Avaliação de Risco Stoch Environ Res DOI 10.1007/s00477-015-1193-3

#### REVEJA



# Aplicação de processo autorregressivo periódico à modelagem das vazões do rio Garonne

Eugen Ursu1 • Jean-Christophe Pereau1

Springer-Verlag Berlim Heidelberg 2015

Resumo A previsão precisa das vazões dos rios é uma das aplicações mais importantes em hidrologia, especialmente para o gerenciamento de sistemas de reservatórios. Para capturar as variações sazonais nas estatísticas de vazão do rio, este artigo desenvolve uma abordagem de modelagem robusta para identificar e estimar o modelo periódico autorregressivo (PAR) na presença de valores discrepantes aditivos. Como os estimadores de mínimos quadrados não são robustos na presença de outliers, sugerimos uma estimativa robusta baseada em autocovariâncias residuais. Um algoritmo genético com critério de informação de Bayes é utilizado para identificar o modelo PAR ótimo. O método é aplicado a dados de vazão média mensal e trimestral (1959-2010) para o rio Garonne no sudoeste da França. Os resultados mostram que a precisão das previsões é melhorada no modelo robusto em relação ao modelo não robusto para os fluxos trimestrais. Ao reduzir o número de parâmetros a serem estimados, o princípio da parcimônia favorece a escolha da abordagem robusta.

Palavras-chave Análise de vazões de rios Séries periódicas Estimação robusta Algoritmos genéticos

#### 1. Introdução

Nos últimos anos tem havido considerável pesquisa no desenvolvimento de modelos de séries temporais com propriedades sazonais ou periódicas em hidrologia e recursos hídricos (Vecchia 1985a). Esses modelos são importantes para o planejamento e

& Eugen Ursu eugen.ursu@u-bordeaux.fr

GREThA UMR-CNRS 5113, Universidade de Bordeaux, Avenue Le´on Duguit, 33608 Pessac CEDEX, França o desenho de políticas de gestão da água. Eles visam fornecer previsões precisas ao gestor do rio para alocar o recurso entre usuários com interesses conflitantes, por exemplo. Essas previsões podem ser obtidas usando uma abordagem de variável única que não precisa modelar os processos físicos do sistema hidrológico (Hipel e McLeod 1994). Quando tal modelo hidrológico existe, modelos de séries temporais podem ser usados para analisar a diferença entre os dados observados e as previsões do modelo hidrológico como em Madsen e Skotner (2005) e Gragne et al. (2014).

Em ambos os casos, há a necessidade de se encontrar um procedimento automático para a identificação e a estimativa dos fatores mais relevantes. modelos de séries temporais vantajosas.

Modelos de séries temporais sazonais, como o autor sazonal

O modelo de média móvel integrada progressiva (SARIMA) desenvolvido
originalmente por Box e Jenkins (1970, cap. 9) tem sido amplamente
estudado na literatura de modelos hidrológicos. Mishra e Desai (2005)
aplicaram com sucesso os modelos ARIMA e ARIMA sazonal para prever
secas com base nos estágios usuais de desenvolvimento do modelo:
identificação, estimativa e verificação de diagnóstico. Fernandez et ai.
(2008) aplicaram um modelo multiplicativo SARIMA para prever o fluxo
mensal em uma pequena bacia hidrográfica no noroeste da Espanha.
Durdu (2010) usou um modelo SARIMA para prever a seca na bacia do rio
Bu"y"k Menderes usando o índice de precipitação padronizado (SPI) como
índice de seca.

No entanto, como apontado por McLeod (1993), os fluxos dos rios para uma determinada estação do ano podem ser estatisticamente semelhantes de ano para ano, mas podem depender intrinsecamente da estação. Esta característica não pode ser capturada por modelos SARMA que representam uma classe de modelos estacionários com grandes autocorrelações de defasagem que são invariantes em relação à estação. Além disso, muitas séries temporais sazonais não podem ser filtradas para alcançar estações de segunda ordem

aridade devido à estrutura de correlação dessas séries temporais

Published online: 05 December 2015

com a estação (Vecchia 1985b). Como a maioria das séries temporais de vazões dos rios satisfazem a propriedade de estacionaridade periódica, significando que suas funções de média e covariância são periódicas em relação ao tempo, justifica-se o uso da modelagem periódica autorregressiva (PAR) em recursos hídricogodelos de séries temporais pode ser encontrada em Maronna et Claramente, um modelo PAR encapsula um modelo autorregressivo separado para cada estação do ano, que são potencialmente diferentes. O modelo PAR também pode ser visto como um caso particular da classe de modelos de média móvel PAR (PARMA) (Jimenez et al. 1989).

O método dos momentos baseado nas equações de Yule-Walker (McLeod 1994) e o método dos mínimos quadrados (LS) no caso univariado (Franses e Paap 2004) são eficientes para estimar modelos PAR. Conforme mencionado por Hipel e McLeod (1994), quando se utilizam os dados sazonais e o modelo para cada estação ao invés dos dados anuais e do modelo associado, pode-se obter ganhos significativos na eficiência dos parâmetros. No entanto, o principal problema na modelagem do PAR depende do número de parâmetros a serem estimados, que varia com a escolha da estação para as vazões dos rios. Passar de levantamentos mensais para dados trimestrais de vazão do rio aumenta tanto o número de modelos quanto o número de parâmetros a serem estimados. Para obter modelos parcimoniosos, é interessante estudar situações em que são introduzidas restrições lineares nos parâmetros de uma determinada estação do ano (Ursu e Duchesne 2009). Também justifica o uso de algoritmo genético com critério de informação de Bayes (BIC) para identificar a ordem ótima do modelo PAR. Ferramentas como algoritmo genético (Koutroumanidis et al. 2009), máquina vetorial de suporte e programação genética (Wang et al. 2014) ou rede neural artificial (Aksoy e Dahamsheh 2009) aparecem como alguns dos inúmeros métodos modernos de simulação e otimização em hidrologia . Uma visão geral desses métodos de simulação e otimização usados na operação de reservatórios é apresentada em Fayaed et al. (2013).

