ALÉM DOS ERROS TIPO I E II EM ESTUDOS DE TENDÊNCIA HIDROMETEOROLÓGICA: UMA ANÁLISE BASEADA EM SIMULAÇÃO MONTE CARLO

Renato Souza Amorim

Aluno de Mestrado (término em 2018) do PTARH/UnB e Coordenador de Projetos de Infraestrura Aquaviária do DNIT

Saulo Aires de Souza

Aluno de Doutorado do PTARH/UnB e Analista de Recursos Hídricos da Agência Nacional de Águas

Dirceu Silveira Reis Junior

Professor do PTARH/UnB

# Introdução

Os extremos hidrometeorológicos são eventos raros; entretanto, as mudanças no clima têm resultado em alterações na sua frequência e intensidade (Seneviratne *et al*., 2012). Seja pela ação de forças externas, como o aumento da emissão de gases do efeito estufa, ou resultado da variabilidade natural do clima, o cenário de alterações visto para algumas variáveis hidrológicas, tais como a vazão e precipitação, varia de local para local. Essa constatação tem fomentado debates acerca da consideração de não-estacionariedade em problemas que vão desde a gestão de recursos hídricos até a elaboração de projetos de engenharia.

Ao empregar métodos baseados na teoria dos valores extremos para o dimensionamento de projetos de estruturas hidráulicas, como diques e reservatórios, é usual assumir que o comportamento da série observada até o momento do projeto se perpetuará ao longo da vida útil do empreendimento, e, consequentemente, que os riscos hidrológicos associados permanecerão os mesmos. Esse pode não ser o caso de algumas localidades, que experimentam mudanças significativas em variáveis hidrológicas, por exemplo, na vazão, em período pós-construção de uma dada infraestrutura. Assim, negligenciar as alterações detectadas pode não ser a melhor alternativa, pois essa omissão pode resultar em danos econômicos maiores do que aqueles previstos ou até mesmo no colapso das estruturas hidráulicas e/ou em perdas de vidas humanas, razão pela qual discute-se a necessidade de se considerar as eventuais alterações observadas nas variáveis hidrológicas

Detectar uma mudança em variável hidromeoterológica equivale a identificar, a partir do emprego de metodologias de base estatística, alterações amostrais significativas em diferentes escalas de tempo e espaço. A tarefa de detecção é conceitualmente simples, mas de difícil aplicação, tanto pelas limitações de disponibilidade de base amostral, quanto pelas dificuldades intrínsecas à própria atividade. Diversos trabalhos de avaliação de estacionariedade de variáveis hidrológicas foram publicados em periódicos internacionais nos últimos anos (Lettenmaier *et al*., 1994; Krakauer e Fung, 2008; Vogel et al., 2011; Alkama *et al*., 2013; Hossain, 2014; Madsen *et al*., 2014; Bayazit, 2015) e no Brasil (Muller *et al*., 1998; Detzel *et al*., 2011; Aires de Souza *et al*., 2013; Alves *et al*., 2013; Araújo e Aires de Souza, 2014; Scofield *et al*., 2014; Rosin *et al*., 2015).

Em geral, estudos de detecção de tendência são normalmente realizados com base em testes de hipóteses paramétricos ou não-paramétricos, sendo a hipótese nula () comumente assumida como estacionária, ou seja, de que série não possui tendência, enquanto a hipótese alternativa () prevê que a mudança está de fato presente na série. A decisão sobre a presença ou não de estacionariedade na amostra é baseada na distribuição da hipótese nula da estatística do teste (*T*) específica para cada tipo de teste.

A distribuição nula é o nome que se dá à distribuição da estatística do teste *T*, quando a hipótese nula é verdadeira. A hipótese nula é geralmente rejeitada quando o valor amostral de *T* se localiza na cauda da distribuição, indicando uma baixa probabilidade de se observar tal amostra, caso essa conjectura fosse realmente verdadeira. Embora não muito provável, é possível observar uma série estacionária, cujo valor de *T* esteja na cauda da distribuição nula (e.g. . Vem daí a definição do erro tipo I em testes de hipótese e sua relação com a definição da região de rejeição do teste, *R*.

Existe uma relação direta entre a região de rejeição *R* da estatística-teste e a probabilidade de se cometer um erro tipo I, que ocorre quando se rejeita erroneamente a hipótese nula. A probabilidade de se cometer esse engano define o nível de significância do teste, *α*, que pode ser compreendido como a materialização do risco que se está disposto a correr em permitir que determinada hipótese seja rejeitada equivocadamente,

P.

Outro erro passível de ser cometido quando se realiza um teste de hipóteses é o de tipo II, definido como a probabilidade de não rejeição da hipótese nula quando a hipótese alternativa é verdadeira,

|  |  |  |
| --- | --- | --- |
|  |  |  |

A probabilidade complementar a , chamada de poder do teste, é outro parâmetro de interesse por representar a propensão de se rejeitar uma hipótese nula falsa, ou seja, reflete o “poder” de detecção real de evidências que levem à rejeição do teste realizado

|  |  |  |
| --- | --- | --- |
|  |  |  |

Mas os erros tipo I e tipo II não são os únicos erros que podem ser cometidos durante a etapa de detecção de mudanças. Mais recentemente, em estudo ligado à área da medicina, iniciou-se uma discussão acerca de outros tipos de erros que podem prejudicar as análises dos resultados obtidos em estudos de detecção a nível local. Trata-se dos erros Tipo S e Tipo M (Gelman e Carlin, 2014).

O erro Tipo S está relacionado com a estimativa equivocada (contrária) do sinal (S) de mudança, ou seja, quando da realização de um teste de hipótese, existe a probabilidade de se obter uma estimativa de tendência crescente (decrescente) quando na verdade a mesma é decrescente (crescente). Esse tipo de erro pode resultar numa visão totalmente distorcida da realidade.

O erro Tipo M, por sua vez, refere-se à superestimativa da magnitude das tendências. Esse tipo de erro ambém pode ser grave, dependendo do grau de superestimativa, que pode chegar, dependendo do caso em questão, a algumas vezes a magnitude verdadeira.

