

## **TEXTO DE DISCUSSÃO Nº 80**

# **X-13ARIMA-SEATS COM R: UM ESTUDO DE CASO PARA A PRODUÇÃO INDUSTRIAL BRASILEIRA**

Pedro Guilherme Costa Ferreira  
José Lisboa Gondin Jr.  
Daiane Marcolino de Mattos

Junho de 2015

# X-13ARIMA-SEATS com R: Um Estudo de Caso para a Produção Industrial Brasileira

Pedro Guilherme Costa Ferreira\*, José Lisboa Gondin Jr.\*, and Daiane Marcolino de  
Mattos\*

*Instituto Brasileiro de Economia (IBRE)*

## Resumo

Esse artigo apresenta o uso do programa de ajuste sazonal X-13ARIMA-SEATS desenvolvido e mantido por US Census Bureau. O programa é mundialmente utilizado por órgãos de estatística e, em alguns lugares, ainda em sua versão anterior. O ajuste sazonal é implementado no software R e aplicado na série temporal do índice de produção industrial geral do Brasil. As ferramentas disponíveis pelo programa permitiram aprimorar o ajuste automático e diagnosticá-lo de forma adequada. Apresenta-se também uma rotina para executar o ajuste sazonal em mais de uma série simultaneamente, uma vez que, geralmente, as instituições que realizam tal procedimento, não o fazem para apenas uma série temporal.

Palavras Chaves: ajuste sazonal, X-13ARIMA-SEATS, produção industrial, R.

## 1 Introdução

Uma série temporal, segundo a decomposição clássica, pode ser decomposta em quatro componentes não observáveis: tendência, sazonalidade, ciclo e erro. A sazonalidade, principal objeto deste estudo, é causada por movimentos oscilatórios de mesma periodicidade que ocorrem em

---

\*Instituto Brasileiro de Economia (IBRE|FGV), R. Barão de Itambi, Botafogo, Rio de Janeiro, Brasil; tel: +55 21 3799-6751; emails: pedro.guilherme@fgv.br; gondin@gmail.com; daiane.mattos@fgv.br.

período intra-anual, como variações climáticas, férias, feriados, entre outros. A ocorrência desses eventos pode levar a conclusões inadequadas a respeito da série temporal em estudo. Por exemplo, a oferta de emprego costuma aumentar no final do ano devido às festas natalinas, isto é, há uma demanda maior por bens e serviços, elevando o nível de contratações de pessoas. Porém, como a maioria das vagas é temporária, geralmente, há diminuição no nível de pessoal ocupado no período seguinte. Para a análise econômica, o importante é detectar a diferença entre o que ocorre periodicamente e o que de fato ocorre de diferente naquele período específico, possibilitando observar a tendência e o ciclo da variável.

Dessa forma, precisa-se de uma ferramenta adequada que consiga remover essa componente (a sazonalidade). A remoção da sazonalidade de uma série temporal é chamada de ajuste sazonal ou dessazonalização. Na literatura, é possível encontrar diversas metodologias/softwarees que permitem a remoção da sazonalidade de uma série temporal, como, por exemplo:

- a. *Seasonal Dummies* (Zellner, 1979);
- b. *Holt-Winters* (Rasmussen, 2004);
- c. *Structural Models* (Harvey and Shepard, 1993; Plosser, 1979; Koopman et al., 2009);
- d. *Dainties* (Fok et al., 2005);
- e. *TRAMO-SEATS* (Gómez and Maravall, 1997; Hungarian Central Statistical Office, 2007);
- f. *X-11, X-11ARIMA, X-12ARIMA* (Shiskin et al., 1967; Findley et al., 1998);
- g. *X-13ARIMA-SEATS* (U.S. Census Bureau, 2015);

Diversos órgãos de estatística relevantes, como IBGE (2015b); Eurostat (2015) and BLS (2015), aplicam o X12-ARIMA ou o X13-ARIMA-SEATS que permitem um ajuste sazonal de qualidade. E, atualmente, alguns softwares facilitam a execução desses programas. É possível citar como exemplo o R (R Core Team, 2015), que possui uma interface aprimorada ao utilizar o RStudio (2015), ambos disponíveis para download de forma gratuita. Um guia sobre como utilizar o X-12ARIMA foi criado por Kowarik et al. (2014). E, de forma semelhante a esses autores, esse artigo propõe o uso do X-13ARIMA-SEATS.