Um segundo problema na estimação de parâmetros de modelos de séries temporais ocorre com a presença de outliers que podem implicar em sérios problemas. Hau e Tong (1989) apresentam uma nova abordagem para detecção de outliers em modelos autorregressivos com dados reais de recursos hídricos. Eles também apresentam dois tipos de outliers que são geralmente discutidos na literatura de séries temporais: inovação e outliers aditivos. Assim como em Ursu e Pereau (2014), este artigo foca em outliers aditivos, uma vez que seu impacto nas estimativas de parâmetros é conhecido por ser mais importante do que outliers de inovação (Li 2004, cap. 4). Em particular, a sensibilidade do método de estimação LS a outliers requer o uso de abordagens robustas (Denby e Martin (1979) para modelos autorregressivos de primeira ordem; Ben et al. (1999) para modelos vetoriais de média móvel autorregressiva (VARMA); Shao (2007) para modelos PAR univariados). Com base no estimador de escala robusta desenvolvido por Ma e Genton (2000), Sarnaglia et al. (2010) analisa suas propriedades em modelos PAR. Dentro

Modelos periódicos de vetor autorregressivo (PVAR), Ursu e Pereau (2014) implementam um método de estimação robusto baseado em autocovariâncias residuais (RA) para lidar com outliers aditivos. Uma extensa revisão sobre o impacto de outliers em al. (2006, cap. 8).

A identificação do modelo é normalmente o aspecto mais difícil do procedimento de construção do modelo. O exame dos gráficos da função de autocorrelação parcial periódica como ferramenta para identificar um modelo periódico requer um alto nível de experiência do usuário (ver, por exemplo, Jimenez et al. 1989). Outra abordagem é usar um critério de seleção automática, como o critério de informação Akaike (AIC) ou o critério de informação Bayesiano (BIC) quando todos os modelos possíveis são examinados. Uma possível dificuldade com este procedimento é o grande número de modelos a serem investigados. A contribuição do artigo para a literatura teórica é fornecer uma metodologia automática capaz de selecionar e estimar modelos PAR na presença de outliers. Além disso, do ponto de vista operacional, este artigo mostra que tal metodologia pode ser aplicada com sucesso em uma escala temporal trimestral mensal mostrando que as previsões de uma etapa com o modelo robusto são sempre melhores do que o modelo de abordagem não robusta com vários critérios de eficiência. Essa abordagem mostra que o número de parâmetros a serem estimados também é menor com o modelo robusto

Este artigo está organizado da seguinte forma. Na Sec. 2, o modelo PAR é introduzido e os estimadores de mínimos quadrados são calculados. Na Sec. 3, uma estimativa robusta na presença de outliers é desenvolvida. A seção 4 ilustra os resultados do estudo de caso do rio Garonne. A Seção 5 oferece algumas considerações finais.

#### 2 modelos periódicos

A classe de modelos PAR estende a classe de modelos autorregressivos (AR), permitindo que os parâmetros autorregressivos variem com as estações do ano. Vale ressaltar que um modelo PAR é formado pela definição de um modelo AR diferente para cada estação do ano. Um modelo PAR com 12 períodos pode ser associado a 12 modelos AR. Note-se que, quando o número de períodos é 1, o modelo PAR passa a ser modelo AR.

Seja Y 1/4 fYt; t 2 Zg ser um processo estocástico PAR dado por

Ynsþm ¼ XpðmÞ / kðmÞYnsþmk þ nsþm: ð1Þ

Para m fixo e valores predeterminados s, a variável aleatória Ynsþm representa a realização durante a m-ésima estação, com m 2 f1; ...;sg, no ano n, n 2 Z. Com dados mensais, o

o valor m ½ 12 é naturalmente selecionado, enquanto que para dados mensais trimestrais m ½ 48. A ordem do modelo autorregressivo na estação m é dada por põmÞ, enquanto /kðmÞ, k ½ 1; ...; pðmÞ, representam os coeficientes do modelo autorregressivo durante a estação m, m ½ 1; ...;s. O processo de erro ½ f t; t2Zg na Eq. (1) corresponde a um ruído br**pació**dico, com Eð tÞ ½ 0 e varð nsþmÞ ½ r2ðmÞ [ 0, m ½ 1; ...;s. O processo aleatório Yt em (1) deve ter média zero.

Salvo indicação em contrário, assumimos que os modelos PAR são estacionários no sentido periódico. A estacionariedade periódica é discutida em Gladyshev (1961). Normalmente, os modelos periódicos usados em recursos hídricos e sistemas ambientais são estacionários, no sentido de que não precisam ser diferenciados para alcançar a estacionaridade (ou, dito de outra forma, os dados não têm raízes unitárias). Em aplicações, as médias sazonais são primeiro removidas da série temporal.

#### 2.1 Identificação e estimativa para modelos PAR

Esta seção resume sem provas o material relevante sobre identificação e estimativa de parâmetros para modelos PAR. Referências que fornecem provas detalhadas estão incluídas no texto.

Várias técnicas de estimação estão disponíveis para modelos periódicos, nomeadamente o método dos mínimos quadrados (Franses e Paap 2004; Lu tkepohl 2005), o método dos momentos baseado na estimação de Yule-Walker (Pagano 1978; Hipel e McLeod 1994), a máxima verossimilhança estimativa (Vecchia 1985a; Eshete e Vandewiele 1992) e o filtro de Kalman (Jimenez et al. 1989).

Considere os dados da série temporal Ynsþm, n  $\frac{1}{4}$  0; 1; ...; N 1, m  $\frac{1}{4}$  1; ...;s com tamanho de amostra n  $\frac{1}{4}$  Ns. Deixar

sejam matrizes aleatórias N 1, N 1 e N pômÞ. Ao definir o vetor pômÞ 1 bômÞ dos parâmetros como:

o modelo PAR pode ser escrito da seguinte forma:

Da Eq. (3), os estimadores de mínimos quadrados (irrestritos e restritos) de bomp podem ser facilmente encontrados. Para mais detalhes, remetemos o leitor para Ursu e Turkman

(2012, sec. 2). Uma vez o estimativas b^ðmÞ ¼

/^ 1ðmÞ; ...; /^ pðmÞðmÞ são obtidos, os resíduos do modelo
são apresentados:

que são bem definidos para n = 0; 1; ...; N 1.

Vários critérios de seleção usando AIC ou BIC podem ser usados para identificação do modelo PAR. Uma maneira possível é usar o critério de seleção BIC separadamente para cada um dos componentes sazonais:

$$\frac{\log \delta N + \log r \wedge 2 \delta m}{\log r \wedge 2 \delta m} > \rho \rho \delta m + \log r \wedge 2 \delta m$$

onde r^ômÞ representa os estimadores de mínimos quadrados de rômÞ, e pômÞ representa o número de parâmetros autorregressivos na estação m McLeod (1994).

Mesmo que este método reduza o número de modelos a serem investigados, o número de modelos possíveis permanece muito alto. O grande número de soluções possíveis para o modelo de seleção PAR sugere que os algoritmos genéticos (AG) podem ser úteis para um exame eficiente do espaço de soluções e seleção da combinação de parâmetros que corresponde ao melhor modelo. O GA combinado com o critério BIC é uma maneira confiável e fácil de identificar modelos PAR (Ursu e Turkman 2012).