Ambos os erros se tornam mais frequentes e relevantes à medida que o poder do teste se torna menor. Esses dois erros não são discutidos na literatura de detecção de tendências em variáveis hidrometeorológicas, apesar da importância de seu entendimento no processo de análise de resultados e no planejamento da tomada de decisões em questões que envolvem a gestão de recursos hídricos, por exemplo. Afinal, tão importante quanto a detecção é a estimativa, de forma precisa, da magnitude da mudança, que irá sinalizar aos gestores sua gravidade e a necessidade ou não de intervenção imediata.

Este capítulo tem o objetivo de avaliar, por meio de estudo de Simulação Monte Carlo, como o tamanho da série histórica (*n*), magnitude da tendência (*b*), coeficiente de variação ( e parâmetro de forma (*κ*) da distribuição generalizada de valores extremos (GEV) afetam os erros do tipo S e M associados ao estimador de magnitude da mudança proposto por Sen (1968) – B\_SEN.

# Erros Tipo S e M: Conceitos e Importância

Os efeitos da alteração na frequência e intensidade de eventos hidrológicos têm motivado a realização de estudos de detecção de tendências para subsidiar o processo de gestão dos recursos hídricos. Diversos fatores influenciam a capacidade de detecção de tendências verdadeiras, tais como a presença de autocorrelação temporal, a questão da multiplicidade de testes em análises regionais, e o fato de que o poder do teste empregado na análise e o grau de tendenciosidade na estimativa da magnitude da tendência dependem de características da série histórica (tamanho, graus de variabilidade e assimetria, distribuição geradora, e magnitude real de tendência) (Amorim, 2018).

Mais recentemente, em estudos ligados à área da medicina, iniciou-se uma discussão acerca de outros tipos de erros, além dos tradicionais tipos I e II, que podem prejudicar as análises dos resultados obtidos em estudos de detecção a nível local e regional. Trata-se dos erros Tipo S e Tipo M, que foram formalmente definidos por Gelman e Carlin (2014) como:

1. *Erro tipo S*: a probabilidade de a estimativa do sinal de mudança apresentar sinal incorreto, dado que o mesmo é estatisticamente significante diferente de zero. Em outras palavras, o Erro Tipo S está relacionado com a estimativa equivocada (contrária) do sinal (S) de mudança, ou seja, quando da realização de um teste de hipótese, existe a probabilidade de se obter uma estimativa de tendência crescente (decrescente) quando na verdade a mesma é decrescente (crescente).

1. Erro Tipo M: relacionado à magnitude da tendência estimada; representa a expectativa do valor absoluto da estimativa dividido pelo tamanho efetivo, se estatisticamente significante diferente de zero.

A inexistência de discussões a respeito dos *Erros do tipo S* e *M* na literatura de detecção de tendências em variáveis hidrometeorológicas e a relevância de seu entendimento no processo de análise de resultados foram os aspectos motivadores para o presente capítulo. Afinal, tão importante quanto a detecção é a estimativa, de forma precisa, da magnitude da mudança, que irá sinalizar aos gestores sua gravidade e a necessidade ou não de intervenção imediata.

# Estudo Monte Carlo

A simulação Monte Carlo pode ser definida como um método estatístico que se baseia em amostragens massivas para a obtenção de resultados numéricos, isto é, repete-se sucessivas simulações um elevado número de vezes para calcular probabilidades heuristicamente. Esse tipo de método é utilizado em simulações estocásticas com diversas aplicações em áreas como a física, matemática e biologia e bastante utilizada na obtenção de aproximações numéricas de funções complexas em que não é viável, ou é mesmo impossível, obter-se uma solução analítica.

Na sequência, apresentam-se a lógica e o passo a passo adotado para a geração de séries com e sem tendências lineares distribuídas conforme GEV, via simulação Monte Carlo; a metodologia empregada na estimativa da magnitude da tendência e, por fim, as métricas definidas para estimativa dos erros do tipo S e tipo M

## Lógica do experimento

A simulação Monte Carlo foi empregada na geração de diferentes amostras aleatórias do tipo GEV. Para tal, a seleção dos valores assumidos para parâmetros ditos variáveis durante o processo MC, *n*, *Cv*, *b* e *κ*, foi pautada naqueles usualmente utilizados/vistos na hidrologia. Diante disso, escolheu-se os tamanhos amostrais, coeficientes de variação, parâmetros de forma e magnitudes de tendência iguais a, respectivamente, (), ), e (. A seleção do parâmetro de forma da GEV baseou-se no estudo desenvolvido por Martins e Stedinger (2000), em que se define o intervalo compreendido entre -0.30 e +0.30 como razoável para a definição dos valores de *κ* em aplicações hidrológicas.

Em posse dos valores previamente arbitrados (*n*, e , passa-se ao processo de geração das séries, como discutido na próxima seção.

## Geração de variáveis GEV com e sem a presença de tendência

### A distribuição GEV

A distribuição Generalizada de Valores Extremos engloba três distribuições frequentemente utilizadas na descrição probabilística de variáveis meteorológicas e hidrológicas máximas e mínimas, denominadas EV1 (Gumbel), EV2 (Fréchet) e EV3 (Weibull), cuja função de distribuição acumulada é dada por:

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

em que ξ, α e κ são os parâmetros de posição, escala e forma, respectivamente. Se (EV2), o suporte da distribuição possui um limite inferior (, mas, se (EV3), nota-se a existência de restrição superior . Para , a GEV converge, no limite, para uma distribuição de Gumbel de suporte ilimitado, ou seja, .

Ao contrário de outras distribuições empregadas na avaliação de extremos, tal como a LP3, a GEV possui expressão analítica para o cálculo de quantis:

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

em que *p* é a probabilidade acumulada e os parâmetros da GEV, ξ, α e κ, podem ser empregados no cômputo dos valores da média (, desvio padrão e coeficiente de assimetria populacionais, a saber:

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

### Geração de variável GEV sem tendência

Diferentemente das equações apresentadas para o cálculo de e , que são dependentes de mais de um parâmetro da distribuição GEV, inclusive sendo restritas à valores de , a expressão utilizada para a estimativa do coeficiente de assimetria possui relação exclusiva com o parâmetro de forma *κ*. Essa dependência única favorece o processo de geração de amostras do tipo GEV a partir de um determinado valor fixado de assimetria populacional.

