comportamento de longo prazo. Também foi aplicada uma dessazonalização, que consiste em subtrair o efeito sazonal da série temporal. Esse ajuste sazonal $[hSA(t)]$ foi obtido a partir do modelo aditivo descrito na equação (7), subtraindo o valor sazonal $[S(t)]$ de $[h(t)]$, conforme mostrado na equação (8).

$$hSA(t) = h(t) - S(t) = T(t) + R(t) \quad (8)$$

A tendência linear (TL) mede a direção e magnitude da mudança das variáveis ao longo do tempo, ou seja, calcula padrões de crescimento ou diminuição ao longo do tempo. A componente Sazonal (Seasonal) é caracterizada por padrões repetitivos em períodos de 12 meses (1 ano), causados por fenômenos naturais de caráter cíclico, tornando-a previsível. Ao separar o efeito sazonal das séries temporais, tem-se uma visão mais clara dos comportamentos a longo prazo, facilitando a identificação dos efeitos de fenômenos climáticos com períodos mais longos, como La Niña ou El Niño, e comportamentos de longo prazo.

3.7.2. Análise de Componentes Principais (PCA)

O PCA, pelas suas siglas em inglês, foi formulado pela primeira vez pelo matemático e estatístico britânico Karl Pearson em 1901. É um método estatístico cujo objetivo é simplificar a complexidade de conjuntos de dados grandes, reduzindo a dimensionalidade das variáveis nos dados e preservando suas informações. Este método aumenta a interpretabilidade e minimiza a perda de informação, criando novas variáveis não correlacionadas que maximizam a variância, chamadas de componentes principais (PC). É considerado um algoritmo de aprendizado não supervisionado dentro do Machine Learning.

Sua aplicação requer a normalização dos dados, garantindo que as variáveis estejam na mesma escala, evitando que uma variável com uma escala maior domine a análise. Estimar as componentes principais (PC) é um problema de autovalores e autovetores. Primeiro, são calculados os autovetores e autovalores da matriz de covariância ou da matriz de correlação dos dados originais. Os autovetores são vetores unitários (ou seja, têm magnitude igual a 1) e representam as principais direções de máxima variância nos dados. Esses autovetores transformam os dados originais no novo espaço de menor dimensionalidade. Os autovalores representam a quantidade total de variância explicada por cada componente principal. Os loadings são obtidos a partir dos autovetores. Cada loading corresponde a um autovetor e mostra como cada variável original contribui para a formação de um componente principal específico. Os valores dos loadings indicam a magnitude e a direção da contribuição de cada variável

para esse componente. Os loadings também são normalizados, de modo que a soma de seus quadrados seja igual a 1, o que ajuda na interpretação de sua importância relativa.

A partir do *downscaling* e da estimativa das anomalias de ΔGWS entre 2002 e 2021 a 0.25° , determinam-se os coeficientes de correlação para cada pixel entre as séries temporais de ΔGWS e cada uma das variáveis hidrometeorológicas a 0.25° . Com esses valores, criam-se mapas de correlação, nos quais é possível identificar domínios ou áreas dentro da ARB onde o comportamento entre as duas variáveis pode ser nulo, diretamente ou indiretamente proporcional. Os valores de correlação desses mapas estarão normalizados com valores entre 1 e -1, facilitando a aplicação da Análise de Componentes Principais.

Por meio deste método estatístico, simplifica-se a complexidade das relações entre ΔGWS e as variáveis hidrometeorológicas, mantendo sua informação, ou seja, o PCA reduz o número de variáveis, agora chamadas de componentes principais, mantendo e explicando aproximadamente a mesma informação das variáveis originais. Este método de aprendizado não supervisionado permite extrair em poucas componentes as principais relações entre as séries temporais de ΔGWS e cada uma das variáveis hidrometeorológicas, que são representadas espacialmente após a transformação dos dados em cada pixel, representando agora novos padrões espaciais associados a essas características. As variações espaciais que melhor explicam a variância entre os coeficientes de correlação estão contidas nas três primeiras componentes principais (PC1, PC2 e PC3), com as quais é obtida uma composição RGB que representa os domínios mais característicos do comportamento de ΔGWS em relação às variáveis hidrometeorológicas. Sobre esses domínios, os resultados da validação são comparados com os poços de monitoramento da RIMAS, identificando as variáveis que tiveram maior influência de acordo com a localização espacial.

3.7.3. Índice de umidade topográfica ou Topographic Wetness Index (TWI)

Com o propósito de identificar a influência da topografia nos ΔGWS , propõe-se a estimativa e análise do TWI. Este índice é uma medida do equilíbrio entre o suprimento de água de uma bacia hidrográfica e o escoamento local, causados pela topografia (Kopecký et al., 2020). Foi proposto por Moore et al., 1991, como uma métrica para estimar a umidade do solo utilizando dados topográficos. Ele mede a variação na umidade do solo com base na topografia de uma área ou bacia hidrográfica. Para sua estimativa, devem-se eliminar os espaços vazios ou pixels ausentes. Calcula-se a área total da bacia hidrográfica (Total catchment area), a largura do fluxo (Flow width) e a inclinação (Slope). Essas três variáveis são integradas no TWI de acordo com a equação (9)

$$TWI = \ln \frac{TCA/FW}{\tan(Slope)} \quad (9)$$

Onde TCA é *Total catchment área*, FW é o *Flow width* e *Slope* está em radianos

Para sua interpretação, o TWI é uma medida indireta da umidade do solo. Quantitativamente, os valores mais altos indicam áreas com maior umidade, enquanto os valores mais baixos correspondem a áreas mais secas. Isso é estabelecido levando em consideração que áreas com menor inclinação e maior acumulação de fluxo retêm maior quantidade de água do que áreas com alta inclinação.

Na estimativa do TWI, foram utilizados dados de elevação do Modelo Digital de Elevação MERIT Hydro-DEM de Yamazaki et al., 2019. Sobre cada uma das variáveis topográficas, assim como o resultado, foi aplicado um processo de agregação com os valores médios para realizar um upscaling de uma resolução espacial de 90m para 0.25° (aproximadamente 27500m). O software SAGA GIS 7.8.2 foi utilizado para a estimativa das variáveis e o cálculo final do TWI

3.7.4. Zoneamento das variações no Δ GWS

Analisar o comportamento de longo e curto prazo das séries temporais de Δ GWS em cada pixel e sua relação com as variáveis hidrometeorológicas, morfológicas, geológicas e a interação com atividades antrópicas é possível através da delimitação de zonas com comportamentos similares. A partir dos mapas do desvio padrão (SD) e tendência linear (LT) das séries temporais de Δ GWS em cada pixel. Foi realizado uma reclassificação em 3 categorias para cada métrica, que, ao se sobrepor, estabelecem um total de 9 subdomínios característicos que resumem o comportamento de Δ GWS entre 2002 e 2021. Por outro lado, os mapas de correlação entre Δ GWS e cada uma das variáveis hidrometeorológicas, bem como a composição RGB dos três principais componentes do PCA, revelam diferentes padrões espaciais ou subdomínios característicos do comportamento entre as variáveis e Δ GWS entre 2002 e 2021. Esses domínios são interpretados e analisados integrando comparações pontuais onde existem poços de monitoramento da RIMAS com valores de inclinação e TWI como parte das características morfológicas, das unidades hidrogeológicas, do tipo de aquífero, bem como os mapas de mudanças na cobertura vegetal do MAPBIOMAS, 2022, e as áreas com atividades antrópicas influentes, como incêndios florestais, desmatamento, construção de hidrelétricas, áreas protegidas e de conservação da RAISG, 2022.

O PCA, pelas suas siglas em inglês, foi formulado pela primeira vez pelo matemático e estatístico britânico Karl Pearson em 1901. É um método estatístico cujo objetivo é simplificar a complexidade de conjuntos de dados grandes, reduzindo a dimensionalidade das variáveis nos dados e preservando suas informações. Este método aumenta a interpretabilidade e minimiza a perda de informação, criando novas variáveis não correlacionadas que maximizam a variância, chamadas de componentes principais (PC). É considerado um algoritmo de aprendizado não supervisionado dentro do Machine Learning.

Sua aplicação requer a normalização dos dados, garantindo que as variáveis estejam na mesma escala, evitando que uma variável com uma escala maior domine a análise. Estimar as componentes

principais (PC) é um problema de autovalores e autovetores. Primeiro, são calculados os autovetores e autovalores da matriz de covariância ou da matriz de correlação dos dados originais. Os autovetores são vetores unitários (ou seja, têm magnitude igual a 1) e representam as principais direções de máxima variância nos dados. Esses autovetores transformam os dados originais no novo espaço de menor dimensionalidade. Os autovalores representam a quantidade total de variância explicada por cada componente principal. Os loadings são obtidos a partir dos autovetores. Cada loading corresponde a um autovetor e mostra como cada variável original contribui para a formação de um componente principal específico. Os valores dos loadings indicam a magnitude e a direção da contribuição de cada variável para esse componente. Os loadings também são normalizados, de modo que a soma de seus quadrados seja igual a 1, o que ajuda na interpretação de sua importância relativa.

3.8. Programas Utilizados

Os Sistemas de Informações Geográficas (SIG ou GIS em inglês) foram os principais softwares utilizados nesta pesquisa. As versões usadas são ArcGIS v10.8 e QGIS v3.22.5 para o processamento de dados espaciais. O geoprocessamento do Modelo Digital de Elevação (MDE ou DEM em inglês) de 90m na estimativa do Índice de Umidade Topográfica (TWI) foi realizado com o SAGA GIS v7.8.2. O treinamento e a validação dos modelos de *machine learning* (ML) foram realizados com Python v3.7.1. A decomposição e a interpretação das séries temporais foram realizadas em Python e Orange Data Mining v3.34. A manipulação de informações tabuladas, criação de gráficos iniciais e processamento de dados foram feitos com o Microsoft Excel. A diagramação e o design das figuras tiveram o apoio do Adobe Illustrator 2020 v24.1.2.

4. RESULTADOS

4.1. ARTIGO CIENTÍFICO: Characterization of groundwater storage changes in the Amazon River Basin based on downscaling of GRACE/GRACE-FO data with machine learning models

Diego Alejandro Satizábal Alarcón¹; Alexandra Vieira Suhogusoff¹; Luiz C. K. M. Ferrari¹

¹Institute of Geosciences, Groundwater Research Center (CEPAS), University of São Paulo (USP), Rua do Lago 562 - Cidade Universitária, 05508-080 São Paulo – SP, Brasil.

Corresponding author: diego.satizabal@usp.br

Published in the Science of the Total Environment journal (STOTEN) 912 (2024) 168958
DOI: <https://doi.org/10.1016/j.scitotenv.2023.168958>

Annex A

Resumo

Resumo: As mudanças no armazenamento de águas subterrâneas na Bacia Amazônica (ARB) desempenham um papel importante no comportamento hidrológico da região, com influência significativa na variabilidade climática e nos ecossistemas da floresta tropical. As missões satelitais GRACE e GRACE-FO fornecem anomalias de gravidade a partir das quais é possível monitorar mudanças no armazenamento de água terrestre, embora em baixa resolução espacial. Este estudo reduziu os dados do GRACE e do GRACE-FO a partir de modelos de aprendizado de máquina de 1° (aproximadamente 110 km) para $0,25^\circ$ (aproximadamente 27,5 km). Ele estimou a variabilidade espaço-temporal das anomalias de armazenamento terrestre e de águas subterrâneas entre 2002 e 2021 para a Bacia Amazônica. Paralelamente, os algoritmos Random Forest e AdaBoost foram comparados e analisados. Os resultados refletiram um bom ajuste dos modelos com um erro muito baixo e uma ligeira superioridade nas previsões obtidas pelo AdaBoost. Nas previsões a $0,25^\circ$, foram identificados padrões espaciais associados à forte influência nas mudanças de armazenamento de alguns rios e montanhas cobertas de neve, bem como um aumento na precisão dos dados escalados dos originais. Também foi obtido um comportamento positivo de longo prazo no armazenamento terrestre e de águas subterrâneas de $14,26 \pm 1,18 \text{ km}^3/\text{ano}$ e $+22,24 \pm 1,18 \text{ km}^3/\text{ano}$, respectivamente. A validação das séries temporais de anomalias de águas subterrâneas para os níveis de água nos poços de monitoramento obteve coeficientes de correlação máximos de 0,85 com níveis de confiança de 0,01. Estes resultados são promissores para a informação satelital na gestão da água, especialmente no monitoramento regional de aquíferos não confinados.

Palavras-chave: Sensoriamento remoto; AdaBoost; Random Forest; Inteligência artificial; Séries temporais; Sistema Aquífero Amazônico; Ocupação do solo.

1. Introduction

The Amazon basin is the largest river basin in the world, characterized by complex hydrological processes that connect high rates of precipitation, extensive floodplains, dense tropical forests, complex topography, and significant variations in freshwater storage and discharge (Fassoni-Andrade et al., 2021). The region is now facing risks under climate and anthropogenic changes, and disturbances in the Amazon water balance could have substantial impacts globally (Jimenez et al., 2021). Different studies and approaches to understanding the basin's hydrological behavior have been conducted, with the groundwater component being one of the least studied, although with increasing interest. Groundwater storage also plays a major role in the Amazon's hydrology and greatly influences on climate variability and rainforest ecosystems (Pokhrel et al., 2013). Strong memory effects of the Amazon groundwater system propagate climate anomalies over the region for several years (Frappart et al., 2019; Miguez-Macho & Fan, 2012; Pfeffer et al., 2014). Researchers (Alves & Wahnfried, 2017; Do Nascimento et al., 2008; Pfeffer et al., 2014; Pimentel & Hamza, 2012; Tomasella et al., 2008) developed other studies on groundwater hydrogeology at different scales, while some of them (Frappart et al., 2019; Miguez-Macho & Fan, 2012; Pokhrel et al., 2013) covered the entire river basin and implemented an integrative model-based approach. Groundwater monitoring networks play a crucial role in validating remote sensing data and comprehending changes in groundwater storage for research. Nonetheless, their installation can be costly and unevenly distributed across regions, posing challenges for studying groundwater storage changes at the regional scale.

Satellite-derived information is invaluable for regions like the Amazon River Basin (ARB), whose extensive size and limited access challenge field data collection. Fassoni-Andrade et al. (2021) extensively review hydrology advances and challenges in the Amazon basin utilizing satellite data.

The Gravity Recovery and Climate Experiment (GRACE) missions, launched in 2002, and their successor, Follow One (GRACE-FO) since 2018, enable global measurements of terrestrial water storage changes. These satellite missions detect small fluctuations in the Earth's gravity field caused by spatiotemporal variations in terrestrial water storage on monthly or sub-monthly time scales (Tapley et al., 2004). Terrestrial Water Storage Changes (Δ TWS) result from these mass variations in terrestrial water, with TWS representing the vertical integration of all water storage components quantified in equivalent centimeters of water column. From a terrestrial water balance perspective, it becomes possible to decompose the complex hydrological variable TWS into Groundwater Storage Changes (Δ GWS) by analyzing changes in other hydrological variables like Soil Moisture (SMS), Snow Water (SWS), Canopy Water (CWS), and surface runoff (Q_s). In regional hydrological studies, applications of GRACE/GRACE-FO have yielded promising results, achieving accuracy levels of less than 1 centimeter in areas exceeding 400,000 square kilometers (Swenson et al., 2003).

Nevertheless, one of its primary limitations lies in the low spatial resolution of its $1 \times 1^\circ$ data, which hinders interpretation at local scales. Frappart & Ramillien (2018) and Pokhrel et al. (2013) employed GRACE for the spatiotemporal analysis of groundwater storage variability in the Amazon basin. Concerning the relationship between hydrogeology and its impact on the spatial distribution of Δ GWS, Castellazi et al. (2016) theoretically postulate that unconfined aquifers exhibit more pronounced localized mass changes than confined and semi-confined aquifers. Hence, comprehending spatiotemporal variations in TWS and GWS has facilitated the tracking of water movement at both planetary and regional scales, allowing the identification of long-term trends. It is important to note that satellite information has uncertainties, including those related to satellite data and assimilation models. The increasing use of artificial intelligence, particularly machine learning, has expanded the possibilities of processing remote sensor information. Applications such as *downscaling*, i.e., methods used to derive high spatial resolution information from low spatial resolution information, have gained popularity due to the possibility of using various algorithms to create nonlinear models that can pick up nonlinear solid relationships between independent and dependent variables. In remote sensing, *downscaling* refers to an increase in spatial resolution or decrease in the pixel size of an image, and upscaling to a decrease or coarsening of spatial resolution (Atkinson and Tate, 2000). In this field, researchers such as Ali et al. (2021), Agarwal et al. (2023), Chen et al. (2019), Khorrami et al. (2023), Seyoum et al. (2019), Pulla et al. (2023), Wang et al. (2023), Yin et al. (2022), Zhang et al. (2021) and Zhong et al. (2021), among others, have applied different Machine Learning (ML) models in *downscaling* GRACE/GRACE-FO derived TWS and GWS anomalies using algorithms such as Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Networks (ANN) or Multiple Linear Regression (MLR), obtaining good correlations with in situ measurements and reaching spatial resolutions with pixels from 1 to 5km on a side. This increase in the spatial resolution of TWS and GWS anomalies contributes to the identification of locally accurate spatial and temporal patterns, an improvement in terms of measurement and representation of the scaled variables, as well as an increase in the amount of information per unit area (number of pixels per unit area). In the case of ARB, this may represent a means to understand the interactions between TWS/GWS with anthropogenic activity and climate dynamics such as climate change.

Considering the importance of the ARB, the ecosystems it harbors, the strong influence it exerts on the regional and global hydrological cycle, and the lack of high-resolution and accurate TWS and GWS information, this study has the following objectives:

- Apply a 1° to 0.25° *downscaling* methodology on GRACE/GRACE-FO TWS anomalies for the Amazon River Basin (ARB) between 2002 and 2021 from ML models and different hydrometeorological and topographic variables.
- Compare Random Forest (RF) and Adaptive Boosting (AB) algorithms in TWS *downscaling*.

- Decompose and validate the GWS anomalies with the monitoring wells from Integrated Groundwater Monitoring Network (RIMAS), supported by the Geological Survey of Brazil (SGB-CPRM).
- Evaluate long-term trends in TWS and GWS changes.

The study's premise for *downscaling* is that interactions with other hydrometeorological variables and local characteristics (elevation, slope, and aspect) of higher resolution condition low-resolution data. Therefore, 43 variables were selected from five different sources, and a sensitivity analysis was conducted. Two ML algorithms based on Boosting and Bagging techniques were defined and compared. Based on the modeling, the scaled and validated GWS data was used for the temporal decomposition of the ARB time series and each pixel. This analysis identified long-term trends and their spatial distribution.

2. Study Area

The Amazon River Basin (ARB) is the largest in the world, with an approximate area of 6.3×10^6 km², distributed among six countries: Brazil (67.9%), Peru (13.0%), Bolivia (11.2%), Colombia (5.5%), Ecuador (1.7%) and Venezuela (0.7%). It comprises several rivers, extensive floodplains, and wetlands (Reis et al., 2019). Located in central and eastern South America, between the latitude and longitude ranges 5°N-20°S and 80°W-50°W, respectively. ARB extends from the Andes Mountains, in the Huayhuash Mountain in Peru, with maximum altitudes of 6,450 m.a.s.l., to the Amazonas River mouth in the Atlantic Ocean, northeast of Brazil. The Amazon River flows approximately 7,100 km to its mouth, with an average flow rate of 2.10×10^5 m³s⁻¹ (Braga et al., 2011). Its main affluents are the Madeira, Negro, Japurá, Tapajós, and Xingú rivers. Its large extent and complexity, characterized by intense hydrological processes, such as high precipitation and evapotranspiration rates, as well as large variations in water storage and river discharge, make the ARB a key player in the global climate system (Gash et al., 2013; Gatti et al., 2021; Nagy et al., 2016). The mean annual precipitation over most of the basin varies between 1,000 and 3,000 mm, with maximum values of 4,000 mm to the northeast and a minimum of 100 mm over the Andes (Haghtalab et al., 2020; Maeda et al., 2017). The Amazon tropical forest also provides a massive water flux to the atmosphere, ranging between 1,000 and 1,500 mm per year (Maeda et al., 2017). Seasonal changes in water, in turn, contribute to the formation of tropical forests (Leite et al., 2012) and sustain biodiversity (Junk, 1977; Junk et al., 2010). The ARB was subdivided into Upper Basins (>1000m), Middle Basins (170m-1000m), and Lower Basins (<170m) based on elevation (Fig. S1).

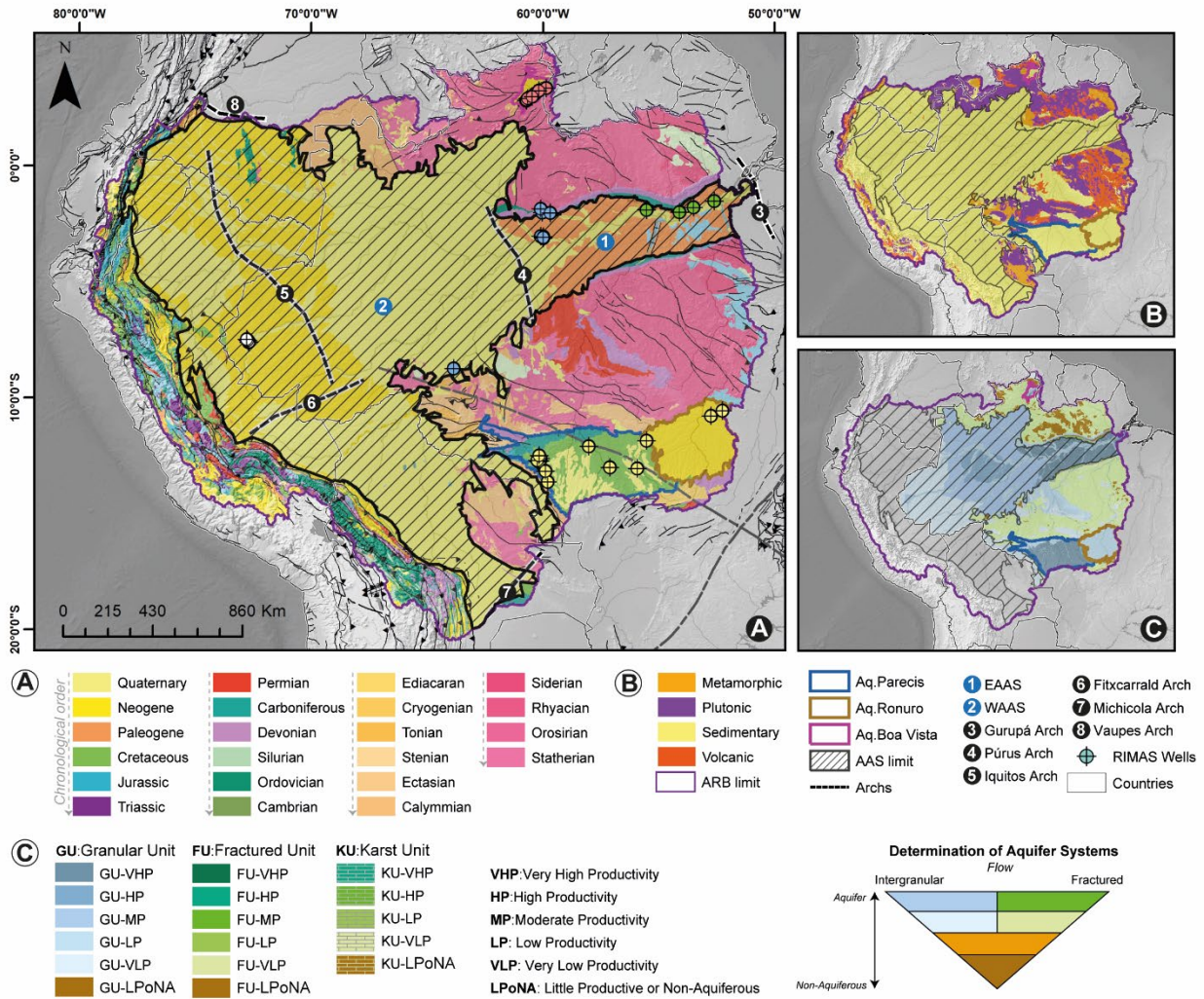


Fig. 1. Study area with a) location of the monitoring wells and b) main chronostratigraphic units extracted from the geological map of South America.

In ARB subsoil occurs the Amazon Aquifer System (AAS), a transboundary aquifer cataloged in 2005 by UNESCO/OAS as one of the largest in the world, which comprises the hydrogeological provinces of the Orinoco and the Amazon (Rosário et al., 2016). There is considerable uncertainty about the AAS extension, as evidenced by the cartography proposed by different authors (ISARM and UNESCO, 2009; Margat, 2007; Anjos et al., 2007; Rosário et al., 2016 and IGRAC, 2015). For this study, the AAS corresponds to the Cenozoic to Cretaceous sedimentary units of the geological map of South America (Gomez et al., 2019), with a total area of $3.1 \times 10^6 \text{ km}^2$, bounded to the west by the sub-Andean fault system, to the North and South, by the ARB watersheds and by the Guiana and Brazilian crystalline rock shields, as shown in Fig. 1. AAS is a regional, heterogeneous, multi-layer and multi-basin flow system, with the same free potentiometric surface connected to the rivers, settled on Paleozoic rocks (diabasic landslides or sedimentary formations) and crystalline basement (ANA, 2015). It develops over the Amazonian craton and comprises Solimões, Acre, Madre de Dios, Beni, Ucayali, Marañon, Oriente, and Putumayo sedimentary basins. These basins are filled with consolidated and unconsolidated

sediments that exhibit stratigraphic similarities from the early Cretaceous to the late Tertiary (Rosário et al., 2016). Despite the differences in the denomination between countries, some of the principal unconfined aquifers that make up the AAS are Alter do Chão, Içá, Tucunaré, Alluvial, as well as the Solimões-Repouso aquitard. The Brazilian National Water Agency (ANA) assumes the hypothesis of the same potentiometric surface between the Cenozoic, Cretaceous, and Paleozoic units, regardless of the nature of the materials, interrelated by lateral variations or short flow circuits through fractures within the Brazilian portion of the AAS (Agência Nacional de Águas ., 2015). The main geological structures bounding the sedimentary basins are the Iquitos, Purus, Gurupá, Fitzcarrald, Vaupes, and Michicola arcs and the sub-Andean fold belt (Fig. 1). The AAS is divided into two regions, the Western Amazon Aquifer System (WAAS) and the Eastern Amazon Aquifer System (EAAS), starting from the Purus Arch (Fig. 1). Other aquifers of regional importance within the ARB are the Parecis and Ronuro aquifers to the southwest and Boa Vista or North.

Within the Brazilian portion of the ARB are 69 monitoring wells of the Integrated Groundwater Monitoring Network (RIMAS) in 26 cities classified within the North, East, South, West, and Central Zones. The validation stage considered 39 of these monitoring wells due to the lack of continuity or strong anomalies in their record. These wells cross various geological materials with depths between 9 and 260m. The cities with the highest number of wells are Manaus, Vilhena, Cruzeiro do Sul, Bonfim, Boa Vista, and Presidente Figueiredo. The number of pixels over which wells are located at 1° is 19 in CSR, 21 in GFZ, 21 in JPL, and 28 at 0.25°.

3. Materials and Methods

3.1. Study design

In the context of this research, an empirical approach was adopted, which relied on implementing AdaBoost and Random Forest algorithms to carry out the *downscaling* of GRACE/GRACE-FO data in the ARB. The models were trained using a set of hydrometeorological and morphological variables from various sources. The study followed the four steps below:

- 1) Initially, the training data was standardized, transforming variables to a 1° resolution and estimating pixels with missing spatial information. Optimal aggregation functions were evaluated based on their generalization ability during upscaling. A correlation matrix was employed to understand the relationships between the variables.

- 2) The analysis focused on identifying the variables that exerted the most significant influence on the models. This involved scrutinizing their correlations with Total Water Storage (TWS) and employing the Permutation Feature Importance Technique (PFIT). Subsequent to adjusting the hyperparameters of the two algorithms, the models underwent training, predictions were generated, and accuracy was assessed in comparison with the original data. To refine the Total Water Storage (TWS) estimation, a

residual correction was implemented. Hypothesis tests were then conducted to evaluate differences between the data before and after the *downscaling* process, particularly concerning the mean.

3) In estimating changes in Groundwater Storage (Δ GWS), progress was achieved by applying the terrestrial water balance method. Temporal series were validated using data obtained from monitoring wells.

4) To comprehensively characterize changes in groundwater storage, a decomposition of the temporal components of Groundwater Storage (GWS) was executed. This allowed for both spatial and temporal insights. A comparative analysis was undertaken between Δ GWS and Δ TWS following the application of the *downscaling* process. This methodological approach comprehensively addressed the challenge of understanding and representing variations in groundwater storage for the ARB.

3.2. Data Source and Pre-processing

The target spatial resolution is 0.25° , and this choice aligns with the resolution of the independent variables. The analysis period spans from March 2002 to December 2021, totaling 236 months of data. The ARB comprises 7,740 pixels, each with a 0.25° resolution, varying from 492 to 881 pixels in 1° resolution, depending on the solution used (CSR, GFZ, and JPL). For pixels with missing values in their datasets, estimation was performed based on the average of surrounding pixels.

As for the working environment, this research was conducted using Python 3.7.1 for training and validating *downscaling* models. Additionally, ArcGIS 10.8 provided support for spatial operations and effective data representation. It is worth noting that the various datasets were acquired on April 4, 2022, through the online platforms associated with the respective programs, as summarized in Table 1 of this study.

Table 1. Summary of datasets and variables used in this study. Where ground-based measurements (GBM), remote sensing products (RS), and land surface models (LSM)

Variable	Mission	Product	Resolution	Data source
Terrestrial Water Storage (TWS) (CSR, JPL, and GFZ)	GRACE/GRACE-FO	RS	1° - Monthly	http://www2.csr.utexas.edu/grace https://podaac.jpl.nasa.gov/ http://grace.jpl.nasa.gov
Albedo				
Average surface skin temperature (AvgSurfT)				
Plant canopy surface water (CanopInt)				
Canopy water evaporation (ECanop)				
Direct evaporation from bare soil (ESoil)				
Evapotranspiration (Evap)				
Downward long-wave radiation flux (LWdown)	GLDAS	LSM	0.25° - Monthly	https://disc.gsfc.nasa.gov/
Net long-wave radiation flux (Lwnet)				
Potential evaporation rate (PotEvap)				
Surface pressure (Psurf)				
Specific humidity (Qair)				
Ground heat flux (Qg)				

4.1. ARTIGO CIENTÍFICO

Sensible heat net flux (Qh)
 Latent heat net flux (Qle)
 Storm surface runoff (Qs)
 Baseflow-groundwater runoff (Qsb)
 Snow melt (Qsm)
 Rain precipitation rate (Rainf)
 Total precipitation rate (RainT)
 Root zone soil moisture (RootMoist)
 Snow depth (SnowDepth)
 Snow precipitation rate (Snowf)
 Soil moisture (0-10 cm) (SM0-10cm)
 Soil moisture (10-40 cm) (SM10-40cm)
 Soil moisture (40-100 cm) (SM40-100cm)
 Soil moisture (100-200 cm) (SM100-200 cm)
 Soil temperature (0-10 cm) (ST0_10cm)
 Soil temperature (10-40 cm) (ST10_40cm)
 Soil temperature (40-100 cm) (ST40_100cm)
 Soil temperature (100-200 cm) (ST100_200cm)
 Downward short-wave radiation flux (SWdown)
 Snow depth water equivalent (SWS)
 Net short wave radiation flux (Swnet)
 Air temperature (Tair)
 Transpiration (Tveg)
 Wind speed (Wind)

Precipitation (Prec_CHIRPS)	CHIRPS	RS and GBM	0.05° - Monthly	https://www.chc.ucsb.edu/data/chirps
Evaporation (Evap_GLEAM)	GLEAM	RS and GBM	0.25° - Monthly	https://www.gleam.eu/
Slope				
Elevation	MERIT Hydro DEM	LSM	90m	http://hydro.iis.u-tokyo.ac.jp/
Aspect				
Water Level (WL)	RIMAS	GB	Point - Daily	http://rimasweb.cprm.gov.br/

3.2.1. GRACE and GRACE-FO

The GRACE and its successor GRACE-FO satellite missions are a partnership between the National Aeronautics and Space Administration (NASA) and the German Research Centre for Geosciences (GFZ). They provide space-time measurements of changes in the Earth's gravitational field (Tapley et al., 2004; Wahr et al., 2004). Gravity anomalies are the product of land mass changes related to the movement of water over land and subsequently transformed to equivalent water columns called Terrestrial Water Storage (TWS) (Cooley et al., 2019). The official GRACE Science Data System continuously provides gravity solutions through three different processing centers: the Jet Propulsion Laboratory (JPL), Center for Space Research at the University of Texas, Austin (CSR), and GeoforschungsZentrum Potsdam (GFZ). The data used in this study were level-3 RL06 Mass Concentration blocks (mascons) from the centers CSR, GFZ, and JPL, downloaded on 04/04/2022. The RL06 data are state-of-the-art in processing mission observations and have the best signal-to-noise ratio

and lower uncertainties (Scanlon et al., 2016; Watkins et al., 2015). For its part, CSR RL06 provided accurate surface-based gridded information that can be used without further processing (Save et al., 2016). On the data, processes such as decorrelation filter, geocenter correction, glacial isostatic adjustment, spatial soothing and leakage correction, coastal resolution improvement (CRI) filters, and ellipsoid landform corrections have been applied (Wiese et al., 2016). On the JPL data, scaling factors were used according to Landerer & Swenson (2012), restoring the original model amplitude. Battery failures in the GRACE mission generated months with gaps starting in 2011.

3.2.2. GLDAS

The Global Land Data Assimilation System (GLDAS) was developed jointly by NASA, Goddard Space Flight Center (GSFC), the National Oceanic and Atmospheric Administration (NOAA), and the National Centers for Environmental Prediction (NCEP). GLDAS provides fields of land surface states and fluxes, integrating satellite- and ground-based observational data products using advanced land surface modeling and data assimilation techniques (Rodell et al., 2004). It includes four Land Surface Models (LSM): Noah, Catchment (CLSM), the Community Land Model (CLM), and the Variable Infiltration Capacity (VIC). There are four versions of the data (1.0, 2.0, 2.1, and 2.2), which have 1.0° or 0.25° spatial resolutions and 3-hourly, daily, and monthly temporal resolutions. This study used 36 hydroclimatic variables of the Noah LSM L4 monthly 0.25 x 0.25° V2.1 dataset. This dataset was selected because of the diversity of hydroclimatic variables, spatial and temporal resolution, and its recurrent use in estimating GWS.

3.2.3. CHIRPS

The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) is a product of the collaboration between the United States Geological Survey (USGS) and the Earth Resources Observation and Science (EROS). CHIRPS provides daily, monthly (every 1, 2, or 3 months), and annual near-global precipitation data and spatial resolution of 0.05° and 0.25°. The dataset builds on previous approaches to interpolation techniques and high-resolution, long periods of record precipitation estimates based on infrared Cold Cloud Duration (CCD) observations (Funk et al., 2015). The primary data used in its creation are 1) the monthly precipitation climatology, CHPClim, 2) quasi-global geostationary thermal infrared (IR) satellite observations, 3) the Tropical Rainfall Measuring Mission (TRMM) 3B42 product from NASA (Huffman et al., 2007; Scheel et al., 2011); 4) atmospheric model rainfall fields from the NOAA Climate Forecast System, version 2 (CFSv2) (Saha et al., 2014; Yuan et al., 2011) and 5) in situ precipitation observations (Funk et al., 2014). This study used precipitation data with a monthly and spatial resolution of 0.05°, obtained from the Climate Hazards Center of the University of California. This dataset assimilates different sources of information from both satellite and field stations, with a high spatial and temporal resolution and good correlation coefficients for the ARB, according to Haghtalab et al. (2020).

3.2.4. GLEAM

The Global Land Evaporation Amsterdam Model (GLEAM) refers to the algorithms used to estimate land evaporation and root-zone soil moisture from satellite information. Since its development in 2011 by Miralles et al. (2011), GLEAM has been regularly revised and optimized, incorporating new geophysical variables observed by satellite (Martens et al., 2017). The latest version of the model (3.6) provides ten different variables at a resolution of 0.25° over two different periods, between 1980 to 2021 (GLEAM v3.6a) and between 2003 to 2021 (GLEAM v3.6b). This study used the Actual Evaporation (E) variable from the GLEAM v3.6a model as part of the data set to perform the *downscaling* model.

3.2.5. MERIT Hydro DEM

MERIT DEM is a Digital Elevation Model (DEM) developed by removing different error components, such as absolute bias, stripe noise, speckle noise, and tree height bias from the existing spaceborne DEMs (SRTM3 v2.1 and AW3D-30m v1) (Yamazaki et al., 2017). MERIT Hydro is a global hydrography dataset derived from MERIT DEM and water body datasets such as G1WBM, GSWO, and OpenStreetMap, which includes Flow Direction, Hydrologically Adjusted Elevations, Upstream Drainage Area, Number of Upstream Drainage Pixels, River Channel Width and Height Above Nearest Drainage (HAND) (Yamazaki et al., 2019). Ground elevation data from the Hydrologically Adjusted Elevations has a spatial resolution of 90 m (0.0008° approx). Slope and aspect values were estimated with support from geographic information systems (GIS) and then transformed to 0.25° by applying an aggregation technique with the average.

3.2.6. RIMAS and SIAGAS

The Integrated Groundwater Monitoring Network (RIMAS) is a Geological Survey of Brazil (SGB-CPRM) project. Since 2009, piezometers have been installed in the principal aquifers of the country, which measure changes in groundwater levels (WL) over time, as well as hydro-chemical parameters. There are currently 409 wells, 69 located within the ARB. The data were used in the validation stage of the results to be compared with anomalies in groundwater storage changes (ΔGWS). The piezometers have a heterogeneous spatial distribution concentrated in some of the main cities, with filtering sections at different depths, as well as temporal information gaps or anomalous records, which are taken into account later. The Geological Survey of Brazil (SGB-CPRM) created the Groundwater Information System (SIAGAS), which consists of an integrated database of exploitation wells throughout the Brazilian territory and is constantly updated. The wells' distribution was used to determine the spatial density in pixels 0.25° . The location and density of SIAGAS wells per unit area (0.25° pixels) were used to analyze results.

3.3. Downscaling Model

3.3.1. Machine Learning Algorithms

Within Artificial Intelligence, Machine Learning (ML) is the discipline that enables a system to learn from data through models. Several learning methods are part of this process, but two supervised learning algorithms are of interest for this study: AdaBoost and Random Forest. The first corresponds to a boosting algorithm, meaning that its operation is sequential to reduce bias, and the second one is a bagging algorithm, which aims to minimize variance. *Downscaling* is a regression problem for which a suitable amount of good-quality data must be available to train the models. The 0.25° data were transformed to 1° (upsampling) for CSR, GFZ, and JPL solutions, using Mean (M), Truncated Mean (MT), and Standard Deviation (SD) as statistics during aggregation. The new data (upsampling at 1°) make up a training set, where each space-time data of GRACE/GRACE-FO of 1° has a data set (of different variables) that describes it and with which each model is trained. The original datasets were separated into test and training data in a 3:7 ratio. The models were then trained and pre-validated, identifying the best statistic for each solution and algorithm. A selection of the variables with the greatest influence on the models was obtained from the analysis between the Pearson correlation coefficients of the independent and dependent variables and the Permutation Feature Importance Technique (PFIT). The hyperparameters were parameters specific to each algorithm, whose adjustment can provide better prediction results. Thus, having established the best statistics for each solution and the variables with the largest generalization capacity, the hyperparameters of each algorithm were adjusted by the trial and error method. Finally, the models were again trained using the totality of the data. Cross-validation by the k-folds method was used during the evaluation of the predictions.

3.3.1.1.1. Random Forest

In 2001, Breiman proposed the Random Forest (RF) model as a new integrated ML algorithm consisting of multiple combinations of Classification and Regression Trees (CART) (Breiman, 2001). This algorithm can be used for both classification and regression problems. Within Machine Learning (ML), RF is a supervised learning algorithm, constituting an ensemble method, i.e., it combines the output of multiple decision trees to obtain an average result. The algorithm generates random subsets of features, reducing the correlation between decision trees and improving the model's accuracy. Three of its main hyperparameters are the number of estimators, the maximum features, and the minimum sample split. Bootstrapping, an iterative resampling technique used to randomly extract multiple samples for training, is applied to the training data, and the remaining samples are grouped as data out of a bag and used to test the model's accuracy. It is a widely used algorithm in GRACE/GRACE-FO *downscaling* (Agarwal et al., 2023; Chen et al., 2019; Seyoum et al., 2019). Some of its advantages are the low possibility of overfitting, detection of the interactions and importance of the variables, methods incorporation to estimate missing values, and stability when using multiple samples, prevailing the average of the voting of the decision trees. RF disadvantages are high training times and poor performance with small datasets.

3.3.1.2. Adaptive Boosting

In 1995, Freund and Schapire proposed the Adaptive Boosting or AdaBoost (AB) model as an ensemble method that creates several simple sequential predictors from a common algorithm (Freund & Schapire, 1995). These predictors are trained with the same data set, receiving the weights derived from the errors made and improving consecutively. In the end, when all predictors have been trained, AB performs a final prediction by estimating a weighted vote or sum to each predictor according to a relative weight (Elith et al., 2008). This algorithm can be used for both classification and regression problems. Some of its main hyperparameters are the number of estimators, the learning rate, and the regression loss function. AB is a simple algorithm that is easy to program and flexible, with high generalization capacity, low error rate, and high accuracy. However, the influence of some weak predictors at outliers can produce overfitting, making it vulnerable to noise.

3.3.2. Downscaling Model Design

AdaBoost (AB) and Random Forest (RF) were defined as the algorithms for nonlinear statistical *downscaling* models for the TWS data. The *downscaling* model design was the product of applying the methodological proposals of Chen et al. (2019) and Ali et al. (2021), extending the training datasets, performing a sensitivity analysis between variables and models, varying the aggregation strategies, and exploring the hyperparameters, aiming to compare both ML models. Initially, the statistical relationships between the hydrometeorological and topographic variables were defined for the Δ TWS of each solution (CSR, GFZ, and JPL) within the ARB at a resolution of $1^\circ \times 1^\circ$. Defining the Δ TWS as the dependent variable and the other variables as independent, their relationships were established in $1^\circ \times 1^\circ$ from the ML models and then applied to the $0.25^\circ \times 0.25^\circ$ variables, obtaining a high-resolution TWS. Fig. 2 presents the flowchart of the *downscaling* process.

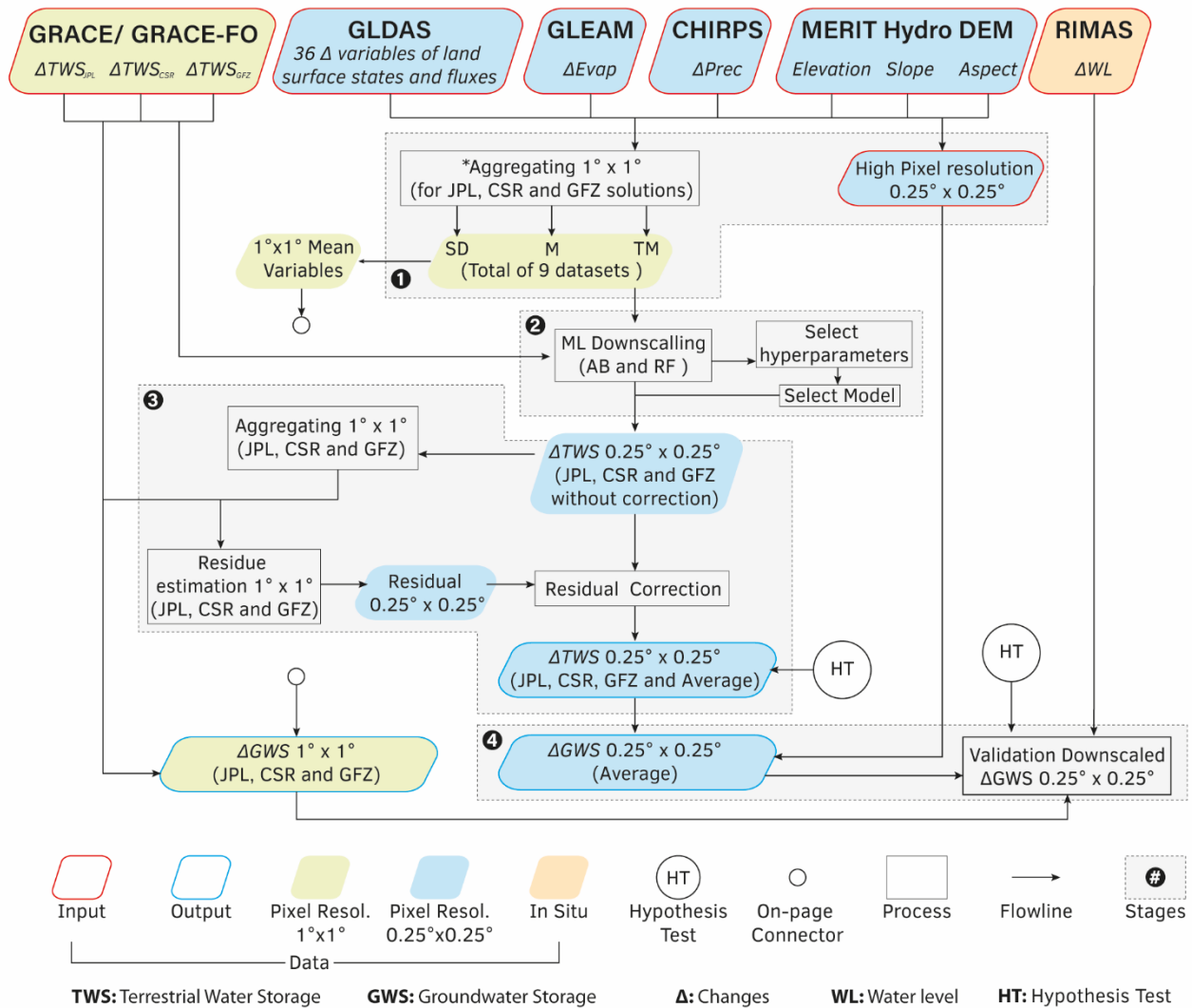


Fig. 2. Flowchart of the *downscaling* model applied in this study.

1) All independent variables (41 in total, where 36 of GLDAS, 1 of GLEAM, 1 of CHIRPS, and 3 of DEM) within the ARB between March 2002 and December 2021 were resampled from their original resolutions to 1° according to the different pixel solution of the CSR, GFZ, and JPL (Fig. S2) programs. The resampling utilized three aggregation strategies, the Standard Deviation (SD), the Mean (M), and the Truncated Mean (TM), obtaining a total of 3 resamples for each ΔTWS solution, with a total of 9 training datasets.

2) During this stage, spatial (ID_Spatial) and temporal (ID_Month) indices were included to facilitate the subsequent interpretation of the time series and their spatial relationship. The models were trained and evaluated using AB and RF algorithms with each of the nine datasets with default hyperparameters, using 70% of the data for training and 30% for validation. The best-performing dataset from each solution (CSR, GFZ, and JPL) was subjected to hyperparameter tuning for both AB and RF, selecting the best models per solution. From a multiplicative analysis of the variables, based on the correlation

coefficients and the Permutation Feature Importance Technique (PFIT), the most influential variables in the model were defined with which a final model was trained, using the totality of each dataset's data.

3) The ΔTWS derived from the AB and RF models with 0.25° resolution were resampled again to 1° with their respective solutions. The residuals between the original and derived ΔTWS at 1° were estimated. Finally, the residuals were added to the 0.25° derived ΔTWS (residual correction). This residual correction ensured that the reduced ΔTWS matched the original data (GRACE/GRACE-FO derived) and corrected for the prediction bias that could result from omitted variables. From the resulting three datasets (CSR, GFZ, and JPL), an AVERAGE ΔTWS was obtained.

4) The ΔGWS for the four resulting datasets (CSR, GFZ, JPL, and AVERAGE) was obtained by the terrestrial water balance method (Equation (5)) using the variables with high pix resolution (0.25°). The resulting ΔGWS time series were compared and validated with the water level anomalies (ΔWL) of RIMAS.

3.3.3. Model performance assessment

The evaluation of the results of the *downscaling* models (predicted data) applied three different metrics: Root-Mean-Square Error (RMSE) [Equation (1)], Mean Absolute Error (MAE) [Equation (2)] and the Pearson correlation coefficient (R) [Equation (3)]. RMSE and MAE values close to 0 represent less difference between the GRACE-derived and predicted data, i.e., higher model accuracy. R varies between -1 and +1, measuring the statistical relationship or association between two variables; values close to ± 1 mean a perfect linear relationship.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}} \quad (1)$$

$$MAE = \frac{1}{n} (\sum_{i=1}^n |X_i - Y_i|) \quad (2)$$

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3)$$

Where,

X_i : original or input value of ΔTWS

Y_i : value obtained by the model

X and Y: average value of each independent variable of ΔTWS

n: total number of samples

Two significance tests were applied using the student's t-test method. The first one was used to compare the time series averages between the scaled pixels and the originals in the different solutions, identifying pixels with significant differences in the mean as a product of the *downscaling*. A second correlation

test between the time series of the RIMAS monitoring wells and the ΔGWS , estimating Pearson correlation coefficients (R) and critical values. The significance levels defined were 0.05 (5%) and 0.01 (1%).

3.3.4. Characteristic Analysis

Reducing the number of training variables decreases computational costs, prevents model overfitting, enhances result interpretability, reduces data noise, and improves generalization capacity. Pearson's correlation coefficients and the Permutation Feature Importance technique (PFIT) purpose is to evaluate the relationship between the independent and dependent variables and their relative contribution to the model and its predictions. The spatiotemporal correlation matrices between variables were obtained after aggregation to a resolution of 1° for the three models resulting from each solution (CSR, GFZ, and JPL). PFIT was applied to the models, which assigned scores to each variable according to its relative ability to improve predictions. PFIT scores are relative, not absolute, estimating the difference in the metrics obtained between the predictions of a model with the data of each variable organized and randomly distributed. The results of these two methods were the variables with the most significant relationship to ΔTWS and their contribution to each model. With these results, training variables are identified and reduced based on their relative contribution to the models, thereby reducing computational costs and improving prediction accuracy.

3.4. Groundwater Storage Anomaly Estimation

The terrestrial water storage (TWS) is the sum of the different terrestrial water storage components constituted by soil moisture, surface runoff, canopy water, and groundwater, as shown in Equation (4). The groundwater term (GWS) in the water balance method is rearranged in Equation (5).

$$\Delta TWS = \Delta GWS + \Delta SMS + \Delta SWS + \Delta CWS + \Delta Qs \quad (4)$$

$$\Delta GWS = \Delta TWS - (\Delta SMS + \Delta SWS + \Delta CWS + \Delta Qs) \quad (5)$$

Where,

ΔTWS : terrestrial water storage changes or anomalies obtained from GRACE and GRACE-FO

ΔGWS : groundwater storage changes or anomalies

ΔSMS : soil moisture storage changes or anomalies obtained from GLDAS

ΔSWS : snow storage changes or anomalies obtained from GLDAS

ΔCWS : canopy water storage changes or anomalies obtained from GLDAS

ΔQs : surface runoff changes obtained from GLDAS

ΔGWS was obtained after *downscaling* the data to 0.25° to validate and compare with the estimates at 1° . The storage components are expressed in volume with units of km^3 and mm of equivalent water height. Within the analysis period (March 2002 to December 2021), there were both spatial and temporal

information gaps in the datasets, which were estimated from the average of nearby pixels (spatial) and with the use of ML methods (temporal).

GLDAS (SWS, CWS, SMS, and Qs) data should be transformed to changes or anomalies, estimated from the same baseline with which GRACE data were calculated (January 2004 to December 2009) (Cooley et al., 2019). The estimation of anomalies at time t was calculated using Equation (6):

$$\Delta\text{Var}(t,i,j)= \text{Var}(t,i,j)-\text{Var}(i,j)_{2004-2009} \quad (6)$$

Where,

Var: variables to be transformed to anomalies (Δ)

t : time

i and j : pixel coordinates

$\text{Var}(i,j)_{2004-2009}$: average between January 2004 and December 2009

3.5. Time series decomposition

The time series of ΔTWS and ΔGWS [$h(t)$] were decomposed into their trend [$T(t)$], seasonal [$S(t)$], and residual [$R(t)$] components applying the nonparametric approach STL (Seasonal Trend decomposition using Loess) from Cleveland et al., (1990), as described in Equation (7).

$$h(t)=T(t)+S(t)+R(t) \quad (7)$$

The trend $T(t)$ or long-term signal can be further divided into linear trend [$T_{\text{lin}}(t)$] and the anomalies to this linear trend, named interannual variability [$T_{\text{inter}}(t)$].

Long-term trends were interpreted after removing the seasonal effect (deseasonalization) from the ΔTWS and ΔGWS time series. The Seasonal Adjustment $h_{\text{SA}}(t)$ was obtained from the additive model described in Equation (8) by subtracting the seasonal $S(t)$ value from $h(t)$, as shown in Equation 5.

$$h_{\text{SA}}(t)=h(t)-S(t)=T(t)+R(t) \quad (8)$$

4. Results

4.1. Correlation of Input Variables

All variables were analyzed using a correlation matrix. This procedure allowed establishing which variables had strong linear associations with ΔTWS derived from GRACE/GRACE-FO and identifying regional relationships with statistical significance ($p > 0.05$). Correlation coefficients were estimated for the entire ARB and each solution (CSR, GFZ, and JPL) at a resolution of 1° (aggregate) during the months with GRACE/GRACE-FO data (204 of 237).

The results were similar despite the differences in the area and pixel distribution of each ΔTWS solution. The absolute averages of each matrix were 0.36 (GFZ), 0.36 (JPL), and 0.35 (CSR). Between ΔTWS and the variables, the highest positive correlations were with Soil moisture (SM100-200cm), Baseflow-groundwater runoff (Qsb), and Root zone soil moisture (RootMoist) with average values of 0.52, 0.50, and 0.47, respectively, and negative with Potential evaporation rate (PotEvap) of -0.43. This represented a strong positive relationship between stored water contents in soil and groundwater with the ΔTWS signal for the whole ARB. Some of the variables that had less association with ΔTWS were Aspect, Elevation, Slope, Snow depth (SnowDepth), Snow depth water equivalent (SWS), Snow melt (Qsm), Snow precipitation rate (Snowf), Surface pressure (Psurf) and Transpiration (Tveg). The correlation coefficients for Slope, Elevation, SWE, and SnowDepth were not statistically significant. The correlation between ΔTWS and Soil Moisture (SM) increased with depth, ranging from 0.43 to 0.60 for depths of 0-10cm and 100-200cm, respectively. Soil moisture (SM) in the shallower portions (0-40cm) presented a greater sensitivity to climatic and energetic variables such as Plant canopy surface water (CanopInt), Canopy water evaporation (ECanop), Evapotranspiration (Evap), Evapotranspiration (Evap_GLEAM), Net long-wave radiation flux (Lwnet) and Specific humidity (Qair). Soil moisture (SM) decreased with increasing Potential evaporation rate (PotEvap), Sensible heat net flux (Qh), Downward short-wave radiation flux (SWdown), Net short-wave radiation flux (Swnet), and Wind speed (Wind). As expected, the greater the depth in the soil, the higher the correlation between moisture and Baseflow-groundwater runoff (Qsb). Concerning precipitation from CHIRPS and GLDAS, they presented moderate correlations with ΔTWS (0.28) and strong correlations with soil moisture (0.48 - 0.67). The latter may be due to the temporal delay between rainfall and aquifer recharge, related to the permeability of the vadose zone. The correlation matrix for the CSR solution is presented in Fig. S3 as an example due to its slight average superiority to the other solutions.

4.2. Feature of Input Variables

The results of the models varying the aggregation function are available in Table S1. According to the highest R and lowest MAE and RMSE, the best results for the GFZ and JPL solutions were obtained using the truncated mean (MT), unlike CSR, in which the mean (M) was slightly superior. This slight superiority in generalization ability was due to the reduction in the effect of outliers in the predictions. The AB models obtained better metrics for RF by approximately 2% in R, -19% in MAE, and -18% in RMSE. Considering both AB and RF models, GFZ solution with MT statistics obtained the best results, whereas JPL solution with DS statistics was the worst.

On these three resulting datasets (CSR_M, GFZ_MT, and JPL_MT), the Permutation Feature Importance Technique (PFIT) was applied. Table 2 summarizes the results and correlation coefficients.

Table 2. Pearson's correlation coefficients and PFIT of the hydrometeorological and topographic variables according to each solution and algorithm. The first column shows the solution used for each variable, numbered 1) CSR_M, 2) GFZ_MT, and 3) JPL_MT.

Solution	Pearson CC with ΔTWS			PFIT						
	CSR	GFZ	JPL	CSR	AB GFZ	JPL	CSR	RF GFZ	JPL	
123	ID_Month	0.072	0.075	0.063	0.0023	0.0028	0.0020	0.0236	0.0252	0.0230
123	ID_Spatial	-0.012	-0.021	-0.024	0.1503	0.1563	0.1763	0.2161	0.2535	0.2745
	Albedo	-0.037	-0.032	-0.035	0.0015	0.0050	0.0040	0.0211	0.0146	0.0201
	AvgSurfT	-0.087	-0.121	-0.102	0.0006	0.0003	0.0005	0.0127	0.0062	0.0089
23	CanopInt	0.302	0.318	0.31	0.0004	0.0002	0.0003	0.0034	0.0028	0.0038
	ECanop	0.235	0.232	0.239	0.0004	0.0002	0.0002	0.0043	0.0031	0.0042
	ESoil	0.14	0.139	0.14	0.0011	0.0023	0.0006	0.0073	0.0152	0.0071
	Evap	0.217	0.22	0.207	0.0001	0.0000	0.0000	0.0004	0.0003	0.0004
	LWdown	0.048	0.047	0.054	0.0008	0.0006	0.0005	0.0075	0.0082	0.0072
12	Lwnet	0.343	0.382	0.351	0.0017	0.0022	0.0011	0.0096	0.0109	0.0073
	PotEvap	-0.416	-0.468	-0.425	0.0005	0.0005	0.0005	0.0038	0.0070	0.0041
123	Psurf	0.008	0.005	0.01	0.0432	0.0345	0.0437	0.0642	0.0354	0.0963
	Qair	0.189	0.199	0.194	0.0020	0.0019	0.0016	0.0092	0.0210	0.0193
123	Qg	-0.239	-0.283	-0.245	0.0026	0.0020	0.0026	0.0145	0.0287	0.0157
123	Qh	-0.301	-0.32	-0.291	0.0019	0.0013	0.0022	0.0108	0.0107	0.0145
123	Qle	0.218	0.22	0.208	0.0058	0.0056	0.0040	0.0299	0.0180	0.0213
13	Qs	0.327	0.343	0.316	0.0004	0.0002	0.0003	0.0032	0.0029	0.0039
123	Qsb	0.498	0.529	0.497	0.0945	0.3114	0.1831	0.5052	0.4225	0.5083
	Qsm	0.005	0.007	0.009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Rainf	0.28	0.294	0.287	0.0001	0.0001	0.0001	0.0016	0.0014	0.0013
	RainfT	0.28	0.295	0.288	0.0001	0.0001	0.0001	0.0017	0.0015	0.0014
	RootMoist	0.461	0.493	0.46	0.0012	0.0005	0.0007	0.0096	0.0079	0.0104
	SnowDepth	-1E-04	0.002	0.004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Snowf	0.004	0.007	0.009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	SM0-10cm	0.433	0.458	0.428	0.0002	0.0001	0.0001	0.0018	0.0026	0.0040
	SM10-40cm	0.475	0.502	0.468	0.0002	0.0001	0.0001	0.0014	0.0015	0.0023
123	SM40-100cm	0.548	0.583	0.541	0.0028	0.0015	0.0086	0.0064	0.0066	0.0166
123	SM100-200 cm	0.598	0.631	0.591	0.2468	0.0312	0.1545	0.1936	0.1019	0.1389
	ST0_10cm	-0.08	-0.112	-0.094	0.0006	0.0008	0.0006	0.0098	0.0163	0.0112
12	ST10_40cm	-0.058	-0.082	-0.067	0.0028	0.0520	0.0099	0.0805	0.1143	0.0984
13	ST40_100cm	-0.046	-0.066	-0.053	0.0212	0.0037	0.0257	0.0126	0.0060	0.0112
3	ST100_200cm	-0.03	-0.044	-0.034	0.0061	0.0023	0.0264	0.0228	0.0119	0.0330
	SWdown	-0.263	-0.307	-0.274	0.0003	0.0001	0.0002	0.0038	0.0041	0.0032
123	SWE	-2E-04	0.002	0.004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
	Swnet	-0.264	-0.307	-0.273	0.0003	0.0001	0.0001	0.0034	0.0025	0.0028
	Tair	-0.116	-0.133	-0.114	0.0010	0.0016	0.0004	0.0081	0.0080	0.0064
	Tveg	0.018	0.022	0.003	0.0036	0.0035	0.0034	0.0307	0.0224	0.0295
2	Wind	-0.133	-0.148	-0.141	0.0012	0.0086	0.0007	0.0083	0.0272	0.0080
123	Precip_CHIPRS	0.277	0.296	0.286	0.0006	0.0002	0.0004	0.0056	0.0043	0.0053
123	Evapo_GLEAM	0.135	0.139	0.139	0.0010	0.0009	0.0009	0.0079	0.0088	0.0101
123	Slope	-0.003	0.003	-0.004	0.0083	0.0092	0.0082	0.0559	0.0520	0.0631
123	Elevation	-0.002	0.002	-0.004	0.0685	0.0159	0.0583	0.1753	0.0845	0.1195
123	Aspect	-0.015	-0.021	-0.009	0.0009	0.0165	0.0016	0.0238	0.0537	0.0168

According to the PFIT scores in the RF models, more variables were sensitive to change their generalizability relative to AB. In both algorithms AB and RF, the variables Baseflow-groundwater runoff (Qsb) and Soil moisture (SM100-200cm) were the only ones with high Pearson and PFIT values. Other variables with a high relative contribution to the different models were Elevation, Surface pressure (Psurf), and Soil temperature (ST10_40cm, ST40_100cm, and ST100_200cm), despite not having high Pearson values with ΔTWS .