Resumimos brevemente nosso procedimento GA para modelagem PAR de subconjunto.

- Representação de string Cada modelo AR de subconjunto é codificado como uma string, cada locus na string é preenchido com 1 se o parâmetro for livre e com 0 se o parâmetro for restrito a zero. Uma vez que uma ordem de pesquisa máxim a ser selecionada, cada corda tem o mesmo comprimento L. Por exemplo, se tomarmos s ¼ 12, m ¼ 1 e põ1Þ ¼ 15, e o modelo

Y12nþ1 ¼ / 6ð1ÞY12n5 þ / 7ð1ÞY12n6 þ 12nþ1

então, a string que representa nosso modelo é 0000011000000000:

Observe que, neste caso, o número de todos os modelos possíveis é 12 215 ¼ 393216.

- População inicial Uma população arbitrária de cromossomas de tamanho
   Np é gerada. Cada cromossomo é codificado como uma cadeia binária de comprimento L, conforme descrito acima. O tamanho da população
   Np e o comprimento do cromossomo L são escolhidos pelo investigador.
- Uma função de aptidão Cada cromossomo é avaliado por meio de uma função de valor real positiva chamada função de aptidão. Como o BICômÞ pode ser negativo, uma

candidato para a função de aptidão é uma transformação exponencial

fjðmÞ ¼ exp fBlCjðmÞ = dg;

onde BICjômÞ representa o valor BICômÞ para o cromossomo jth na população atual e d é uma constante de escala. Para outra função de aptidão apropriada, remetemos o leitor para Gaetan (2000).

 Gerando uma nova população Uma nova população de cromossomos potenciais é criada, utilizando operadores evolutivos como: seleção, cruzamento e mutação. Este ciclo continua até que o número máximo de gerações Ng seja atingido, ou até que uma condição de parada seja alcançada.

Para muitas variações do AG básico e explicações detalhadas, ver Goldberg (1989), Mitchell (1996), barragem de Sivanan e Deepa (2008).

#### 2.2 Previsão com modelo PAR

A previsão com modelos PAR procede da mesma forma que os modelos AR padrão. O objetivo é obter uma previsão com o menor erro possível, levando à previsão de erro quadrático médio mínimo (MMSE). A previsão do MMSE é dada por sua expectativa condicional (Hipel e McLeod 1994).

Assumindo que as observações e inovações são conhecidas até o n-ésimo ano e m-ésima estação, toma-se a expectativa condicional da eq. (1) para obter a previsão do MMSE Y^nsþmðlÞ, onde Y^nsþmðlÞ é interpretado como a previsão l-passo à frente na origem da previsão t ¼ ns þ m. Por exemplo, a previsão de 1 passo à frente feita na origem t ¼ ns þ m é Y^nsþmð1Þ ¼ E Y½ nsþmþ1jYnsþm; Ynsþm1; ... ¼ /1 ðmÞYnsþm þ /2ðmÞYnsþm1 þ þ /pðmÞYnsþmpþ1

A representação causal dos modelos PAR (Uesu e Duchesne 2009, Eq. 5) pode ser usada para calcular intervalos de confiança para previsões, mas isso está além do escopo deste artigo. A melhor referência geral para intervalos de confiança em modelos periódicos são Hipel e McLeod (1994, cap. 15) e Anderson et al. (2013). O desempenho de previsão de vários modelos de séries temporais usados na análise de vazão de rios é apresentado em Noakes et al. (1985). Os resultados sugerem que os modelos PAR fornecem as previsões mais precisas.

#### 3 Modelagem robusta de modelos PAR

Como é bem conhecido, os métodos de estimação podem ser seriamente afetados na presença de outliers aditivos que são os maiores erros de registro (Bustos e Yohai 1986; Shao 2007). Outliers aditivos referem-se a um processo PAR com probabilidade 1 x e um processo PAR mais um erro com probabilidade x. A ocorrência de outliers é geralmente pequena (x 0:05).

Estimadores robustos baseados em autocovariâncias robustas para modelos ARMA foram propostos por Bustos e Yohai (1986). Sua metodologia foi estendida para modelos PAR multivariados por Ursu e Pereau (2014). O sistema de equações obtido em Ursu e Pereau (2014, Eqs. (6) e (9)) pode ser facilmente adaptado para processos PAR.

Portanto, para reduzir a influência dos resíduos suspeitos de serem discrepantes, os resíduos ^nsþm definidos na Seç. 2 são substituídos por seus resíduos modificados ~nsþm definidos como:

onde w representa uma função ímpar e limitada e r^ômÞ é um estimador robusto para rômÞ. Uma escolha usual para a função w é a função de Huber:

wH;kðxÞ ¼ sgnðxÞ min fjxj; kg;

onde k é uma constante e sgnðxÞ é a função signum. Geralmente, um algoritmo iterativo é proposto para modelos ARMA em Bustos e Yohai (1986) e para modelos PAR em Ursu e Pereau (2014).

Em todas as simulações não tivemos problemas de convergência com o algoritmo iterativo como em Ursu e Pereau (2014). Outros trabalhos recentes sobre robustez em séries temporais periódicas incluem o estimador de modelos PAR propostos por Sarnaglia et al. (2010). Além disso, uma estimativa robusta para modelos PAR foi discutida em Shao (2007).

#### 4 Estudo de caso: o rio Garonne

O modelo PAR é aplicado às vazões fluviais médias mensais e às vazões fluviais médias trimestrais do rio Garonne, localizado no sudoeste da França. Esse rio

desce desde a sua nascente em Espanha até ao Oceano Atlântico ao longo de 647 km. É o principal contribuinte para o Gironde Estuário que é o principal sistema fluvial-estuarino europeu. O rio Garonne é o terceiro maior rio da França em termos de vazão, com uma área de captação de 51.500 km2 em Tonneins, a saída da bacia hidrográfica.

O rio Garonne é administrado pela Adour-Garonne Water Agency, responsável pelo planejamento e desenho de políticas hídricas relacionadas a várias questões como abastecimento urbano de água, energia hidrelétrica, gestão de irrigação, controle de enchentes e secas, poluição, proteção de peixes ameaçados, áreas úmidas e conservação de habitats (Baker e Vervier 2004; Muylaert et al. 2009; Oeurng et al. 2011; Maire et al.

