Revista Brasileira de Climatologia
ISSN: 2237-8642 (Eletrônica)

VALIDAÇÃO DE DIFERENTES BASES DE DADOS DE PRECIPITAÇÃO NAS BACIAS HIDROGRÁFICAS DO SAPUCAÍ E SÃO FRANCISCO

TORRES, Francisca Lanai Ribeiro - lanai.torres@hotmail.com
Universidade Federal de Itajubá / UNIFEI

FERREIRA, Glauber Willian de Souza - glauber_ferreira@unifei.edu.br
Universidade Federal de Itajubá / UNIFEI

KUKI, Cassia Akemi Castro - cassiakuki@yahoo.com.br
Universidade Federal de Itajubá / UNIFEI

VASCONCELLOS, Bruna Tayla Cabral de - tayla_bruna@yahoo.com.br
Universidade Federal de Itajubá / UNIFEI

FREITAS, Aline Araújo de - alinearaujof1020@gmail.com
Universidade Federal de Itajubá / UNIFEI

SILVA, Paola do Nascimento - paoola.silva@hotmail.com
Universidade Federal de Itajubá / UNIFEI

SOUZA, Christie André - christie@unifei.edu.br
Universidade Federal de Itajubá / UNIFEI

REBOITA, Michelle Simões - reboita@unifei.edu.br
Universidade Federal de Itajubá / UNIFEI

Submetido em: 10/05/2020
Aceito para publicação em: 20/07/2020
Publicado em: 11/08/2020
DOI: http://dx.doi.org/10.5380/abclima.v27i0.73634

RESUMO: A densidade da rede de instrumentos de medição de precipitação no Brasil ainda precisa ser ampliada, a fim dessa variável ser representada adequadamente nas bacias hidrográficas. Neste cenário, uma alternativa para os estudos que necessitam de séries temporais de precipitação é a utilização de dados em pontos de grade (chamados de análises). Estes dados podem resultar da interpolação de medições *in situ* juntamente com estimativas por satélite. Assim, o objetivo do presente estudo é validar análises de precipitação mensal e diária nas bacias do São Francisco e do Sapucaí mediante comparação com dados observados *in situ*. Em escala mensal, foram validadas as análises do CPC, GPCC, GPCP e MERGE-CPTEC, enquanto em escala diária validaram-se as análises do CPC e do MERGE-CPTEC. A aplicação de diferentes medidas estatísticas mostrou que as análises do CPC e MERGE-CPTEC apresentam, em geral, melhores resultados nas comparações, possibilitando sua utilização em escala mensal. Em escala diária, embora o MERGE-CPTEC possua uma performance ligeiramente melhor do que a do CPC, este último também pode ser utilizado satisfatoriamente nos estudos.

PALAVRAS-CHAVE: Precipitação; Bacia do São Francisco; Bacia do Sapucaí; Validação.

VALIDATION OF DIFFERENT PRECIPITATION DATASETS IN THE SAPUCAÍ AND SÃO FRANCISCO HYDROGRAPHIC BASINS

ABSTRACT: The density of the precipitation measurement network in Brazil still needs to be expanded, in order of this variable to be adequately represented in the hydrographic basins. In this scenario, an alternative for studies that require precipitation time series is

the use of data in grid points (called analyzes). These data can result from the interpolation of measurements *in situ* in conjunction with satellite estimates. Thus, the objective of the present study is to validate analyzes of monthly and daily precipitation in the São Francisco and Sapucaí basins through comparison with data observed *in situ*. On a monthly scale, CPC, GPCC, GPCP and MERGE-CPTEC analyzes were validated, while on a daily scale, validation of CPC and MERGE-CPTEC analyzes was done. The application of different statistical measures showed that the analyzes of the CPC and MERGE-CPTEC present, in general, better results in the comparisons, allowing their use on a monthly scale. On a daily scale, although MERGE-CPTEC performs slightly better than CPC, the latter one can also be used satisfactorily on the studies.

KEYWORDS: Precipitation; São Francisco Basin; Sapucaí Basin; Validation

1. INTRODUÇÃO

Como a precipitação é o principal parâmetro de entrada no ciclo hidrológico, sua variabilidade influencia diretamente a resposta de uma bacia hidrográfica (TUCCI, 2002; ZENG et al., 2008; ESPINOZA et al., 2011; MARENGO et al., 2012; CALDEIRA et al., 2015). Portanto, bases de dados históricas de precipitação são essenciais para o planejamento e o gerenciamento dos recursos hídricos e o dimensionamento de estruturas hidráulicas e sistemas de drenagem. Contudo, nem toda bacia hidrográfica conta com uma densa rede de estações meteorológicas e/ou hidrológicas. Fato que decorre da localização em áreas sem acessibilidade e/ou dos custos envolvidos para a manutenção dessas estações (FENSTERSEIFER, 2013; VICENTE et al., 2014). Some-se a isso a ausência de um planejamento adequado e de políticas públicas destinadas à ampliação das redes de obtenção de dados meteorológicos/hidrológicos. No Brasil, o Instituto Nacional de Meteorologia (INMET - http://www.inmet.gov.br/portal/) possui uma malha com mais de 400 estações meteorológicas convencionais e automáticas (http://www.inmet.gov.br/portal/index.php?r=home/page&page=sm_previsao_tempo), mas que ainda não é suficiente para a representação de todo o território nacional. A Agência Nacional de Águas (ANA) também possui uma rede de pluviômetros com cerca de 2767 estações que monitoram a chuva (http://www.snirh.gov.br/hidroweb/apresentacao), que é mais densa do que a do INMET. Mesmo assim, algumas bacias brasileiras não possuem um número adequado de estações de medição e/ou séries históricas extensas (http://www.snirh.gov.br/hidroweb). Além disso, muitas das estações da ANA não possuem dados atualizados em tempo real.

Uma alternativa para os estudos que necessitam de medidas de precipitação é a utilização de bases de dados provenientes de estimativas de satélites, de dados medidos *in situ* interpolados em pontos de grade regular (chamados de análises) e/ou da combinação de ambos. Entretanto, esses dados podem não apresentar uma boa representação da chuva, uma vez que estimativas de satélite possuem incertezas (XIE e ARKIN, 1997; HUFFMAN et al., 2001; TIAN e PETERS-LIDARD, 2010). Tais incertezas podem ser decorrentes dos instrumentos de medição e das suas resoluções espaciais e temporais (CUI et al., 2016). Franchito et al. (2009) indicaram que as estimativas do *Tropical Rainfall Measuring Mission* (TRMM) apresentam vieses sazonais em algumas regiões do Brasil. Por exemplo, comparando-se às climatologias de dados observados em estações meteorológicas, no trimestre junho-agosto (JJA) há condições mais secas nas regiões Nordeste e Centro-Oeste, enquanto no trimestre setembro-novembro (SON) há condições mais

chuvosas na região Sul. Já Collischonn et al. (2007), Rozante et al. (2010), Pereira et al. (2013) e Danelichen et al. (2013) indicaram que, em alguns setores da América do Sul, o sensor TRMM tende a superestimar a precipitação sobre o continente em relação aos dados de estações meteorológicas. Rozante et al. (2010) avaliaram o MERGE-CPTEC na América do Sul por meio de comparação com dados de superfície interpolados, mas removendo aleatoriamente 10% das estações meteorológicas. Em áreas com alta densidade de estações meteorológicas, o desempenho do MERGE-CPTEC é equivalente ao dos dados observados interpolados, mas em áreas com observações esparsas, o MERGE-CPTEC apresentou resultados superiores (o que foi testado com a exclusão de algumas estações meteorológicas). Villanueva-Baez et al. (2018) avaliaram seis estimativas de precipitação por satélite (TRMM 3B42v7, TRMM 3B42RT, CHIRPSv2, CMORPHv1, PERSIANN-CDR e MSWEPv2) em diferentes regiões da América do Sul e recomendaram a necessidade de validar esses dados com médias *in situ* antes de aplicá-los a qualquer estudo. Silva et al. (2019) também compararam as estimativas de precipitação por satélite (TRMM, CHIRPS e GPCP) para diferentes setores da América do Sul, especialmente para o Sul e Sudeste brasileiros. Quando comparados aos dados observados nas estações meteorológicas automáticas, os conjuntos de dados GPCP e TRMM apresentaram melhor correlação com os dados mensais, enquanto o CHIRPS demonstrou melhor ajuste linear aos dados diários. Cassalho et al. (2020) fizeram uma avaliação de estimativas de satélite (MERGE-CPTEC e TMPA 3B42v7 NASA) para a Bacia do Alto Paraíba do Sul. Quando comparado ao TMPA 3B42V7 na escala da bacia hidrográfica, o MERGE-CPTEC tem uma correlação maior com os dados observados. Para a região amazônica, Abe e Novo (2019) compararam os dados de precipitação do TRMM e do MERGE-CPTEC com dados observados *in situ*. Os resultados indicaram que os dados mensais foram os que tiveram melhor similaridade com os dados de estação, mas na escala diária, o MERGE apresentou maior concordância com os dados *in situ* do que o TRMM. Embora as estimativas de precipitação por satélite apresentem alguns vieses, a maioria dos estudos citados menciona que, em decorrência da alta resolução temporal e espacial, aquelas podem ser usadas em substituição aos dados diretamente observados nas estações meteorológicas para estudos sobre a variabilidade da precipitação em diversas escalas. Porém, também se recomenda que sempre que possível, seja realizada uma análise de consistência e validação dos dados estimados.

Os estudos sobre validação de conjuntos de dados de precipitação também têm sido realizados para avaliar tanto produtos decorrentes da interpolação apenas de dados *in situ* como da combinação de produtos de satélite com observações de superfície. Limberger e Silva (2018), ao compararem análises de precipitação do *Climate Research Unit* (CRU), *Global Precipitation Climatology Centre* (GPCC), *National Centers for Environmental Prediction* (NCEP) e *National Center for Atmospheric Research* (NCAR) na região da bacia amazônica brasileira, verificaram que as melhores correlações com os dados observados foram as análises do GPCC. Santos et al. (2017) consideraram a análise do Global *Precipitation Climatology Project* (GPCP) como uma das mais similares aos dados observados para a Amazônia, uma vez que elas consistem na combinação de dados de satélite e estações pluviométricas. Assim, os autores compararam essas análises com dados mensais do GPCC, CPC e ERA-Interim. Nesse estudo, verificou-se que o padrão sazonal de precipitação

ISSN: 2237-8642 (Eletrônica)

se mostrou equivalente nos quatro conjuntos de dados, porém ocorreram diferenças no volume da precipitação sazonal.

No Nordeste brasileiro, Marengo (2007), Santos et al. (2012) e ANA (2015) abordam a variabilidade da recorrência de secas ou inundações, que estão diretamente associadas a anomalias da circulação atmosférica. Para o Semiárido nordestino, essa variabilidade é especialmente grave com a recorrência de situações de secas, pois comprometem o abastecimento rural e urbano, bem como a geração de energia elétrica (MARENGO et al., 2016). Um exemplo é apresentado em Cunha et al. (2019), que estudaram episódios de secas entre 2011 e 2019. Segundo os autores, a seca de 2011-2017 foi a mais severa das últimas décadas. O período de seca entre 2003 e 2013, em especial, apresentou condições severas que fizeram com que muitos municípios tivessem racionamento e decretassem emergência ou estado de calamidade pública (ANA, 2015). Por esses motivos, as secas são objeto de preocupação e, com o cenário de mudanças climáticas, seus impactos podem ser mais graves. Assim, destaca-se a necessidade de estudos hidroclimáticos na região nordestina, o que demanda dados de precipitação.

Próximo à bacia do São Francisco, mas no Estado de Minas Gerais, encontra-se também a bacia do Sapucaí, que apresenta um histórico de inundações. O IGAM (2010) menciona que as cidades de Itajubá, Piranguinho, Santa Rita do Sapucaí e Pouso Alegre sofrem com esses eventos por estarem localizadas na planície de inundação do rio principal. Por exemplo, de acordo com Moraes (2003), o município de Itajubá sofre com esses eventos desde 1821, com destaque para aqueles ocorridos em 1874, 1929, 1945 e 2000 devido à magnitude e danos provocados. Pinheiro (2005) ainda cita que, em 2005, devido à iminência de transbordamento do rio Sapucaí na estação base de estudo localizada em Itajubá, foi instalado um Sistema de Monitoramento de Enchentes, vinculado à Universidade Federal de Itajubá (UNIFEI), a fim de fornecer dados de nível dos rios em tempo real. Em 2015, esse sistema foi desativado, mas em meados de 2018 voltou a funcionar após uma série de modificações em sua estrutura. Os dados desse sistema são disponibilizados em meteorologia.unifei.edu.br (REBOITA et al., 2019). Outros estudos relacionados a eventos de inundações também foram realizados para a bacia do Sapucaí, mediante a aplicação de modelos regressivos para a previsão de cheias (REIS et al., 2016; REBOITA et al., 2019) ou de modelos chuva-vazão (MARTINS et al., 2019), além daqueles que discutem sobre a influência da circulação atmosférica para a ocorrência de eventos extremos de precipitação (CAMPOS et al., 2012; REBOITA et al., 2017).

As bacias do São Francisco e do Sapucaí são integrantes do projeto internacional "Análise da pressão hidrológica com base em Modelos Globais", estabelecido entre Brasil e Suécia (CONFAP-WATER-JPI). Este projeto, além de analisar as características físicas das bacias, também se destina a auxiliar os usuários/tomadores de decisão por meio de indicadores das condições das bacias. A primeira etapa do estudo envolve a caracterização climática das bacias, o que demanda dados de precipitação. Diante ao exposto, o objetivo do estudo é validar dados de precipitação provenientes das análises do MERGE-CPTEC, CPC, GPCC e GPCP, para o período de 1998 a 2018, mediante comparações com dados medidos em estações meteorológicas nas bacias do São Francisco e Sapucaí. Se tais bases apresentarem uma boa representação dos dados medidos *in situ*, poderão ser utilizadas com maior confiança em

ISSN: 2237-8642 (Eletrônica)

estudos climáticos em áreas sem estações meteorológicas. Ressalta-se aqui que é de conhecimento que algumas das estações meteorológicas podem ser assimiladas por tais análises (como as do CPCP e GPCC), o que pode implicar em um bom resultado quando da comparação dessas com as séries medidas *in situ*.

2 MATERIAL E MÉTODOS

2.1 ÁREA DE ESTUDO

2.1.1 BACIA DO SÃO FRANCISCO

A Bacia do São Francisco abrange sete unidades federativas do Brasil (Minas Gerais, Bahia, Pernambuco, Alagoas, Sergipe, Goiás e Distrito Federal, totalizando mais de 500 municípios), de forma que sua área é distribuída entre as regiões Sudeste, Centro-Oeste e, principalmente, Nordeste. A bacia conta com uma área de drenagem de 638.323 km², o que representa 8% do território nacional, sendo dividida em quatro sub-bacias **(Figura 1)**: a) Alto São Francisco, no sudoeste de Minas Gerais, que compreende a nascente do rio na Serra da Canastra até os municípios de Pirapora e Montes Claros, no centro-norte do estado; b) Médio São Francisco, que se estende do município de Pirapora até Remanso, no interior da Bahia, onde está localizado o reservatório de Sobradinho; c) Submédio São Francisco, de Remanso até Paulo Afonso, na Bahia, constituindo-se na divisa natural entre esse estado e Pernambuco, até alcançar o limite com Alagoas; d) Baixo São Francisco, que corresponde à região a partir do município de Paulo Afonso até a foz, localizada entre os estados de Alagoas e Sergipe (MMA, 2006). Cada sub-bacia apresenta características distintas em relação ao clima, solo e vegetação.