RF and AB models with the same solution (CSR, GFZ, or JPL) used the same variables. According to the results, 22 of the 41 original variables were selected. For each model, 19 variables were used, 15 equal between solutions. The variables Precip_CHIRPS, Evapo_GLEAM, and SWE_inst were included in all models due to their hydrological relevance. There were seven variables that differentiate the training data between models.

4.3. Models accuracy

The hyperparameters are parameters of each algorithm whose adjustment can provide better prediction results. The AB and RF algorithms were optimized by manipulating some of their hyperparameters, establishing maximum and minimum values, and evaluating their results through the trial and error method. To obtain the best results, for AB, the hyperparameters were a number of estimators of 500, a learning rate of 1.0, and a loss function of regression square. In the case of RF, the number of estimators was 100, the maximum features were all, and the minimum sample split was eight. Table S2 summarizes the results of the models with the adjustment of the hyperparameters.

The predictions were evaluated by cross-validation by the k-folds method on a total of 6 models. A comparison between the GRACE-derived Δ TWS values and an average of those obtained for each solution is presented in the scatter plots in Fig. 3.

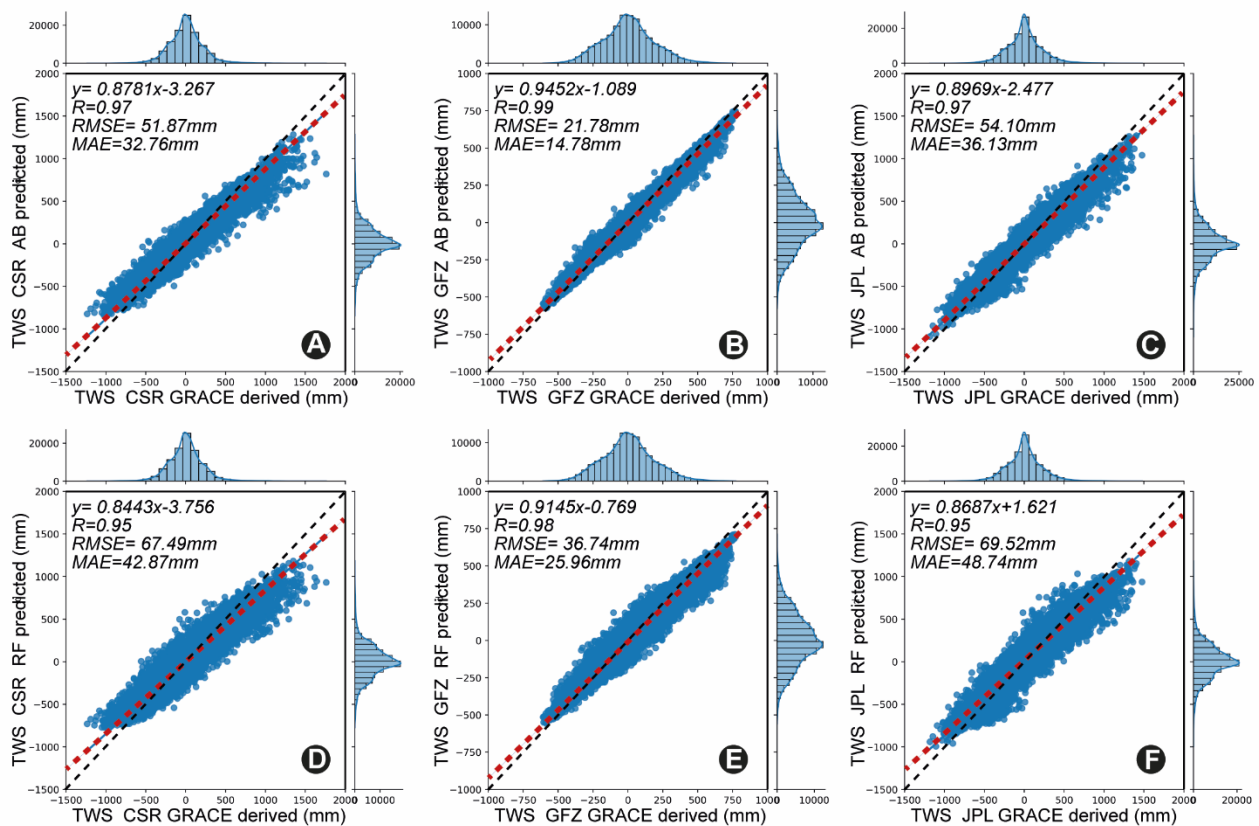


Fig. 3. Scatter plots of between GRACE/GRACE-FO derived Δ TWS values and model predictions for a) CSR with AB, b) GFZ with AB, c) JPL with AB, d) CSR with RF, e) GFZ with RF, and f) JPL with RF.

GRACE-derived Δ TWS values on the X-axis, and the model predicted data of Δ TWS on the Y-axis. All datasets' histograms and density functions tended to be symmetric or slightly positive. The CSR (Figure 3 a,d) and JPL (Fig. 3 c,f) solutions were leptokurtic, unlike GFZ (Fig. 3 b,e), which was mesokurtic, indicating a better distribution between values and a lower presence of extreme values. As for the models, the best-performing solutions were GFZ, followed by CSR and JPL. Of the models obtained with the AB algorithm (Figure 6a,b,c), the best results were GFZ with a correlation of 0.99, an RMSE of 2.178mm, and an MAE of 1.478mm, followed by CSR and slightly below JPL. In the case of the RF models (Fig. 3 d,e,f), the best results were GFZ with a correlation of 0.98, an RMSE of 3.674cm, and an MAE of 2.596cm. That is to say that the AB models, on average, had a higher R of 0.02, a lower RMSE of 1.53cm, and a lower MAE of 1.13cm for the RF models, indicating a slight superiority in the AB algorithm in the ability to predict Δ TWS values for the ARB. A new average dataset was obtained from the three predictions of each solution by the AB models, which was analyzed and validated. The information increase resulting from *downscaling* introduced derived features from the training variables, leading to greater spatial and temporal variance in TWS and, consequently, in GWS. As shown in Fig. S4, after *downscaling*, spatial variance increases, causing higher values for the same period. New spatial patterns become apparent, enabling the recognition of novel data characteristics. The average dataset aggregates the variance increments from various solutions, enhancing spatial patterns.

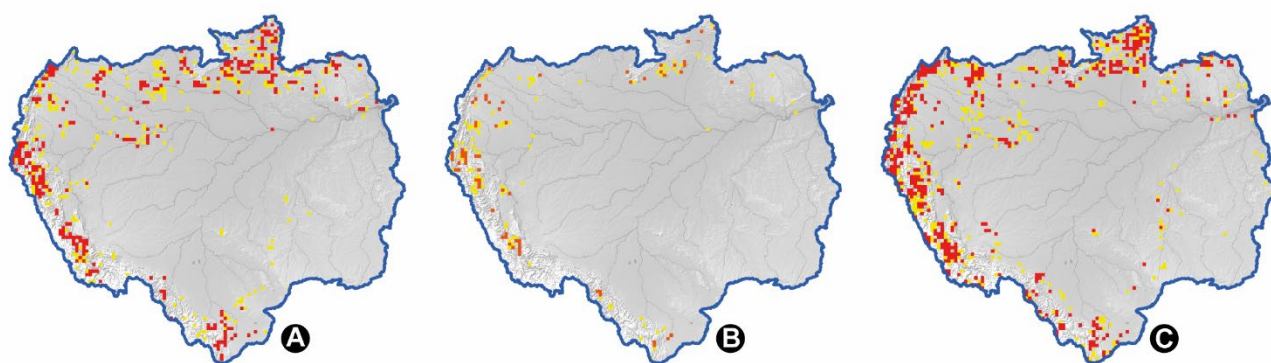


Fig. 4. Pixels whose time series differ from the mean before and after scaling for: a) CSR, b) GFZ, and c) JPL for the AdaBoost models. In yellow, pixels have confidence levels of 5%, and in red, 1%.

Hypothesis testing for the mean between the TWS time series of the GRACE/GRACE-FO derived data and the data predicted by the ML models for each pixel of each solution allowed to identify those pixels and areas where this gain in spatial and temporal variance has generated statistically significant changes (Fig. 4). In the three solutions (CSR, GFZ, and JPL), the pixels where the null hypothesis was not accepted, in other words, the mean value of their time series before and after *downscaling* was different, were distributed mainly over the Andes and North of the ARB. Some isolated pixels can also be identified, such as one located over Manaus, the only one with statistical significance in the presence of RIMAS monitoring wells. The solution with the lowest number of pixels with statistical significance was GFZ, followed by CSR and JPL. These changes in the time series resulted from the increase in

spatial variance caused during *downscaling*, broadening the data representation and allowing the identification of new local effects in the Δ TWS and GWS signals.

4.4. Δ GWS data validation

From highest to lowest, Pearson correlation coefficients between the time series of the average Δ WL measured by the RIMAS monitoring wells and the downscaled Δ GWS anomalies were obtained in the East, Central, South, North, and West zones with maximum values of 0.79, 0.76, 0.73, 0.58 and 0.23, respectively. This showed that the time series resulting from Δ GWS *downscaling* had a statistically significant correlation with the monitoring wells in the different zones of the ARB. An overlay and spatial comparison between the Δ GWS time series and the Δ WL recorded by some of the RIMAS monitoring wells is presented in Fig. 5.

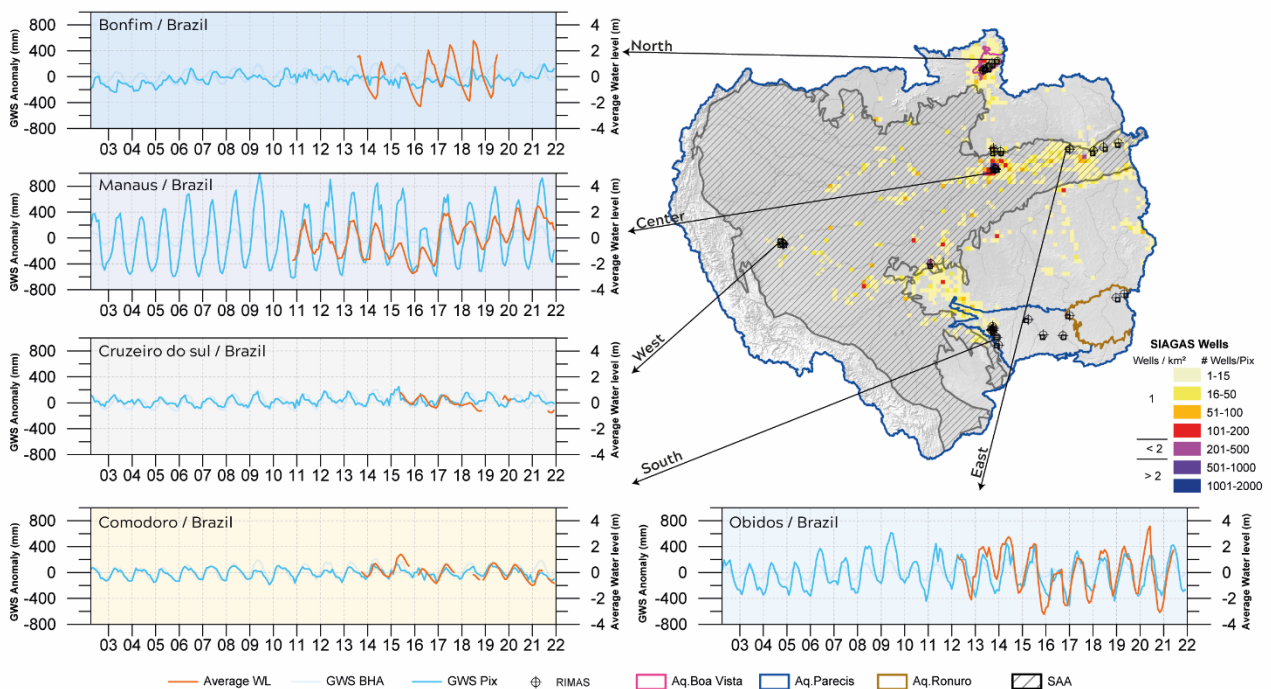


Fig. 5. Comparison between the Δ GWS AVERAGE time series and the RIMAS Δ WL for pixels in Bonfim, Manaus, Cruzeiro do Sul, Comodoro, and Obidos. The density of SIAGAS exploitation wells per unit area (pixel 0.25°) are in color scales from yellow to purple. The background of the time series corresponds to blue for a positive trend (surplus), yellow for a negative trend (deficit), and gray for a neutral trend or constant behavior. The AAS is in a dashed line with a black outline.

Most of the exploration wells present in the SIAGAS database in the study area are found in the SAA, Parecis, Ronuro, and Boa Vista aquifers. High concentrations per unit area were appreciable in different cities, Manaus being the most relevant, with a number of exploitation wells in the range of 1,001 to 1,700 for an area of approximately 770 km^2 (pixel of 0.25°). This could have had implications in the Δ WLs, such as discrepancies with the regional behavior of the time series or stronger local depletions that could interfere with the monitoring well logs, although a coherent behavior could be seen with the scaled Δ GWS data and the ARB average values. The variability between the scaled Δ GWS time series and the Δ WLs were similar, high in areas such as the Center and East with SD between 422-235mm and

1.6-1.2m, respectively, low in the West and South with SD values of 71-64mm and 0.5-0.4m. In the North, this variability was slightly different, with high variations of the WL (1.32m) compared to the GWS anomalies (82mm), which in turn presented an irregular seasonality. The long-term behavior of the periods in which they coincided was comparatively similar, with positive trends in the North and Central zones, constant in the West and East, and negative in the South. Fig. 6 shows the spatial representation of the correlation coefficients and their statistical significance.

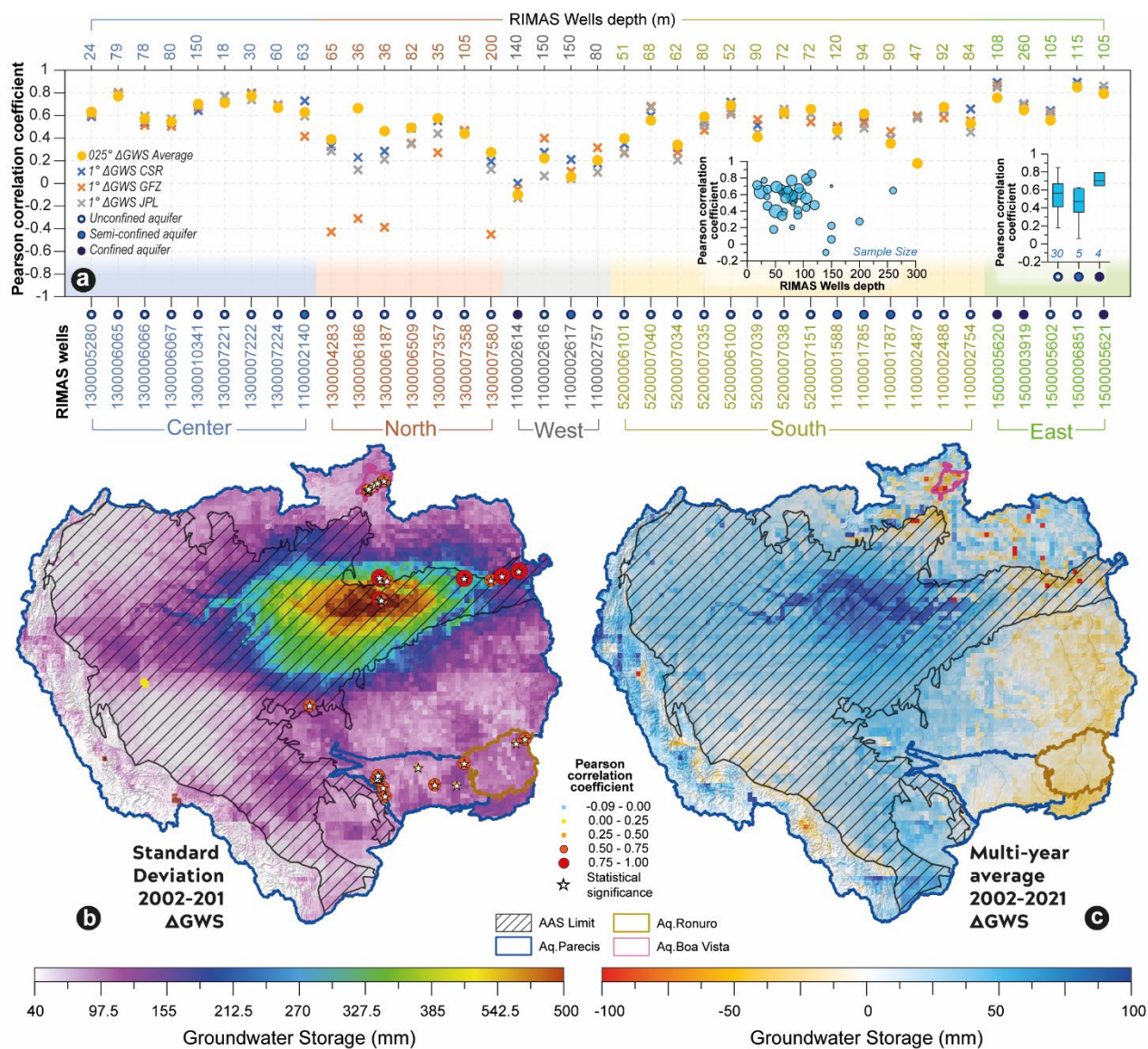


Fig. 6. Spatialization of the correlation coefficients, where a) comparison of the correlation coefficients for the RIMAS well zones, their depths, and aquifer types; b) standard deviation of the time series of ΔGWS , together with the correlation coefficients of the RIMAS wells and their statistical significance; and c) a multiyear average of ΔGWS between 20002 and 2021.

Most RIMAS wells are located at depths shallower than 120m and within locally unconfined aquifers. The scatter plot in Fig. 6a shows higher correlations with shallow wells; nevertheless, in the eastern zone, a 260m deep well within a confined aquifer recorded a correlation coefficient of 0.65. The number of samples and their distributions are unclear enough to differentiate the behavior between RIMAS wells

installed in unconfined, semi-confined, and confined aquifers. Regarding the the standard deviation (SD) of the time series per pixel of ΔGWS , the highest magnitudes are observed over the Alter do Chão and Içá aquifers and alluvial deposits (AAS). Spatial patterns related to geology, topography, and the influence of surface water bodies can be appreciated.

4.5. Spatio-temporal characterization of ΔGWS

The temporal decomposition of the ΔGWS signal into its seasonal, trend, and residual components provides a deeper understanding of water storage variations over short and long time scales. By isolating the seasonal component, variations linked to recurrent natural phenomena occurring throughout the year (short-term) become evident. The trend component imparts significant insights into medium-term changes associated with low-frequency seasonal effects such as El Niño and La Niña. Meanwhile, the linear trend quantifies long-term changes, offering valuable information about growth or decline patterns.

The spatial representation of ΔGWS anomalies (Fig. 7) reveals maximum and minimum extremes in the central region of the ARB, specifically in the Alter do Chão and Içá aquifers (for TWS see Fig. S6). During positive peaks, spatial patterns emerge in the floodplains of the Mamirauá Natural Reserve and the Purus, Madeira, and Tapajós rivers, underscoring the substantial influence of changes in surface water storage on ΔGWS . Positive anomalies prevail during the months of MAMJ, while negative anomalies occur during SOND. In the Northern region, anomalies exhibit maximum negatives during DJFM and maximum positives during JJAS, with a phase lag of approximately 2 to 3 months compared to the Southern region of the ARB. These storage patterns are influenced by seasonal variations in precipitation, which, in turn, are linked to the Intertropical Convergence Zone (ITCZ) migration and the South American monsoon system.

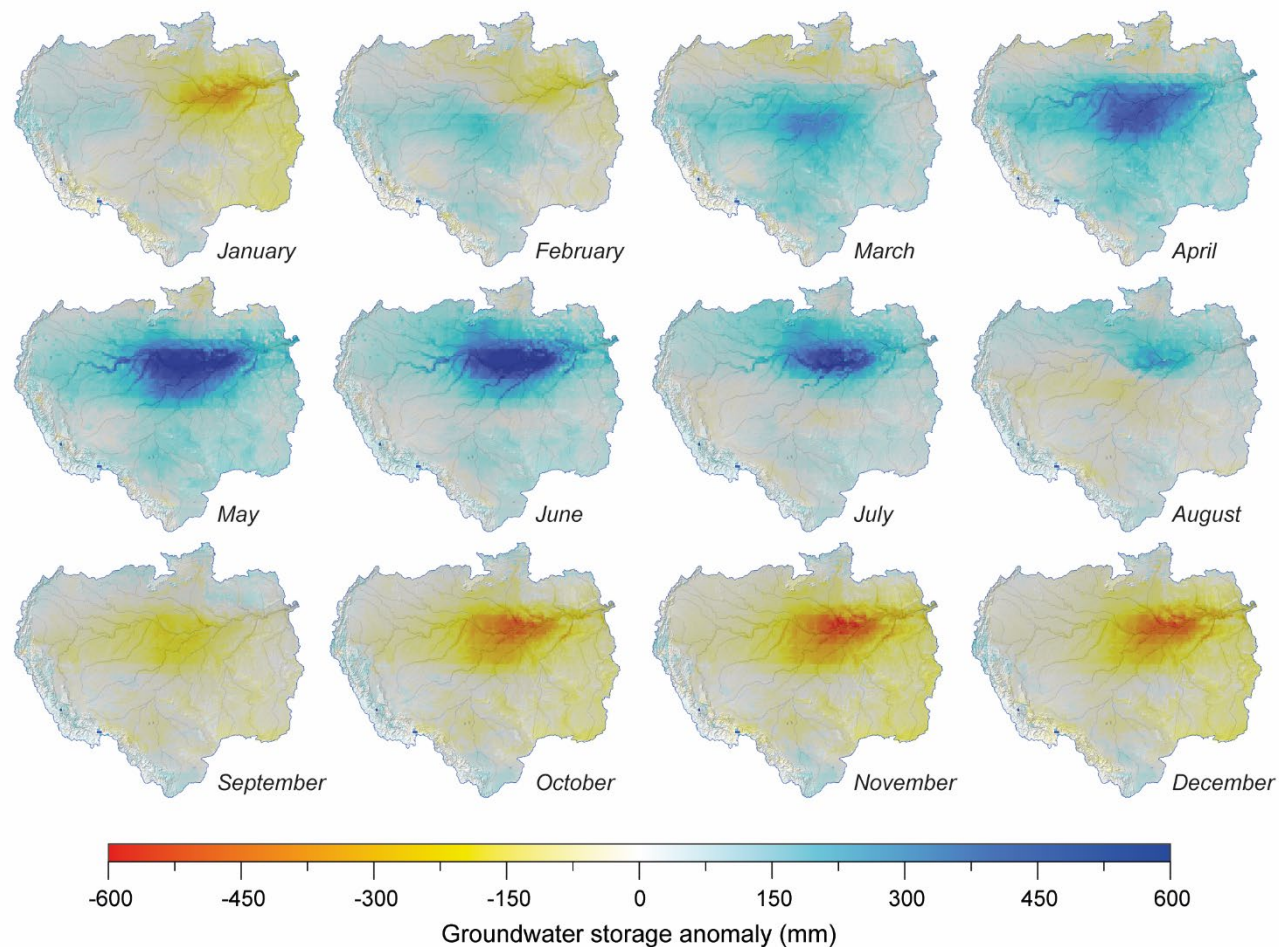


Fig. 7. Spatio-temporal distribution of multiyear seasonality of ΔGWS in the ARB between 2002 and 2021.

The ΔGWS time series in the different solutions and at 1° and 0.25° resolutions showed a positive linear trend, superior to those of the ΔTWS . The average value between April 2002 and December 2021 for a spatial resolution at 1° was $+21.49 \pm 0.85 \text{ km}^3/\text{year}$ and at 0.25° , $+22.24 \pm 1.18 \text{ km}^3/\text{year}$. At both resolutions, the solution with smaller linear trends was GFZ. After *downscaling*, all the solutions increased the spatial variance and their linear trend by 6.5% for CSR, 1.6% for GFZ, and 2.1% for JPL. During the time analyzed, the ARB would have an average total gain of $432.94 \pm 26.7 \text{ km}^3$ of stored groundwater. Table 3 summarizes the linear trends of the ΔGWS for the ARB (for TWS see Table S3).

Table 3. Linear trends of ΔGWS in the ARB for the different solutions and resolutions between 2002 and 2021.

Resolution	Period	Solution	Area_ARB (km ²)	Rate (mm/year)	Rate (km ³ /year)	Total (km ³)
1°	2002-2021	CSR	6089336.3	+3.575	+21.772	+429.996
	2002-2021	GFZ	6089336.3	+3.367	+20.500	+404.870
	2002-2021	JPL	6089336.3	+3.647	+22.209	+438.623
0.25°	2002-2021	CSR	6089336.3	+3.811	+23.204	+458.280
	2002-2021	GFZ	6089336.3	+3.422	+20.836	+411.517
	2002-2021	JPL	6089336.3	+3.725	+22.685	+448.031
	2002-2021	Prom	6089336.3	+3.653	+22.242	+439.276

The El Niño events of 2010, 2015, and 2016 strongly negatively influenced the GWS anomalies. Fig. 8 shows the time series of the average Δ GWS values for the ARB (for TWS in Fig. S5). The standard deviation and interquartile ranges tended to be higher during extreme events; however, a widening was evident from 2012 onwards. This represented a greater spatial variance between positive and negative values. On the other hand, the decrease in SWS in snowy areas in southern Peru translated into significant positive Δ GWS anomalies. These changes were not necessarily representative since a portion of these melted masses reached the rivers, with respect to the portion infiltrated and effectively converted into recharge or Δ GWS.

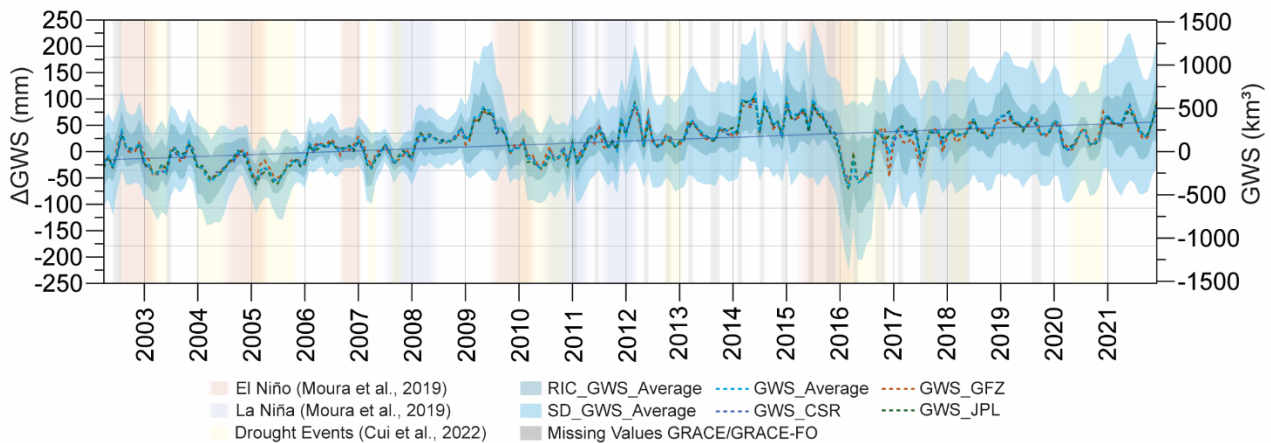


Fig. 8. Comparison of the Seasonal Adjustment components in the Δ GWS time series between the original data derived from GRACE/GRACE-FO and the AVERAGE scaled with AB.

The spatiotemporal distribution of the average ARB values is presented in Fig. 9 (for TWS see Fig. S7). Globally, a positive trend was evident over the lower basin towards the central, eastern, and southern zones within the AAS and its margins. Over the Andes, some pixels can be observed South of Peru, especially in the Vilcanota mountain range in the Ausangate, Suyuparina, Quisoquipina, and Quelccaya snow-capped mountains, where the strong influence of the SWS descent generated strong positive anomalies in the GWS. Negative anomalies developed over the Alter do Chão, Parecis, Ronuro, and Boa Vista aquifers and the transnational zones between the Andes and the lower basin, East of Peru and Ecuador.

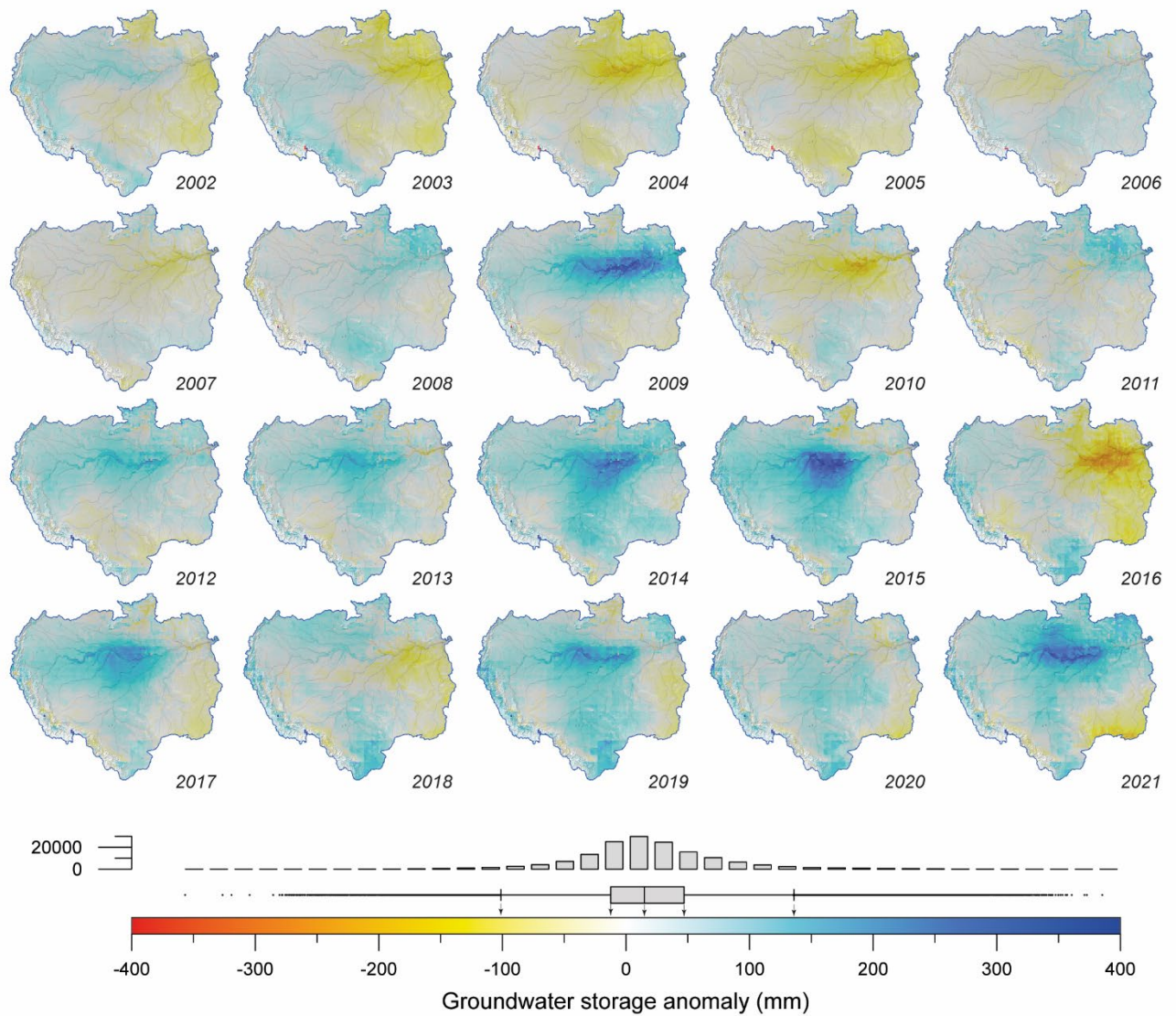


Fig. 9. Spatiotemporal distribution of average annual GWS anomalies within the ARB between 2002 and 2021.

4.6. ΔTWS and scaled ΔGWS comparison

Correlation coefficients were estimated between the ΔWL time series of the RIMAS monitoring wells with ΔTWS and ΔGWS at resolutions of 0.25° with the AVERAGE dataset and 1° with the GFZ solution. The time series of the RIMAS pixels were the average of the ΔWL anomalies of the wells contained in that pixel, and the RIMAS wells were by analyzing each well individually. Fig. 10 presents the boxplots of the Pearson values of the ΔGWS and ΔTWS time series at 0.25° and 1° (GFZ) with the RIMAS Wells and Pixel averages.

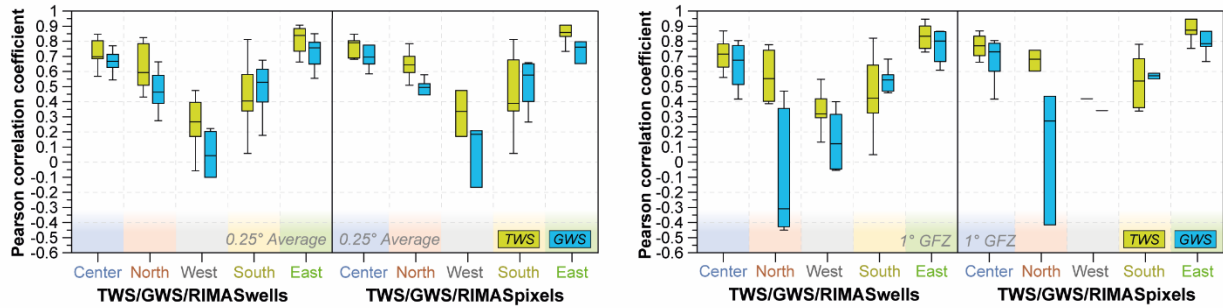


Fig. 10. Boxplots of the Pearson correlation coefficients between the time series of the Δ WL of the RIMAS monitoring wells and the Δ GWS and Δ TWS anomalies for the different zones of the ARB. The average values of the wells contained per pixel and the individual values of each well are compared on the left with the anomalies of the AVERAGE dataset at 0.25° and on the right with GFZ at 1° .

The correlations with the Δ WL RIMASwells were similar at both resolutions, being more accurate at 0.25° and particularly with the GWS anomalies. Something similar occurred with the pixel averages of the wells, where the differences were smaller, but were still more accurate at 0.25° . This demonstrated an improvement in representativeness between satellite data and monitoring wells post-downscaling. Most of the Δ GWS coefficients were moderately lower than the Δ TWS coefficients, unlike in the southern zone, where the median was slightly higher. This difference was generated by the variance and noise acquired during the decomposition, which can be caused by the lack of continuity in some time series, the effect of local depletion and a lower number of measurements. A comparison of the correlation coefficients between the components of the RIMAS Δ WL time series with Δ TWS and Δ GWS average at 0.25° for the different monitoring wells and zones of the ARB is presented in the radial graph of Fig. 11.

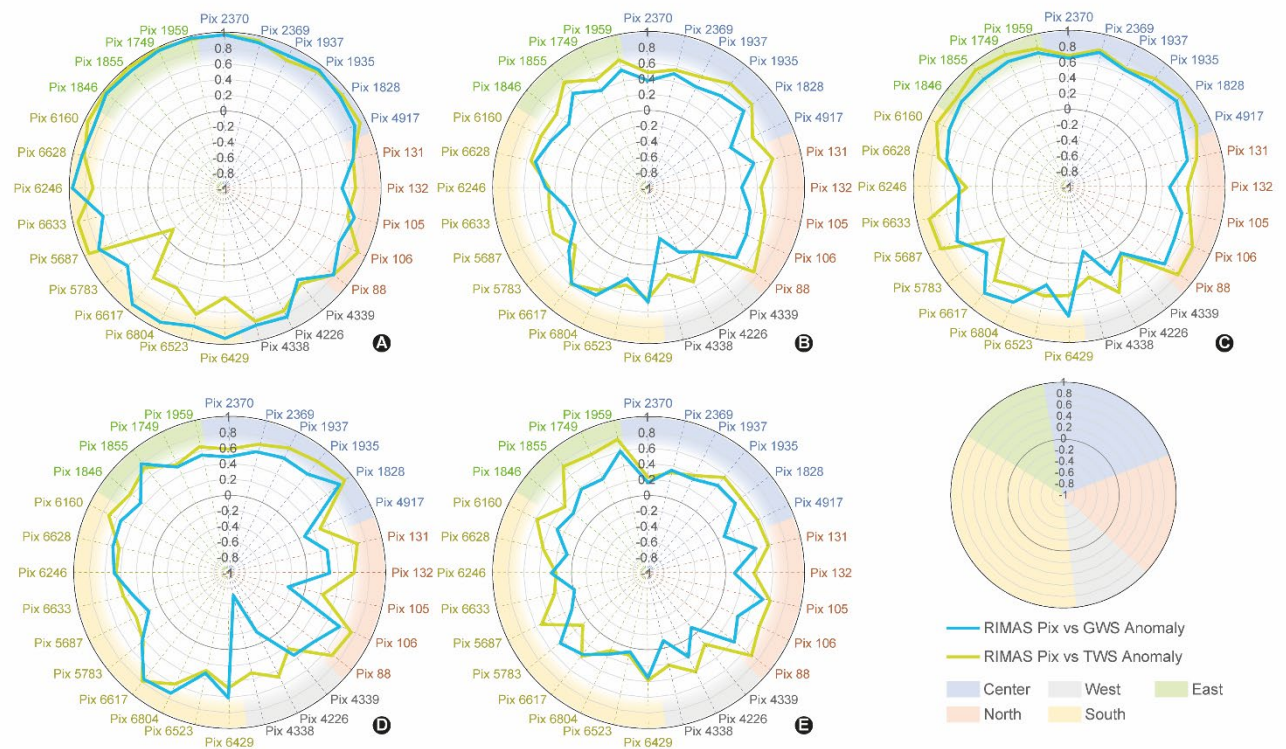


Fig. 11. Radial graph with the correlation coefficients between the RIMAS Δ WL with the Δ GWS and Δ TWS AVERAGE anomalies for the components of a) Seasonality, b) Seasonal Adjustment, c) Total, d) Trend, and e) Residual.

An interesting aspect of this analysis was the increase in the correlation coefficients in the seasonality of the Δ GWS anomalies scaled with the RIMAS monitoring data. In the East, North, and Central zones, there was not much difference to Δ TWS; however, an increase was evident in the West and South zones. This increase in the Seasonality correlation between Δ GWS and RIMAS Δ WL pix indicated an improvement in the representativeness of the post-decomposition data. As for the long-term behavior through the trend, there were significant differences to the North (average difference of 0.32) and West (average difference of 0.58), where the monitoring wells were more correlated with TWS. The original time series and the Seasonal Adjustment coefficients were similar, although a slight superiority of TWS over GWS prevailed.

When interpreting the variations in linear trends over time, there were significant differences between Δ TWS and Δ GWS in response to climatic phenomena such as El Niño and La Niña. Notably, during 2016, there was a substantial reduction in the rates, being much more pronounced in Δ TWS, indicating a greater sensitivity to climate actors in surface waters to groundwater (Fig. 12). From 2017 to 2021, the linear trends stabilized with GWS magnitudes on the order of 23 km³/year compared to the 12 km³/year of Δ TWS. Although at a much lower rate, albeit relatively constant, Δ GWS continued to decrease from magnitudes of 28.8 km³/year in 2016 to 22.2 km³/year in 2021, which, if it continues, could reach a neutral or negative trend between 2038 and 2043.

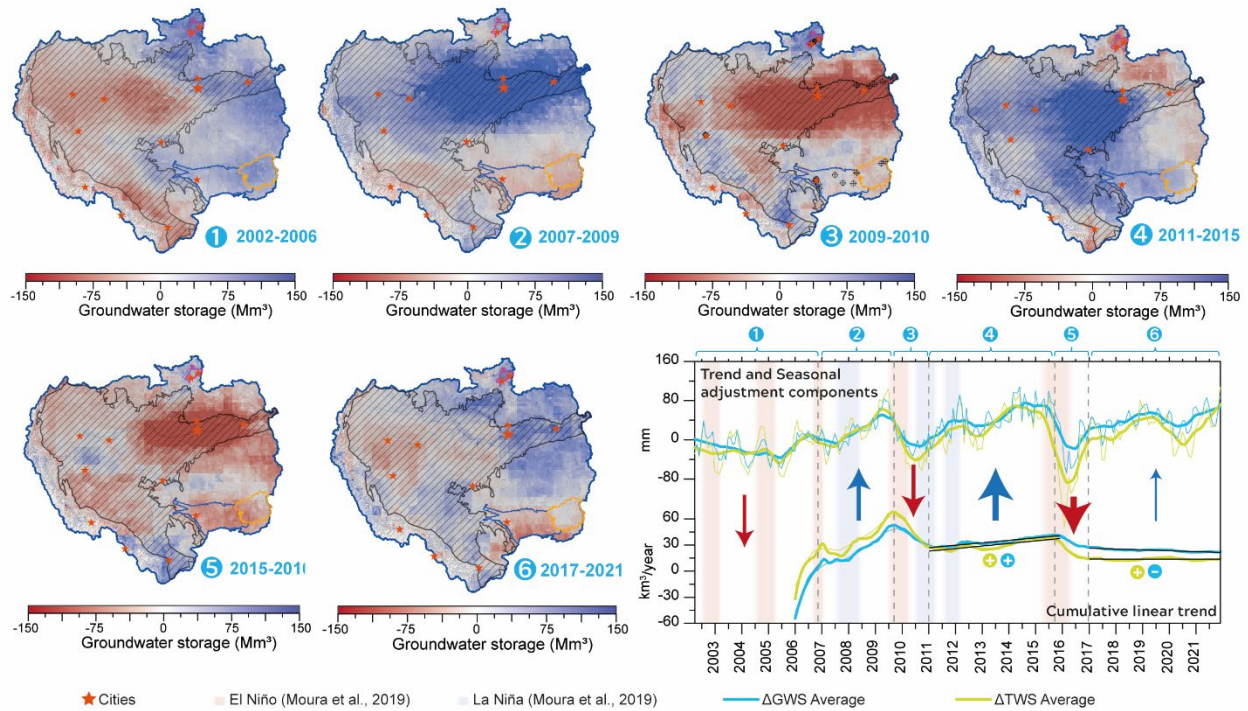


Fig. 12. Long-term behavior of the ΔTWS and ΔGWS time series of the ARB, where a) Trend and Seasonal adjustment, and b) variations in linear trends over time for the Trend and Seasonal adjustments.

Since the El Niño phenomenon of 2015-2016, changes were generated in the rates of growth and decrease in the ARB water storage changes, going from positive to negative, especially those of ΔGWS . Fig. S4 shows the strong influence of the phenomenon within the Alter do Chão Aquifer; however, as shown in Fig. 12, the behavior from 2017 to 2021 is positive, showing a prompt recovery. The areas with the greatest losses are within the ARB's middle basin. Some RIMAS wells, such as those in São Jose do Xingu and Presidente Figueiredo, corroborated this negative behavior in long-term trends. Fig. 13 compares the linear trends in the ΔGWS anomalies scaled by pixel for the ARB.

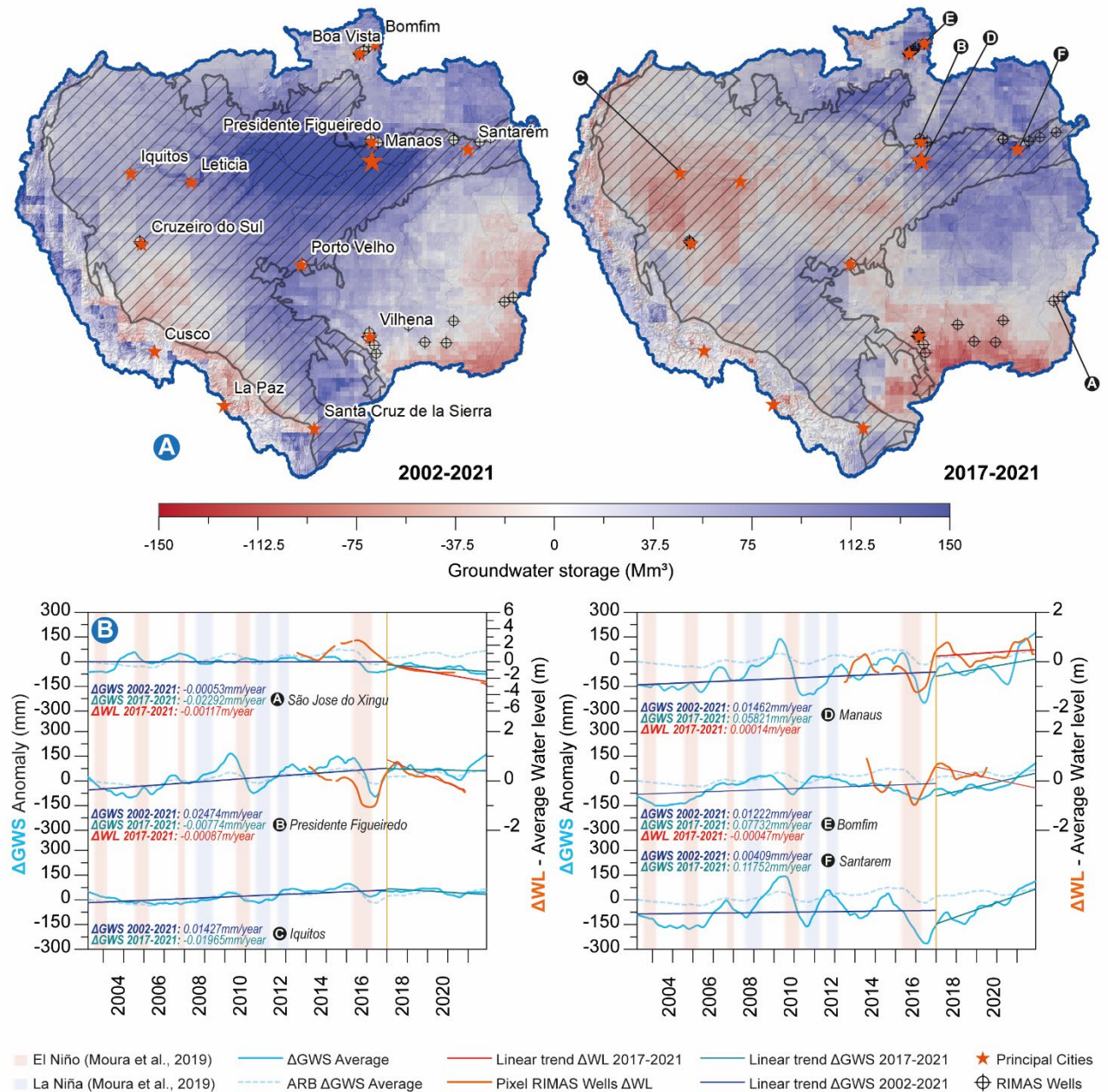


Fig. 13. Long-term ΔGWS behavior of the ARB, where a) Spatialization of GWS derived from the linear trend between 2002-2021 and 2017-2021, and b) Comparison of the time series of ΔGWS and ΔWL pixels along with their linear trends after 2017 and between 2002 and 2021.

5. Discussion

5.1. Downscaling with machine learning algorithms

This research generated a TWS dataset for various GRACE/GRACE-FO solutions, along with an average of GWS, at a spatial resolution of 0.25° and a monthly temporal resolution without temporal gaps in the ARB. It applied machine learning models that adapted to the region's specific conditions and characteristics, considering hydrometeorological and morphological variables. This allowed for obtaining more accurate predictions for the ARB, reflecting local patterns, effects, and relationships. It's

worth noting that in this process, the inherent uncertainties of the models and the remote sensors used as sources of information were also assimilated. Consistent with the findings of Fassoni-Andrade et al. (2021), variables such as precipitation and evaporation in the ARB show high spatial and temporal variability. This is compounded by the region's complex topography and varied vegetation cover, making it challenging to obtain precise and reliable measurements through remote sensing methods. It evaluated several precipitation datasets used in the ARB, such as TRMM (Tropical Rainfall Measuring Mission) and GPM (Global Precipitation Measurement Mission), and found that CHIRPS exhibited a strong correlation with in-situ measurement stations in the ARB, as reported by Haghtalab et al. (2020).

Additionally, GLEAM data demonstrated reasonable and consistent accuracy (Gomis-Cebolla et al., 2019; Martens et al., 2017; and Zhang et al., 2019), though some limitations were identified in areas with high soil evaporation, as noted by Fassoni-Andrade et al. (2021). Regarding GLDAS, it utilized 16 of its 36 variables. This dataset combines the assimilation of in-situ observation data and remote sensors with hydrological models and is widely used in decomposing the GRACE/GRACE-FO signal. The diversity of information sources granted exploring and identifying the most relevant variables and their relative contribution to the models' generalization capacity. This approach enabled the models to be trained with variables that contributed diverse characteristics to the predictions.

In terms of accuracy, the evaluation metrics, including R, MAE, and RMSE, yielded results similar to those obtained by Chen et al. (2019) and Ali et al. (2021) when employing similar methodologies and algorithms, despite differences in the extent and complexity of the study locations. However, a slight advantage in prediction accuracy was observed when using the AB algorithm compared to RF, possibly due to its higher sensitivity to extreme values. This study analyzed 39 monitoring wells, and strong correlations (exceeding 0.6) were found in 4 out of the five examined zones, validating the downscaled results. The wells' location, depth, and lithological characteristics provided key information for characterizing local behaviors in the Alter do Chão Aquifer System (AAS) and the Parecis, Ronuro, and Boa Vista aquifers. It's important to note that correlations decreased at greater well depths, suggesting they may be related to confined or semi-confined portions of the aquifers.

Furthermore, the downscaled process improved data accuracy, as evidenced by the increased well correlations compared to 1° data. When comparing well correlations with GWS and TWS at a 0.25° temporal resolution across different temporal components (Fig. 11), it was observed that some regions exhibited higher correlations with TWS than with GWS. However, the seasonal component of GWS displayed the highest correlations with all the wells.

In summary, the results of this research solution the utility of *downscaling* using machine learning algorithms to enhance data accuracy in the ARB, particularly in the context of aquifer monitoring and the interactions of hydrometeorological variables in the region. These findings have significant

implications for understanding and managing water resources in the ARB and can contribute to future research in hydrology and groundwater monitoring.

5.2. GWS Spatial-Temporal Characteristics

During the estimation of GWS, the changes in surface water storage obtained from GLDAS data underestimate the true magnitude. Nevertheless, this underestimation can be addressed by applying methodologies such as those proposed by Frappart et al. (2008, 2012) and Ndehedehe & Ferreira (2020). These methodologies combine the extent of surface waters with altimetric time series of water levels over rivers and floodplains. Furthermore, when examining the seasonality maps (Fig. 7), annual averages (Fig. 9), and standard deviation (Fig. 6) of GWS, spatial patterns can be identified in which extreme magnitudes are associated with geological units, particularly the major aquifers in the ARB, as well as floodplains, surface water bodies, and some topographic features. The high seasonality in these regions is linked to recharge patterns due to precipitation and the unconfined nature of aquifers and some residual soils in the ARB, which aligns with the findings of Hu et al. (2017).

The seasonal component accounts for the major proportion of both TWS and GWS. Monthly values obtained using the method of Cleveland et al. (1990) repeat periodically. However, the original signal exhibits variations in amplitude and frequency over time. An analysis of the power spectrum of this original seasonality may provide additional interpretative criteria to discern changes at different time intervals. Within these methodologies, the application of Least Squares Cross-Wavelet Analysis (LSCWA) or Cross-Wavelet Transform (XWT) can be employed, enabling the identification of periods during which the time series exhibit significant seasonality and their temporal lags (Ghaderpour et al., 2023; Li et al., 2020).

As shown in Fig. 12, the linear trends over time indicate that starting from the El Niño event in 2009-2010 positive GWS rates exceed those of TWS. However, from 2017 onward, GWS rates consistently decreased, in contrast to TWS, which remains relatively constant. These magnitudes and directions of the linear trends align with the research of Ndehedehe et al. (2020) and Scanlon et al. (2018). In the maps of linear trends for the period 2002-2021, it can be observed that regions with negative values are located over the Ronuro and Parecis Aquifers, areas heavily impacted by human activities such as burning, deforestation, and extensive agriculture. In the Xingu River sub-basin, negative trends are related to disruptions in the aquifer and residual soil recharge, where deforestation leads to increased surface runoff, reduced evapotranspiration, and consequently, decreased infiltration, in line with the findings of Panday et al. (2015). Spatial analysis of the linear trends for different periods indicates a high susceptibility of the Içã and Alter do Chão aquifers to El Niño and La Niña events. The continual population growth, its demand for water (primarily from underground sources), and the development of various activities like burning, deforestation, the construction of hydroelectric dams, and mining, among

others, disrupt the hydrological cycle and the ecosystem services of the ARB (Latrubesse et al., 2017; Neill et al., 2011; Nobre et al., 2009).

A pixel located over the city of Manaus, which exhibits statistically significant differences compared to the original GRACE data, shows a negative average of GWS between 2002 and 2021 (Fig. S8). Furthermore, this pixel has a higher density of extraction wells recorded by SIAGAS and significant correlations (>0.6) with the RIMAS monitoring wells. This may indicate an improvement in the representation of hydrological phenomena at local scales, achieved through the *downscaling* process, thereby enhancing the possibilities for monitoring unconfined aquifers using remote sensing data.

5.3. Future Research

Over time, the incorporation of more extensive and higher-quality time series will lead to a substantial improvement in model accuracy. The division of the ARB into subdomains, such as the upper, middle, and lower basins, can facilitate the separation and better assimilation of local climatic effects. Including categorical variables, such as geology and soil type (which remain constant over time), or indicators of human activity, like vegetation cover, well density, fires, deforestation, and population density (which vary over time), is essential for future research. It is imperative to consider the implementation of new *downscaling* methodologies with an even more precise target resolution, reaching levels of 0.05° or even 0.01° (equivalent to 5.5 km or 1.1 km, respectively), or even 0.004° (corresponding to 500 meters), as achieved with approaches like those of Zhang et al. (2021) and Zhong et al. (2021). These methodologies should be based on morphometric and multispectral information, utilizing indices such as NDVI, NDWI, TWI, or TPI, among others.

Furthermore, efforts should be made to integrate tools that allow the joint modeling of groundwater and surface water. Additionally, new monitoring well data provided by the governments of Colombia, Ecuador, Peru, and Bolivia should be incorporated. To further enrich the research, it is necessary to complement and compare the results obtained with additional datasets, such as GIEMS, GHWS, and other GLDAS models like V2.2, which includes Δ TWS anomalies at a daily resolution of 0.5° . Furthermore, ML or Deep Learning models can be implemented in restoring TWS/GWS signals for periods without GRACE/GRACE-FO data, such as before 2002, or to enhance their forecasting capabilities, as suggested in (Dannouf et al., 2022; Kalu et al., 2022, 2023). These actions will significantly contribute to advancing our understanding of hydrology in the ARB and strengthen the foundations for the sustainable management of water resources in the region.

6. Conclusions

GRACE/GRACE-FO records allowed the exploration of new research on large-scale GWS variations. During this study, two ML models were compared in Δ TWS *downscaling*, evaluating multiple independent hydrometeorological variables. The results showed a slight superiority of the models based

on AB algorithms over those based on RF algorithms. *Downscaling* from 1° to 0.25° was successfully achieved, decomposing the GWS anomalies between 2002 and 2021 for the ARB.

The error of the predictions was very low, showing good consistency between the data before and after *downscaling*. It determined the short and long-term behavior of the time series, obtaining linear trends for ΔTWS of $+14.26 \pm 1.18 \text{ km}^3/\text{year}$ and for ΔGWS of $+22.24 \pm 1.18 \text{ km}^3/\text{year}$ for all ARB. With an increase or gain in groundwater storage between 2002 and 2021 of $432.94 \pm 26.7 \text{ km}^3$ for the ARB. The validation with the RIMAS monitoring wells showed promising results with maximum correlation coefficients between the ΔWL with ΔGWS of 0.85 and ΔTWS of 0.90. The maximum correlations between ΔWL with ΔGWS and ΔTWS were 0.79/0.92 in the East, 0.77/0.84 in the Center, 0.73/0.88 in the South, 0.57/0.79 in the North and 0.23/0.49 in the West of the ARB. Spatial patterns were also identified in the anomalies related to the strong influence of some water bodies and topographic features, such as snow-capped mountains in southern Peru, which would not have been possible to identify at a lower spatial resolution. Differences in the SD and linear trend of the predicted or scaled ΔTWS and ΔGWS show the spatial and temporal variance in behavior between the upper (Andes) and lower (plains) basins, which can be taken into consideration to define domains before training future *downscaling* models.

Downscaling using AB effectively broadened the spatial resolution of the GRACE and GRACE-FO-derived data and improved their accuracy. Hypothesis testing to the mean between the scaled pixels (0.25°) and the original resolution (1°) in the three solutions (CSR, GFZ, and JPL) evidenced significant changes derived from *downscaling*, mainly to the North of the ARB, in the Andes region and isolated pixels such as one located over the city of Manaus. These significant changes resulted from gains in spatial and temporal variance resulting from *downscaling*, in which features of the training variables were assimilated into the ML models. Extreme periods caused by El Niño and La Niña phenomena were identified in years such as 2002, 2009, 2012, and 2016, highlighting the high vulnerability of some aquifers of the AAS to climatic effects. Future research should consider the impacts of anthropogenic activity at regional and local scales. The high correlation in the analyzed areas between the ΔWL and ΔGWS promised results regarding using GRACE/GRAC-FO data in the regional monitoring of unconfined aquifers, becoming a valuable tool for water management in large areas with difficult access.

Funding: This research was supported by São Paulo Research Foundation (FAPESP) [grant 2014/00844-7] and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) [Finance Code 001].

Data Availability

The TWS datasets for each solution (CSR, GFZ, and JPL) obtained using the AB algorithm, as well as the average TWS and GWS for the ARB, are provided in the repository. The files are stored in geodatabases that contain raster data for the 237 analyzed months. Data archived in the Mendely.data

repository which can be accessed using the following link (<https://data.mendeley.com/preview/2jp6c3gdh8?a=8ab99768-5bd7-41b1-be02-0f4a6796ee60>) (Satizábal-Alarcón, et al., 2023).