Contudo, o coeficiente de variação (, que é uma estatística populacional que também deve ser controlada durante a obtenção aleatória de séries temporais, depende dos três parâmetros distributivos da GEV, ξ, α e κ. Isso pode ser mostrado a partir da manipulação das equações acima apresentadas, sendo inclusive possível reescrevê-las de modo a correlacioná-lo a dois fatores, o parâmetro de forma e uma variável adimensional :

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

Então, para gerar amostras aleatórias baseadas em uma distribuição GEV de coeficientes de assimetria e variação (, primeiramente determina-se , com base na Eq. (3), e utilizando-se a Eq. (4). Com os valores de e definidos, pode-se, por exemplo, escolher e . Salienta-se que essa última parte é arbitrária, pois qualquer escolha de e α que resultem em proverá um conjunto de amostras distribuídas segundo uma GEV oriundas de população de . Entretanto, para a realização de Simulação Monte Carlo, escolhe-se valores de e que resultem em . Informa-se que o uso do μ unitário está relacionada à geração de séries temporais não-estacionárias, mais especificamente à magnitude das tendências adotadas neste trabalho.

Logo, por conveniência, e sem perder a generalidade, e devem ser calculados de modo a ter :

|  |  |  |
| --- | --- | --- |
|  |  | (5) |
|  |  | (6) |

Em posse dos valores previamente arbitrados (*n*, *κ* e *b*), gerar amostras estacionárias independentes e distribuídas conforme uma GEV é relativamente simples, dada a existência de expressão analítica para a determinação dos quantis. Isso permite que seja empregado o método da Transformação Inversa, que possui dois passos metodológicos básicos:

1. Gerar *p* utilizando-se uma distribuição uniforme entre 0 e 1 ;
2. Calcular com base na Equação 4.2

Mais especificamente, para gerar uma série do tipo GEV com *n* observações de uma determinada população com coeficientes de assimetria e de variação (), são necessários seguir as seguintes etapas metodológicas:

1. Calcular utilizando-se da última função na Eq. (3) para ;
2. Determinar utilizando-se da Eq. (4) com e ;
3. Calcular os valores de e – por conveniência, o cálculo foi realizado utilizando as Equações (5) e (6);
4. Gerar *R* (por ex. 10,000) amostras do tipo GEV seguindo os dois passos executivos básicos do método da Transformação Inversa; para tal, emprega-se e para a estimativa de , por meio da Eq. (2).

### Geração de variável GEV com tendência

No caso da geração de amostras do tipo GEV com a presença de tendências lineares, um termo não-estacionário é adicionado à série independente obtida por meio dos procedimentos anteriormente descritos. Supondo-se como sendo o conjunto das observações independentes e distribuídas segundo uma GEV de parâmetros populacionais e média igual a unidade (*μ* = 1), a série monotônica pode ser expressa por:

(7)

em que *b* é a magnitude da tendência ou a taxa da mudança em relação à média por unidade de tempo. Por exemplo, se *b* =0.01 e a série é de base anual, então as taxas de variação por ano e por década saõ de 1% e 10%, respectivamente. Esclarece-se que caso fossem utilizados valores de μ diferentes da unidade, por exemplo μ = 10, o termo não-estacionário na E. (7), b×t, representaria um percentual de mudança médio pequeno, o que poderia levar a não-percepção de tendência quando de fato a mesma está presente na série.

Salienta-se que essa configuração permite a extrapolação dos resultados a serem obtidos a partir da realização da simulação Monte Carlo proposta, visto que, para um dado valor de *b*, que representa a taxa de mudança relativa da média por unidade de tempo, e , associado aos valores populacionais e da série, conclusões sobre o desempenho dos métodos de detecção podem ser atribuídas a quaisquer outras séries que possuam as mesmas características em termos de , e percentual de mudança em relação à média (*b*).

## Estimativa da magnitude da tendência

Alternativamente ao coeficiente de regressão *β*, obtido pelo método dos mínimos quadrados (OLS), Sen (1968) sugere o emprego de estimador não-paramétrico baseado no Coeficiente de Correlação Tau de Kendall para a determinação da magnitude de tendências, a fim de evitar a influência de pontos amostrais discrepantes na estimativa de *b*.

O estimador proposto por Sen (1968), Sen (b), prevê que seja calculada para cada um dos pares de observação amostrais possíveis , a chamada inclinação do par ordenado, , dada por:

(9)

em que *i* e *j* são os índices de tempo das posições amostrais ocupadas, respectivamente, por e , sendo que , sendo . o tamanho da série.

Formalmente, define-se o estimador de Sen (*b*) como sendo a mediana de todos calculados, a saber:

(10)

## Métricas para estimativa dos erros do tipo S e tipo M

### Erro tipo S

O erro Tipo S ocorre quando a estimativa da tendência de mudança possui sinal contrário ao do valor populacional da tendência. O que se deseja determinar é a probabilidade da estimativa de tendência de mudança apresentar sinal incorreto, dado que a tendência foi considerada estatisticamente diferente de zero. A estimativa da probabilidade de se obter um sinal trocado é obtida por meio da taxa, calcualda pela equação abaixo, tendo como base as amostras geradas no estudo Monte Carlo:

|  |  |
| --- | --- |
|  | (11) |

em que N\_(Erro tipo S) é o número de vezes em que o sinal da tendência estimado por B\_SEN é contrário ao assumido durante o processo de geração serial e N\_significativos é o número de séries declaradas como não estacionárias após a aplicação do teste de Mann-Kendall para um dado conjunto de parâmetros *n*, *b*, *Cv* e *κ*, sendo b≠0;

3.4.2. Erro tipo M

O erro tipo M ocorre quando a magnitude da tendência é incorretamente estimada, dado que a presença de mudança foi considerada estatisticamente diferente de zero. Duas métricas são empregadas aqui para avaliar essa questão. A primeira mede o viés relativo dessa estimativa, determinado da seguinte forma,

() (12)

em que é a média dos valores das estimativas de b ao longo de todas as simulações, para cada conjunto de valores populacionais de *n*, *b*, *Cv* e *κ*, sendo *b* o valor populacional da tendência.