ISSN: 2237-8642 (Eletrônica)

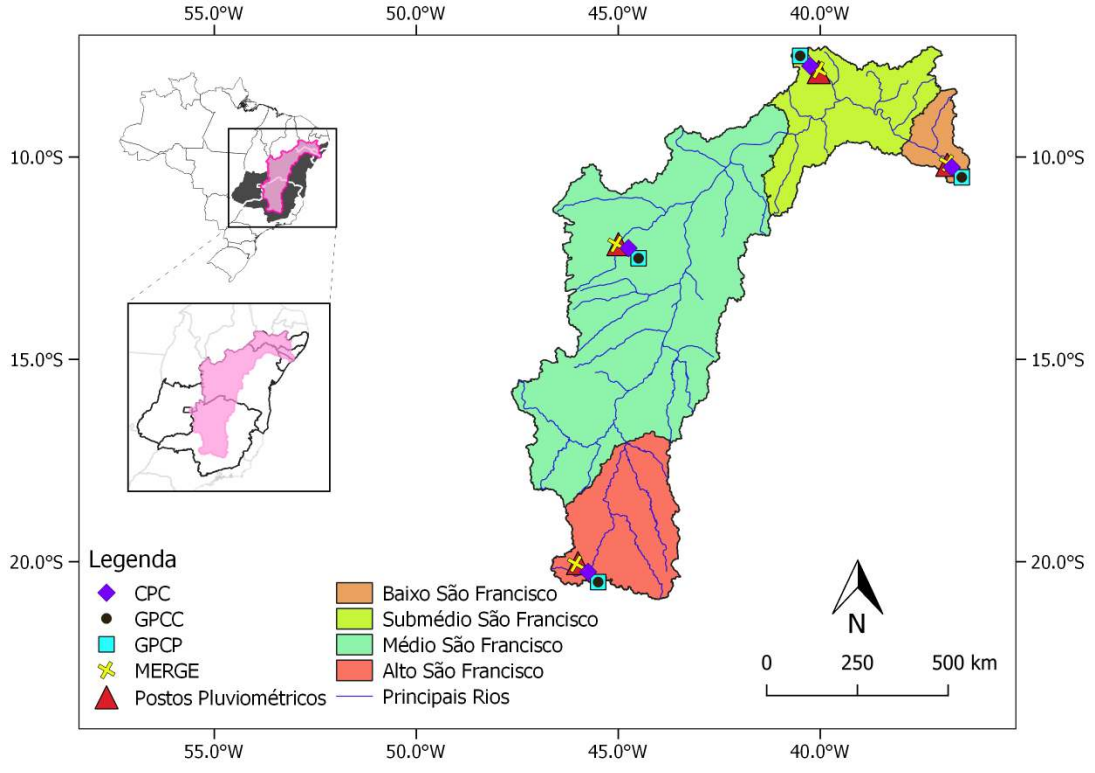


Figura 1 - Delimitação da bacia do São Francisco e localização das estações meteorológicas utilizadas no estudo (triângulos vermelhos), bem como os pontos de grade de diferentes bases de dados de precipitação (losango roxo para CPC, círculo preto para GPCC, retângulo azul para GPCP e cruzeta amarela para o MERGE-CPTEC).

A bacia do São Francisco apresenta grande diversidade ambiental, contemplando fragmentos de diferentes tipos de vegetação: Cerrado, Floresta Atlântica, mata seca, Caatinga, mangues e vegetação litorânea. O Cerrado cobre boa parte da área da bacia, predominando nos trechos Alto e Médio São Francisco, enquanto a Caatinga prevalece no Médio e Submédio São Francisco, onde as condições de clima são mais severas. Já a floresta tropical, cujos fragmentos remanescentes ainda não foram utilizados para agricultura e pastagens, localiza-se no Alto São Francisco. Nas margens do rio São Francisco, na porção oeste do Médio São Francisco, onde a umidade é mais elevada, coexistem a Floresta Estacional Decidual Montana, encontrada em regiões de transição de biomas e caracterizada por ser uma mata densa, com a ocorrência de duas estações climáticas distintas e bem demarcadas e pela perda de mais de 50% de suas folhas na estação seca (IBGE, 2012; LEMES, 2019), e a Caatinga. Por fim, a vegetação litorânea (plantas que crescem nas restingas, nos manguezais, nas dunas e nas praias) existe no trecho Baixo São Francisco. Os tipos de solo na bacia do São Francisco são variados devido às múltiplas origens geológicas, desde o período Pré-Cambriano até o Quaternário. Com isso, há diferentes drenabilidades e texturas, variando entre arenosa e argilosa (MMA, 2006).

Os aspectos climáticos que atuam na região da bacia do São Francisco são controlados por um conjunto de sistemas dinâmicos da atmosfera como: Zona de Convergência Intertropical (ZCIT), Zona de Convergência do Atlântico Sul (ZCAS), Vórtices Ciclônicos de Altos Níveis (VCANs) e, na escala interanual,

ISSN: 2237-8642 (Eletrônica)

a influência dos fenômenos El Niño e La Niña (REBOITA et al., 2010; SANTOS et al., 2011; ASSIS et al., 2015). A precipitação média anual da bacia é de 1.036 mm e a chuva anual pode variar de 350 mm, no Submédio São Francisco (no Semiárido nordestino), até 2.000 mm, no Alto São Francisco. A elevada temperatura média anual, que varia entre 23 e 27 °C, justifica os elevados níveis de evapotranspiração média anual. Estes variam de 1.000 mm (no Alto São Francisco) a 1.550 mm (no trecho Submédio), atingindo o valor de 2.700 mm na parte semiárida, que é mais suscetível à ocorrência de prolongados eventos extremos de seca e estiagens (MMA, 2006).

Segundo Santos et al. (2011) e Assis et al. (2015), o padrão comportamental da duração e intensidade da precipitação na região semiárida do São Francisco sofreu alterações. Por exemplo, até a década de 1980, era mais comum a ocorrência de anomalias positivas de precipitação, mas posteriormente, os anos de estiagens se tornaram mais evidentes, especialmente entre 1990-2010.

2.1.2 BACIA DO SAPUCAÍ

A Bacia do Sapucaí está localizada na região Sudeste, abrangendo os Estados de Minas Gerais e São Paulo, e integra a bacia do Rio Grande. A bacia conta com uma área de drenagem de 9.465 km², sendo dividida em três regiões **(Figura 2)**: a) Alto Sapucaí, compreende a nascente do rio em Campos do Jordão (no estado de São Paulo) até o município de Wenceslau Brás (no interior de Minas Gerais); b) Médio Sapucaí, que se estende de Wenceslau Brás até Pouso Alegre; c) Baixo Sapucaí, de Pouso Alegre até sua foz, onde se localiza a represa de Furnas (IGAM, 2010).

ISSN: 2237-8642 (Eletrônica)

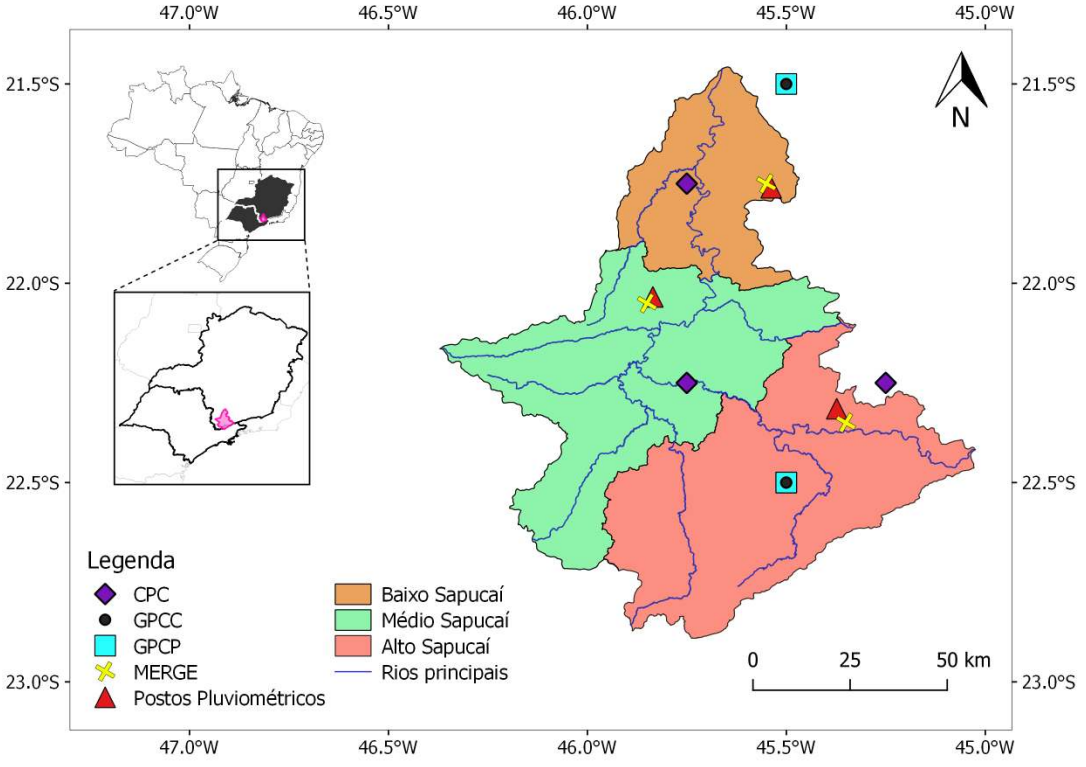


Figura 2 – Idem à **Figura 1**, mas para a bacia do Sapucaí. Na sub-bacia do Médio Sapucaí, os símbolos das análises do GPCC e GPCP não são exibidos pois, para essa região, os pontos de grade mais próximos correspondem aos mesmos utilizados na sub-bacia do Alto Sapucaí (localizados em 22,5ºS-45,5ºW).

A bacia do Sapucaí está inserida nas regiões de domínio da Mata Atlântica e do Cerrado. Na região de Mata Atlântica, há grande diversidade de solos e ecossistemas florestais, enquanto no Cerrado predomina a vegetação rasteira, arbórea e arbustiva. Devido à urbanização e ao uso do solo para pastagens e agriculturas, grande parte da vegetação foi removida. No trecho Alto Sapucaí, há presença e maior conservação da vegetação ciliar em relação aos trechos Médio e Baixo Sapucaí, nos quais há predominância de pastagens e culturas. Os tipos de solo na bacia do Sapucaí são variados e estão relacionados diretamente aos biomas e à topografia local, de modo que há diferentes porosidades e drenabilidades na bacia (IGAM, 2010).

Devido às características topográficas e aos biomas que compõem a bacia do Sapucaí, o clima varia de acordo com o trecho a ser analisado. De acordo com a classificação climática de Köppen, baseada nos registros de precipitação e temperatura, o clima da bacia do Sapucaí pode ser caracterizado como Cwa (clima subtropical com inverno seco e verão quente) e Cwb (clima subtropical com inverno seco e verão moderadamente quente) (REBOITA et al., 2015; REIS et al., 2018; MARTINS et al., 2018). Por exemplo, na nascente, o clima sofre influência da Serra da Mantiqueira, onde a temperatura média anual é de 13,4 °C e a precipitação média anual é superior a 1.500 mm. Os maiores acumulados de precipitação ocorrem no verão, com média de cerca de 900 mm, e os menores no inverno, com média em torno de 190 mm (REIS et al., 2017). Nas demais áreas, a temperatura média anual varia entre 18 e 19 °C e a precipitação é ligeiramente inferior a 1.500 mm.

ISSN: 2237-8642 (Eletrônica)

Pela influência do volume das chuvas associadas a episódios de Zona de Convergência do Atlântico Sul (REBOITA et al., 2015, 2017) e também ao uso do solo relacionado à urbanização sobre a planície de inundação do rio, as regiões de menor declividade da bacia do Sapucaí são aquelas que estão mais suscetíveis à ocorrência de inundações, como os municípios de Itajubá, Santa Rita do Sapucaí e Pouso Alegre (IGAM, 2010).

2.2 DADOS

2.2.1 DADOS MEDIDOS IN SITU

Nesse estudo, selecionou-se uma estação com dados de precipitação diária medidos *in situ* para representar cada sub-bacia do São Francisco e do Sapucaí. As informações de localização dessas estações encontram-se no **Quadro 1** e nas **Figuras 3** e **4**. O ideal seria a obtenção de estações na área central de cada sub-bacia, mas isso nem sempre é possível, uma vez que pode não existir estação de medição ou as séries disponíveis apresentarem inconsistências. Por isso, a **Figura 1** mostra algumas estações nas bordas das sub-bacias. Para a bacia do São Francisco, foram selecionadas informações provenientes de estações meteorológicas convencionais, disponibilizadas pelo INMET no Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP). Para a região da bacia do Sapucaí, foram utilizados dados de precipitação disponibilizados pela ANA no portal HIDROWEB, dado que as séries do INMET são menos extensas e possuem muitas falhas. O período selecionado para o estudo foi de 1998 até 2018. A precipitação diária do INMET que é registrada em um determinado dia corresponde ao período das 12:00 UTC do dia anterior às 12:00 UTC do dia atual (das 09 às 09 h local). Já a ANA, acumula a precipitação das 07 às 07 h local (ANA, 2014). Portanto, há uma diferença de 2 horas em relação ao INMET, conforme apresentado em Back et al. (2012) e Back e Pola (2016).

Quadro 1 – Estações meteorológicas do HIDROWEB e BDMEP. As estações destacadas em vermelho referem-se àquelas utilizadas para preenchimento de falhas.

Bacia	Sub-bacia	Código (OMM)	Localização	Latitude (º)	Longitude (º)
São Francisco	Alto	83582 2045001	Bambuí – MG Bambuí – MG	-20,03000 -20,02110	-46,00000 -45,96610
	Médio	83236 1245004	Barreiras – BA Fazenda Redenção – BA	-12,15000 -12,13472	-45,00000 -45,10420
	Submédio	82753 839009	Ouricuri – PE Jacaré – PE	-7,90000 -8,26280	-40,04000 -39,84720
	Baixo	83097 1036005	Propriá – SE Penedo – AL	-10,19000 -10,28500	-36,86000 -36,55640
Sapucaí	Alto	2245088 2245064	Maria da Fé – MG Delfim Moreira – MG	-22,31470 -22,50970	-45,37310 -45,28670
	Médio	2245089 2245000	Silvianópolis – MG Santa Rita do Sapucaí – MG	-22,03420 -22,25080	-45,83530 -45,70890
	Baixo	2145017	Monsenhor Paulo – MG	-21,76030	-45,53780

ISSN: 2237-8642 (Eletrônica)

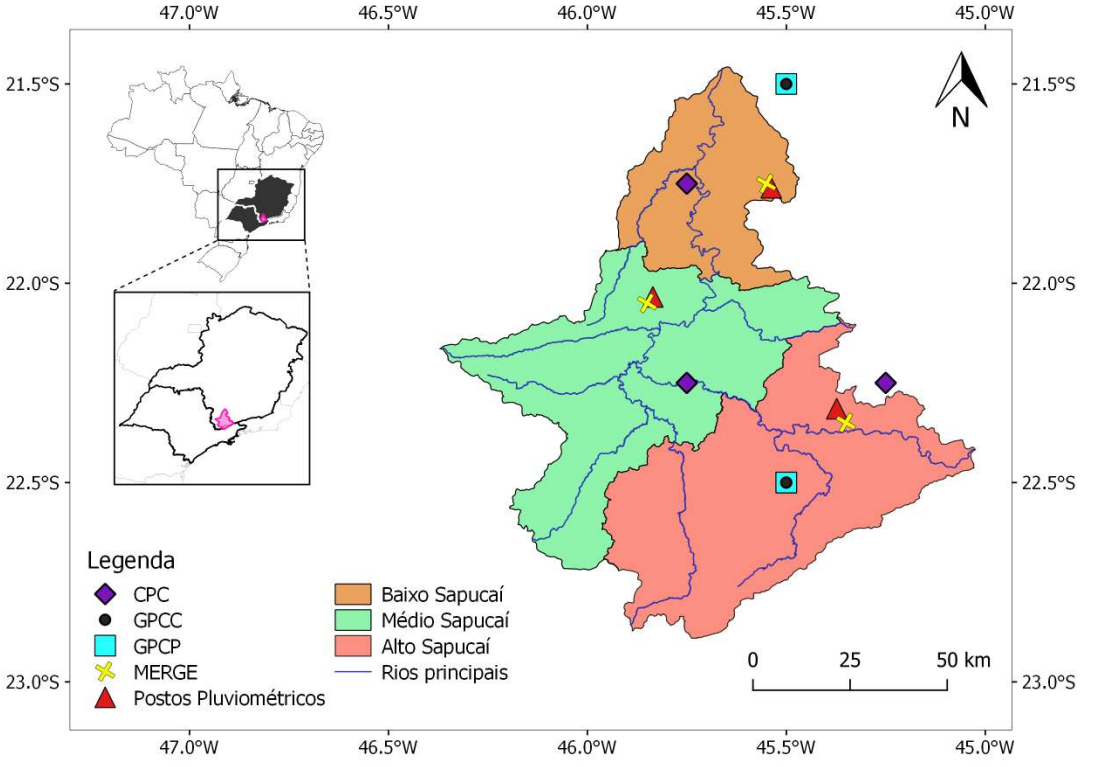


Figura 3 - Delimitação da bacia do São Francisco e localização das Estações meteorológicas do HIDROWEB e BDMEP. As estações representadas por pontos vermelhos referem-se aos postos-base, enquanto aquelas utilizadas para preenchimento de falhas são representadas por pontos verdes.

