XVIII SIMPÓSIO DE RECURSOS HÍDRICOS DO NORDESTE
XVIII ENCONTRO DE RECURSOS HÍDRICOS EM SERGIPE

SiADH-Plu: MÓDULO PARA ANÁLISE DE DADOS PLUVIOMÉTRICOS DO SISTEMA PARA ANÁLISE DE DADOS HIDROLÓGICOS

Henrique Dantas Borges[1]; Elton Vinícius Krauss[1]; Gabriel Meldau Lemos[1]; Gabriel Reis de Vasconcelos[1]; João Carlos Carvalho[1]; Taiara Sampaio Santos[1]; Taína Martins Magalhães[1] & Walszon Terllizzie A. Lopes[1]

RESUMO – Este trabalho apresenta o SiADH-Plu, módulo pluviométrico do Sistema para Análise de Dados Hidrológicos, desenvolvido em Python para apoiar a triagem e a revisão de séries de chuva. O sistema padroniza registros mensais e diários, seleciona estações de apoio, estima uma chuva de referência regional e aplica indicadores de consistência, tais como: outlier climatológico, consistência espacial e temporal, consistência diária, concordância entre fontes e homogeneidade. A abordagem combina sinalização automática com produtos gráficos e interface de revisão. Em uma aplicação com 54 estações do Ceará, foram avaliados 23.814 registros mensais e 706.841 registros diários. A classificação final indicou 90,1% dos meses como consistentes e 9,9% com algum nível de prioridade de revisão. Os resultados mostram que o SiADH-Plu organiza evidências hidrológicas de forma reprodutível e facilita a inspeção de casos prioritários.

ABSTRACT – This paper presents SiADH-Plu, the rainfall module of the System for Hydrological Data Analysis, implemented in Python to support screening and review of rainfall time series. The system standardizes monthly and daily records, selects support stations, estimates regional reference rainfall and applies indicators for climatological outliers, spatial and temporal consistency, daily consistency, agreement between data sources and homogeneity. The approach combines automatic flagging with graphical products and a review interface. In an application with 54 stations in Ceará, Brazil, 23,814 monthly records and 706,841 daily records were evaluated. The final classification indicated 90.1% of monthly records as consistent and 9.9% with some level of review priority. The results show that SiADH-Plu organizes hydrological evidence in a reproducible way and supports the inspection of priority cases.

Palavras-Chave – controle de qualidade; séries pluviométricas; consistência espacial.

1 Agência Nacional de Águas e Saneamento Básico – ANA. Setor Policial, Área 5, Quadra 3, Bloco L. Fone: (61) 2109 5476. E-mail: henrique.borges@ana.gov.br; elton.krauss@ana.gov.br; gabriel.lemos@ana.gov.br; gabriel.vasconcelos@ana.gov.br; joao.carvalho@ana.gov.br; taiara.santos@ana.gov.br; taina.magalhaes@ana.gov.br; walszon@ana.gov.br.

1 - INTRODUÇÃO

Séries pluviométricas são usadas no dimensionamento de obras hidráulicas, estudos de disponibilidade hídrica, análise de cheias e secas, operação de reservatórios, calibração de modelos e acompanhamento de extremos. Nessas aplicações, um erro isolado pode afetar estatísticas de frequência, totais sazonais e comparações regionais. A padronização de observações hidrológicas é premissa recorrente nos guias da Organização Meteorológica Mundial (WORLD METEOROLOGICAL ORGANIZATION, 2008), mas a inspeção visual completa torna-se pouco viável em bases extensas.

A qualidade de um dado de chuva não pode ser definida por uma única regra. Um total elevado pode ser evento extremo real; um valor comum pode ser incoerente com estações de apoio; zeros podem indicar estiagem, falha de observação ou ausência de lançamento; e uma mudança persistente pode estar associada a deslocamento, equipamento, operador ou ambiente local. Por isso, sistemas automatizados de controle de qualidade combinam testes de outlier, consistência temporal e consistência espacial (DURRE et al., 2010).

Esse cuidado é crítico em precipitação, pela alta variabilidade espacial e temporal. Em dados telemétricos, são comuns valores excessivos, falsos zeros, inconsistência entre sensores e baixa concordância espacial (OŚRÓDKA et al., 2022). Assim, a sinalização automática deve ser entendida como triagem técnica, não como descarte definitivo.