2013). As autoridades fluviais visam controlar a gestão quantitativa da água para alcançar um bom estado da água de acordo com os requisitos ecológicos impostos pela Diretiva-Quadro Europeia da Água. Para garantir o bom funcionamento do sistema econômico e ecológico, a Agência Adour-Garonne estabelece em diferentes estações de medição do rio alguns valores mínimos de limiar medidos em metros cúbicos por segundo (m3/s) para as vazões do rio. Esses limites devem ser alcançados todos os anos durante os períodos de maré baixa entre 1º de julho e 31 de outubro. Para cumprir esses objetivos de vazão, políticas de armazenamento de água que consistem em medidas de economia de água e lançamentos de água de barragens e reservatórios devem ser implementadas a montante das estações de medição. Em um contexto de maior pressão hídrica aumentada pela mudança climática (Caballero et al. 2007, Boe et al. 2009), Tisseuil et al. (2010) mostram redução significativa nos fluxos fluviais em diferentes estações de medição do rio Garonne e em particular em Tonneins. Com base em um modelo de chuva-vazão acoplado a um sistema de gestão de reservatórios. Hendrickx e Sauguet (2013) mostram que será necessário o enchimento antecipado dos reservatórios no inverno para garantir que os lançamentos de água durante os períodos quentes cumpram os objetivos de vazão.

As medidas de vazão são registradas na última estação de medição do rio Garonne em Tonneins. Os dados são obtidos a partir de medições diárias de vazão em metros cúbicos por segundo (m3/s) de janeiro de 1959 a dezembro de 2010 (DIREN Banque Hydro, monitoramento de água francês). Os fluxos de dados diários são então transformados em dados mensais, respectivamente dados mensais trimestrais, consistindo em fluxos médios para um mês, respectivamente do 1º ao 7º, do 8º ao 15º, do 16º ao 22º, e do 23º ao final do mês como em Hipel e McLeod (1994). Em Tonneins, o valor limite do caudal diário é igual a 110 m3/s para a estação seca entre 1 de julho a 31 de outubro que corresponde aos períodos 25 a 40.

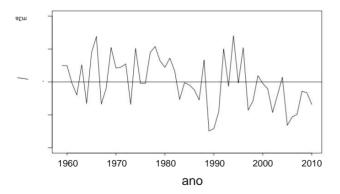


Fig. 1 Gráfico de vazão média anual do rio Garonne em Tonneins entre 1959 e 2010 em metros cúbicos por segundo (m3/s)

A Figura 1 mostra os fluxos anuais do rio Garonne entre 1959–2010. Ele mostra que vários episódios de seca severa ocorreram em 1989–1990 e em 2005. Esta figura também mostra que os fluxos anuais permaneceram abaixo da média de 600 m3/s por vários anos durante a última década.

Para capturar o padrão periódico das vazões dos rios, são analisadas as séries de vazões médias mensais e trimestrais. As Tabelas 1 e 2 mostram, respectivamente, a média amostral, mediana e desvio padrão para cada série de fluxo. Um gráfico parcial de pesquisas mensais e trimestrais entre 1980 e 2000 é dado na Fig. 2.

A Figura 2 mostra que uma flutuação periódica nas médias e variâncias é claramente exibida. Para dados mensais, os fluxos dos rios são mais altos em fevereiro e muito mais baixos em agosto. Para dados trimestrais, a maioria dos picos ocorreu durante o 4º, 5º, 6º e 7º período e a maioria dos vales ocorreu no 30º, 31º, 32º e 33º período. A vazão média trimestral máxima foi observada no 48º período em 1959 como 4059 m3/s, enquanto uma vazão média trimestral mínima de 55,86 m3/s foi registrada no 27º período em 2003. Os períodos mais secos para dados mensais e trimestrais estão em agosto (8º mês) e no período 31, respectivamente, enquanto fevereiro (2º mês) e o 4º período são os períodos mais úmidos para dados mensais e trimestrais.

As Figuras 3 e 4 mostram o padrão da série no período 1959-2010. Para os períodos mais secos, observamos que a média móvel de 10 anos aumentou de 1959 a 1975, depois se estabilizou entre 1975 e 1980, seguida de uma queda de 1980 a 1990. Em 1990, a média móvel de 10 anos ficou em torno de 110 m3/s e parece permanecer estável. Um comportamento semelhante é observado para os períodos mais úmidos. As mudanças nas condições hidrológicas médias nas vazões do rio Garonne podem ser explicadas por vários fatores relacionados a mudanças naturais após as secas severas de 1989-1990 ou atividades humanas com aumento da agricultura irrigada e crescimento urbano. Desde meados da década de 1990, a estabilização da média móvel de 10 anos em torno do valor de 110 m3/s pode ser explicada pela implementação de políticas de gestão da água.

A Figura 5 mostra quantos períodos as vazões trimestrais permanecem abaixo do valor limite de 110 m3/s.

Entre 1959 e 2010, ocorre com maior frequência nos períodos 28 a 33, com um máximo de 20 vezes no período 31.

O último ano do conjunto de dados (12 meses ou 48 observações) foi omitido para a seleção e estimativa do modelo periódico. Para garantir que os resíduos do modelo ajustado sejam aproximadamente normalmente distribuídos e homocedásticos, uma transformação Box-Cox foi usada d (Eshete e Vandewiele 1992; McLeod e Gweon 2013). A análise Box-Cox mostra que os resíduos são

Tabela 1 Média da amostra, mediana e desvio padrão do

série de vazão média mensal

| Período | Quer dizer | Mediana | SD     | Período | Quer dizer | Mediana | SD     |
|---------|------------|---------|--------|---------|------------|---------|--------|
| 1       | 927,89     | 854,70  | 493,66 | 7       | 258,06     | 235,60  | 159,46 |
| 2       | 984,70     | 876,25  | 477,15 | 8       | 161,63     | 146,85  | 81,52  |
| 3       | 824,19     | 810,80  | 321,06 | 9       | 212.13     | 175,10  | 118,54 |
| 4       | 840,39     | 808,65  | 331,90 | 10      | 344,47     | 274,75  | 226,29 |
| 5       | 798,71     | 756,15  | 320,76 | 11      | 524,22     | 438,10  | 322,96 |
| 6       | 527,83     | 482,65  | 249,42 | 12      | 823,28     | 721,80  | 600,96 |

O período 1 corresponde a janeiro. Todas as medidas estão em metros cúbicos por segundo (m3/s)