Appendix A

The Supplementary Material contains three tables (Table S1 to Table S3) and eight additional figures (Fig. S1 to Fig. S8), mentioned in the article.

7. References

- Agarwal, V., Akyilmaz, O., Shum, C. K., Feng, W., Yang, T. Y., Forootan, E., Syed, T. H., Haritashya, U. K., & Uz, M. (2023). Machine learning-based downscaling of GRACE-estimated groundwater in Central Valley, California. *Science of the Total Environment*, 865. <https://doi.org/10.1016/j.scitotenv.2022.161138>
- Agência Nacional de Águas - ANA. (2015). Avaliação dos Aquíferos das Bacias Sedimentares da Província Hidrogeológica Amazonas no Brasil (escala 1:1.000.000) e Cidades Pilotos (escala 1:50.000) - Volume III – Hidrogeologia e Modelo Numérico de Fluxo da PHA no Brasil.
- Ali, S., Liu, D., Fu, Q., Cheema, M. J. M., Pham, Q. B., Rahaman, M. M., Dang, T. D., & Anh, D. T. (2021). Improving the resolution of grace data for spatiotemporal groundwater storage assessment. *Remote Sensing*, 13(17). <https://doi.org/10.3390/rs13173513>
- Alves Saraiva, M., & Daniel Wahnfried, I. (2017). Avaliação da sustentabilidade do uso da água do Aquífero Alter do Chão na zona urbana de Manaus.
- Anjos, Nelson. D. F. R. D., Miletto, Michela., Donoso, M. Concepción., Aureli, Alice., van de Gun, Jac., Tujchneider, Ofelia., Rivera, Alfonso., & UNESCO. (2007). Sistemas acuíferos transfronterizos en las Américas, evaluación preliminar. UNESCO.
- Braga, B., Varella, P., & Gonçalves, H. (2011). Transboundary water management of the Amazon Basin. *International Journal of Water Resources Development*, 27(3), 477–496. <https://doi.org/10.1080/07900627.2011.595382>
- Breiman, L. (2001). Random Forests. 45, 5–32.
- Chen, L., He, Q., Liu, K., Li, J., & Jing, C. (2019). Downscaling of GRACE-derived groundwater storage based on the random forest model. *Remote Sensing*, 11(24). <https://doi.org/10.3390/rs11242979>
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition. *J. Off. Stat*, 6(1), 3-73. Cooley, S. S., & Landerer, F. W. (2019). GRACE L-3 Product user handbook. Jet Propulsion Laboratory California Institute of Technology.

- Cui, L., Yin, M., Huang, Z., Yao, C., Wang, X., & Lin, X. (2022). The Drought Events over the Amazon River Basin from 2003 to 2020 Detected by GRACE/GRACE-FO and Swarm Satellites. *Remote Sensing*, 14(12). <https://doi.org/10.3390/rs14122887>
- do Nascimento, N. R., Fritsch, E., Bueno, G. T., Bardy, M., Grimaldi, C., & Melfi, A. J. (2008). Podzolization as a deferralitization process: Dynamics and chemistry of ground and surface waters in an Acrisol - Podzol sequence of the upper Amazon Basin. *European Journal of Soil Science*, 59(5), 911–924. <https://doi.org/10.1111/j.1365-2389.2008.01049.x>
- Dannouf, R., Yong, B., Ndehedehe, C. E., Correa, F. M., & Ferreira, V. (2022). Boosted Regression Tree Algorithm for the Reconstruction of GRACE-Based Terrestrial Water Storage Anomalies in the Yangtze River Basin. *Frontiers in Environmental Science*, 10. <https://doi.org/10.3389/fenvs.2022.917545>
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. In *Journal of Animal Ecology* (Vol. 77, Issue 4, pp. 802–813). <https://doi.org/10.1111/j.1365-2656.2008.01390.x>
- Fassoni-Andrade, A. C., Fleischmann, A. S., Papa, F., Paiva, R. C. D. de, Wongchuig, S., Melack, J. M., Moreira, A. A., Paris, A., Ruhoff, A., Barbosa, C., Maciel, D. A., Novo, E., Durand, F., Frappart, F., Aires, F., Abrahão, G. M., Ferreira-Ferreira, J., Espinoza, J. C., Laipelt, L., ... Pellet, V. (2021). Amazon Hydrology From Space: Scientific Advances and Future Challenges. In *Reviews of Geophysics* (Vol. 59, Issue 4). John Wiley and Sons Inc. <https://doi.org/10.1029/2020RG000728>
- Frappart, F., Papa, F., Famiglietti, J. S., Prigent, C., Rossow, W. B., & Seyler, F. (2008). Interannual variations of river water storage from a multiple satellite approach: A case study for the Rio Negro River basin. *Journal of Geophysical Research Atmospheres*, 113(21). <https://doi.org/10.1029/2007JD009438>
- Frappart, F., Papa, F., Güntner, A., Tomasella, J., Pfeffer, J., Ramillien, G., Emilio, T., Schiatti, J., Seoane, L., da Silva Carvalho, J., Medeiros Moreira, D., Bonnet, M. P., & Seyler, F. (2019). The spatio-temporal variability of groundwater storage in the Amazon River Basin. *Advances in Water Resources*, 124, 41–52. <https://doi.org/10.1016/j.advwatres.2018.12.005>
- Frappart, F., Papa, F., Santos Da Silva, J., Ramillien, G., Prigent, C., Seyler, F., & Calmant, S. (2012). Surface freshwater storage and dynamics in the Amazon basin during the 2005 exceptional drought. *Environmental Research Letters*, 7(4). <https://doi.org/10.1088/1748-9326/7/4/044010>

- Frappart, F., & Ramillien, G. (2018). Monitoring groundwater storage changes using the Gravity Recovery and Climate Experiment (GRACE) satellite mission: A review. In *Remote Sensing* (Vol. 10, Issue 6). MDPI AG. <https://doi.org/10.3390/rs10060829>
- Freund, Y., & Schapire, R. E. (1995, March). A decision-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory* (pp. 23-37). Berlin, Heidelberg: Springer Berlin Heidelberg. Funk, C. C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D., Romero, B. E., Husak, G. J., Michaelsen, J. C., & Verdin, A. P. (2014). A Quasi-Global Precipitation Time Series for Drought Monitoring. U.S. Department of the Interior. U.S. Geological Survey, Data Series 832. <https://doi.org/10.3133/ds832>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, 2. <https://doi.org/10.1038/sdata.2015.66>
- Gatti, L. v., Basso, L. S., Miller, J. B., Gloor, M., Gatti Domingues, L., Cassol, H. L. G., Tejada, G., Aragão, L. E. O. C., Nobre, C., Peters, W., Marani, L., Arai, E., Sanches, A. H., Corrêa, S. M., Anderson, L., von Randow, C., Correia, C. S. C., Crispim, S. P., & Neves, R. A. L. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature*, 595(7867), 388–393. <https://doi.org/10.1038/s41586-021-03629-6>
- Ghaderpour, E., Mazzanti, P., Mugnozza, G. S., & Bozzano, F. (2023). Coherency and phase delay analyses between land cover and climate across Italy via the least-squares wavelet software. *International Journal of Applied Earth Observation and Geoinformation*, 118. <https://doi.org/10.1016/j.jag.2023.103241>
- Haghtalab, N., Moore, N., Heerspink, B. P., & Hyndman, D. W. (2020). Evaluating spatial patterns in precipitation trends across the Amazon basin driven by land cover and global scale forcings. *Theoretical and Applied Climatology*, 140(1–2), 411–427. <https://doi.org/10.1007/s00704-019-03085-3>
- Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., & Wolff, D. B. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1), 38–55. <https://doi.org/10.1175/JHM560.1>
- ISARM, U. I. (2009). Atlas of Transboundary Aquifers: Global maps, regional cooperation and local inventories. *World water week*, 326..

- Jimenez, J. C., Marengo, J. A., Alves, L. M., Sulca, J. C., Takahashi, K., Ferrett, S., & Collins, M. (2021). The role of ENSO flavours and TNA on recent droughts over Amazon forests and the Northeast Brazil region. *International Journal of Climatology*, 41(7), 3761–3780. <https://doi.org/10.1002/joc.6453>
- Kalu, I., Ndehedehe, C. E., Okwuashi, O., Eyoh, A. E., & Ferreira, V. G. (2022). A new modelling framework to assess changes in groundwater level. *Journal of Hydrology: Regional Studies*, 43. <https://doi.org/10.1016/j.ejrh.2022.101185>
- Kalu, I., Ndehedehe, C. E., Okwuashi, O., Eyoh, A. E., & Ferreira, V. G. (2023). Reconstructing terrestrial water storage anomalies using convolution-based support vector machine. *Journal of Hydrology: Regional Studies*, 46. <https://doi.org/10.1016/j.ejrh.2023.101326>
- Khorrami, B., Ali, S., & Gündüz, O. (2023). Investigating the Local-scale Fluctuations of Groundwater Storage by Using Downscaled GRACE/GRACE-FO JPL Mascon Product Based on Machine Learning (ML) Algorithm. *Water Resources Management*, 37(9), 3439–3456. <https://doi.org/10.1007/s11269-023-03509-w>
- Cooley, S. S., & Landerer, F. W. (2019). GRACE L-3 Product user handbook. Jet Propulsion Laboratory California Institute of Technology.
- Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, 48(4). <https://doi.org/10.1029/2011WR011453>
- Latrubesse, E. M., Arima, E. Y., Dunne, T., Park, E., Baker, V. R., D’Horta, F. M., Wight, C., Wittmann, F., Zuanon, J., Baker, P. A., Ribas, C. C., Norgaard, R. B., Filizola, N., Ansar, A., Flyvbjerg, B., & Stevaux, J. C. (2017). Damming the rivers of the Amazon basin. In *Nature* (Vol. 546, Issue 7658, pp. 363–369). Nature Publishing Group. <https://doi.org/10.1038/nature22333>
- Leite, C. C., Costa, M. H., Soares-Filho, B. S., & de Barros Viana Hissa, L. (2012). Historical land use change and associated carbon emissions in Brazil from 1940 to 1995. *Global Biogeochemical Cycles*, 26(2). <https://doi.org/10.1029/2011GB004133>
- Li, Y., Wen, Y., Lai, H., & Zhao, Q. (2020). Drought response analysis based on cross wavelet transform and mutual entropy. *Alexandria Engineering Journal*, 59(3), 1223–1231. <https://doi.org/10.1016/j.aej.2020.01.050>
- Maeda, E. E., Ma, X., Wagner, F. H., Kim, H., Oki, T., Eamus, D., & Huete, A. (2017). Evapotranspiration seasonality across the Amazon Basin. *Earth System Dynamics*, 8(2), 439–454. <https://doi.org/10.5194/esd-8-439-2017>
- Margat, J. (2007). Great aquifer systems of the world. In *Aquifer Systems Management: Darcy's Legacy in a World of Impending Water Shortage* (pp. 105-116). CRC Press. Martens, B., Miralles, D. G.,

- Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., & Verhoest, N. E. C. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5), 1903–1925. <https://doi.org/10.5194/gmd-10-1903-2017>
- Miguez-Macho, G., & Fan, Y. (2012). The role of groundwater in the Amazon water cycle: 1. Influence on seasonal streamflow, flooding and wetlands. *Journal of Geophysical Research Atmospheres*, 117(15). <https://doi.org/10.1029/2012JD017539>
- Miralles, D. G., Holmes, T. R. H., de Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, 15(2), 453–469. <https://doi.org/10.5194/hess-15-453-2011>
- Moura, M. M., dos Santos, A. R., Pezzopane, J. E. M., Alexandre, R. S., da Silva, S. F., Pimentel, S. M., de Andrade, M. S. S., Silva, F. G. R., Branco, E. R. F., Moreira, T. R., da Silva, R. G., & de Carvalho, J. R. (2019). Relation of El Niño and La Niña phenomena to precipitation, evapotranspiration and temperature in the Amazon basin. *Science of the Total Environment*, 651, 1639–1651. <https://doi.org/10.1016/j.scitotenv.2018.09.242>
- Ndehedehe, C. E., & Ferreira, V. G. (2020). Assessing land water storage dynamics over South America. *Journal of Hydrology*, 580. <https://doi.org/10.1016/j.jhydrol.2019.124339>
- Neill, C., Chaves, J. E., Biggs, T., Deegan, L. A., Elsenbeer, H., Figueiredo, R. O., Germer, S., Johnson, M. S., Lehmann, J., Markewitz, D., & Piccolo, M. C. (2011). Runoff sources and land cover change in the Amazon: An end-member mixing analysis from small watersheds. *Biogeochemistry*, 105(1), 7–18. <https://doi.org/10.1007/s10533-011-9597-8>
- Nobre, P., Malagutti, M., Urbano, D. F., De Almeida, R. A. F., & Giarolla, E. (2009). Amazon deforestation and climate change in a coupled model simulation. *Journal of Climate*, 22(21), 5686–5697. <https://doi.org/10.1175/2009JCLI2757.1>
- Pfeffer, J., Seyler, F., Bonnet, M. P., Calmant, S., Frappart, F., Papa, F., Paiva, R. C. D., Satgé, F., & Silva, J. S. da. (2014). Low-water maps of the groundwater table in the central Amazon by satellite altimetry. *Geophysical Research Letters*, 41(6), 1981–1987. <https://doi.org/10.1002/2013GL059134>
- Pimentel, E. T., & Hamza, V. M. (2012). Indications of regional scale groundwater flows in the Amazon Basins: Inferences from results of geothermal studies. *Journal of South American Earth Sciences*, 37, 214–227. <https://doi.org/10.1016/j.jsames.2012.03.007>
- Pokhrel, Y. N., Fan, Y., Miguez-Macho, G., Yeh, P. J. F., & Han, S. C. (2013). The role of groundwater in the Amazon water cycle: 3. Influence on terrestrial water storage computations and comparison

- with GRACE. *Journal of Geophysical Research Atmospheres*, 118(8), 3233–3244. <https://doi.org/10.1002/jgrd.50335>
- Pulla, S. T., Yasarer, H., & Yarbrough, L. D. (2023). GRACE Downscaler: A Framework to Develop and Evaluate Downscaling Models for GRACE. *Remote Sensing*, 15(9). <https://doi.org/10.3390/rs15092247>
- Reis, V., Hermoso, V., Hamilton, S. K., Bunn, S. E., Fluet-Chouinard, E., Venables, B., & Linke, S. (2019). Characterizing seasonal dynamics of Amazonian wetlands for conservation and decision making. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 29(7), 1073–1082. <https://doi.org/10.1002/aqc.3051>
- Rodell, B. M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., & Toll, D. (2004). The Global Land Data Assimilation System. *American Meteorological Society*, 381–394. <https://doi.org/10.1>
- Rosário, F. F. do, Custodio, E., & Silva, G. C. da. (2016). Hydrogeology of the Western Amazon Aquifer System (WAAS). *Journal of South American Earth Sciences*, 72, 375–386. <https://doi.org/10.1016/j.jsames.2016.10.004>
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y. T., Chuang, H. Y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., & Becker, E. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, 27(6), 2185–2208. <https://doi.org/10.1175/JCLI-D-12-00823.1>
- Satizabal, Diego; Suhogusoff, Alexandra; Ferrari, Luiz (2023), “Dataset: Characterization of groundwater storage changes in the Amazon River Basin based on downscaling of GRACE/GRACE-FO data with machine learning models”, Mendeley Data, V1, doi: 10.17632/2jp6c3gdh8.1
- Save, H., Bettadpur, S., & Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, 121(10), 7547–7569. <https://doi.org/10.1002/2016JB013007>
- Scanlon, B. R., Zhang, Z., Save, H., Wiese, D. N., Landerer, F. W., Long, D., Longuevergne, L., & Chen, J. (2016). Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resources Research*, 52(12), 9412–9429. <https://doi.org/10.1002/2016WR019494>
- Scheel, M. L. M., Rohrer, M., Huggel, Ch., Santos Villar, D., Silvestre, E., & Huffman, G. J. (2011). Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrology and Earth System Sciences*, 15(8), 2649–2663. <https://doi.org/10.5194/hess-15-2649-2011>

- Seyoum, W. M., Kwon, D., & Milewski, A. M. (2019). Downscaling GRACE TWSA data into high-resolution groundwater level anomaly using machine learning-based models in a glacial aquifer system. *Remote Sensing*, 11(7). <https://doi.org/10.3390/rs11070824>
- Swenson, S., Wahr, J., & Milly, P. C. D. (2003). Estimated accuracies of regional water storage variations inferred from the Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research*, 39(8). <https://doi.org/10.1029/2002WR001808>
- Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., & Watkins, M. M. (2004). GRACE Measurements of Mass Variability in the Earth System. *Science*, 305(5683), 503–505. <https://doi.org/10.1126/science.1099192>
- Tomasella, J., Hodnett, M. G., Cuartas, L. A., Nobre, A. D., Waterloo, M. J., & Oliveira, S. M. (2008). The water balance of an Amazonian micro-catchment: The effect of interannual variability of rainfall on hydrological behaviour. *Hydrological Processes*, 22(13), 2133–2147. <https://doi.org/10.1002/hyp.6813>
- Wahr, J., Swenson, S., Zlotnicki, V., & Velicogna, I. (2004). Time-variable gravity from GRACE: First results. *Geophysical Research Letters*, 31(11). <https://doi.org/10.1029/2004GL019779>
- Wang, J., Xu, D., & Li, H. (2023). Constructing GRACE-Based 1 km Resolution Groundwater Storage Anomalies in Arid Regions Using an Improved Machine Learning Downscaling Method: A Case Study in Alxa League, China. *Remote Sensing*, 15(11). <https://doi.org/10.3390/rs15112913>
- Watkins, M. M., Wiese, D. N., Yuan, D. N., Boening, C., & Landerer, F. W. (2015). Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth*, 120(4), 2648–2671. <https://doi.org/10.1002/2014JB011547>
- Wiese, D. N., Landerer, F. W., & Watkins, M. M. (2016). Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resources Research*, 52(9), 7490–7502. <https://doi.org/10.1002/2016WR019344>
- Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M. (2019). MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset. *Water Resources Research*, 55(6), 5053–5073. <https://doi.org/10.1029/2019WR024873>
- Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, 44(11), 5844–5853. <https://doi.org/10.1002/2017GL072874>
- Yin, W., Zhang, G., Liu, F., Zhang, D., Zhang, X., & Chen, S. (2022). Improving the spatial resolution of GRACE-based groundwater storage estimates using a machine learning algorithm and

hydrological model. *Hydrogeology Journal*, 30(3), 947–963. <https://doi.org/10.1007/s10040-021-02447-4>

Yuan, X., Wood, E. F., Luo, L., & Pan, M. (2011). A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophysical Research Letters*, 38(13). <https://doi.org/10.1029/2011GL047792>

Yuan, X., Wood, E. F., Luo, L., & Pan, M. (2011). A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophysical Research Letters*, 38(13). <https://doi.org/10.1029/2011GL047792>

Zhang, J., Liu, K., & Wang, M. (2021). Downscaling groundwater storage data in China to a 1-km resolution using machine learning methods. *Remote Sensing*, 13(3). <https://doi.org/10.3390/rs13030523>

Zhong, D., Wang, S., & Li, J. (2021). Spatiotemporal downscaling of grace total water storage using land surface model outputs. *Remote Sensing*, 13(5), 1–19. <https://doi.org/10.3390/rs13050900>

4.2. Efeitos do clima e das atividades antropogênicas nos Δ GWS da ARB

Após a implementação da metodologia de *downscaling* dos dados de Δ TWS derivados do GRACE/GRACE-FO com algoritmos de ML, estimando e validando o Δ GWS a 0.25° com os poços de monitoramento do RIMAS, descritos no item 4.1 (Satizábal-Alarcón et al., em submissão), deseja-se saber quais são os efeitos espaço-temporais das variáveis hidrometeorológicas, morfológicas e das atividades antropogênicas com o Δ GWS a curto e longo prazo para a ARB. Para isso, foi realizada uma caracterização do comportamento médio da ARB em geral, uma descrição dos padrões espaciais e temporais das principais variáveis hidrometeorológicas, a estimativa e a análise do índice topográfico de umidade (TWI) e uma análise de componentes principais (PCA) com base nos mapas de correlação entre as séries temporais de Δ GWS e cada uma das variáveis. Por fim, foram analisadas algumas das principais atividades antrópicas com potencial para afetar o ciclo hidrológico e as relacionamos com o Δ GWS a curto e longo prazo.

4.2.1. Decomposição das séries temporais

A decomposição temporal do sinal de GRACE/GRACE-FO em Sazonal (Seasonal), Tendência (Trend) e Residual facilita a interpretação das variações no armazenamento a curto e longo prazo. Na estimativa do Ajuste Sazonal (SA), ao subtrair apenas a componente sazonal do sinal original, mantêm-se todas as variações explicadas pelas demais variáveis. Ao comparar as componentes temporais das anomalias de Δ TWS e Δ GWS para a ARB (Figura 14), observa-se que as maiores variações são explicadas pelo comportamento a curto prazo (Sazonal), seguidas pela tendência (Tendência) e a residual, em menor medida. Esta componente apresenta um desvio padrão (SD) para a ARB de 140.3 mm para Δ TWS e de 78.4 mm para Δ GWS. A componente de Tendência (Trend) tem um SD de 34.4 mm para Δ TWS e de 29.6 mm para GWS. A componente Residual é a parte da série temporal que não pode ser explicada pelas outras componentes e representa padrões ou características não capturadas, que podem ser fenômenos não identificados. Neste caso, o SD é de 25.6 mm para Δ TWS e de 19.2 mm para GWS. Ao comparar as variações das componentes temporais em relação ao sinal original obtido a partir dos valores médios da ARB, observa-se que 95.5%/91.2% do sinal de Δ TWS/ Δ GWS é explicado pela componente Sazonal, 31.9%/4.3% pela componente de Tendência e 18.2%/22.4% pela componente Residual. Em relação ao sinal de Δ TWS em relação a Δ GWS, o segundo explica 58.4% do sinal total (original), 55.8% da componente Sazonal, 86.1% da componente de Tendência e 75.2% da componente Residual. Isso indica que as mudanças em Δ GWS têm a maior influência sobre o sinal de Δ TWS em comparação com Δ SMS, Δ CWS, Δ SWS e Δ Qs, tanto a curto como a longo prazo. No entanto, a influência do sinal de Δ GWS no comportamento a longo prazo é maior.

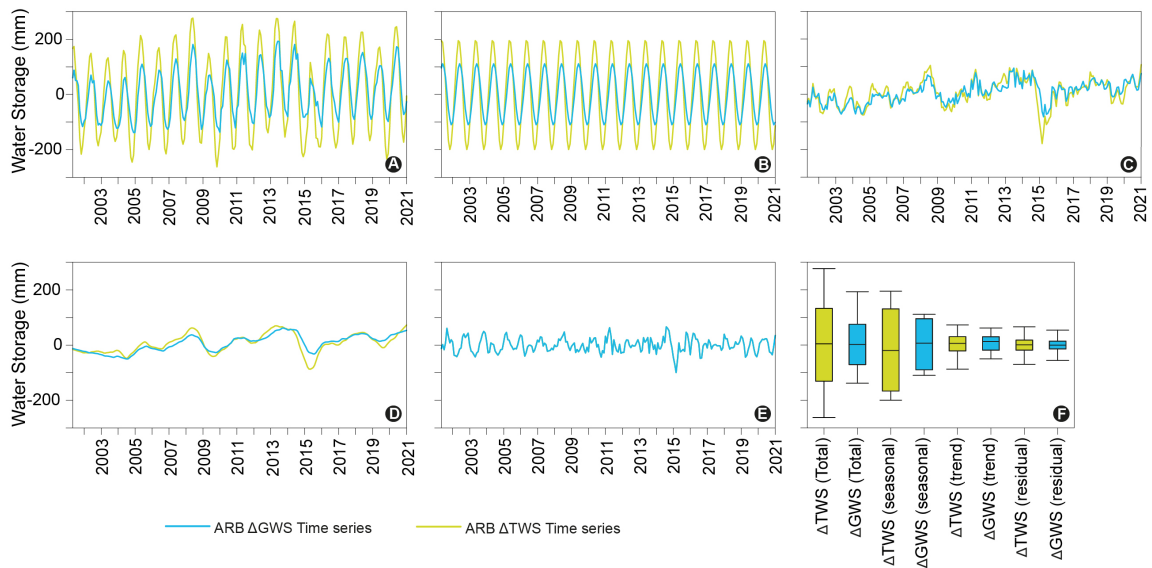


Figura 14. Componentes temporais do sinal Δ TWS e Δ GWS para o ARB. No eixo Y são apresentadas as mudanças no armazenamento de água em mm. A) série temporal original, total ou derivada do GRACE/GRACE-FO, B) componente sazonal, C) ajuste sazonal, D) componente de tendência, E) componente residual e F) valor médio do boxplot da série temporal para o ARB.

Comportamento médio da ARB

O comportamento da ARB descrito pelos valores médios das principais variáveis hidrológicas permite identificar a influência dos fenômenos climáticos, a relação entre as variáveis e quantificar as principais mudanças ao longo do tempo. A integração das variações espaciais e temporais das séries temporais proporciona critérios para identificar a influência de fatores como topografia, vegetação, geologia ou atividades antrópicas no ciclo hidrológico. A decomposição das séries temporais médias para a ARB (Figura 15-A) em sua componente de tendência mostra como os efeitos do El Niño e La Niña repercutem nas variações do armazenamento de água. As variáveis com maiores magnitudes são Δ TWS, Δ GWS e Δ SMS com variações de ± 80 mm, seguidas por Δ SWS de ± 1 mm, Δ CWS de ± 0.08 mm e, finalmente, Δ Qs de ± 0.04 mm. O fenômeno El Niño com maiores efeitos para a ARB foi o de 2015-2016, gerando os maiores déficits nas mudanças de armazenamento de água entre 2002 e 2021. O evento El Niño de 2009-2010 teve influência regional, embora com menor magnitude, no entanto, coincide com uma forte queda no Δ CWS, que se estende até 2021 sem mostrar uma recuperação significativa. A redução no Δ CWS ocorreu ao longo de 2 anos, iniciando em 2009 e se estabilizando em 2011, evento que coincide com uma queda no Δ SMS. A vegetação tem o papel de interceptar a água da chuva, facilitando sua evaporação ou transferência gradual para o solo. Uma possível causa de uma queda progressiva no Δ CWS é o avanço do desmatamento e queimadas na Amazônia, aspecto que pode ser acelerado com o aumento dos incêndios florestais agravados pelas secas de 2010. A vegetação desempenha uma função protetora do solo; quando é removida, o solo fica exposto diretamente à radiação solar, aumentando a evaporação da água nele contida e reduzindo consequentemente a umidade do solo. Essa redução na capacidade de retenção da água também é afetada pelo aumento do escoamento superficial e diminuição das taxas de recarga. De acordo com Coe et al., 2013, o desmatamento em

grande escala resulta em diminuição das taxas de precipitação em áreas próximas e até mesmo distantes. Ao detalhar o comportamento dos eventos extremos sofridos pela ARB, podem ser observados desfasamentos entre a precipitação, ΔTWS , ΔGWS , ΔSMS e ΔCWS . Em relação à precipitação, a resposta do ΔCWS é a mais próxima, seguida por ΔQs , ΔSMS , ΔTWS e, finalmente, ΔGWS . Isso é explicado pelo avanço da água ao longo do ciclo hidrológico a partir da precipitação, sendo inicialmente interceptada pela vegetação, fluindo para o solo, para os rios pelo escoamento e percolando lentamente para os aquíferos. Um dos efeitos de longo prazo mais representativos é a constante redução no ΔSWS a partir de 2005 nas montanhas localizadas nos Andes peruanos e bolivianos. Entre 2002 e 2021, ΔTWS e ΔGWS têm uma taxa de aumento linear de 2.341mm/ano e 3.653mm/ano, enquanto ΔSMS tem uma taxa de queda de -1.200mm/ano, ΔCWS de -0.003mm/ano e ΔSWS de -0.108mm/ano.

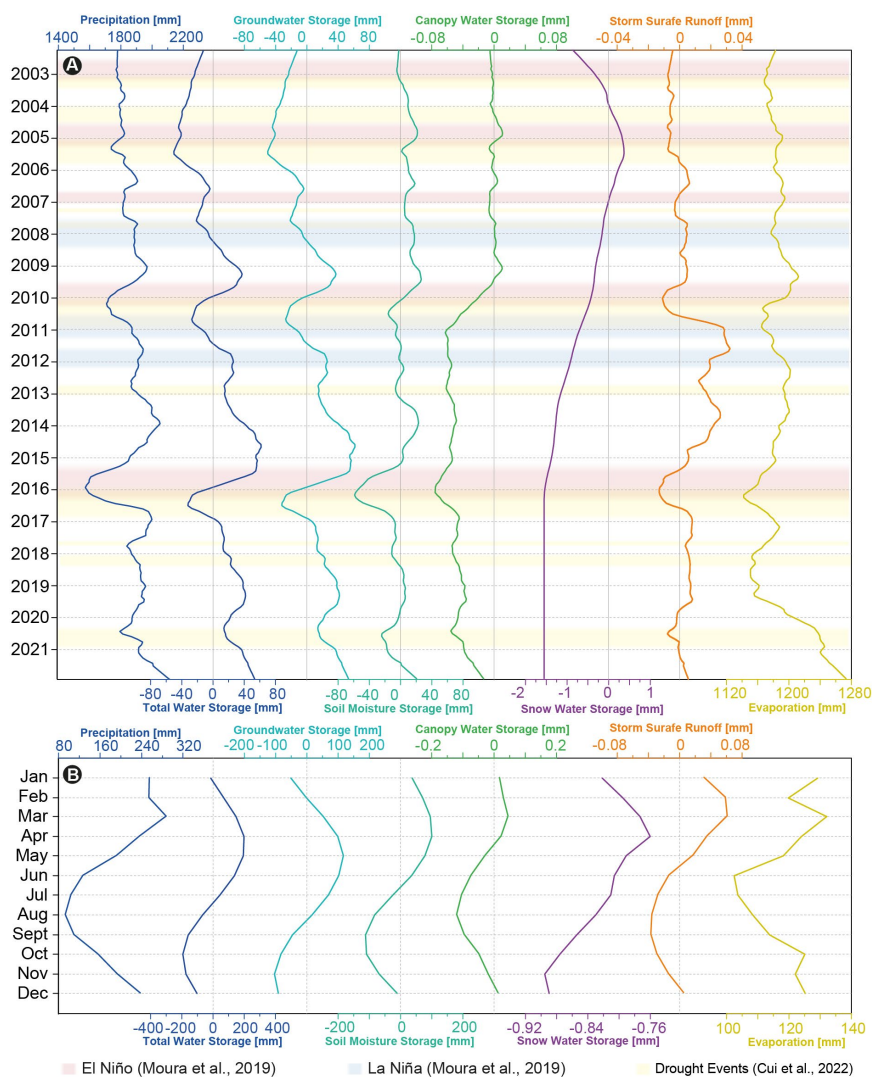


Figura 15. Série temporal média das principais variáveis hidrológicas da ARB. Onde A) Componente de tendência e B) Componente sazonal. Em vermelho, eventos de El Niño, em azul, eventos de La Niña de acordo com Moura et al., 2019 e em amarelo

A sazonalidade das variáveis (Figura 15-B) contém as mudanças de armazenamento com as maiores magnitudes das ST. Os máximos positivos ocorrem durante o primeiro semestre para todas as

variáveis e os máximos negativos durante o segundo semestre. Cronologicamente, os máximos positivos foram alcançados em março para Prec com 286mm, Evapo com 132mm, Δ Qs com 0.06mm e Δ CWS com 0.04mm, em abril para Δ TWS com 198mm e Δ SMS com 100mm, e finalmente em maio para Δ GWS com 117mm. Os máximos negativos são atingidos em junho para Evap com 102mm, em agosto para Prec com 92mm e Δ CWS com -0.12mm, em setembro para Δ Qs com -0.03mm e Δ SMS -112mm, em outubro para Δ TWS com -193mm e, finalmente, em novembro para Δ SWS com -0.89 e Δ GWS com -103mm.

A representação espacial dos valores médios multianuais para a ARB captura padrões de mudança a longo prazo das variáveis. A Figura 16 mostra que Δ TWS e Δ GWS na região central da AAS sobre o aquífero Içá, no meio das planícies de inundação, apresentam fortes anomalias positivas. A influência dos rios na região central é mais clara e intensa em Δ TWS. O centro e o leste do aquífero Alter do Chão apresentam valores negativos em ambas as variáveis, sendo moderadamente mais fortes em Δ TWS. As imediações da fronteira oeste da AAS apresentam anomalias moderadamente positivas de Δ TWS, embora baixas ou nulas para Δ GWS. A fronteira sudeste da ARB, sobre a sub-bacia do rio Xingu, o aquífero Ronuro e parte do aquífero Parecis, Δ TWS e Δ GWS são moderadamente negativos, comportamento que é representativo durante o período de 2002-2021. Dentro do domínio do aquífero Boa Vista, ao norte da ARB, as anomalias de Δ TWS apresentam valores fortemente negativos, em contraste com anomalias moderadamente negativas de Δ GWS. Nas áreas circundantes a este aquífero, prevalecem as magnitudes negativas de Δ TWS, ao contrário de Δ GWS que apresenta magnitudes nulas. Na região dos Andes, predominam as anomalias positivas de Δ TWS, com magnitudes moderadas, exceto por alguns pixels moderadamente ou fortemente negativos, localizados em áreas florestais que não parecem guardar relação com aspectos geográficos, geológicos ou climáticos específicos. As anomalias de Δ GWS nos Andes são moderadamente positivas na porção norte, sobre o Equador e a Colômbia. Em direção à porção central e sul, as médias são nulas e algumas regiões negativas são localizadas sobre afluentes do rio Ucayali. Anomalias fortemente positivas são encontradas em algumas montanhas no sul do Peru. A faixa da Bolívia é caracterizada por valores nulos e algumas anomalias negativas.

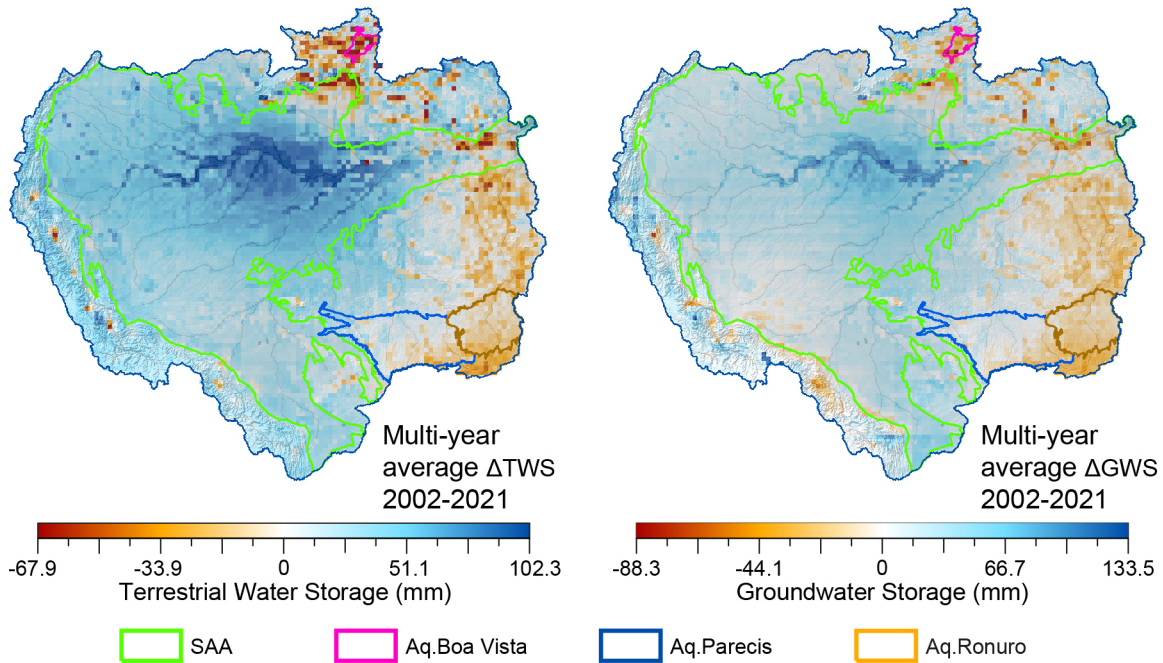


Figura 16. Média multianual das mudanças no armazenamento total (ΔTWS) e de água subterrânea (ΔGWS) do ARB.

As variáveis com as maiores magnitudes no comportamento médio da ARB, e sua influência em ΔTWS e ΔGWS , são a umidade do solo (SM), a precipitação (Prec) e a evaporação (Evap). As médias multianuais de SM (Figura 17-A) revelam padrões espaciais causados pela assimilação do mapa de classe de textura do solo USDA de Reynolds et al.,2000, utilizado na estimativa do GLDAS. Apesar da alta heterogeneidade nos dados, a região centro-oeste da AAS apresenta as maiores magnitudes (800 a 1000mm anuais). O aquífero Ronuro caracteriza-se por apresentar valores baixos (<500mm anuais). As magnitudes da bacia alta (Andes) são contrastantemente mais baixas do que as da bacia média e baixa (400 a 600mm anuais). Em relação às precipitações (Figura 17-B), os valores mais altos (>4000mm anuais) ocorrem ao NO da ARB, sobre território da Colômbia e do Brasil. As zonas de transição entre a bacia alta e média apresentam alguns "hot spots" ao NO e CO da ARB. As maiores altitudes nos Andes centrais e do sul têm as menores precipitações. O comportamento espacial da evaporação (Figura 17-C) assemelha-se à precipitação. O leito central do rio Amazonas ao C e E da ARB atinge valores médios de cerca de 2200mm anuais. As zonas C e N apresentam valores altos (>1400mm anuais), que diminuem progressivamente em direção ao sul, chegando a valores baixos (<800mm anuais). As maiores altitudes nos Andes são caracterizadas por apresentar valores baixos (<600mm anuais). O regime climático sobre o aquífero Boa Vista apresenta precipitações e evaporações moderadamente baixas (1300mm e 800mm anuais, respectivamente).

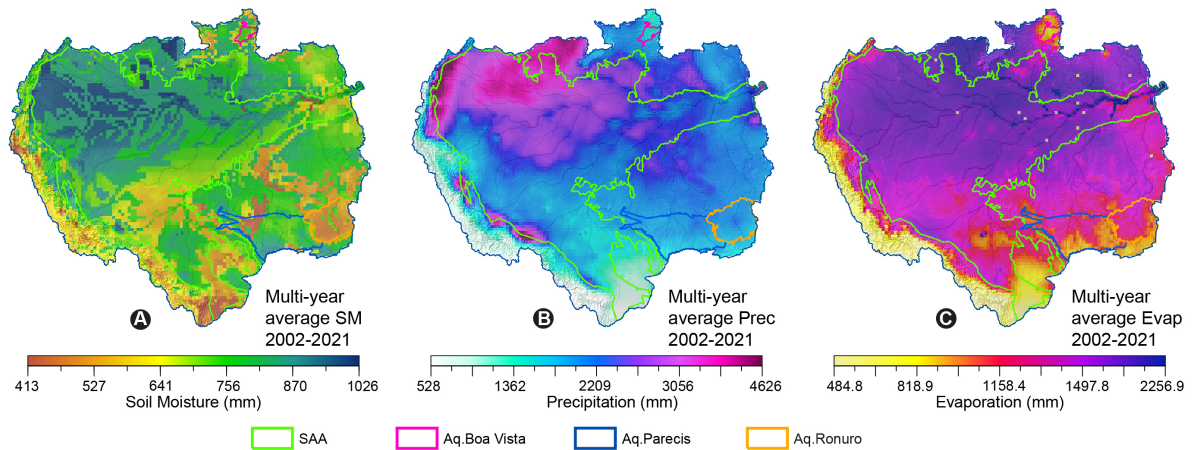


Figura 17. Média multianual de: A) umidade do solo (SM), B) precipitação (Prec) e C) evaporação (Evap) da ARB.

Resumidamente, ao NO da AAS, apresentam-se as maiores anomalias de SM (1000mm), Prec (4200mm) e valores altos de Evap (1650mm). Nas zonas de transição entre a bacia alta e média, localizam-se alguns "hot spots" de precipitação denominados chuvas orográficas, causados por mudanças na altitude do terreno e pela contribuição substancial da transpiração da floresta. Esses abundantes fluxos de água para a atmosfera gerados pelas florestas eventualmente retornam à terra como precipitação reciclada e contribuem com cerca de 30% da chuva da bacia (Fassoni-Andrade et al., 2021).

Os menores valores da bacia média e baixa ocorrem ao sur e sudeste da ARB, sobre o aquífero Ronuro e no território da Bolívia, ao sur dos Llanos de Moxos. O aquífero Boa Vista tem um regime climático que contrasta negativamente com relação às áreas adjacentes, com valores anuais de SM de 725mm, Prec de 1500mm e Evap de 800mm. O aquífero Alter do Chão, a oeste, apresenta maiores valores de SM entre 900-750mm, Prec entre 2600-2000mm e Evap entre 2100-1400mm, que diminuem em direção ao E, com SM entre 750-550, Prec entre 2300-1700mm e Evap entre 2200-1100mm.

A região da bacia alta ou dos Andes tende a diminuir os valores de Prec e Evap com a altitude, atingindo os valores mínimos de 528mm e 484mm, respectivamente. Em relação à SM, seu comportamento é mais heterogêneo, associado a mudanças na inclinação e geologia que facilitam a formação ou perda de solo, bem como variações locais de menor magnitude na Prec e Evap. Os mesmos padrões de distribuição entre os valores máximos e mínimos são observados nas 3 variáveis. Apesar das heterogeneidades entre os solos, os comportamentos médios obedecem à variabilidade na precipitação e evaporação, associada, por sua vez, a fatores como elevação, inclinação, latitude e geologia.

Sazonalidade das ATWS e AGWS

No que diz respeito às mudanças sazonais médias de Δ TWS para a ARB, elas atingem seu máximo positivo em abril, com um valor de 194.8 mm, e seu mínimo em outubro, com -199.5 mm. Ao representar as mudanças sazonais de Δ TWS para cada um dos pixels da ARB (Figura 2), o intervalo de magnitudes das anomalias aumenta, com valores extremamente positivos de 791.6 mm e negativos de -

750.8 mm. Os valores extremos máximos e mínimos são localizados na região central da ARB, sobre a AAS nos aquíferos Alter do Chão e Iça. Durante os máximos positivos, podem ser observados padrões espaciais sobre as planícies de inundação na reserva natural Mamirauá e nos rios Juruá, Purus, Madeira e Tapajós, evidenciando a forte influência das mudanças no armazenamento de água superficial sobre ΔTWS . As anomalias fortemente positivas abrangem os meses de MAMJ e negativas de SONDJ. Com magnitudes menores, a região Norte apresenta anomalias máximas negativas durante DJFM e positivas durante JJAS, mostrando uma diferença de fase de aproximadamente 2 a 3 meses em relação à região Sul da ARB. As mudanças no armazenamento ao longo do ano apresentam um padrão de deslocamento de Sul a Norte, causado pelo efeito das chuvas. De acordo com Marengo et al (2016), a variabilidade sazonal da precipitação na ARB é influenciada pela migração da Zona de Convergência Intertropical (ITCZ) e pelo sistema de monções da América do Sul. A ITCZ é uma faixa de baixa pressão que se move para o Norte e Sul do equador ao longo do ano. Assim, durante a estação seca (ASONDJ), predominam os ventos alísios ao Sudeste, levando ar seco do Atlântico para a ARB, reduzindo as precipitações e aumentando a evaporação. Por outro lado, durante a estação úmida (FMAMJJ), a zona de convergência intertropical se desloca para o Norte, transportando massas de ar úmido e aumentando a precipitação na região central.

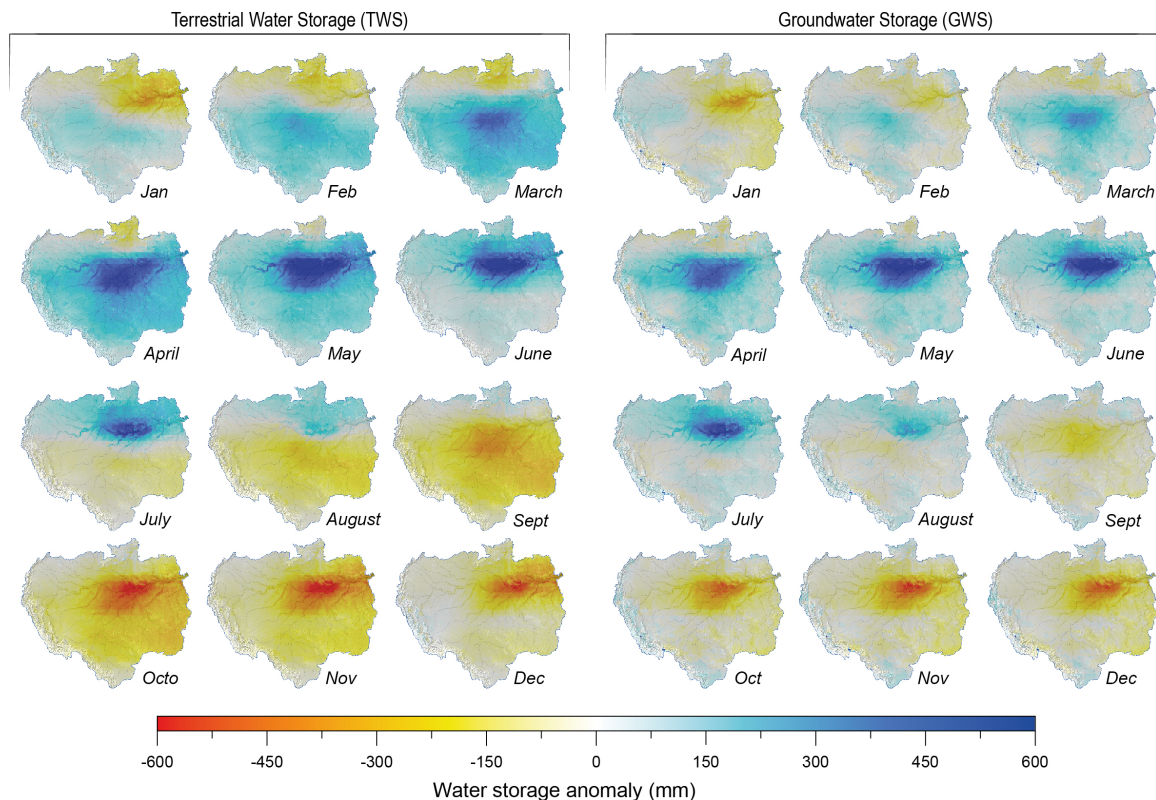


Figura 18. Distribuição espacial dos componentes sazonais de ΔTWS e ΔGWS para ARB.

Em geral, as variações sazonais de ΔGWS mostram um comportamento semelhante ao de ΔTWS . As mudanças sazonais médias de ΔGWS atingem seu máximo positivo de 111.1 mm no mês de maio e

o máximo negativo de -109.7 mm no mês de novembro, mostrando um desfase de um mês com os máximos de Δ TWS. Isso é atribuído a um tempo médio de trânsito da água dentro da ARB durante a recarga de solos e aquíferos. As variações espaciais para cada pixel mostram que os valores extremos positivos e negativos estão distribuídos sobre o centro da AAS, além de alguns pixels localizados sobre a cordilheira dos Andes. A forte influência do Δ SWS em alguns picos nevados dos Andes gera anomalias não representativas de Δ GWS. Entre os valores médios, podem ser identificados alguns padrões espaciais associados à localização dos aquíferos de Parecis e Ronuro ao Sudeste e do aquífero Boa Vista ao Norte. Da mesma forma, variações médias são observadas no Oeste, na depressão de Ucamara e nos Llanos de Moxos, caracterizados pela presença permanente de áreas inundáveis e corpos de água.

4.2.2. Variabilidade hidrológica

TWS

Na Figura 19, resumem-se aspectos como o comportamento sazonal, a tendência (Trend), tendência linear e desvio padrão das séries temporais de Δ TWS para a ARB. Em relação ao AAS, os aquíferos Iça e Alter do Chão apresentam as maiores mudanças sazonais, atingindo os valores máximos nos meses de AMJ nas regiões central e leste. Nessas mesmas áreas, compostas pelos pixels 2906, 2369, 2064 e 1749, encontram-se os maiores desvios padrão, com valores superiores a 400 mm. As tendências lineares mostram um ganho nas anomalias com valores superiores a 12 mm/ano, valor que diminui em direção ao leste, alcançando taxas de 2.7 mm/ano. Isso parece estar influenciado pelo aumento da inclinação das montanhas que limitam o rio Amazonas na porção central e leste do aquífero Alter do Chão. Com base na tendência (Trend) dessa região, é clara a alta susceptibilidade aos fenômenos El Niño e La Niña, como os ocorridos em 2010 e 2016. A oeste do AAS, nos pixels 421, 2648, 2774, 4339 e 5805, a variabilidade é menor do que na zona central e leste da ARB. A sazonalidade é maior em direção ao equador, diminuindo em direção às latitudes norte e sul. Os máximos positivos são alcançados nos meses de FM ao sul, deslocando-se para os meses de JJ ao norte, de acordo com as estações chuvosas. As variações de armazenamento de água ao longo dos rios Ucayali, Juruá e Purus são claramente representadas nos padrões espaciais do desvio padrão e da tendência linear. Na bacia média da ARB, entre Colômbia e Equador, o desvio padrão de Δ TWS é o mais baixo dentro do AAS. Na fronteira oeste do AAS, as tendências lineares mostram taxas nulas a moderadamente negativas (-1.9 mm/ano). A região compreendida entre os pixels 4339 e 5805, localizada no território peruano, compreende vários valores negativos e é caracterizada por inclinações médias de 3°, que são moderadamente superiores às dos outros pixels, influenciando a inclinação média dos cursos dos rios e seus caudais. Ao contrário da zona central do AAS, o El Niño e as secas de 2010 tiveram um impacto maior do que os ocorridos em 2016. Sobre o domínio dos Andes, apresentam-se as menores variâncias sazonais da ARB, atingindo valores máximos positivos nos meses de FMA e menores em SON, variando entre 113.5 e -104.7 mm. Comparativamente, os Andes são o domínio com os menores desvios padrão da ARB, apresentando tendências lineares nulas, ligeiramente negativas e ligeiramente positivas. Os

efeitos do degelo progressivo dos nevados localizados sobre os pixels 6851 e 7393 estão registrados na tendência linear e tendência, com episódios críticos durante 2005, 2010 e 2016. No sudeste da ARB, apresentam-se tendências lineares fortemente negativas nos aquíferos Parecis e Ronuro, com taxas que atingem 5.8 mm/ano. Essas áreas correspondem às cabeceiras das bacias dos rios Xingu e Tapajós, lugares fortemente afetados por incêndios, desmatamentos e construção de algumas barragens, de acordo com RASIG, 2022, onde predominam coberturas vegetais como agricultura, pastagens e áreas campestres (MapBiomias, 2022). A pressão exercida pela atividade antrópica resulta em aumento do escoamento superficial, redução das taxas de infiltração e recarga de aquíferos, com fortes efeitos a médio e longo prazo no armazenamento terrestre de água. As sub-bacias dos rios Xingu e Iriri são as mais afetadas, onde predominam pixels com tendências lineares moderadamente negativas a nulas. As séries temporais de tendência nessas áreas mostram flutuações constantes até o El Niño e as secas de 2015-2016, quando ocorre aumento das perdas ou déficits de Δ TWS, resultando em tendência linear negativa até 2021. Esse evento climático teve forte impacto nos Δ TWS dessa região da ARB, com efeitos a longo prazo que ainda não foram recuperados até 2021. Ao norte, sobre o Aquífero Boa Vista, os máximos positivos são obtidos nos meses de JJA e máximos negativos nos meses de FM.

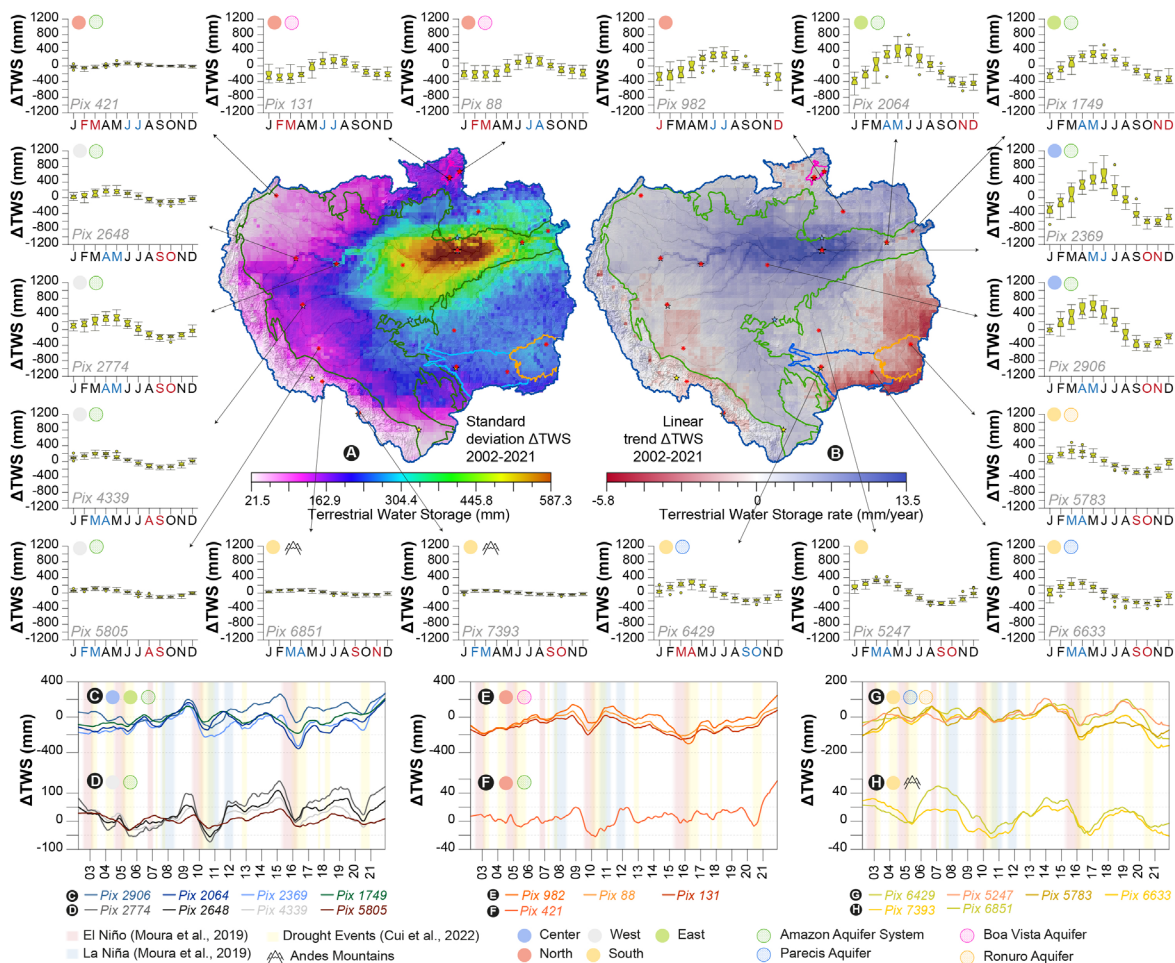


Figura 19. Representação espaço-temporal da variabilidade de Δ TWS para a ARB. Onde: A) SD de Δ TWS entre 2002-2021, B) tendência linear de Δ TWS entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente de tempo da sazonalidade entre 2002 e 2021. No

eixo temporal dos bloxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha.

GWS

O comportamento regional do desvio padrão apresentado na Figura 20-A mostra coerência entre a magnitude das mudanças sazonais dos pixels analisados. Os Δ GWS com desvios padrão maiores que 280 mm entre 2002 e 2021 estão claramente localizados dentro do AAS, entre os aquíferos Iça e Alter do Chão. Dentro desta região, os valores mais altos, em torno de 500 mm, distribuem-se ao longo dos eixos dos rios que constituem as planícies de inundação do rio Amazonas e algumas áreas adjacentes. Esses valores podem ser interpretados como: a) áreas altamente suscetíveis a mudanças no armazenamento de água superficial, onde as variações nos níveis dos rios repercutem diretamente nos Δ GWS; e b) locais onde os efeitos de eventos climáticos extremos adquirem maiores magnitudes, gerando grandes reduções no armazenamento de água subterrânea, que são rapidamente recuperados, tanto no comportamento sazonal quanto no de tendência. Esse comportamento pode estar fortemente relacionado aos coeficientes de armazenamento dos solos e aquíferos da área. A oeste, dentro do AAS, observam-se duas áreas com desvio padrão em torno de 130 mm, correspondentes aos Δ GWS influenciados pelas planícies de Moxos e Ucamara, bem como duas regiões com desvio padrão muito baixo em torno dos Pixels 421 e 5805. Alguns dos mesmos padrões espaciais se repetem em relação à tendência linear entre 2002 e 2021 na Figura 20-B, onde as regiões centrais da ARB, entre os Pixels 2906, 2369 e 2064, apresentam taxas positivas com valores máximos de 14.4 mm/ano. Em contraste, na região oeste do AAS, na direção de norte a sul, observam-se taxas principalmente próximas a zero ou ligeiramente positivas até chegar à zona próxima ao pixel 4339, a partir da qual as taxas de mudança se tornam negativas. Esse comportamento se estende ao sul ao longo da zona de sopé. A tendência registrada na Figura 20-D e F mostra uma recuperação para os Pixels 2774, 2648 e 4339 até meados de 2015, momento em que El Niño e uma sucessão de secas invertem essa tendência, resultando em um déficit ou perda de Δ GWS. A tendência linear moderadamente negativa pode ser observada na tendência do Pix 5805. Na Colômbia, sobre o Pixel 421, ocorre o oposto, uma forte recuperação do Δ GWS após a ocorrência dos fenômenos La Niña de 2011 e 2012.

Sobre os Andes, os desvios padrão geralmente são inferiores a 100 mm, exceto em locais com presença de nevados, onde os SWE afetam os valores de Δ GWS. As medições de SWE obtidas por GLDAS em alguns pixels sobre os Andes, onde se localizam alguns nevados, têm magnitudes altas. Estas magnitudes devem ser comparadas com medições in situ, no entanto, são elevadas em relação a outras variáveis hidrológicas. De acordo com a sazonalidade e tendência do Pixel 6851, desenvolvem-se Δ GWS com magnitudes que oscilam entre 2000 mm, medidas que não são representativas. Isso ocorre durante o balanço hídrico, ao subtrair anomalias fortemente negativas de Δ SWE do valor de Δ TWS, resultando em anomalias fortemente positivas de Δ GWS. Esses dados não são representativos na realidade, já que a grande maioria dessa coluna de água descongelada é transformada em escoamento.