O SiADH-Plu, módulo pluviométrico do Sistema para Análise de Dados Hidrológicos, desenvolvido em Python, transforma esse conjunto de verificações em fluxo operacional: organiza evidências, atribui classes de prioridade, preserva motivos de sinalização e oferece interface para revisão técnica. A proposta não é automatizar a decisão hidrológica, mas tornar a etapa de triagem mais transparente, repetível e auditável.

O objetivo deste artigo é apresentar o SiADH-Plu como ferramenta de apoio à avaliação da qualidade de séries pluviométricas, com ênfase no processamento hidrológico-estatístico, nos produtos gerados e no uso por equipes técnicas. O artigo privilegia a lógica dos indicadores e sua interpretação operacional, evitando detalhes de implementação que não são necessários para o usuário final do sistema.

2 - MATERIAL E MÉTODOS

2.1 - Seleção de estações de apoio e referência regional

No fluxo implementado em Python, o ponto central do SiADH-Plu é comparar cada estação analisada com um conjunto de estações de apoio. A aplicação apresentada neste trabalho utilizou 54

estações pluviométricas localizadas no estado do Ceará, com registros nas bases de dados hidrológicos mantidas pela ANA no âmbito da Rede Hidrometeorológica Nacional.

A seleção de apoios considera distância, altitude quando disponível, sobreposição temporal, falhas e correlação entre desvios mensais em relação ao ciclo sazonal local, definidos adiante como anomalias mensais, conforme a Equação (1). Quando há poucos apoios adequados, o raio de busca é ampliado; as estações candidatas devem apresentar meses em comum e comportamento compatível. Essa lógica segue a prática de comparar registros pluviométricos com séries regionalmente coerentes em procedimentos de controle de qualidade (DURRE et al., 2010; OŚRÓDKA et al., 2022).

$$S_j = w_j \cdot r_j \quad (1)$$

em que S_j é a pontuação do apoio j, w_j é o peso associado à distância e r_j é a correlação positiva entre a estação analisada e o apoio.

Com os apoios selecionados, a chuva de referência é calculada como média ponderada dos valores dos apoios, conforme Equação (2). A referência regional é interpretada como estimativa de comparação, condicionada à qualidade, proximidade e coerência temporal dos apoios selecionados.

$$P_{ref} = (\Sigma\, S_j \cdot P_{apoio,j}) \,/\, (\Sigma\, S_j) \quad (2)$$

em que P_{ref} é a chuva de referência, $P_{apoio,j}$ é a chuva observada no apoio j e S_j é a pontuação desse apoio. O sistema também registra quantos apoios entraram no cálculo, quais foram usados e qual foi a sobreposição temporal, permitindo revisar a robustez da comparação.

2.2 - Padronização dos dados e fluxo de processamento

O fluxo operacional começa pela leitura dos metadados, registros mensais e, quando disponíveis, registros diários. As datas são convertidas para referência mensal e os totais tratados em milímetros. Quando o total mensal está ausente, mas há dados diários suficientes, o acumulado é calculado a partir dos dias.

Após a padronização e a seleção dos apoios, o sistema calcula a referência regional, aplica sinalizações de controle de qualidade, executa testes de homogeneidade e consolida resultados tabulares, figuras e dados para a interface. Esse arranjo segue procedimentos automatizados de controle de qualidade que combinam testes independentes sem transformar a etapa automática em decisão final (DURRE et al., 2010). A separação entre processamento em lote e revisão visual permite reproduzir a análise e examinar cada caso com apoio gráfico e documental.

2.3 - Indicadores de qualidade

Os indicadores de qualidade do SiADH-Plu foram agrupados por tipo de evidência: climatológica, espacial, temporal, diária e de concordância entre fontes. Cada indicador produz uma

sinalização associada à estação, ao mês e ao motivo técnico, e a classificação final combina essas evidências em uma prioridade de revisão.

Na síntese mensal de prioridade, os critérios de maior peso são o outlier climatológico mensal e a consistência espacial. O primeiro avalia se o total observado é incomum para o histórico da própria estação; o segundo verifica se esse total é coerente com a referência regional calculada a partir dos apoios. Os demais indicadores qualificam a natureza da inconsistência e ajudam a interpretar a sinalização.