A segunda métrica procura avaliar fornecer uma ideia da variância do estimador de *b*. Nesse caso, a diferença entre valores de magnitude da tendência correspondentes aos percentis de 97.5% e 2.5% da totalidade dos *b’s* estimados por B\_SEN (para um dado conjunto de valores populacionais para *n*, *Cv*, *κ* e *b*) foi a maneira adotada de expressar a variabilidade das estimativas constatada. Essa diferença foi denominada *Largura Relativa do Erro Tipo M* e, após a sua divisão pelo valor da magnitude da tendência verdadeiro, definiu-se a métrica de avaliação dos erros do tipo M, a saber:

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  | (13) |

onde é a métrica indireta de avaliação da variância do erro tipo M, denominada Largura Relativa do Erro Tipo M, são os valores de valor da magnitude de mudança estimados por B\_SEN associados aos percentis de, respectivamente, 97.5 e 2.5% para um dado *n, Cv,* κ e *b*, sendo também .

# Discussão dos Resultados

A Figura 1 mostra a dependência da probabilidade do Erro Tipo S para séries do tipo GEV ao se variar o tamanho amostral (), coeficiente de variação ( e a magnitude da tendência decenal relativa positiva (), mantendo-se fixo apenas o parâmetro de forma . Assim, pode-se constatar que a proporção de séries que tiveram o seu sinal estimado por B\_SEN contrário àquele assumido durante o processo de geração serial pode ser expressiva em diversas situações, ficando claro que a probabilidade do Erro Tipo S é decrescente à medida que se eleva o tamanho da amostra (*n*). Em outras palavras, ao se gerar 10,000 séries do tipo GEV, via MC, para um dado *κ*, *Cv*, *b* fixos e diversos n, a taxa de equívocos de sinal estimados pela metodologia proposta por SEN será tão menor quanto maior o comprimento da série. Por exemplo, para *Cv* = 1.00 e *b* = 2.00%, o caso mais extremo dentre os considerados neste trabalho, a probabilidade de se obter uma tendência com sinal trocado é de 36.4% para *n* = 20, 18% para *n* = 40, 8% para *n* = 60, 2.5% para *n* =80 e menos de 1% para uma série de tamanho *n* = 100.

Naturalmente, os resultados apresentados acima estão relacionados com uma combinação extrema de *Cv* e *b*, sendo outras combinações possíveis de serem encontradas na prática. Por exemplo, um comportamento similar ao observado para o tamanho amostral nos casos mencionados acima ocorre com o aumento da magnitude da tendência, em que a relação negativa fica evidente ao se constatar que as barras da Figura 1 diminuem de magnitude da esquerda para a direita, ou seja, no sentido de crescimento da tendência relativa decenal. Cita-se a diminuição no valor da probabilidade do Erro Tipo S em amostras cujo *n* = 40 e *Cv* =1.0, passando de 18.4%, quando *b* = 2.00%, para 6.3% quando *b* = 4.00%, o que mostra claramente que o problema tem em uma de suas causas a relação entre *Cv* e *b*.

Além disso, a Figura 1 evidencia o crescimento da probabilidade do Erro Tipo S com o aumento do coeficiente de variação, o que pode ser visto para amostras, por exemplo, de tamanho *n*=20 e *b* = 2.00% e que tem o *Cv* aumentado de 0.20 (0.076) para 0.60 (0.295) e, posteriormente para 1.00 (0.364). Intuitivamente, essa relação entre *Cv* e o erro em questão era esperada, dado que o aumento do ruído serial proporciona maior variabilidade à série, conferindo a possibilidade de aumento no número de equívocos do sentido de sinal de mudança. Ratificando a impressão anteriormente relatada, a Figura 2 mostra que, mesmo para diferentes valores do parâmetro de forma (*κ*), a relação positiva entre *Cv* e probabilidade do Erro Tipo S permanece, desde que os demais fatores intervenientes estudados sejam mantidos constantes. Diz-se isso ao observar que, para *n* = 50, *b*=2.00% e *k* =0.3, os valores das ordenadas para *Cv* =0.2 e 1.0 são, respectivamente, iguais a 0.010 e 0.285.

Em relação à Figura 2, nota-se que os efeitos de aumento da magnitude de *b*, de maneira análoga ao que ocorre para o tamanho amostral, resultam em decréscimos no valor da probabilidade do Erro Tipo S observado para diferentes *Cv*’s simulados, ou seja, a relação de dependência entre essas variáveis é negativa.

A relação positiva existente entre o parâmetro de forma da GEV (*κ*) e a probabilidade de se cometer equívocos na indicação do sinal de tendência estimada por B\_SEN, é facilmente constatada na Figura 2 ao se verificar que os menores e maiores valores de probabilidade do Erro Tipo S estão associados, respectivamente, ao menor (-0.3) e maior (0.3) *κ* simulados, para *n*, *Cv* e *b* fixos. Traduzindo em números, tem-se que a diferença observada entre os Erros do Tipo S para *κ* = -0.30 e *κ* = 0.30, considerando-se *n* = 50, *Cv* = 0.60, *b* = 2.00% fixos, é de 13.2%. As relações do parâmetro de forma e a taxa de erro em análise com a variação da magnitude da tendência (*b*) e o coeficiente de variação (*Cv*) são, respectivamente, crescentes e decrescentes.

Adicionalmente, o que chama a atenção é que, independente das variáveis intervenientes testadas, *n*, *Cv*, *b* e *κ*, existem combinações possíveis entre elas que resultam em percentuais de sinais com indicação equivocada em torno de 30%.