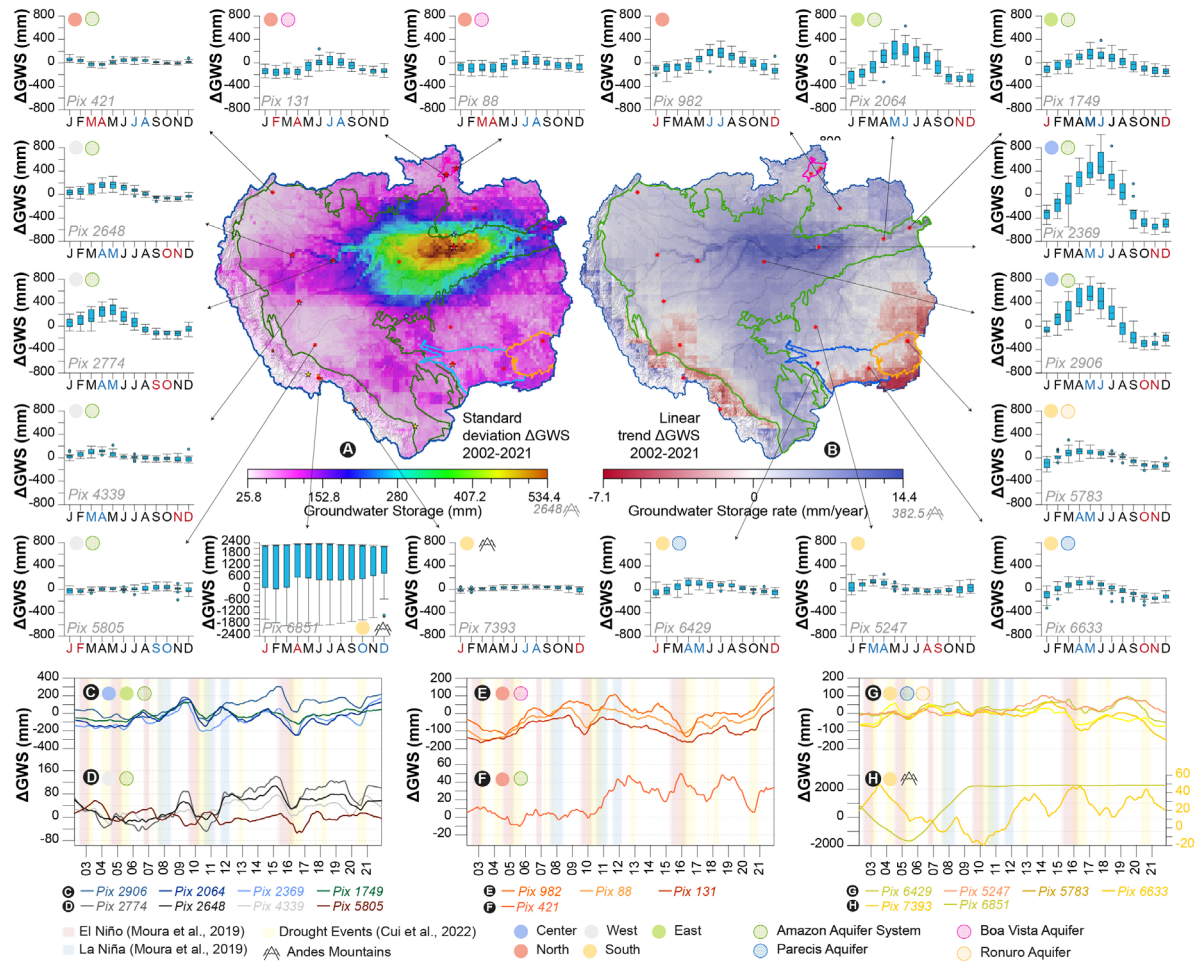


Figura 20. Representação espaço-temporal da variabilidade de ΔGWS para a ARB. Onde: A) SD de ΔGWS entre 2002-2021, B) tendência linear de ΔGWS entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente temporal da sazonalidade entre 2002 e 2021. No eixo temporal dos boxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha.

Ao comparar as representações de SD com ΔTWS , observam-se menores variações nos ΔGWS das regiões compreendidas pelos crátons Norte e Sul, os aquíferos Parecis, Ronuro e Boa Vista, demonstrando a forte influência dos ΔGWS do AAS na região central da ARB em relação a outras regiões. Os solos residuais presentes nos domínios cratônicos que limitam as imediações norte e sul do AAS desenvolvem ΔGWS com SD moderadas, as quais diminuem progressivamente. Os padrões espaciais na SD sobre os aquíferos Parecis e Ronuro estão relacionados à distribuição de sedimentos terciários, cuja porosidade e permeabilidade teórica devem ser moderadamente superiores às das rochas circundantes.

As tendências lineares fortemente negativas (-7.1 mm/ano) concentram-se no limite sudeste da ARB, sobre os limites sudeste dos aquíferos Parecis e Ronuro, onde a intensa pressão de atividades antrópicas, como queimadas, desmatamento e agricultura extensiva, reduziu a recarga de solos/aquíferos e aumentou a susceptibilidade aos efeitos climáticos.

Ao norte da ARB, os Pixels 131 e 88 sobre o aquífero Boa Vista mostram grandes mudanças de Δ GWS na tendência (Figura 20-E), com magnitudes semelhantes às do AAS central, embora com variações sazonais menores em comparação. A La Niña de 2011 e 2012 marcou sua recuperação máxima, seguida de uma perda progressiva de Δ GWS até o El Niño e as secas de 2016, momento em que começou uma recuperação constante até o ano de 2021. Este último assemelha-se a um comportamento sazonal com um período de 10 a 12 anos nos Δ GWS associado aos fenômenos de El Niño e La Niña.

Umidade do solo

O aquífero de Ronuro é a área com maior SD (190 mm), com valores médios multianuais de aproximadamente 500 mm e uma TL moderadamente negativa. Valores semelhantes de SD estendem-se sobre o aquífero Parecis, os crátons Norte e Sul, além da porção Leste do AAS, no aquífero Alter do Chão. As menores SD (19 mm) distribuem-se no nordeste da ARB e nas maiores altitudes dos Andes. Grande parte da ARB apresenta TL moderadas (-10 a -15 mm/ano) a fortemente negativas (<-15 mm/ano). As áreas com TL positivas são encontradas na porção centro-leste do AAS, parte do cráton Norte e nas zonas de transição entre a bacia média e alta, ao sul do Peru e ao norte da Bolívia (Figura 21). Esse comportamento generalizado e com taxas tão elevadas está relacionado às mudanças climáticas e aos fortes efeitos do El Niño e das secas.

A componente de tendência mostra os declínios prolongados no tempo, com forte influência dos fenômenos do El Niño e das secas.

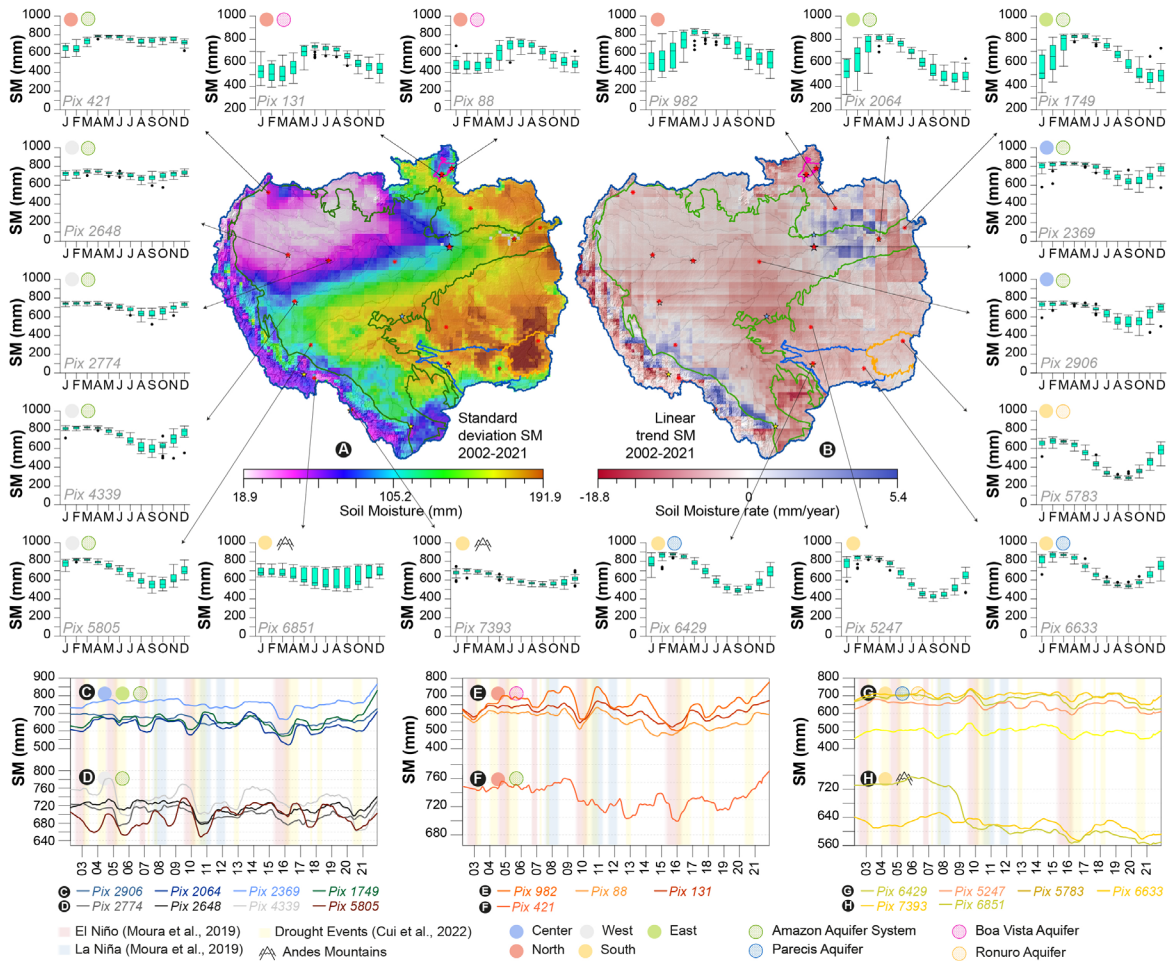


Figura 21. Representação espaço-temporal da variabilidade da umidade do solo (SM) para o ARB. Onde: A) SD da SM entre 2002-2021, B) tendência linear da SM entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente temporal da sazonalidade entre 2002 e 2021. No eixo temporal dos boxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha.

Água da Neve

Os glaciares das picos nevados com fortes mudanças no sudoeste dentro da ARB estão localizados no sul do Peru no pixel 6851 e na Bolívia no pixel 7393. Conhecidos como o nevado Ausangate na cordilheira de Vilcanota (Peru) e o Huayna Potosí, sobre a cordilheira Real ao norte da cidade de La Paz (Bolívia). Vários outros nevados se estendem ao longo dos Andes peruanos e bolivianos, no entanto, não estão registrados nos dados do GLDAS. O comportamento sazonal do pixel 6851, embora atinja valores máximos nos meses de AM e mínimos em ND, é fortemente influenciado por uma tendência negativa (Figura 22). Os boxplots mostram SD altas, com máximos em torno de 4000 mm e médias próximas a 0. A TL máxima é alcançada nesta área com -380.5 mm/ano. A tendência mostra uma recuperação no sudoeste até o início de 2006, momento em que começa um declínio acelerado até 2009, onde se estabiliza perto de 0 mm. No pixel 7393, a sazonalidade é mais heterogênea com variações inferiores a 2 mm. A TL é de -0.043 mm/ano e a tendência mostra um declínio contínuo, com pequenos episódios de recuperação em 2004, 2015 e 2018. O descongelamento progressivo dos nevados é causado

peelo aumento da temperatura média atribuído ao aquecimento global (Mark, et al., 2017). A radiação solar é absorvida pelo gelo e neve dos glaciares, elevando sua temperatura e promovendo o derretimento. Fatores como a altitude e inclinação podem influenciar os efeitos da radiação solar sobre os glaciares.

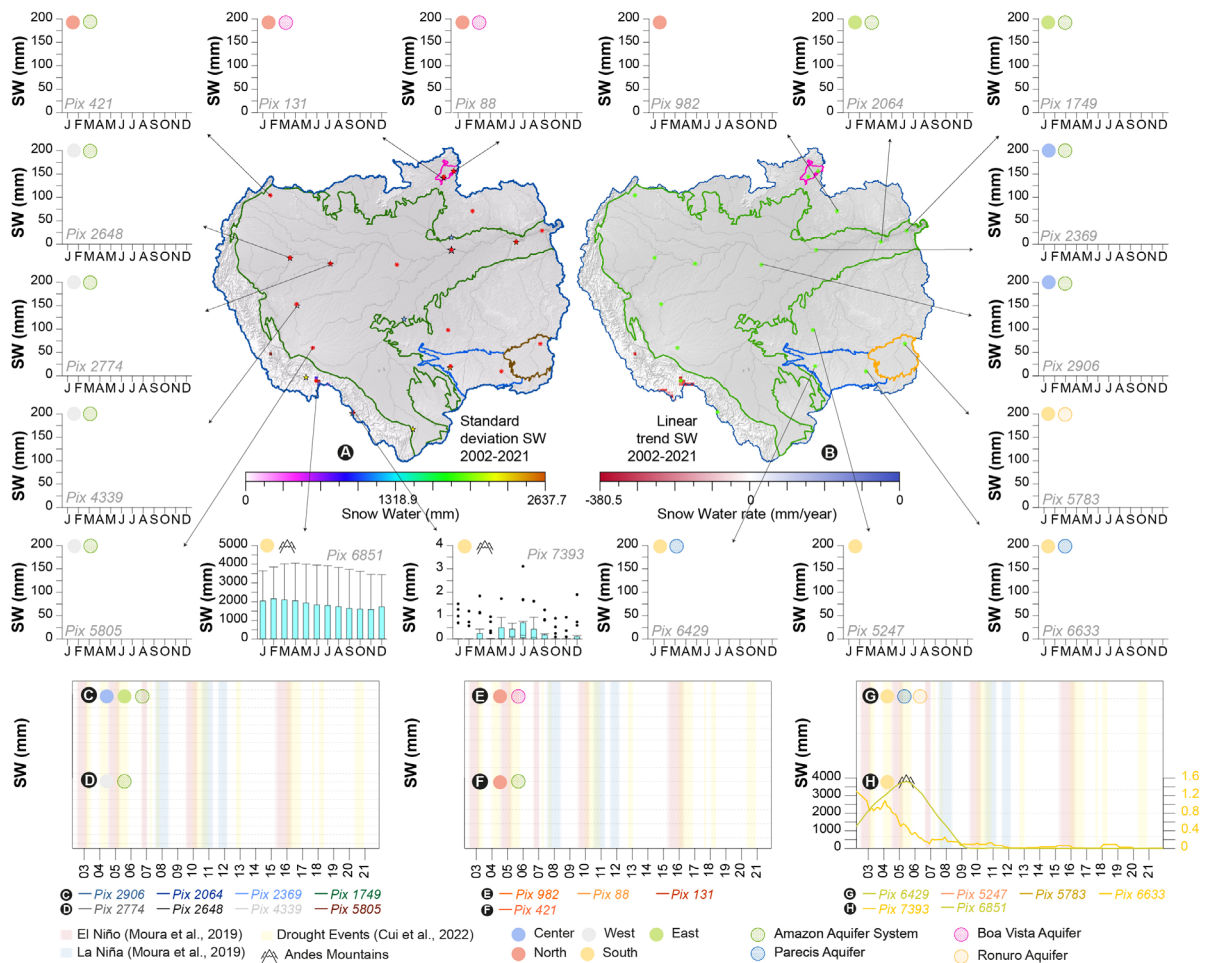


Figura 22. Representação espaço-temporal da variabilidade da água da neve (SW) para o ARB. Onde: A) SD de SW entre 2002-2021, B) tendência linear de SW entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente temporal da sazonalidade entre 2002 e 2021. No eixo temporal dos bloxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha.

Precipitação

As maiores precipitações ocorrem ao NO da ARB com valores médios de 4500 mm. Nessas mesmas áreas, o SD é inferior a 180 mm com TL negativas de -3 mm/ano. Na fronteira O do AAS e na transição entre a bacia média e alta, ocorrem "hot spots" de precipitação, conhecidos como chuvas orográficas (Fassoni-Andrade et al., 2021). Nestes locais, o SD é alto e alguns mostram TL negativas. Nos aquíferos de Parecis e Ronuro, as precipitações médias multianuais são de aproximadamente 1500 mm e o SD de 250 mm. O aquífero Parecis apresenta um ponto focal com TL negativas. De acordo com Coe et al., 2013, o desmatamento em grande escala pode gerar padrões de alteração na precipitação. As altas SD sobre o cratão S estão associadas ao intenso desmatamento e queimadas nessas áreas.

As séries temporais de tendência mostram como os efeitos dos fenômenos de El Niño causam acentuadas diminuições na precipitação em toda a ARB. No oeste, no pixel 2648, os fenômenos de La Niña de 2011 e 2012 marcaram o início de uma recuperação nas taxas de precipitação da região. Isso é consistente com as TL fortemente positivas e a localização da depressão de Ucamara. O pixel 421, localizado ao norte no território da Colômbia, mostra o maior déficit de chuva associado ao fenômeno de El Niño.

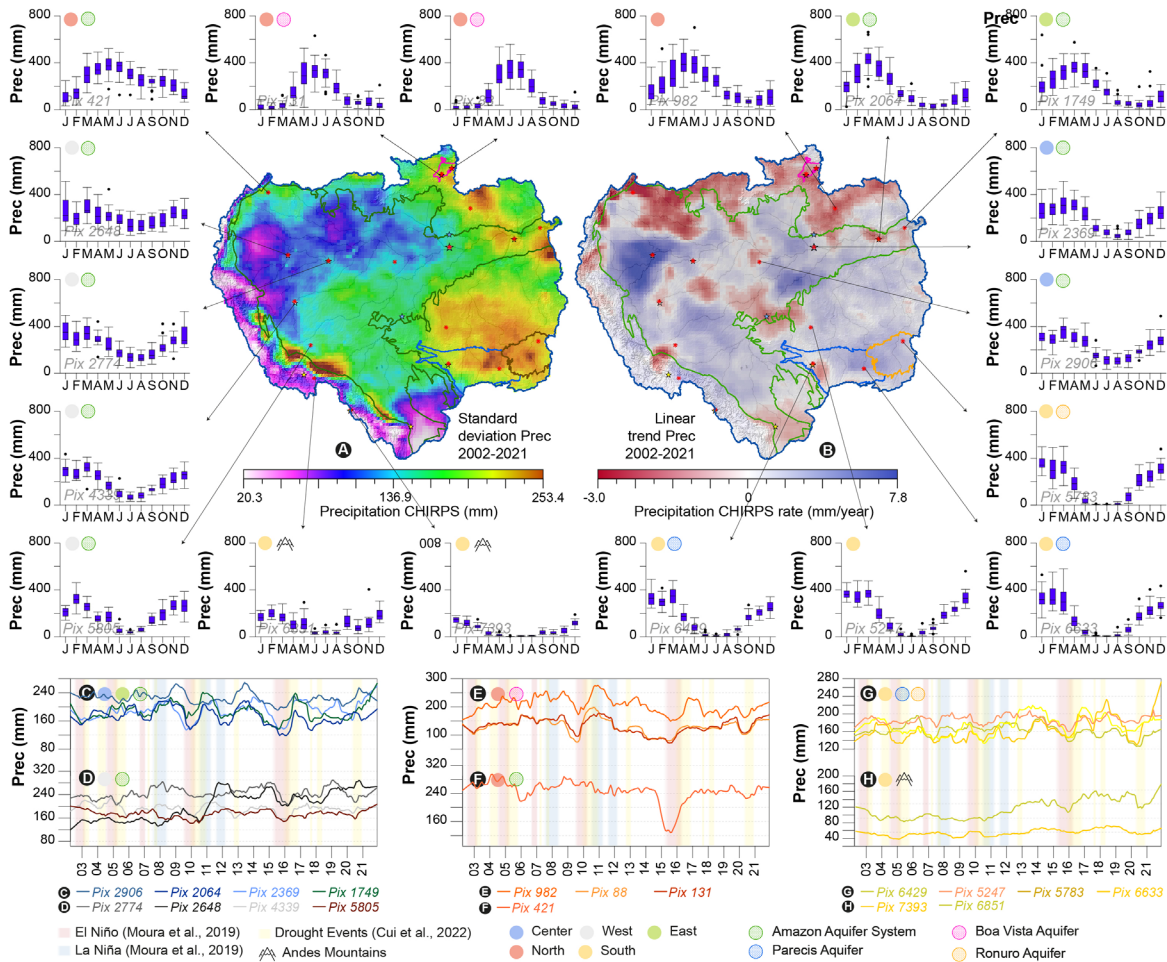


Figura 23. Representação espaço-temporal da variabilidade de precipitação do CHIRPS para o ARB. Onde: A) SD de TWS entre 2002-2021, B) tendência linear de TWS entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente de tempo de sazonalidade entre 2002 e 2021. No eixo temporal dos bloxplots, os meses com valores máximos positivos são mostrados em fonte azul e os meses com valores máximos negativos em fonte vermelha.

Evaporação

Dentro da sub-bacia do rio Xingu, Panday et al (2015), argumenta que o desmatamento tem causado uma redução de 3% na evapotranspiração e um aumento de 6% no escoamento, afetando negativamente o equilíbrio hídrico da região. Essas reduções moderadas podem ser identificadas nas TL negativas sobre a sub-bacia. Em muitas das áreas com baixos valores médios multianuais, ocorreram

incêndios e desmatamento, algumas das quais correspondem a TL negativas e maiores SD. Isso parece indicar alterações como redução nas taxas de evaporação e maior variabilidade (SD) resultantes de mudanças na cobertura vegetal e maior exposição do solo. As porções C, NO e O da bacia baixa apresentam as menores SD, o que está associado às formações florestais e, em certa medida, à precipitação, como reguladoras da variabilidade na evaporação. O domínio dos Andes e, particularmente, as maiores altitudes, também apresentam baixas SD (Figura 24).

As ST de tendência das zonas C e O não mostram grandes variações, como é bem representado pela SD. Para o norte, as variações são mais evidentes, com tendências ligeiramente negativas e uma forte queda durante os fenômenos de El Niño. Para o sul, as magnitudes das variações nas ST diminuem ligeiramente, sem episódios particularmente altos ou baixos.

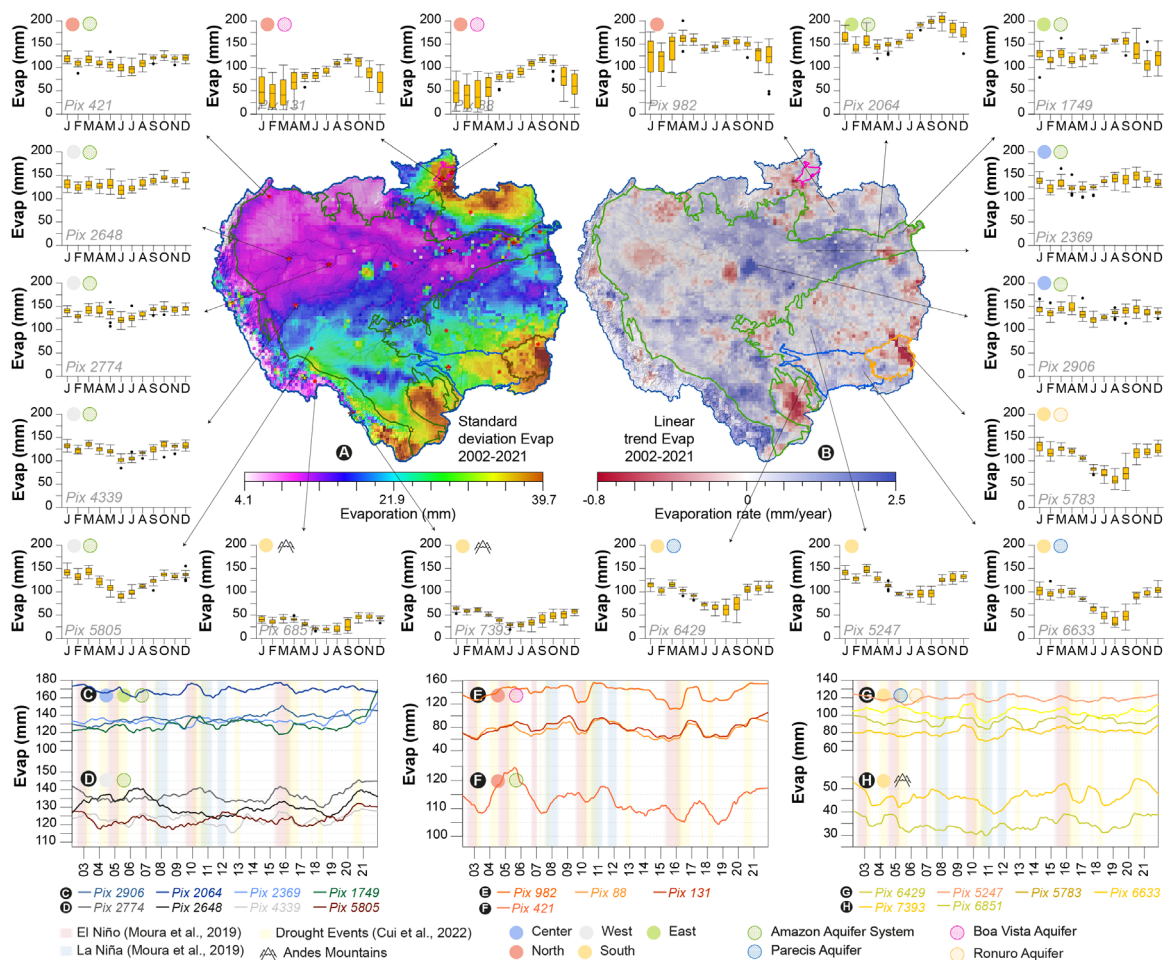


Figura 24. Representação espaço-temporal da variabilidade da evaporação para o ARB. Onde: A) SD da Evap entre 2002-2021, B) tendência linear da Evap entre 2002-2021, C,D,E,F,G,H) componente de tendência (Trend) dos pixels analisados. Os boxplots de cada pixel analisado representam os valores do componente temporal da sazonalidade entre 2002 e 2021.

4.2.3. Relação das componentes sazonal e total entre Δ GWS e outras variáveis

Compreender as relações espaço-temporais entre as principais variáveis hidrológicas em relação ao Δ GWS é de grande ajuda na interpretação dos dados. Essas relações com o Δ GWS facilitam a identificação de padrões de comportamento temporal em curto e longo prazo, bem como a definição

espacial de zonas cujas variações são diretamente ou indiretamente proporcionais ao ΔGWS . Esses critérios são necessários para propor hipóteses e fundamentar os resultados. O coeficiente de correlação de Pearson permite caracterizar essas relações. Deve-se levar em conta que as ST totais incluem as componentes estacionais, de tendência e residual. A componente sazonal armazena as maiores variações ou magnitudes dos ΔGWS , embora esteja limitada à compreensão temporal em curto prazo. As correlações estacionais são obtidas a partir de 12 valores mensais. Uma correlação sazonal direta (0 meses) pode ser afetada por um desfasamento (lag) ou atraso nos valores máximos e mínimos da ST (Figura 25). Esse atraso é causado por múltiplos fatores de grande interesse na interpretação.

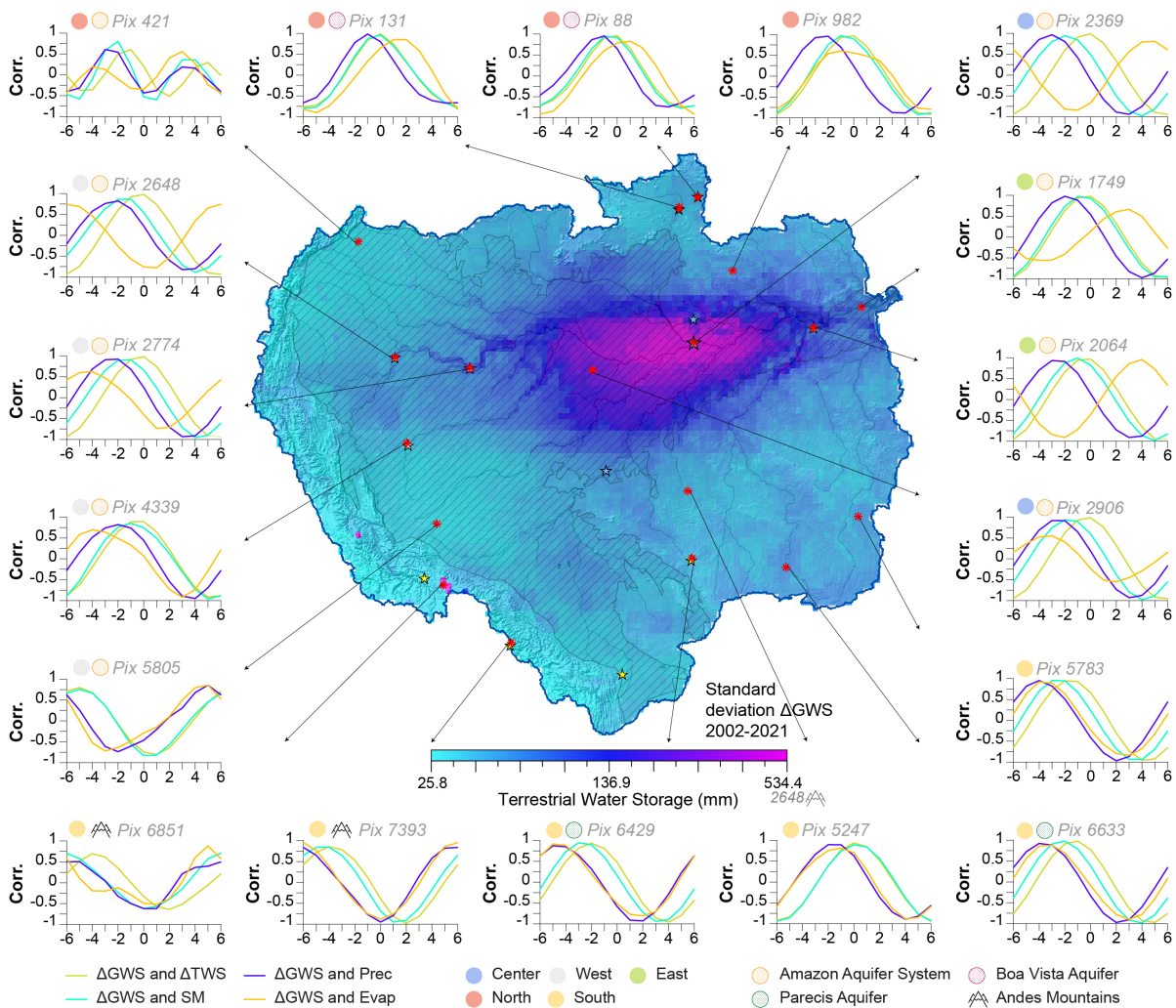


Figura 25. Variações na correlação sazonal entre o ΔGWS e as principais variáveis hidrológicas para os pixels de análise no ARB.

Para compreender a dependência entre o desfasamento (lag) sazonal e a máxima correlação sazonal, realizou-se uma análise de fase. Para essa análise, foram deslocados os valores estacionais em intervalos de desfasamentos positivos e negativos de 1 mês, obtendo diferentes correlações. Na Figura 26 estão representadas as correlações diretas das ST totais (Figura 26-A, B, C e D), as máximas correlações estacionais (Figura 26-E, F, G e H) e os desfasamentos (lag) (Figura 26-I, J, K e L). As correlações diretas com valores fortemente positivos ou negativos são interpretadas como locais onde

as ST têm altas associações com as componentes estacionais, de tendência e residual. O sinal positivo indica que são diretamente proporcionais.

Em termos estacionais, os valores máximos (crestas) e mínimos (vales) coincidem temporalmente, ou seja, estão em fase. Os sinais negativos indicam que as variáveis têm relações inversamente proporcionais, ou seja, uma aumenta à medida que a outra diminui. Sazonalmente, isso se expressa quando os valores máximos de uma variável coincidem com os valores mínimos da outra. Assume-se, então, que a componente de tendência e residual guardam as mesmas relações (e sentidos?) expressas nas correlações diretas das ST totais; no entanto, uma análise de correlação de tendência, TL e residual separadamente pode fornecer novas descobertas sobre as relações e interações espaço-temporais entre as variáveis. Uma análise de fase sobre as ST totais entre -12 e 12 meses pode ajudar a compreender o comportamento a médio prazo.

Quanto aos efeitos estacionais, sua avaliação é feita por meio da correlação sazonal máxima e do desfasamento. A correlação sazonal máxima mostra o quanto a correspondência existe entre os valores máximos e mínimos das variáveis em períodos de um ano (12 meses), independentemente do desfasamento. O desfasamento mostra quanto tempo e em que direção foi necessário deslocar a componente sazonal da variável analisada para alcançar a máxima correlação sazonal. O desfasamento pode ser interpretado em relação à correlação direta como as áreas onde o efeito sazonal pode interferir no sentido e magnitude das correlações. Um alto desfasamento pode inverter o sentido das correlações diretas. Do ponto de vista analítico, o sentido e a magnitude dos desfasamentos podem ser atribuídos a diferentes fatores condicionantes entre as variáveis, como propriedades hidráulicas dos materiais (solos ou rochas), heterogeneidade nos materiais geológicos (anisotropia, granulometria, mudanças de fácies, etc.), fatores topográficos (inclinação, elevação, relevo relativo, curvatura, etc.), efeito das coberturas vegetais e suas mudanças ou atividades antrópicas. A sensibilidade em detectar esses fatores condicionantes em pixels de 0.25° depende, por sua vez, da escala e magnitude em que eles se apresentam/desenvolvem.

As correlações diretas com Δ TWS (Figura 25-A) são predominantemente positivas para toda a ARB, fortemente positivas (>0.9) na bacia baixa e sobre o AAS, exceto nas áreas próximas ao pixel 5805. Os domínios cratônicos N e S, bem como os aquíferos Parecis e Ronuro, têm correlações positivas moderadamente altas (0.7 a 0.9). O domínio dos Andes tem variações mais heterogêneas, com valores extremos negativos (-1). Quanto à correlação sazonal, as áreas amarelas apresentam correlações praticamente nulas (<0.3). Isso pode ser interpretado como: 1) influência nula da componente sazonal nessas áreas, 2) correlação nula na ST total ou 3) correlações explicadas pelas componentes de tendência ou residual. Os desfasamentos de 0 meses ocorrem sobre o AAS em sua porção central de O a L, o aquífero Boa Vista, o cratão N e ao S dos Llanos de Moxos. Essas áreas combinam altas correlações estacionais e totais, mostrando uma forte conexão entre Δ TWS e uma resposta praticamente imediata a

curto e longo prazo. As áreas com desfases de -1 a -3 meses distribuem-se sobre a bacia média, onde os máximos Δ GWS ocorrem 1 a 3 meses após os Δ TWS. Sobre os Andes e nas proximidades do aquífero Boa Vista, ocorrem desfases com valores extremos. Maiores desfases são associados a maiores altitudes e inclinações. Também é associada uma forte influência dos materiais geológicos e efeitos antrópicos ao SE da ARB.

As correlações diretas máximas positivas com SM têm valores de 0.76 e distribuem-se ao E, C, N e O da ARB, nos domínios do AAS, aquífero Boa Vista e cratões N e S. Zonas com fortes correlações estacionais e desfases entre -3 e +2 meses. Heterogeneidades nos mecanismos de recarga de Δ GWS em relação à SM podem explicar as magnitudes e sentidos dos desfases estacionais. O aquífero Ronuro e a sub-bacia do rio Xingu contêm correlações diretas suaves a moderadamente positivas (0 a 0.5), alta correlação sazonal (> 0.8) e desfases de -1 a -4 meses, supondo uma recarga de Δ GWS causada pela SM com velocidades menores na porção sul em relação à norte. Em áreas com valores positivos entre 1 e 4 meses, atribui-se uma maior dependência de SM em relação ao Δ GWS. Os valores extremos de 5 meses ou -5 e -6 meses são desfases que podem ser provenientes do ano imediatamente anterior ou com desfases superiores a 6 meses.

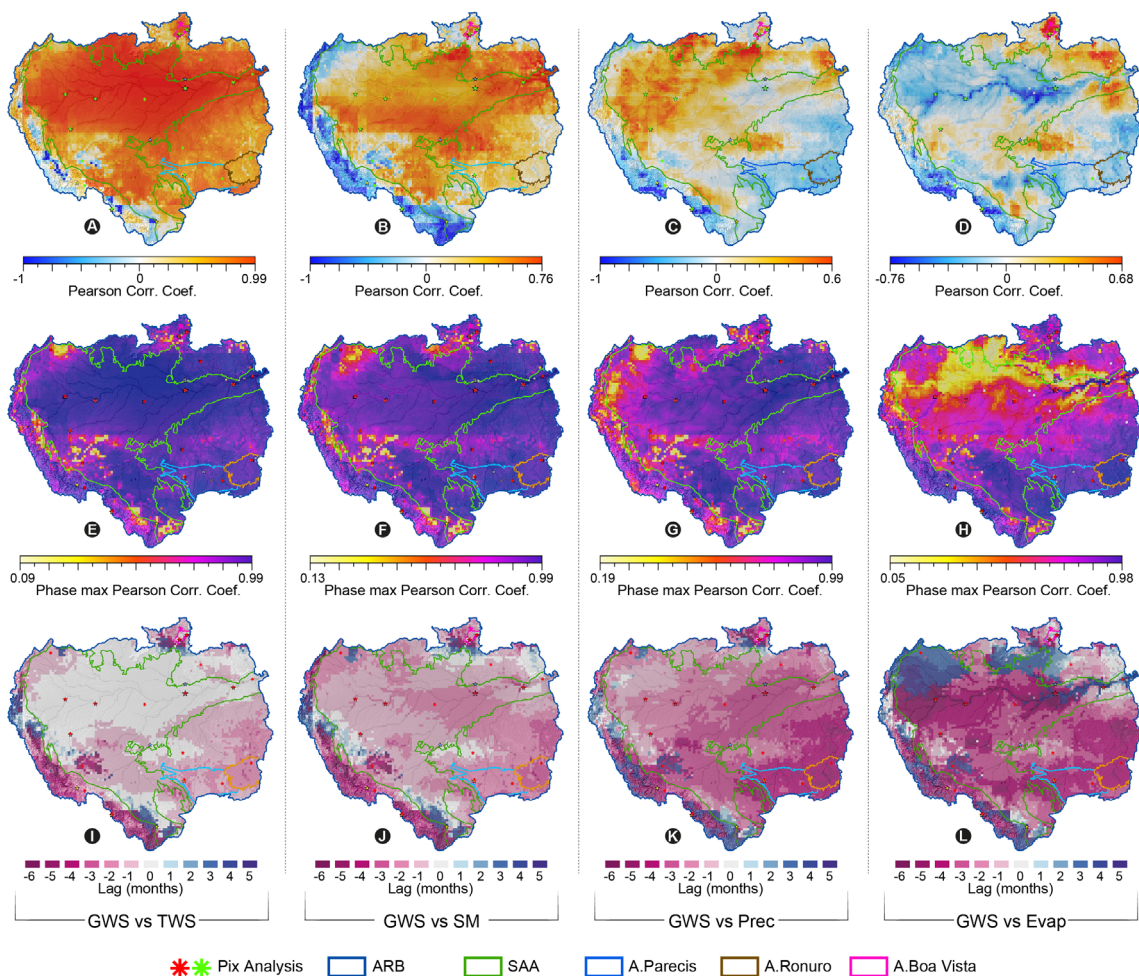


Figura 26. Mapas de correlação entre GWS e TWS, SM, Prec e Evap para a ARB. Onde A, B, C e D) correlação entre TS total, E, F, G e H) correlação sazonal máxima (fase) e I, J, K e L) defasagem para atingir a correlação sazonal máxima.

No que diz respeito à precipitação, as maiores correlações diretas positivas (entre 0.3 e 0.6) distribuem-se sobre as porções N, NO e O do AAS, sobre o aquífero Boa Vista e na sub-bacia menor do rio Jari. As correlações fortemente negativas são localizadas em 4 zonas sobre o domínio dos Andes, nas porções de Peru e Bolívia. Algumas dessas zonas (Andes) têm elevações acima de 2500m com presença de geleiras, pastagens, coberturas não florestais ou sem vegetação. E algumas delas têm elevações <2500m com vegetação florestal. Estas últimas têm precipitação média anual muito baixa (<700mm). A sazonalidade é moderada a forte (0.5 a 0.99) em grande parte da bacia baixa e média da ARB, com máximos na zona C e E, bem como nos Llanos de Moxos. As zonas com maior sazonalidade nos Andes ocorrem ao S do Peru e N da Bolívia. Os locais com correlações estacionais baixas a nulas (0.19 a 0.3) são amplamente encontrados entre os limites do AAS e os Andes a O, especialmente entre Colômbia e Equador. A forma e os padrões de distribuição espacial nos valores de correlação sazonal nula nesta região aumentam progressivamente para o Leste. Deve-se ter em mente que esta é a região com as maiores precipitações médias multianuais (>3500mm) dentro da ARB. Dentro da fração correspondente ao AAS e levando em consideração a topografia (elevação e inclinação) como critério morfológico, pode-se inferir que isso pode ser causado por facies mais argilosas, onde a recarga é muito baixa e não interfere nos Δ GWS. No entanto, a avaliação de outros fatores no ciclo hidrológico, como evaporação e escoamento superficial, pode fornecer outras explicações.

Os desfasamentos em geral para esta variável mostram uma predominância em sentido negativo, com valores entre -4 e -2 meses na bacia baixa. As zonas com menor desfasamento ocorrem no aquífero Boa Vista, nos limites NO, O e SO do AAS e sobre o cratão N, onde as mudanças estacionais da precipitação repercutem nos Δ GWS entre 1 e 2 meses depois. Os aquíferos Parecis e Ronuro têm um desfasamento de -3 a -5 meses, com correlação máxima de fase moderada a alta (0.7 a 0.9) e correlação direta com valores negativos baixos a moderados (-0.15 a -0.6). Isso significa que as correlações diretas nessas zonas são fortemente influenciadas pelo comportamento sazonal. Ou seja, a precipitação nos aquíferos é fortemente sazonal, com desfasamentos que invertem o sinal da correlação direta.

A evaporação é a variável com maiores desfasamentos na ARB, sendo os mais comuns entre -4 a -6 meses e 4 a 5 meses. As áreas com menores desfasamentos (entre -2 a 2 meses) correspondem a áreas com correlações diretas positivas, moderadas a moderadamente fortes (0.3 a 0.68). São áreas onde a correlação sazonal máxima também é moderada a moderadamente forte (0.3 a 0.75), evidenciando alta sazonalidade. Essas áreas distribuem-se ao E, NE, S e SO da ARB, nos domínios do AAS, cratônicos N e S, bem como no aquífero Boa Vista. Nessas zonas, ao contrário do que ocorre com a precipitação, as correlações estacionais são mais próximas das correlações diretas.

Dentro da ARB, as correlações estacionais máximas distribuem-se nas regiões mais altas dos Andes, nos Llanos de Moxos, aquífero Ronuro e sobre o curso do rio Amazonas, em seu percurso sobre o aquífero Alter do Chão. São áreas com altos desfasamentos e correlações diretas negativas com

magnitudes moderadamente fortes (-0.5 a -0.76). Ou seja, nessas áreas, as mudanças na evaporação têm um efeito retardado nos ΔGWS , onde em períodos de alta evaporação, a recarga de ΔGWS diminui com efeitos entre 4 a 6 meses seguintes.

4.2.4. TWI

Para compreender como a topografia pode influenciar os ΔGWS , estimou-se o Índice Topográfico de Umidade (TWI). Trata-se de uma medida amplamente utilizada em hidrologia para avaliar padrões de umidade do solo de acordo com as características topográficas da área. Com base no modelo digital de elevação MERIT Hydro-DEM, foram obtidos os parâmetros de largura de fluxo (Flow width), inclinação (slope) e área de captação específica (Specific catchment área) (Figura 27).

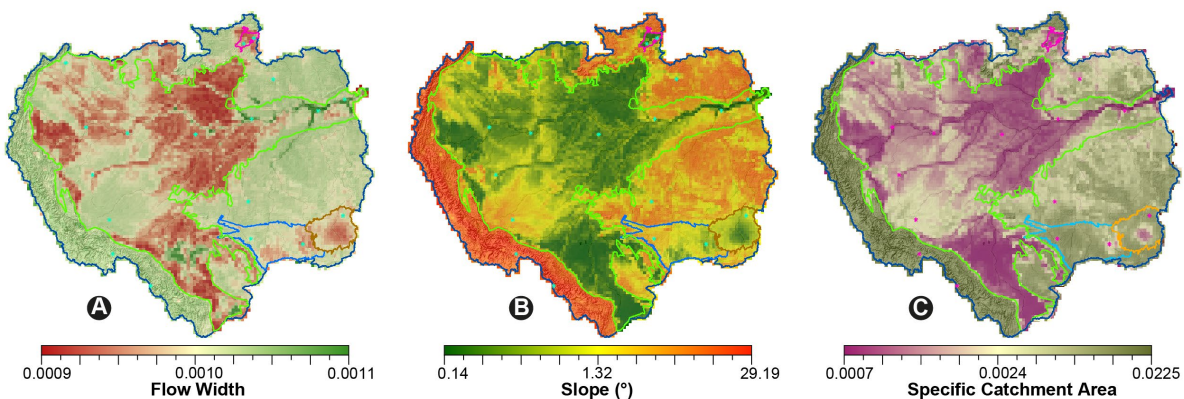


Figura 27. Parâmetros topográficos derivados do MERIT Hydro-DEM para a ARB em uma resolução espacial de 0.25°. Onde A) Flow width (FW), B) Slope em graus, C) Specific catchment área (SCA)

O processamento dos dados foi realizado em uma resolução espacial original de 90m, aplicando uma agregação com o valor médio para fazer um upscaling para 0.25° (aproximadamente 27500m). Uma comparação dos valores de TWI para a ARB (Figura 28) em sua resolução espacial original (90m) e escalada (0.25° ou 27500m aproximadamente) destaca as principais características espaciais. A bacia alta (> 1000m) ou domínio dos Andes apresenta valores médios de FW médios (0.001), altas inclinações (> 25°) e SCA altas, resultando em TWI baixos. De forma generalizada, esse domínio montanhoso apresenta fortes variações topográficas, algumas das quais correspondem a serras nevadas e vulcões. Abriga rios jovens com seções longitudinais de alta inclinação, seções transversais com vales em V e forte erosão de fundo, próprias de geformas denudacionais. O desenvolvimento de solos é intenso, embora suscetíveis a processos erosivos. As altas inclinações se manifestam em canais com pequenas áreas que contribuem para seu fluxo, dificultando a acumulação de umidade no solo.

Na bacia média e baixa, localiza-se o AAS. Em seu interior, no limite O e SO, as depressões de Mamirauá e Llanos de Moxos, bem como na região central, os TWI são altos (TWI>0). São zonas extensas, com valores médios de FW, SCA e inclinação muito baixos. Essas características facilitam a convergência de fluxos de água superficial e subsuperficial. Amplas planícies de inundação e baixa inclinação se expressam em rios meândricos e trançados, características de um ambiente agradacional.

Em contraste, duas áreas com valores fortemente negativos estão presentes entre a Colômbia e o Equador, e entre os pixels 4339 e 5805. Essas duas áreas possuem inclinações moderadas e alta densidade de drenagem, características atribuídas a uma textura predominantemente argilosa dos materiais. Na região CE e E do AAS, no aquífero Alter do Chão, os valores altos de TWI se concentram no leito do rio Amazonas, diminuindo rapidamente em direção às suas margens. Região onde as planícies de inundação nas margens do leito são menos extensas, rodeadas por um relevo colinoso de maior elevação e inclinação. Os traços morfológicos do aquífero Alter do Chão são consideravelmente mais pronunciados do que os do aquífero Iça. Os padrões espaciais de TWI formados por valores altos, médios e baixos dentro do AAS são semelhantes à SD e TL de Δ GWS.

Os aquíferos Parecis e Ronuro apresentam valores altos de TWI (-3 a -1). Dentro do aquífero Parecis, eles estão localizados em áreas de maior elevação, baixa inclinação e menor FW, especialmente em direção ao seu limite S. Essas áreas formam uma meia ou planalto com coberturas vegetais campestres e agricultura extensiva, onde as TL dos Δ GWS são fortemente negativas (-6 mm/ano). Da mesma forma, o aquífero Ranuro possui valores altos em sua área central, com baixas elevações, inclinações, SCA e FW, sobre o leito dos rios. Grandes extensões com agricultura extensiva estão localizadas em altitudes mais altas e coberturas com formações florestais nas zonas mais baixas (altos TWI). As áreas com agricultura coincidem com os pixels com maiores TL negativas (-6 mm/ano). Em termos gerais, esses aquíferos correspondem às cabeceiras das bacias dos rios Xingu e Tapajós, com boas condições morfológicas e hidrogeológicas para armazenar umidade no solo e recarregar os aquíferos. No entanto, de acordo com os valores de SD e TL de Δ GWS, a pressão contínua ao longo do tempo sobre as coberturas vegetais e a agricultura extensiva é atribuída à diminuição no armazenamento de água subterrânea. Esse fenômeno é causado pelo aumento da escorrência superficial, menor infiltração, maior temperatura superficial do solo e maior evaporação, o que diminui as taxas de recarga dos aquíferos, com efeitos apreciáveis a longo prazo.

O aquífero Boa Vista tem uma elevação média de 90m acima do nível do mar, com inclinações médias baixas ($<1^\circ$), SCA baixas (0.0007) e FW baixos (0.0009), refletidos em valores moderadamente altos de TWI (-1.3). Essa região forma uma depressão topográfica, rodeada pelo escudo guianense com expressões topográficas de maior elevação e inclinação, entre as quais se encontram os Tepuyes. Essa região é uma savana, com coberturas vegetais de formação campestre e algumas áreas agrícolas. Os valores médios anuais de precipitação (1400mm) e evaporação (800mm) são contrastantemente mais baixos do que nas regiões florestais. Os solos nas savanas são geralmente finos e pobres em nutrientes, favorecendo o crescimento de pastagens e vegetação rasteira. Em relação aos Δ GWS, apresentam SD <95 mm e TL ligeiramente positivas (4mm/ano), no entanto, TL de Δ TWS moderadamente negativas (-1.3mm/ano) nos pixels localizados sobre o leito do rio Branco.

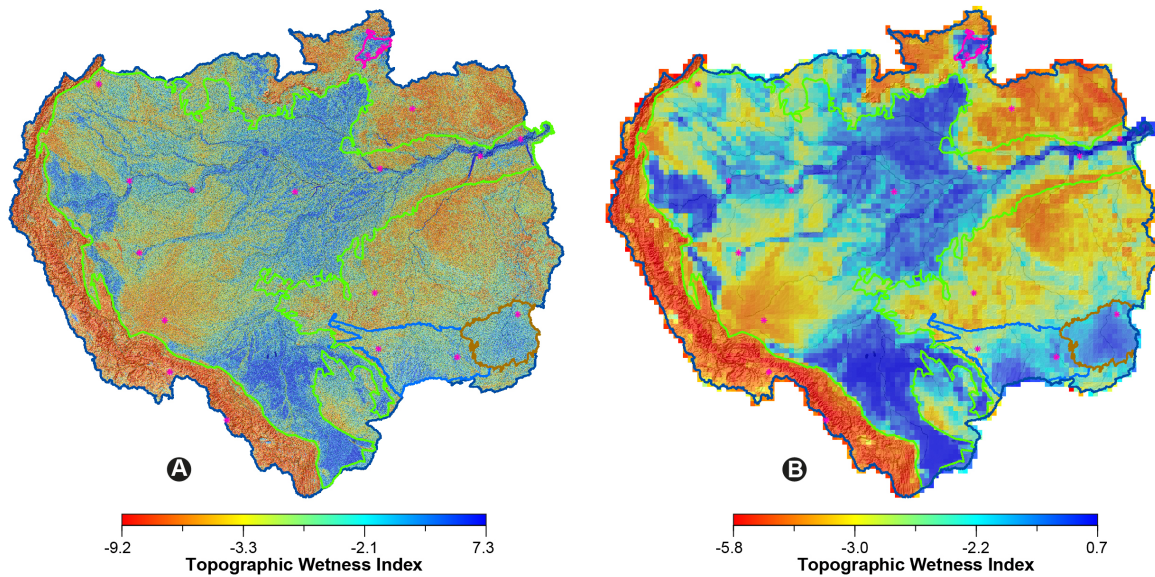


Figura 28. Índice de umidade topográfica (TWI) do ARB. Onde a resolução espacial do TWI é: A) 90m e B) 0.25° (27500m aproximadamente).

Nos domínios cratônicos N e S, os valores de TWI são particularmente altos devido a valores moderados a altos de inclinação, SCA e FW. No entanto, o desenvolvimento de solos com a capacidade de armazenar e transportar água é captado em Δ GWS. Essa relação é um pouco mais heterogênea e difícil de interpretar, mas alguns padrões espaciais com a SD de Δ GWS são apreciáveis.

Embora o TWI seja concebido como um índice, abrange muitas das características do relevo. Características que, de acordo com a matriz de correlação espaço-temporal (Figura 31), guardam correlações moderadas com AvgSurfT (0.5), LWdown (0.5), Psurf (0.5), Qair (0.4), ST0-200cm (0.5 a 0.6) e Tair (0.5). Mudanças na temperatura do solo (AvgSurfT e ST0-200cm) são causadas por fatores como o tempo de exposição à luz direta do sol, aspecto que pode variar de acordo com a altitude, inclinação e direção das encostas. Maiores altitudes têm temperaturas menores e menores altitudes têm temperaturas maiores, o que explicaria o sentido da correlação com TWI. Por outro lado, a pressão superficial ou atmosférica (Psurf) está em função da elevação, diminuindo a coluna de ar (pressão) à medida que aumenta a altitude. Essas mudanças de pressão causadas pela topografia podem influenciar a circulação do ar e, conseqüentemente, os sistemas de alta e baixa pressão, afetando os padrões climáticos (incluindo a umidade relativa ou Qair). Essas correlações indicam como as características hereditárias do relevo no TWI interferem na variabilidade climática e, conseqüentemente, nas mudanças de umidade do solo e Δ GWS.

Os contrastes morfológicos entre unidades geológicas são resultado dos processos de denudação e agraduação, em maior ou menor medida relacionados à resistência dos materiais, estruturas geológicas e intensidade dos agentes geomorfológicos ao longo do tempo. Apesar da diversidade e complexidade dos traços topográficos e morfológicos da ARB, as variações espaciais na magnitude do TWI estão intrinsecamente relacionadas aos processos hidrológicos superficiais e subterrâneos. Os contrastes entre

valores altos e baixos de TWI coincidem com os limites de vários aquíferos, revelando a forte influência da geologia na morfologia e comportamento hidrológico. Os aquíferos Parecis, Ronuro e Boa Vista apresentam fortes traços morfológicos e hidrogeológicos para a retenção de umidade, no entanto, sua natureza não confinada os torna altamente susceptíveis aos efeitos do clima e atividades antrópicas. Pesquisas mais detalhadas sobre a influência da topografia, morfologia e geologia em relação às variações do Δ GWS podem fornecer novos conhecimentos para compreender as relações hidrogeomorfológicas com as águas subterrâneas para a ARB.

4.2.5. Atividade antrópica

A ARB é um território extenso com diversidade de processos complexos que, em escala regional, são difíceis de caracterizar. Com base nas informações geográficas compartilhadas pela RAISG, 2022, e MapBiomias, 2022, alguns aspectos regionais relacionados às atividades humanas foram integrados. Extensas áreas de desmatamento e queimadas desenvolvem-se ao sudeste e sul da ARB, sobre as planícies de Moxos na Bolívia, sobre o cratão S e os aquíferos Parecis e Ronuro (Figura 29). Muitas dessas áreas correspondem a territórios indígenas, áreas protegidas e até mesmo zonas inundáveis onde se desenvolvem os pântanos. Áreas onde formações florestais e savânicas foram extensivamente transformadas em pastagens e agricultura. Sobre o aquífero Boa Vista, esse comportamento na transformação dos usos do solo se desenvolve com maior intensidade, especialmente em áreas baixas, locais que coincidem com maiores TWI. Alguns focos, não menos importantes, mas regionalmente menos proeminentes, ocorrem na Colômbia e em várias partes do Aquífero Alter do Chão.

Apesar de medidas de proteção, como a definição de áreas protegidas naturais e territórios indígenas, parecerem delimitar os padrões de incêndios em algumas partes da ARB, existem alguns locais onde, apesar disso, os incêndios têm sido promovidos. Nem todo incêndio tem origem antropogênica, mas claramente sua principal causa é na ARB.

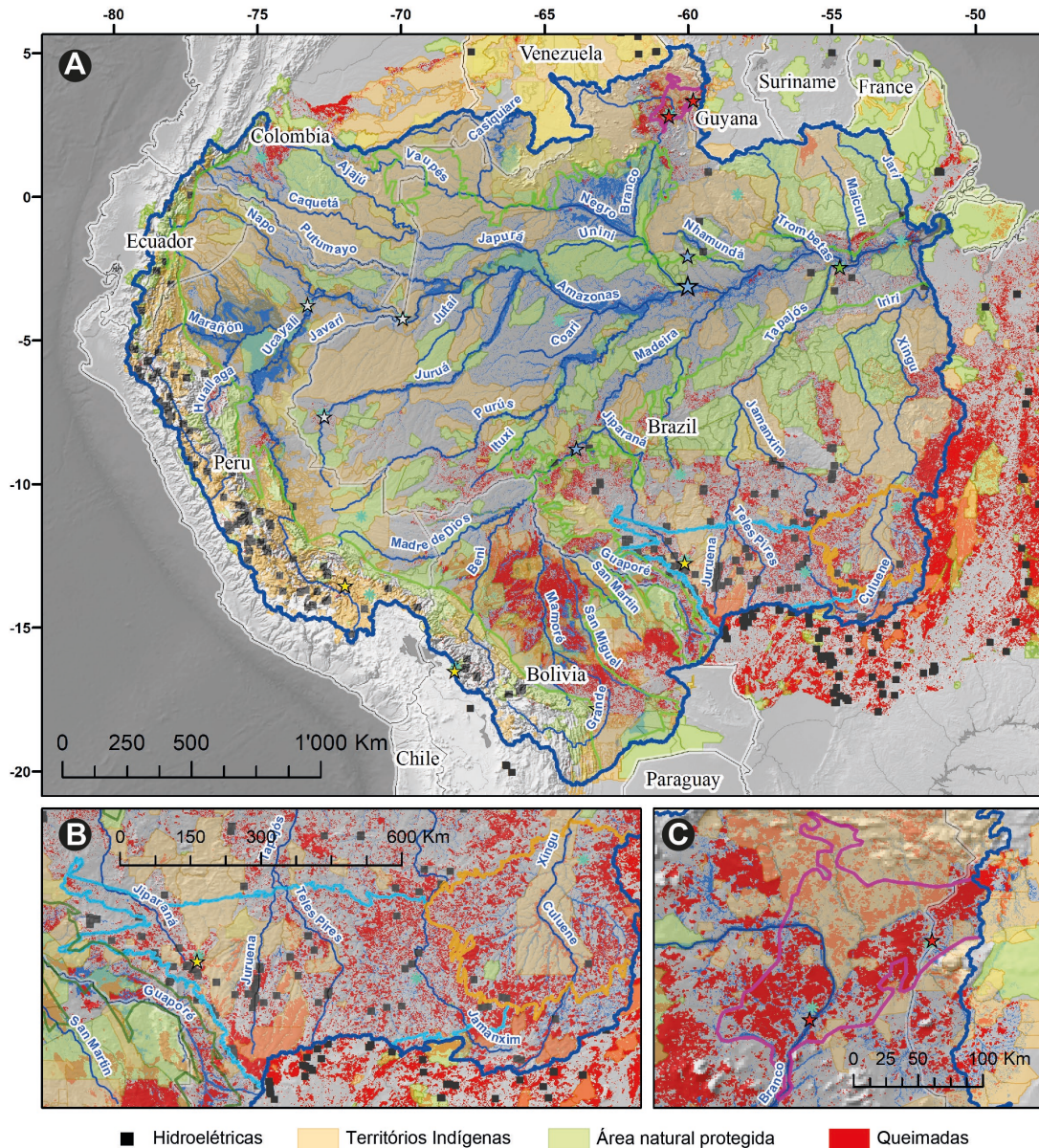


Figura 29. Localizam-se as hidrelétricas, áreas protegidas e zonas de incêndio em relação aos principais aquíferos da ARB. Onde A) ARB completa, B) aquíferos Parecis e Ronuro, e C) aquífero Boavista. Dados da RAISG, 2022

A maior concentração de hidrelétricas se localiza nos Andes, devido às suas características topográficas. Os dados de precipitação indicam uma maior variabilidade na precipitação sobre a região da hidrelétrica de Belo Monte. Essas alterações no ciclo hidrológico devem repercutir nos ΔGWS , no entanto, alterações fortes não são apreciáveis. Os efeitos mais claros das transformações dos usos do solo (Figura 30) ocorrem nas sub-bacias dos rios Xingu e na sub-bacia alta do rio Tapajós, sobre os aquíferos Parecis e Ronuro. Nessas áreas, a pressão constante e até acelerada na transformação de florestas em agricultura e pastagens reduziram as taxas de recarga dos aquíferos, promovidas por reduções na infiltração, aumento da temperatura do solo, evaporação e aumento da escorrência, que geraram TL de -7 mm/ano nos ΔGWS . A distribuição das maiores TL é observada nas áreas elevadas dessas sub-bacias, sobre áreas de recarga e trânsito das águas subterrâneas. Com o aumento da população, avanço da fronteira agrícola e indústria madeireira (legal e ilegal), esses processos podem se

acelerar. Os avanços no entendimento regional do SAA devem esclarecer os mecanismos de recarga e suas áreas de maior vulnerabilidade, de modo que, de forma conjunta entre governos, possam ser estabelecidos mecanismos de proteção, levando em consideração a natureza transfronteiriça.