Para reduzir o efeito de diferenças climatológicas entre estações, alguns indicadores trabalham com anomalias mensais. A anomalia é definida como a diferença entre a chuva observada e a média histórica da própria estação no mesmo mês do ano, conforme a Equação (3).

$$A_{est} = P_{est} - P_{mes} \quad (3)$$

em que Aest representa a anomalia mensal, Pest é a chuva observada no mês e P̄mes é a média histórica da estação para o mesmo mês do ano. Assim, estações de regimes pluviométricos distintos são comparadas pelo afastamento em relação ao seu próprio padrão sazonal.

O outlier climatológico mensal identifica totais excepcionalmente altos ou baixos para o mês-calendário da própria estação. O critério usa um escore robusto, calculado pela diferença entre o total observado e a mediana histórica do mês, normalizada por uma medida robusta de dispersão, conforme a Equação (4).

$$Z_R = (P_{est} - M_{mes}) / D_{mes} \quad (4)$$

em que ZR é o escore robusto, Mmes é a mediana histórica do mês-calendário e Dmes é a dispersão robusta estimada. O mês é sinalizado quando o escore e a diferença absoluta em milímetros ultrapassam limiares definidos, reduzindo alertas associados a variações pequenas ou pouco relevantes.

A consistência temporal é avaliada na sequência mensal da própria estação, sem depender diretamente dos apoios. O spike temporal é entendido como pico mensal isolado em relação aos meses adjacentes; a persistência indica repetição ou variação muito baixa por meses consecutivos; e a anomalia sazonal aponta distribuição anual incompatível com o ciclo climatológico esperado. Esses testes complementam os indicadores principais ao identificar padrões que não dependem apenas do total mensal.

A consistência espacial é calculada em duas etapas. Primeiro, estima-se o erro espacial como a diferença absoluta entre a chuva observada na estação e a referência regional obtida pelos apoios, conforme a Equação (5).

$$E_{esp} = |P_{est} - P_{ref}| \quad (5)$$

em que Eesp é o erro espacial entre estação e referência. Esse indicador é central porque expressa diretamente a divergência entre o valor observado e a chuva esperada a partir dos apoios selecionados. Em seguida, o erro espacial é comparado à tolerância estimada para a estação e o mês analisado, conforme a Equação (6).

$$E_{esp} > L_{esp} \quad (6)$$

em que Lesp representa a tolerância espacial. A tolerância evita que pequenas diferenças, esperadas em chuva por sua alta variabilidade local, sejam tratadas como inconsistência relevante.

O zero suspeito, ou falso zero, é tratado como caso específico de incoerência espacial. Ele ocorre quando a estação registra chuva nula ou residual enquanto a referência regional indica precipitação relevante no mesmo mês. Esse indicador é especialmente útil em bases telemétricas, nas quais falhas de medição, transmissão ou registro podem produzir ausência indevida de precipitação (OŚRÓDKA et al., 2022).

Nos dados diários, a revisão busca três tipos de evidência: valores diários extremos, persistência de valores repetidos e incompatibilidade na contagem de dias chuvosos. Essa escala ajuda a interpretar a composição do total mensal, especialmente quando o acumulado parece plausível, mas resulta de poucos dias extremos, registros repetidos ou distribuição diária incompatível com estações próximas.

Quando há fonte telemétrica compatível, os registros são agregados ao mês e comparados à fonte convencional. O indicador de concordância entre fontes considera cobertura, dias válidos e diferença entre totais mensais, permitindo identificar divergências entre registros independentes e ajustar a prioridade de revisão.

2.4 - Homogeneidade e pontos de mudança

A homogeneidade busca pontos de mudança, ou quebras, ao longo da série. Diferentemente dos indicadores mensais, orienta a revisão de períodos: uma quebra estimada serve como referência para comparar o comportamento antes e depois da mudança.