Tabela 2 Média da amostra, mediana e desvio padrão do

fluxo médio trimestral Series

| Período | Quer dizer | Mediana | SD     | Período | Quer dizer | Mediana | SD     |
|---------|------------|---------|--------|---------|------------|---------|--------|
| 1       | 898,87     | 738,21  | 749,83 | 25      | 334,77     | 283,57  | 198,61 |
| 2       | 864,82     | 682,19  | 574,70 | 26      | 306,38     | 236,38  | 273,73 |
| 3       | 907,50     | 822,93  | 536,55 | 27      | 227,25     | 207.21  | 125,45 |
| 4       | 1022,39    | 911,33  | 639,42 | 28      | 179,41     | 161,72  | 87,35  |
| 5       | 1001,22    | 795,86  | 675,40 | 29      | 166,54     | 145,43  | 106,39 |
| 6       | 998,22     | 778,56  | 608,97 | 30      | 164,21     | 145,62  | 84,26  |
| 7       | 995,64     | 833,93  | 559,70 | 31      | 150,34     | 130,29  | 72,73  |
| 8       | 938,86     | 795,71  | 507,46 | 32      | 164,31     | 151,00  | 92,44  |
| 9       | 843,37     | 754,29  | 388,77 | 33      | 171,39     | 157,36  | 91,53  |
| 10      | 778,47     | 707,50  | 361,49 | 34      | 179,45     | 156,81  | 86,64  |
| 11      | 809,78     | 695,36  | 494,55 | 35      | 226,81     | 173,57  | 199,56 |
| 12      | 861,10     | 722,28  | 463,38 | 36      | 267,63     | 177,00  | 221,72 |
| 13      | 810,89     | 793,29  | 387,19 | 37      | 288,85     | 212.07  | 233,54 |
| 14      | 821,18     | 686,50  | 433,57 | 38      | 316,05     | 276,94  | 222,85 |
| 15      | 829,67     | 764,93  | 371,26 | 39      | 352,91     | 285,21  | 210,20 |
| 16      | 894,76     | 782,81  | 469,93 | 40      | 406,41     | 299,56  | 327,58 |
| 17      | 867,98     | 707,86  | 458,46 | 41      | 441,59     | 308,57  | 382,77 |
| 18      | 835,76     | 777,50  | 407,20 | 42      | 515,26     | 367,38  | 352,86 |
| 19      | 768,96     | 678,14  | 379,08 | 43      | 550,07     | 470,36  | 399,15 |
| 20      | 735,05     | 683,44  | 349,19 | 44      | 582,86     | 462,31  | 403,36 |
| 21      | 630,39     | 596,71  | 258,53 | 45      | 779,49     | 533,71  | 678,62 |
| 22      | 612,22     | 506,50  | 387,94 | 46      | 835,27     | 563,38  | 756,48 |
| 23      | 484,57     | 399,00  | 256,52 | 47      | 829,46     | 579,00  | 680,99 |
| 24      | 391,56     | 351,75  | 196,47 | 48      | 841,88     | 610,22  | 735,66 |

Todas as medidas estão em metros cúbicos por segundo (m3/s)

mais simétrico quando os dados de vazão são transformados em logaritmos naturais.

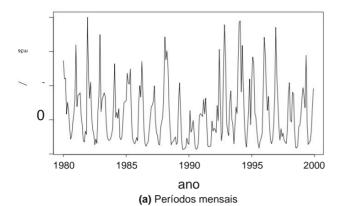
## 4.1 Modelagem não robusta de fluxos fluviais

Na modelagem de caso não robusto, os dados foram centralizados por subtraindo as médias sazonais da série logarítmica. O modelo PAR é ajustado à série truncada. o

número de modelos AR usados no modelo PAR é igual ao número de temporada associado à escolha do conjunto de dados. Isto oferece 12 modelos AR para dados mensais e 48 modelos para dados trimestrais. Observe que com o modelo mensal,

o número de modelos possíveis a serem estimados é de aproximadamente 4 105 e aumenta para 16 105 com dados mensais trimestrais. Este grande número de soluções possíveis sugerem o uso de técnicas de AG para reduzir o espaço de soluções e selecionar a combinação de parâmetros que dá o melhor modelo. Para cada estação, os parâmetros do modelo AR identificado são estimados usando o mínimos quadrados com restrições lineares.

Usando métodos GA, apenas 29 e 112 parâmetros têm foi estimado para os 12 e 48 modelos AR diferentes, respectivamente. O modelo mais complicado para fluxos logarítmicos mensais trimestrais é obtido para a 48ª temporada



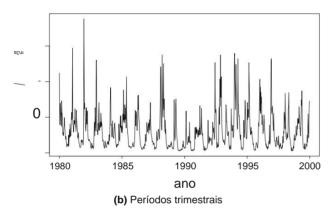


Fig. 2 Gráfico parcial dos fluxos médios mensais e trimestrais de Rio Garonne em Tonneins entre 1980 e 2000 em metro cúbico por segundo (m3/s)

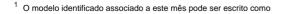
onde foi identificado um modelo AR com nove parâmetros.

Com fluxos logarítmicos mensais, o número de
parâmetros é seis para o mês de novembro.1 Um modelo PAR
parcimonioso pode ser obtido usando um único
modelo para todas as estações em um determinado grupo e, portanto, o
o número de parâmetros em um modelo PAR está diminuindo (Hipel e
McLeod 1994, cap. 14).

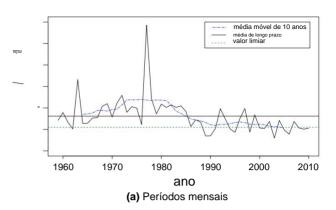
Os modelos propostos para Garonne são então usados para gerar previsões um passo à frente para fluxo logarítmico Series. Para obter as previsões no domínio não transformado tomamos a transformação inversa Box-Cox da frente

lançamentos no domínio transformado. A Figura 7 mostra a previsão e os dados observados para o último ano do conjunto de dados (2010). A Figura 1 mostra que a vazão do rio em 2010 é de cerca de 400 m3/s que é inferior à média de todo

amostra. Pode ser considerado como um ano seco.



onde /^ \_1 ½ 0:823, /^ \_4 ½ 0:518, /^ \_ \_5 ½ 0:444, /^ \_8 ½ 0:224, /^ \_ \_11  $^{^{\times}}$  0:146 e /^ \_ \_14 ½ 0:335.



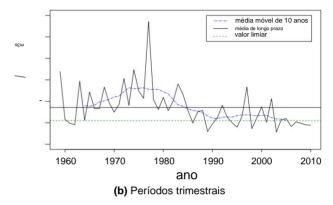
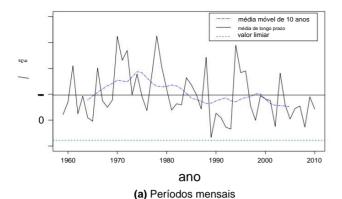


Fig. 3 Os fluxos mensais (agosto) e trimestrais mais secos (31 período) do rio Garonne em Tonneins de 1959 a 2010. A longa a média do prazo (para agosto) é 161,27 m3/s e 150,34 m3/s para o período 31.