Figura – Probabilidade do Erro Tipo S (sinal) dado que a estação foi declarada não-estacionária (α = 0.05) para κ = -0.30 e diferentes valores de e Cv = 0.2 (a), 0.6 (b) e 1.0 (c)



Figura - Probabilidade do Erro Tipo S (sinal) dado que a estação foi declarada não-estacionária (α = 0.05) para n=50 e diferentes valores de e κ = -0.3 (a), 0.0 (b) e 0.3.0 (c)



Figura - Probabilidade do Erro Tipo S (sinal) dado que a estação foi declarada não-estacionária (α = 0.05) para *n=50* e diferentes valores de e Cv = 0.2 (a), 0.6 (b) e 1.0 (c)

Os resultados apresentados até aqui mostram que, em séries com alta variabilidade anual e com magnitude de tendência relativamente pequena – características que resultam em baixo poder de um teste estatístico (depende também do parâmetro de forma da GEV) –, a probabilidade de se obter uma estimativa da tendência com sinal trocado é relativamente alta para amostras com tamanhos geralmente constatados para as estações brasileiras, normalmente inferiores a 60 anos.

A Figura 4 mostra que, para poder do teste com valores inferiores a 0.15, a probabilidade do Erro Tipo S pode ser bastante elevada. Esse fato é um complicador em estudos de detecção, pois pode resultar, por exemplo, em situações em que duas estações próximas venham a ser identificadas como não-estacionárias, porém uma com tendência crescente e outra com tendência decrescente, levantando dúvidas sobre a análise realizada e causando dificuldades na interpretação dos resultados. Como será visto na sequência, esse problema pode ser exacerbado porque o estimador da magnitude da tendência tende a superestimar a real magnitude quando o poder do teste é baixo.



Figura – Curva que expressa a relação entre o poder do teste e a probabilidade do Erro Tipo S | significativo (%)

Os resultados do estudo Monte Carlo permitiram avaliar também o grau de tendenciosidade e a variância do estimador B\_SEN para diferentes características da série. A Figura 5 relaciona o poder do teste com a tendenciosidade relativa do estimador B\_SEN quando a série é declarada não-estacionária pelo teste MK. Salienta-se que a tendenciosidade relativa é expressa pela razão entre o valor esperado do estimador e o valor real (populacional) da magnitude, . Pode-se observar que o estimador tende a superestimar a magnitude real, especialmente em situações em que o poder do teste é baixo. Por exemplo, em casos em que o poder do teste é da ordem de 0.50, a estimativa da magnitude da tendência pode ser, em média, 40% maior do que a magnitude real, chegando à 100% (dobro) para valores de poder iguais a 0.20. Esclarece-se que distorções como essas podem comprometer significativamente a interpretação dos resultados e as decisões de gestão de recursos hídricos ou de dimensionamento de estruturas de engenharia que venham a ser baseadas em tais achados.

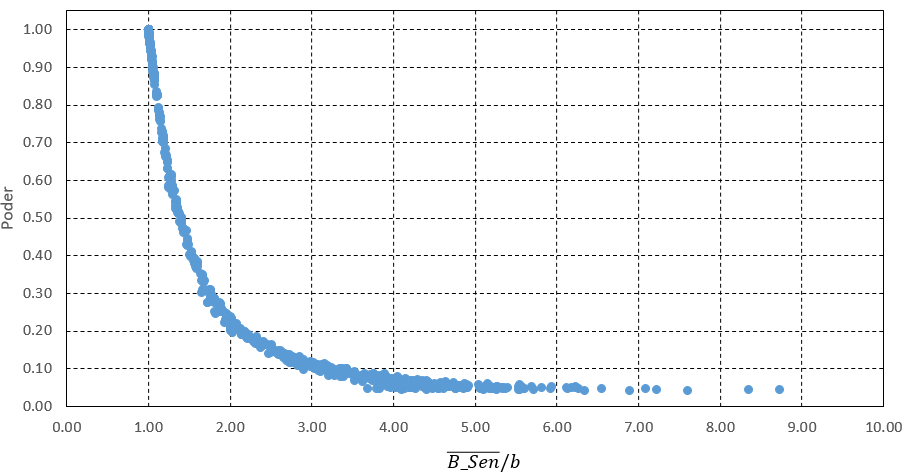


Figura - Curva que expressa a relação entre o poder do teste e o valor médio estimado por B\_SEN dividido por b ()

Além da tendenciosidade relativa do estimador, foi feita uma análise da variância do mesmo, utilizando para isso os percentis de 2.5% e 97.5% (IC) dos valores de B\_SEN estimados para as 10,000 séries geradas para valores fixos de *n*, *Cv* e *κ*. Na verdade, os valores apresentados e discutidos a seguir consideram apenas os casos em que a estação foi declarada não-estacionária, utilizando-se como métrica de avaliação a denominada Largura relativa do Erro Tipo M |Significativas, que nada mais é do que a diferença entre os valores dos percentis dividida pelo valor *b* para as séries significativas. Os resultados obtidos para essa avaliação são apresentados nas Figura 6 e Figura 7, em que se pode observar que, para algumas situações de baixo poder do teste, a Largura relativa do Erro Tipo M pode chegar a valores bastante elevados, 20, por exemplo.

Analisando a Figura 6, que apresenta a relação existente entre o tamanho amostral e a largura do intervalo de confiança obtida para um valor fixo de *κ* = -0.30 e *b* =0.006 e diferentes magnitudes de coeficiente de variação (*Cv* = [0.2 (0.2) 1.0], nota-se que, à medida que o número de anos de dados disponíveis para uma amostra se eleva, o intervalo de confiança se estreita. Já para o coeficiente de variação, observa-se um comportamento contrário: a elevação da sua magnitude resulta em alargamento do IC, ou seja, maior variabilidade nos valores estimados para a tendência pelo B\_SEN. A influência de *Cv* sobre a largura de IC é mais pronunciada em amostras de tamanho reduzido (), visto que o aumento do coeficiente de variação, por exemplo, de 0.20 para 1.00, resulta em diferenças que chegam a quase 25 e 1,1 vez o valor de *b* simulado para, respectivamente, *n* = 20 e *n* = 100.