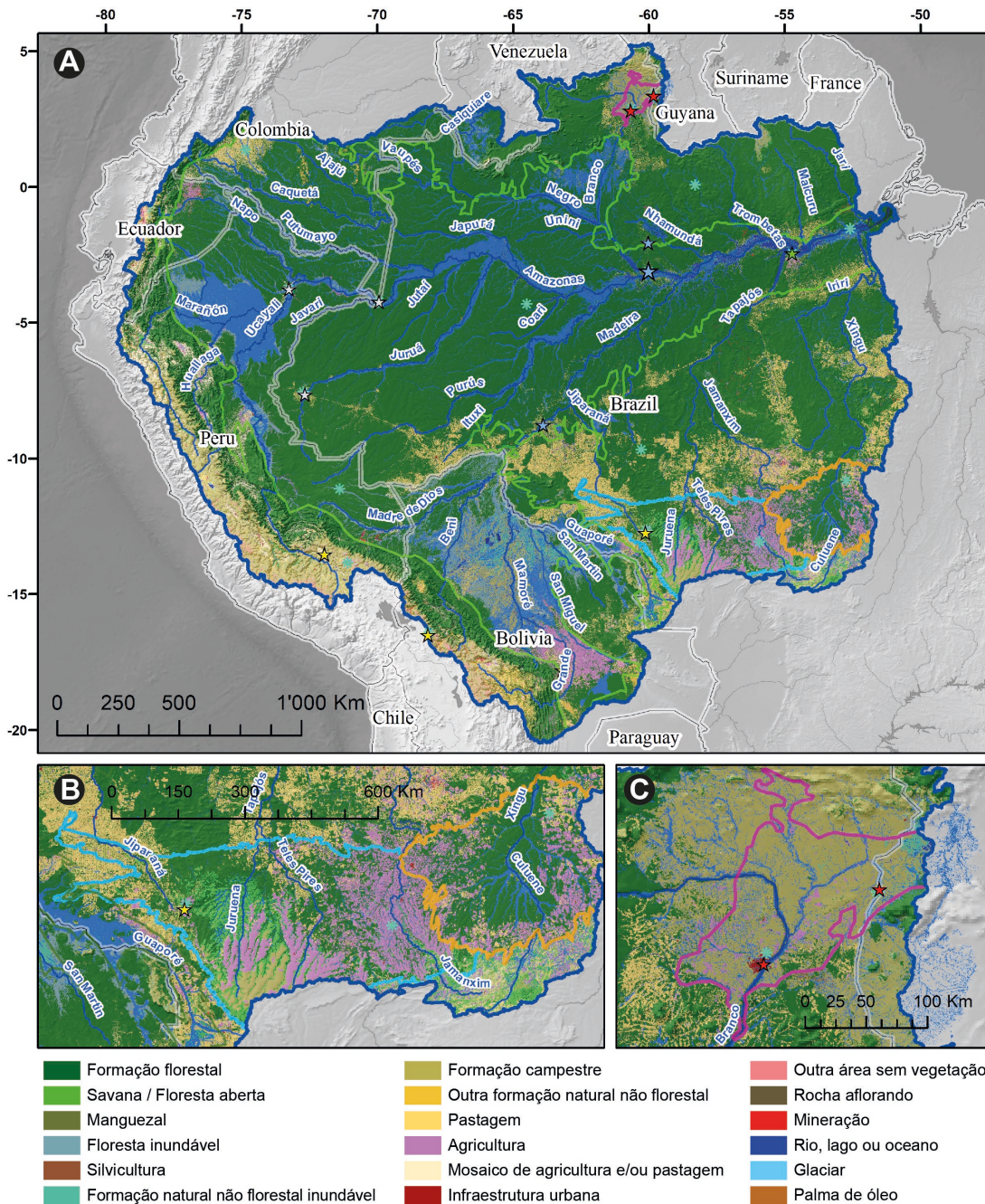


Figura 30. Vegetação da ARB. Onde A) ARB completa, B) aquíferos Parecis e Ronuro, e C) aquífero Boavista. Dados do MapBiomias, 2022.

4.2.6. Correlações espaço-temporais e com Δ GWS

A matriz de correlações espaço-temporais (Figura 31) descreve de forma geral as relações entre as 44 variáveis analisadas para a ARB. Foram incluídas variáveis morfométricas que, apesar de serem quantitativas, são constantes ou fixas e não variam ao longo do tempo de análise. Cada coeficiente de

correlação é estimado para a nuvem de pontos gerada para todos os pixels de 0.25° (7740), para cada mês (237), entre as duas variáveis, ou seja, um total de 1834380 pontos que contêm padrões espaciais e temporais da ARB. Temporalmente, foram utilizadas as ST totais, uma vez que elas abrangem a totalidade do sinal. Os resultados mostram que as maiores correlações com ΔGWS são ΔTWS (0.8), Qsb (0.3), $SM100-200$ (0.3), $RootMoist$ (0.2), $PotEvap$ (-0.2), $SM40-100cm$ (0.2) e $SM10-40cm$ (0.2). Embora à primeira vista essas correlações sejam baixas, elas estão acima do nível de significância de 1%. Variáveis como precipitação ($Precip_CHIRPS$, $Rainf$, $RainfT$), escoamento superficial (Qs), evaporação potencial ($PotEvap$), fluxo líquido de radiação de onda longa ($Lwnet$), fluxo líquido de calor latente (Qh), evaporação de água do dossel ($ECanop$) e água superficial do dossel ($CanpInt$) apresentam fortes correlações com a umidade do solo ($RootMoist$ e $SM0-200cm$) e, conseqüentemente, com os ΔGWS .

4.2. Efeitos do clima e das atividades antropogênicas nos Δ GWS da ARB

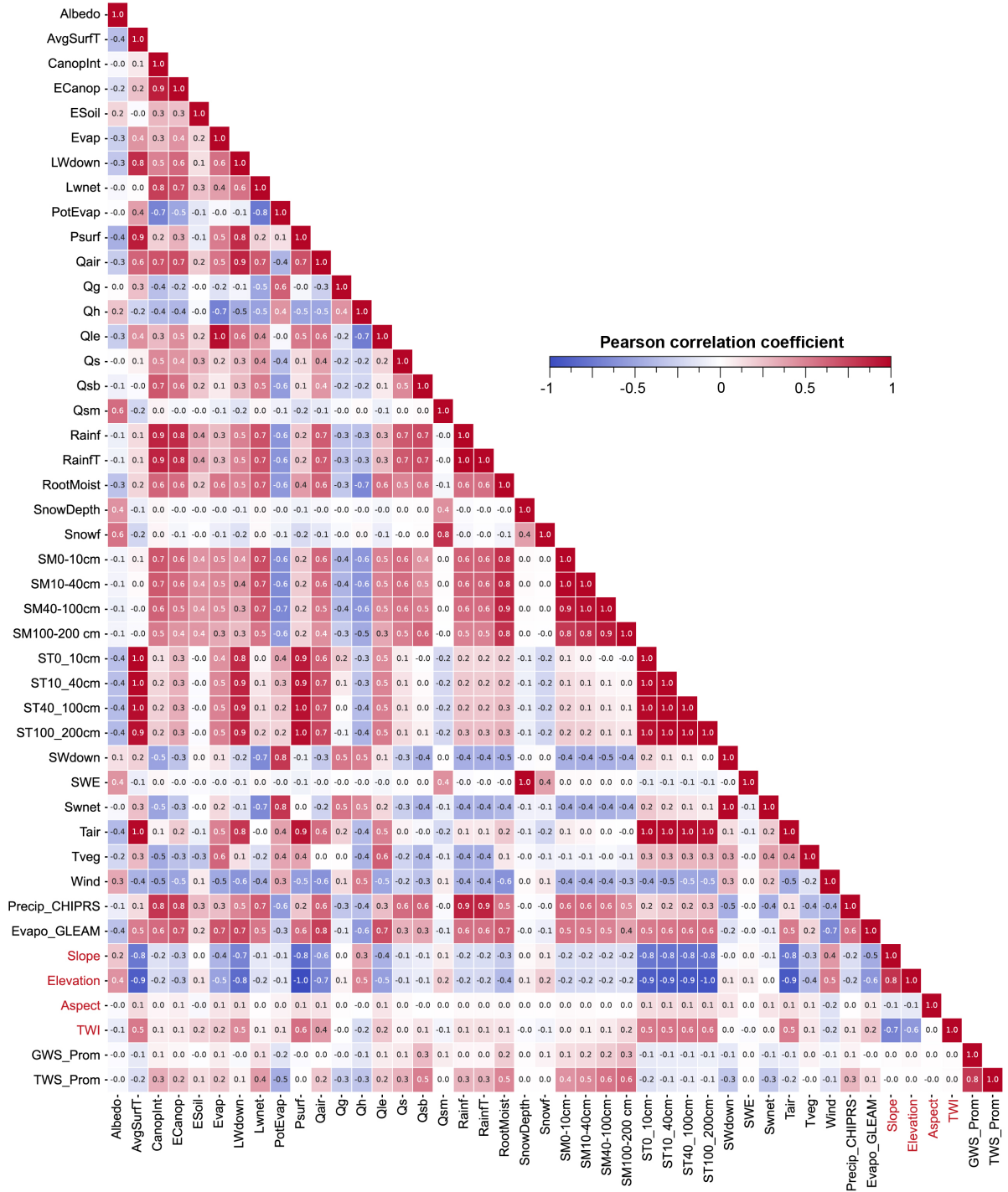


Figura 31. Matriz de correlação espaço-temporal para ARB. Variáveis constantes ou fixas em fonte vermelha.

Para compreender melhor como essas correlações espaço-temporais se comportam espacialmente, foram obtidas as correlações entre as ST totais de cada variável em relação aos Δ GWS e foram representadas espacialmente (Figura 32). As correlações diretas entre as ST totais são fortemente explicadas pela sazonalidade ou comportamento a curto prazo, o que torna difícil a interpretação dos comportamentos a longo prazo. Correlações positivas altas são diretamente proporcionais, onde se assume que existem desfases de -1, 0 ou 1 mês. Correlações negativas altas são inversamente proporcionais, causadas por desfases de -5 a -6 meses ou 5 a 6 meses. Embora as ST totais contenham

em seu sinal as componentes de tendência e residual, sua menor magnitude torna difícil identificá-las em uma correlação direta. Uma análise comparativa de cada componente temporal das ST aplicada entre as diferentes variáveis em relação aos Δ GWS contribuiria para caracterizar com maior detalhe as relações espaço-temporais dentro da ARB no médio e longo prazo.

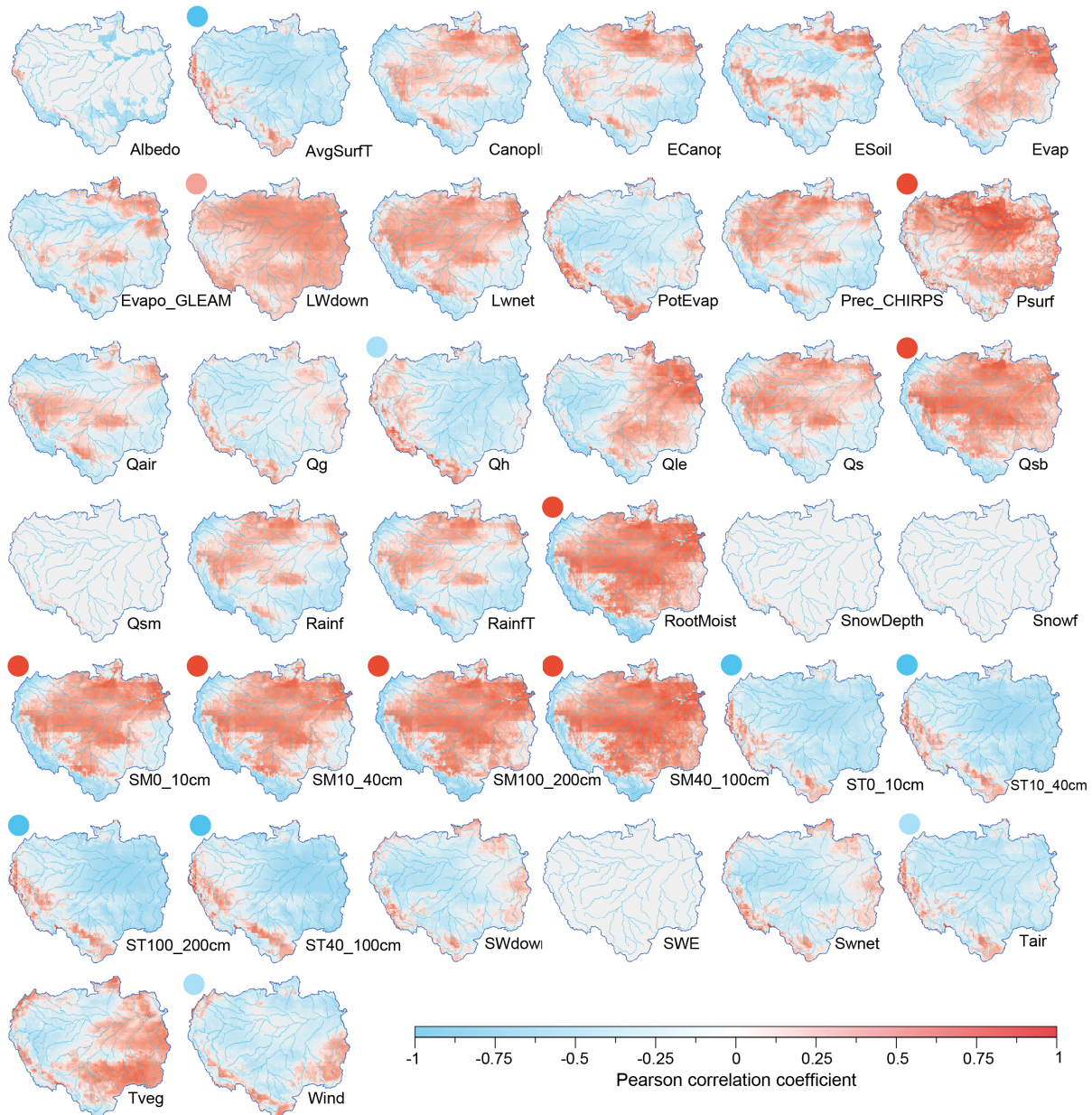


Figura 32. Mapas de correlação entre o GWS TS e as 38 variáveis espaço-temporais para o ARB.

A distribuição dos valores de correlação nas ST totais dos Δ GWS em relação a cada uma das variáveis (Figura 33) mostra que as variáveis com correlações predominantemente positivas ou negativas apresentam relações com sentidos opostos entre os domínios da bacia alta em relação à bacia média e baixa. Isso é interpretado como uma alta correlação em termos gerais com a variável (por exemplo, AvgSurfT, SM0-200cm ou ST0-200cm), embora com diferenças significativas de fase entre os domínios. A variável de pressão atmosférica é particularmente interessante porque pode causar

pequenas variações no nível freático de aquíferos não confinados (livres). Quando a pressão atmosférica aumenta, o nível freático diminui (Freeze e Cherry, 1979). A magnitude dessas mudanças também é afetada pela textura do solo e seu coeficiente de armazenamento. Esperaríamos encontrar correlações fortemente negativas sobre os principais aquíferos, no entanto, as correlações são positivas. Os padrões espaciais das correlações positivas se assemelham, em algumas áreas, aos limites dos aquíferos. Embora o sentido das correlações não seja coerente analiticamente, uma análise da SD e LT da Psurf pode fornecer mais características espaciais e quantitativas que podem ser atribuídas a mudanças na textura do solo ou em suas propriedades hidráulicas. Variáveis como a umidade do solo (SM0-200cm) são diretamente proporcionais aos Δ GWS para a bacia baixa, enquanto variáveis como a temperatura do solo (ST0-200cm) são inversamente proporcionais. Esses comportamentos são teoricamente coerentes.

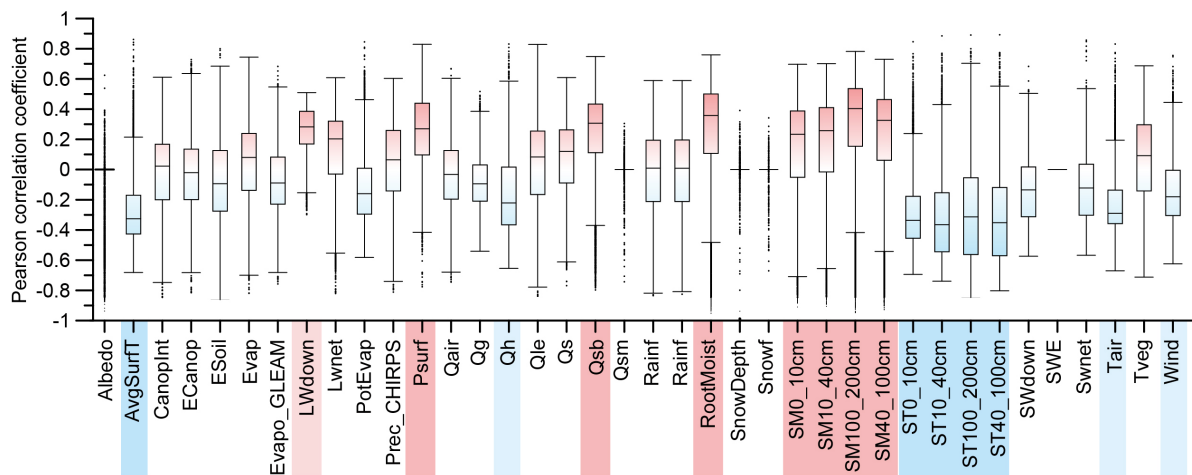


Figura 33. Boxplot dos valores de correlação entre o GWS e as 38 variáveis espaço-temporais para o ARB.

4.2.7. PCA Pearson

A Análise de Componentes Principais (PCA) é realizada utilizando as correlações entre Δ GWS e as variáveis espaço-temporais como dimensões (Figura 32). Os componentes principais (PC) são as novas variáveis, resultado de combinações lineares das variáveis originais (correlações), que respeitam as direções de maior variância entre os dados e resumem as principais relações entre si. A aplicação dessa análise permite resumir em poucos componentes as relações entre as diferentes variáveis em relação aos Δ GWS dentro da ARB, permitindo identificar (zonificar) áreas com comportamentos semelhantes ou homogêneos. A representação espacial na composição RGB dos componentes principais PC1 (Vermelho), PC2 (Verde) e PC3 (Azul) facilita a compreensão das relações lineares com Δ GWS dentro da ARB.

Um total de 38 variáveis são utilizadas, com 7740 pixels ou dados por variável. Considerando que os coeficientes de correlação variam entre -1 e 1, não foi realizada uma normalização das variáveis. A proporção de variância e o número de componentes principais são apresentados na Figura 34. A variância explicada individualmente por PC1 é de 55%, por PC2 é de 21% e por PC3 é de 8%, com um

total de variância explicada para as três primeiras PC de 85%. A interpretação dos padrões e características das PC1, PC2 e PC3 facilita a compreensão e definição de zonas na composição RGB..

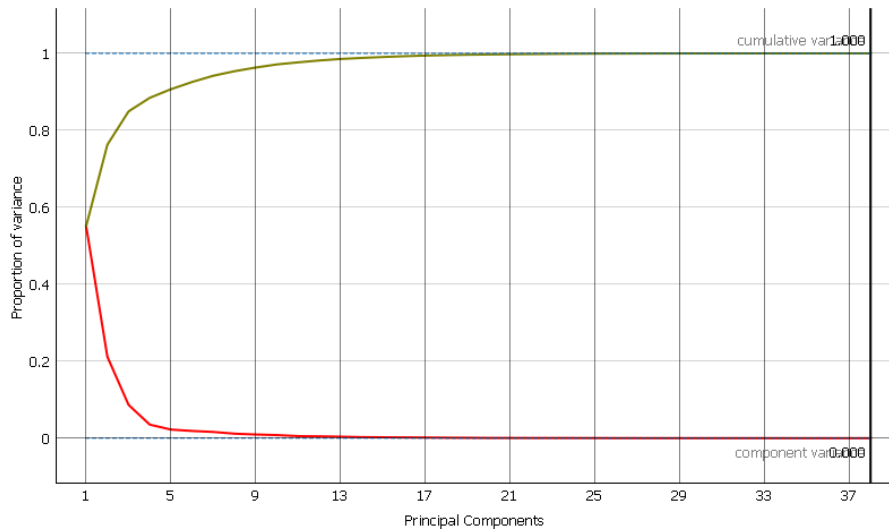


Figura 34. Proporção da variância cumulativa explicada.

As variáveis localizadas na periferia dos diagramas de dispersão são aquelas com maiores contribuições para a variância explicada pelas PC (Figura 35). Os loadings das variáveis de temperatura do solo (ST10-200cm) contribuem para a variância explicada pelas PC1 e PC2. Essas três variáveis (ST10-40cm, ST40-100cm e ST100-200cm) também têm os maiores loadings positivos do PC2. A transpiração (Tveg) é o loading negativo de maior magnitude do PC2, com uma contribuição importante também para a variância do PC3. A evaporação de GLEAM (Evapo_GLEAM) e de GLDAS (Evap), bem como o fluxo líquido de calor latente (Qle), são loadings com magnitudes importantes nas três componentes, além de serem os máximos positivos do PC3. Outras variáveis que influenciam nas três componentes, embora com magnitudes menores, são: a pressão atmosférica (Psurf), a taxa de evaporação potencial (PotEvap), o fluxo líquido de calor sensível (Qh) e a umidade específica (Qair). As variáveis com contribuições mínimas ou nulas estão associadas às mudanças no armazenamento de água em níveis (SWS), como SnowDepth, Qsm, Snowf e SWE.

Os loadings com maiores magnitudes do PC1 são explicados pelos conteúdos de água armazenados no solo (RootMoist, SM0_10cm, SM10-40, SM40_100, SM100_200cm e Qsb) e dossel das plantas (CanopInt), bem como pelo fluxo líquido de radiação de onda longa (Lwnet). Essas variáveis explicam as principais mudanças nos Δ GWS dentro da ARB, corroborando a forte influência da água armazenada nos solos, na vegetação e nas trocas de energia superficial. Em relação aos maiores loadings do PC2, eles estão associados à precipitação medida por GLDAS e CHIRPS (Ranf, RanfT, Prec_CHIRPS), à transpiração (Tveg) e à temperatura do solo entre 10 a 200cm (ST10_40cm, ST40_100cm e ST100_200cm). Com 21% da variância explicada, a temperatura do solo, a precipitação e a água condensada pelas plantas são fatores que influenciam significativamente os Δ GWS, variáveis fortemente influenciadas pelas mudanças climáticas e atividades humanas dentro e fora da ARB. O PC3

é explicado pela Evaporação (Evap, Evap_GLEAM e ESoil), a temperatura superficial (Tair e AvgSurfT) e fluxos de radiação térmica (SWdown, Swnet e Qle).

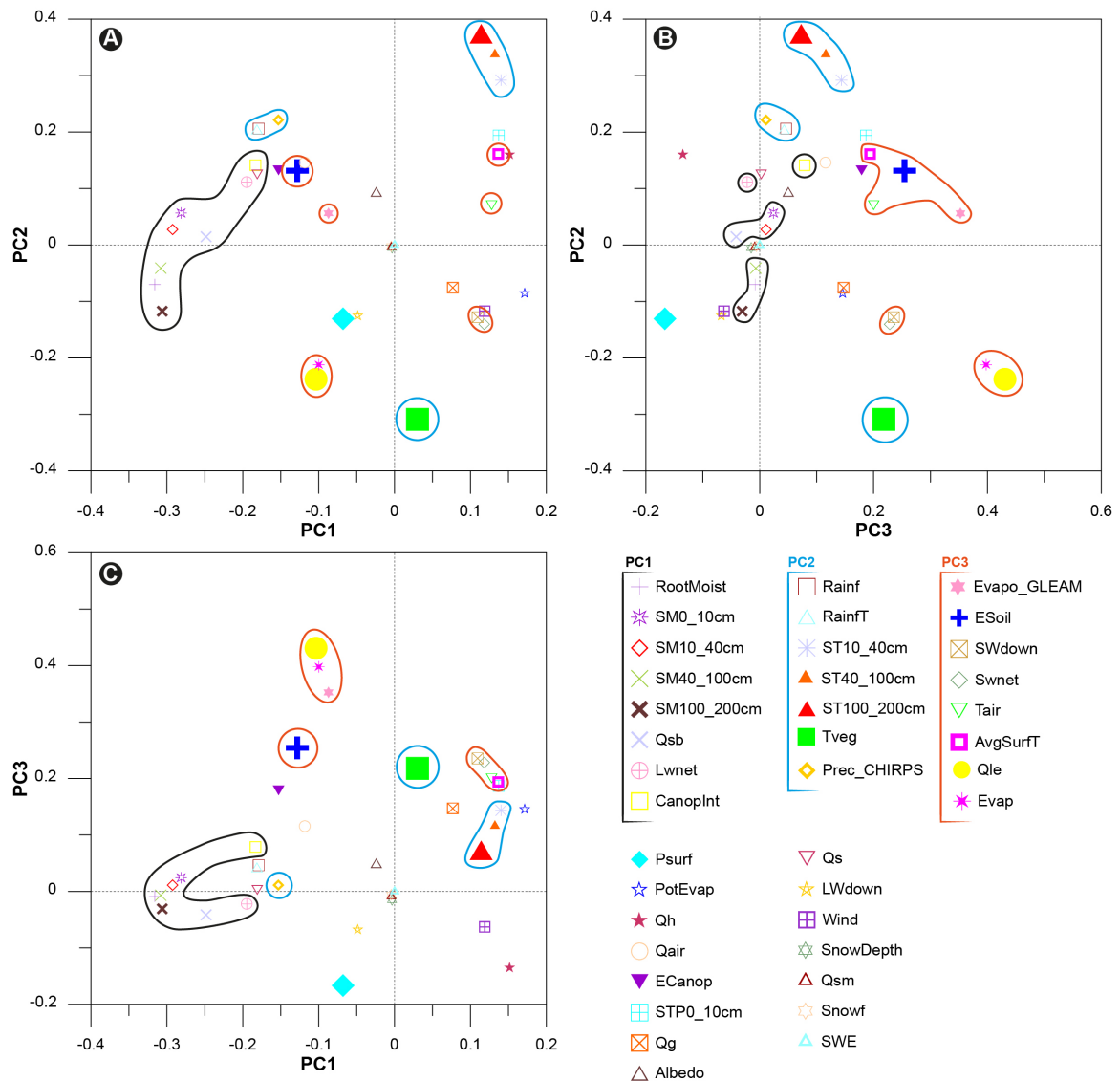


Figura 35. Gráfico de dispersão dos loadings de PC1, PC2 e PC3. A magnitude extrema das cargas de PC1 é destacada na linha preta, a de PC2 na linha azul e a de PC3 na linha vermelha. Onde: A) PC1 com PC2, B) PC2 com PC3 e C) PC1 com PC3.

PC1

A transformação dos dados do PC1 na Figura 36 representa em azul os valores negativos e em vermelho os valores positivos. O sentido e magnitude dos loadings representam correlações positivas ou negativas das variáveis com Δ GWS. Para cada variável, quando os loadings e os valores transformados têm o mesmo sinal, trata-se de correlações positivas, e quando têm sinais diferentes, existem correlações negativas. Nas zonas 1, 2 e 3, assim como no aquífero Boa Vista, existem correlações fortemente positivas entre Δ GWS e as variáveis de umidade do solo (de 0 a 200cm), RootMoist, Qsb e Lwnet, embora moderadamente negativas com a PotEvap, Qh e as ST0_200. Nessas áreas, as mudanças de Δ GWS são fortemente influenciadas pela umidade do solo, uma variável

suscetível a fatores como topografia, cobertura vegetal, precipitação e evaporação. Os padrões espaciais das zonas N, NE e O da ARB no PC1 obedecem a padrões topográficos, como a elevação e inclinação do terreno. A zona S, embora guarde relação topográfica, está fortemente alterada pela intensa pressão antrópica. Ao contrário do que ocorre ao norte, onde os contrastes nos valores transformados são coerentes com os limites dos aquíferos, a leste, na zona 2, apesar de existirem heterogeneidades topográficas e geológicas, as correlações são fortemente positivas. A magnitude da SD e LT do ΔGWS nas zonas 2 e 3 é baixa a moderada (130mm e 7mm/ano, respectivamente), no entanto, está fortemente correlacionada com a umidade do solo, associada, portanto, à heterogeneidade regional das propriedades hidráulicas dos solos. A zona C do AAS, sobre as planícies de inundação, estende-se com valores ligeiramente negativos a nulos, onde se assume que existe uma maior dependência entre ΔGWS e os níveis de água superficial dos rios. As zonas 4, correspondentes a Ucamará e Llanos de Moxos, depressões topográficas onde confluem vários rios, os dados são moderadamente negativos, limitados ao norte e sul por características topográficas com valores moderados a altos positivos. Sobre a bacia alta, no domínio dos Andes, os dados transformados são os mais altos, guardando correlações positivas entre ΔGWS e PotEvap, Qh, AvgSurfT e ST(0-200cm), e negativas com a umidade do solo. Nos aquíferos Parecis e Ronuro, predominam os valores positivos, os quais são interpretados como: 1) correlações negativas com a umidade do solo, associadas a aumentos nos ΔGWS derivados de quedas na SM, ou um desfase sazonal moderado (2 a 1 meses ou -1 a -2 meses) causando aumentos nos ΔGWS um tempo depois dos aumentos na SM, e 2) moderada sensibilidade dos ΔGWS à evaporação, temperatura do solo e ao fluxo de calor. Esse fenômeno estende-se em grande parte da sub-bacia do rio Xingu.

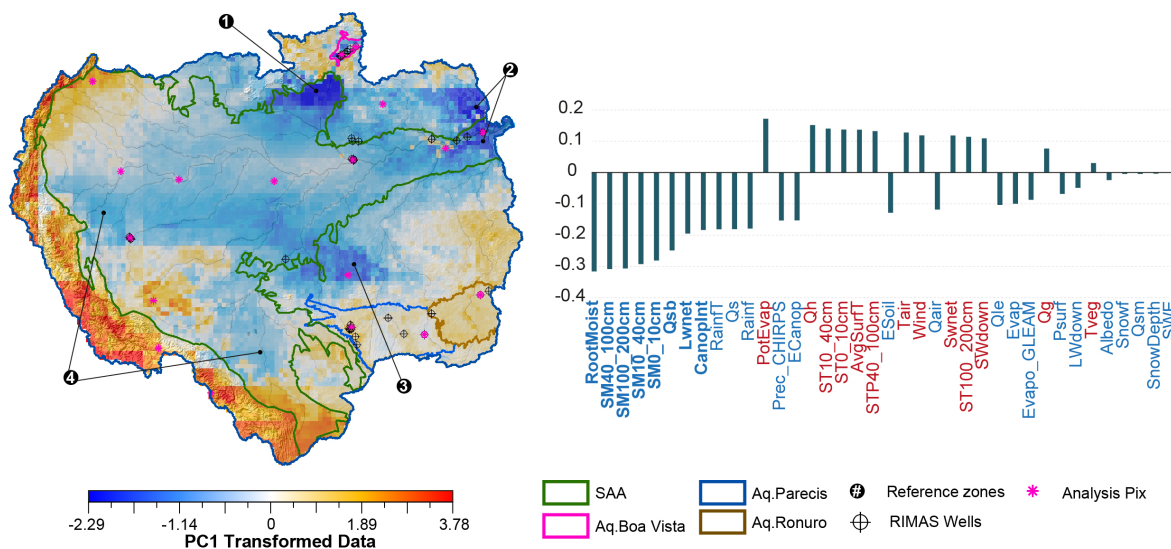


Figura 36. Mapa dos dados transformados para PC1 e gráfico de barras com as magnitudes dos loadings. As cores das fontes correspondem ao azul nas magnitudes negativas e ao vermelho nas magnitudes.

PC2

As variáveis com maiores loadings positivos para o PC2 são a temperatura do solo (ST10-200cm) e a precipitação (Prec_CHIRPS, Rainf e RainfT), com sentidos negativos estão a transpiração (Tveg), o fluxo neto de calor latente (Qle) e a Evaporação (Evap de GLDAS, que inclui a evapotranspiração). A transformação dos dados mostra espacialmente que os máximos negativos se distribuem no AAS C e E (aquífero Alter do Chão), parte dos domínios cratônicos N e S, a sub-bacia do rio Xingu, aquífero Parecis, aquífero Ronuro, aquífero Boa Vista e algumas áreas dentro da cordilheira dos Andes colombianos, peruanos e bolivianos (zonas 1 e 2)(Figura 37). Nessas áreas, os Δ GWS apresentam fortes correlações negativas com a temperatura do solo nas profundidades entre 40 e 200cm, no entanto, a transpiração na bacia baixa e média, nas porções NE, E, SE e S da ARB, é moderada a fortemente positiva. Nos Andes, a transpiração é mais heterogênea, com correlações negativas nas proximidades do pixel 7393, embora fortemente positivas nas zonas C e N, a partir do pixel 6851. Esse comportamento nos dados de transpiração é atribuído a mudanças na cobertura vegetal de acordo com a altitude e latitude (geobotânica), bem como às transformações na vegetação decorrentes da atividade humana. Nos aquíferos Parecis e Ronuro, onde a LT do Δ GWS é negativa, as correlações inversamente proporcionais com a temperatura do solo evidenciam sua alta suscetibilidade superficial, bem como moderados retardos na recarga pela precipitação. Nas zonas 4, onde se distribuem as maiores magnitudes positivas, existem inclinações moderadas e alta densidade de drenagens, associadas a materiais de textura mais argilosa. No Peru (zona 4 N), existe uma depressão onde passa o rio Ucayali e está localizada a reserva natural Imiria, uma região suscetível a ter mudanças a curto prazo de Δ TWS e Δ GWS com a precipitação.

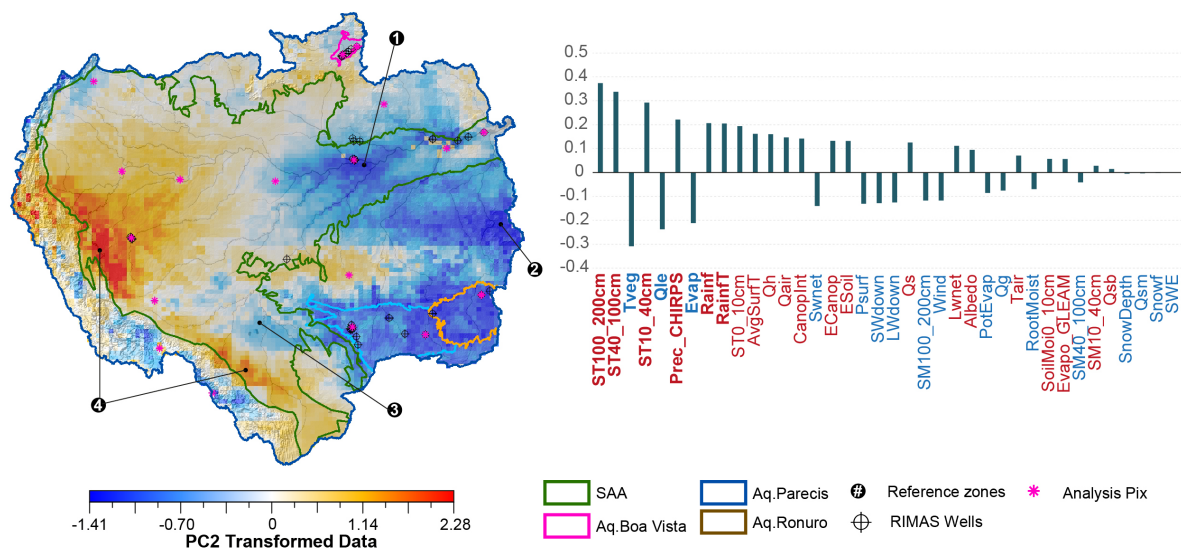


Figura 37. Mapa dos dados transformados para PC2 e gráfico de barras com as magnitudes dos loadings. As cores das fontes correspondem ao azul para magnitudes negativas e ao vermelho para magnitudes positivas.

PC3

Explicado de forma significativa pelo fluxo de calor latente (Qle) e pela evaporação (Evap e Evapo_GLEAM). Os valores máximos positivos são encontrados sobre as zonas 1 (Bolívia) e 2 (Cordilheira Huaytapallana), valores intermediários no cratônico N e aquífero Boa Vista (Figura 38). Correlação nem sempre significa causalidade. Para interpretar seu significado nos domínios da bacia baixa e média, assume-se que as correlações positivas são distribuídas em áreas onde o desfasamento é mínimo a nulo com a evaporação, mas moderadamente negativo com a precipitação (Figura 25 e Figura 26). Isso significa que os incrementos nos Δ GWS derivados da precipitação com desfasagem de -2 meses coincidem com os máximos sazonais de evaporação. Isso não significa, então, que uma maior evaporação gere incrementos nos Δ GWS. Essa explicação propõe que os dados transformados positivos são áreas onde: 1) ocorrem incrementos de Δ GWS com a precipitação em pequenos desfasamentos negativos (-2 meses), 2) a diferença temporal entre os máximos sazonais de precipitação e evaporação é pequena e 3) onde os incrementos de Δ GWS e evaporação estão em fase. Nos locais com dados transformados negativos, ocorre o mesmo fenômeno, mas com os máximos sazonais negativos das variáveis (correlações fortemente negativas).

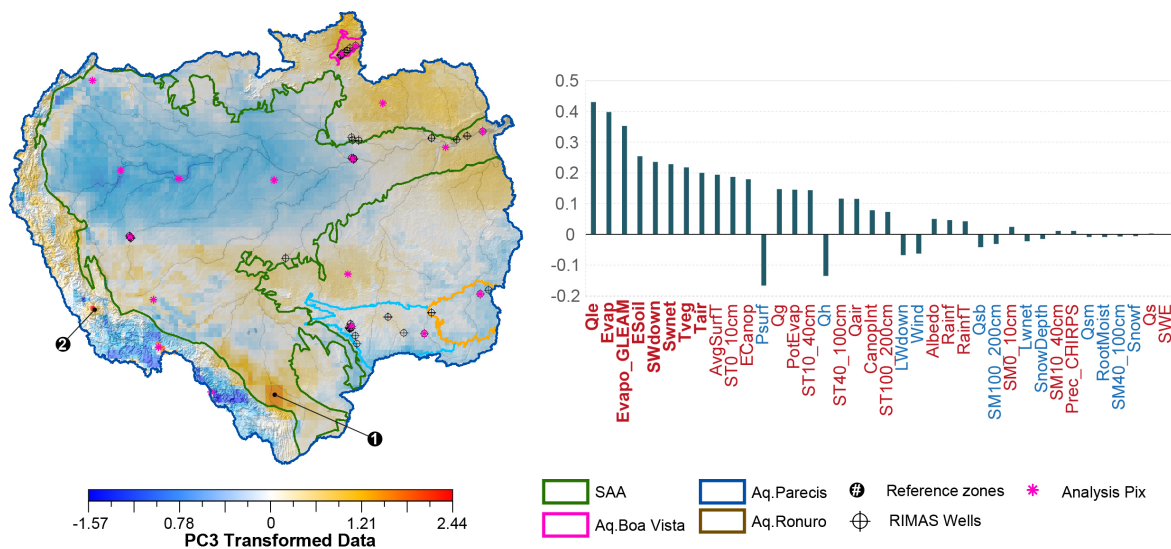


Figura 38. Mapa dos dados transformados para PC3 e gráfico de barras com as magnitudes dos loadings. As cores das fontes correspondem ao azul para magnitudes negativas e ao vermelho para magnitudes positivas.

Composição RGB

A combinação das três principais componentes na composição RGB atribui cores de acordo com a magnitude dos dados transformados. Essa representação espacial facilita a compreensão das PC que tiveram maior ou menor variância explicada em relação ao Δ GWS, bem como a magnitude e sentido das correlações com as variáveis (Figura 39). As relações lineares (corr.) não expressam a magnitude nos Δ GWS, nem são facilmente interpretáveis a longo prazo, para o que deve ser interpretada independentemente dessas métricas. Nas áreas com cores predominantemente vermelhas (R), a variação nos Δ GWS é explicada predominantemente pelo PC1. Ou seja, são lugares onde as variações imediatas no

Δ GWS (sem considerar o desfasamento) são inversamente proporcionais às variáveis SM0-200cm (umidade do solo no intervalo de 0 a 200cm de profundidade), RootMoist (umidade do solo na zona radicular), Qsb (fluxo de base de água subterrânea), Lwnet (fluxo líquido de radiação de onda longa) e CanopInt (água superficial do dossel). Em cor verde (G), encontram-se as maiores magnitudes dos dados transformados do PC2. Nestas áreas, as variações do Δ GWS são diretamente proporcionais a ST10-200cm (temperatura do solo entre 10 e 200cm de profundidade) e Prec_CHIRPS (precipitação), e inversamente proporcionais a Qle (fluxo líquido de calor latente) e Tveg (transpiração). Em azul, o PC3, onde os valores altos representam correlações diretamente proporcionais entre as variações do Δ GWS e Qle, Evapo (evaporação), ESoil (evaporação direta do solo), SWdown (fluxo de radiação de onda curta para baixo), Tveg (transpiração) e Tair (temperatura do ar).

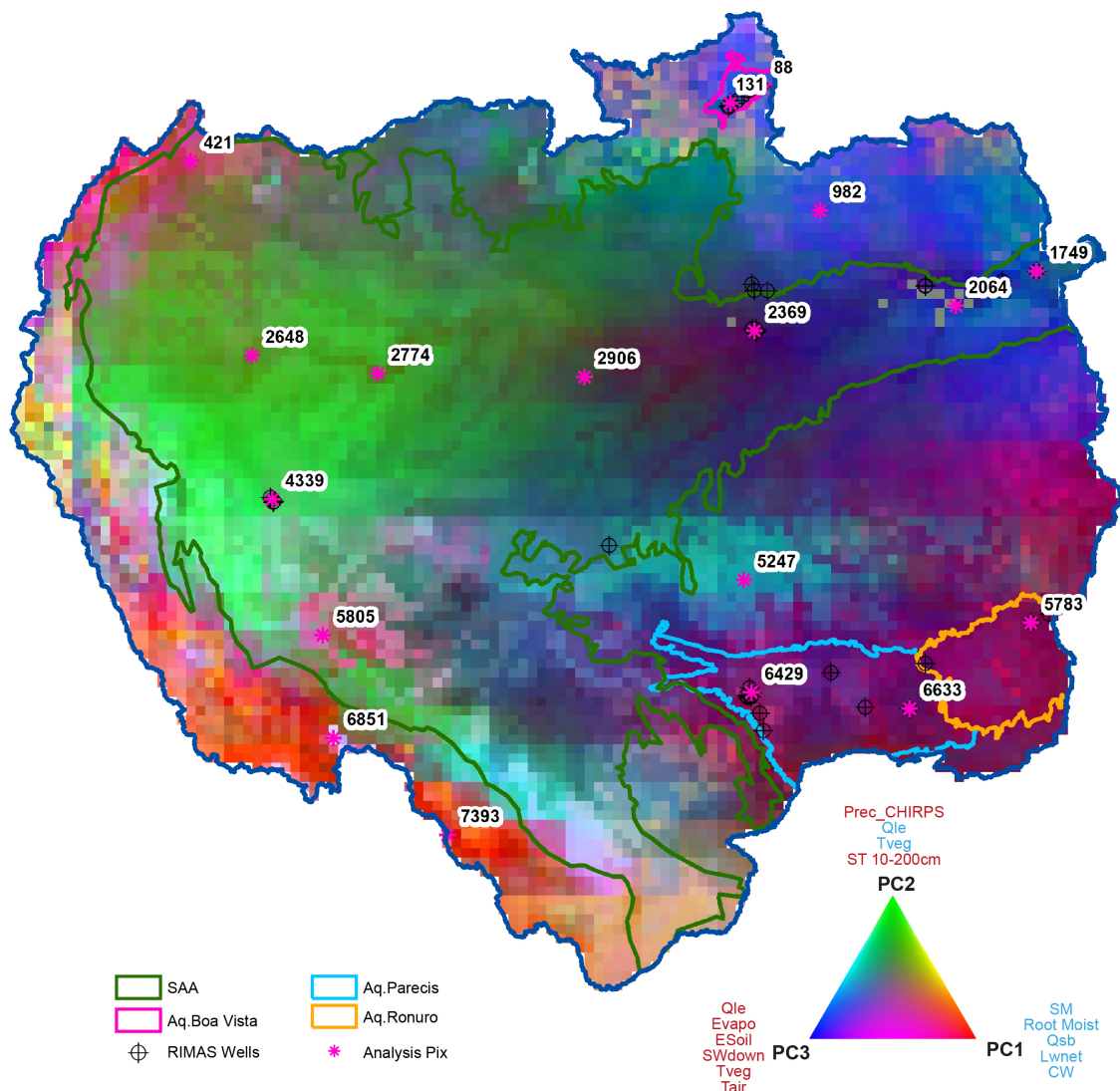


Figura 39. Composição RGB de PC1, 2 e 3 obtida a partir da correlação entre a TS total do GWS e as variáveis hidrometeorológicas. No diagrama ternário, a cor da fonte indica a direção do loading: azul é negativo e vermelho é positivo.

Nos pixels de análise dos aquíferos Parecis e Ronuro, predominam magnitudes fortemente negativas do PC2, moderadas a altas de P1 e moderadas a baixas de P3 (Figura 40). Esses valores se

estendem sobre parte do cráton S e a sub-bacia do rio Xingu. Sobre eles estão localizados os pixels de análise 6429, 6633 e 5783. Nessas regiões, aumentos na temperatura do solo podem provocar a evaporação da água contida no solo, diminuindo os Δ GWS. Apesar de não ser um efeito esperado ser imediato (sem defasagem), as correlações aumentam com a profundidade, tendo maior efeito direto nas águas subterrâneas do que na SM, e reduzindo o retardo causado pela profundidade. Ambos os aquíferos têm alta (>0.75) correlação sazonal máxima em relação à SM, Prec e Evap. No entanto, as defasagens são entre -2 a -3 meses com a SM, -4 a -5 meses com a Prec e de -3 a -4 meses com a Evap. A forte correlação positiva com Tveg (transpiração) pode ser causada pela tendência (TL) de diminuir no Δ GWS e o avanço da fronteira agrícola (maiores incêndios e diminuição na transpiração), padrão que caracteriza esta região da ARB.

O aquífero Boa Vista, algumas regiões do cráton N e S, assim como a porção E do AAS no aquífero Alter do Chão, apresentam cor azul, com predominância do PC1 de magnitudes negativas. Dentro desse domínio estão localizados os pixels 131, 88, 982, 1749 e 2064. As mudanças na umidade do solo influenciam diretamente os Δ GWS, aumentando sua correlação com a profundidade. Também apresentam fortes correlações negativas com a temperatura do solo (ST10-200cm), que aumenta em magnitude com a profundidade. Correlações negativas moderadas a fortemente negativas com AvgSurfT e Tair, variáveis que contribuem para a evaporação da água contida nos solos e, conseqüentemente, nos Δ GWS. Uma forte correlação positiva com Evap (GLDAS e GLEAM), tendo em conta que são áreas com correlações sazonais máximas altas (>0.75) e defasagens nulas a baixas (-1 a 0 meses). Essas correlações indicam que as mudanças nos Δ GWS estão mais relacionadas às mudanças na SM, tornando-as mais suscetíveis aos efeitos da temperatura no ar e no solo. Diferenças nas coberturas dos pixels localizados nas savanas do aquífero Boa Vista deveriam ter influência sobre variáveis como evaporação ou transpiração, no entanto, não são muito contrastantes em relação aos outros pixels.

A porção CE e E do AAS apresenta cores verdes, com predominância do PC. Nessas áreas estão localizados os pixels 2906, 2774, 2648 e 4339. Correlações positivas fortes com a umidade do solo e negativas com Tair, SWdown, Swnet e AvgSurfT. A baixa influência das precipitações, as moderadas pendentes e os altos valores de SD nos Δ GWS são associados a zonas de trânsito com baixa recarga. Em sua superfície, desenvolvem-se extensas formações florestais, mudando para bosques inundáveis sobre a depressão de Ucamará.

Sobre os Andes, o comportamento é heterogêneo, com prevalência das cores laranja e vermelha em direção ao Sul, sobre os pixels 6851 e 7393. Nessas áreas, as correlações são mais intensas com fortes diferenças entre as variáveis. O pixel 6851 não possui um registro de Δ GWS representativo, impedindo uma interpretação correta de seus dados. O pixel 7393 tem correlações fortemente negativas

com a precipitação e umidade do solo, para ambos os casos atribuídos a altos desfases (-5 a -6 meses), pois são áreas com alta correlação sazonal máxima (Figura 26).

Um domínio particularmente delimitado pelo contraste em sua composição RGB é apresentado dentro do cráton S, nas imediações do pixel 5247, com cores azul claro. É uma área com forte dependência positiva entre ΔGWS e a umidade do solo, precipitação, água do dossel, mas também com a evaporação. Possui alta a moderada correlação sazonal máxima (>0.5) com ΔTWS , SM, Prec e Evap (Figura 26). O padrão espacial que define esse subdomínio se repete nos mapas de desfase (Figura 26), mostrando valores nulos a baixos. Assim como em domínios anteriores, o maior desfase ocorre com a evaporação, causando uma inversão no sentido da correlação. É, portanto, uma área onde as mudanças no ΔGWS são fortemente influenciadas pela umidade do solo, precipitação e evaporação, localizada sobre áreas de cráton cobertas por formação florestal e extensas regiões agrícolas. Algumas estruturas geológicas regionais atravessam esse subdomínio, ao que é atribuída uma porosidade secundária que repercute nos ΔGWS .

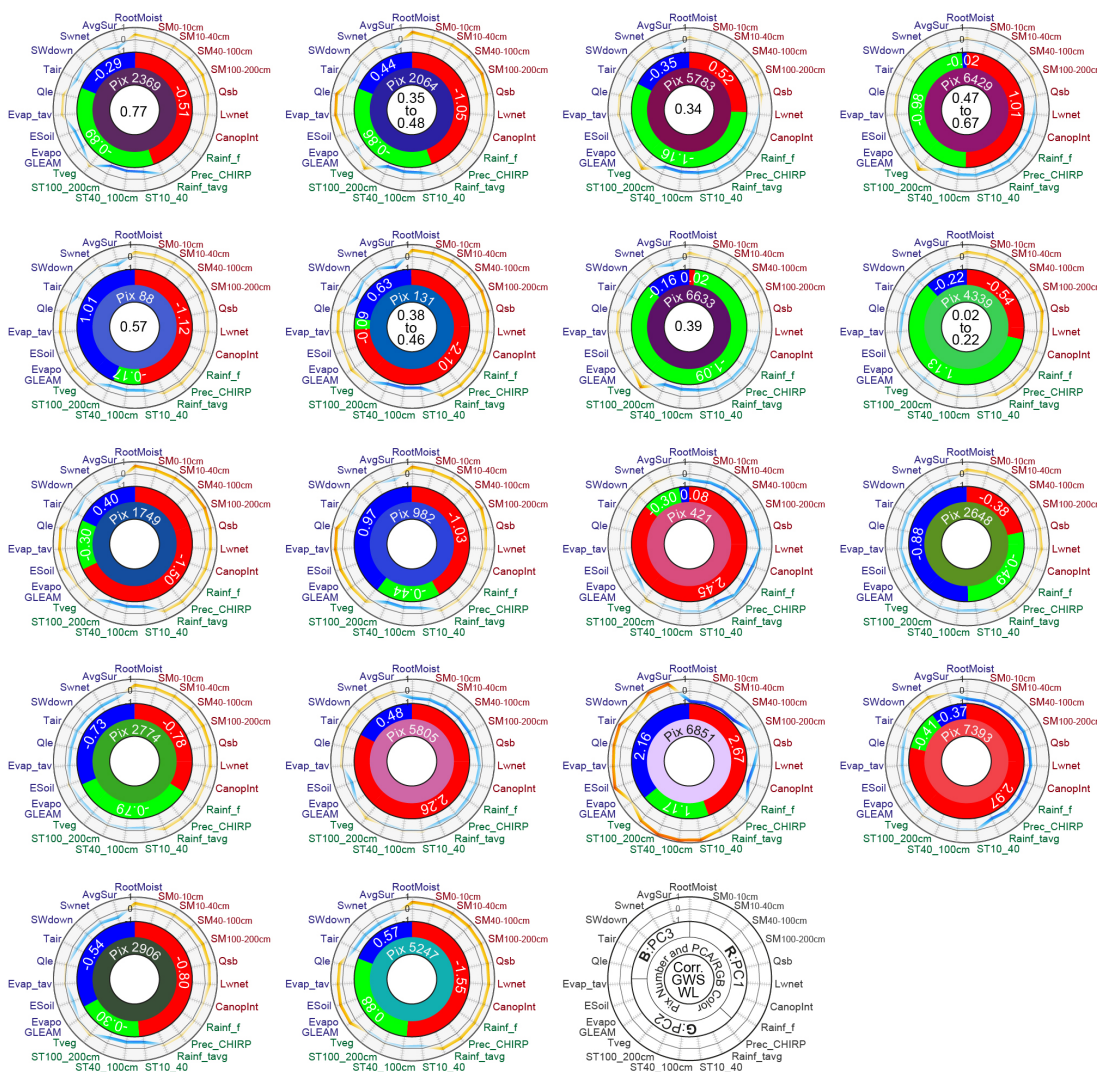


Figura 40. Diagrama resumido da magnitude das CPs, correlações entre GWS, variáveis hidrometeorológicas e WL dos poços de monitoramento do RIMAS, para pixels de análise a 0.25° da ARB.

A sobreposição das PC1, 2 e 3 com as correlações entre ΔGWS e os níveis de água dos poços de monitoramento RIMAS não parece mostrar uma explicação clara em suas diferenças de magnitude (Figura 41). Os poços com menores correlações (0.02 a 0.22), localizados sobre o AAS na zona O, têm uma dominância da PC2. Suas correlações são fracas na grande maioria das variáveis, com valores moderadamente altos e positivos para SM100-200cm e Qsb.

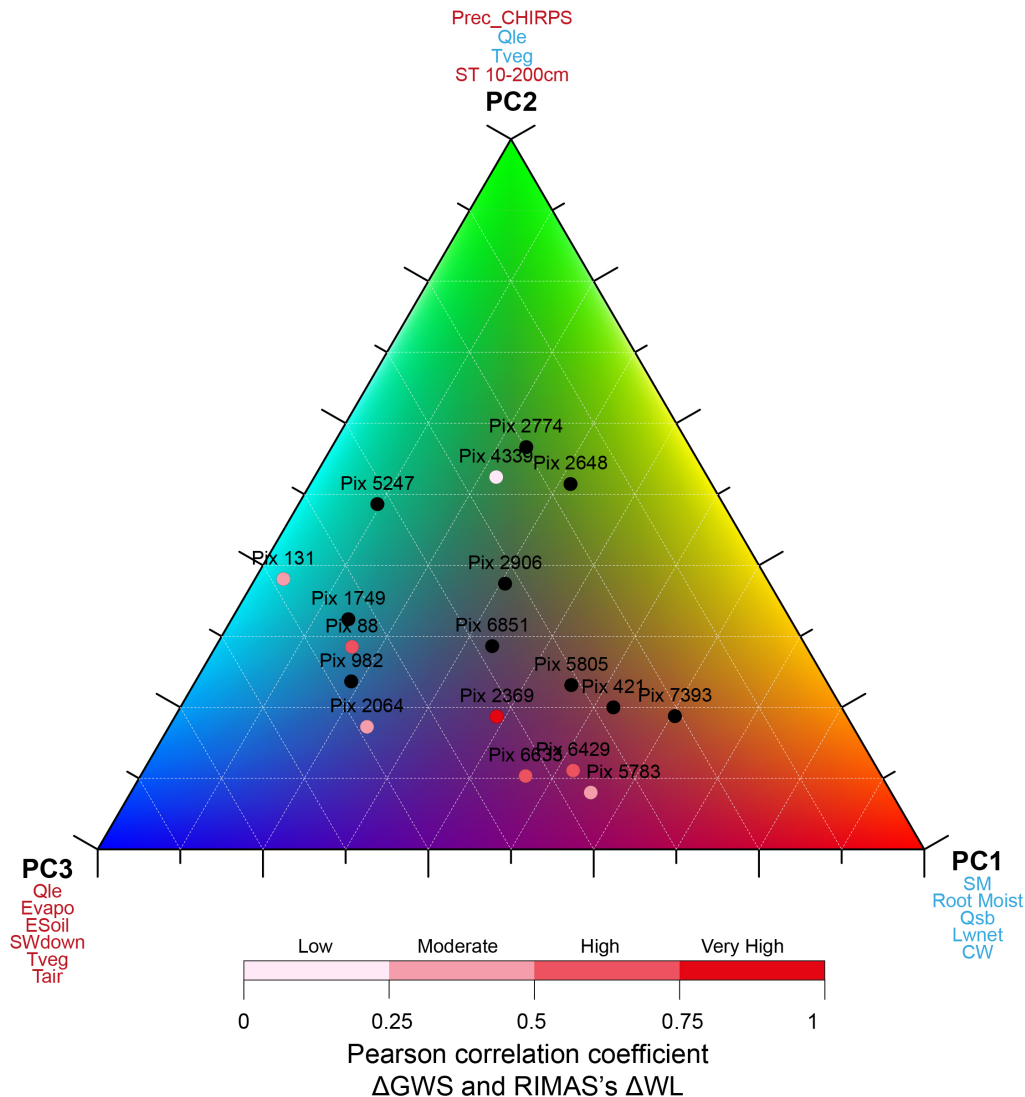


Figura 41. Diagrama ternário de PC1, 2 e 3, sobreposto às correlações entre os STs totais dos WLs dos poços RIMAS e o GWS a 0.25°.

4.2.8. Zoneamento de ΔGWS

Os padrões espaciais e temporais necessários para compreender os ΔGWS dentro da ARB são descritos pelo comportamento a longo prazo através da tendência linear (TL), e pela magnitude das variações na ST, descrita pelo desvio padrão (SD) (Figura 42). Cada variável é reclassificada em três faixas, diferenciando uma TL fortemente positiva, fortemente negativa e nula ou muito baixa. Da mesma forma, são consideradas variações com SD baixo, moderado e alto. A combinação dessas classes resulta em 9 domínios que resumem o comportamento dos ΔGWS e facilitam a identificação de padrões

espaciais. O domínio 9, com SD alto e TL fortemente positiva, concentra-se sobre o AAS nos aquíferos Iça e Alter do Chão, estendendo-se lateralmente sobre os crátons N e S. Acredita-se que, sobre os crátons, desenvolvam-se solos residuais profundos com capacidade de armazenar e transportar água. Na sub-bacia do rio Xingu e nos aquíferos Parecis e Ronuro, são apresentados domínios do tipo 7, 8, 4 e 5 com TL negativas a nulas e SD baixas a altas, demonstrando os efeitos acumulados negativos do desmatamento, queimadas e desenvolvimento de agricultura extensiva, bem como a influência da hidrelétrica de Belo Monte na sub-bacia. Em direção à margem NO e O do AAS, os domínios mostram claramente os efeitos da topografia na bacia média, com domínios do tipo 2 e 3. Essas áreas apresentam baixas SD causadas por declives moderados, materiais argilosos e alta densidade de drenagem. Nessas zonas, podem se desenvolver taxas de recarga baixas a nulas, associadas a fluxos verticais a subverticais. Os Andes apresentam principalmente domínios do tipo 1, 2, 3 e 4. Com predominância do domínio 3, e áreas com domínios 1, 4 e 5 ao sul do Peru e ao norte da Bolívia.