# 4.2 Modelagem robusta de fluxos fluviais

A Figura 6 mostra que várias observações podem ser identificadas como outliers nos boxplots sazonais (de acordo com o e regra da metade do intervalo interquartil). Os forasteiros parecem ser mais numerosos em fluxos logarítmicos trimestrais do que em fluxos logarítmicos mensais no período mais úmido (de períodos 3 a 7), mas também no final dos períodos mais secos (dos períodos 35 a 42).

Ao contrário do caso pouco robusto, os dados foram centrados subtraindo as medianas sazonais do logarítmico séries em vez de médias sazonais. Como indicado por Shao (2007), as medianas sazonais são preferidas devido ao menor impacto dos outliers nas medianas. O procedimento robusto descrito no Sec. 3 é então aplicado. Conforme mencionado no Sec. 2, a melhor abordagem para identificar os parâmetros AR exigido em cada temporada para o modelo PAR é usar o técnicas de algoritmo genético. Vale ressaltar que o número de parâmetros estimados para mensal e fluxos logarítmicos trimestrais foram 23 e 108 respectivamente. O modelo AR mais complicado (por um período) implica cinco parâmetros (em novembro) e seis parâmetros (27º período) para o logarítmico mensal e dados trimestrais, respectivamente.



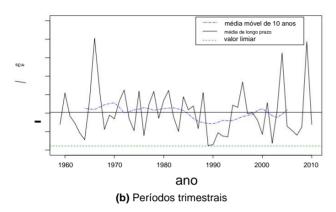


Fig. 4 Os fluxos mensais (fevereiro) e trimestrais mais úmidos (4º período) do rio Garonne em Tonneins de 1959 a 2010. O a média de longo prazo para fevereiro é 984,7 m3/s e 1022,39 m3/s para período 4

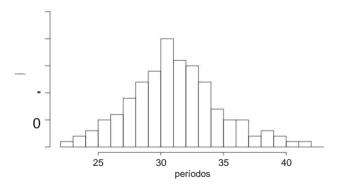
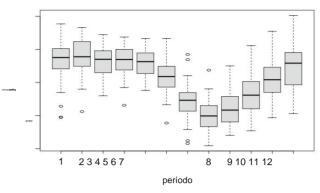


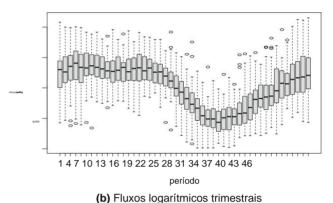
Fig. 5 Histograma dos fluxos trimestrais mais secos do rio Garonne em Toneladas entre 1959 e 2010

Os modelos robustos para dados de Garonne são usados para gerar previsões de um passo à frente para a série de fluxo médio. Uma vez novamente a transformação inversa Box-Cox foi aplicado. Previsões robustas e os dados observados para 2010 são mostrados na Fig. 7.

A precisão de previsão dos modelos não robustos e robustos é avaliada em relação às seguintes medidas: a erro quadrático médio (RMSE), o erro absoluto médio



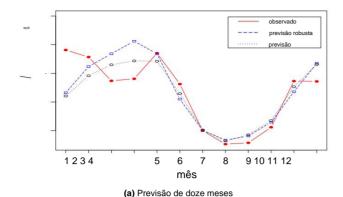
### (a) Fluxos logarítmicos mensais

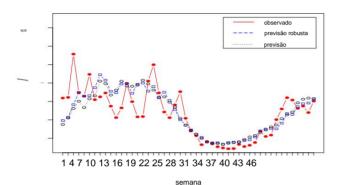


(b) Fluxos logaritmicos trimestra

Fig. 6 Box-plots do logarítmico mensal e trimestral fluxos do rio Garonne em Tonneins

(MAE) e o erro percentual absoluto médio (MAPE). Essas medidas são explicitamente definidas em Hyndman e Koehler (2006). O MAE e o RMSE entre o modelo proposto e os dados observados são calculados no mesmas unidades dos dados observados. Um valor menor indica um melhor desempenho do modelo. A medida MAPE é baseada em erros percentuais. Esses critérios devem ser interpretados apenas como uma indicação para qual modelo tem melhor desempenho, mas não afirmação pode ser feita a partir desta comparação. Para testar o hipótese nula de nenhuma diferença na precisão do modelos propostos, um teste de classificação assinada de Wilcoxon para dados podem ser usados (Noakes et al. 1985). Além do RMSE, medidas MAE e MAPE, fornecemos o índice relativo de concordância (rd), o coeficiente de determinação (R2) e a eficiência de Nash-Sutcliffe (NSe) para melhor comparação. Essas medidas são definidas em Krause et al. (2005). O rd índice varia entre 0 e 1. Um valor de 1 indica um correspondência perfeita e 0 indica nenhuma concordância. o intervalo de NSe situa-se entre 1 e 1. Essencialmente, quanto mais próximo para 1, o modelo mais preciso é. O alcance de R2 está entre 0 e 1 e um coeficiente mais alto é um indicador de um melhor modelo. Todas as medidas são calculadas usando o pacotes hydroGOF e previsão na R. Para outro





(b) Previsão de quarenta e oito períodos

Fig. 7 Os dados observados para os primeiros 51 anos aparecem em linha contínua (em vermelho), previsões robustas estão em linha tracejada (em azul) e previsões não robustas estão em linha pontilhada (em preto). Os dados observados para o próximo ano (52) não foram utilizados na previsão. Em uma previsão são para os 12 meses enquanto, b trata das previsões ao longo dos 48 trimestres. Os fluxos estão em m3/s

medidas de precisão de previsão nos referimos a Hyndman e Koehler (2006), Krause et al. (2005).