Essa lógica dialoga diretamente com a curva de dupla massa, método gráfico tradicional de análise de consistência em que os acumulados de uma estação são comparados aos acumulados de um padrão regional; mudanças na inclinação indicam alteração na relação entre as séries (SEARCY; HARDISON, 1960). No SiADH-Plu, a curva de dupla massa permanece disponível para inspeção visual, enquanto os testes de homogeneidade formalizam esse mesmo tipo de indício por meio de técnicas estatísticas aplicadas à relação entre estação e referência.

O procedimento é aplicado preferencialmente sobre diferenças entre estação e referência regional, reduzindo o efeito de variações climáticas comuns. Também usa segmentos com conjunto

estável de apoios, evitando confundir mudança no conjunto de apoio com alteração da estação principal.

Na execução foram aplicados testes como Buishand Range, Von Neumann, Pettitt e SNHT, métodos clássicos de homogeneidade e detecção de mudança (VON NEUMANN, 1941; PETTITT, 1979; BUISHAND, 1982; ALEXANDERSSON, 1986). O resultado indica onde procurar mudança persistente, não uma decisão automática de aceite ou rejeição.

2.5 - Classes finais e rastreabilidade

Cada estação-mês recebe uma classe de prioridade de revisão: classe 1, consistente; classe 2, atenção; classe 3, suspeito; e classe 4, crítico.

A classe 1 indica ausência de evidência relevante de problema; a classe 2, atenção por evidência fraca ou isolada; a classe 3, suspeita que merece inspeção técnica; e a classe 4, situação crítica, como grande divergência espacial ou combinação forte de evidências. Na prática, sinais fracos elevam pouco a prioridade; outlier climatológico, consistência espacial e acúmulo de motivos aumentam a classe final.

A classificação final preserva os motivos que originaram cada sinalização, como outlier climatológico mensal, zero suspeito, desvio espacial, inconsistência diária, divergência convencional-telemetria ou consistência temporal. Essa rastreabilidade apoia a decisão técnica sobre aceitar, corrigir, investigar ou descartar o registro e é coerente com abordagens que preservam índices de qualidade para usos posteriores (OŚRÓDKA et al., 2022).

2.6 - Interface de revisão e produtos de saída

Além dos resultados tabulares, o SiADH-Plu inclui interface gráfica que reúne os dados processados por estação, mês e tipo de evidência. As visualizações incluem série temporal, dispersão estação-referência, curva de dupla massa, triagem, tabelas, resumo de estações, dados diários, mapa e boletins de apoio.

O painel de triagem lista os meses sinalizados e permite acessar diretamente o respectivo ponto no gráfico. A interface reúne gráficos, tabelas e registros de apoio para que a equipe técnica consulte, em um mesmo ambiente, a série mensal, a referência regional, os apoios, dados diários, telemetria e notas de revisão.

As saídas estruturadas incluem sínteses tabulares, arquivo geográfico, relatório em PDF, figuras por estação e dados para a interface, permitindo filtrar resultados por estação, mês, classe e motivo de sinalização.

3 - RESULTADOS E DISCUSSÃO

A aplicação analisada considerou 54 estações pluviométricas localizadas no Ceará. O fluxo incluiu padronização mensal e diária, seleção de apoios, referência regional, indicadores, homogeneidade, classificação mensal e revisão na interface. O período coberto foi de outubro de 1910 a setembro de 2025, com 23.814 registros mensais avaliados, 23.197 meses com referência espacial e 706.841 registros diários carregados.

Tabela 1 – Síntese da execução analisada.

Item	Resultado
Recorte espacial	Estado do Ceará
Estações analisadas	54
Período analisado	10/1910 a 09/2025
Registros mensais avaliados	23.814
Registros diários carregados	706.841
Apoios por estação alvo	3 a 5; média = 4,43
Distância dos apoios	média = 54,0 km; mediana = 50,4 km
Sobreposição temporal dos apoios	mediana = 250 meses
Sinalizações de controle de qualidade	5.013
Homogeneidade	166 resultados de teste; 31 quebras; 15 estações com quebra
Classes de prioridade de revisão	classe 1 = 21.446; classe 2 = 585; classe 3 = 675; classe 4 = 1.108

A classificação indicou 90,1% dos meses na classe 1, sem evidência relevante de inconsistência, e 9,9% em classes de atenção, suspeita ou crítica. Com isso, a revisão é concentrada em um subconjunto de meses com evidência objetiva de inconsistência, mantendo os demais registros disponíveis para inspeção quando necessário. Esse cuidado é particularmente importante em séries de chuva, nas quais extremos reais podem ser sinalizados por critérios estatísticos, sobretudo em regiões de alta variabilidade espacial (DURRE et al., 2010).