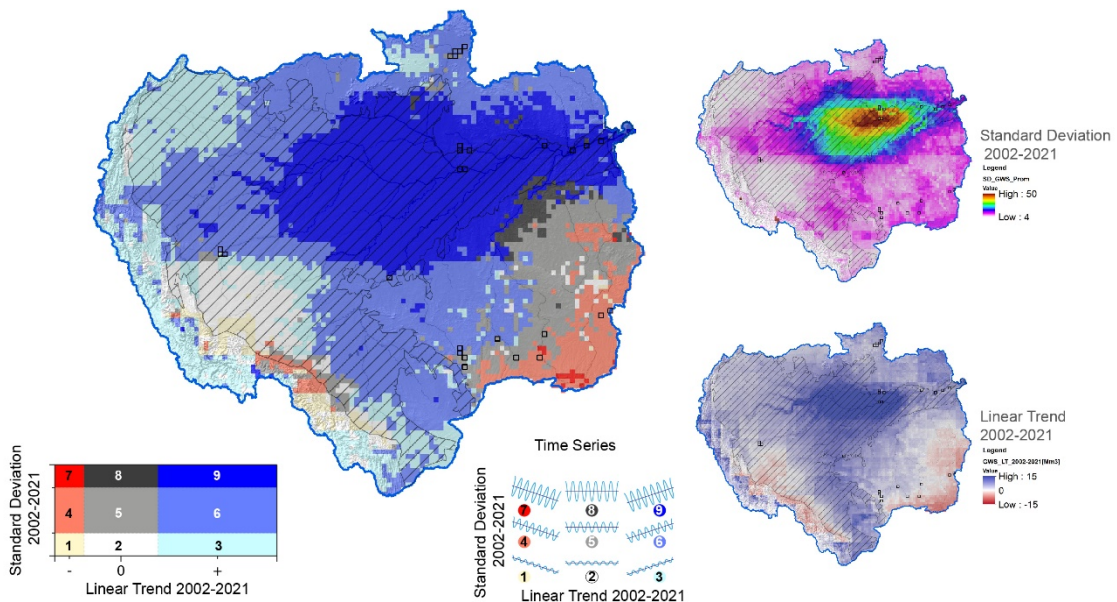


Figura 42. Zonamento do Δ GWS de acordo com SD e TL para o ARB.

5. CONCLUSÕES

As médias das Séries Temporais (ST) de ΔTWS e ΔGWS para a ARB apresentam uma forte influência dos fenômenos El Niño e La Niña. A decomposição temporal das ST totais para a ARB mostra que a magnitude do componente Residual é significativa tanto para ΔTWS (18.2%) quanto para ΔGWS (22.4%). Essas magnitudes podem estar relacionadas a vários fenômenos que merecem ser identificados e estudados para compreender a integralidade dos processos que interferem nas mudanças de armazenamento na ARB. Os padrões espaciais obtidos das ST por meio de métricas como valores médios plurianuais, desvio padrão e tendência linear proporcionam uma visão abrangente do comportamento espacial e temporal das variáveis a médio e longo prazo. Da mesma forma, os componentes temporais de tendência e sazonalidade permitem identificar comportamentos a curto e longo prazo.

Identificar e descrever as relações entre variáveis a partir do coeficiente de correlação de Pearson entre as ST totais é um processo complexo. Os efeitos da sazonalidade podem ser descritos por meio de uma análise de fase, no entanto, estabelecer relações de causalidade é uma tarefa interpretativa. A análise de fase permitiu distinguir áreas onde não há correlação sazonal, podendo ser explicadas pelas componentes de tendência ou residual. Essa análise facilitou delimitar as áreas onde há uma boa correlação sazonal e o número de meses de defasagem com ΔGWS . Essa defasagem pode ter múltiplas implicações interpretativas, dificultando a análise dos resultados. A maior dificuldade consiste em diferenciar se a relação direta ou indiretamente proporcional é causada por uma defasagem na sazonalidade entre variáveis ou seu comportamento a longo prazo, levando em consideração a diversidade de cenários climáticos dentro da ARB. Futuras pesquisas podem estimar as correlações entre variáveis a partir das componentes de tendência, o que proporcionaria uma compreensão a longo prazo, excluindo os efeitos da sazonalidade.

O Índice de Declividade do Terreno (TWI) permitiu identificar áreas onde as características morfológicas favorecem a acumulação de umidade no solo. Os padrões espaciais do TWI estão fortemente relacionados aos do desvio padrão de ΔGWS e ΔTWS . Essas relações podem ser estudadas em escalas de detalhe muito maiores, permitindo identificar os efeitos locais da topografia nas variações de armazenamento de água subterrânea (ΔGWS), proporcionando insumos potenciais para futuros modelos de downsclaiing ou modelos de previsão do nível da água.

O avanço da fronteira agrícola, o desmatamento e as queimadas extensivas nas cabeceiras das bacias hidrográficas e/ou zonas de recarga de aquíferos têm afetado significativamente os ΔGWS . Nos aquíferos Parecis e Ronuro, as tendências lineares negativas mais expressivas para a ARB foram observadas em áreas onde ocorrem coberturas agrícolas e pastagens, decorrentes de queimadas e desmatamento ao longo do tempo. Da mesma forma, o aquífero Boa Vista, apesar de não registrar tendências negativas, é uma área fortemente impactada ao longo do tempo.

A interpretação das variações de armazenamento de água subterrânea (ΔGWS) e sua relação com outras variáveis hidrometeorológicas e morfológicas fornecem critérios que contribuem para o conhecimento regional das águas subterrâneas. Além disso, esses critérios contribuem para a formulação do modelo hidrogeológico regional do SAA. A integração das anomalias de ΔTWS e ΔGWS à modelagem numérica regional das águas subterrâneas e superficiais integradas é uma área de pesquisa promissora.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abdalati, W., Zwally, H. J., Bindschadler, R., Csatho, B., Farrell, S. L., Fricker, H. A. (2010). The ICESat-2 laser altimetry mission. *Proceedings of the IEEE*, 98(5), 735–751. <https://doi.org/10.1109/jproc.2009.2034765>
- Abraham, J. D., & Cannia, J. C. (2011). Airborne electromagnetic surveys for 3d geological mapping. USGS Staff, 506. Retrieved from <https://digitalcommons.unl.edu/usgsstaffpub/506>
- Abreu, F. A. M.; Cavalcante, I. N.; Matta, M. A. S. (2013). O sistema aquífero Grande Amazônia – SAGA: um imenso potencial de água subterrânea no Brasil. In: III Congresso Internacional Meio Ambiente Subterrâneo, 2013, São Paulo - SP. Anais. São Paulo: ABAS.
- Adams, K. H., Reager, J. T., Rosen, P., Wiese, D. N., Farr, T. G., Rao, S., Haines, B. J., Argus, D. F., Liu, Z., Smith, R., Famiglietti, J. S., & Rodell, M. (2022). Remote Sensing of Groundwater: Current Capabilities and Future Directions. In *Water Resources Research* (Vol. 58, Issue 10). John Wiley and Sons Inc. <https://doi.org/10.1029/2022WR032219>
- Agarwal, V., Akyilmaz, O., Shum, C. K., Feng, W., Yang, T. Y., Forootan, E., Syed, T. H., Haritashya, U. K., & Uz, M. (2023). Machine learning based downscaling of GRACE-estimated groundwater in Central Valley, California. *Science of the Total Environment*, 865. <https://doi.org/10.1016/j.scitotenv.2022.161138>
- Agarwal, V., Akyilmaz, O., Shum, C. K., Feng, W., Yang, T. Y., Forootan, E., Syed, T. H., Haritashya, U. K., & Uz, M. (2023). Machine learning-based downscaling of GRACE-estimated groundwater in Central Valley, California. *Science of the Total Environment*, 865. <https://doi.org/10.1016/j.scitotenv.2022.161138>
- Agência Nacional de Águas - ANA. (2015). Avaliação dos Aquíferos das Bacias Sedimentares da Província Hidrogeológica Amazonas no Brasil (escala 1:1.000.000) e Cidades Pilotos (escala 1:50.000) - Volume III – Hidrogeologia e Modelo Numérico de Fluxo da PHA no Brasil.
- Aguiar, C. J. B.; Mourão, M. A. A. (2012). Projeto Rede Integrada de Monitoramento das Águas Subterrâneas: relatório diagnóstico Aquífero Alter do Chão no Estado do Amazonas – Bacia Sedimentar do Amazonas. Belo Horizonte: CPRM. Vol. 6. 30p.
- Ahmed, M., Sultan, M., Wahr, J., & Yan, E. (2014). The use of GRACE data to monitor natural and anthropogenic induced variations in water availability across Africa. *Earth-Science Reviews*, 136, 289–300. <https://doi.org/10.1016/j.earscirev.2014.05.009>
- Ali, S., Liu, D., Fu, Q., Cheema, M. J. M., Pham, Q. B., Rahaman, M. M., Dang, T. D., & Anh, D. T. (2021). Improving the resolution of grace data for spatiotemporal groundwater storage assessment. *Remote Sensing*, 13(17). <https://doi.org/10.3390/rs13173513>
- Alley, W. M., & Konikow, L. F. (2015). Bringing GRACE Down to Earth. *Groundwater*, 53(6), 826–829. <https://doi.org/10.1111/gwat.12379>
- Alves Saraiva, M. (2015) ‘Evolução do nível de água do Aquífero Alter Do Chão no município de Manaus, Amazonas, ao longo de quatro décadas’.
- Amelung, F., Galloway, D. L., Bell, J. W., Zebker, H. A., & Lacznia, R. J. (1999). Sensing the ups and downs of Las Vegas: InSAR reveals structural control of land subsidence and aquifer-system deformation. *Geology*, 27(6), 483–486. [https://doi.org/10.1130/0091-7613\(1999\)027%3C0483:STUADO%3E2.3.CO;2](https://doi.org/10.1130/0091-7613(1999)027%3C0483:STUADO%3E2.3.CO;2)

- Anjos, Nelson. D. F. R. D., Miletto, Michela., Donoso, M. Concepción., Aureli, Alice., van de Gun, Jac., Tujchneider, Ofelia., Rivera, Alfonso., & UNESCO. (2007). Sistemas acuíferos transfronterizos en las Américas, evaluación preliminar. Unesco
- Argus, D. F., Heflin, M. B., Peltzer, G., Webb, F. H., & Crampe, F. (2005). Interseismic strain accumulation and anthropogenic motion in metropolitan Los Angeles. *Journal of Geophysical Research*, 101(B4), B04401. <https://doi.org/10.1029/2003JB002934>
- Atkinson, P. and N. Tate. (2000). Spatial Scale Problems and Geostatistical Solutions: A Review, *Professional Geographer*, vol. 52, no. 4, pp. 607-623
- Baby, P., Calderón, Y., Brusset, S., Roddaz, M., Brichau, S., Eude, A., Calves, G., Quispe, A., Ramirez, L., Bandach, A., Bolaños, R., Hurtado, C., Louterbach, M., Espurt, N. (2018). The Peruvian sub-Andean foreland basin system: structural overview, geochronologic constraints, and unexplored plays. In: AAPG Memoir, vol. 117. American Association of Petroleum Geologists, pp. 91e119. <https://doi.org/10.1306/13622118M1173767>.
- Baby, P., M. Rivadeneira y R. Barragán. (2004). Introducción, en: P. Baby, M. Rivadeneira y R. Barragán, eds., *La Cuenca Oriente: Geología y Petróleo: Travaux de l'Institut Français des Études Andines* (Lima), 144, p. 13-20.
- Bawden, G. W., Thatcher, W., Stein, R. S., Hudnut, K. W., & Peltzer, G. (2001). Tectonic contraction across Los Angeles after removal of groundwater pumping effects. *Nature*, 412(6849), 812–815. <https://doi.org/10.1038/35090558>
- Braga, B., Varella, P., & Gonçalves, H. (2011). Transboundary water management of the Amazon Basin. *International Journal of Water Resources Development*, 27(3), 477–496. <https://doi.org/10.1080/07900627.2011.595382>
- Breiman, L. (2001). *Random Forests*. 45, 5–32.
- Bürgmann, R., Hilley, G., Ferretti, A., & Novali, F. (2006). Resolving vertical tectonics in the San Francisco Bay Area from permanent scatterer InSAR and GPS analysis. *Geology*, 34(3), 221–224. <https://doi.org/10.1130/G22064.1>
- Caputo, M. V. (2014). BACIA DO AMAZONAS: ESTRATIGRAFIA, TECTÔNICA E MAGMATISMO. <https://www.researchgate.net/publication/311193310>
- Caputo, M. V.; Rodrigues, R.; Vasconcelos, D. N. H. (1972). Nomenclatura Estratigráfica da Bacia do Amazonas: Histórico e Atualização. In: XXVI Congresso Brasileiro de Geologia, Belém – PA. Anais. Belém: SBG. Vol. 3, p.35-46.
- Carlotto, V.; Cárdenas, J.; Jaimes, F.; Romero, D.; Valdivia, W. (2008). Nuevos datos en la litoestratigrafía y tectónica Mio-Pleistocena en el Occidente Amazónico: Sur-Este Perú. In: 14° Congreso Peruano de Geología y 13° Congreso Latinoamericano de Geología, 2008, Lima - Peru.
- Castellazzi, P., Longuevergne, L., Martel, R., Rivera, A., Brouard, C., & Chaussard, E. (2018). Quantitative mapping of groundwater depletion at the water management scale using a combined GRACE/InSAR approach. *Remote Sensing of Environment*, 205(December 2017), 408–418. <https://doi.org/10.1016/j.rse.2017.11.025>
- Castellazzi, P., Martel, R., Galloway, D. L., Longuevergne, L., & Rivera, A. (2016). Assessing groundwater depletion and dynamics using GRACE and InSAR: Potential and limitations. *Groundwater*, 54(6), 768–780. <https://doi.org/10.1111/gwat.12453>

- Chaussard, E., Amelung, F., Abidin, H., & Hong, S. H. (2013). Sinking cities in Indonesia: ALOS PALSAR detects rapid subsidence due to groundwater and gas extraction. *Remote Sensing of Environment*, 128, 150–161. <https://doi.org/10.1016/j.rse.2012.10.015>
- Chen, L., He, Q., Liu, K., Li, J., & Jing, C. (2019). Downscaling of GRACE-derived groundwater storage based on the random forest model. *Remote Sensing*, 11(24). <https://doi.org/10.3390/rs11242979>
- Chen, X., & Hu, Q. (2004). Groundwater influences on soil moisture and surface evaporation. *Journal of Hydrology*, 297(1–4), 285–300. <https://doi.org/10.1016/j.jhydrol.2004.04.019>
- Chen, Z., Zheng, W., Yin, W., Li, X., Zhang, G., & Zhang, J. (2021). Improving the spatial resolution of grace-derived terrestrial water storage changes in small areas using the machine learning spatial downscaling method. *Remote Sensing*, 13(23). <https://doi.org/10.3390/rs13234760>
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, 6(No 1.), 3–76.
- Coe, M. T., Brando, P. M., Deegan, L. A., Macedo, M. N., Neill, C., & Silvério, D. v. (2017). The Forests of the Amazon and Cerrado Moderate Regional Climate and Are the Key to the Future. *Tropical Conservation Science*, 10. <https://doi.org/10.1177/1940082917720671>
- Coe, M. T., Macedo, M. N., Brando, P. M., Lefebvre, P., Panday, P., & Silvério, D. (2016). The Hydrology and Energy Balance of the Amazon Basin (pp. 35–53). https://doi.org/10.1007/978-3-662-49902-3_3
- Costa, M. H., Botta, A., & Cardille, J. A. (2003). Effects of large-scale changes in land cover on the discharge of the Tocantins River, Southeastern Amazonia. *Journal of Hydrology*, 283(1–4), 206–217. [https://doi.org/10.1016/S0022-1694\(03\)00267-1](https://doi.org/10.1016/S0022-1694(03)00267-1)
- Cruz, N.M.C. (1984). Palinologia do Linhito do Solimões, Estado do Amazonas. 2º Symposium Amazônico, Manaus, Anais: 473-480.
- Cui, L., Yin, M., Huang, Z., Yao, C., Wang, X., & Lin, X. (2022). The Drought Events over the Amazon River Basin from 2003 to 2020 Detected by GRACE/GRACE-FO and Swarm Satellites. *Remote Sensing*, 14(12). <https://doi.org/10.3390/rs14122887>
- Cunha, P.R.C. (2000). Análise estratigráfica dos sedimentos eomesodevonianos da estratigrafia de sequências no interior cratônico. Dissertation. Universidade Federal do Rio Grande do Sul.
- de Paiva, R. C. D., Buarque, D. C., Collischonn, W., Bonnet, M. P., Frappart, F., Calmant, S., & Bulhões Mendes, C. A. (2013). Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. *Water Resources Research*, 49(3), 1226–1243. <https://doi.org/10.1002/wrcr.20067>
- Dehghan-Soraki, Y., Sharifikia, M., & Sahebi, M. R. (2015). A comprehensive interferometric process for monitoring land deformation using ASAR and PALSAR satellite interferometric data. *GIScience and Remote Sensing*, 52(1), 58–77. <https://doi.org/10.1080/15481603.2014.989774>
- Dewar, N., & Knight, R. (2020). Estimation of the top of the saturated zone from airborne electromagnetic data. *Geophysics*, 85(5), EN63–EN76. <https://doi.org/10.1190/geo2019-0539.1>
- Dias, L. C. P., Macedo, M. N., Costa, M. H., Coe, M. T., & Neill, C. (2015). Effects of land cover change on evapotranspiration and streamflow of small catchments in the Upper Xingu River Basin, Central Brazil. *Journal of Hydrology: Regional Studies*, 4(PB), 108–122. <https://doi.org/10.1016/j.ejrh.2015.05.010>

- do Nascimento, N. R., Fritsch, E., Bueno, G. T., Bardy, M., Grimaldi, C., & Melfi, A. J. (2008). Podzolization as a deferralitization process: Dynamics and chemistry of ground and surface waters in an Acrisol - Podzol sequence of the upper Amazon Basin. *European Journal of Soil Science*, 59(5), 911–924. <https://doi.org/10.1111/j.1365-2389.2008.01049.x>
- Eiras, J. F.; Wanderley Filho, J. R. (2006). *Bacias sedimentares brasileiras – Bacia do Solimões*. Fundação Paleontológica Phoenix, ano 8, nº 85. 6p.
- Eiras, J.F., Kinoshita E.M. (1990). Geologia e perspectivas petrolíferas da Bacia do Tacutu. In: Gabaglia, G.P.R, Milani, E.J (Eds), *Origem e Evolução de Bacias Sedimentares*. Petrobras, Rio de Janeiro, pp 197-220.
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. In *Journal of Animal Ecology* (Vol. 77, Issue 4, pp. 802–813). <https://doi.org/10.1111/j.1365-2656.2008.01390.x>
- Encarnacion, I.B.M. (2008). *Modelado de los sistemas petroleros en las cuencas subandinas del Perú (Cuencas: Madre de Dios, Ucayali, Huallaga, Santiago y Marañon)*. Dissertation. Universidad Nacional de Ingeniería, Perú.
- Fassoni-Andrade, A. C., Fleischmann, A. S., Papa, F., Paiva, R. C. D. de, Wongchuig, S., Melack, J. M., Moreira, A. A., Paris, A., Ruhoff, A., Barbosa, C., Maciel, D. A., Novo, E., Durand, F., Frappart, F., Aires, F., Abrahão, G. M., Ferreira-Ferreira, J., Espinoza, J. C., Laipelt, L., ... Pellet, V. (2021). Amazon Hydrology From Space: Scientific Advances and Future Challenges. In *Reviews of Geophysics* (Vol. 59, Issue 4). John Wiley and Sons Inc. <https://doi.org/10.1029/2020RG000728>
- Feng, W., Zhong, M., Lemoine, J. M., Biancale, R., Hsu, H. T., & Xia, J. (2013). Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements. *Water Resources Research*, 49(4), 2110–2118. <https://doi.org/10.1002/wrcr.20192>
- Ferreira, A. N. P.; Lima, C. F.; Cardoso, F. B. F.; Oliveira Junior, H. J.; Kettelhut, J. T. S. (2006). *Panorama da gestão de aquíferos transfronteiriços – enfoque nas regiões centro-oeste e norte do Brasil*. VIII Congresso Latino americano de Hidrologia Subterrânea e Expo Água.
- Ferretti, A., Savio, G., Barzaghi, R., Borghi, A., Musazzi, S., Novali, F. (2007). Submillimeter accuracy of InSAR time series: Experimental validation. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5), 1142–1153. <https://doi.org/10.1109/TGRS.2007.894440>
- Foroumandi, E., Nourani, V., Jeanne Huang, J., & Moradkhani, H. (2023). Drought monitoring by downscaling GRACE-derived terrestrial water storage anomalies: A deep learning approach. *Journal of Hydrology*, 616. <https://doi.org/10.1016/j.jhydrol.2022.128838>
- França, H. P. M.; Manoel Filho, J.; Duarte, R. X.; André, H. O. (2005). *Estudos Hidrogeológicos da Cidade de Manaus*. Relatório Final, vol.3, Shaft Consultoria Ltda. / Águas do Amazonas S.A, Manaus.
- Frappart, F., & Ramillien, G. (2018). Monitoring groundwater storage changes using the Gravity Recovery and Climate Experiment (GRACE) satellite mission: A review. In *Remote Sensing* (Vol. 10, Issue 6). MDPI AG. <https://doi.org/10.3390/rs10060829>
- Frappart, F., Papa, F., Famiglietti, J. S., Prigent, C., Rossow, W. B., & Seyler, F. (2008). Interannual variations of river water storage from a multiple satellite approach: A case study for the Rio Negro River basin. *Journal of Geophysical Research Atmospheres*, 113(21). <https://doi.org/10.1029/2007JD009438>

- Frappart, F., Papa, F., Güntner, A., Tomasella, J., Pfeffer, J., Ramillien, G., Emilio, T., Schiatti, J., Seoane, L., da Silva Carvalho, J., Medeiros Moreira, D., Bonnet, M. P., & Seyler, F. (2019). The spatio-temporal variability of groundwater storage in the Amazon River Basin. *Advances in Water Resources*, 124, 41–52. <https://doi.org/10.1016/j.advwatres.2018.12.005>
- Frappart, F., Papa, F., Santos Da Silva, J., Ramillien, G., Prigent, C., Seyler, F., & Calmant, S. (2012). Surface freshwater storage and dynamics in the Amazon basin during the 2005 exceptional drought. *Environmental Research Letters*, 7(4). <https://doi.org/10.1088/1748-9326/7/4/044010>
- Freeze, R.A. and Cherry, J.A. (1979) *Groundwater*. Prentice Hall Inc., Englewood Cliffs, New Jersey, 36-38.
- Freund, Y., & Schapire, R. E. (1995). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting.
- Funk, C. C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D., Romero, B. E., Husak, G. J., Michaelsen, J. C., & Verdin, A. P. (2014). A Quasi-Global Precipitation Time Series for Drought Monitoring. U.S. Department of the Interior. U.S. Geological Survey, Data Series 832. <https://doi.org/10.3133/ds832>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, 2. <https://doi.org/10.1038/sdata.2015.66>
- Galvão, P. H. F. (2011). Caracterização hidrogeológica dos aquíferos da Formação Solimões na base operacional geólogo Pedro de Moura, Província Petrolífera de Urucu/AM. 2011. 148f. Dissertação (Mestrado em Geociências) – Universidade Federal de Pernambuco – UFPE, Recife.
- Gao, H., Tang, Q. H., Ferguson, C. R., Wood, E. F., & Lettenmaier, D. P. (2010). Estimating the water budget of major US river basins via remote sensing. *International Journal of Remote Sensing*, 31(14), 3955–3978. <https://doi.org/10.1080/01431161.2010.483488>
- Gatti, L. v., Basso, L. S., Miller, J. B., Gloor, M., Gatti Domingues, L., Cassol, H. L. G., Tejada, G., Aragão, L. E. O. C., Nobre, C., Peters, W., Marani, L., Arai, E., Sanches, A. H., Corrêa, S. M., Anderson, L., von Randow, C., Correia, C. S. C., Crispim, S. P., & Neves, R. A. L. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature*, 595(7867), 388–393. <https://doi.org/10.1038/s41586-021-03629-6>
- Giroto, M., De Lannoy, G. J. M., Reichle, R. H., & Rodell, M. (2016). Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model. *Water Resources Research*, 52, 4164–4183. <https://doi.org/10.1002/2015wr018417>
- Gonçalves, F.T., C.A. Mora, F. Córdoba, E.C. Kairuz y B.N. Giraldo. (2002). Petroleum generation and migration in the Putumayo Basin, Colombia: insights from an organic geochemistry and basin modeling study in the foothills. *Marine and Petroleum Geology*, v.19, p. 711-725.
- Grayson, R. B., & Ladson, D. A. R. (1991). DIGITAL TERRAIN MODELLING: A REVIEW OF HYDROLOGICAL, GEOMORPHOLOGICAL, AND BIOLOGICAL APPLICATIONS. In *HYDROLOGICAL PROCESSES* (Vol. 5).
- Haghtalab, N., Moore, N., Heerspink, B. P., & Hyndman, D. W. (2020). Evaluating spatial patterns in precipitation trends across the Amazon basin driven by land cover and global scale forcings. *Theoretical and Applied Climatology*, 140(1–2), 411–427. <https://doi.org/10.1007/s00704-019-03085-3>

- Han, S. C., Sauber, J., Luthcke, S. B., Ji, C., & Pollitz, F. F. (2008). Implications of postseismic gravity change following the great 2004 Sumatra-Andaman earthquake from the regional harmonic analysis of GRACE intersatellite tracking data. *Journal of Geophysical Research*, 113(B11), B11413. <https://doi.org/10.1029/2008JB005705>
- Hartmann, A., Liu, Y., Olarinoye, T., Berthelin, R., & Marx, V. (2020). Integrating field work and large-scale modeling to improve assessment of karst water resources. *Hydrogeology Journal*, 29(1), 315–329. <https://doi.org/10.1007/s10040-020-02258-z>
- Heerspink, B. P., Kendall, A. D., Coe, M. T., & Hyndman, D. W. (2020). Trends in streamflow, evapotranspiration, and groundwater storage across the Amazon Basin linked to changing precipitation and land cover. *Journal of Hydrology: Regional Studies*, 32. <https://doi.org/10.1016/j.ejrh.2020.100755>
- Higley, D.K. (2001). The Putumayo-Oriente-Marañón Province of Colombia, Ecuador, and Peru Mesozoic-Cenozoic and Paleozoic Petroleum Systems, U.S.G.S., Digital Data Series 63, 40 p.
- Hoorn, C. (1993). Marine incursions and the influence of Andean tectonics on Miocene depositional history of northwestern Amazonia: results of a palynostratigraphic study. *Palaeogeography, Palaeoclimatology, Palaeoecology*, vol.105, p.267- 309.
- Houborg, R., Rodell, M., Li, B., Reichle, R., & Zaitchik, B. F. (2012). Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, 48(7), W07525. <https://doi.org/10.1029/2011WR011291>
- Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., & Wolff, D. B. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1), 38–55. <https://doi.org/10.1175/JHM560.1>
- Hurtado, C., Baby, P., Calderon, Y., & Gobbo, E. (2018). Análisis estructural y estratigráfico de la Subcuenca Pachitea, zona noroccidental de la cuenca Ucayali-Perú. *Boletín de La Sociedad Geológica Del Perú*, N8, 039–054. <https://www.researchgate.net/publication/323573771>
- Hurtado, C., Baby, P., Calderon, Y., Brusset, S., & Viveen, W. (2022). The Candamo antiformal stack, Madre de Dios Basin, Peru. In *Andean Structural Styles: A Seismic Atlas* (pp. 313–319). Elsevier. <https://doi.org/10.1016/B978-0-323-85175-6.00024-9>
- Hwang, C., Yang, Y., Kao, R., Han, J., Shum, C. K., Galloway, D. L. (2016). Time-varying land subsidence detected by radar altimetry: California, Taiwan and north China. *Scientific Reports*, 6(1), 1–12. <https://doi.org/10.1038/srep28160>
- INPE - Instituto Nacional de Pesquisas Espaciais. (2021) Programa de Monitoramento da Floresta Amazônica Brasileira por Satélite (PRODES). Acessível em: <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>, atualizado em 19/11/2021.
- ISARM and UNESCO. (2009). ATLAS OF TRANSBOUNDARY AQUIFERS Global maps, regional cooperation and local inventories. <http://www.sg-guarani.org>
- Jackson, T. J. (2002). Remote sensing of soil moisture: Implications for groundwater recharge. *Hydrogeology Journal*, 10(1), 40–51. <https://doi.org/10.1007/s10040-001-0168-2>
- Jimenez, J. C., Marengo, J. A., Alves, L. M., Sulca, J. C., Takahashi, K., Ferrett, S., & Collins, M. (2021). The role of ENSO flavours and TNA on recent droughts over Amazon forests and the

- Northeast Brazil region. *International Journal of Climatology*, 41(7), 3761–3780. <https://doi.org/10.1002/joc.6453>
- Jyolsna, P. J., Kambhammettu, B. V. N. P., & Gorugantula, S. (2021). Application of random forest and multi-linear regression methods in downscaling GRACE derived groundwater storage changes. *Hydrological Sciences Journal*, 66(5), 874–887. <https://doi.org/10.1080/02626667.2021.1896719>
- Karunakalage, A., Sarkar, T., Kannaujiya, S., Chauhan, P., Pranjali, P., Taloor, A. K., & Kumar, S. (2021). The appraisal of groundwater storage dwindling effect, by applying high resolution downscaling GRACE data in and around Mehsana district, Gujarat, India. *Groundwater for Sustainable Development*, 13. <https://doi.org/10.1016/j.gsd.2021.100559>
- Kirsch, R. (2006). Hydrogeophysical properties of permeable and low permeable rocks. In R. Kirsch (Ed.), *Groundwater geophysics – a tool for hydrogeology* (pp. 1–22). Springer.
- Knight, R., Smith, R., Asch, T., Abraham, J., Cannia, J., Viezzoli, A., & Fogg, G. (2018). Mapping aquifer systems with airborne electromagnetics in the Central Valley of California. *Groundwater*, 56(6), 893–908. <https://doi.org/10.1111/gwat.12656>
- Kopecký, M., Macek, M., & Wild, J. (2021). Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition. *Science of the Total Environment*, 757. <https://doi.org/10.1016/j.scitotenv.2020.143785>
- Kuo, C. Y., Cheng, Y. J., Lan, W. H., & Kao, H. C. (2015). Monitoring vertical land motions in southwestern Taiwan with retracked TOPEX/Poseidon and Jason-2 satellite altimetry. *Remote Sensing*, 7(4), 3808–3825. <https://doi.org/10.3390/rs70403808>
- Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, 48(4). <https://doi.org/10.1029/2011WR011453>
- Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, 48(4). <https://doi.org/10.1029/2011WR011453>
- Landerer, F.W. and Cooley, S. S. (2021). GRACE D-103133 Gravity Recovery and Climate Experiment Level-3 Data Product User Handbook. 1–58. https://podaac-tools.jpl.nasa.gov/drive/files/allData/gracefo/docs/GRACE-FO_L3_Handbook_JPL-D-103133_20190327.pdf
- Latrubesse, E. M., Arima, E. Y., Dunne, T., Park, E., Baker, V. R., D’Horta, F. M., Wight, C., Wittmann, F., Zuanon, J., Baker, P. A., Ribas, C. C., Norgaard, R. B., Filizola, N., Ansar, A., Flyvbjerg, B., & Stevaux, J. C. (2017). Damming the rivers of the Amazon basin. In *Nature* (Vol. 546, Issue 7658, pp. 363–369). Nature Publishing Group. <https://doi.org/10.1038/nature22333>
- Leite, C. C., Costa, M. H., Soares-Filho, B. S., & de Barros Viana Hissa, L. (2012). Historical land use change and associated carbon emissions in Brazil from 1940 to 1995. *Global Biogeochemical Cycles*, 26(2). <https://doi.org/10.1029/2011GB004133>
- Li, Z. W., Xu, W. B., Feng, G. C., Hu, J., Wang, C. C., Ding, X. L., & Zhu, J. J. (2012). Correcting atmospheric effects on InSAR with MERIS water vapour data and elevation-dependent interpolation model. *Geophysical Journal International*, 189(2), 898–910. <https://doi.org/10.1111/j.1365-246X.2012.05432.x>
- Li, Z., Cao, Y., Wei, J., Duan, M., Wu, L., Hou, J., & Zhu, J. (2019). Time-series InSAR ground deformation monitoring: Atmospheric delay modeling and estimating. *Earth-Science Reviews*, 192, 258–284. <https://doi.org/10.1016/j.earscirev.2019.03.008>

- Llovel, W., Willis, J. K., Landerer, F. W., & Fukumori, I. (2014). Deep-ocean contribution to sea level and energy budget not detectable over the past decade. *Nature Climate Change*, 4(11), 1031–1035. <https://doi.org/10.1038/nclimate2387>
- Long, D., Scanlon, B. R., Longuevergne, L., Sun, A. Y., Fernando, D. N., & Save, H. (2013). GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas. *Geophysical Research Letters*, 40(13), 3395–3401. <https://doi.org/10.1002/grl.50655>
- Louterbach, M., Roddaz, M., Antoine, P.O., Marivaux, L., Adnet, S., Bailleul, J., Dantas, E., Santos, R.V., Chemale, F., Baby, P., Sanchez, C., Calderon, Y. (2018). Provenance record of late Maastrichtian-early Palaeocene Andean Mountain building in the Amazonian retroarc foreland basin (Madre de Dios basin, Peru). *Terra. Nova* 30 (1), 17e23. <https://doi.org/10.1111/ter.12303>.
- Luthcke, S. B., Sabaka, T. J., Loomis, B. D., Arendt, A. A., McCarthy, J. J., & Camp, J. (2013). Antarctica, Greenland and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution. *Journal of Glaciology*, 59(216), 613–631. <https://doi.org/10.3189/2013JoG12J147>
- Maeda, E. E., Ma, X., Wagner, F. H., Kim, H., Oki, T., Eamus, D., & Huete, A. (2017). Evapotranspiration seasonality across the Amazon Basin. *Earth System Dynamics*, 8(2), 439–454. <https://doi.org/10.5194/esd-8-439-2017>
- Maia, M. A. M.; Marmos, J. L. (2010). Geodiversidade do Estado do Amazonas: programa geologia do Brasil - levantamento da geodiversidade. Manaus: CPRM. 275p.
- Maia, R. G. N.; Godoy, H. K.; Yamaguti, H. S.; Moura, P. A.; Costa, F. S. F.; Holanda, M. A.; Costa, J. A. (1977). Projeto Carvão no Alto Solimões: Relatório final. Manaus: DNPM/CPRM. vol.11
- Marengo, J. A., Williams, E. R., Alves, L. M., Soares, W. R., & Rodriguez, D. A. (2016). Extreme Seasonal Climate Variations in the Amazon Basin: Droughts and Floods (pp. 55–76). https://doi.org/10.1007/978-3-662-49902-3_4
- Margat, J. (2007). Great aquifer systems of the world. *Aquifer Systems Management: Darcy's Legacy in a World of Impending Water Shortage: Selected Papers on Hydrogeology* 10, 105.
- Mark, B. G., French, A., Baraer, M., Carey, M., Bury, J., Young, K. R., Polk, M. H., Wigmore, O., Lagos, P., Crumley, R., McKenzie, J. M., & Lautz, L. (2017). Glacier loss and hydro-social risks in the Peruvian Andes. *Global and Planetary Change*, 159, 61–76. <https://doi.org/10.1016/j.gloplacha.2017.10.003>
- Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., & Verhoest, N. E. C. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5), 1903–1925. <https://doi.org/10.5194/gmd-10-1903-2017>
- Matos de Abreu, F. de A., Nazareno Cavalcante, I., & da Sila Matta, M. A. (2013). O sistema aquífero grande Amazônia – SAGA: um imenso potencial de água subterrânea no Brasil.
- Maxwell, R. M., Chow, F. K., & Kollet, S. J. (2007). The groundwater–land–surface–atmosphere connection: Soil moisture effects on the atmospheric boundary layer in fully-coupled simulations. *Advances in Water Resources*, 30(12), 2447–2466. <https://doi.org/10.1016/j.advwatres.2007.05.018>
- Miguez-Macho, G., & Fan, Y. (2012). The role of groundwater in the Amazon water cycle: 1. Influence on seasonal streamflow, flooding and wetlands. *Journal of Geophysical Research Atmospheres*, 117(15). <https://doi.org/10.1029/2012JD017539>

- Milewski, A. M., Thomas, M. B., Seyoum, W. M., & Rasmussen, T. C. (2019). Spatial downscaling of GRACE TWSA data to identify spatiotemporal groundwater level trends in the upper Floridan aquifer, Georgia, USA. *Remote Sensing*, 11(23). <https://doi.org/10.3390/rs11232756>
- Miralles, D. G., Holmes, T. R. H., de Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, 15(2), 453–469. <https://doi.org/10.5194/hess-15-453-2011>
- Miro, M. E., & Famiglietti, J. S. (2018). Downscaling GRACE remote sensing datasets to high-resolution groundwater storage change maps of California's Central Valley. *Remote Sensing*, 10(1). <https://doi.org/10.3390/rs10010143>
- Moura, M. M., dos Santos, A. R., Pezzopane, J. E. M., Alexandre, R. S., da Silva, S. F., Pimentel, S. M., de Andrade, M. S. S., Silva, F. G. R., Branco, E. R. F., Moreira, T. R., da Silva, R. G., & de Carvalho, J. R. (2019). Relation of El Niño and La Niña phenomena to precipitation, evapotranspiration and temperature in the Amazon basin. *Science of the Total Environment*, 651, 1639–1651. <https://doi.org/10.1016/j.scitotenv.2018.09.242>
- NASA - National Aeronautics and Space Administration. (2017). README Document for the Tropical Rainfall Measurement Mission (TRMM). Goddard Earth Sciences Data and Information Services Center (GES DISC), 69.
- Ndehedehe, C. E., & Ferreira, V. G. (2020). Assessing land water storage dynamics over South America. *Journal of Hydrology*, 580. <https://doi.org/10.1016/j.jhydrol.2019.124339>
- Neill, C., Chaves, J. E., Biggs, T., Deegan, L. A., Elsenbeer, H., Figueiredo, R. O., Germer, S., Johnson, M. S., Lehmann, J., Markewitz, D., & Piccolo, M. C. (2011). Runoff sources and land cover change in the Amazon: An end-member mixing analysis from small watersheds. *Biogeochemistry*, 105(1), 7–18. <https://doi.org/10.1007/s10533-011-9597-8>
- Nie, W., Zaitchik, B. F., Rodell, M., Kumar, S. V., Anderson, M. C., & Hain, C. (2018). Groundwater withdrawals under drought: Reconciling GRACE and Land Surface Models in the United States High Plains Aquifer. *Water Resources Research*, 48(3), 317. <https://doi.org/10.1029/2017WR022178>
- Nobre, P., Malagutti, M., Urbano, D. F., De Almeida, R. A. F., & Giarolla, E. (2009). Amazon deforestation and climate change in a coupled model simulation. *Journal of Climate*, 22(21), 5686–5697. <https://doi.org/10.1175/2009JCLI2757.1>
- Nogueira, A. C. R.; Silveira, R.; Guimarães, J. T. F. (2013). Neogene- Quaternary sedimentary and paleovegetation history of the eastern Solimões Basin, central Amazon region. *Journal of South American Earth Sciences*, vol.46, p.89-99.
- O'Connell, Y., Brown, C., Henry, T., Morrison, L., & Daly, E. (2020). Quantitative assessment of groundwater resources using airborne electromagnetic remote sensing. *Journal of Applied Geophysics*, 175, 103990. <https://doi.org/10.1016/j.jappgeo.2020.103990>
- Paine, J. G., & Minty, R. S. (2005). Airborne hydrogeophysics. In Y. Rubin & S. S. Hubbard (Eds.), *Hydrogeophysics* (pp. 333–357). Springer
- Palacky, G. J. (1993). Use of airborne electromagnetic methods for resource mapping. *Advances in Space Research*, 13(11), 5–14. [https://doi.org/10.1016/0273-1177\(93\)90196-i](https://doi.org/10.1016/0273-1177(93)90196-i)

- Panday, P. K., Coe, M. T., Macedo, M. N., Lefebvre, P., & Castanho, A. D. de A. (2015). Deforestation offsets water balance changes due to climate variability in the Xingu River in eastern Amazonia. *Journal of Hydrology*, 523, 822–829. <https://doi.org/10.1016/j.jhydrol.2015.02.018>
- Pascal, C., Ferrant, S., Selles, A., Maréchal, J. C., Paswan, A., & Merlin, O. (2022). Evaluating downscaling methods of GRACE (Gravity Recovery and Climate Experiment) data: A case study over a fractured crystalline aquifer in southern India. *Hydrology and Earth System Sciences*, 26(15), 4169–4186. <https://doi.org/10.5194/hess-26-4169-2022>
- Pedreira, A. J., & Bahia, R. B. C. (2004). *Estratigrafia e Evolução da Bacia dos Parecis*.
- Peltier, W. R., Argus, D. F., & Drummond, R. (2018). Comment on the paper by Purcell et al. 2016 entitled 'An assessment of ICE-6G_C (VM5a) glacial isostatic adjustment model. *Journal of Geophysical Research: Solid Earth*, 122(1), 450–487. <https://doi.org/10.1002/2014jb011176>
- Pereira, L. A. C.; Cajazeiras, C. C. A. (2012). Caracterização dos Principais Aquíferos da Região Sudoeste da Amazônia Ocidental. In: XVII Congresso Brasileiro de Águas Subterrâneas, Bonito - MS. Anais. Bonito: ABAS.
- Pfeffer, J., Seyler, F., Bonnet, M. P., Calmant, S., Frappart, F., Papa, F., Paiva, R. C. D., Satgé, F., & Silva, J. S. da. (2014). Low-water maps of the groundwater table in the central Amazon by satellite altimetry. *Geophysical Research Letters*, 41(6), 1981–1987. <https://doi.org/10.1002/2013GL059134>
- Pimentel, E. T., & Hamza, V. M. (2012). Indications of regional scale groundwater flows in the Amazon Basins: Inferences from results of geothermal studies. *Journal of South American Earth Sciences*, 37, 214–227. <https://doi.org/10.1016/j.jsames.2012.03.007>
- Pokhrel, Y. N., Fan, Y., Miguez-Macho, G., Yeh, P. J. F., & Han, S. C. (2013). The role of groundwater in the Amazon water cycle: 3. Influence on terrestrial water storage computations and comparison with GRACE. *Journal of Geophysical Research Atmospheres*, 118(8), 3233–3244. <https://doi.org/10.1002/jgrd.50335>
- Pulla, S. T., Yasarer, H., & Yarbrough, L. D. (2023). GRACE Downscaler: A Framework to Develop and Evaluate Downscaling Models for GRACE. *Remote Sensing*, 15(9). <https://doi.org/10.3390/rs15092247>
- Radambrasil. (1977). *Geologia. Levantamento de recursos naturais*. Min. Minas Ener., DNPM, 14: 49-66
- Rakhit, A. (2002). Marañon Basin Petroleum Hydrogeology Study. Report for PERUPETRO, p. 57. <http://www.perupetro.com.pe>.
- Reager, J. T., Thomas, A. C., Sproles, E. A., Rodell, M., Beaudoin, H. K., Li, B., & Famiglietti, J. S. (2015). Assimilation of GRACE terrestrial water storage observations into a land surface model for the assessment of regional flood potential. *Remote Sensing*, 7(11), 14663–14679. <https://doi.org/10.3390/rs71114663>
- Reis, V., Hermoso, V., Hamilton, S. K., Bunn, S. E., Fluet-Chouinard, E., Venables, B., & Linke, S. (2019). Characterizing seasonal dynamics of Amazonian wetlands for conservation and decision making. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 29(7), 1073–1082. <https://doi.org/10.1002/aqc.3051>
- Reynolds, C. A., Jackson, T. J., & Rawls, W. J. (2000). Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon

- databases and continuous pedotransfer functions. *Water Resources Research*, 36(12), 3653–3662. <https://doi.org/10.1029/2000WR900130>
- Rodell, B. M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., & Toll, D. (2004). The Global Land Data Assimilation System. *American Meteorological Society*, 381–394. <https://doi.org/10.1>
- Rodell, M., Chen, J., Kato, H., Famiglietti, J. S., Nigro, J., & Wilson, C. R. (2007). Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal*, 15(1), 159–166. <https://doi.org/10.1007/s10040-006-0103-7>
- Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaulieu, H. K., Landerer, F. W., & Lo, M. H. (2018). Emerging trends in global freshwater availability. *Nature*, 557(7707), 651–659. <https://doi.org/10.1038/s41586-018-0123-1>
- Rodell, Matthew, Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, 460(7258), 999–1002. <https://doi.org/10.1038/nature08238>
- Rosário, F. F. (2011) O sistema aquífero cretáceo multicamada Tikuna: subunidade do sistema aquífero amazonas. Tese de Doutorado. UFRJ.
- Rosário, F. F. do, Custodio, E., & Silva, G. C. da. (2016). Hydrogeology of the Western Amazon Aquifer System (WAAS). *Journal of South American Earth Sciences*, 72, 375–386. <https://doi.org/10.1016/j.jsames.2016.10.004>
- Rowlands, D. D., Luthcke, S. B., Klosko, S. M., Lemoine, F. G., Chinn, D. S., McCarthy, J. J., et al. (2005). Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements. *Geophysical Research Letters*, 32(4). <https://doi.org/10.1029/2004GL021908>
- Rui, H. L., & Beaulieu, H. (2020). README Document for NASA GLDAS Version 2 Data Products. Goddard Earth Sciences Data and Information Services Center (GES DISC), 16(1), 1–32. https://hydro1.gesdisc.eosdis.nasa.gov/data/GLDAS/README_GLDAS2.pdf
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y. T., Chuang, H. Y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., & Becker, E. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, 27(6), 2185–2208. <https://doi.org/10.1175/JCLI-D-12-00823.1>
- Sahour, H., Sultan, M., Vazifedan, M., Abdelmohsen, K., Karki, S., Yellich, J. A., Gebremichael, E., Alshehri, F., & Elbayoumi, T. M. (2020). Statistical applications to downscale GRACE-derived terrestrial water storage data and to fill temporal gaps. *Remote Sensing*, 12(3). <https://doi.org/10.3390/rs12030533>
- Santiago, A.F., Santos, J.O.S., Maia, R.G.N. (1980). Estratigrafia Preliminar da Bacia Sedimentar do Alto Tapajós. In: Congresso Brasileiro de Geologia, 31, Camboriú. Anais, SBG, vol. 2. pp. 786–797.
- Sasgen, I., Konrad, H., Ivins, E. R., Van Den Broeke, M. R., Bamber, J. L., Martinec, Z., & Klemann, V. (2013). Antarctic ice-mass balance 2003 to 2012: Regional reanalysis of GRACE satellite gravimetry measurements with improved estimate of glacial-isostatic adjustment based on GPS uplift rates. *Cryosphere*, 7(5), 1499–1512. <https://doi.org/10.5194/tc-7-1499-2013>
- Save, H., Bettadpur, S., & Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, 121(10), 7547–7569. <https://doi.org/10.1002/2016JB013007>

- Save, H., Bettadpur, S., & Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, 121(10), 7547–7569. <https://doi.org/10.1002/2016JB013007>
- Scanlon, B. R., Longuevergne, L., & Long, D. (2012). Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA. *Water Resources Research*, 48(4), 1–9. <https://doi.org/10.1029/2011WR011312>
- Scanlon, B. R., Zhang, Z., Save, H., Wiese, D. N., Landerer, F. W., Long, D., Longuevergne, L., & Chen, J. (2016). Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resources Research*, 52(12), 9412–9429. <https://doi.org/10.1002/2016WR019494>
- Scanlon, Bridget R., Zhang, Z., Save, H., Sun, A. Y., Schmied, H. M., Van Beek, L. P. H., Wiese, D. N., Wada, Y., Long, D., Reedy, R. C., Longuevergne, L., Döll, P., & Bierkens, M. F. P. (2018). Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proceedings of the National Academy of Sciences of the United States of America*, 115(6), E1080–E1089. <https://doi.org/10.1073/pnas.1704665115>
- Schaller, H.; Vasconcelos, D. N.; Castro, J. C. (1971). Estratigrafia preliminar da Bacia Sedimentar da Foz do Amazonas. In: 25º Congresso Brasileiro de Geologia, 1971, São Paulo – SP. Anais. São Paulo: SBG, vol. 3, p.189-202.
- Scheel, M. L. M., Rohrer, M., Huggel, Ch., Santos Villar, D., Silvestre, E., & Huffman, G. J. (2011). Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrology and Earth System Sciences*, 15(8), 2649–2663. <https://doi.org/10.5194/hess-15-2649-2011>
- Serviço Geológico Do Brasil - CPRM. (1998). Mapa Hidrogeológico do Estado de Rondônia: 1:1.000.000 - Programa Recursos Hídricos: texto explicativo. Porto Velho: CPRM. 32 p.
- Serviço Geológico Do Brasil - CPRM. (2002). Projeto estudos hidrogeológicos da Região Metropolitana de Belém e adjacências. Belém: CPRM. 93p
- Seyoum, W. M., Kwon, D., & Milewski, A. M. (2019). Downscaling GRACE TWSA data into high-resolution groundwater level anomaly using machine learning-based models in a glacial aquifer system. *Remote Sensing*, 11(7). <https://doi.org/10.3390/rs11070824>
- Seyoum, W. M., Kwon, D., & Milewski, A. M. (2019). Downscaling GRACE TWSA data into high-resolution groundwater level anomaly using machine learning-based models in a glacial aquifer system. *Remote Sensing*, 11(7). <https://doi.org/10.3390/rs11070824>
- Siemon, B., Christiansen, A. V., & Auken, E. (2009). A review of helicopter-borne electromagnetic methods for groundwater exploration. *Near Surface Geophysics*, 7(5–6), 629–646. <https://doi.org/10.3997/1873-0604.2009043>
- Siqueira, L. P.; Teixeira, L. B. (1993). Bacia dos Parecís: nova fronteira exploratória da Petrobrás. In: CONGRESSO INTERNACIONAL DA SOCIEDADE BRASILEIRA DE GEOFÍSICA, 3., 1993, Rio de Janeiro. Resumos Expandidos Rio de Janeiro: SBGf. P. 168-170.
- Smith, R. G., Knight, R., Chen, J., Reeves, J. A., Zebker, H. A., Farr, T., & Liu, Z. (2017). Estimating the permanent loss of groundwater storage in the southern San Joaquin Valley, California. *Water Resources Research*, 53(3), 2133–2148. <https://doi.org/10.1002/2016wr019861>
- Souza, E. (2009). Projeto ECO-03: Caracterização Hidrogeológica de duas áreas da BOGPM (Pólo Arara e Complexo Vitória-Régia) para o Uso Sustentável da Água Subterrânea e Diagnóstico

- Geoambiental de Duas Áreas de Disposição de Resíduos de Perfuração (Urucu/AM) – 2º Relatório Técnico Trimestral. Amazonas: FADESP/UFPA.
- Swenson, S., & Wahr, J. (2002). Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity. *Journal of Geophysical Research*, 107(B9), ETG-3–ETG3-13. <https://doi.org/10.1029/2001JB000576>
- Swenson, S., Wahr, J., & Milly, P. C. D. (2003). Estimated accuracies of regional water storage variations inferred from the Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research*, 39(8). <https://doi.org/10.1029/2002WR001808>
- Tancredi, A. C. F. N. S. (1996). Recursos Hídricos Subterrâneos de Santarém: fundamentos para uso e proteção. 1996. 153f. Tese (Doutorado em Geociências) - Universidade Federal do Pará (UFPA), Belém.
- Tao, H., Al-Sulttani, A. H., Salih, S. Q., Mohammed, M. K. A., Khan, M. A., Beyaztas, B. H., Ali, M., Elsayed, S., Shahid, S., & Yaseen, Z. M. (2023). Development of high-resolution gridded data for water availability identification through GRACE data downscaling: Development of machine learning models. *Atmospheric Research*, 291. <https://doi.org/10.1016/j.atmosres.2023.106815>
- Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., & Watkins, M. M. (2004). GRACE Measurements of Mass Variability in the Earth System. *Science*, 305(5683), 503–505. <https://doi.org/10.1126/science.1099192>
- Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., Sasgen, I., Famiglietti, J. S., Landerer, F. W., Chambers, D. P., Reager, J. T., Gardner, A. S., Save, H., Ivins, E. R., Swenson, S. C., Boening, C., Dahle, C., Wiese, D. N., Dobslaw, H., ... Velicogna, I. (2019). Contributions of GRACE to understanding climate change. *Nature Climate Change*, 9(5), 358–369. <https://doi.org/10.1038/s41558-019-0456-2>
- Terzaghi, K. (1925). Structure and volume of voids in soils, translated from *Erdbaummechanik auf Bodenphysikalischer Grundlage*. In *From theory to practice in soil mechanics*. John Wiley.
- Tian, S., Tregoning, P., Renzullo, L. J., van Dijk, A. I. J. M., Walker, J. P., Pauwels, V. R. N., & Allgeyer, S. (2017). Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals. *Water Resources Research*, 53, 1820–1840. <https://doi.org/10.1002/2016WR019641>
- Toczeck, A., Schmitt, R. da S., da Silva Braga, M. A., & de Miranda, F. P. (2019). Tectonic evolution of the Paleozoic Alto Tapajós intracratonic basin - A case study of a fossil rift in the Amazon Craton. *Journal of South American Earth Sciences*, 94. <https://doi.org/10.1016/j.jsames.2019.102225>
- Tomasella, J., Hodnett, M. G., Cuartas, L. A., Nobre, A. D., Waterloo, M. J., & Oliveira, S. M. (2008). The water balance of an Amazonian micro-catchment: The effect of interannual variability of rainfall on hydrological behaviour. *Hydrological Processes*, 22(13), 2133–2147. <https://doi.org/10.1002/hyp.6813>
- UNESCO. (2009). ATLAS OF TRANSBOUNDARY AQUIFERS Global maps, regional cooperation and local inventories ISARM Programme UNESCO-IHP. <http://www.sg-guarani.org>
- Verma, K., & Katpatal, Y. B. (2020). Groundwater Monitoring Using GRACE and GLDAS Data after Downscaling Within Basaltic Aquifer System. *Groundwater*, 58(1), 143–151. <https://doi.org/10.1111/gwat.12929>

- Vishwakarma, B. D., Zhang, J., & Sneeuw, N. (2021). Downscaling GRACE total water storage change using partial least squares regression. *Scientific Data*, 8(1), 1–13. <https://doi.org/10.1038/s41597-021-00862-6>
- Wahr, J., Molenaar, M., & Bryan, F. (1998). Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research*, 103(B12), 30205–30229. <https://doi.org/10.1029/98JB02844>
- Wahr, J., Swenson, S., Zlotnicki, V., & Velicogna, I. (2004). Time-variable gravity from GRACE: First results. *Geophysical Research Letters*, 31(11). <https://doi.org/10.1029/2004GL019779>
- Wanderley Filho, J. R.; Costa, B. S. (1991). Contribuição a Evolução Estrutural da Bacia do Amazonas e sua Relação com o Embasamento. In: Simpósio de Geologia da Amazônia, Belém, p.244-259.
- Watkins, M. M., Wiese, D. N., Yuan, D. N., Boening, C., & Landerer, F. W. (2015). Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth*, 120(4), 2648–2671. <https://doi.org/10.1002/2014JB011547>
- Watkins, M. M., Wiese, D. N., Yuan, D. N., Boening, C., & Landerer, F. W. (2015). Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth*, 120(4), 2648–2671. <https://doi.org/10.1002/2014JB011547>
- Wiese, D. N., Landerer, F. W., & Watkins, M. M. (2016). Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resources Research*, 52(9), 7490–7502. <https://doi.org/10.1002/2016WR019344>
- Wiese, D. N., Watkins, M. M., Yuan, D. N., Boening, C., & Landerer, F. W. (2016). The JPL RL05M GRACE mascon solution: Status, updates, and future prospects. *AGUFM*, G13A-1087.
- Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M. (2019). MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset. *Water Resources Research*, 55(6), 5053–5073. <https://doi.org/10.1029/2019WR024873>
- Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, 44(11), 5844–5853. <https://doi.org/10.1002/2017GL072874>
- Yin, W., Zhang, G., Han, S. C., Yeo, I. Y., & Zhang, M. (2022). Improving the resolution of GRACE-based water storage estimates based on machine learning downscaling schemes. *Journal of Hydrology*, 613. <https://doi.org/10.1016/j.jhydrol.2022.128447>
- Yuan, X., Wood, E. F., Luo, L., & Pan, M. (2011). A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophysical Research Letters*, 38(13). <https://doi.org/10.1029/2011GL047792>
- Zaitchik, B. F., Rodell, M., & Reichle, R. H. (2008). Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin. *Journal of Hydrometeorology*, 9(3), 535–548. <https://doi.org/10.1175/2007JHM951.1>
- Zalán, P. V., & Siguehiko, N. (2007). Bacia do Marajó. *Boletim de Geociências - Petrobras*.
- Zalán, P. V.; Matsuda, N. S. (2007). Bacia do Marajó. *Boletim de Geociências da PETROBRAS*, vol. 15, p.311-310.

- Zhang, J., Liu, K., & Wang, M. (2021a). Downscaling groundwater storage data in China to a 1-km resolution using machine learning methods. *Remote Sensing*, 13(3). <https://doi.org/10.3390/rs13030523>
- Zhang, J., Liu, K., & Wang, M. (2021b). Downscaling groundwater storage data in China to a 1-km resolution using machine learning methods. *Remote Sensing*, 13(3). <https://doi.org/10.3390/rs13030523>
- Zhong, D., Wang, S., & Li, J. (2021). Spatiotemporal downscaling of grace total water storage using land surface model outputs. *Remote Sensing*, 13(5), 1–19. <https://doi.org/10.3390/rs13050900>
- Zuo, J., Xu, J., Chen, Y., & Li, W. (2021). Downscaling simulation of groundwater storage in the Tarim River basin in northwest China based on GRACE data. *Physics and Chemistry of the Earth*, 123(June), 103042. <https://doi.org/10.1016/j.pce.2021.103042>

ANEXO A



Characterization of groundwater storage changes in the Amazon River Basin based on downscaling of GRACE/GRACE-FO data with machine learning models

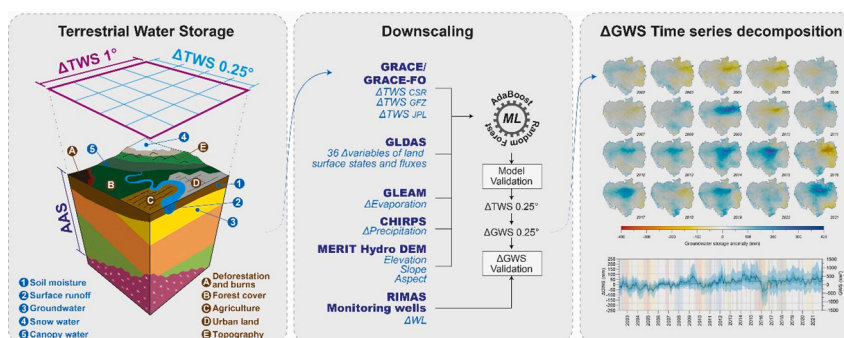
Diego Alejandro Satizábal-Alarcón*, Alexandra Suhogusoff, Luiz Carlos Ferrari

Institute of Geosciences, Groundwater Research Center (CEPAS), University of São Paulo (USP), Rua do Lago 562 - Cidade Universitária, 05508-080 São Paulo, SP, Brazil

HIGHLIGHTS

- The research identified long-term spatial and temporal patterns in Δ GWS of the basin.
- A new dataset of TWS and GWS at 0.25° for ARB is provided.
- The Amazon Aquifer System is strongly susceptible to El Niño and La Niña events.
- The downscaled data obtained improvements in the accuracy of the original data.
- Since the 2016 droughts, GWS prom of the ARB present a decline in the linear trends.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Christian Herrera

Keywords:

Remote sensing
AdaBoost
Random Forest
Artificial intelligence
Time series
Amazon aquifer system
Land cover

ABSTRACT

Groundwater storage changes in the Amazon River Basin (ARB) play an important role in the hydrological behavior of the region, with significant influence on climate variability and rainforest ecosystems. The GRACE and GRACE-FO satellite missions provide gravity anomalies from which it is possible to monitor changes in terrestrial water storage, albeit at low spatial resolution. This study downscaled GRACE and GRACE-FO data from machine learning models from 1° (110 km approx) to 0.25° (27.5 km approx). It estimated the spatio-temporal variability of terrestrial and groundwater storage anomalies between 2002 and 2021 for the Amazon River Basin. In parallel, the Random Forest and AdaBoost algorithms were compared and analyzed. The results reflected a good fit of the models with a very low error and a slight superiority in the predictions obtained by AdaBoost. On the predictions at 0.25° , spatial patterns associated with the strong influence on storage changes of some rivers and snow-capped mountains were identified, as well as an increase in the accuracy of the scaled data of the original ones. Positive long-term behavior was also obtained in terrestrial and groundwater storage of $14.26 \pm 1.18 \text{ km}^3/\text{yr}$ and $+ 22.24 \pm 1.18 \text{ km}^3/\text{yr}$, respectively. Validation of the time series of groundwater anomalies to water levels in the monitoring wells obtained maximum correlation coefficients of 0.85 with confidence levels of 0.01. These results are promising for satellite information in water management, especially

* Corresponding author.