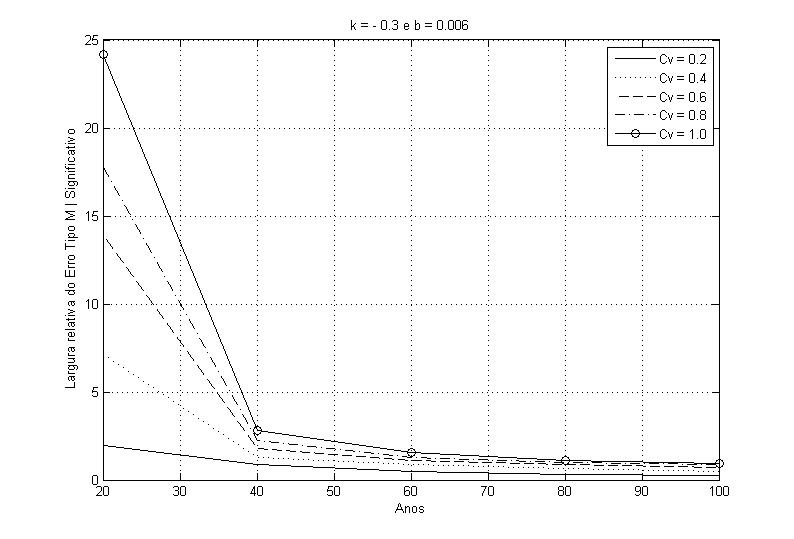


Figura -Largura relativa do IC do Erro tipo M, dado que a estação foi declarada não-estacionária (alpha = 0.05) para *κ* = -0.3, *b* =0.006 e diferentes valores de *Cv* =[0.2 (0.2) 1.0] e *n* = [20 40 60 80 100].

A Figura 7 mostra a interdependência constatada entre o parâmetro de forma (*κ*) e a largura do IC da estimativa da magnitude da tendência por meio B\_SEN, adotando-se um nível de confiança de 5%. Fixando-se os valores de *b* = 0.006 e do *Cv* = 0.6, construiu-se, para cada um dos κ avaliados ([-0.3, -0.1, 0, +0.1, +0.3]), curvas que expressam a relação entre as variáveis procuradas e plotadas nos pares de eixos coordenados. É interessante notar que, assim como ocorre para o *Cv*, o IC se estreita à medida que se aumenta o tamanho amostral, atingindo praticamente o valor esperado de b para séries com . Já o parâmetro de forma tem efeitos mais severos para amostras de tamanho reduzido (, para e , para ) e atua de maneira significativa na dispersão dos resultados de B\_SEN obtidos, sendo, para alguns casos, superiores a 25 vezes o modulo de *b*.

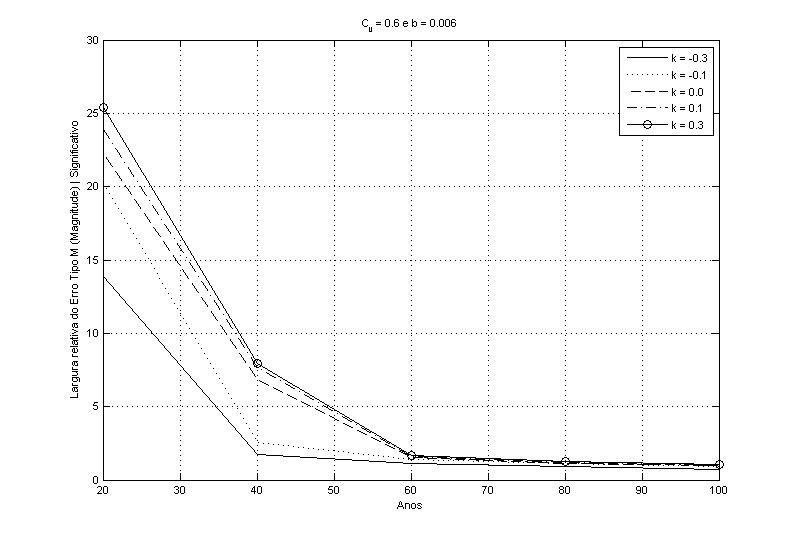


Figura - Largura relativa do IC do Erro tipo M dado que a estação foi declarada não-estacionária (alpha = 0.05) para *Cv* = 0.6, *b* = 0.006 e diferentes valores de *κ* =[-0.3 -0.1 0 0.1 0.3] e *n* = [20 40 60 80 100].

# Considerações Finais

A probabilidade de se estimar a magnitude da tendência com o sinal trocado (erro Tipo S) empregando o estimador proposto por Sen (1968), muito utilizado em estimativas de magnitudes de mudança, não é desprezível. Por exemplo, verificou-se que a probabilidade do erro Tipo S pode variar de 5% a 45% para valores de poder do teste inferiores a 10%, valor esse observado quando a série é distribuída por uma distribuição Gumbel (*κ* = 0), com 30 anos de dados, *Cv* = 1, e tendência monotônica de 10% por década em relação à média da série. Esse fato já havia sido evidenciado por Gelman e Carlin (2014), que apontaram para chances próximas de 40% ou mais de acometimento de erros tipo S para situações de resultado significativos localmente no nível de significância de 5%. Os mesmos autores pontuam ainda que a probabilidade de um erro tipo S não é a mesma que *o valor p* ou nível de significância, o que é desconhecido até mesmo por usuários avançados de técnicas estatísticas.

A estimativa de magnitude de tendência com sinal trocado pode causar interpretações equivocadas sobre os resultados de análise de tendência e pode resultar em decisões equivocadas sobre investimentos em adaptação. Ainda em relação à magnitude da tendência, os resultados da simulação Monte Carlo apontam que o estimador proposto por Sen (1968) tende a superestimar a magnitude (erro Tipo M) quando o poder do teste MK é baixo. O valor estimado da magnitude da tendência, em média, pode ser 1,5 vezes maior que o valor real quando o poder do teste é igual a 0,40, podendo ainda ser até 5 vezes maior do que o valor real, em média, quando o poder do teste é igual a 0,05.