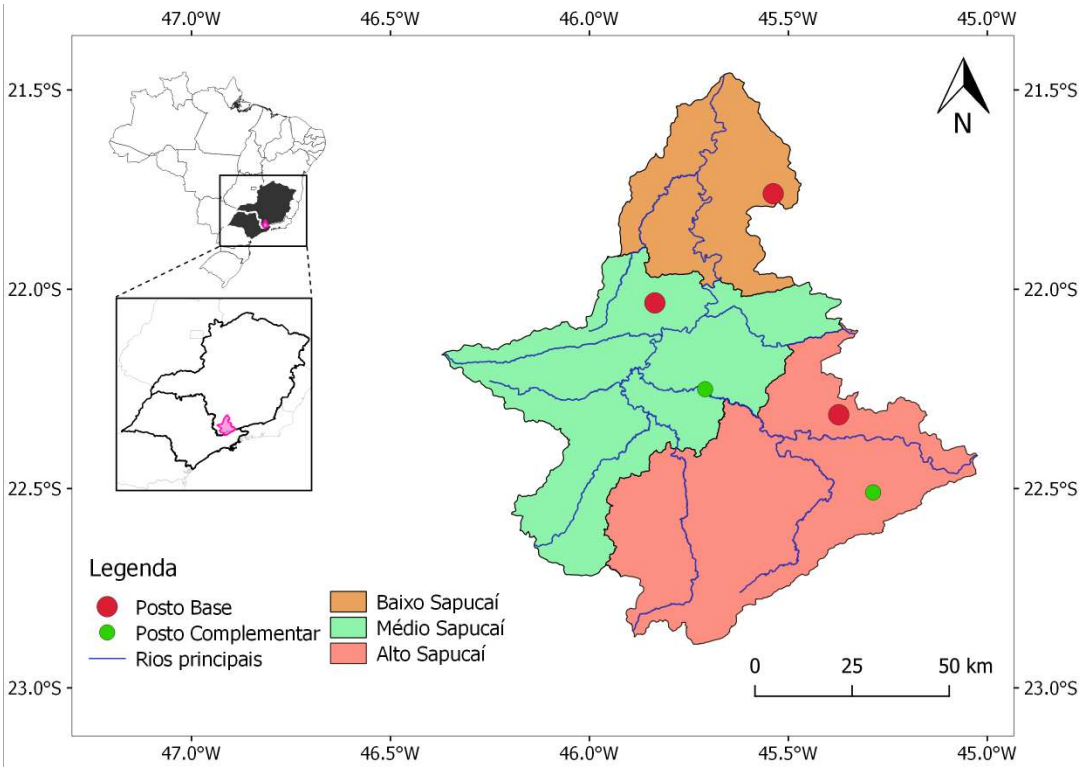


Figura 4 – Idem à **Figura 3**, mas para a bacia do Sapucaí.

2.2.2 ANÁLISES DE PRECIPITAÇÃO

2.2.2.1 Merge-CPTEC

O produto MERGE foi criado pelo Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), do Instituto Nacional de Pesquisas Espaciais (INPE). Esse produto é obtido a partir da combinação entre a precipitação observada e a precipitação estimada por satélite. Inicialmente, para a geração e disponibilização desse produto de previsão, o CPTEC utilizava as estimativas de precipitação do sensor *Tropical Rainfall Measuring Mission* (TRMM), cujos dados eram produzidos e operacionalizados com frequência temporal de 3 horas e resolução horizontal de 0,25°. As saídas do MERGE-CPTEC-TRMM eram disponibilizadas diariamente, tendo como variável a precipitação acumulada em 24 horas, cobrindo a América do Sul (82,8ºW-34ºW e 52,2ºS-12,2ºN) (ROZANTE et al., 2010). O sensor TRMM estava a bordo do satélite lançado em 27 de novembro de 1997 e desativado em 8 de abril de 2015, (NASA, 2015; SERRÃO et al., 2016). Por esse motivo, o CPTEC passou a utilizar as estimativas de precipitação do Global Precipitation Measurement (GPM), lançado em 27 de fevereiro de 2014 (NASA, 2014). Essas estimativas são denominadas de GPM-IMERG e são produzidas com frequência temporal de 30 minutos e resolução espacial de 0,1°. O MERGE-CPTEC-GPM também é disponibilizado diariamente, tendo como variável a precipitação acumulada em 24 horas (das 12:00 às 12:00 UTC), cobrindo a América do Sul e oceanos adjacentes (90°W-26°W e 57°S-06°N). O MERGE-CPTEC-GPM possui duas saídas. Em uma, denominada *Early*, utilizam-se os primeiros dados recebidos e possui atraso de aproximadamente 4 horas desde o recebimento desses dados, obtendo-se a estimativa inicial. A

outra, denominada *Late*, é processada após a chegada de mais dados e tem atraso de aproximadamente 12 horas.

Nesse estudo são validados os dados *Early*, pois esse conjunto está sendo utilizado para dirigir previsões com um modelo hidrológico dentro das atividades do projeto CONFAP-WATER-JPI sendo, portanto, necessária a validação dos dados. As informações do GPM-IMERG estão disponibilizadas no sítio eletrônico <http://ftp.cptec.inpe.br/modelos/tempo/CRNG/MERGE/DAILY/>, cujas subpastas identificam as saídas diárias e horárias desde 2014. Em suma, para este estudo, foram utilizados os dados diários de precipitação do MERGE-CPTEC-GPM (chamado apenas de MERGE-CPTEC ao longo do estudo) no período de abril de 2014 a dezembro de 2018, com resolução espacial de 0,1°.

2.2.2.2 CPC

O conjunto de dados diários de precipitação CPC *Gauge-Based Analysis of Global Daily Precipitation* (CPC-Global) se trata de um produto oriundo do programa CPC *Unified Precipitation Project*, estabelecido pelo *National Oceanic and Atmospheric Administration*/*Climate Prediction Center* (NOAA/CPC). Tal projeto se destina à combinação de todas as fontes de informações disponíveis no CPC, juntamente com a utilização de técnicas de interpolação para a criação de um conjunto unificado de produtos de precipitação com robustez qualitativa e quantitativa (SUN et al., 2018). Nesse programa, dados de precipitação diária (acumulados das 12:00 às 12:00 UTC) de cerca de 30.000 estações de superfície são coletados de diversas fontes, como o sistema global de telecomunicações da Organização Meteorológica Mundial (OMM), redes cooperativas de observação, além de agências meteorológicas nacionais e internacionais. O controle de qualidade é efetuado por meio de comparações com registros históricos e informações independentes provenientes de medições de estações de superfície, observações de radares e satélites, bem como previsões de modelos numéricos (CHEN et al., 2008).

As informações diárias da precipitação do CPC estão disponibilizadas no sítio eletrônico <https://ftp.cpc.ncep.noaa.gov/precip/CPC_UNI_PRCP/GAUGE_GLB/>. Para este estudo, foram utilizados dados diários de precipitação do CPC do período de 1998 a 2018, com resolução espacial de 0,5º.

2.2.2.3 GPCC

O GPCC *Full Data version 2018* é um produto do *Global Precipitation Climatology Centre* (GPCC), programa operado pelo Serviço Meteorológico Alemão. Para a criação desse produto, foram analisadas e interpoladas informações de mais de 79.000 estações de superfície de diferentes modalidades espalhadas pelo globo, cujos dados de totais diários de precipitação foram fornecidos por serviços meteorológicos e hidrológicos nacionais e internacionais, bancos de dados regionais e globais, bem como dados do sistema de telecomunicações da OMM (SCHAMM et al., 2014; SCHNEIDER et al., 2018). Salienta-se que esse conjunto utiliza apenas dados observados em superfície para a construção das análises. Ressalta-se também que, desde abril de 2001, os acumulados diários de precipitação são calculados considerando-se

ISSN: 2237-8642 (Eletrônica)

o intervalo de tempo de 06:50 do dia anterior às 06:50 CET (*Central European Time*) do dia em questão (comunicação pessoal de Drª Annika Jahnke-Bornemann, pesquisadora do *Integrated Climate Data Center* (ICDC)/Universidade de Hamburgo; DWD *Climate Data*, 2020).

As informações de precipitação são disponibilizadas no sítio eletrônico <https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-daily_v2018_doi_download.html>. Para o presente estudo, utilizaram-se dados diários de precipitação com resolução espacial de 1°, para o período de 1998 a 2018.

2.2.2.4 GPCP

O conjunto de dados de precipitação do *Global Precipitation Climatology Project* (GPCP) foi estabelecido pelo *World Climate Research Programme* (WCRP). Neste estudo, utilizou-se o produto GPCP *One-Degree Daily Precipitation Data Set Ver*sion 1.2 (GPCP-1DD), no qual estimativas de precipitação foram computadas por meio de produtos oriundos do canal infravermelho de satélites geoestacionários (HUFFMAN et al., 2001). Estimativas de precipitação baseadas em satélites são calculadas mediante relações empíricas entre a temperatura do topo de nuvens (medida por radiômetros de infravermelho) e taxas de precipitação. Podem também ser estimadas por instrumentos de micro-ondas que medem diretamente o espalhamento de radiação ascendente, bem como a emissão térmica de gotas de chuva e hidrometeoros, associando essas informações à taxa de precipitação (JOYCE et al., 2004; CUI et al., 2016).

Desde o lançamento do TRMM no final de 1997, técnicas e algoritmos de estimativas de precipitação têm sido modificados e aprimorados para combinar observações de diferentes canais dos satélites, a fim de fornecer informações de precipitação em tempo real (CUI et al., 2016). Ajustes e melhorias como a transição de dados do produto *TIROS Operational Vertical Sounder* (TOVS) para o produto *Atmospheric Infrared Sounder* (AIRS) têm auxiliado o GPCP-1DD a prover estimativas confiáveis de precipitação. Entretanto, as técnicas baseadas em satélites são associadas a erros resultantes da amostragem de dados de diferentes instrumentos e satélites, estimativas imprecisas de algoritmos, bem como de características dos próprios instrumentos. Assim, a acurácia das estimativas do GPCP-1DD deve ser determinada mediante comparações cuidadosas com conjuntos de dados com alta resolução espacial e temporal (CUI et al., 2016).

Como os dados de infravermelho são fornecidos em intervalos de três horas (00, 03, …, 21 UTC), o conjunto de dados de precipitação é melhor representado quando considerado o período que começa às 22:30 UTC do dia anterior até as 22:30 UTC do dia em questão (comunicação pessoal de Dr. Howard J. Diamond, administrador do *Climate Science Program*/NOAA*'s Air Resources Laboratory*; HUFFMAN e BOLVIN, 2013).

As informações de precipitação são disponibilizadas no sítio eletrônico <https://www1.ncdc.noaa.gov/pub/data/gpcp/daily-v1.2/data/>. Para este estudo, foram utilizados dados diários de precipitação com resolução espacial de 1º, do período de 1998 a 2015.

O **Quadro 2** apresenta uma síntese de informações das análises de precipitação utilizadas no presente estudo.

Quadro 2 – Análises de precipitação utilizadas no estudo.

Produto	Período	Resolução Horizontal	Link	Referência
MERGE-CPTEC	2014-2018	0,1º	http://ftp.cptec.inpe.br/modelos/tempo/CRNG/MERGE/DAILY/	Rozante et al. (2010)
CPC Gauge-Based Analysis of Global Daily Precipitation	1998-2018	0,5º	https://ftp.cpc.ncep.noaa.gov/precip/CPC_UNI_PRCP/GAUGE_GLB/	Chen et al. (2008)
GPCC Full Data Daily Version 2018	1998-2018	1º	https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-daily_v2018_doi_download.html	Schamm et al. (2014)
GPCP One-Degree Daily Precipitation Data Set Version 1.2	1998-2015	1º	https://www1.ncdc.noaa.gov/pub/data/gpcp/daily-v1.2/data/	Huffmann et al. (2001)

2.3 MÉTODOS

Das bases de dados mencionadas, foram extraídas séries temporais dos pontos de grade mais próximos às estações. A representação dos pontos está nas **Figuras 1** e **2**.

No caso das séries temporais de dados diários do INMET e ANA, o total mensal de precipitação, corresponde ao acumulado de todos os dias do mês. Portanto, seguindo a WMO – N^{o} 100 (2018), não pode haver nenhum dia faltante para se ter o total mensal. Como alguns meses apresentaram falhas de dados diários, adotou-se como metodologia no presente estudo preencher as falhas diárias se um dado mês possuía até 5 dias de dados faltantes, mais do que isso, o mês era excluído do estudo. As normas da WMO – N^{o} 100 (2018) para o cálculo de normais climatológicas recomenda a exclusão do mês já na falta de um dia faltante. Porém, aqui como não são calculadas normais climatológicas e para não serem descartados muitos meses, a metodologia foi mais flexível. Assim, para os meses que apresentaram cinco ou menos dados diários faltantes, a falha foi substituída por dados de precipitação da estação mais próxima (REBOITA et al., 2015). No caso de falha nas estações do INMET, foram utilizados dados da ANA (esse método foi usado mesmo sabendo-se que o acumulado diário da ANA é realizado com duas horas de diferença em relação ao INMET, como descrito na seção anterior) que estão destacados em vermelho no **Quadro 1**. Para a sub-bacia do Baixo Sapucaí, não foi possível aplicar a metodologia de preenchimento de falhas, pois não havia estações próximas. Dessa forma, os meses com falhas de dados diários foram descartados.

A fim de comparar a precipitação observada nas estações do INMET e ANA com as estimadas pelo CPC (dado *in situ*), GPCC (dado *in situ*), GPCP (estimativa de satélites) e MERGE-CPTEC (combinação de estimativa do GPM e

dado *in situ*), foram utilizadas diferentes medidas estatísticas, como mostra o **Quadro 3**. Essas medidas foram calculadas em três etapas:

1) precipitação mensal no período de 1998 a 2018: para essa seção, foram geradas séries temporais de acumulados mensais para o período de 1998 a 2018, exceto para o GPCP que só possui dados até 2015. Nessa análise, o MERGE-CPTEC não foi incluído, pois sua série compreende o período de 2014 a 2018, que foi anomalamente seco em partes da região Nordeste do Brasil (MARENGO et al., 2018; CUNHA et al., 2019). Portanto, se fosse incluído nas estatísticas, haveria subestimativa da precipitação (essa análise foi realizada previamente e um exemplo será mostrado na seção Resultados – **Figura 9**);

2) precipitação mensal no período de 2014 a 2018: essa análise é para permitir a inclusão do MERGE-CPTEC;

3) precipitação diária no período de 2014 a 2018: visto que os acumulados diários de todas as análises não foram realizados no mesmo intervalo temporal, não foi possível a comparação com os dados de estações, uma vez que os acumulados são das 12:00 às 12:00 UTC nas estações do INMET e das 10:00 às 10:00 UTC nas estações da ANA. Portanto, nessa seção as estatísticas são calculadas apenas considerando-se as análises do CPC e MERGE-CPTEC e as estações do INMET e ANA.

Quadro 3 – Medidas estatísticas adotadas para comparação entre a precipitação observada e a das análises.