Outras previsões mensais podem ser derivadas da agregação das previsões trimestrais. Eles estão obtido pela média dos quatro períodos de cada mês. A Tabela 3 apresenta as medidas RMSE, MAE, MAPE, rd, R2 e NSe para mensal, trimestral e previsões trimestrais agregadas. Os resultados mostram que o modelo robusto é melhor em relação a todos os critérios para a dados trimestrais e trimestrais agregados. Em pelo contrário, o modelo não robusto parece ter melhor desempenho

Tabela 3 Precisão de uma etapa previsões de fluxos do rio Garonne

| Critério | Por mês     |         | Trimestral  |         | Ag. trimestral | Ag. trimestral |  |
|----------|-------------|---------|-------------|---------|----------------|----------------|--|
|          | Não robusto | Robusto | Não robusto | Robusto | Não robusto    | Robusto        |  |
| MAPA     | 20,37       | 22,65   | 31.05       | 29,96   | 24,34          | 23,39          |  |
| MUITO DE | 91,47       | 103,38  | 128,32      | 123,52  | 101,37         | 95,64          |  |
| RMSE     | 121,44      | 139,87  | 175,46      | 169,56  | 133,43         | 125,26         |  |
| rd       | 0,93        | 0,92    | 0,85        | 0,86    | 0,90           | 0,92           |  |
| R2       | 0,73        | 0,67    | 0,56        | 0,59    | 0,68           | 0,72           |  |
| NSe      | 0,72        | 0,63    | 0,53        | 0,57    | 0,67           | 0,71           |  |

para dados mensais. Isso pode ser explicado pelo baixo número de outliers (igual a 11), mas mesmo neste caso o robusto modelo pode ser preferido devido ao princípio da parcimônia. Os resultados mostram que o número de estimativas é menor no caso robusto. Além disso, o comportamento das estimativas robustas foi na maioria dos casos razoável, mesmo que os valores discrepantes sejam ausente como mostrado por Ursu e Pereau (2014) usando simulação.

#### Conclusões

A previsão precisa das vazões dos rios é uma das importantes aplicações em hidrologia, especialmente para a gestão de sistemas de reservatórios. A precisão do análise de previsão é importante para a gestão do rio autoridades para atingir os objetivos de fluxo mínimo durante o estações mais secas. Com base em uma abordagem de modelagem robusta para a identificação e a estimativa de séries temporais PAR modelo, este artigo fornece uma aplicação para o Garonne Rio no período 1959-2010. Para lidar com o problema de um grande número de parâmetros necessários para serem estimados especialmente com modelos trimestrais, um sistema automático método usando rotinas genéticas foi desenvolvido. Resultados mostram que a detecção de outliers é maior com dados de fluxo mensal trimestral do que com dados mensais, implicando melhor estimadores robustos do que os estimadores de mínimos quadrados (não robustos). Os resultados mostram que as previsões de 1 ano são melhores em modelos robustos trimestrais do que em modelos não robustos. As projeções trimestrais agregadas também mostram melhores desempenho do que as previsões mensais no robusto análise.

Nossa análise também sugere integrar a evolução algoritmo para otimização da operação do reservatório. Como mostrado por Wang et al. (2014), o uso de métodos mais sofisticados como algoritmo genético ou vetor de suporte máquina pode melhorar a qualidade da previsão hidrológica sobre os modelos autoregressivos clássicos. Pesquisa futura lidar com modelos de limiares periódicos deve ser feito para melhorar o ajuste e a previsão do fluxo do rio. Em particular, medidas de dados sugerem uma mudança estrutural no regime hidrológico por volta do ano de 1990. Isso requer

desenvolver uma nova abordagem de modelagem PAR com pontos de comutação endógenos como em Koutroumanidis et al. (2009).

Para abordar a questão das propriedades homocedásticas dos resíduos de vazão, que é uma questão importante na modelagem hidrológica, pesquisas futuras são necessárias para implementar um teste portmanteau robusto em modelos PAR. Esse teste robusto existe para modelos de séries temporais ARMA (Li 1988), mas não para modelos PAR. McLeod (1994) desenvolveu um teste não robusto em modelos PAR. As propriedades dos resíduos também podem ser melhor capturadas combinando modelos periódicos com a teoria de valores extremos como em Tesfaye et al. (2006).

Agradecimentos Os autores agradecem a dois pareceristas anônimos por suas valiosas e construtivas observações. Os autores agradecem a A. Coynel da Universidade de Bordeaux e unidade EPOC por seus comentários. Os dados de fluxos de Garona estão disponíveis aos autores mediante solicitação (eugen.ursu@u-bordeaux.fr). Os dados são arquivados no DIREN Banque Hydro, monitoramento de água francês. Este estudo foi realizado com o apoio financeiro da Agência Nacional de Pesquisa Francesa (ANR) como parte do projeto ADAPTEAU (ANR-11-CEPL-008) e no âmbito do Cluster of Excellence COTE (ANR-10-LABX-45).

Reclamação com padrões éticos

Conflito de interesse Os autores declaram não ter conflito de interesse.

#### Referências

- Aksoy H, Dahamsheh A (2009) Modelos de redes neurais artificiais para previsão de precipitação mensal na Jordânia. Stoch Environ Res Risk Assessment 23:917–931 Anderson PA, Meerschaert MM, Zhang K (2013) Previsão com
- intervalos de previsão para modelos periódicos de média móvel autorregressiva. J Tempo Ser Anal 34:187–193
- Baker MA, Vervier P (2004) Variabilidade hidrológica, fornecimento de matéria orgânica e desnitrificação no ecossistema do rio garonne.

  Freshw Biol 49(2):181–190 Ben
- MG, Martinez EJ, Yohai VJ (1999) Estimação robusta em modelos de séries temporais de média móvel de vetor autoregressivo. J Time Ser Anal 20:381–399 Boe J, Terray L, Martin E, Habets F (2009) Mudanças projetadas em
- componentes do ciclo hidrológico em bacias hidrográficas francesas durante o século XXI. Water Resour Res 45:W08426 Box GEP, Jenkins GM (1970) Análise de séries temporais, previsão e
  - ao controle. Holden Day, São Francisco
- Bustos OH, Yohai VJ (1986) Estimativas robustas para modelos ARMA. J Am Stat Assoc 81:155–168
- Caballero Y, Voirin-Morel S, Habets F (2007) Sensibilidade hidrológica da bacia do rio Adour-Garonne às mudanças climáticas. Water Resour Res 43(7):W07448 Denby L, Martin RD (1979) Estimativa robusta da primeira ordem
  - parâmetro autorregressivo. J Am Stat Assoc 74:140-146
- Durdu OF (2010) Aplicação de modelos estocásticos lineares para previsão de seca na bacia do rio Buu"y"k Menderes, oeste da Turquia.