Nesse contexto, o objetivo desse artigo é apresentar como dessazonalizar uma série temporal usando o X13-ARIMA-SEATS com interface no R/RStudio. Para atingir esse objetivo, além dessa parte introdutória, o artigo está organizado da seguinte forma: na seção 2 tem-se uma visão geral do X-13ARIMA-SEATS. Na seção 3, é apresentado passo-a-passo um algoritmo para dessazonalizar uma série temporal com o X-13 no R. Na seção 4, aplica-se o algoritmo de ajuste sazonal à série temporal do índice de produção industrial geral do Brasil. E, por fim, na seção 5, conclui-se o trabalho e discute-se os possíveis trabalhos futuros.

## 2 X-13ARIMA-SEATS

O X13-ARIMA-SEATS, criado em julho de 2012, é um programa de ajuste sazonal desenvolvido por U.S Census Bureau com o apoio do Bank of Spain. O programa é a junção dos softwares X12-ARIMA e TRAMO/SEATS com melhorias. As melhorias incluem uma variedade de novos diagnósticos que ajudam o usuário a detectar e corrigir inadequações no ajuste. Além de incluir diversas ferramentas que superaram problemas de ajuste e permitiram um aumento na quantidade de séries temporais econômicas que podem ser ajustadas de maneira adequada (U.S. Census Bureau, 2015).

Um procedimento contido no X-13ARIMA-SEATS que merece destaque é o pré-ajuste da série temporal, isto é, uma correção antes de ser feito, de fato, o ajuste sazonal. Alguns eventos atípicos e/ou não sazonais como, por exemplo, efeitos do calendário (trading days, working day, moving holidays, etc), greves, catástrofes, entre outros, podem afetar o padrão sazonal da série temporal e, conseqüentemente, gerar um ajuste de qualidade inferior. O tratamento desses eventos (pré-ajuste) deve ser feito, se necessário. Um exemplo da aplicação de ajuste sazonal a partir do X-13ARIMA-SEATS com a utilização de trading days pode ser encontrado em Livsey et al. (2014).

### 3 Algoritmo de ajuste sazonal usando o R

Na seção 3.1, a seguir, apresentam-se os passos para realizar o ajuste sazonal em uma série temporal. Na seção posterior, é proposta uma rotina desenvolvida no software R para o ajuste sazonal de múltiplas séries simultaneamente. E na seção 3.3, tem-se os comandos que precisam ser executados no R antes do ajuste sazonal.

#### 3.1 SSA: Single Seasonal Adjustment

Para executar o ajuste sazonal no R em apenas uma série (SSA), serão seguidos os quatro passos abaixo, e suas definições são dadas logo em seguida.

1. Análise Gráfica;
2. Execução o X-13ARIMA-SEATS no modo automático;
3. Avaliação do ajuste sazonal em (2);
4. Correção do ajuste sazonal em (2) (se necessário).

A análise gráfica de uma série temporal permite visualizar suas características para uma boa modelagem, por exemplo: seu padrão sazonal, quebras estruturais, possíveis outliers, se há necessidade (e possibilidade) de usar transformação logarítmica nos dados.

O X13, basicamente, funciona em duas etapas: pré-ajuste e ajuste sazonal. Na primeira, o software pode corrigir a série de efeitos determinísticos. É nesta etapa que o usuário pode especificar, por exemplo, outliers e efeitos do calendários (Páscoa, Carnaval, etc). Na segunda etapa, é feito o ajuste sazonal de fato. A execução do programa no modo automático pode trazer um ajuste sazonal de boa qualidade. O programa no modo automático verifica, entre outras coisas, se há necessidade de transformação log nos dados; se existem outliers (additive, level shift e temporary change); a ordem do modelo ARIMA; se há efeitos de calendário). Essas verificações automáticas podem poupar o tempo do usuário e ajudá-lo na escolha de um bom modelo, principalmente na etapa do pré-ajuste. No entanto, este modelo precisa ser avaliado e o X-13ARIMA-SEATS fornece algumas ferramentas<sup>1</sup> para essa finalidade:

---

<sup>1</sup>Há uma gama de recursos oferecidos pelo X-13ARIMA-SEATS que não ainda não foram explorados nesse trabalho. Mais informações ver X13-ARIMA-SEATS Reference Manual Accessible HTML Output Version (U.S.

- QS statistic: Verifica a existência de sazonalidade em uma série temporal.