Estatística	**Descrição**	**Equação**
Erro Absoluto Médio, em inglês *Mean Absolute Error* (MAE)	Mede a variabilidade no erro de um conjunto de dados em relação a outro por meio da magnitude das diferenças entre eles, não levando em consideração a subestimação ou superestimação do erro (MONTGOMERY et al., 2008). De acordo com Hallak e Filho (2011), o melhor valor de MAE é aquele próximo de zero, indicando melhor representação de um conjunto de dados.	$MAE = \frac{1}{n}\sum_{t=1}^{n}\lvert\hat{y}_t - y_t\rvert$ onde: n: Número de amostras t: Período y_t: Valor observado na estação meteorológica no período t $\hat{y}_t$: Valor dos dados obtidos de análises em pontos de grade no período t
Raiz do Erro Médio Quadrático, em inglês *Root Mean Square Error* (RMSE)	Mede a variabilidade no erro de duas séries temporais. Para isso, somam-se os erros obtidos a partir do quadrado da diferença entre o valor observado nas estações meteorológicas e os dados de análises em pontos de grade (MONTGOMERY et al., 2008). De acordo com Hallak e Filho (2011), esse critério estatístico é comumente usado para expressar a acurácia dos resultados numéricos com a vantagem de que esse parâmetro apresenta valores de erro na mesma	$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}[\hat{y}_t - y_t]^2}$ onde: n: Número de amostras t: Período y_t: Valor observado na estação meteorológica no período t

	dimensão da variável analisada.	$\hat{y}_t$: Valor dos dados obtidos de análises em pontos de grade no período t
Coeficiente de Correlação de Pearson (r)	Mede o grau de correlação entre as variáveis dos conjuntos de dados de análises em pontos de grade e os dados observados em estações meteorológicas na série temporal, assumindo valores dentro do intervalo de -1 a 1. Quando r=1, a correlação entre as variáveis é perfeita, isto é, ambas as séries crescem ou diminuem ao mesmo tempo; r=-1 indica correlação perfeita, mas negativa, de modo que as variáveis são inversamente proporcionais; e r=0 indica que as variáveis são linearmente independentes (MONTGOMERY et al., 2008). De acordo com Hinkle et al. (2003), os intervalos de correlação podem ser interpretados da seguinte forma: intervalo de 0,9 a 1,0 (-0,9 a -1,0) indica correlação positiva (negativa) muito alta; de 0,7 a 0,9 (-0,7 a -0,9) representa correlação positiva (negativa) alta; de 0,5 a 0,7 (-0,5 a -0,7) sugere correlação positiva (negativa) moderada; de 0,3 a 0,5 (-0,3 a -0,5) indica correlação positiva (negativa) fraca; de 0,0 a 0,3 (0,0 a -0,3) representa correlação positiva (negativa) muito fraca.	$$r = \frac{\sum_{t=1}^{n}[(y_t - \bar{y})(\hat{y}_t - \bar{\hat{y}})]^2}{\sqrt{\sum_{t=1}^{n}[y_t - \bar{y}]^2}\sqrt{\sum_{t=1}^{n}[\hat{y}_t - \bar{\hat{y}}]^2}}$$ onde: n: Número de amostras t: Período y_t: Valor observado na estação meteorológica no período t $\bar{y}$: Média dos valores observados $\hat{y}_t$: Valor dos dados obtidos de análises em pontos de grade no período t $\bar{\hat{y}}$: Média dos valores previstos
Coeficiente de Determinação (R^2)	Avalia o ajuste dos dados em pontos de grade em relação ao conjunto de dados observados nas estações meteorológicas. Isto é, a proporção da variação dos dados em pontos de grade pode ser explicada pela variação total dos dados observados (MONTGOMERY et al., 2008). Morettin e Bussab (2010) complementam que esse critério varia entre 0 e 1, de modo que quanto maior o valor de R^2, melhor a qualidade de ajuste do conjunto de dados de análise aos valores observados nas estações meteorológicas.	$$R^2 = \frac{[\sum_{t=1}^{n}(y_t - \bar{y})(\hat{y}_t - \bar{\hat{y}})]^2}{\sum_{t=1}^{n}[y_t - \bar{y}]^2 \sum_{t=1}^{n}[\hat{y}_t - \bar{\hat{y}}]^2}$$ onde: n: Número de amostras t: Período y_t: Valor observado na estação meteorológica no período t $\bar{y}$: Média dos valores observados $\hat{y}_t$: Valor dos dados obtidos de análises em pontos de grade no período t $\bar{\hat{y}}$: Média dos valores previstos
Coeficiente de Eficiência de Nash-Sutcliffe, em inglês *Nash-Sutcliffe Efficiency* (NSE)	Compara a similaridade entre dois conjuntos de dados. O NSE pode variar de -∞ a 1 (NASH e SUTCLIFFE, 1970). Moriasi et al. (2007) apresentaram a seguinte classificação para esse critério estatístico: intervalo de 0,75 a 1,0 indica muito bom desempenho; de 0,65 a 0,75 representa bom	$$NSE = 1 - \frac{\sum_{t=1}^{n}[\hat{y}_t - y_t]^2}{\sum_{t=1}^{n}[y_t - \bar{y}]^2}$$ onde: n: Número de amostras t: Período

	desempenho; de 0,50 a 0,65 indica desempenho satisfatório; e menor que 0,5 indica desempenho insatisfatório.	y_t: Valor observado na estação meteorológica no período t $\bar{y}$: Média dos valores observados na estação meteorológica $\hat{y}_t$: Valor dos dados obtidos de análises em pontos de grade no período t
Índice de Willmott (*d*)	Descreve a precisão das análises em pontos de grade em relação aos dados observados em estações meteorológicas. Esse índice foi proposto por Willmott (1981) e sofreu modificações ao longo do tempo. Sua versão refinada apresenta melhorias em relação às propostas anteriores e flexibilidade de aplicação (WILLMOTT et al., 2012). Esse índice pode variar de 0 a 1, de modo que valores de *d* próximos a 0 indicam discordância entre valores dos dados em pontos de grade e observados, enquanto valores de *d* próximos a 1 indicam melhor ajuste da análise.	$$d = \begin{cases} 1 - \frac{\sum_{t=1}^{n}\lvert\hat{y}_t - y_t\rvert}{2\sum_{t=1}^{n}\lvert y_t - \bar{y}\rvert}, \\ quando \sum_{t=1}^{n}\lvert\hat{y}_t - y_t\rvert \le 2\sum_{t=1}^{n}\lvert y_t - \bar{y}\rvert \\ \frac{2\sum_{t=1}^{n}\lvert y_t - \bar{y}\rvert}{\sum_{t=1}^{n}\lvert\hat{y}_t - y_t\rvert} - 1, \\ quando \sum_{t=1}^{n}\lvert\hat{y}_t - y_t\rvert > 2\sum_{t=1}^{n}\lvert y_t - \bar{y}\rvert \end{cases}$$ onde: n: Número de amostras t: Período y_t: Valor observado na estação meteorológica no período t $\bar{y}$: Média dos valores observados na estação meteorológica $\hat{y}_t$: Valor dos dados obtidos de análises em pontos de grade no período t
Erro Médio, em inglês *Mean Error* (ME), também chamado de *bias*	Estima o valor esperado do desvio entre dois conjuntos de dados (MONTGOMERY et al., 2008), medindo a tendência das análises em pontos de grade em superestimar ou subestimar os dados observados em estações meteorológicas. Segundo Hallak e Filho (2011), os erros individuais positivos e negativos de mesma magnitude podem se cancelar na somatória do cálculo do viés, levando à subestimação do erro das simulações.	$$ME = \frac{1}{n}\sum_{t=1}^{n}(\hat{y}_t - y_t)$$ onde: n: Número de amostras t: Período y_t: Valor observado na estação meteorológica no período t $\hat{y}_t$: Valor dos dados obtidos de análises em pontos de grade no período t

3. RESULTADOS E DISCUSSÃO

3.1 COMPARAÇÃO MENSAL 1998-2018

Nessa seção, os acumulados mensais de precipitação do CPC, GPCP e do GPCC são validados mediante apresentação gráfica e parâmetros estatísticos calculados em relação aos dados observados *in situ* no período de 1998 a 2018 (**Figuras 5** e **6**). Todas as estatísticas são calculadas com base nas séries

ISSN: 2237-8642 (Eletrônica)

temporais originais, isto é, sem remoção do ciclo anual e de tendências, já que não se tem como finalidade uma análise de anomalias.

De forma geral, as análises (dados em pontos de grade) de precipitação conseguem reproduzir o padrão das séries temporais da precipitação observada em ambas as bacias (**Figuras 5** e **6**). Na Bacia do São Francisco (**Figura 5**), os mínimos de precipitação ocorrem no inverno, exceto no caso da sub-bacia Baixo São Francisco (**Figura 5d**), em que o período seco ocorre na primavera. Na bacia do Sapucaí, a precipitação também apresenta os menores totais mensais nos meses de inverno (**Figura 6**).

As diferenças entre as análises e os dados observados de precipitação são melhor reveladas por meio de medidas estatísticas. Nesse sentido, a **Tabela 1** mostra o MAE, RMSE, r, R^2, NSE, *d* e *bias* para cada sub-bacia e análise. Além disso, os melhores resultados foram destacados em azul e os piores em vermelho. Para a maioria das sub-bacias do São Francisco (Médio, Submédio e Baixo) e do Sapucaí (Alto e Baixo), o MAE indica que o CPC apresenta erros (11,23, 19,28, 18,23, 26,99 e 26,16, respectivamente) inferiores àqueles calculados para as outras análises. Já para os casos do Alto São Francisco e Médio Sapucaí é o GPCC que se destaca, apresentando os menores erros (23,38 e 33,36, respectivamente), mesmo essa análise estando distante da estação do Médio Sapucaí. Em geral, o GPCP possui recorrentemente os maiores MAEs, exceto no Submédio São Francisco, em que o GPCC é a análise com pior desempenho. Quando dois conjuntos de dados são muito similares, o MAE tende a zero (Hallak e Filho, 2011), mas aqui os valores obtidos foram maiores do que zero, indicando presença de diferenças entre as séries comparadas.

O RMSE é uma medida estatística que se diferencia do MAE por amplificar a relevância das maiores diferenças entre os dados estimados e observados. Apesar dessa distinção, observa-se que os resultados numéricos do RMSE estão de acordo com o que foi obtido para o MAE. O RMSE indica que o CPC possui os menores erros para o Médio e Baixo São Francisco e para o Alto e Baixo Sapucaí (19,76, 30,02, 42,37 e 40,67, respectivamente), enquanto apresenta pior desempenho para o caso do Submédio São Francisco (40,35) e Médio Sapucaí (60,59). Além disso, observa-se que o GPCP apresenta os maiores valores de RMSE para o Alto, Médio e Baixo São Francisco (50,86, 32,86 e 64,06, respectivamente), enquanto o GPCC para o Alto e Baixo Sapucaí (44,26 e 48,64, respectivamente). Similarmente ao caso do MAE, os valores obtidos para a estatística RMSE são considerados elevados em todos os casos quando se usa como parâmetro de comparação o valor ótimo definido por Hallak e Filho (2011), que é RMSE igual a zero.

Quando se trata da correlação entre as análises e as observações, verifica-se que o CPC, o GPCC e o GPCP representam bem o aumento ou diminuição da precipitação mensal acumulada, e que as análises estão sempre em concordância com o comportamento dos valores medidos nas estações. De fato, quando é aplicada a classificação definida por Hinkle et al. (2003) nos resultados numéricos, observam-se correlações muito altas e superiores a 0,9 para todas as análises em cinco sub-bacias (Alto, Médio e Baixo São Francisco e Alto e Baixo Sapucaí), com exceção do GPCP no Baixo São Francisco, que apresenta correlação moderada de 0,66. Já no caso das outras duas sub-bacias (Submédio São Francisco e Médio Sapucaí), os valores da estatística r ficam acima de 0,8, sendo, portanto, classificadas como tendo fortes correlações.

ISSN: 2237-8642 (Eletrônica)

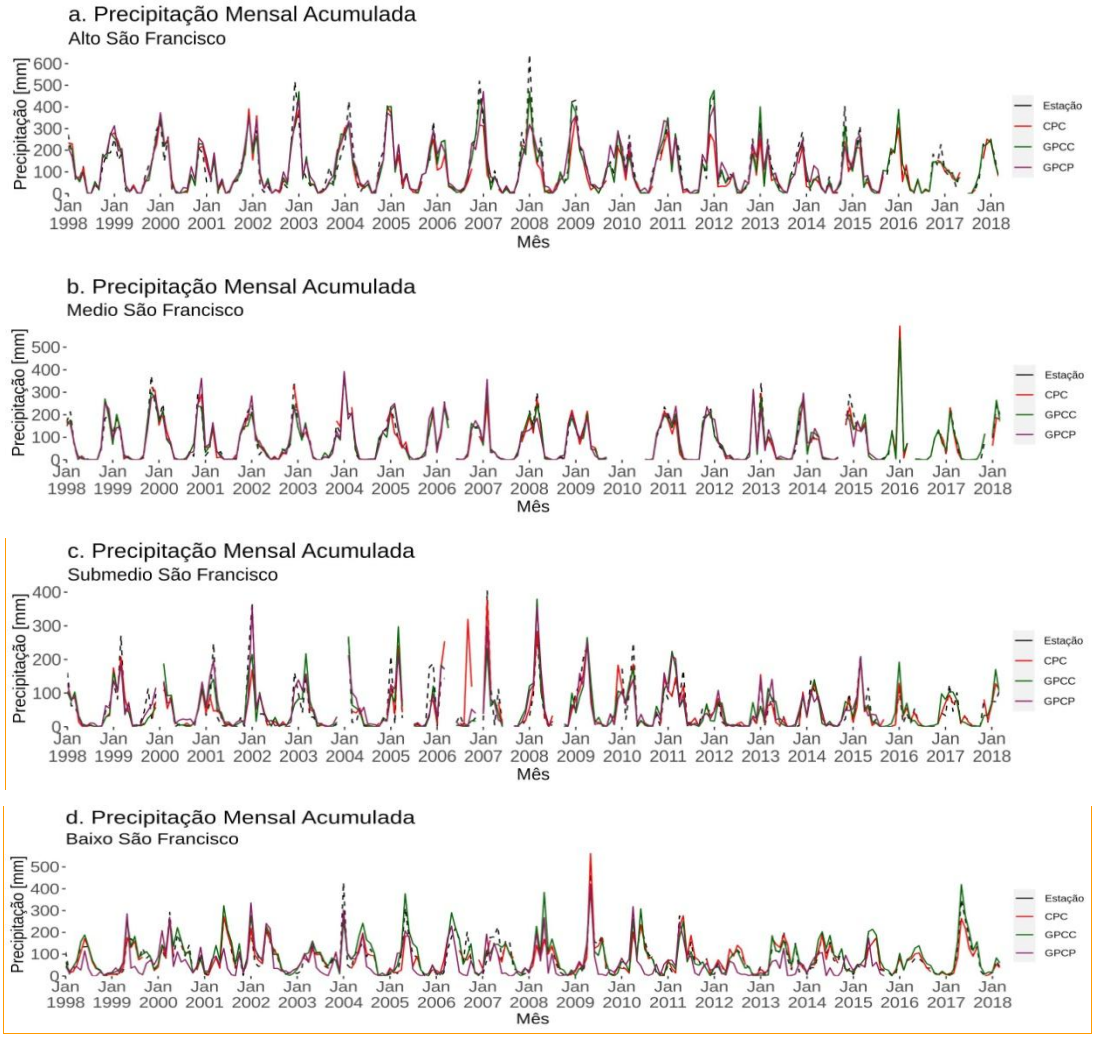


Figura 5 – Séries históricas de precipitação mensal acumulada das estações meteorológicas e das análises (CPC, GPCC e GPCP) entre os anos de 1998 e 2018 na bacia do São Francisco.