A Tabela 2 reúne os resultados por indicador. A coluna de interpretação resume a lógica hidrológica de cada sinalização: comparação com o histórico local, apoios, coerência temporal, consistência diária, concordância entre fontes ou mudança persistente. Como um mês pode acumular motivos, os quantitativos não são mutuamente exclusivos.

Tabela 2 – Resultados por indicador de qualidade.

Indicador	Interpretação	Quantitativo
Outlier climatológico mensal	Total muito distante do histórico local do mesmo mês; $Z_R = (P_{est} - M_{mes}) / D_{mes}$.	197 meses
Spike temporal	Valor mensal isolado em relação aos meses adjacentes da própria estação.	113 meses
Persistência mensal	Repetição ou baixa variação em sequência mensal.	215 meses
Anomalia sazonal	Distribuição anual incompatível com o ciclo climatológico esperado.	138 ocorrências
Consistência espacial	Diferença entre estação e chuva estimada pelos apoios, expressa por Eesp = |Pest - Pref|.	1.004 sinalizações
Zero suspeito (falso zero)	Chuva nula ou quase nula na estação com chuva regional relevante; $P_{est} \approx 0$ e $P_{ref} \geq L_{zero}$.	319 meses
Divergência convencional-telemetria	Diferença entre totais mensais de fontes independentes; $\Delta_{fontes} = P_{conv} - P_{tel}$.	202 sinalizações
Outlier diário	Precipitação diária muito elevada em relação ao padrão observado.	716 registros diários
Persistência diária	Repetição de valor diário em sequência.	7 ocorrências
Inconsistência em dias chuvosos	Contagem mensal de dias chuvosos incompatível com o padrão esperado.	19 meses
Homogeneidade	Mudança persistente no comportamento relativo da estação; testes aplicados sobre diferenças estação-referência.	166 testes; 31 quebras; 15 estações

Entre os indicadores mensais, a consistência espacial tem papel estruturante. O confronto com a referência regional aproxima a revisão do raciocínio hidrológico usado em análise de consistência: um total mensal é interpretado à luz do comportamento dos apoios, não apenas do histórico isolado da própria estação. Essa leitura também dá contexto aos outliers climatológicos mensais, pois ajuda a distinguir um extremo potencialmente regional de um valor isolado na estação.

Os indicadores temporais e diários qualificam a natureza da inconsistência. Picos isolados, persistências e anomalias sazonais apontam padrões que podem passar despercebidos em filtros baseados apenas no total mensal. Os indicadores diários, por sua vez, ajudam a interpretar a composição do acumulado, separando meses plausíveis por total, mas suspeitos pela distribuição interna dos dias chuvosos.

A comparação convencional-telemetria foi aplicada em 33 estações, com 1.696 meses elegíveis e 202 sinalizações. Esse confronto é útil quando as fontes apresentam cobertura suficiente para comparar registros independentes e identificar divergências que alterem a prioridade de revisão.