  Avaliação de Risco Stoch Environ Res 24:1145–1162
- Eshete Z, Vandewiele GL (1992) Comparação de multicomponentes não-gaussianos e modelos autorregressivos periódicos para o fluxo do rio. Stoch Hydrol Hydraul 6:223-238

- Fayaed SS, El-Shafie A, Jaafar O (2013) Simulação de sistemas de reservatórios e técnicas de otimização. Avaliação de Risco Stoch Environ Res 27:1751–1772
- Fernandez C, Vega JA, Fonturbel T, Jimenez E (2008) Previsão de séries temporais de seca de fluxo de corrente: um estudo de caso em uma pequena bacia hidrográfica no noroeste da Espanha. Avaliação de Risco Stoch Environ Res 23:1063–1070
- Franses PH, Paap R (2004) Modelos periódicos de séries temporais. Oxford
  University Press, Oxford Gaetan C (2000) Identificação do modelo ARMA do
  subconjunto usando algoritmos genéticos. J Time Ser Anal 21:559–570 Gladyshev
  EG (1961) Sequências aleatórias correlacionadas periodicamente. Sov Math
- 2:385–388 Goldberg DE (1989) Algoritmos genéticos em busca, otimização e aprendizado de máquina. Addison-Wesley, Boston Gragne A, Sharma A,
- Mehrotra R, Alfredsen K (2014) Melhorando a previsão de influxo em reservatórios hidrelétricos por meio de uma estrutura de modelagem complementar. Hydrol
- Earth Syst Sci 11:12063–12101 Hau MC, Tong H (1989) Um método prático para detecção de valores discrepantes na modelagem de séries temporais autorregressivas. Stoch Hydrol Hydraul 3:241–260
- Hendrickx F, Sauquet E (2013) Impacto do aquecimento climático na gestão da água para a bacia do rio arge (França). Hydrol Sci J 58(4):1–17
- Hipel KW, McLeod AI (1994) Modelagem de séries temporais de recursos hídricos e sistemas ambientais. Elsevier, Amsterdam Hyndman RJ, Koehler AB (2006)
- Outro olhar sobre as medidas de precisão das previsões. Int J Forecast 22:679–688

  Jimenez C, McLeod AI, Hippel KW (1989) Estimativa de filtro de Kalman para
- modelos periódicos de média móvel autorregressiva. Stoch Hydrol Hydraul 3:227–240 Koutroumanidis T, Sylaios G, Zafeiroiou E, Tsihrintzis V (2009)
- Modelagem genética para a previsão ótima de séries temporais hidrológicas: aplicação no Rio Nestos. J Hydrol 368:156–164 Krause P, Boyle DP, Base F (2005) Comparação de diferentes critérios de eficiência para avaliação de modelos hidrológicos. Adv Geosci 5:89–97
- Li WK (1988) Um teste de ajuste em modelagem robusta de séries temporais. Biometrika 75:355–361 Li
- WK (2004) Verificações de diagnóstico em séries temporais. Chapman & Hall/CRC, Nova lorque
- Lu"tkepohl H (2005) Nova introdução à análise de séries temporais múltiplas. Springer, Berlin Ma
- Y, Genton MG (2000) Estimação altamente robusta da função de variância autoco. J Time Ser Anal 21:663–684 Madsen H, Skotner C (2005) Atualização de
- estado adaptável em rio em tempo real agora prevendo um procedimento combinado de filtragem e previsão de erros. J Hidrol 308(1):302–312
- Maire A, Buisson L, Biau S, Canal J, Lafaille P (2013) Uma estrutura multifacetada de diversidade para priorizar a conservação de assembleias de peixes. Ecol Indic 34:450–459 Maronna RA, Martin RD, Yohai VJ (2006) Estatísticas
- robustas: teoria e métodos. Wilev. Nova York
- McLeod AI (1993) Parcimônia, adequação do modelo e autocorrelação periódica na previsão de séries temporais. Int Stat Rev 61:387–393 McLeod AI (1994) Verificação de diagnóstico de autoregressão periódica
  - modelos com aplicações. J Tempo Ser Anal 15:221-233
- McLeod AI, Gweon H (2013) Dessazonalização ideal para mensal e séries temporais geofísicas diárias. J Environ Stat 4:1–11 Mishra
- AK, Desai VR (2005) Previsão de seca usando modelos estocásticos. Avaliação de Risco Stoch Environ Res 19:326–339
- Mitchell M (1996) Uma introdução aos algoritmos genéticos. MIT Press, Cambridge

- Muylaert K, Sanchez-Perez JM, Teissier S, Sauvage S, Dauta A, Vervier P (2009) Eutrohisation e seu efeito sobre as concentrações de Si dissolvido no rio Garonne (França). J Limnol 68(2):368–374
- Noakes DJ, McLeod Al, Hipel KW (1985) Previsão de séries temporais mensais de fluxo de rios. Int J Previsão 1:179–190 Oeurng C, Sauvage S, Coynel A,
- Maneux E, Etcheber H, Sanchez Perez JM (2011) Transporte fluvial de sedimentos suspensos e carbono orgânico durante eventos de inundação em uma grande bacia agrícola no sudoeste da França. Hydrol Process 25:2365–2378 Pagano M (1978) Em autorregressões periódicas e múltiplas. Ann Stat 6:1310–1317
- Sarnaglia AJQ, Reisen VA, Le'vy-Leduc C (2010) Estimativa robusta de processos autorregressivos periódicos na presença de outliers aditivos. J Multivar Anal 101:2168–2183
- Shao Q (2007) Estimativa robusta para tempo autorregressivo periódico Series. J Tempo Ser Anal 29:251–263
- Sivanandam SN, Deepa SN (2008) Introdução aos algoritmos genéticos. Springer, Berlin
- Tesfaye YG, Meerschaert MM, Anderson PL (2006) Identificação de modelos periódicos de média móvel autorregressiva e seus

- aplicação à modelação de caudais fluviais. Recursos Hídricos Res 42:1-11
- Tisseuil C, Vrac M, Lek S, Wade A (2010) Redução estatística de fluxo do rio. J Hydrol 385: 279-291
- Ursu E, Duchesne P (2009) Sobre modelagem e verificação diagnóstica de modelos vetoriais periódicos autorregressivos de séries temporais. J Tempo Ser Anal 30:70–96
- Ursu E, Pereau JC (2014) Modelagem robusta de séries temporais autoregressivas vetoriais periódicas. J Stat Plan Inference 155:93–106 Ursu E, Turkman KF
- (2012) Identificação periódica de modelo autorregressivo usando algoritmo genético. J Time Ser Anal 33:398–405 Vecchia AV (1985a) Modelagem periódica de
- média móvel autorregressiva (PARMA) com aplicações em recursos hídricos. Water Resour Bull 21:721–730 Vecchia AV (1985b) Estimativa de máxima verossimilhança para modelos periódicos de média móvel autorregressiva.
- Technometrics 27:375–384 Wang Y, Guo S, Chen H, Zhou Y (2014) Estudo comparativo de métodos de previsão de fluxo mensal para o voir do reservatório de três desfiladeiros. Avaliação de Risco Stoch Environ Res 28:555–570