ISSN: 2237-8642 (Eletrônica)

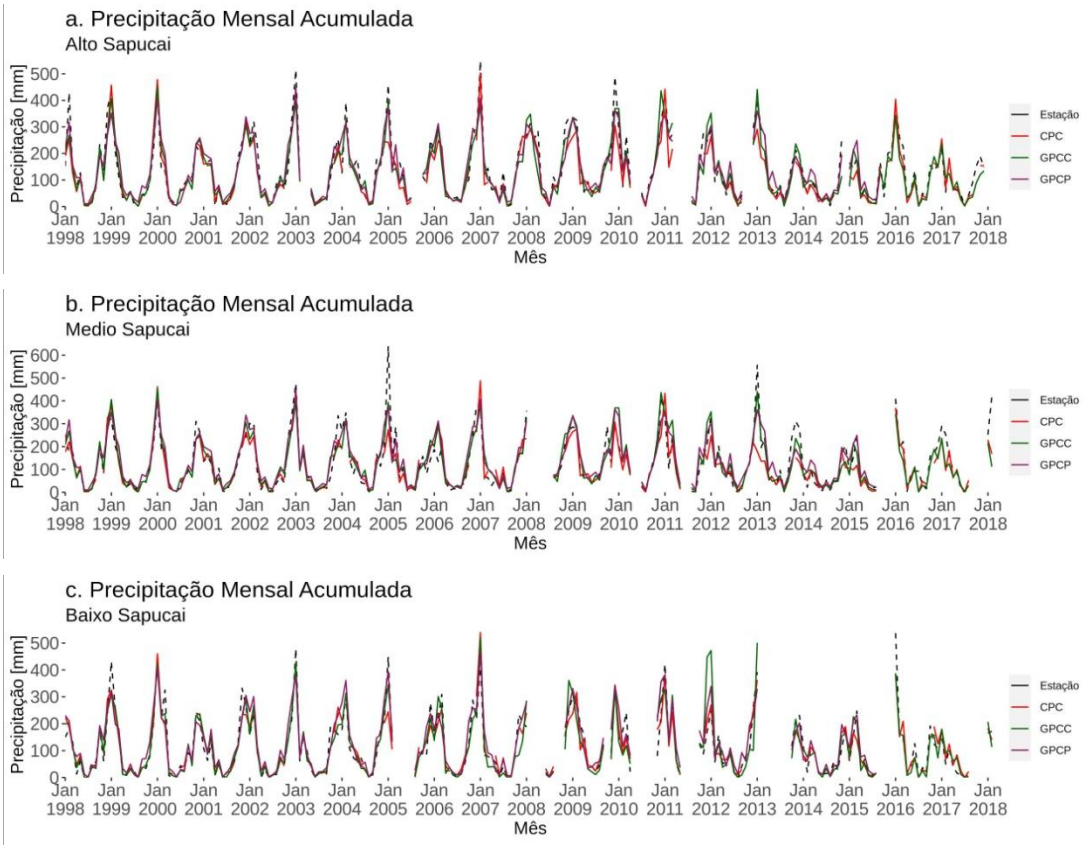


Figura 6 – Idem à **Figura 5**, mas para a bacia do Sapucaí.

A qualidade do ajuste de cada uma das análises é avaliada por meio do R^2. Esse coeficiente indica qual é a proporção em que a quantidade de variância de um conjunto de dados pode explicar a de um outro conjunto. Segundo Morettin e Bussab (2010), quanto mais próximo de 100% a estatística R^2 estiver, melhor é a qualidade do ajuste. Os resultados obtidos indicam que o CPC apresenta a melhor capacidade explicativa para os dados observados do Médio e Baixo São Francisco e do Alto e Baixo Sapucaí (96,2%, 85,4%, 90,3% e 86,9%, respectivamente). Por outro lado, o GPCP apresenta os menores coeficientes para o caso do Alto, Médio e Baixo São Francisco (83,1%, 87,6% e 44,3%, respectivamente), com destaque para essa última sub-bacia em que o valor da estatística R^2 é reduzido.

O NSE é uma estatística capaz de indicar quantitativamente a proximidade entre os valores estimados e observados. Se for aplicada a classificação definida por Moriasi et al. (2007) para o NSE, observa-se que de forma unânime todas as análises, nas sub-bacias do Alto e Médio São Francisco e Alto e Baixo Sapucaí, apresentam muito bom desempenho na representação dos dados observados, enquanto no caso das outras sub-bacias há predominância de bom desempenho. O CPC demonstra o melhor desempenho para o Médio e Baixo São Francisco e Alto e Baixo Sapucaí (0,96, 0,85, 0,87 e 0,86, respectivamente), enquanto o GPCP apresenta o pior desempenho para o Alto, Médio e Baixo São Francisco (0,82, 0,87 e 0,30, respectivamente), havendo novamente uma indicação clara de queda elevada de acurácia no caso do Baixo São Francisco, em que o GPCP apresenta desempenho insatisfatório.

Além do NSE, foi utilizada uma segunda estatística para a análise da concordância entre os dados em pontos de grade e observados, o *d*. Os resultados obtidos para esse índice concordam com aqueles obtidos para o NSE. O CPC continua apresentando os melhores resultados, destacando-se em todas as sub-bacias, com exceção do Alto São Francisco em que o GPCC se sobressai com 0,87. Inversamente, o GPCP apresenta os piores resultados para o Alto, Médio e Baixo São Francisco e Baixo Sapucaí (0,83, 0,87, 0,61 e 0,80, respectivamente). Apesar da ocorrência de algumas exceções, as análises apresentaram índices superiores a 0,8 e próximos ao ideal definido por Willmott et al. (2012), que é de *d* igual a 1.

Tabela 1 – Estatísticas calculadas para avaliação do desempenho de cada análise (CPC, GPCC e GPCP) de precipitação mensal acumulada com relação às medições feitas nos postos pluviométricos das sub-bacias consideradas, no período de 1998 a 2018. Resultados em azul (vermelho) indicam melhor (pior) desempenho.

Bacia	Sub-bacia	Análise	MAE	RMSE	r	R^2	NSE	*d*	*bias*
São Francisco	Alto	CPC	28,94	48,92	0,92	0,859	0,84	0,85	-12,72
		GPCC	23,38	37,45	0,95	0,903	0,90	0,87	-2,00
		GPCP	33,07	50,86	0,91	0,831	0,82	0,83	4,58
	Médio	CPC	11,23	19,76	0,98	0,962	0,96	0,93	-2,47
		GPCC	13,33	22,11	0,97	0,949	0,94	0,91	0,23
		GPCP	19,45	32,86	0,93	0,876	0,87	0,87	1,79
	Submédio	CPC	19,28	40,35	0,83	0,689	0,68	0,82	-0,77
		GPCC	25,07	39,42	0,83	0,693	0,68	0,76	0,83
		GPCP	24,06	37,14	0,86	0,749	0,74	0,78	2,76
	Baixo	CPC	18,23	30,02	0,92	0,854	0,85	0,84	-0,40
		GPCC	25,61	37,05	0,90	0,813	0,77	0,78	11,25
		GPCP	46,30	64,06	0,66	0,443	0,30	0,61	-20,76
Sapucaí	Alto	CPC	26,99	42,37	0,95	0,903	0,87	0,86	-18,00
		GPCC	30,48	44,26	0,92	0,860	0,85	0,83	-9,18
		GPCP	30,95	42,79	0,93	0,874	0,86	0,83	0,28
	Médio	CPC	33,64	60,59	0,86	0,752	0,73	0,82	-12,22
		GPCC	33,36	53,01	0,88	0,790	0,79	0,81	0,16
		GPCP	34,56	50,85	0,90	0,819	0,80	0,81	13,50
	Baixo	CPC	26,16	40,67	0,93	0,869	0,86	0,85	0,46
		GPCC	29,58	48,64	0,90	0,814	0,80	0,82	-1,85
		GPCP	33,81	45,42	0,91	0,844	0,82	0,80	13,71

Legendas: MAE – *Mean Absolute Error* (Erro Absoluto Médio); RMSE – *Root Mean Square Error* (Raiz do Erro Médio Quadrático); r – Coeficiente de Correlação de Pearson; R^2 - Coeficiente de Determinação; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; *d* – Índice de Willmott; *bias* – viés.

A última estatística apresentada na **Tabela 1** é o *bias*. Quando seu valor é diferente de zero, há indicação de tendência das análises subestimarem (*bias* com sinal negativo) ou superestimarem (*bias* com sinal positivo) os valores de precipitação mensal acumulada. Portanto, tal estatística é utilizada. Quando o *bias* é igual ou muito próximo a zero, não é possível afirmar que a série histórica de dados da análise não possui viés. De fato, em trechos distintos de uma

ISSN: 2237-8642 (Eletrônica)

mesma série de dados podem ocorrer subestimação e superestimação de precipitação, que acabam não sendo detectadas pela estatística *bias* devido ao cancelamento matemático entre tais vieses de sinais opostos. Em função disso, nota-se que quando o *bias* é igual a zero não é possível chegar a uma conclusão a respeito da existência de viés. Dados os fatos apresentados, Hallak e Filho (2011) indicam a aplicação de tal estatística não como medida de acurácia, mas sim como um indicador da existência de tendência à superestimação ou subestimação. Os resultados revelam que para o CPC não existe um padrão único aparente de subestimação ou superestimação para todas as sub-bacias. De fato, observam-se valores com módulos pequenos e sinais alternados, com exceção de três casos (Alto São Francisco, Alto e Médio Sapucaí) que apresentam *bias* expressivo (-12,72, -18,00 e -12,22, respectivamente). Quando se trata do GPCC, há ocorrência de *bias* com valor positivo no Baixo São Francisco (11,25), negativo no Alto Sapucaí (-9,18) e inconclusivo para o restante das sub-bacias. Por fim, o GPCP apresenta valores positivos de *bias* para o Alto São Francisco e Médio e Baixo Sapucaí (4,58, 13,50 e 13,71, respectivamente), negativo para o Baixo São Francisco (-20,76) e inconclusivo para as demais sub-bacias. É importante destacar que, de forma geral, tais valores elevados de *bias* decorrem do fato de que essa estatística foi calculada utilizando-se a precipitação mensal acumulada e não a média diária de cada mês. Portanto, os desvios verificados correspondem aos vieses acumulados ao longo de vários dias.

Com o propósito de mostrar de forma mais detalhada o ciclo anual de precipitação nas sub-bacias, foi calculada a média mensal climatológica com base nos acumulados mensais. Exceto a sub-bacia do Baixo São Francisco, todas as demais mostram o verão como o período mais úmido do ano e inverno o mais seco. Esse padrão responde ao sistema de monção da América do Sul (REBOITA et al., 2010; MARENGO et al., 2012; ASHFAQ et al., 2020). Já na sub-bacia Baixo São Francisco, os máximos ocorrem nos meses de abril a julho (meados do outono-inverno) devido à influência dos ventos do Anticiclone Subtropical do Atlântico Sul (ASAS) que estão mais intensos nesta época do ano nas cercanias do litoral nordestino (REBOITA et al., 2016). As análises reproduzem bem a fase do ciclo anual registrado nas estações. As **Figuras 7** e **8** confirmam isso visualmente e também reforçam outros dois resultados obtidos por meio das estatísticas: a) o CPC apresenta o melhor desempenho para a maior parte das sub-bacias, exceto em três sub-bacias em que subestima a precipitação (Alto São Francisco e Alto e Médio Sapucaí; ver *bias* na **Tabela 1** e **Figuras 7** e **8**); e b) o GPCP sofre uma queda expressiva de desempenho na sub-bacia do Baixo São Francisco.

As **Figuras 7** e **8** também mostram o padrão espacial da precipitação média anual com base nos dados do GPCC nas duas bacias em estudo. Ressalta-se que não é objetivo do estudo comparar a distribuição espacial da precipitação anual dos diferentes conjuntos de dados. Por isso, só um conjunto de dados é apresentado.

ISSN: 2237-8642 (Eletrônica)

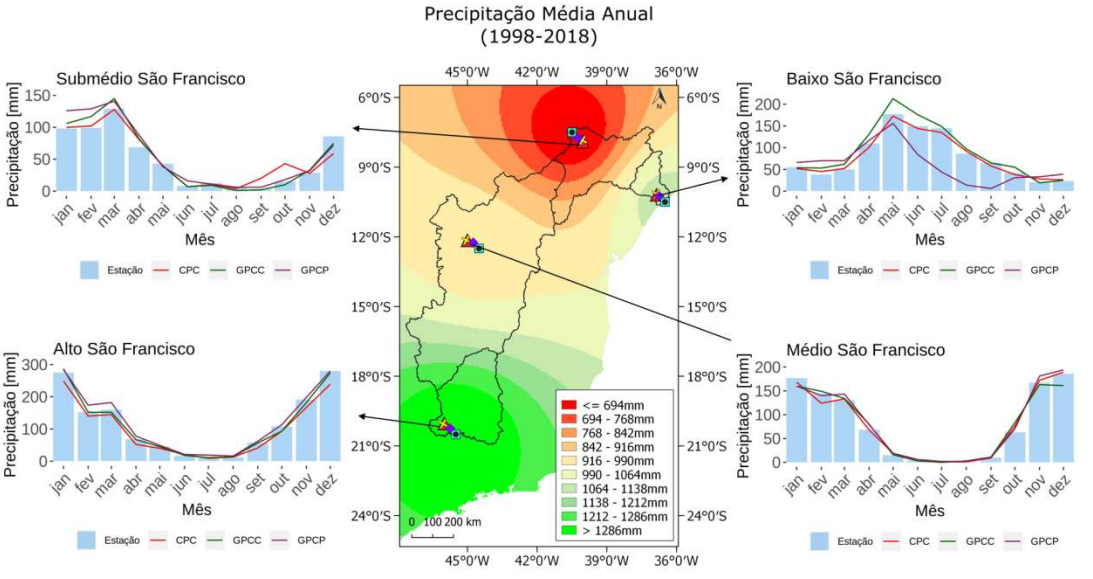


Figura 7 – Mapa com a precipitação média anual da análise do GPCC (1998-2018) e gráficos com o ciclo anual da precipitação observada (barras) e das análises (linhas: CPC, GPCC e GPCP), entre os anos de 1998 e 2018 na Bacia do São Francisco.

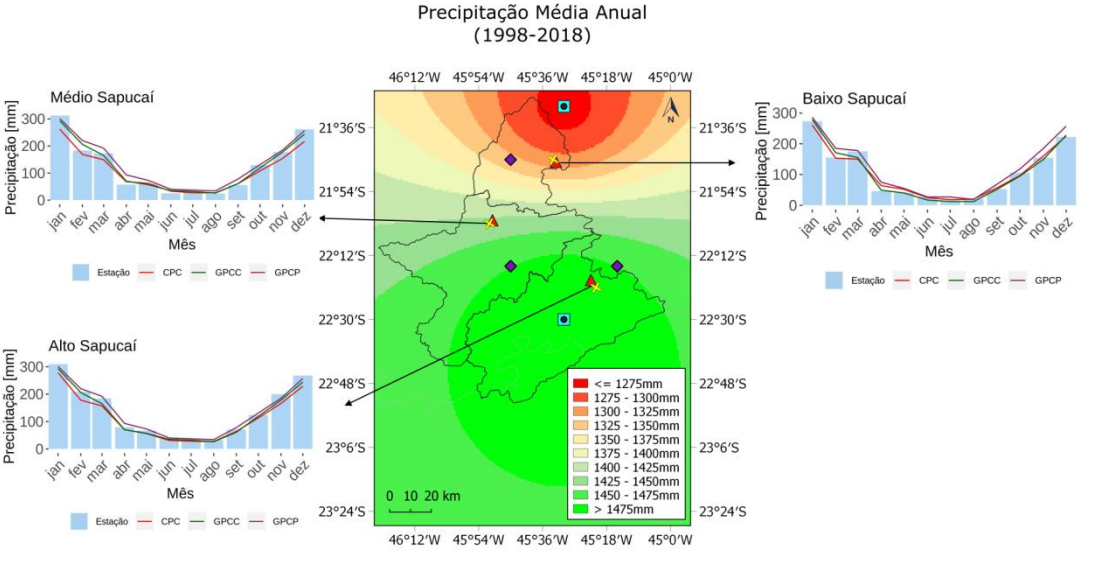


Figura 8 – Idem à **Figura 7**, mas para a Bacia do Sapucaí.