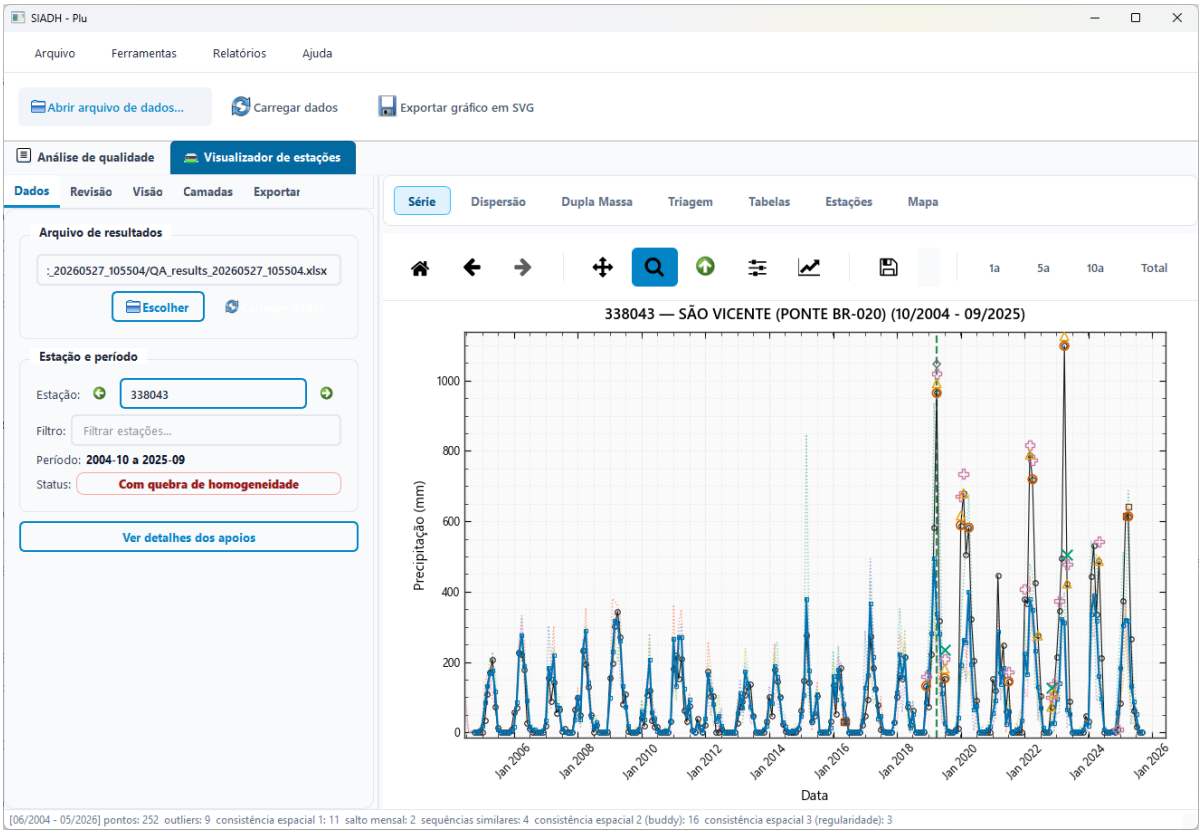


Figura 1 – Interface de revisão do SiADH-Plu com navegação por estação, painel lateral de controles e visualização da série temporal com marcadores diagnósticos.

A interface permite selecionar um mês com sinalização de algum indicador, e, então, o analista visualiza série, referência regional, apoios, telemetria, dados diários, mapa. O módulo de revisão permite que seja indicado se aquele mês deverá ser confirmado como suspeito, ou não.