Esses resultados mostram também que, em séries com alta variabilidade anual e com magnitude de tendência relativamente pequena – características que resultam em baixo poder de um teste estatístico (depende também do parâmetro de forma *κ* da GEV) –, a probabilidade de se obter uma estimativa da tendência com sinal trocado é relativamente alta para amostras com tamanhos geralmente constatados para as estações brasileiras, normalmente inferiores a 60 anos

Referências

AghaKouchak, A. *et al.* (Eds.), 2013. *Extremes in a Changing Climate: Detection, Analysis and Uncertainty*, Water in Science Technology Library, Volume 65. Springer, New York, NY.

ANA (Agência Nacional de Águas), 2016. Mudanças Climáticas e Recursos Hídricos: Avaliação e Diretrizes para Adaptação. Agência Nacional de Água, Brasília.

ASCE, 2018. *Climate-Resilient Infrastructure: Adaptive Design and Risk Management.* ASCE manuals and reports on engineering practice no. 140.  Reston, VA. American Society of Civil Engineers.

Ball, J. Babister, M., Nathan, R., Weeks, W., Weinmann, E., Retalick, E., Testoni, I. (2016). Australian Rainfall and Runof: A Guide to Flood Estimation.

Bayazit, M., Fernandez, B. & Salas, J., 2001. Return Period and Risk of Hydrologic Events. I: Mathematical Formulation. *Journal of Hydrologic Engineering*, 6, p.358.

Borgomeo, E. et al., 2018. Risk, Robustness and Water Resources Planning Under Uncertainty. *Earth's Future*, 6(3), pp.468–487.

Borgomeo, E., Hall, J. Fung, F., Watts, G., Colquhoun, K., Lambert, C. (2014). Risk-based water resources planning: incorporating probabilistic nonstationary climate uncertainties. *Water Resoures Reasearch*, 50, 6850-6873, doi: 10.1002/2014WR015558.

Brown, C. et al., 2011. A Decision-Analytic Approach to Managing Climate Risks: Application to the Upper Great Lakes1. *JAWRA Journal of the American Water Resources Association*, 47(3), pp.524–534.

Burn, D., Fan, L., Bell, G. (2008). Identification and quantification of streamflow trends on the Canadian Prairies, *Hydrological Sciences Journal*, 53:3, 538-549, DOI: 10.1623/hysj.53.3.538

Coch, A., Luis Mediero, L. (2016). Trends in low flows in Spain in the period 1949–2009, *Hydrological Sciences Journal*, 61:3, 568-584, DOI: 10.1080/02626667.2015.1081202

Craig, R. K. (2010). Stationarity Is Dead - Long live transformation: Five principles for climate change adaptation law. *Harvard Environ*, Rev*.*, 34 (1), pp. 9-75.

England, J.F., Jr., Cohn, T.A., Faber, B.A., Stedinger, J.R., Thomas, W.O., Jr., Veilleux, A.G., Kiang, J.E., and Mason, R.R., Jr. (2018). Guidelines for determining flood flow frequency—Bulletin 17C: U.S. Geological Survey Techniques and Methods, book 4, chap. B5, 148 p.

Franks, S.W. (2014). Flood risk in Eastern Australia: climate variability and change. In: Beven, K., Hall, J. (Eds.). Applied uncertainty analysis for flood risk management. Imperial College Press, London.

Galloway, G.E., 2011. If Stationarity is Dead, What Do We Do Now?1. *JAWRA Journal of the American Water Resources Association*, 47(3), pp.563–570.

He, H., Cheng, Y., Cheng, N., Lu, Y., Singh, S. (2017). Variability of Hydrological Processes and Systems in a Changing Environment. In: Vinjay, S. (Ed.), Handbook of Applied Hydrology, Second Edition. McGraw-Hill Education.

Hirsch, R.M., 2011. A Perspective on Nonstationarity and Water Management1. *JAWRA Journal of the American Water Resources Association*, 47(3), pp.436–446.

IPCC (2013). Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. ISBN 978-1-107-66182-0.

Kaatz, L., 2010. *Decision Support Planning Methods: Incorporating Climate Change Uncertainties into Water Planning*,

Kiang, J., Olsen, J., Waskom, R. (2011). Introduction to the Featured Collection on ‘‘Nonstationarity, Hydrologic Frequency Analysis, and Water Management.’’ *Journal of the American Water Resources Association* (JAWRA) 47(3):433-435. DOI: 10.1111/j.1752-1688.2011.00551.x

Koutsoyianis, D., Montanari, A. (2014). Negligent killing of scientific concepts: The stationarity case. *Hydrol. Sci. J.*, 60, pp. 1174–1183.

Kundzewicz, Z.W., 2011. Nonstationarity in Water Resources - Central European Perspective1. *JAWRA Journal of the American Water Resources Association*, 47(3), pp.550–562.

Lins, H. F., Cohn, T. A. (2011). Stationarity: Wanted dead or alive? *J. Am. Water Resour. Assoc.*, 47, pp. 475–480.

Mangini, W., Viglione, A., Hall, J., Hundecha, Y., Ceola, S., Montanari, A., Rogger, M., Salinas, J.L., Borzì, I., Parajka, J. (2018). Detection of trends in magnitude and frequency of flood peaks across Europe, *Hydrological Sciences Journal*, 63:4, 493-512, DOI: 10.1080/02626667.2018.1444766

McPhail, C. et al., 2018. Robustness Metrics: How Are They Calculated, When Should They Be Used and Why Do They Give Different Results? *Earth's Future*, 6(2), pp.169–191.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M.; Kundzewicz, Z. W.; Lettenmaier, D. P., Stouffer, R. J., Dettinger, M. D., Krysanova, V. (2015). On Critiques of ‘‘Stationarity is Dead: Whither Water Management?’’. *Water Resour. Res.,* 51, pp. 7785–7789.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M.; Kundzewicz, Z. W.; Lettenmaier, D. P., Stouffer, R. J. (2008). Stationarity is dead: Whiter water management? *Science*, 319, pp. 573–574.

Mondal, A., Daniel, D., 2019. Return Levels under Nonstationarity: The Need to Update Infrastructure Design Strategies. *Journal of Hydrologic Engineering*, 24 (1): 04018060.