3.2 COMPARAÇÃO MENSAL 2014-2018

Com o intuito de avaliar o desempenho das análises MERGE-CPTEC, CPC e GPCC ao longo de anos mais recentes (2014-2018), foram calculadas as estatísticas mostradas na **Tabela 2**, onde os melhores resultados foram destacados em azul e os piores em vermelho. Nessa análise, o GPCP não foi incluído, pois sua série temporal se estende somente até 2015. Os estudos realizados nesta seção não desconsideram o ciclo sazonal de precipitação nas sub-bacias do São Francisco e do Sapucaí.

ISSN: 2237-8642 (Eletrônica)

O MAE indica que o MERGE-CPTEC se destaca positivamente, apresentando os menores erros para os casos das sub-bacias Submédio e Baixo São Francisco e Alto e Médio Sapucaí (10,93, 9,47, 12,08 e 37,26, respectivamente), enquanto o CPC apresenta o melhor desempenho no Médio São Francisco e Baixo Sapucaí (9,18 e 24,24, respectivamente) e o GPCC apenas no Alto São Francisco (16,88). Ademais, é importante destacar que, de forma geral, os valores assumidos pela estatística MAE são expressivamente superiores ao valor ótimo definido por Hallak e Filho (2011), que é MAE igual a zero.

O RMSE indica que o MERGE-CPTEC possui os menores erros para as sub-bacias Baixo São Francisco e Alto e Médio Sapucaí (16,12, 17,97 e 61,93, respectivamente), enquanto cada um dos conjuntos CPC e GPCC apresentam melhor desempenho em duas sub-bacias diferentes. A saber: Médio e Submédio São Francisco e Alto São Francisco e Baixo Sapucaí, respectivamente. Similarmente ao caso do MAE, os valores obtidos para a estatística RMSE são considerados elevados em todos os casos quando se usa como parâmetro de comparação o valor ótimo definido por Hallak e Filho (2011), que é RMSE igual a zero.

Quando é aplicada a classificação definida por Hinkle et al. (2003) nos valores calculados para a estatística r, notam-se correlações muito altas e superiores a 0,9 para todas as análises em cinco sub-bacias (Alto, Médio e Baixo São Francisco e Alto e Baixo Sapucaí). Já no caso das outras duas sub-bacias (Submédio São Francisco e Médio Sapucaí), o valor da estatística r de cada análise fica acima de 0,8, o que indica alta correlação.

A qualidade do ajuste de cada uma das análises é avaliada por meio do R^2. Segundo Morettin e Bussab (2010), quanto mais próximo de 100% a estatística R^2 estiver, melhor é a qualidade do ajuste. Os resultados obtidos indicam que o CPC apresenta o melhor desempenho nas sub-bacias Médio e Submédio São Francisco e Médio Sapucaí (98,0%, 84,7% e 74,0%, respectivamente) e o GPCC no Alto e Baixo São Francisco e Baixo Sapucaí (93,8%, 95,6% e 86,6%, respectivamente). Já o MERGE-CPTEC demonstra bom resultado apenas no Alto Sapucaí (95,8%).

A fim de avaliar a capacidade das análises em representar as observações, foram utilizadas as medidas NSE e *d*. Com relação ao NSE, se for aplicada a classificação definida por Moriasi et al. (2007), observa-se que todas as análises nas sub-bacias Alto, Médio e Baixo São Francisco e Alto e Baixo Sapucaí apresentam muito bom desempenho (1,00>NSE>0,75). Enquanto nas sub-bacias restantes predominam análises com bom desempenho (0,75>NSE>0,65) e desempenho satisfatório (0,65>NSE>0,50) em representar os dados medidos *in situ*. Entretanto, o CPC continua apresentando os melhores resultados para as sub-bacias Médio e Submédio São Francisco e Médio Sapucaí (0,97, 0,84 e 0,67), o MERGE-CPTEC no Baixo São Francisco e Alto Sapucaí (0,95 para ambas as sub-bacias) e o GPCC no Alto São Francisco e Baixo Sapucaí (0,93 e 0,86, respectivamente).

Para o índice *d*, o MERGE-CPTEC exibe os melhores resultados para o caso das sub-bacias Submédio e Baixo São Francisco e Alto e Médio Sapucaí (0,85, 0,91, 0,91 e 0,79, respectivamente). Com relação às exceções, o CPC apresenta o melhor desempenho no Médio São Francisco e Baixo Sapucaí (0,94 e 0,82, respectivamente), enquanto o GPCC no Alto São Francisco e Baixo

ISSN: 2237-8642 (Eletrônica)

Sapucaí (0,89 e 0,82, respectivamente). Apesar da ocorrência de algumas exceções, as análises apresentaram índices superiores a 0,8 e próximos ao ideal definido por Willmott et al. (2012), que é de *d* igual a 1.

Com relação ao *bias*, os resultados obtidos para as análises revelam que não existe um padrão único aparente de subestimação ou superestimação em todas as sub-bacias. Na verdade, observam-se valores com módulos pequenos e sinais alternados, com algumas exceções no São Francisco: para o caso do MERGE-CPTEC e CPC no Alto e Médio São Francisco (13,73 e -9,73 para o MERGE-CPTEC, e -11,43 e -6,50 para o CPC, respectivamente), e para o GPCC no Submédio e Baixo São Francisco (5,89 e 20,27, respectivamente). Já na bacia do Sapucaí, as exceções ocorrem em duas sub-bacias (Alto e Médio Sapucaí) para o caso de todas as análises.

As **Figuras 9** e **10** mostram o ciclo anual da precipitação no período de 2014-2018 e uma comparação com o período 1998-2018. O padrão do ciclo anual segue o descrito na seção anterior, porém o interessante aqui são os menores valores médios mensais obtidos no período de 2014-2018 na sub-bacia do Submédio São Francisco. Isso está relacionado às secas descritas, por exemplo, em Marengo et al. (2018) e Cunha et al. (2019).

Tabela 2 – Estatísticas calculadas para avaliação do desempenho de cada análise (MERGE, CPC e GPCC) de precipitação mensal acumulada com relação às medições feitas nos postos pluviométricos das sub-bacias consideradas, no período de 2014 a 2018. Resultados em azul (vermelho) indicam melhor (pior) desempenho.

Bacia	Sub-bacia	Análise	MAE	RMSE	r	R^2	NSE	*d*	*bias*
São Francisco	Alto	MERGE	20,79	30,58	0,96	0,932	0,90	0,86	13,73
		CPC	20,52	38,01	0,94	0,900	0,86	0,87	-11,43
		GPCC	16,88	26,26	0,96	0,938	0,93	0,89	-1,66
	Médio	MERGE	13,75	29,97	0,96	0,937	0,92	0,91	-9,73
		CPC	9,18	17,99	0,99	0,980	0,97	0,94	-6,50
		GPCC	15,38	25,21	0,97	0,952	0,94	0,90	-0,79
	Submédio	MERGE	10,93	22,81	0,87	0,768	0,75	0,85	2,14
		CPC	11,49	18,31	0,92	0,847	0,84	0,84	0,48
		GPCC	19,10	29,06	0,83	0,690	0,60	0,73	5,89
	Baixo	MERGE	9,47	16,12	0,97	0,953	0,95	0,91	-1,27
		CPC	13,02	20,91	0,97	0,940	0,91	0,88	-2,62
		GPCC	20,77	28,49	0,97	0,956	0,84	0,81	20,27
Sapucaí	Alto	MERGE	12,08	17,97	0,97	0,958	0,95	0,91	-1,17
		CPC	21,52	29,13	0,95	0,916	0,89	0,86	-13,34

		GPCC	30,35	42,44	0,88	0,791	0,75	0,79	-14,88
	Médio	MERGE	37,26	61,93	0,83	0,700	0,67	0,79	-16,72
		CPC	38,77	62,68	0,86	0,740	0,67	0,78	-24,62
		GPCC	39,55	66,17	0,83	0,698	0,62	0,78	-25,80
	Baixo	MERGE	27,60	38,92	0,92	0,854	0,84	0,80	6,08
		CPC	24,24	40,30	0,91	0,842	0,83	0,82	1,40
		GPCC	24,80	36,74	0,93	0,866	0,86	0,82	-3,84

Legendas: MAE – *Mean Absolute Error* (Erro Absoluto Médio); RMSE – *Root Mean Square Error* (Raiz do Erro Médio Quadrático); r – Coeficiente de Correlação de Pearson; R^2 - Coeficiente de Determinação; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; *d* – Índice de Willmott; *bias* – viés.

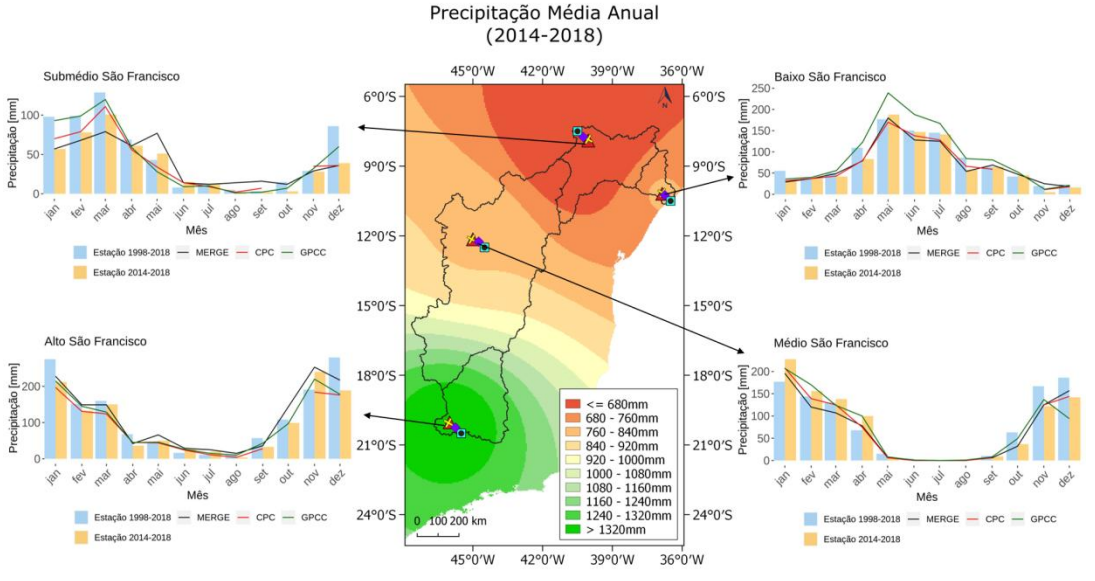


Figura 9 – Mapa com a precipitação média anual do MERGE-CPTEC e gráficos com o ciclo anual da precipitação observada (barras azuis indicam o período 1998-2018 e barras laranjas o período 2014-2018) e das análises (MERGE-CPTEC, CPC e GPCC), entre os anos de 2014 e 2018 na bacia do São Francisco.

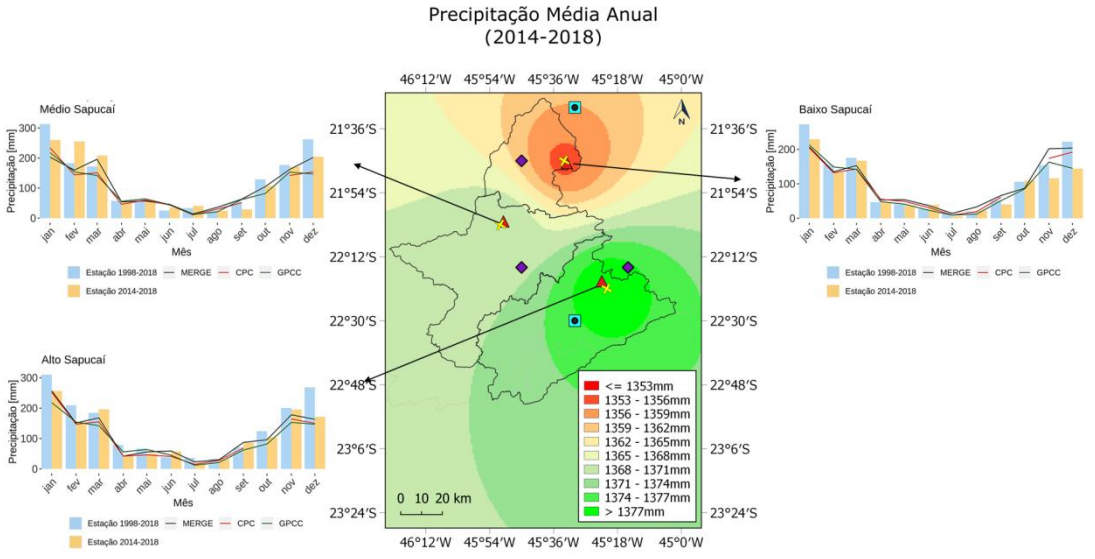


Figura 10 – Idem à **Figura 9**, mas para a bacia do Sapucaí.

3.3 COMPARAÇÃO DIÁRIA

Além de avaliar o comportamento mensal dos conjuntos de dados, é importante verificar se as análises, cujos dados históricos são originalmente diários, cumprem o papel a que se dispõem que é o de fornecer valores diários estimados em concordância com os observados.

Neste estudo, o GPCC e o GPCP foram desconsiderados em razão de utilizarem um período distinto dos dados observados para acumular diariamente a precipitação (detalhes foram apresentados na seção Metodologia). Um trecho das séries temporais é mostrado nas **Figuras 11** e **12**. A diferença entre as análises e as observações é obtida através de medidas estatísticas aplicadas no período de 2014 a 2018 (**Tabela 3**). A fim de facilitar a visualização dos principais valores obtidos, foram destacados em azul os melhores resultados e em vermelho os piores.

O MAE (**Tabela 3**) indica que os erros absolutos médios de cada análise em relação aos dados observados não ultrapassam 3,5 mm. Os baixos erros verificados estão bem próximos ao valor ótimo definido por Hallak e Filho (2011), que é MAE igual a zero. Em todas as sub-bacias, com exceção do Médio São Francisco e Baixo Sapucaí, o MERGE-CPTEC apresenta o melhor desempenho. Apesar do MERGE-CPTEC apresentar números melhores que o CPC, os erros associados às duas análises são muito próximos.