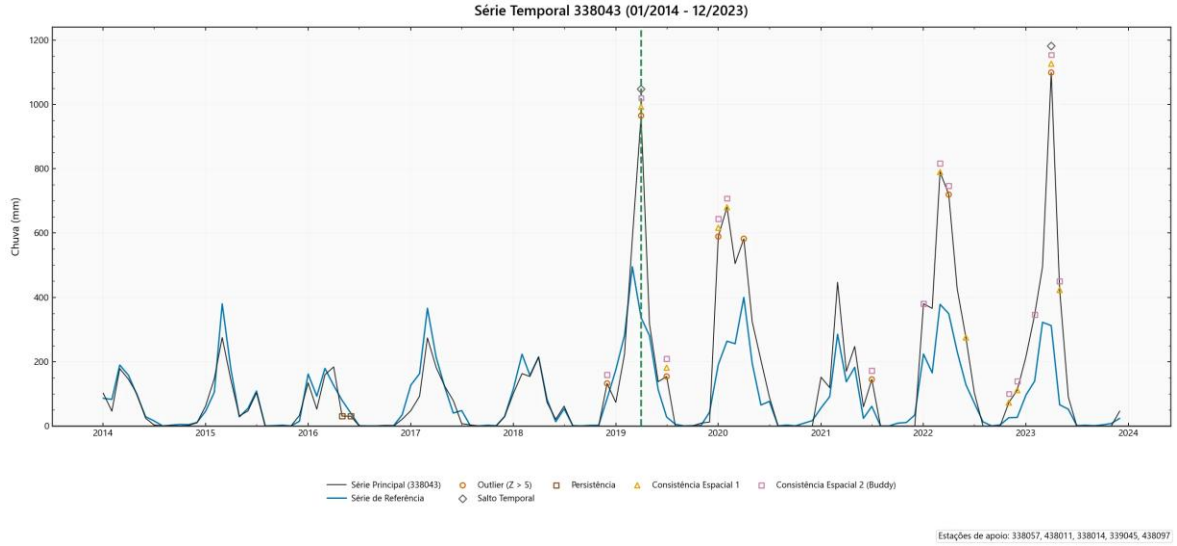


Figura 2 – Série temporal mensal da estação 338043 no período 2014-2023, com chuva observada, referência regional e marcadores de sinalização.

Além da série temporal, o SiADH-Plu exporta curva de dupla massa, dispersão estação-referência, mapas de apoio, tabelas diagnósticas e relatórios, permitindo avaliar a consistência sob perspectivas temporal, acumulada, espacial e tabular.

A homogeneidade complementa a revisão mensal ao apontar mudanças persistentes no comportamento relativo da estação. Na execução, 31 resultados apresentaram data de quebra e 15 estações tiveram indicação de quebra no resumo de homogeneidade. A leitura conjunta com a curva de dupla massa facilita a comparação entre períodos antes e depois da mudança e ajuda a direcionar a consulta a metadados, alterações operacionais e possíveis modificações na área da estação.

4 - CONCLUSÕES

O SiADH-Plu integra padronização de dados, seleção de apoios, referência regional, outlier climatológico mensal, consistência temporal, espacial, diária e entre fontes, homogeneidade, classificação mensal e interface de revisão.

Sua principal contribuição é integrar evidências antes dispersas em um processo rastreável, no qual a revisão avança da triagem mensal para a análise gráfica, diária, espacial e documental de cada caso relevante.

Na execução analisada, o SiADH-Plu processou 54 estações, 23.814 registros mensais e 706.841 registros diários; 90,1% dos meses foram classificados como consistentes e 9,9% direcionados à revisão.

A interpretação dos casos suspeitos e críticos deve considerar dados diários, apoios, telemetria, metadados, boletins e informações sobre extremos registrados no período.

5 - REFERÊNCIAS

ALEXANDERSSON, H. A homogeneity test applied to precipitation data. Journal of Climatology, v. 6, n. 6, p. 661-675, 1986. DOI: 10.1002/joc.3370060607.

BUISHAND, T. A. Some methods for testing the homogeneity of rainfall records. Journal of Hydrology, v. 58, n. 1-2, p. 11-27, 1982. DOI: 10.1016/0022-1694(82)90066-X.

DURRE, I.; MENNE, M. J.; GLEASON, B. E.; HOUSTON, T. G.; VOSE, R. S. Comprehensive automated quality assurance of daily surface observations. Journal of Applied Meteorology and Climatology, v. 49, n. 8, p. 1615-1633, 2010. DOI: 10.1175/2010JAMC2375.1.

OŚRÓDKA, K.; OTOP, I.; SZTURC, J. Automatic quality control of telemetric rain gauge data providing quantitative quality information (RainGaugeQC). Atmospheric Measurement Techniques, v. 15, p. 5581-5597, 2022. DOI: 10.5194/amt-15-5581-2022.

PETTITT, A. N. A non-parametric approach to the change-point problem. Journal of the Royal Statistical Society. Series C (Applied Statistics), v. 28, n. 2, p. 126-135, 1979. DOI: 10.2307/2346729.

SEARCY, J. K.; HARDISON, C. H. Double-mass curves. Washington, DC: U.S. Government Printing Office, 1960. U.S. Geological Survey Water-Supply Paper 1541-B.

VON NEUMANN, J. Distribution of the ratio of the mean square successive difference to the variance. The Annals of Mathematical Statistics, v. 12, n. 4, p. 367-395, 1941. DOI: 10.1214/aoms/1177731677.

WORLD METEOROLOGICAL ORGANIZATION. Guide to Hydrological Practices, Volume I: Hydrology – From Measurement to Hydrological Information. 6. ed. Geneva: WMO, 2008. WMO-No. 168.