E-mail address: diego_satizabal@hotmail.com (D.A. Satizábal-Alarcón).

in regional monitoring of unconfined aquifers. The obtained data is stored in a dedicated repository (Satizábal-Alarcón et al., 2023).

1. Introduction

The Amazon basin is the largest river basin in the world, characterized by complex hydrological processes that connect high rates of precipitation, extensive floodplains, dense tropical forests, complex topography, and significant variations in freshwater storage and discharge (Fassoni-Andrade et al., 2021). The region is now facing risks under climate and anthropogenic changes, and disturbances in the Amazon water balance could have substantial impacts globally (Jimenez et al., 2021). Different studies and approaches to understanding the basin's hydrological behavior have been conducted, with the groundwater component being one of the least studied, although with increasing interest. Groundwater storage also plays a major role in the Amazon's hydrology and greatly influences on climate variability and rainforest ecosystems (Pokhrel et al., 2013). Strong memory effects of the Amazon groundwater system propagate climate anomalies over the region for several years (Frappart et al., 2019; Miguez-Macho and Fan, 2012; Pfeffer et al., 2014). Researchers (Alves Saraiva and Daniel Wahnfried, 2017; Do Nascimento et al., 2008; Pfeffer et al., 2014; Pimentel and Hamza, 2012; Tomasella et al., 2008) developed other studies on groundwater hydrogeology at different scales, while some of them (Frappart et al., 2019; Miguez-Macho and Fan, 2012; Pokhrel et al., 2013) covered the entire river basin and implemented an integrative model-based approach. Groundwater monitoring networks play a crucial role in validating remote sensing data and comprehending changes in groundwater storage for research. Nonetheless, their installation can be costly and unevenly distributed across regions, posing challenges for studying groundwater storage changes at the regional scale.

Satellite-derived information is invaluable for regions like the Amazon River Basin (ARB), whose extensive size and limited access challenge field data collection. Fassoni-Andrade et al. (2021) extensively review hydrology advances and challenges in the Amazon basin utilizing satellite data.

The Gravity Recovery and Climate Experiment (GRACE) missions, launched in 2002, and their successor, Follow One (GRACE-FO) since 2018, enable global measurements of terrestrial water storage changes. These satellite missions detect small fluctuations in the Earth's gravity field caused by spatiotemporal variations in terrestrial water storage on monthly or sub-monthly time scales (Tapley et al., 2004). Terrestrial Water Storage Changes (Δ TWS) result from these mass variations in terrestrial water, with TWS representing the vertical integration of all water storage components quantified in equivalent centimeters of water column. From a terrestrial water balance perspective, it becomes possible to decompose the complex hydrological variable TWS into Groundwater Storage Changes (Δ GWS) by analyzing changes in other hydrological variables like Soil Moisture (SMS), Snow Water (SWS), Canopy Water (CWS), and surface runoff (Qs). In regional hydrological studies, applications of GRACE/GRACE-FO have yielded promising results, achieving accuracy levels of <1 cm in areas exceeding 400,000 km² (Swenson et al., 2003).

Nevertheless, one of its primary limitations lies in the low spatial resolution of its $1 \times 1^\circ$ data, which hinders interpretation at local scales. Frappart and Ramillien (2018) and Pokhrel et al. (2013) employed GRACE for the spatiotemporal analysis of groundwater storage variability in the Amazon basin. Concerning the relationship between hydrogeology and its impact on the spatial distribution of Δ GWS, Castellazzi et al. (2016) theoretically postulate that unconfined aquifers exhibit more pronounced localized mass changes than confined and semi-confined aquifers. Hence, comprehending spatiotemporal variations in TWS and GWS has facilitated the tracking of water movement at

both planetary and regional scales, allowing the identification of long-term trends. It is important to note that satellite information has uncertainties, including those related to satellite data and assimilation models. The increasing use of artificial intelligence, particularly machine learning, has expanded the possibilities of processing remote sensor information. Applications such as downscaling, i.e., methods used to derive high spatial resolution information from low spatial resolution information, have gained popularity due to the possibility of using various algorithms to create nonlinear models that can pick up nonlinear solid relationships between independent and dependent variables. In remote sensing, downscaling refers to an increase in spatial resolution or decrease in the pixel size of an image, and upscaling to a decrease or coarsening of spatial resolution (Atkinson and Tate, 2000). In this field, researchers such as Ali et al. (2021), Agarwal et al. (2023), Chen et al. (2019), Khorrami et al. (2023), Seyoum et al. (2019), Pulla et al. (2023), Wang et al. (2023), Yin et al. (2022), Zhang et al. (2021) and Zhong et al. (2021), among others, have applied different Machine Learning (ML) models in downscaling GRACE/GRACE-FO derived TWS and GWS anomalies using algorithms such as Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Networks (ANN) or Multiple Linear Regression (MLR), obtaining good correlations with in situ measurements and reaching spatial resolutions with pixels from 1 to 5 km on a side. This increase in the spatial resolution of TWS and GWS anomalies contributes to the identification of locally accurate spatial and temporal patterns, an improvement in terms of measurement and representation of the scaled variables, as well as an increase in the amount of information per unit area (number of pixels per unit area). In the case of ARB, this may represent a means to understand the interactions between TWS/GWS with anthropogenic activity and climate dynamics such as climate change.

Considering the importance of the ARB, the ecosystems it harbors, the strong influence it exerts on the regional and global hydrological cycle, and the lack of high-resolution and accurate TWS and GWS information, this study has the following objectives:

- Apply a 1° to 0.25° downscaling methodology on GRACE/GRACE-FO TWS anomalies for the Amazon River Basin (ARB) between 2002 and 2021 from ML models and different hydrometeorological and topographic variables.
- Compare Random Forest (RF) and Adaptive Boosting (AB) algorithms in TWS downscaling.
- Decompose and validate the GWS anomalies with the monitoring wells from Integrated Groundwater Monitoring Network (RIMAS), supported by the Geological Survey of Brazil (SGB-CPRM).
- Evaluate long-term trends in TWS and GWS changes.

The study's premise for downscaling is that interactions with other hydrometeorological variables and local characteristics (elevation, slope, and aspect) of higher resolution condition low-resolution data. Therefore, 43 variables were selected from five different sources, and a sensitivity analysis was conducted. Two ML algorithms based on Boosting and Bagging techniques were defined and compared. Based on the modeling, the scaled and validated GWS data was used for the temporal decomposition of the ARB time series and each pixel. This analysis identified long-term trends and their spatial distribution. The data collection is stored in a dedicated repository (Satizábal-Alarcón et al., 2023).

2. Study area

The Amazon River Basin (ARB) is the largest in the world, with an

approximate area of 6.3×10^6 km², distributed among six countries: Brazil (67.9 %), Peru (13.0 %), Bolivia (11.2 %), Colombia (5.5 %), Ecuador (1.7 %) and Venezuela (0.7 %). It comprises several rivers, extensive floodplains, and wetlands (Reis et al., 2019). Located in central and eastern South America, between the latitude and longitude ranges 5°N-20°S and 80°W-50°W, respectively. ARB extends from the Andes Mountains, in the Huayhuash Mountain in Peru, with maximum altitudes of 6450 m.a.s.l., to the Amazonas River mouth in the Atlantic Ocean, northeast of Brazil. The Amazon River flows approximately 7100 km to its mouth, with an average flow rate of 2.10×10^5 m³s⁻¹ (Braga et al., 2011). Its main affluents are the Madeira, Negro, Japurá, Tapajós, and Xingú rivers. Its large extent and complexity, characterized by intense hydrological processes, such as high precipitation and evapotranspiration rates, as well as large variations in water storage and river discharge, make the ARB a key player in the global climate system (Gash et al., 2013; Gatti et al., 2021; Nagy et al., 2016). The mean annual precipitation over most of the basin varies between 1000 and 3000 mm, with maximum values of 4000 mm to the northeast and a minimum of 100 mm over the Andes (Haghtalab et al., 2020; Maeda et al., 2017). The Amazon tropical forest also provides a massive water flux to the atmosphere, ranging between 1000 and 1500 mm per year (Maeda et al., 2017). Seasonal changes in water, in turn, contribute to the formation of tropical forests (Leite et al., 2012) and sustain biodiversity (Junk, 1997; Junk et al., 2010). The ARB was subdivided into Upper Basins (>1000 m), Middle Basins (170 m–1000 m), and Lower Basins (<170 m) based

on elevation (Fig. S1).

In ARB subsoil occurs the Amazon Aquifer System (AAS), a trans-boundary aquifer cataloged in 2005 by UNESCO/OAS as one of the largest in the world, which comprises the hydrogeological provinces of the Orinoco and the Amazon (do Rosário et al., 2016). There is considerable uncertainty about the AAS extension, as evidenced by the cartography proposed by different authors (ISARM, U. I., 2009; Margat, 2007; Anjos et al., 2007; do Rosário et al., 2016; IGRAC, 2015). For this study, the AAS corresponds to the Cenozoic to Cretaceous sedimentary units of the geological map of South America (Gómez et al. 2019), with a total area of 3.1×10^6 km², bounded to the west by the sub-Andean fault system, to the North and South, by the ARB watersheds and by the Guiana and Brazilian crystalline rock shields, as shown in Fig. 1. AAS is a regional, heterogeneous, multi-layer and multi-basin flow system, with the same free potentiometric surface connected to the rivers, settled on Paleozoic rocks (diabasic landslides or sedimentary formations) and crystalline basement (ANA, 2015). It develops over the Amazonian craton and comprises Solimões, Acre, Madre de Dios, Beni, Ucayali, Marañon, Oriente, and Putumayo sedimentary basins. These basins are filled with consolidated and unconsolidated sediments that exhibit stratigraphic similarities from the early Cretaceous to the late Tertiary (do Rosário et al., 2016). Despite the differences in the denomination between countries, some of the principal unconfined aquifers that make up the AAS are Alter do Chão, Içá, Tucunará, Alluvial, as well as the Solimões-Repouso aquitard. The Brazilian National Water Agency

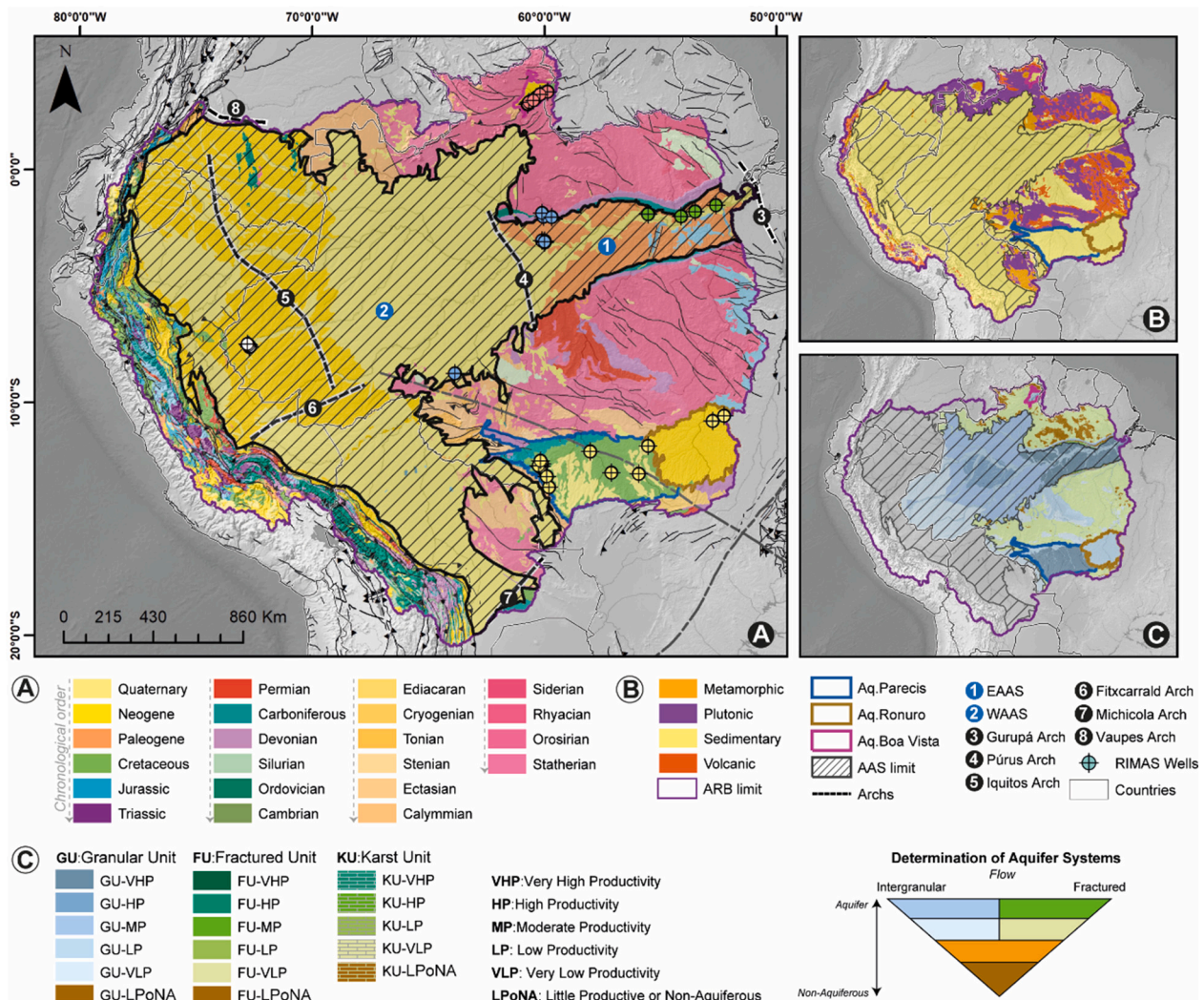


Fig. 1. Study area with a) location of the monitoring wells and b) main chronostratigraphic units extracted from the geological map of South America.

(ANA) assumes the hypothesis of the same potentiometric surface between the Cenozoic, Cretaceous, and Paleozoic units, regardless of the nature of the materials, interrelated by lateral variations or short flow circuits through fractures within the Brazilian portion of the AAS (Agência Nacional de Águas, 2015). The main geological structures bounding the sedimentary basins are the Iquitos, Purus, Gurupá, Fitzcarrald, Vaupes, and Michicola arcs and the sub-Andean fold belt (Fig. 1). The AAS is divided into two regions, the Western Amazon Aquifer System (WAAS) and the Eastern Amazon Aquifer System (EAAS), starting from the Purus Arch (Fig. 1). Other aquifers of regional importance within the ARB are the Parecis and Ronuro aquifers to the southwest and Boa Vista or North.

Within the Brazilian portion of the ARB are 69 monitoring wells of the Integrated Groundwater Monitoring Network (RIMAS) in 26 cities classified within the North, East, South, West, and Central Zones. The validation stage considered 39 of these monitoring wells due to the lack of continuity or strong anomalies in their record. These wells cross various geological materials with depths between 9 and 260 m. The cities with the highest number of wells are Manaus, Vilhena, Cruzeiro do Sul, Bonfim, Boa Vista, and Presidente Figueiredo. The number of pixels over which wells are located at 1° is 19 in CSR, 21 in GFZ, 21 in JPL, and 28 at 0.25°.

3. Materials and methods

3.1. Study design

In the context of this research, an empirical approach was adopted, which relied on implementing AdaBoost and Random Forest algorithms to carry out the downscaling of GRACE/GRACE-FO data in the ARB. The models were trained using a set of hydrometeorological and morphological variables from various sources. The study followed the four steps below:

- 1) Initially, the training data was standardized, transforming variables to a 1° resolution and estimating pixels with missing spatial information. Optimal aggregation functions were evaluated based on their generalization ability during upscaling. A correlation matrix was employed to understand the relationships between the variables.
- 2) The analysis focused on identifying the variables that exerted the most significant influence on the models. This involved scrutinizing their correlations with Total Water Storage (TWS) and employing the Permutation Feature Importance Technique (PFIT). Subsequent to adjusting the hyperparameters of the two algorithms, the models underwent training, predictions were generated, and accuracy was assessed in comparison with the original data. To refine the Total Water Storage (TWS) estimation, a residual correction was implemented. Hypothesis tests were then conducted to evaluate differences between the data before and after the downscaling process, particularly concerning the mean.
- 3) In estimating changes in Groundwater Storage (Δ GWS), progress was achieved by applying the terrestrial water balance method. Temporal series were validated using data obtained from monitoring wells.
- 4) To comprehensively characterize changes in groundwater storage, a decomposition of the temporal components of Groundwater Storage (GWS) was executed. This allowed for both spatial and temporal insights. A comparative analysis was undertaken between Δ GWS and Δ TWS following the application of the downscaling process. This methodological approach comprehensively addressed the challenge of understanding and representing variations in groundwater storage for the ARB.

3.2. Data source and pre-processing

The target spatial resolution is 0.25°, and this choice aligns with the resolution of the independent variables. The analysis period spans from

March 2002 to December 2021, totaling 236 months of data. The ARB comprises 7740 pixels, each with a 0.25° resolution, varying from 492 to 881 pixels in 1° resolution, depending on the solution used (CSR, GFZ, and JPL). For pixels with missing values in their datasets, estimation was performed based on the average of surrounding pixels.

As for the working environment, this research was conducted using Python 3.7.1 for training and validating downscaling models. Additionally, ArcGIS 10.8 provided support for spatial operations and effective data representation. It is worth noting that the various datasets were acquired on April 4, 2022, through the online platforms associated with the respective programs, as summarized in Table 1 of this study.

3.2.1. GRACE and GRACE-FO

The GRACE and its successor GRACE-FO satellite missions are a partnership between the National Aeronautics and Space Administration (NASA) and the German Research Centre for Geosciences (GFZ). They provide space-time measurements of changes in the Earth's gravitational field (Tapley et al., 2004; Wahr et al., 2004). Gravity anomalies are the product of land mass changes related to the movement of water over land and subsequently transformed to equivalent water columns called Terrestrial Water Storage (TWS) (Cooley and Landerer, 2019a, 2019b). The official GRACE Science Data System continuously provides gravity solutions through three different processing centers: the Jet Propulsion Laboratory (JPL), Center for Space Research at the University of Texas, Austin (CSR), and Geoforschungszentrum Potsdam (GFZ). The data used in this study were level-3 RL06 Mass Concentration blocks (mascons) from the centers CSR, GFZ, and JPL, downloaded on 04/04/2022. The RL06 data are state-of-the-art in processing mission observations and have the best signal-to-noise ratio and lower uncertainties (Scanlon et al., 2016; Watkins et al., 2015). For its part, CSR RL06 provided accurate surface-based gridded information that can be used without further processing (Save et al., 2016). On the data, processes such as decorrelation filter, geocenter correction, glacial isostatic adjustment, spatial soothing and leakage correction, coastal resolution improvement (CRI) filters, and ellipsoid landform corrections have been applied (Wiese et al., 2016). On the JPL data, scaling factors were used according to Landerer and Swenson (2012), restoring the original model amplitude. Battery failures in the GRACE mission generated months with gaps starting in 2011.

3.2.2. GLDAS

The Global Land Data Assimilation System (GLDAS) was developed jointly by NASA, Goddard Space Flight Center (GSFC), the National Oceanic and Atmospheric Administration (NOAA), and the National Centers for Environmental Prediction (NCEP). GLDAS provides fields of land surface states and fluxes, integrating satellite- and ground-based observational data products using advanced land surface modeling and data assimilation techniques (Rodell et al., 2004). It includes four Land Surface Models (LSM): Noah, Catchment (CLSM), the Community Land Model (CLM), and the Variable Infiltration Capacity (VIC). There are four versions of the data (1.0, 2.0, 2.1, and 2.2), which have 1.0° or 0.25° spatial resolutions and 3-hourly, daily, and monthly temporal resolutions. This study used 36 hydroclimatic variables of the Noah LSM L4 monthly 0.25° × 0.25° V2.1 dataset. This dataset was selected because of the diversity of hydroclimatic variables, spatial and temporal resolution, and its recurrent use in estimating GWS.

3.2.3. CHIRPS

The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) is a product of the collaboration between the United States Geological Survey (USGS) and the Earth Resources Observation and Science (EROS). CHIRPS provides daily, monthly (every 1, 2, or 3 months), and annual near-global precipitation data and spatial resolution of 0.05° and 0.25°. The dataset builds on previous approaches to interpolation techniques and high-resolution, long periods of record precipitation estimates based on infrared Cold Cloud Duration (CCD)

Table 1

Summary of datasets and variables used in this study. Where ground-based measurements (GBM), remote sensing products (RS), and land surface models (LSM).

Variable	Mission	Product	Resolution	Data source
Terrestrial Water Storage (TWS) (CSR, JPL, and GFZ)	GRACE/GRACE-FO	RS	1° - Monthly	http://www2.csr.utexas.edu/grace https://podaac.jpl.nasa.gov/ http://grace.jpl.nasa.gov https://disc.gsfc.nasa.gov/
Albedo	GLDAS	LSM	0.25° - Monthly	
Average surface skin temperature (AvgSurfT)				
Plant canopy surface water (CanopInt)				
Canopy water evaporation (ECanop)				
Direct evaporation from bare soil (ESoil)				
Evapotranspiration (Evap)				
Downward long-wave radiation flux (LWdown)				
Net long-wave radiation flux (Lwnet)				
Potential evaporation rate (PotEvap)				
Surface pressure (Psurf)				
Specific humidity (Qair)				
Ground heat flux (Qg)				
Sensible heat net flux (Qh)				
Latent heat net flux (Qle)				
Storm surface runoff (Qs)				
Baseflow-groundwater runoff (Qsb)				
Snow melt (Qsm)				
Rain precipitation rate (Rainf)				
Total precipitation rate (RainfT)				
Root zone soil moisture (RootMoist)				
Snow depth (SnowDepth)				
Snow precipitation rate (Snowf)				
Soil moisture (0–10 cm) (SM0_10 cm)				
Soil moisture (10–40 cm) (SM10_40 cm)				
Soil moisture (40–100 cm) (SM40_100 cm)				
Soil moisture (100–200 cm) (SM100_200 cm)				
Soil temperature (0–10 cm) (ST0_10cm)				
Soil temperature (10–40 cm) (ST10_40cm)				
Soil temperature (40–100 cm) (ST40_100cm)				
Soil temperature (100–200 cm) (ST100_200cm)				

Table 1 (continued)

Variable	Mission	Product	Resolution	Data source
Downward short-wave radiation flux (SWdown)				
Snow depth water equivalent (SWS)				
Net short wave radiation flux (Swnet)				
Air temperature (Tair)				
Transpiration (Tveg)				
Wind speed (Wind)				
Precipitation (Prec_CHIRPS)	CHIRPS	RS and GBM	0.05° - Monthly	https://www.chc.ucsb.edu/data/chirps
Evaporation (Evap_GLEAM)	GLEAM	RS and GBM	0.25° - Monthly	https://www.gleam.eu/
Slope	MERIT	LSM	90 m	http://hydro.iis.u-tokyo.ac.jp/
Elevation	Hydro			
Aspect	DEM			
Water Level (WL)	RIMAS	GB	Point - Daily	http://ri.masweb.cprm.gov.br/

observations (Funk et al., 2015). The primary data used in its creation are 1) the monthly precipitation climatology, CHPClim, 2) quasi-global geostationary thermal infrared (IR) satellite observations, 3) the Tropical Rainfall Measuring Mission (TRMM) 3B42 product from NASA (Huffman et al., 2007; Scheel et al., 2011); 4) atmospheric model rainfall fields from the NOAA Climate Forecast System, version 2 (CFSv2) (Saha et al., 2014; Yuan et al., 2011) and 5) in situ precipitation observations (Funk et al., 2014). This study used precipitation data with a monthly and spatial resolution of 0.05°, obtained from the Climate Hazards Center of the University of California. This dataset assimilates different sources of information from both satellite and field stations, with a high spatial and temporal resolution and good correlation coefficients for the ARB, according to Haghtalab et al. (2020).

3.2.4. GLEAM

The Global Land Evaporation Amsterdam Model (GLEAM) refers to the algorithms used to estimate land evaporation and root-zone soil moisture from satellite information. Since its development in 2011 by Miralles et al. (2011), GLEAM has been regularly revised and optimized, incorporating new geophysical variables observed by satellite (Martens et al., 2017). The latest version of the model (3.6) provides ten different variables at a resolution of 0.25° over two different periods, between 1980 and 2021 (GLEAM v3.6a) and between 2003 and 2021 (GLEAM v3.6b). This study used the Actual Evaporation (E) variable from the GLEAM v3.6a model as part of the data set to perform the downscaling model.

3.2.5. MERIT Hydro DEM

MERIT DEM is a Digital Elevation Model (DEM) developed by removing different error components, such as absolute bias, stripe noise, speckle noise, and tree height bias from the existing spaceborne DEMs (SRTM3 v2.1 and AW3D-30 m v1) (Yamazaki et al., 2017). MERIT Hydro is a global hydrography dataset derived from MERIT DEM and water body datasets such as G1WBM, GSWO, and OpenStreetMap, which includes Flow Direction, Hydrologically Adjusted Elevations, Upstream Drainage Area, Number of Upstream Drainage Pixels, River Channel Width and Height Above Nearest Drainage (HAND) (Yamazaki et al., 2019). Ground elevation data from the Hydrologically Adjusted Elevations has a spatial resolution of 90 m (0.0008° approx). Slope and aspect values were estimated with support from geographic information systems (GIS) and then transformed to 0.25° by applying an aggregation technique with the average.

3.2.6. RIMAS and SIAGAS

The Integrated Groundwater Monitoring Network (RIMAS) is a Geological Survey of Brazil (SGB-CPRM) project. Since 2009, piezometers have been installed in the principal aquifers of the country, which measure changes in groundwater levels (WL) over time, as well as hydro-chemical parameters. There are currently 409 wells, 69 located within the ARB. The data were used in the validation stage of the results to be compared with anomalies in groundwater storage changes (Δ GWS). The piezometers have a heterogeneous spatial distribution concentrated in some of the main cities, with filtering sections at different depths, as well as temporal information gaps or anomalous records, which are taken into account later. The Geological Survey of Brazil (SGB-CPRM) created the Groundwater Information System (SIAGAS), which consists of an integrated database of exploitation wells throughout the Brazilian territory and is constantly updated. The wells' distribution was used to determine the spatial density in pixels 0.25° . The location and density of SIAGAS wells per unit area (0.25° pixels) were used to analyze results.

3.3. Downscaling model

3.3.1. Machine learning algorithms

Within Artificial Intelligence, Machine Learning (ML) is the discipline that enables a system to learn from data through models. Several learning methods are part of this process, but two supervised learning algorithms are of interest for this study: AdaBoost and Random Forest. The first corresponds to a boosting algorithm, meaning that its operation is sequential to reduce bias, and the second one is a bagging algorithm, which aims to minimize variance. Downscaling is a regression problem for which a suitable amount of good-quality data must be available to train the models. The 0.25° data were transformed to 1° (upsampling) for CSR, GFZ, and JPL supports, using Mean (M), Truncated Mean (MT), and Standard Deviation (SD) as statistics during aggregation. The new data (upsampling at 1°) make up a training set, where each space-time data of GRACE/GRACE-FO of 1° has a data set (of different variables) that describes it and with which each model is trained. The original datasets were separated into test and training data in a 3:7 ratio. The models were then trained and pre-validated, identifying the best statistic for each support and algorithm. A selection of the variables with the greatest influence on the models was obtained from the analysis between the Pearson correlation coefficients of the independent and dependent variables and the Permutation Feature Importance Technique (PFIT). The hyperparameters were parameters specific to each algorithm, whose adjustment can provide better prediction results. Thus, having established the best statistics for each support and the variables with the largest generalization capacity, the hyperparameters of each algorithm were adjusted by the trial and error method. Finally, the models were again trained using the totality of the data. Cross-validation by the k-folds method was used during the evaluation of the predictions.

3.3.1.1.1. Random Forest. In 2001, Breiman proposed the Random Forest (RF) model as a new integrated ML algorithm consisting of multiple combinations of Classification and Regression Trees (CART) (Breiman, 2001). This algorithm can be used for both classification and regression problems. Within Machine Learning (ML), RF is a supervised learning algorithm, constituting an ensemble method, i.e., it combines the output of multiple decision trees to obtain an average result. The algorithm generates random subsets of features, reducing the correlation between decision trees and improving the model's accuracy. Three of its main hyperparameters are the number of estimators, the maximum features, and the minimum sample split. Bootstrapping, an iterative resampling technique used to randomly extract multiple samples for training, is applied to the training data, and the remaining samples are

grouped as data out of a bag and used to test the model's accuracy. It is a widely used algorithm in GRACE/GRACE-FO downscaling (Agarwal et al., 2023; Chen et al., 2019; Seyoum et al., 2019). Some of its advantages are the low possibility of overfitting, detection of the interactions and importance of the variables, methods incorporation to estimate missing values, and stability when using multiple samples, prevailing the average of the voting of the decision trees. RF disadvantages are high training times and poor performance with small datasets.

3.3.1.2. Adaptive boosting. In 1995, Freund and Schapire proposed the Adaptive Boosting or AdaBoost (AB) model as an ensemble method that creates several simple sequential predictors from a common algorithm (Freund and Schapire, 1995). These predictors are trained with the same data set, receiving the weights derived from the errors made and improving consecutively. In the end, when all predictors have been trained, AB performs a final prediction by estimating a weighted vote or sum to each predictor according to a relative weight (Elith et al., 2008). This algorithm can be used for both classification and regression problems. Some of its main hyperparameters are the number of estimators, the learning rate, and the regression loss function. AB is a simple algorithm that is easy to program and flexible, with high generalization capacity, low error rate, and high accuracy. However, the influence of some weak predictors at outliers can produce overfitting, making it vulnerable to noise.

3.3.2. Downscaling model design

AdaBoost (AB) and Random Forest (RF) were defined as the algorithms for nonlinear statistical downscaling models for the TWS data. The downscaling model design was the product of applying the methodological proposals of Chen et al. (2019) and Ali et al. (2021), extending the training datasets, performing a sensitivity analysis between variables and models, varying the aggregation strategies, and exploring the hyperparameters, aiming to compare both ML models. Initially, the statistical relationships between the hydrometeorological and topographic variables were defined for the Δ TWS of each program (CSR, GFZ, and JPL) within the ARB at a resolution of $1^\circ \times 1^\circ$. Defining the Δ TWS as the dependent variable and the other variables as independent, their relationships were established in $1^\circ \times 1^\circ$ from the ML models and then applied to the $0.25^\circ \times 0.25^\circ$ variables, obtaining a high-resolution TWS. Fig. 2 presents the flowchart of the downscaling process.

- 1) All independent variables (41 in total, where 36 of GLDAS, 1 of GLEAM, 1 of CHIRPS, and 3 of DEM) within the ARB between March 2002 and December 2021 were resampled from their original resolutions to 1° according to the different pixel supports of the CSR, GFZ, and JPL (Fig. S2) programs. The resampling utilized three aggregation strategies, the Standard Deviation (SD), the Mean (M), and the Truncated Mean (TM), obtaining a total of 3 resamples for each Δ TWS solution, with a total of 9 training datasets.
- 2) During this stage, spatial (ID_Spatial) and temporal (ID_Month) indices were included to facilitate the subsequent interpretation of the time series and their spatial relationship. The models were trained and evaluated using AB and RF algorithms with each of the nine datasets with default hyperparameters, using 70 % of the data for training and 30 % for validation. The best-performing dataset from each support (CSR, GFZ, and JPL) was subjected to hyperparameter tuning for both AB and RF, selecting the best models per support. From a multiplicative analysis of the variables, based on the correlation coefficients and the Permutation Feature Importance Technique (PFIT), the most influential variables in the model were defined with which a final model was trained, using the totality of each dataset's data.
- 3) The Δ TWS derived from the AB and RF models with 0.25° resolution were resampled again to 1° with their respective supports. The

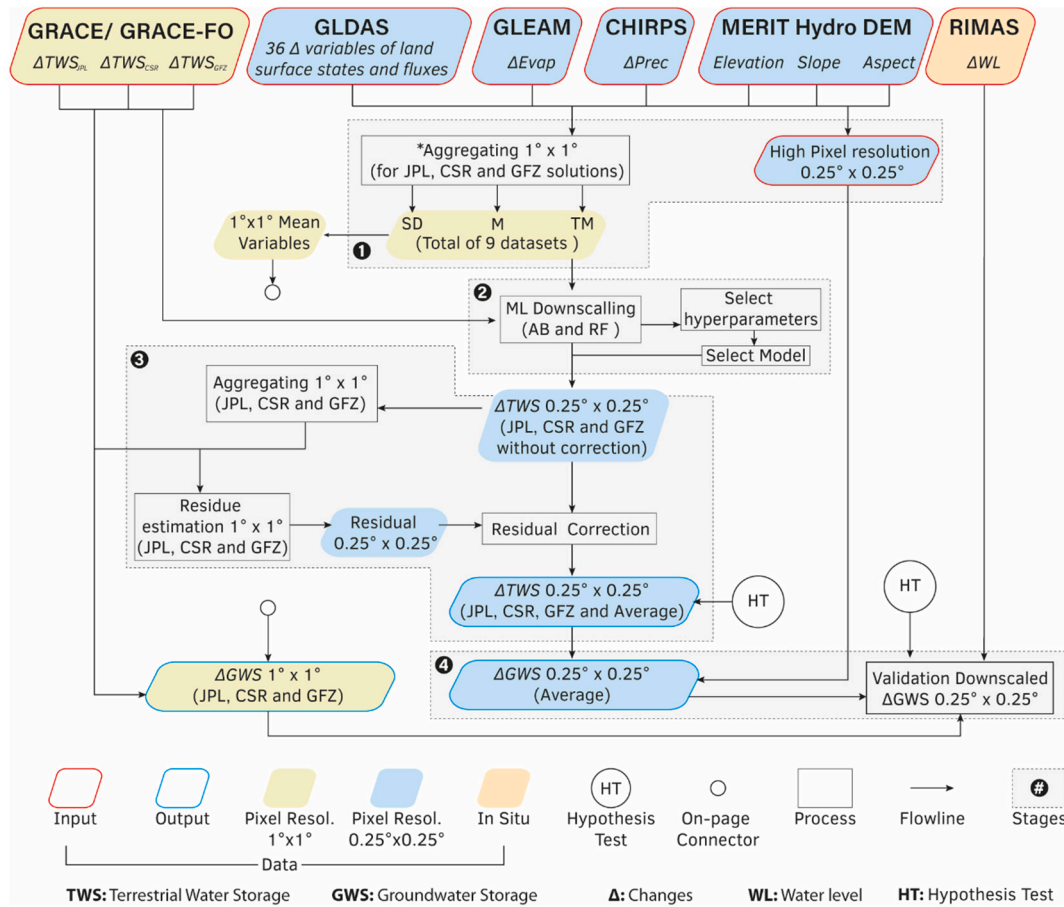


Fig. 2. Flowchart of the downscaling model applied in this study.

residuals between the original and derived ΔTWS at 1° were estimated. Finally, the residuals were added to the 0.25° derived ΔTWS (residual correction). This residual correction ensured that the reduced ΔTWS matched the original data (GRACE/GRACE-FO derived) and corrected for the prediction bias that could result from omitted variables. From the resulting three datasets (CSR, GFZ, and JPL), an AVERAGE ΔTWS was obtained.

- 4) The ΔGWS for the four resulting datasets (CSR, GFZ, JPL, and AVERAGE) was obtained by the terrestrial water balance method [Eq. (5)] using the variables with high pix resolution (0.25°). The resulting ΔGWS time series were compared and validated with the water level anomalies (ΔWL) of RIMAS. The dataset resulted is stored in a dedicated repository (Satizábal-Alarcón et al., 2023).

3.3.3. Model performance assessment

The evaluation of the results of the downscaling models (predicted data) applied three different metrics: Root-Mean-Square Error (RMSE) [Eq. (1)], Mean Absolute Error (MAE) [Eq. (2)] and the Pearson correlation coefficient (R) [Eq. (3)]. RMSE and MAE values close to 0 represent less difference between the GRACE-derived and predicted data, i.e., higher model accuracy. R varies between -1 and $+1$, measuring the statistical relationship or association between two variables; values close to ± 1 mean a perfect linear relationship.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}} \quad (1)$$

$$MAE = \frac{1}{n} \left(\sum_{i=1}^n |X_i - Y_i| \right) \quad (2)$$

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3)$$

Where, X_i : original or input value of ΔTWS . Y_i : value obtained by the model. \bar{X} and \bar{Y} : average value of each independent variable of ΔTWS . n : total number of samples.

Two significance tests were applied using the student's t -test method. The first one was used to compare the time series averages between the scaled pixels and the originals in the different supports, identifying pixels with significant differences in the mean as a product of the downscaling. A second correlation test between the time series of the RIMAS monitoring wells and the ΔGWS , estimating Pearson correlation coefficients (R) and critical values. The significance levels defined were 0.05 (5 %) and 0.01 (1 %).

3.3.4. Characteristic analysis

Reducing the number of training variables decreases computational costs, prevents model overfitting, enhances result interpretability, reduces data noise, and improves generalization capacity. Pearson's correlation coefficients and the Permutation Feature Importance technique (PFIT) purpose is to evaluate the relationship between the independent and dependent variables and their relative contribution to the model and its predictions. The spatiotemporal correlation matrices between variables were obtained after aggregation to a resolution of 1° for the three models resulting from each support (CSR, GFZ, and JPL). PFIT was applied to the models, which assigned scores to each variable according to its relative ability to improve predictions. PFIT scores are relative, not absolute, estimating the difference in the metrics obtained between the

predictions of a model with the data of each variable organized and randomly distributed. The results of these two methods were the variables with the most significant relationship to ΔTWS and their contribution to each model. With these results, training variables are identified and reduced based on their relative contribution to the models, thereby reducing computational costs and improving prediction accuracy.

3.4. Groundwater storage anomaly estimation

The terrestrial water storage (TWS) is the sum of the different terrestrial water storage components constituted by soil moisture, surface runoff, canopy water, and groundwater, as shown in Eq. (4). The groundwater term (GWS) in the water balance method is rearranged in Eq. (5).

$$\Delta TWS = \Delta GWS + \Delta SMS + \Delta SWS + \Delta CWS + \Delta Qs \quad (4)$$

$$\Delta GWS = \Delta TWS - (\Delta SMS + \Delta SWS + \Delta CWS + \Delta Qs) \quad (5)$$

Where, ΔTWS : terrestrial water storage changes or anomalies obtained from GRACE and GRACE-FO. ΔGWS : groundwater storage changes or anomalies. ΔSMS : soil moisture storage changes or anomalies obtained from GLDAS. ΔSWS : snow storage changes or anomalies obtained from GLDAS. ΔCWS : canopy water storage changes or anomalies obtained from GLDAS. ΔQs : surface runoff changes obtained from GLDAS.

ΔGWS was obtained after downscaling the data to 0.25° to validate and compare with the estimates at 1° . The storage components are expressed in volume with units of km^3 and mm of equivalent water height. Within the analysis period (March 2002 to December 2021), there were both spatial and temporal information gaps in the datasets, which were estimated from the average of nearby pixels (spatial) and with the use of ML methods (temporal).

GLDAS (SWS, CWS, SMS, and Qs) data should be transformed to changes or anomalies, estimated from the same baseline with which GRACE data were calculated (January 2004 to December 2009) (Cooley and Landerer, 2019a, 2019b). The estimation of anomalies at time t was calculated using Eq. (6):

$$\Delta \text{Var}(t, i, j) = \text{Var}(t, i, j) - \text{Var}(i, j)_{2004-2009} \quad (6)$$

Where, Var : variables to be transformed to anomalies (Δ). t : time. i and j : pixel coordinates. $\text{Var}(i, j)_{2004-2009}$: average between January 2004 and December 2009.

3.5. Time series decomposition

The time series of ΔTWS and ΔGWS [$h(t)$] were decomposed into their trend [$T(t)$], seasonal [$S(t)$], and residual [$R(t)$] components applying the nonparametric approach STL (Seasonal Trend decomposition using Loess) from Cleveland et al. (1990), as described in Eq. (7).

$$h(t) = T(t) + S(t) + R(t) \quad (7)$$

The trend $T(t)$ or long-term signal can be further divided into linear trend [$T_{lin}(t)$] and the anomalies to this linear trend, named interannual variability [$T_{inter}(t)$].

Long-term trends were interpreted after removing the seasonal effect (deseasonalization) from the ΔTWS and ΔGWS time series. The Seasonal Adjustment $hSA(t)$ was obtained from the additive model described in Eq. (8) by subtracting the seasonal $S(t)$ value from $h(t)$, as shown in Eq. 5.

$$hSA(t) = h(t) - S(t) = T(t) + R(t) \quad (8)$$

4. Results

4.1. Correlation of input variables

All variables were analyzed using a correlation matrix. This procedure allowed establishing which variables had strong linear associations with ΔTWS derived from GRACE/GRACE-FO and identifying regional relationships with statistical significance ($p > 0.05$). Correlation coefficients were estimated for the entire ARB and each support (CSR, GFZ, and JPL) at a resolution of 1° (aggregate) during the months with GRACE/GRACE-FO data (204 of 237).

The results were similar despite the differences in the area and pixel distribution of each ΔTWS support. The absolute averages of each matrix were 0.36 (GFZ), 0.36 (JPL), and 0.35 (CSR). Between ΔTWS and the variables, the highest positive correlations were with Soil moisture (SM100–200 cm), Baseflow-groundwater runoff (Qsb), and Root zone soil moisture (RootMoist) with average values of 0.52, 0.50, and 0.47, respectively, and negative with Potential evaporation rate (PotEvap) of -0.43 . This represented a strong positive relationship between stored water contents in soil and groundwater with the ΔTWS signal for the whole ARB. Some of the variables that had less association with ΔTWS were Aspect, Elevation, Slope, Snow depth (SnowDepth), Snow depth water equivalent (SWS), Snow melt (Qsm), Snow precipitation rate (Snowf), Surface pressure (Psurf) and Transpiration (Tveg). The correlation coefficients for Slope, Elevation, SWE, and SnowDepth were not statistically significant. The correlation between ΔTWS and Soil Moisture (SM) increased with depth, ranging from 0.43 to 0.60 for depths of 0–10 cm and 100–200 cm, respectively. Soil moisture (SM) in the shallower portions (0–40 cm) presented a greater sensitivity to climatic and energetic variables such as Plant canopy surface water (CanopInt), Canopy water evaporation (ECanop), Evapotranspiration (Evap), Evapotranspiration (Evap_GLEAM), Net long-wave radiation flux (Lwnet) and Specific humidity (Qair). Soil moisture (SM) decreased with increasing Potential evaporation rate (PotEvap), Sensible heat net flux (Qh), Downward short-wave radiation flux (SWdown), Net short-wave radiation flux (Swnet), and Wind speed (Wind). As expected, the greater the depth in the soil, the higher the correlation between moisture and Baseflow-groundwater runoff (Qsb). Concerning precipitation from CHIRPS and GLDAS, they presented moderate correlations with ΔTWS (0.28) and strong correlations with soil moisture (0.48–0.67). The latter may be due to the temporal delay between rainfall and aquifer recharge, related to the permeability of the vadose zone. The correlation matrix for the CSR support is presented in Fig. S3 as an example due to its slight average superiority to the other supports.

4.2. Feature of input variables

The results of the models varying the aggregation function are available in Table S1. According to the highest R and lowest MAE and RMSE, the best results for the GFZ and JPL supports were obtained using the truncated mean (MT), unlike CSR, in which the mean (M) was slightly superior. This slight superiority in generalization ability was due to the reduction in the effect of outliers in the predictions. The AB models obtained better metrics for RF by approximately 2% in R, -19% in MAE, and -18% in RMSE. Considering both AB and RF models, GFZ support with MT statistics obtained the best results, whereas JPL support with DS statistics was the worst.

On these three resulting datasets (CSR_M, GFZ_MT, and JPL_MT), the Permutation Feature Importance Technique (PFIT) was applied. Table 2 summarizes the results and correlation coefficients.

According to the PFIT scores in the RF models, more variables were sensitive to change their generalizability relative to AB. In both algorithms AB and RF, the variables Baseflow-groundwater runoff (Qsb) and Soil moisture (SM100–200 cm) were the only ones with high Pearson and PFIT values. Other variables with a high relative contribution to the different models were Elevation, Surface pressure (Psurf), and Soil

Table 2

Pearson's correlation coefficients and PFIT of the hydrometeorological and topographic variables according to each support and algorithm. The first column shows the support used for each variable, numbered 1) CSR_M, 2) GFZ_MT, and 3) JPL_MT.

Support	Pearson CC with ΔTWS			PFIT						
	CSR	GFZ	JPL	CSR	AB GFZ	JPL	CSR	RF GFZ	JPL	
123	ID_Month	0.072	0.075	0.063	0.0023	0.0028	0.0020	0.0236	0.0252	0.0230
123	ID_Spatial	-0.012	-0.021	-0.024	0.1503	0.1563	0.1763	0.2161	0.2535	0.2745
	Albedo	-0.037	-0.032	-0.035	0.0015	0.0050	0.0040	0.0211	0.0146	0.0201
	AvgSurfT	-0.087	-0.121	-0.102	0.0006	0.0003	0.0005	0.0127	0.0062	0.0089
23	CanopInt	0.302	0.318	0.31	0.0004	0.0002	0.0003	0.0034	0.0028	0.0038
	ECanop	0.235	0.232	0.239	0.0004	0.0002	0.0002	0.0043	0.0031	0.0042
	ESoil	0.14	0.139	0.14	0.0011	0.0023	0.0006	0.0073	0.0152	0.0071
	Evap	0.217	0.22	0.207	0.0001	0.0000	0.0000	0.0004	0.0003	0.0004
	LWdown	0.048	0.047	0.054	0.0008	0.0006	0.0005	0.0075	0.0082	0.0072
12	Lwnet	0.343	0.382	0.351	0.0017	0.0022	0.0011	0.0096	0.0109	0.0073
	PotEvap	-0.416	-0.468	-0.425	0.0005	0.0005	0.0005	0.0038	0.0070	0.0041
123	Psurf	0.008	0.005	0.01	0.0432	0.0345	0.0437	0.0642	0.0354	0.0963
	Qair	0.189	0.199	0.194	0.0020	0.0019	0.0016	0.0092	0.0210	0.0193
123	Qg	-0.239	-0.283	-0.245	0.0026	0.0020	0.0026	0.0145	0.0287	0.0157
123	Qh	-0.301	-0.32	-0.291	0.0019	0.0013	0.0022	0.0108	0.0107	0.0145
123	Qle	0.218	0.22	0.208	0.0058	0.0056	0.0040	0.0299	0.0180	0.0213
13	Qs	0.327	0.343	0.316	0.0004	0.0002	0.0003	0.0032	0.0029	0.0039
123	Qsb	0.498	0.529	0.497	0.0945	0.3114	0.1831	0.5052	0.4225	0.5083
	Qsm	0.005	0.007	0.009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Rainf	0.28	0.294	0.287	0.0001	0.0001	0.0001	0.0016	0.0014	0.0013
	RainfT	0.28	0.295	0.288	0.0001	0.0001	0.0001	0.0017	0.0015	0.0014
	RootMoist	0.461	0.493	0.46	0.0012	0.0005	0.0007	0.0096	0.0079	0.0104
	SnowDepth	-1E-04	0.002	0.004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Snowf	0.004	0.007	0.009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	SM0-10cm	0.433	0.458	0.428	0.0002	0.0001	0.0001	0.0018	0.0026	0.0040
	SM10-40cm	0.475	0.502	0.468	0.0002	0.0001	0.0001	0.0014	0.0015	0.0023
123	SM40-100cm	0.548	0.583	0.541	0.0028	0.0015	0.0086	0.0064	0.0066	0.0166
123	SM100-200 cm	0.598	0.631	0.591	0.2468	0.0312	0.1545	0.1936	0.1019	0.1389
	ST0_10cm	-0.08	-0.112	-0.094	0.0006	0.0008	0.0006	0.0098	0.0163	0.0112
12	ST10_40cm	-0.058	-0.082	-0.067	0.0028	0.0520	0.0099	0.0805	0.1143	0.0984
13	ST40_100cm	-0.046	-0.066	-0.053	0.0212	0.0037	0.0257	0.0126	0.0060	0.0112
3	ST100_200cm	-0.03	-0.044	-0.034	0.0061	0.0023	0.0264	0.0228	0.0119	0.0330
	SWdown	-0.263	-0.307	-0.274	0.0003	0.0001	0.0002	0.0038	0.0041	0.0032
123	SWE	-2E-04	0.002	0.004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
	Swnet	-0.264	-0.307	-0.273	0.0003	0.0001	0.0001	0.0034	0.0025	0.0028
	Tair	-0.116	-0.133	-0.114	0.0010	0.0016	0.0004	0.0081	0.0080	0.0064
	Tveg	0.018	0.022	0.003	0.0036	0.0035	0.0034	0.0307	0.0224	0.0295
2	Wind	-0.133	-0.148	-0.141	0.0012	0.0086	0.0007	0.0083	0.0272	0.0080
123	Precip_CHIRPS	0.277	0.296	0.286	0.0006	0.0002	0.0004	0.0056	0.0043	0.0053
123	Evapo_GLEAM	0.135	0.139	0.139	0.0010	0.0009	0.0009	0.0079	0.0088	0.0101
123	Slope	-0.003	0.003	-0.004	0.0083	0.0092	0.0082	0.0559	0.0520	0.0631
123	Elevation	-0.002	0.002	-0.004	0.0685	0.0159	0.0583	0.1753	0.0845	0.1195
123	Aspect	-0.015	-0.021	-0.009	0.0009	0.0165	0.0016	0.0238	0.0537	0.0168

temperature (ST10_40cm, ST40_100cm, and ST100_200cm), despite not having high Pearson values with ΔTWS.

RF and AB models with the same support (CSR, GFZ, or JPL) used the same variables. According to the results, 22 of the 41 original variables were selected. For each model, 19 variables were used, 15 equal between supports. The variables Precip_CHIRPS, Evapo_GLEAM, and SWE_inst were included in all models due to their hydrological relevance. There were seven variables that differentiate the training data between models.

4.3. Models accuracy

The hyperparameters are parameters of each algorithm whose adjustment can provide better prediction results. The AB and RF algorithms were optimized by manipulating some of their hyperparameters, establishing maximum and minimum values, and evaluating their results through the trial and error method. To obtain the best results, for

AB, the hyperparameters were a number of estimators of 500, a learning rate of 1.0, and a loss function of regression square. In the case of RF, the number of estimators was 100, the maximum features were all, and the minimum sample split was eight. Table S2 summarizes the results of the models with the adjustment of the hyperparameters.

The predictions were evaluated by cross-validation by the k-folds method on a total of 6 models. A comparison between the GRACE-derived ΔTWS values and an average of those obtained for each support is presented in the scatter plots in Fig. 3.

GRACE-derived ΔTWS values on the X-axis, and the model predicted data of ΔTWS on the Y-axis. All datasets' histograms and density functions tended to be symmetric or slightly positive. The CSR (Fig. 3 a,d) and JPL (Fig. 3 c,f) supports were leptokurtic, unlike GFZ (Fig. 3 b,e), which was mesokurtic, indicating a better distribution between values and a lower presence of extreme values. As for the models, the best-performing supports were GFZ, followed by CSR and JPL. Of the models obtained with the AB algorithm (Fig. 6 a,b,c), the best results

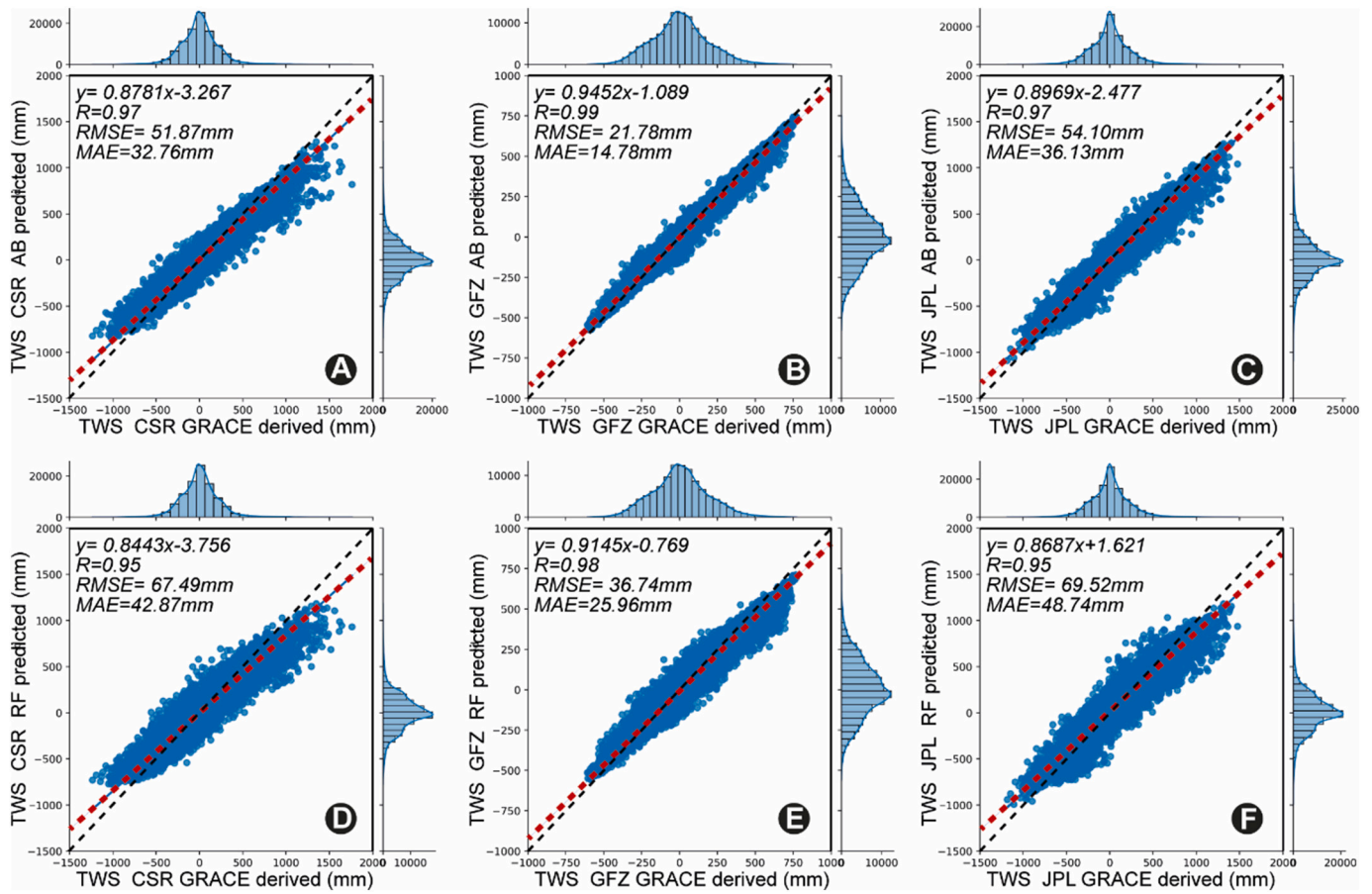


Fig. 3. Scatter plots of between GRACE/GRACE-FO derived Δ TWS values and model predictions for a) CSR with AB, b) GFZ with AB, c) JPL with AB, d) CSR with RF, e) GFZ with RF, and f) JPL with RF.

were GFZ with a correlation of 0.99, an RMSE of 2.178 mm, and an MAE of 1.478 mm, followed by CSR and slightly below JPL. In the case of the RF models (Fig. 3 d,e,f), the best results were GFZ with a correlation of 0.98, an RMSE of 3.674 cm, and an MAE of 2.596 cm. That is to say that the AB models, on average, had a higher R of 0.02, a lower RMSE of 1.53 cm, and a lower MAE of 1.13 cm for the RF models, indicating a slight superiority in the AB algorithm in the ability to predict Δ TWS values for the ARB. A new average dataset was obtained from the three predictions of each support by the AB models, which was analyzed and validated. The information increase resulting from downscaling introduced derived features from the training variables, leading to greater spatial and temporal variance in TWS and, consequently, in GWS. As shown in Fig. S4, after downscaling, spatial variance increases, causing higher

values for the same period. New spatial patterns become apparent, enabling the recognition of novel data characteristics. The average dataset aggregates the variance increments from various supports, enhancing spatial patterns.

Hypothesis testing for the mean between the TWS time series of the GRACE/GRACE-FO derived data and the data predicted by the ML models for each pixel of each support allowed to identify those pixels and areas where this gain in spatial and temporal variance has generated statistically significant changes (Fig. 4). In the three supports (CSR, GFZ, and JPL), the pixels where the null hypothesis was not accepted, in other words, the mean value of their time series before and after downscaling was different, were distributed mainly over the Andes and North of the ARB. Some isolated pixels can also be identified, such as one located

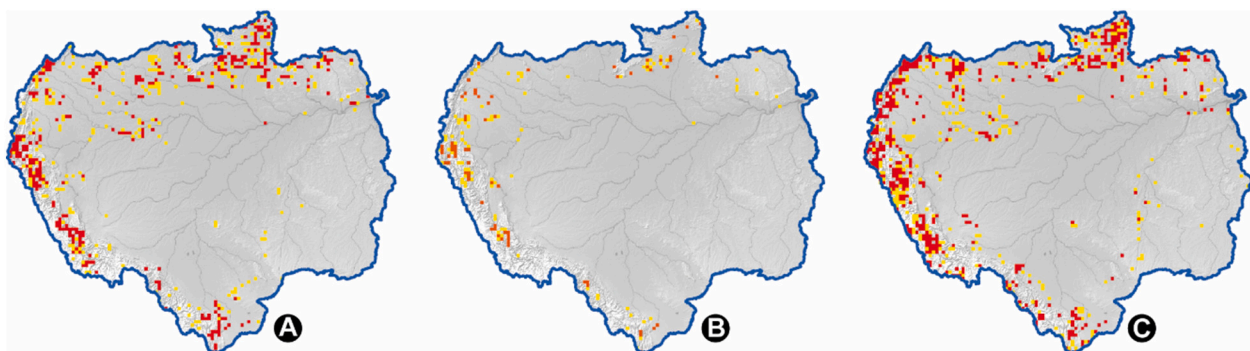


Fig. 4. Pixels whose time series differ from the mean before and after scaling for: a) CSR, b) GFZ, and c) JPL for the AdaBoost models. In yellow, pixels have confidence levels of 5 %, and in red, 1 %.

over Manaus, the only one with statistical significance in the presence of RIMAS monitoring wells. The support with the lowest number of pixels with statistical significance was GFZ, followed by CSR and JPL. These changes in the time series resulted from the increase in spatial variance caused during downscaling, broadening the data representation and allowing the identification of new local effects in the Δ TWS and GWS signals.

4.4. Δ GWS data validation

From highest to lowest, Pearson correlation coefficients between the time series of the average Δ WL measured by the RIMAS monitoring wells and the downscaled Δ GWS anomalies were obtained in the East, Central, South, North, and West zones with maximum values of 0.79, 0.76, 0.73, 0.58 and 0.23, respectively. This showed that the time series resulting from Δ GWS downscaling had a statistically significant correlation with the monitoring wells in the different zones of the ARB. An overlay and spatial comparison between the Δ GWS time series and the Δ WL recorded by some of the RIMAS monitoring wells is presented in Fig. 5.