ISSN: 2237-8642 (Eletrônica)

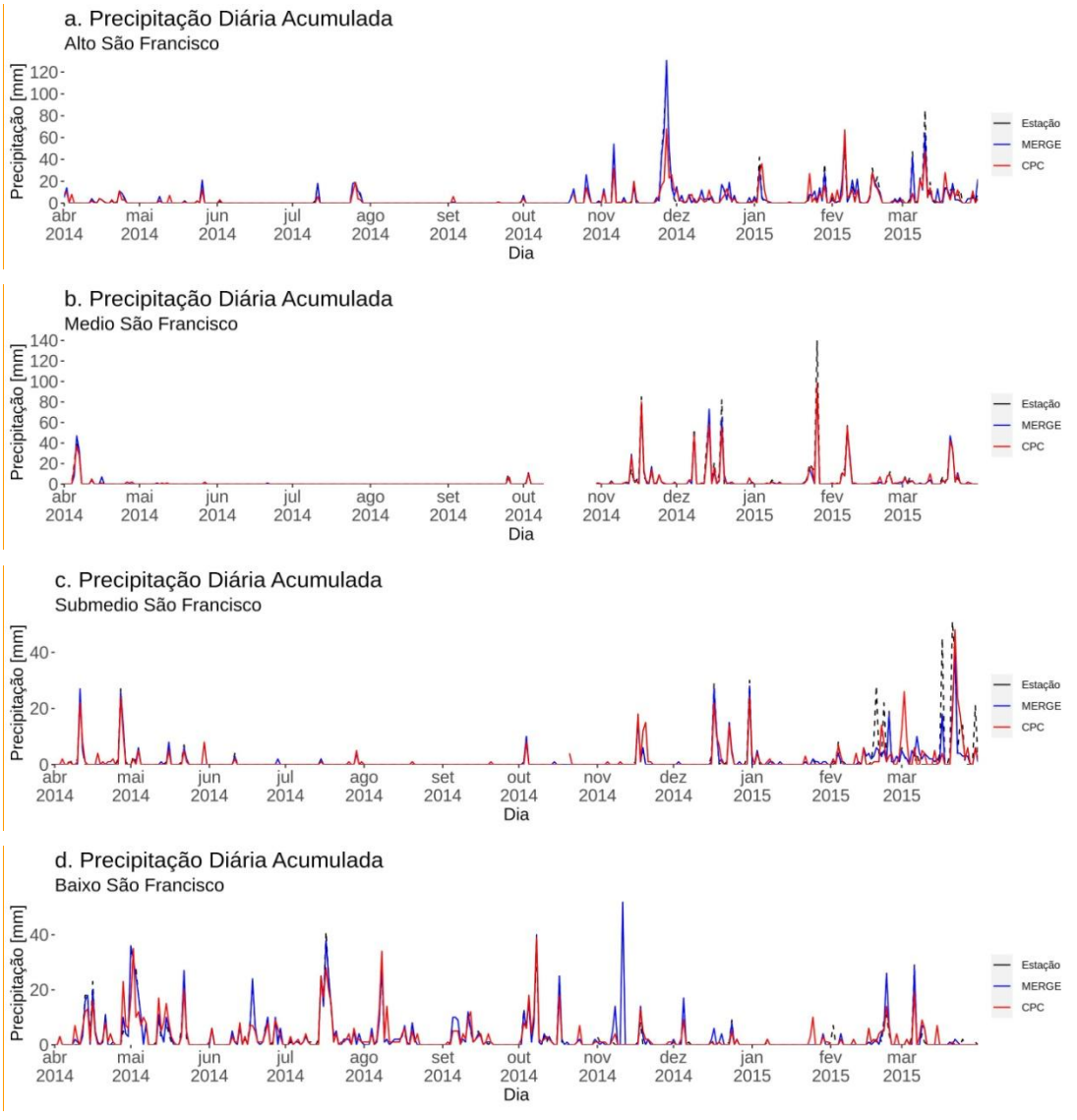


Figura 11 – Séries históricas de precipitação diária acumulada observada (medida nas estações) e das análises (MERGE-CPTEC e CPC), entre os anos de 2014 e 2015 na bacia do São Francisco.

ISSN: 2237-8642 (Eletrônica)

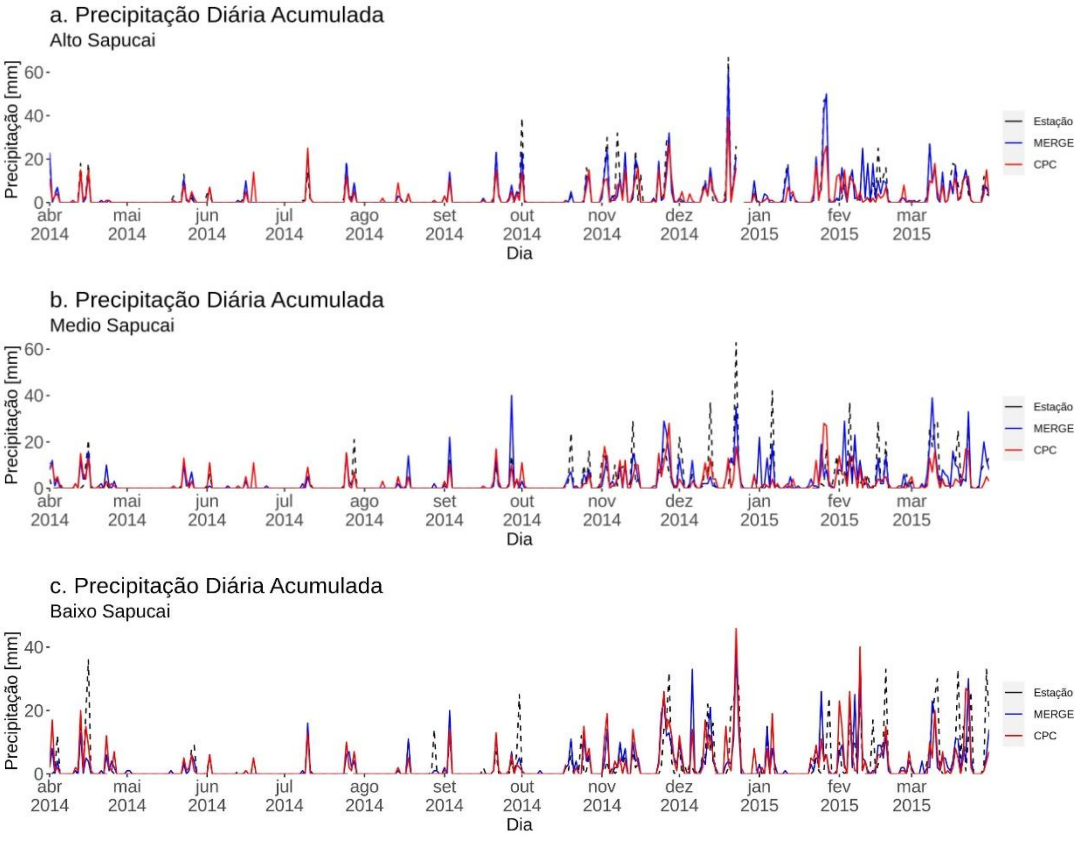


Figura 12 – Idem à **Figura 11**, mas para a bacia do Sapucaí.

Os valores calculados para o RMSE concordam com o MAE, uma vez que o MERGE-CPTEC continua exibindo um melhor desempenho que o do CPC na maioria das sub-bacias (exceto o Médio São Francisco, em que o RMSE é 2,82). Ademais, os valores obtidos para a estatística RMSE são baixos em quase todos os casos, quando se usa como parâmetro de comparação o valor ótimo definido por Hallak e Filho (2011), que é RMSE igual a zero.

Quando se trata da correlação entre os dados estimados e observados, verifica-se que o desempenho das análises varia bastante. As correlações altas ou muito altas que foram observadas na análise mensal não se repetem com tanta frequência na análise diária. Como a variabilidade dos dados tende a ser maior com a redução do período de acumulação da precipitação, torna-se mais difícil os dados estimados captarem corretamente os picos de precipitação observados nas estações. Isso implica diretamente em uma queda generalizada no desempenho das análises, principalmente nas sub-bacias do Médio e Baixo Sapucaí que apresentam correlações moderadas. Nas sub-bacias do Alto, Submédio e Baixo São Francisco e Alto Sapucaí, observam-se correlações altas para todas as análises, enquanto no Médio São Francisco as correlações são muito altas. Ademais, o MERGE-CPTEC se destaca por apresentar um desempenho maior que o do CPC na maioria das sub-bacias (exceto Médio São Francisco, em que a correlação é do MERGE-CPTEC de 0,95).

Assim como foi verificado para o caso da correlação, a qualidade do ajuste das análises aos dados observados, representada pela estatística R^2, sofre queda significativa de desempenho quando a avaliação passa a ser com

ISSN: 2237-8642 (Eletrônica)

dados diários, principalmente no Sapucaí (Médio e Baixo Sapucaí). Em diversos casos, os valores calculados para R^2 foram expressivamente inferiores ao valor ótimo indicado por Morettin e Bussab (2010), que é R^2 igual a 100%. Esse comportamento decorre das análises terem maior dificuldade em captar corretamente os picos diários de precipitação observada, o que pode ser uma consequência da suavização que ocorre nos dados devido à interpolação espacial. Sugere-se que estudos futuros investiguem os efeitos da interpolação espacial sobre os dados diários de precipitação. Além disso, observa-se que o MERGE-CPTEC continua apresentando os melhores resultados para a maior parte das sub-bacias (exceto Médio São Francisco, em que R^2 é igual a 91,5%).

Com relação ao NSE e *d*, o MERGE-CPTEC apresenta o melhor desempenho para todas as sub-bacias, exceto para os casos do Médio e Baixo São Francisco na estatística NSE (0,91 e 0,72, respectivamente) e Médio São Francisco e Baixo Sapucaí no caso da estatística *d* (0,90 e 0,71, respectivamente). Seguindo a classificação estabelecida por Moriasi et al. (2007) para o NSE, pode-se afirmar que as análises apresentam bom desempenho para as sub-bacias Alto e Baixo São Francisco; desempenho muito bom para o caso do Médio São Francisco; e insatisfatório para o Médio e Baixo Sapucaí. Com relação à estatística *d*, nota-se que, para os casos das sub-bacias Médio e Baixo Sapucaí, os valores calculados para as duas análises se distanciam do valor ideal definido por Willmott et al. (2012), que é de *d* igual a 1.

A última estatística apresentada na **Tabela 3** é o *bias*. Os resultados obtidos para as análises revelam que não existe um padrão aparente de subestimação ou superestimação da precipitação. Os valores de *bias* calculados para cada análise e sub-bacia estão muito próximos de zero, o que significa que não é possível interpretar estatisticamente os resultados considerando-se apenas os valores calculados para esse parâmetro. No entanto, observa-se por meio da análise conjunta de *bias* e MAE que as diferenças entre valores estimados e observados é realmente pequena. Tais resultados sugerem que não existem tendências nos conjuntos de dados das análises.

Tabela 3 – Estatísticas calculadas para avaliação do desempenho de cada análise (MERGE-CPTEC e CPC) de precipitação diária acumulada com relação às medições feitas nos postos pluviométricos das sub-bacias consideradas, no período de 2014 a 2018. Resultados em azul (vermelho) indicam melhor (pior) desempenho.

Bacia	Sub-bacia	Análise	MAE	RMSE	r	R^2	NSE	*d*	*bias*
São Francisco	Alto	MERGE	1,36	4,51	0,89	0,805	0,76	0,86	0,44
		CPC	1,71	5,20	0,83	0,701	0,69	0,82	-0,39
	Médio	MERGE	0,81	2,82	0,95	0,915	0,91	0,90	-0,29
		CPC	0,67	2,33	0,97	0,945	0,94	0,92	-0,19
	Submédio	MERGE	0,75	3,00	0,86	0,748	0,73	0,85	-0,23
		CPC	0,95	3,34	0,76	0,589	0,58	0,76	0,02
	Baixo	MERGE	0,72	3,14	0,88	0,784	0,72	0,89	0,29
		CPC	1,14	3,17	0,86	0,752	0,75	0,82	-0,07
Sapucaí	Alto	MERGE	1,21	3,20	0,92	0,862	0,85	0,88	-0,05
		CPC	1,97	4,54	0,84	0,715	0,71	0,80	-0,43
	Médio	MERGE	3,12	8,34	0,60	0,369	0,35	0,73	-0,72
		CPC	3,27	8,66	0,56	0,313	0,30	0,71	-0,96
	Baixo	MERGE	2,38	5,70	0,69	0,485	0,43	0,71	0,17
		CPC	2,38	5,79	0,67	0,461	0,41	0,72	0,02

Legendas: MAE – *Mean Absolute Error* (Erro Absoluto Médio); RMSE – *Root Mean Square Error* (Raiz do Erro Médio Quadrático); r – Coeficiente de Correlação de Pearson; R^2 - Coeficiente de Determinação; NSE – Coeficiente de Eficiência de Nash-Sutcliffe; *d* – Índice de Willmott; *bias* – viés.

4. CONCLUSÕES

Este estudo teve como objetivo avaliar a performance de dados em pontos de grade (chamados de análises) em representar a precipitação medida *in situ* nas bacias hidrográficas do Sapucaí e São Francisco por meio de parâmetros estatísticos (MAE, RMSE, r, R^2, NSE, *d* e *bias*).

O estudo foi dividido em três partes:

a) totais mensais de precipitação no período de 1998 a 2018: aqui foram avaliadas as análises do CPC, GPCC e GPCP. As estatísticas mostraram que os três conjuntos de dados possuem desempenho satisfatório na representação dos dados mensais observados, com destaque para o CPC que apresenta o melhor desempenho em quatro sub-bacias (Médio e Baixo São Francisco e Alto e Baixo Sapucaí). O GPCP tem bom desempenho em duas sub-bacias (Submédio São Francisco e Médio Sapucaí), enquanto o GPCC em apenas uma sub-bacia (Alto São Francisco);

b) totais mensais de precipitação no período de 2014 a 2018: como o MERGE-CPTEC possui apenas dados no período mencionado, aqui foram avaliadas as médias mensais dele, do CPC e do GPCC. Nenhuma análise se sobressai com resultados expressivamente superiores aos das outras. As estatísticas indicam que o MERGE-CPTEC apresenta o melhor desempenho em duas sub-bacias (Baixo São Francisco e Alto Sapucaí), enquanto o CPC em três (Médio e Submédio São Francisco e Médio Sapucaí) e o GPCC em duas (Alto São Francisco e Baixo Sapucaí) e;

c) totais diários de precipitação no período de 2014 a 2018: aqui foram avaliadas as séries temporais dos acumulados diários de precipitação do MERGE-CPTEC e do CPC por estes serem os conjuntos de dados que acumulam a precipitação das 12:00 às 12:00 UTC, seguindo o padrão das estações meteorológicas do INMET. O MERGE-CPTEC apresentou desempenho superior ao CPC em todas as sub-bacias, com exceção de uma (Médio São Francisco).

Por fim, ressalta-se que para estudos futuros cujo objetivo envolva o uso de dados mensais, as análises do CPC e MERGE-CPTEC podem ser úteis. Já em estudos diários, o MERGE-CPTEC possui uma performance ligeiramente melhor do que a do CPC. Entretanto, ambas podem ser utilizadas nos estudos.

5. AGRADECIMENTOS

Os autores agradecem aos centros que disponibilizam os dados utilizados no estudo e ao auxílio financeiro da CAPES, FAPEMIG e CNPq. Os autores também agradecem aos editores e revisores da Revista Brasileira de Climatologia pelos valiosos comentários e sugestões.

6. REFERÊNCIAS

ABE, C. A., NOVO, E. M. L. M. Comparação entre dados de precipitação do TRMM e do MERGE/CPTEC para uma bacia da região amazônica. In: XIX SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 19., 2019, Santos. *Anais* [...].

ANA - Agência Nacional de Águas. Cartilha do observador - Pluviometria. Brasília: ANA, 2« ed., 2014. 23 p.

ANA – Agência Nacional de Águas. Conjuntura dos recursos hídricos no Brasil. Brasília: ANA, 2015. 164 p.

ASHFAQ, M., CAVAZOS, T., REBOITA, M. S., TORRE-ALAVEZ, J. A., IM, E. S., OLUSEGUN, C. F., ALVES, L., KEY, K., ADENIYI, M. O., TALL, M., SYLLA, M. B., MEHMOOD, S., ZAFAR, Q., DAS, S., DIALLO, I., COPPOLA, E., GIORGI, F. Robust late 21st Century shift in the regional monsoon in RegCM-CORDEX simulations. *Climate Dynamics*, Heidelberg, 2020.

ASSIS, J. M. O., SOUZA, W. M., SOBRAL, M. C. Análise climática da precipitação no Submédio da bacia do rio São Francisco com base no índice de anomalia de chuva. *Revista Brasileira de Ciências Ambientais*, Rio de Janeiro, n. 36, p. 115-127, 2015.

BACK, A. J., OLIVEIRA, J. L. R., HENN. A. Relação entre precipitações intensas de diferentes durações para desagregação da chuva diária em Santa Catarina. *Revista Brasileira de Engenharia Agrícola e Ambiental*, Campina Grande, v. 16, n. 4, p. 391-398, 2012.

BACK, A. J., POLA, A. C. Relações entre precipitações intensas de diferentes durações de Videira, Santa Catarina. *Revista Tecnologia e Ambiente*, Criciúma, v. 22, p. 125-135, 2016.

CALDEIRA, T. L., BESKOW, S., MELLO, C. R., FARIA, L. C., SOUZA, M. R., GUEDES, H. A. S. Modelagem probabilística de eventos de precipitação extrema no estado do Rio Grande do Sul. *Revista Brasileira de Engenharia Agrícola e Ambiental,* Campina Grande, v. 19, n. 3, p. 197-203, 2015.

CAMPOS, B., CALHEIROS, A., CARVALHO, V. S. B. A influência dos Vórtices Ciclônicos de Altos Níveis em eventos extremos de precipitação na bacia do Rio Sapucaí, MG. In: XVII CONGRESSO BRASILEIRO DE METEOROLOGIA, 2012, Gramado. *Anais* [...].