Montanari, A, Young, G., Savenije, H., Hughes, D., Wagener, T., Ren, L., Koutsoyiannis, D, et al. 2013. Panta Rhei—Everything Flows’: Change in Hydrology and Society—the IAHS Scientific Decade 2013–2022. *Hydrological Sciences Journal.* 58 (6): 1256–75. doi:10.1080/02626667.2013.809088.

NAS (The National Academies of Science, Engineering and Medicine), 2018. *Future Water Priorities for the Nation: Directions for the U.S. Geological Survey Water Mission Area*, Washington, D.C.: National Academies Press.

NERC (1975). Flood Studies Report. Natural Environment Research Center.

Obeysekera, J. & Salas, J.D., 2014. Quantifying the Uncertainty of Design Floods under Nonstationary Conditions. *Journal of Hydrologic Engineering*, 19(7), pp.1438–1446.

Prosdocimi, I., Kjeldsen, T.R. & Svensson, C., 2014. Non-stationarity in annual and seasonal series of peak flow and precipitation in the UK. *Natural Hazards and Earth System Science*, 14(5), pp.1125–1144.

Rahman, A., Haddad, K., Kuczera, G., Weinmann, E. (2016). Regional Flood Methods. In: Commmonwealth of Australia, *Australian Rainfall and Runoff: A Guide to Flood Estimation* (Chapter 3 in Book 3). Retrived from http://book.arr.org.au.s3-website-ap-southeast-2.amazonaws.com.

Ray, P.A. et al., 2018. Growth of the Decision Tree: Advances in Bottom-Up Climate Change Risk Management. *JAWRA Journal of the American Water Resources Association*, 26(12), pp.1376–18.

Razavi, S. & Vogel, R., 2018. Prewhitening of hydroclimatic time series? Implications for inferred change and variability across time scales. *Journal of Hydrology*, 557, pp.109–115.

Read, L.K. & Vogel, R.M., 2015. Reliability, return periods, and risk under nonstationarity. *Water Resources Reasearch*, 51(8), pp.6381–6398.

Read, L.K. & Vogel, R.M., 2016. Hazard function theory for nonstationary natural hazards. *Natural Hazards and Earth System Science*, 16(4), pp.915–925.

Robson, A., Reed, D. (1999). Flood Estimation Handbook. Volume 3. Institute of Hydrology, Wallinford.

Rootzén, H. & Katz, R.W., 2013. Design Life Level: Quantifying risk in a changing climate. *Water Resources Research*, 49(9), pp.5964–5972.

Rosbjerg, D., 2017. Optimal adaptation to extreme rainfalls in current and future climate. *Water Resources Research*, 53(1), pp.535–543.

Rosner, A., Vogel, R.M. & Kirshen, P.H., 2014. A risk-based approach to flood management decisions in a nonstationary world. *Water Resources Research*, 50(3), pp.1928–1942.

Salas, J.D., Obeysekera, J. (2014). Revisiting the Concepts of Return Period and Risk for Nonstationary Hydrologic Extreme Events. *Journal of Hydrologic Engineering*, 19(3), pp.554–568.

Salas, J.D., Obeysekera, J. & Vogel, R.M., 2018. Techniques for assessing water infrastructure for nonstationary extreme events: a review. *Hydrological Science Journal*, 63(3), pp.325–352.

Schlef, K.E. et al., 2018. A General Methodology for Climate-Informed Approaches to Long-Term Flood Projection-Illustrated With the Ohio River Basin. *Water Resources Research*, 54(11), pp.9321–9341.

Serago, J.M. & Vogel, R.M., 2018. Parsimonious nonstationary flood frequency analysis. *Advances in Water Resources*, 112, pp.1–16.

Serinaldi, F., Kilsby, C. G. (2015). Stationarity is undead: Uncertainty dominates the distribution of extremes. *Adv. Water Resour.*, 77, pp. 17–36.

Serinaldi, F., Kilsby, C.G. & Lombardo, F., 2018. Untenable nonstationarity\_ An assessment of the fitness for purpose of trend tests in hydrology. *Advances in Water Resources*, 111, pp.132–155.

Spence, C.M. & Brown, C.M., 2016. Nonstationary decision model for flood risk decision scaling. *Water Resources Research*, 52(11), pp.8650–8667.

Stedinger, J.R. & Crainiceanu, C.M., 2001. Climate Variability and Flood-Risk Management. In: Proceedings of the Conference on Risk-Based Decisionmaking in Water Resources, October 15-20, 2000, Santa Barbara, California. Reston, VA: American Society of Civil Engineers, pp. 77–86. Doi: [10.1061/40577(306)7](https://doi.org/10.1061/40577(306)7).

Stedinger, J.R. & Griffis, V.W., 2011. Getting From Here to Where? Flood Frequency Analysis and Climate1. *JAWRA Journal of the American Water Resources Association*, 47(3), pp.506–513.

Sveinsson, O.G.B. et al., 2003. Modeling the Dynamics of Long-Term Variability of Hydroclimatic Processes. *Journal of Hydrometeorolgy,* Vol. 4, pp. 489-505.

Vogel, R.M., Castellarin, A., 2017. *Risk, Reliability and Return Periods and Hydrologic Design*. In: Vijay P. Singh (Ed.), Handbook of Applied Hydrology. Second Edition. McGraw-Hill Education, New York, NY.

Vogel, R.M., Rosner, A. & Kirshen, P.H., 2013. Brief communication “Likelihood of societal preparedness for global change.” *Natural Hazards and Earth System Sciences Discussions*, 1(1), pp.1–13.

Vogel, R.M., Yaindl, C., Walter, M., (2011). Nonstationarity: Flood Magnification and Recurrence Reduction Factors in the United States. *JAWRA Journal of the American Water Resources Association*, 47(3), pp.464–474.

Waage, M.D. & Kaatz, L., 2011. Nonstationary Water Planning: An Overview of Several Promising Planning Methods1. *Journal of the American Water Resources Association*, 47(3), pp.535–540.