Most of the exploration wells present in the SIAGAS database in the study area are found in the SAA, Parecis, Ronuro, and Boa Vista aquifers. High concentrations per unit area were appreciable in different cities, Manaus being the most relevant, with a number of exploitation wells in the range of 1001 to 1700 for an area of approximately 770 km² (pixel of 0.25°). This could have had implications in the Δ WLs, such as discrepancies with the regional behavior of the time series or stronger local depletions that could interfere with the monitoring well logs, although a coherent behavior could be seen with the scaled Δ GWS data and the ARB average values. The variability between the scaled Δ GWS time series and the Δ WLs were similar, high in areas such as the Center and East with SD between 422 and 235 mm and 1.6–1.2 m, respectively, low in the West and South with SD values of 71–64 mm and 0.5–0.4 m. In the North, this variability was slightly different, with high variations of the

WL (1.32 m) compared to the GWS anomalies (82 mm), which in turn presented an irregular seasonality. The long-term behavior of the periods in which they coincided was comparatively similar, with positive trends in the North and Central zones, constant in the West and East, and negative in the South. Fig. 6 shows the spatial representation of the correlation coefficients and their statistical significance.

Most RIMAS wells are located at depths shallower than 120 m and within locally unconfined aquifers. The scatter plot in Fig. 6a shows higher correlations with shallow wells; nevertheless, in the eastern zone, a 260 m deep well within a confined aquifer recorded a correlation coefficient of 0.65. The number of samples and their distributions are unclear enough to differentiate the behavior between RIMAS wells installed in unconfined, semi-confined, and confined aquifers. Regarding the standard deviation (SD) of the time series per pixel of Δ GWS, the highest magnitudes are observed over the Alter do Chão and Içá aquifers and alluvial deposits (AAS). Spatial patterns related to geology, topography, and the influence of surface water bodies can be appreciated.

4.5. Spatio-temporal characterization of Δ GWS

The temporal decomposition of the Δ GWS signal into its seasonal, trend, and residual components provides a deeper understanding of water storage variations over short and long time scales. By isolating the seasonal component, variations linked to recurrent natural phenomena occurring throughout the year (short-term) become evident. The trend component imparts significant insights into medium-term changes associated with low-frequency seasonal effects such as El Niño and La Niña. Meanwhile, the linear trend quantifies long-term changes, offering valuable information about growth or decline patterns.

The spatial representation of Δ GWS anomalies (Fig. 7) reveals maximum and minimum extremes in the central region of the ARB, specifically in the Alter do Chão and Içá aquifers (for TWS see Fig. S6). During positive peaks, spatial patterns emerge in the floodplains of the

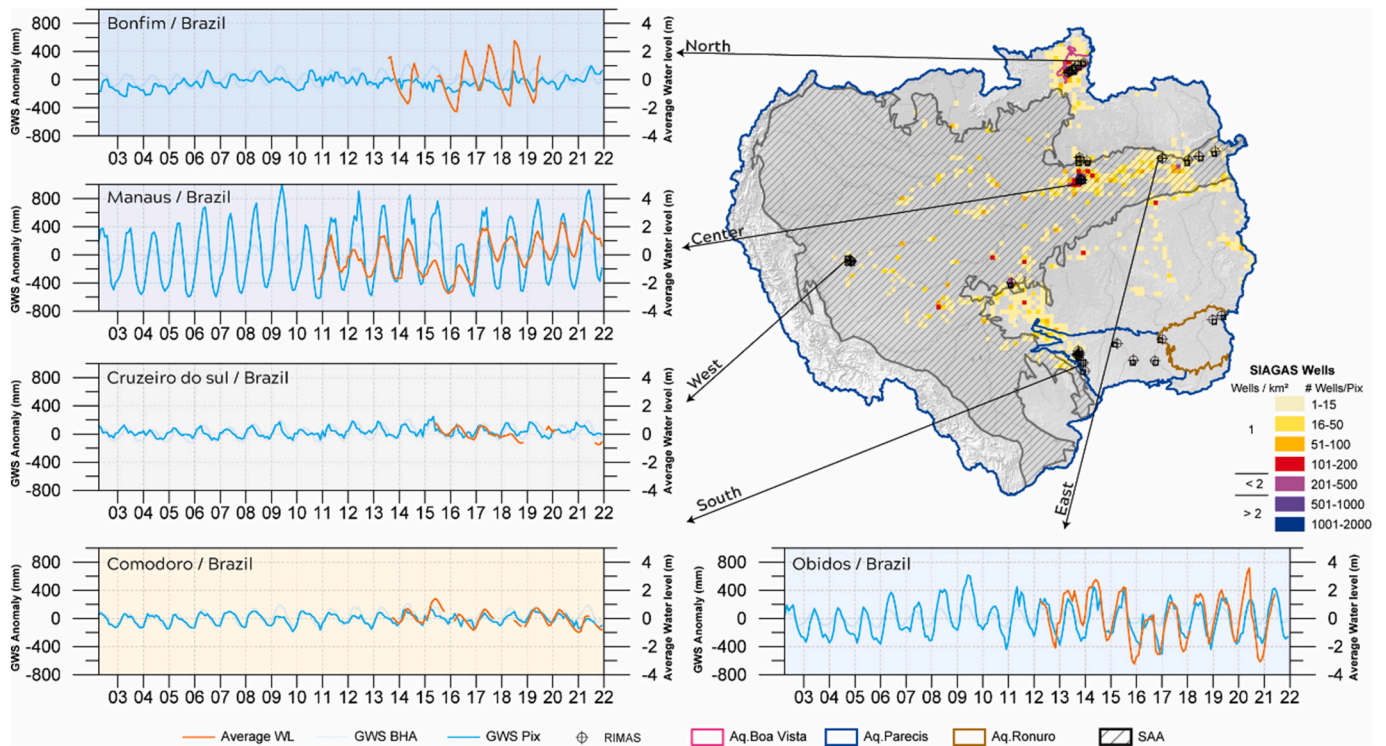


Fig. 5. Comparison between the Δ GWS AVERAGE time series and the RIMAS Δ WL for pixels in Bonfim, Manaus, Cruzeiro do Sul, Comodoro, and Obidos. The density of SIAGAS exploitation wells per unit area (pixel 0.25°) are in color scales from yellow to purple. The background of the time series corresponds to blue for a positive trend (surplus), yellow for a negative trend (deficit), and gray for a neutral trend or constant behavior. The AAS is in a dashed line with a black outline.

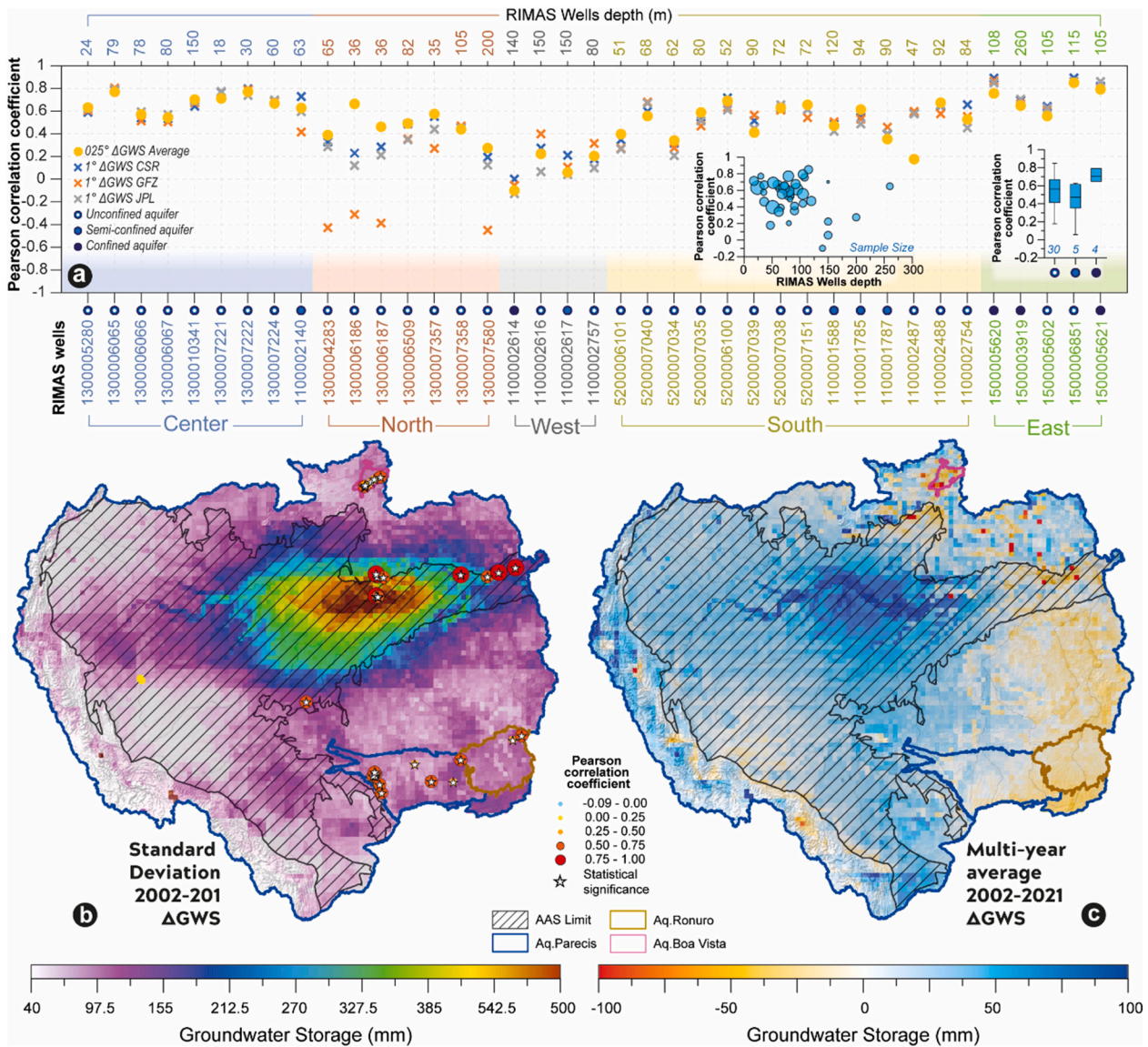


Fig. 6. Spatialization of the correlation coefficients, where a) comparison of the correlation coefficients for the RIMAS well zones, their depths, and aquifer types; b) standard deviation of the time series of ΔGWS , together with the correlation coefficients of the RIMAS wells and their statistical significance; and c) a multiyear average of ΔGWS between 20,002 and 2021.

Mamirauá Natural Reserve and the Purus, Madeira, and Tapajós rivers, underscoring the substantial influence of changes in surface water storage on ΔGWS . Positive anomalies prevail during the months of MAMJ, while negative anomalies occur during SOND. In the Northern region, anomalies exhibit maximum negatives during DJFM and maximum positives during JJAS, with a phase lag of approximately 2 to 3 months compared to the Southern region of the ARB. These storage patterns are influenced by seasonal variations in precipitation, which, in turn, are linked to the Intertropical Convergence Zone (ITCZ) migration and the South American monsoon system.

The ΔGWS time series in the different supports and at 1° and 0.25° resolutions showed a positive linear trend, superior to those of the ΔTWS . The average value between April 2002 and December 2021 for a spatial resolution at 1° was $+21.49 \pm 0.85 \text{ km}^3/\text{year}$ and at 0.25° , $+22.24 \pm 1.18 \text{ km}^3/\text{year}$. At both resolutions, the support with smaller linear trends was GFZ. After downscaling, all the supports increased the spatial variance and their linear trend by 6.5 % for CSR, 1.6 % for GFZ, and 2.1 % for JPL. During the time analyzed, the ARB would have an average total gain of $432.94 \pm 26.7 \text{ km}^3$ of stored groundwater. Table 3 summarizes the linear trends of the ΔGWS for the

ARB (for TWS see Table S3).

The El Niño events of 2010, 2015, and 2016 strongly negatively influenced the GWS anomalies. Fig. 8 shows the time series of the average ΔGWS values for the ARB (for TWS in Fig. S5). The standard deviation and interquartile ranges tended to be higher during extreme events; however, a widening was evident from 2012 onwards. This represented a greater spatial variance between positive and negative values. On the other hand, the decrease in SWS in snowy areas in southern Peru translated into significant positive ΔGWS anomalies. These changes were not necessarily representative since a portion of these melted masses reached the rivers, with respect to the portion infiltrated and effectively converted into recharge or ΔGWS .

The spatiotemporal distribution of the average ARB values is presented in Fig. 9 (for TWS see Fig. S7). Globally, a positive trend was evident over the lower basin towards the central, eastern, and southern zones within the AAS and its margins. Over the Andes, some pixels can be observed South of Peru, especially in the Vilcanota mountain range in the Ausangate, Suyuparina, Quisoquipina, and Quelccaya snow-capped mountains, where the strong influence of the SWS descent generated strong positive anomalies in the GWS. Negative anomalies developed

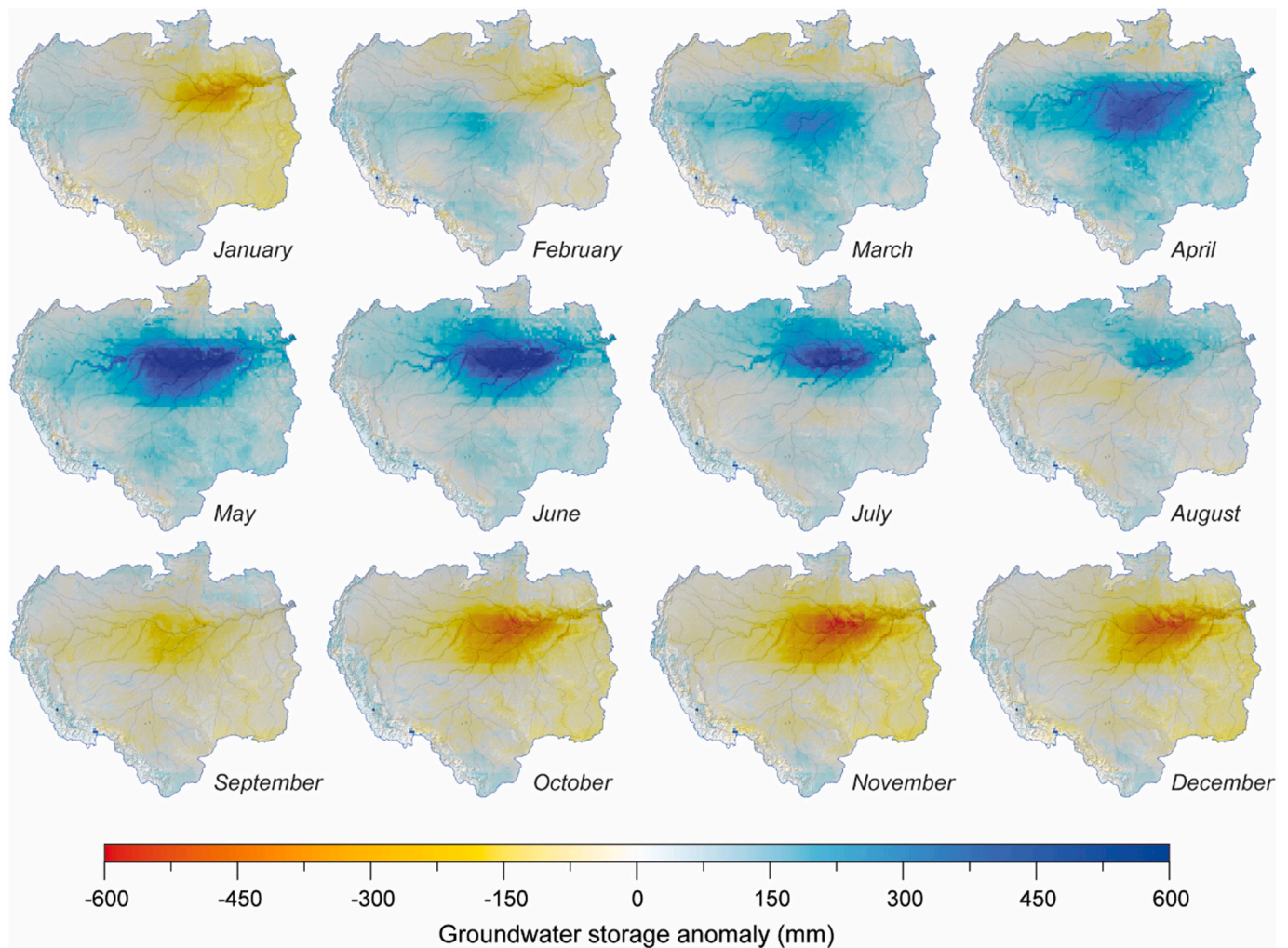


Fig. 7. Spatio-temporal distribution of multiyear seasonality of ΔGWS in the ARB between 2022 and 2021.

Table 3

Linear trends of ΔGWS in the ARB with an area of 6,089,336.3 km² for the different supports and resolutions between 2022 and 2021.

Resolution	Support	Rate (mm/year)	Rate (km ³ /year)	Total (km ³)
1°	CSR	+3.575	+21.772	+429.996
	GFZ	+3.367	+20.500	+404.870
	JPL	+3.647	+22.209	+438.623
0.25°	CSR	+3.811	+23.204	+458.280
	GFZ	+3.422	+20.836	+411.517
	JPL	+3.725	+22.685	+448.031
	Prom	+3.653	+22.242	+439.276

over the Alter do Chão, Parecis, Ronuro, and Boa Vista aquifers and the transnational zones between the Andes and the lower basin, East of Peru and Ecuador.

4.6. ΔTWS and scaled ΔGWS comparison

Correlation coefficients were estimated between the ΔWL time series of the RIMAS monitoring wells with ΔTWS and ΔGWS at resolutions of 0.25° with the AVERAGE dataset and 1° with the GFZ support. The time series of the RIMAS pixels were the average of the ΔWL anomalies of the wells contained in that pixel, and the RIMAS wells were by analyzing each well individually. Fig. 10 presents the boxplots of the Pearson values of the ΔGWS and ΔTWS time series at 0.25° and 1° (GFZ) with the RIMAS Wells and Pixel averages.

The correlations with the ΔWL RIMAS wells were similar at both resolutions, being more accurate at 0.25° and particularly with the GWS anomalies. Something similar occurred with the pixel averages of the wells, where the differences were smaller, but were still more accurate at 0.25°. This demonstrated an improvement in representativeness between satellite data and monitoring wells post-downscaling. Most of the ΔGWS coefficients were moderately lower than the ΔTWS coefficients, unlike in the southern zone, where the median was slightly higher. This difference was generated by the variance and noise acquired during the decomposition, which can be caused by the lack of continuity in some time series, the effect of local depletion and a lower number of measurements. A comparison of the correlation coefficients between the components of the RIMAS ΔWL time series with ΔTWS and ΔGWS average at 0.25° for the different monitoring wells and zones of the ARB is presented in the radial graph of Fig. 11.

An interesting aspect of this analysis was the increase in the correlation coefficients in the seasonality of the ΔGWS anomalies scaled with the RIMAS monitoring data. In the East, North, and Central zones, there was not much difference to ΔTWS ; however, an increase was evident in the West and South zones. This increase in the Seasonality correlation between ΔGWS and RIMAS ΔWL pix indicated an improvement in the representativeness of the post-decomposition data. As for the long-term behavior through the trend, there were significant differences to the North (average difference of 0.32) and West (average difference of 0.58), where the monitoring wells were more correlated with TWS . The original time series and the Seasonal Adjustment coefficients were

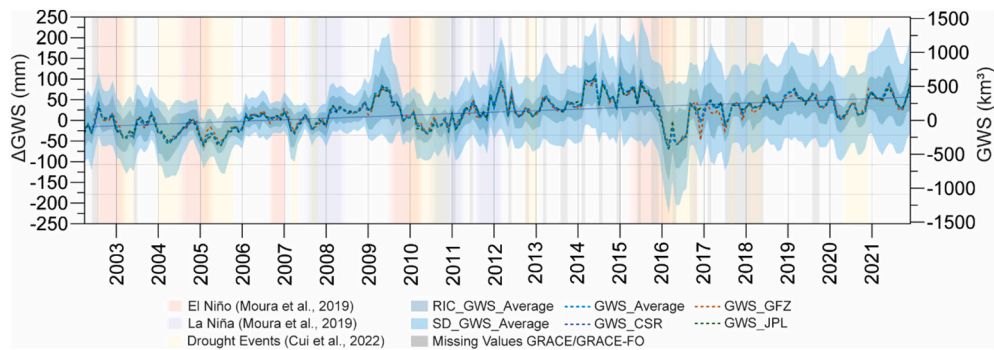


Fig. 8. Comparison of the Seasonal Adjustment components in the ΔGWS time series between the original data derived from GRACE/GRACE-FO and the AVERAGE scaled with AB. El Niño and La Niña phenomena obtained from Moura et al. (2019) and main drought events in Cui et al. (2022).

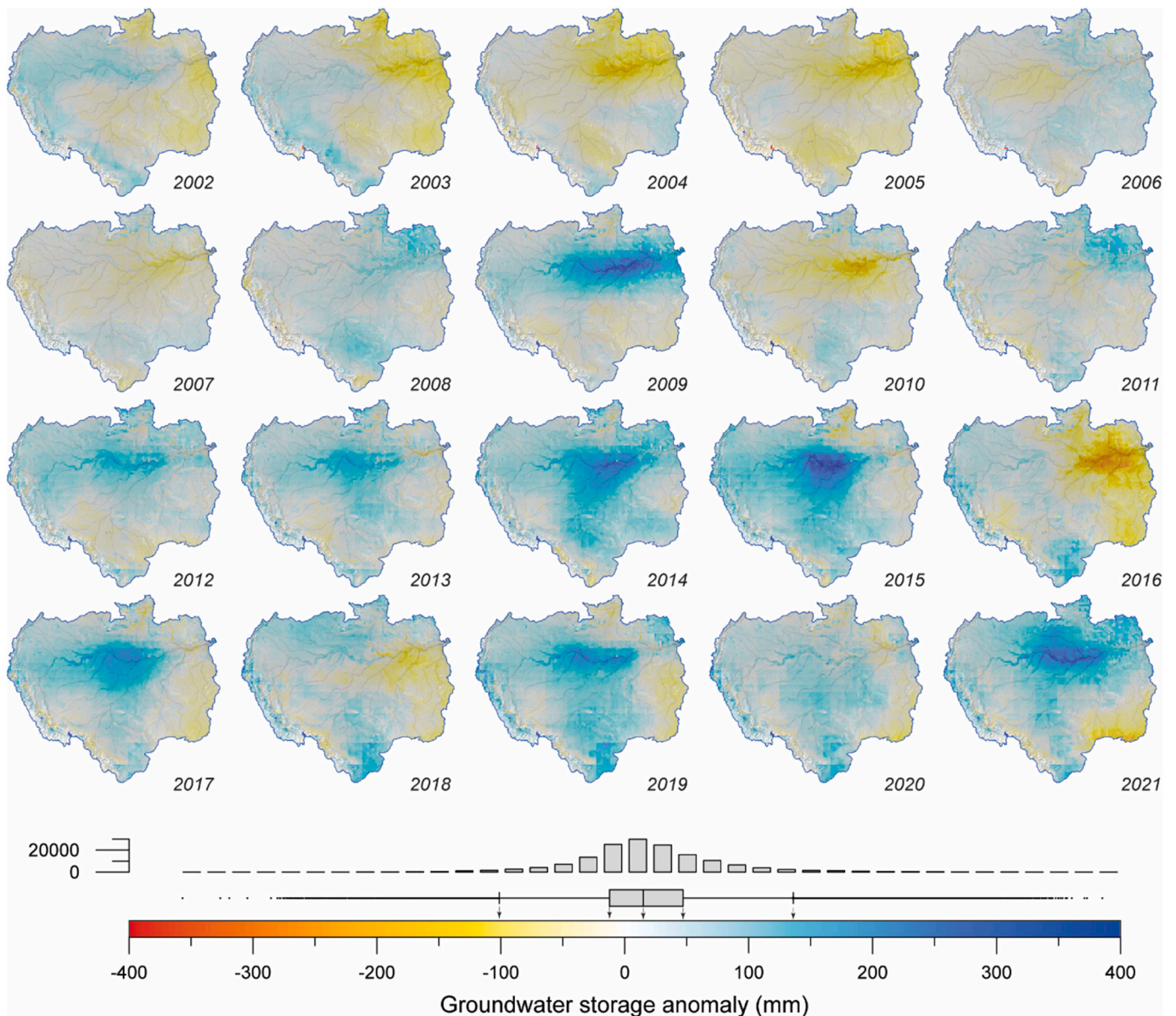


Fig. 9. Spatiotemporal distribution of average annual GWS anomalies within the ARB between 2002 and 2021.

similar, although a slight superiority of TWS over GWS prevailed.

When interpreting the variations in linear trends over time, there were significant differences between ΔTWS and ΔGWS in response to

climatic phenomena such as El Niño and La Niña. Notably, during 2016, there was a substantial reduction in the rates, being much more pronounced in ΔTWS , indicating a greater sensitivity to climate actors in

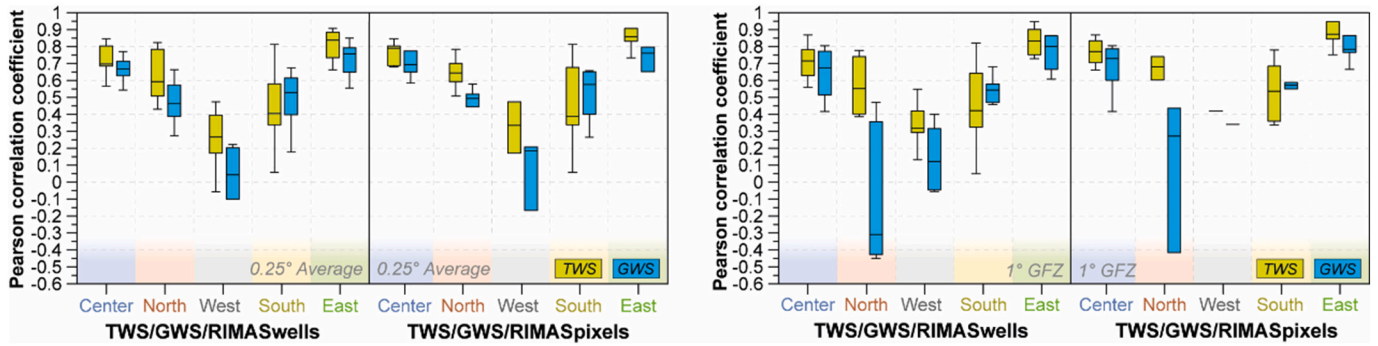


Fig. 10. Boxplots of the Pearson correlation coefficients between the time series of the Δ WL of the RIMAS monitoring wells and the Δ GWS and Δ TWS anomalies for the different zones of the ARB. The average values of the wells contained per pixel and the individual values of each well are compared on the left with the anomalies of the AVERAGE dataset at 0.25° and on the right with GFZ at 1°.

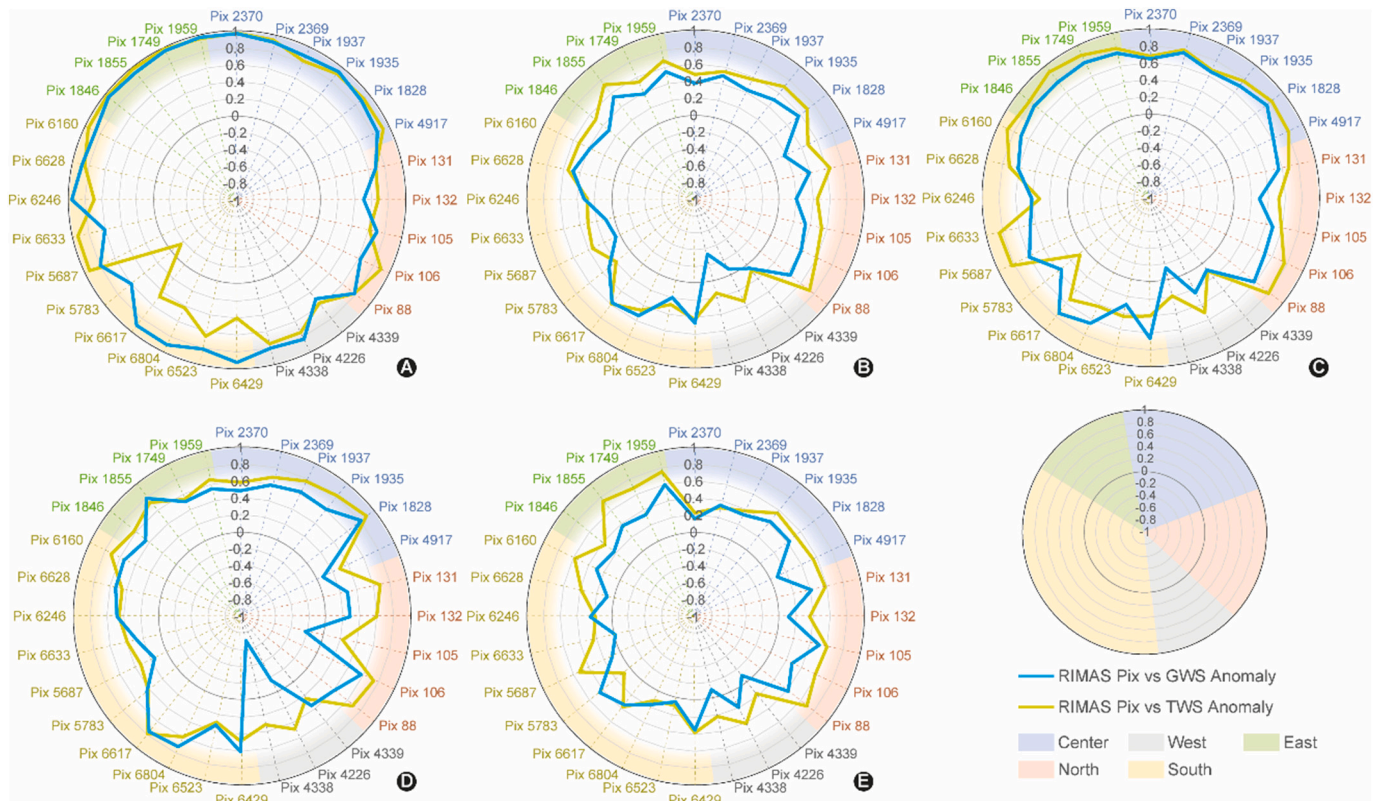


Fig. 11. Radial graph with the correlation coefficients between the RIMAS Δ WL with the Δ GWS and Δ TWS AVERAGE anomalies for the components of a) Seasonality, b) Seasonal Adjustment, c) Total, d) Trend, and e) Residual.

surface waters to groundwater (Fig. 12). From 2017 to 2021, the linear trends stabilized with GWS magnitudes on the order of 23 km³/year compared to the 12 km³/year of Δ TWS. Although at a much lower rate, albeit relatively constant, Δ GWS continued to decrease from magnitudes of 28.8 km³/year in 2016 to 22.2 km³/year in 2021, which, if it continues, could reach a neutral or negative trend between 2038 and 2043.

Since the El Niño phenomenon of 2015–2016, changes were generated in the rates of growth and decrease in the ARB water storage changes, going from positive to negative, especially those of Δ GWS. Fig. S4 shows the strong influence of the phenomenon within the Alter do Chão Aquifer; however, as shown in Fig. 12, the behavior from 2017 to 2021 is positive, showing a prompt recovery. The areas with the greatest losses are within the ARB’s middle basin. Some RIMAS wells, such as those in São Jose do Xingu and Presidente Figueiredo, corroborated this negative behavior in long-term trends. Fig. 13 compares the

linear trends in the Δ GWS anomalies scaled by pixel for the ARB.

5. Discussion

5.1. Downscaling with machine learning algorithms

This research generated a TWS dataset for various GRACE/GRACE-FO solutions, along with an average of GWS, at a spatial resolution of 0.25° and a monthly temporal resolution without temporal gaps in the ARB. It applied machine learning models that adapted to the region’s specific conditions and characteristics, considering hydrometeorological and morphological variables. This allowed for obtaining more accurate predictions for the ARB, reflecting local patterns, effects, and relationships. It is worth noting that in this process, the inherent uncertainties of the models and the remote sensors used as sources of information were

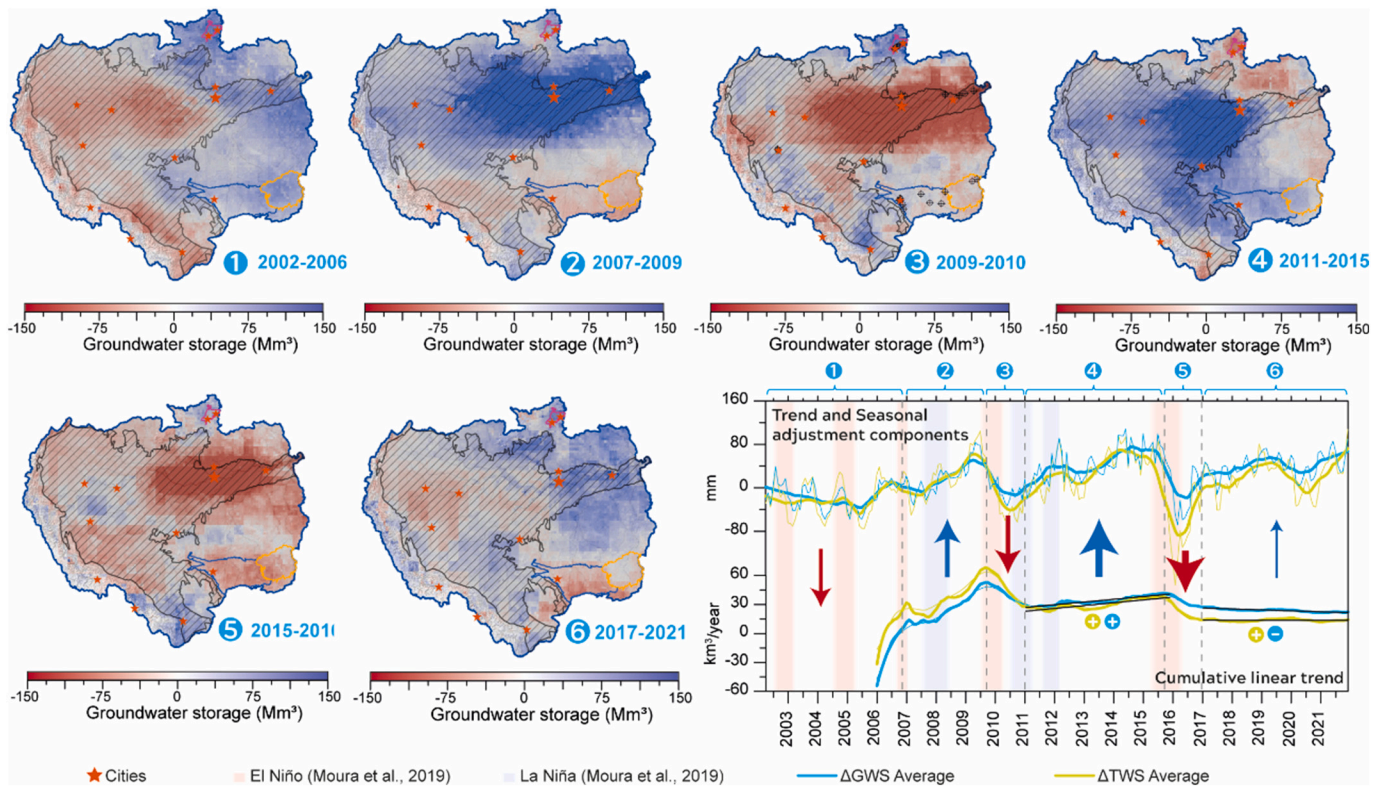


Fig. 12. Long-term behavior of the ΔTWS and ΔGWS time series of the ARB, where a) Trend and Seasonal adjustment, and b) variations in linear trends over time for the Trend and Seasonal adjustments.

also assimilated. Consistent with the findings of [Fassoni-Andrade et al. \(2021\)](#), variables such as precipitation and evaporation in the ARB show high spatial and temporal variability. This is compounded by the region's complex topography and varied vegetation cover, making it challenging to obtain precise and reliable measurements through remote sensing methods. It evaluated several precipitation datasets used in the ARB, such as TRMM (Tropical Rainfall Measuring Mission) and GPM (Global Precipitation Measurement Mission), and found that CHIRPS exhibited a strong correlation with in-situ measurement stations in the ARB, as reported by [Haghtalab et al. \(2020\)](#).

Additionally, GLEAM data demonstrated reasonable and consistent accuracy ([Gomis-Cebolla et al., 2019; Martens et al., 2017; Zhang et al., 2019](#)), though some limitations were identified in areas with high soil evaporation, as noted by [Fassoni-Andrade et al. \(2021\)](#). Regarding GLDAS, it utilized 16 of its 36 variables. This dataset combines the assimilation of in-situ observation data and remote sensors with hydrological models and is widely used in decomposing the GRACE/GRACE-FO signal. The diversity of information sources granted exploring and identifying the most relevant variables and their relative contribution to the models' generalization capacity. This approach enabled the models to be trained with variables that contributed diverse characteristics to the predictions.

In terms of accuracy, the evaluation metrics, including R, MAE, and RMSE, yielded results similar to those obtained by [Chen et al. \(2019\)](#) and [Ali et al. \(2021\)](#) when employing similar methodologies and algorithms, despite differences in the extent and complexity of the study locations. However, a slight advantage in prediction accuracy was observed when using the AB algorithm compared to RF, possibly due to its higher sensitivity to extreme values. This study analyzed 39 monitoring wells, and strong correlations (exceeding 0.6) were found in 4 out of the five examined zones, validating the downscaled results. The wells' location, depth, and lithological characteristics provided key information for characterizing local behaviors in the Alter do Chão Aquifer System (AAS) and the Parecis, Ronuro, and Boa Vista aquifers. It's important to

note that correlations decreased at greater well depths, suggesting they may be related to confined or semi-confined portions of the aquifers.

Furthermore, the downscaled process improved data accuracy, as evidenced by the increased well correlations compared to 1° data. When comparing well correlations with GWS and TWS at a 0.25° temporal resolution across different temporal components ([Fig. 11](#)), it was observed that some regions exhibited higher correlations with TWS than with GWS. However, the seasonal component of GWS displayed the highest correlations with all the wells.

In summary, the results of this research support the utility of downscaling using machine learning algorithms to enhance data accuracy in the ARB, particularly in the context of aquifer monitoring and the interactions of hydrometeorological variables in the region. These findings have significant implications for understanding and managing water resources in the ARB and can contribute to future research in hydrology and groundwater monitoring.

5.2. GWS spatial-temporal characteristics

During the estimation of GWS, the changes in surface water storage obtained from GLDAS data underestimate the true magnitude. Nevertheless, this underestimation can be addressed by applying methodologies such as those proposed by [Frappart et al. \(2008, 2012\)](#) and [Ndehedehe and Ferreira \(2020\)](#). These methodologies combine the extent of surface waters with altimetric time series of water levels over rivers and floodplains. Furthermore, when examining the seasonality maps ([Fig. 7](#)), annual averages ([Fig. 9](#)), and standard deviation ([Fig. 6](#)) of GWS, spatial patterns can be identified in which extreme magnitudes are associated with geological units, particularly the major aquifers in the ARB, as well as floodplains, surface water bodies, and some topographic features. The high seasonality in these regions is linked to recharge patterns due to precipitation and the unconfined nature of aquifers and some residual soils in the ARB, which aligns with the findings of [Hu et al. \(2017\)](#).

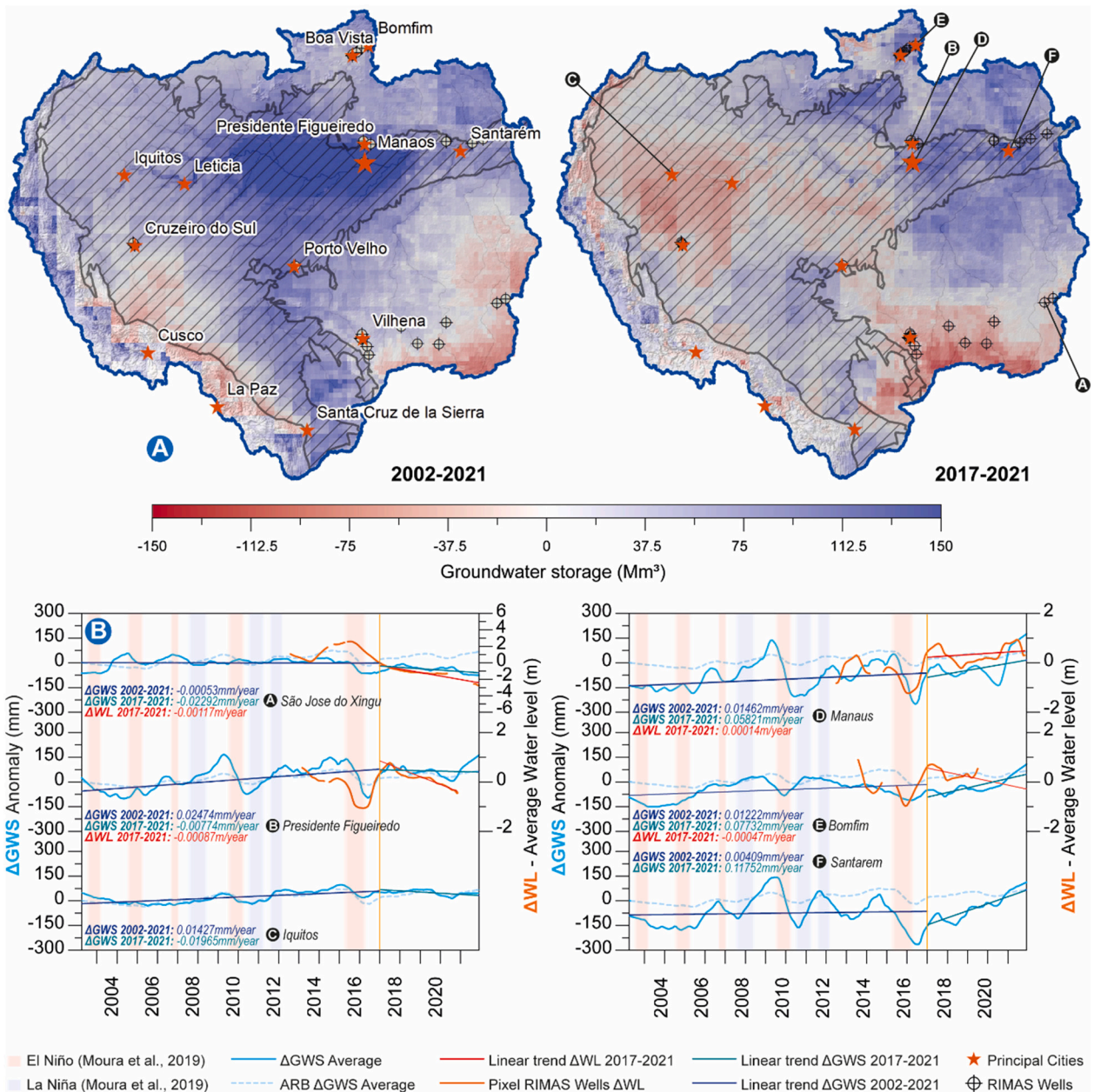


Fig. 13. Long-term ΔGWS behavior of the ARB, where a) Spatialization of GWS derived from the linear trend between 2002 and 2021 and 2017–2021, and b) Comparison of the time series of ΔGWS and ΔWL pixels along with their linear trends after 2017 and between 2002 and 2021.

The seasonal component accounts for the major proportion of both TWS and GWS. Monthly values obtained using the method of Cleveland et al. (1990) repeat periodically. However, the original signal exhibits variations in amplitude and frequency over time. An analysis of the power spectrum of this original seasonality may provide additional interpretative criteria to discern changes at different time intervals. Within these methodologies, the application of Least Squares Cross-Wavelet Analysis (LSCWA) or Cross-Wavelet Transform (XWT) can be employed, enabling the identification of periods during which the time series exhibit significant seasonality and their temporal lags (Ghadepour et al., 2023; Li et al., 2020).

As shown in Fig. 12, the linear trends over time indicate that starting from the El Niño event in 2009–2010 positive GWS rates exceed those of TWS. However, from 2017 onward, GWS rates consistently decreased, in contrast to TWS, which remains relatively constant. These magnitudes and directions of the linear trends align with the research of Ndehedehe and Ferreira (2020) and Scanlon et al. (2016). In the maps of linear

trends for the period 2002–2021, it can be observed that regions with negative values are located over the Ronuro and Parecis Aquifers, areas heavily impacted by human activities such as burning, deforestation, and extensive agriculture. In the Xingu River sub-basin, negative trends are related to disruptions in the aquifer and residual soil recharge, where deforestation leads to increased surface runoff, reduced evapotranspiration, and consequently, decreased infiltration, in line with the findings of Panday et al. (2015). Spatial analysis of the linear trends for different periods indicates a high susceptibility of the Içã and Alter do Chão aquifers to El Niño and La Niña events. The continual population growth, its demand for water (primarily from underground sources), and the development of various activities like burning, deforestation, the construction of hydroelectric dams, and mining, among others, disrupt the hydrological cycle and the ecosystem services of the ARB (Latrubesse et al., 2017; Neill et al., 2011; Nobre et al., 2009).

A pixel located over the city of Manaus, which exhibits statistically significant differences compared to the original GRACE data, shows a

negative average of GWS between 2002 and 2021 (Fig. S8). Furthermore, this pixel has a higher density of extraction wells recorded by SIAGAS and significant correlations (>0.6) with the RIMAS monitoring wells. This may indicate an improvement in the representation of hydrological phenomena at local scales, achieved through the downscaling process, thereby enhancing the possibilities for monitoring unconfined aquifers using remote sensing data.

5.3. Future research

Over time, the incorporation of more extensive and higher-quality time series will lead to a substantial improvement in model accuracy. The division of the ARB into subdomains, such as the upper, middle, and lower basins, can facilitate the separation and better assimilation of local climatic effects. Including categorical variables, such as geology and soil type (which remain constant over time), or indicators of human activity, like vegetation cover, well density, fires, deforestation, and population density (which vary over time), is essential for future research. It is imperative to consider the implementation of new downscaling methodologies with an even more precise target resolution, reaching levels of 0.05° or even 0.01° (equivalent to 5.5 km or 1.1 km, respectively), or even 0.004° (corresponding to 500 m), as achieved with approaches like those of Zhang et al. (2021) and Zhong et al. (2021). These methodologies should be based on morphometric and multi-spectral information, utilizing indices such as NDVI, NDWI, TWI, or TPI, among others.

Furthermore, efforts should be made to integrate tools that allow the joint modeling of groundwater and surface water. Additionally, new monitoring well data provided by the governments of Colombia, Ecuador, Peru, and Bolivia should be incorporated. To further enrich the research, it is necessary to complement and compare the results obtained with additional datasets, such as GIEMS, GHWS, and other GLDAS models like V2.2, which includes Δ TWS anomalies at a daily resolution of 0.5° . Furthermore, ML or Deep Learning models can be implemented in restoring TWS/GWS signals for periods without GRACE/GRACE-FO data, such as before 2002, or to enhance their forecasting capabilities, as suggested in (Dannouf et al., 2022; Kalu et al., 2022, 2023). These actions will significantly contribute to advancing our understanding of hydrology in the ARB and strengthen the foundations for the sustainable management of water resources in the region.

6. Conclusions

GRACE/GRACE-FO records allowed the exploration of new research on large-scale GWS variations. During this study, two ML models were compared in Δ TWS downscaling, evaluating multiple independent hydrometeorological variables. The results showed a slight superiority of the models based on AB algorithms over those based on RF algorithms. Downscaling from 1° to 0.25° was successfully achieved, decomposing the GWS anomalies between 2002 and 2021 for the ARB.

The error of the predictions was very low, showing good consistency between the data before and after downscaling. It determined the short and long-term behavior of the time series, obtaining linear trends for Δ TWS of $+14.26 \pm 1.18 \text{ km}^3/\text{year}$ and for Δ GWS of $+22.24 \pm 1.18 \text{ km}^3/\text{year}$ for all ARB. With an increase or gain in groundwater storage between 2002 and 2021 of $432.94 \pm 26.7 \text{ km}^3$ for the ARB. The validation with the RIMAS monitoring wells showed promising results with maximum correlation coefficients between the Δ WL with Δ GWS of 0.85 and Δ TWS of 0.90. The maximum correlations between Δ WL with Δ GWS and Δ TWS were 0.79/0.92 in the East, 0.77/0.84 in the Center, 0.73/0.88 in the South, 0.57/0.79 in the North and 0.23/0.49 in the West of the ARB. Spatial patterns were also identified in the anomalies related to the strong influence of some water bodies and topographic features, such as snow-capped mountains in southern Peru, which would not have been possible to identify at a lower spatial resolution.

Differences in the SD and linear trend of the predicted or scaled Δ TWS and Δ GWS show the spatial and temporal variance in behavior between the upper (Andes) and lower (plains) basins, which can be taken into consideration to define domains before training future downscaling models.

Downscaling using AB effectively broadened the spatial resolution of the GRACE and GRACE-FO-derived data and improved their accuracy. Hypothesis testing to the mean between the scaled pixels (0.25°) and the original resolution (1°) in the three supports (CSR, GFZ, and JPL) evidenced significant changes derived from downscaling, mainly to the North of the ARB, in the Andes region and isolated pixels such as one located over the city of Manaus. These significant changes resulted from gains in spatial and temporal variance resulting from downscaling, in which features of the training variables were assimilated into the ML models. Extreme periods caused by El Niño and La Niña phenomena were identified in years such as 2002, 2009, 2012, and 2016, highlighting the high vulnerability of some aquifers of the AAS to climatic effects. Future research should consider the impacts of anthropogenic activity at regional and local scales. The high correlation in the analyzed areas between the Δ WL and Δ GWS promised results regarding using GRACE/GRACE-FO data in the regional monitoring of unconfined aquifers, becoming a valuable tool for water management in large areas with difficult access.

Funding

This research was supported by São Paulo Research Foundation (FAPESP) [grant 2014/00844-7] and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) [Finance Code 001].

CRedit authorship contribution statement

Diego Alejandro Satizábal-Alarcón: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Alexandra Suhogusoff:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Writing – review & editing. **Luiz Carlos Ferrari:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The TWS datasets for each solution (CSR, GFZ, and JPL) obtained using the AB algorithm, as well as the average TWS and GWS for the ARB, are provided in the repository. The files are stored in geodatabases that contain raster data for the 237 analyzed months. Data archived in the Mendely.data repository which can be accessed using the following link (<https://data.mendeley.com/preview/2jpb6c3gdh8?a=8ab99768-5bd7-41b1-be02-0f4a6796ee60>) (Satizábal-Alarcón et al., 2023).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2023.168958>.

- observations. *Hydrol. Earth Syst. Sci.* 15 (2), 453–469. <https://doi.org/10.5194/hess-15-453-2011>.
- Moura, M.M., dos Santos, A.R., Pezzopane, J.E.M., Alexandre, R.S., da Silva, S.F., Pimentel, S.M., de Andrade, M.S.S., Silva, F.G.R., Branco, E.R.F., Moreira, T.R., da Silva, R.G., de Carvalho, J.R., 2019. Relation of El Niño and La Niña phenomena to precipitation, evapotranspiration and temperature in the Amazon basin. *Sci. Total Environ.* 651, 1639–1651. <https://doi.org/10.1016/j.scitotenv.2018.09.242>.
- Nagy, L., Artaxo, P., Forsberg, B.R., 2016. *Interactions Between Biosphere, Atmosphere and Human Land Use in the Amazon Basin*, 1st ed. Springer. <https://doi.org/10.1007/978-3-662-49902-3>.
- Ndehedehe, C.E., Ferreira, V.G., 2020. Assessing land water storage dynamics over South America. *J. Hydrol.* 580 <https://doi.org/10.1016/j.jhydrol.2019.124339>.
- Neill, C., Chaves, J.E., Biggs, T., Deegan, L.A., Elsenbeer, H., Figueiredo, R.O., Germer, S., Johnson, M.S., Lehmann, J., Markewitz, D., Piccolo, M.C., 2011. Runoff sources and land cover change in the Amazon: an end-member mixing analysis from small watersheds. *Biogeochemistry* 105 (1), 7–18. <https://doi.org/10.1007/s10533-011-9597-8>.
- Nobre, P., Malagutti, M., Urbano, D.F., De Almeida, R.A.F., Giarolla, E., 2009. Amazon deforestation and climate change in a coupled model simulation. *J. Clim.* 22 (21), 5686–5697. <https://doi.org/10.1175/2009JCLI2757.1>.
- Panday, P.K., Coe, M.T., Macedo, M.N., Lefebvre, P., Castanho, A.D. de A., 2015. Deforestation offsets water balance changes due to climate variability in the Xingu River in eastern Amazonia. *J. Hydrol.* 523, 822–829. <https://doi.org/10.1016/j.jhydrol.2015.02.018>.
- Pfeffer, J., Seyler, F., Bonnet, M. P., Calmant, S., Frappart, F., Papa, F., Paiva, R. C. D., Satgé, F., & Silva, J. S. da. (2014). Low-water maps of the groundwater table in the central Amazon by satellite altimetry. *Geophys. Res. Lett.*, 41(6), 1981–1987. doi: <https://doi.org/10.1002/2013GL059134>.
- Pimentel, E.T., Hamza, V.M., 2012. Indications of regional scale groundwater flows in the Amazon basins: inferences from results of geothermal studies. *J. S. Am. Earth Sci.* 37, 214–227. <https://doi.org/10.1016/j.jsames.2012.03.007>.
- Pokhrel, Y.N., Fan, Y., Miguez-Macho, G., Yeh, P.J.F., Han, S.C., 2013. The role of groundwater in the Amazon water cycle: 3. Influence on terrestrial water storage computations and comparison with GRACE. *J. Geophys. Res. Atmos.* 118 (8), 3233–3244. <https://doi.org/10.1002/jgrd.50335>.
- Pulla, S.T., Yasarer, H., Yarbrough, L.D., 2023. GRACE Downscaler: a framework to develop and evaluate downscaling models for GRACE. *Remote Sens.* 15 (9) <https://doi.org/10.3390/rs15092247>.
- Reis, V., Herosmo, V., Hamilton, S.K., Bunn, S.E., Fluet-Chouinard, E., Venables, B., Linke, S., 2019. Characterizing seasonal dynamics of Amazonian wetlands for conservation and decision making. *Aquat. Conserv. Mar. Freshwat. Ecosyst.* 29 (7), 1073–1082. <https://doi.org/10.1002/aqc.3051>.
- Rodell, B.M., Houser, P.R., Jambor, U., Gottschalk, J., Mitchell, K., Meng, C., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D., Toll, D., 2004. The global land data assimilation system. *Am. Meteorol. Soc.* 381–394.
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.T., Chuang, H.Y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., Becker, E., 2014. The NCEP climate forecast system version 2. *J. Clim.* 27 (6), 2185–2208. <https://doi.org/10.1175/JCLI-D-12-00823.1>.
- Satizábal-Alarcón, D.A., Suhogusoff, A., Ferrari, L., 2023. Dataset: characterization of groundwater storage changes in the Amazon River Basin based on downscaling of GRACE/GRACE-FO data with machine learning models. *Mendeley Data V1*. <https://doi.org/10.17632/2jpc3gdh8.1>.
- Save, H., Bettadpur, S., Tapley, B.D., 2016. High-resolution CSR GRACE RL05 mascons. *J. Geophys. Res. Solid Earth* 121 (10), 7547–7569. <https://doi.org/10.1002/2016JB013007>.
- Scanlon, B.R., Zhang, Z., Save, H., Wiese, D.N., Landerer, F.W., Long, D., Longuevergne, L., Chen, J., 2016. Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resour. Res.* 52 (12), 9412–9429. <https://doi.org/10.1002/2016WR019494>.
- Scheel, M.L.M., Rohrer, M., Huggel, Ch., Santos Villar, D., Silvestre, E., Huffman, G.J., 2011. Evaluation of TRMM multi-satellite precipitation analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrol. Earth Syst. Sci.* 15 (8), 2649–2663. <https://doi.org/10.5194/hess-15-2649-2011>.
- Seyoum, W.M., Kwon, D., Milewski, A.M., 2019. Downscaling GRACE TWSA data into high-resolution groundwater level anomaly using machine learning-based models in a glacial aquifer system. *Remote Sens.* 11 (7) <https://doi.org/10.3390/rs11070824>.
- Swenson, S., Wahr, J., Milly, P.C.D., 2003. Estimated accuracies of regional water storage variations inferred from the gravity recovery and climate experiment (GRACE). *Water Resour. Res.* 39 (8) <https://doi.org/10.1029/2002WR001808>.
- Tapley, B.D., Bettadpur, S., Ries, J.C., Thompson, P.F., Watkins, M.M., 2004. GRACE measurements of mass variability in the earth system. *Science* 305 (5683), 503–505. <https://doi.org/10.1126/science.1099192>.
- Tomasella, J., Hodnett, M.G., Cuartas, L.A., Nobre, A.D., Waterloo, M.J., Oliveira, S.M., 2008. The water balance of an Amazonian micro-catchment: the effect of interannual variability of rainfall on hydrological behaviour. *Hydrol. Process.* 22 (13), 2133–2147. <https://doi.org/10.1002/hyp.6813>.
- Wahr, J., Swenson, S., Zlotnicki, V., Velicogna, I., 2004. Time-variable gravity from GRACE: first results. *Geophys. Res. Lett.* 31 (11) <https://doi.org/10.1029/2004GL019779>.
- Wang, J., Xu, D., Li, H., 2023. Constructing GRACE-based 1 km resolution groundwater storage anomalies in arid regions using an improved machine learning downscaling method: A case study in Alxa league, China. *Remote Sens.* 15 (11) <https://doi.org/10.3390/rs15112913>.
- Watkins, M.M., Wiese, D.N., Yuan, D.N., Boening, C., Landerer, F.W., 2015. Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *J. Geophys. Res. Solid Earth* 120 (4), 2648–2671. <https://doi.org/10.1002/2014JB011547>.
- Wiese, D.N., Landerer, F.W., Watkins, M.M., 2016. Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resour. Res.* 52 (9), 7490–7502. <https://doi.org/10.1002/2016WR019344>.
- Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J.C., Sampson, C.C., Kanae, S., Bates, P.D., 2017. A high-accuracy map of global terrain elevations. *Geophys. Res. Lett.* 44 (11), 5844–5853. <https://doi.org/10.1002/2017GL072874>.
- Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P.D., Allen, G.H., Pavelsky, T.M., 2019. MERIT hydro: A high-resolution global hydrography map based on latest topography dataset. *Water Resour. Res.* 55 (6), 5053–5073. <https://doi.org/10.1029/2019WR024873>.
- Yin, W., Zhang, G., Liu, F., Zhang, D., Zhang, X., Chen, S., 2022. Improving the spatial resolution of GRACE-based groundwater storage estimates using a machine learning algorithm and hydrological model. *Hydrogeol. J.* 30 (3), 947–963. <https://doi.org/10.1007/s10040-021-02447-4>.
- Yuan, X., Wood, E.F., Luo, L., Pan, M., 2011. A first look at climate forecast system version 2 (CFSv2) for hydrological seasonal prediction. *Geophys. Res. Lett.* 38 (13) <https://doi.org/10.1029/2011GL047792>.
- Zhang, Y., Kong, D., Gan, R., Chiew, F.H.S., McVicar, T.R., Zhang, Q., Yang, Y., 2019. Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017. *Remote Sens. Environ.* 222, 165–182. <https://doi.org/10.1016/j.rse.2018.12.031>.
- Zhang, J., Liu, K., Wang, M., 2021. Downscaling groundwater storage data in China to a 1-km resolution using machine learning methods. *Remote Sens.* 13 (3) <https://doi.org/10.3390/rs13030523>.
- Zhong, D., Wang, S., Li, J., 2021. Spatiotemporal downscaling of grace total water storage using land surface model outputs. *Remote Sens.* 13 (5), 1–19. <https://doi.org/10.3390/rs13050900>.