CASSALHO, F., REIS, J. B. C., RENNÓ, C. D., SILVA, B. C. Hydrologic validation of MERGE precipitation products over anthropogenic watershed. *Water*, Basel, v. 12, n. 5, p. 1-21, 2020.

CHEN, M., SHI, W., XIE, P., SILVA, V. B. S., KOUSKY, V. E., HIGGINS, R. W., JANOWIAK, J. E. Assessing objective techniques for gauge-based analyses of global daily precipitation. *Journal of Geophysical Research Atmospheres*, Washington, v. 113, p. 1-13, 2008.

COLLISCHONN, B., ALLASIAM D., COLLISCHONN, W., TUCCI, C. E. M. T. Desempenho do satélite TRMM na estimativa de precipitação sobre a bacia do Paraguai Superior. *Revista Brasileira de Cartografia*, Monte Carmelo, v. 59, n. 1, p. 93-99, 2007.

CUI, W., DONG, X., XI, B., STENZ, R. Comparison of the GPCP 1DD precipitation product and NEXRAD Q2 precipitation estimates over the continental United States. *Journal of Hydrometeorology*, Boston, v. 17, p. 1837-1853, 2016.

CUNHA, A. P. M. A.; ZERI, M.; LEAL, K. D.; COSTA, L.; CUARTAS, L. A.; MARENGO, J. A.; TOMASELLA, J.; VIEIRA, R. M.; BARBOSA, A. A.; CUNNINGHAM, C.; GARCIA, J. V. C.; BROEDEL, E.; ALVALÁ, R.; RIBEIRO-NETO, G. Extreme drought events over Brazil from 2011 to 2019. *Atmosphere*, Basel, v. 10, n. 11, p. 642, 2019.

DANELICHEN, V. H. M., MACHADO, N. G., SOUZA, M. C., BIUDES, M. S. The satellite performance in estimating rainfall over the Midwest Region of Brazil. *Revista Brasileira de Climatologia*, Curitiba, v. 12, p. 22-31, 2013.

DWD - Deutscher Wetterdienst. Integrated Climate Data Center – ICDC. Hamburg, 2020. Disponível em: https://icdc.cen.uni-hamburg.de/en/dwd-station.html. Acesso em: 29 abr. 2020.

ESPINOZA, J. C., RONCHAIL, J.; GUYOT, J. L.; JUNQUAS, C.; VAUCHEL, P.; LAVADO, W.; DRAPEAU, G., POMBOSA, R. Climate variability and extreme drought in the upper Solimões River (western Amazon Basin): understanding the exceptional 2010 drought. *Geophysical Research Letters*, Washington, v. 38, p. L13406, 2011.

FENSTERSEIFER, C. A. *Qualidade das Estimativas de Precipitações Derivadas de Satélites na Bacia do Alto Jacuí – RS*. 2013. 126 p. Dissertação (Mestrado em Engenharia Civil e Ambiental) – Centro de Tecnologia, Universidade Federal de Santa Maria, Santa Maria, 2013.

FRANCHITO, S. H., RAO, B., VASQUES, A. C., SANTO, C. M. E., CONFORTE, J. C. Validation of TRMM precipitation radar monthly rainfall estimates over Brazil. *Journal of Geophysical Research Atmospheres*, Washington, v. 114, p. 1-9, 2009.

HALLAK, R., FILHO, A. J. P. Metodologia para análise de desempenho de simulações de sistemas convectivos na região metropolitana de São Paulo com o modelo ARPS: sensibilidade a variações com os esquemas de advecção e assimilação de dados. *Revista Brasileira de Meteorologia*, São José dos Campos, v. 26, n. 4, p. 591-608, 2011.

HINKLE D. E., WIERSMA W., JURS, S. G. *Applied Statistics for the Behavioral Sciences*. 5« ed. Boston: Houghton Mifflin, 2003. 756 p.

HUFFMAN, G. J., ADLER, R. F., MORRISSEY, M. M., BOLVIN, D. T., CURTIS, S., JOYCE, R., MCGAVOCK, B., SUSSKIND, J. Global precipitation at one-degree daily resolution from multisatellite observations. *Journal of Hydrometeorology*, Boston, v. 2, n. 1, p. 36-50, 2001.

HUFFMAN, G. J., BOLVIN, D. T. *Version 1.2 GPCP One-Degree daily precipitation data set documentation*, 2013.

IBGE - Instituto Brasileiro de Geografia e Estatística. Manual técnico da vegetação brasileira. 2« ed. Rio de Janeiro: IBGE, 2012. 271 p.

IGAM – Instituto Mineiro de Gestão das Águas. Plano diretor de recursos hídricos da bacia hidrográfica do Rio Sapucaí. Resumo executivo. Belo Horizonte: IGAM, p. 10-43, 2010.

JOYCE, R. J., JANOWIAK, J. E., ARKIN, P. A., XIE, P. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, Boston, v. 5, n. 3, p. 487-503, 2004.

LEMES, M. U. V., REBOITA, M. S., TORRES, R. R. Elementos fisiográficos e a precipitação no estado de São Paulo. In: PINHEIRO, L. S.; GORAYEB, A. *Geografia Física e as Mudanças Globais*. Fortaleza: Editora UFC, 2019.

LIMBERGER, L., SILVA, M. E. S. Precipitação observada na Amazônia brasileira: redes convencionais e dados da reanálise I do NCEP/NCAR, CRU E GPCC. *Revista Brasileira de Climatologia*, Curitiba, v. 22, p. 20–37, 2018.

MARENGO, J. A. *Mudanças climáticas globais e seus efeitos sobre a biodiversidade: caracterização do clima atual e definição das alterações climáticas para o território brasileiro ao longo do século XXI*. 2« ed. Brasília: Ministério do Meio Ambiente (Série Biodiversidade, v. 26), 2007. 212 p.

MARENGO, J. A., TOMASELLA, J., SOARES, W. R., ALVES, L. M., NOBRE, C. A. Extreme climatic events in the Amazon Basin. *Theoretical and Applied Climatology*, Wien, v. 107, p. 73-85, 2012.

MARENGO, J. A.; CUNHA, A. P.; ALVES, L. M. A seca de 2012-15 no semiárido do Nordeste do Brasil no contexto histórico. *Revista Climanálise*, São José dos Campos, v. 3, p. 49-54, 2016.

MARENGO, J. A., ALVES, L. M., ALVALA, R. C. S., CUNHA, A. P., CRITO, S., MORAES, O. L. L. Climatic characteristics of the 2010-2016 drought in the semiarid Northeast Brazil region. *Anais da Academia Brasileira de Ciências*, Rio de Janeiro, v. 90, n. 2, p. 1973-1985, 2018.

MARTINS, F. B., GONZAGA, G., SANTOS, D. F., REBOITA, M. S. Classificação climática de Köppen e de Thornthwaite para Minas Gerais: cenário atual e projeções futuras. *Revista Brasileira de Climatologia*, Curitiba, v. 14, p. 129-156, 2018.

MARTINS, C. M. S., SILVA, B. C., PONS, N. A. D. Estimativa de cheias em bacias hidrográficas com base em previsões de precipitação por conjunto. *Revista Brasileira de Geografia Física,* Recife, v. 12, n. 5, p. 1713-1729, 2019.

MMA – Ministério do Meio Ambiente. Caderno da região hidrográfica do São Francisco. Brasília: MMA, 2006. 152 p.

MONTGOMERY, D. C., JENNINGS, C. L., KULAHCI, M. *Introduction to Time Series Analysis and Forecasting.* 2« ed. New Jersey: John Wiley & Sons, 2008. 469 p.

MORAES, J. M. Rio Sapucaí - Vereda da visão de uma alvorada. ACR & Associados. 2003.

MORETTIN, P. A., BUSSAB, W. O. *Estatística Básica*. 6« ed. São Paulo: Editora Saraiva, 2010. 210 p.

MORIASI, D. N., ARNOLD, J. G., LIEW, M. W. V., BINGNER, R. L., HARMEL, R. D., VEITH, T. L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agricultural and Biological Engineers*, St. Joseph, v. 50, n. 3, p. 885-900, 2007.

NASA. NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. GPM Launch. 2014. Disponível em: <https://gpm.nasa.gov/launch>. Acesso em: 19 de junho de 2020.

NASA. NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. TRMM Home Page. 2015. Disponível em: <https://gpm.nasa.gov/TRMM>. Acesso em: 19 de junho de 2020.

NASH, J. E., SUTCLIFFE, J. V. River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology*, Amsterdam, v. 10, n. 3, p. 282-290, 1970.

PEREIRA, G.; SILVA, M. E. S.; MORAES, E. C., CARDOZO, F. S. Avaliação dos dados de precipitação estimados pelo satélite TRMM para o Brasil. *Revista Brasileira de Recursos Hídricos*, Porto Alegre, vol. 18, n. 3, p. 139-148, 2013.

PINHEIRO, M. V. *Avaliação técnica e histórica das enchentes em Itajubá - MG.* Dissertação (Mestrado em Engenharia de Energia). 2005. 104 p. Universidade Federal de Itajubá, Itajubá. 2005.

REBOITA, M. S., GAN, M. A., ROCHA, R. P., AMBRIZZI, T. Regimes de precipitação na América do Sul: uma revisão bibliográfica. *Revista Brasileira de Meteorologia*, São José dos Campos, v. 25, n. 2, p. 185-204, 2010.

REBOITA, M. S., RODRIGUES, M., SILVA, L. F., ALVES, M. A. Aspectos climáticos do estado de Minas Gerais. *Revista Brasileira de Climatologia*, Curitiba, v. 17, p. 206-226, 2015.

REBOITA, M. S., RODRIGUES, M., ARMANDO, R. P., FREITAS, C., MARTINS, D., MILLER, G. Causas da semiaridez do sertão nordestino. *Revista Brasileira de Climatologia*, Curitiba, v. 19, p. 254-277, 2016.

REBOITA, M. S., MARIETTO, D. M. G., SOUZA, A., BARBOSA, M. Caracterização atmosférica quando da ocorrência de eventos extremos de chuva na região sul de Minas Gerais. *Revista Brasileira de Climatologia*, Curitiba, v. 21, p. 20-37, 2017.

REBOITA, M. S., SILVA, B. C., VITORIANO, M. Modelos de regressão aplicados à previsão de nível fluviométrico no Rio Sapucaí em Itajubá - MG. *Anuário do Instituto de Geociências do Rio de Janeiro*, Rio de Janeiro, v. 3, p. 217-229, 2019.

REIS, A. L., SILVA, M. S., REGIS, M. V., SILVEIRA, W. W., SOUZA, A. C., REBOITA, M. S., CARVALHO, V. S. B. Climatologia e eventos extremos de precipitação no estado de Minas Gerais. *Revista Brasileira de Geografia Física*, Recife, v. 11, n. 2, p. 652-660, 2018.

REIS, J. B. C., PONS, N. A. D., LOPES, E. S. S. Monitoramento e alerta de inundação no município de Itajubá (MG) por regressão polinomial. *Revista Geociências*, Rio Claro, v. 35, n. 1, p. 134-148, 2016.

REIS, J. B. C., RENNÓ, C. D., LOPES, E. S. S. Validation of satellite rainfall products over a mountainous watershed in a humid subtropical climate region of Brazil. *Remote Sensing*, Basel, v. 9, n. 1240 , 2017.

ROZANTE, J. R., MOREIRA, D. S., GONÇALVES, L. G. G., VILA, D. A. Combining TRMM and surface observations of precipitation: technique and validation over South America. *Weather and Forecasting*, Boston, v. 25, p. 885-894, 2010.

SANTOS, E., MATOS, H., ALVARENGA, J., SALES, M. C. L. A seca no Nordeste no ano de 2012: relato sobre a estiagem na região e o exemplo de prática de convivência com o semiárido no distrito de Iguaçu/Canindé-CE. *Revista Geonorte*, Manaus, v. 1, n. 5, p. 819-830, 2012.

SANTOS, E. P., CORREIA, M. F., ARAGÃO, M. R. S., SILVA, F. D. S. Eventos extremos de chuva e alterações no regime hidrológico da bacia hidrográfica do rio São Francisco: uma aplicação do índice RAI (Rainfall Anomaly Index). *Engenharia Ambiental: Pesquisa e Tecnologia*, Espírito Santo do Pinhal, v. 8, n. 2, p. 315-330, 2011.

SANTOS, S. R. Q., SANSIGOLO, C. A., NEVES, T. T. A. T., SANTOS, A. P. P. Variabilidade sazonal da precipitação na Amazônia: validação da série de precipitação mensal do GPCC. *Revista Brasileira de Geografia Física*, Recife, v. 10, n. 6, p. 1275–1291, 2017.

SCHAMM, K., ZIESE, M., BECKER, A., FINGER, P., MEYER-CRISTOPHER, A., SCHNEIDER, U., SCHRÖDER, M., STENDER, P. Global gridded precipitation over land: a description of the new GPCC First Guess Daily product. *Earth System Science Data Discussions*, Katlenberg-Lindau, v. 6, 49-60, 2014.

SCHNEIDER, U., FINGER, P., MEYER-CHRISTOFFER, A., ZIESE, M., BECKER, A. Global precipitation analysis products of the GPCC. Deutscher Wetterdienst, 2018.

SERRÃO, E. A. O., WANZELER, R. T. S., SANTOS, C. A., GONÇALVES, L. J. M., LIMA, A. M. M., ROCHA, E. J. P. Avaliação estatística entre as estimativas de precipitação da constelação GPM com TRMM: uma análise da bacia hidrográfica do Rio Solimões. *Revista Brasileira de Climatologia*, Curitiba, v. 18, p. 256-275, 2016.

SILVA, C. B., SILVA, M. E. S., AMBRIZZI, T., TOMMASELLI, J. T. G., PATUCCI, N. N., MATAVELI, G. A. V., LIMA, B. S., CORREA, W. C. Precipitação na América do Sul - Dados obtidos em estações meteorológicas automáticas e sistemas orbitais. *Revista Brasileira de Climatologia*, Curitiba, v. 25, p. 54–79, 2019.

SUN, Q., MIAO, C., DUAN, Q., ASHOURI, H., SOROOSHIAN, S., HSU, K.-L. A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Reviews of Geophysics*, Washington, v. 56, p. 79-107, 2018.

TIAN, Y., LIDARD, C. D. P. A global map of uncertainties in satellite-based precipitation measurements. *Geophysical Research Letters*, Washington, v. 37, n. 24, L24007, 2010.

TUCCI, C. E. M. Impactos da variabilidade climática e do uso do solo nos recursos hídricos. Brasília: Fórum Brasileiro de Mudanças Climáticas, 2002. 150 p.

VICENTE, M. R., SANTOS, J. A., SANTOS, R. M., LEITE, C. V. Comparação entre métodos de estimativa da evapotranspiração de referência para o Vale do Jequitinhonha. *Global Science and Technology*, Rio Verde, v. 7, n. 2, p. 106-118, 2014.

VILLANUEVA-BAEZ, O. M., ZAMBRANO-BIGIARINI, M., RIBBE, L., NAUDITT, A., GIRALDO-OSORIO, J. D., THINH, N. X. Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America. *Atmospheric Research*, Amsterdam, v. 213, p. 34–50, 2018.

WILLMOTT, C. J. On the validation of models. *Physical Geography*, Philadelphia, v. 2, p. 184–194, 1981.

WILLMOTT, C. J., ROBESON, S. M., MATSUURA, K. A refined index of model performance. *International Journal of Climatology*, New York, v. 32, p. 2088-2094, 2012.

WMO n° 100, 2018. Guide to Climatological Practices. Secretariat of the World Meteorological Organization, Geneva, 2018.

XIE, P., ARKIN, P. A. Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society*, Boston, v. 78, p. 2539–2558, 1997.

ZENG, N., YOON, J., MARENGO, J. A., SUBRAMANIAM, A., NOBRE, C. A., MARIOTTI, A., NEELIN, J. D. Causes and impacts of the 2005 Amazon drought. *Environmental Research Letters*, Bristol, v. 3, n. 1, p. 014002, 2008.