A tabela 1, a seguir, resume em quais séries temporais o programa calcula o teste de sazonalidade. Em um bom ajuste sazonal, o diagnóstico dado pela estatística QS, permitiria concluir indícios de sazonalidade somente na série original e não nas restantes.

É importante saber que se a série possui menos de 8 anos de dados mensais (96 observações), o teste de sazonalidade é executado apenas em toda a série temporal; caso contrário, é executado na série temporal inteira e também nos últimos oito anos.

Codificação	Significado
qsori	série original
qsorievadj	série original corrigida por outliers
qsrtd	resíduos do modelo SARIMA
qssadj	série com ajuste sazonal
qssadjevadj	série com ajuste sazonal corrigida por outliers
qsirr	componente irregular
qsirrevadj	componente irregular corrigida por outliers

Tabela 1: Séries temporais disponíveis para o diagnóstico dado pela estatística QS

- Ljung-Box statistic: O teste de Ljung and Box (1978) verifica a existência de autocorrelação em uma série temporal. O X-13 mostra o resultado desse teste aplicado aos resíduos do modelo SARIMA estimado na defasagem 24. Espera-se que os resíduos não sejam autocorrelacionados (hipótese nula).
- Shapiro-Wilk statistic: O teste de Shapiro and Wilk (1965) verifica se a distribuição de um conjunto de dados é normal. O X-13 mostra o resultado desse teste aplicado aos resíduos do modelo SARIMA estimado. Espera-se que os resíduos sigam distribuição normal (hipótese nula).
- Gráfico SI ratio: Útil para verificar se a decomposição das componentes da série temporal foi feita adequadamente. Espera-se que os fatores sazonais acompanhem o SI

---

Census Bureau, 2015).

(componentes sazonal e irregular agregadas<sup>2</sup>, indicando que o SI não é dominado pela componente irregular.

- Gráfico Spectral<sup>3</sup>: é outra ferramenta que alerta se a série temporal possui influência de efeitos sazonais e de trading days. O gráfico é feito para a série original, para a série com ajuste sazonal (se de fato o ajuste sazonal foi executado), para a série da componente irregular e para os resíduos do modelo SARIMA. Se o objetivo realizar um ajuste sazonal na série temporal, então é esperado a identificação de efeitos sazonais no gráfico spectral da série original. E se o ajuste sazonal foi feito adequadamente, espera-se que tais efeitos não sejam encontrados nas séries disponíveis restantes.

Após a análise de todas as ferramentas de diagnóstico, caso alguma não conformidade seja detectada no modelo automático, o modelo deve ser reajustado e diagnosticado novamente. Algumas alterações que podem ajudar a melhorar o ajuste são: rever a necessidade de transformação nos dados, isso pode estabilizar a variância; modificar a ordem do modelo SARIMA; inserir ou retirar outliers e/ou variáveis de regressão, etc.

### 3.2 MSA: Multiple Seasonal Adjustment

Geralmente, as instituições que fazem ajuste sazonal não o fazem apenas para uma série temporal. Com o objetivo de otimizar esse procedimento, foi desenvolvida uma rotina no software R, nomeada de **MSA** (Multiple Seasonal Adjustment), em que o ajuste sazonal utilizando o X-13ARIMA-SEATS é aplicado a mais de uma série temporal simultaneamente.

Basicamente, a função **MSA()** utiliza o pacote **seasonal** desenvolvido por Christoph Sax (2015a; 2015b) para fazer o ajuste sazonal nas séries, no entanto, sua saída exporta os outputs apresentados na seção 3.1, a série temporal original, a série ajustada, entre outras séries disponíveis.

O código da função está disponível em <https://github.com/pedrocostaferreira>.

---

<sup>2</sup>Se for utilizada a decomposição aditiva (sem transformação log) então SI é a soma da componente sazonal e da componente irregular S+I. Caso contrário, usa-se a multiplicação: S×I

<sup>3</sup>Veja mais detalhes sobre o gráfico spectral em (U.S. Census Bureau, 2015, chapter 6).

### 3.3 Pré-algoritmo

Os quatro passos citados na seção 3.1 são sugeridos para um ajuste sazonal de qualidade. Para executá-lo no R, no entanto, alguns procedimentos precisam ser executados primeiramente:

1. Definir diretório de trabalho;
2. Baixar X-13ARIMA-SEATS;
3. Informar o local do arquivo executável baixado em (2);
4. Instalar e carregar pacote **seasonal**;
5. Verificar se todas as etapas foram executadas corretamente.

A seguir, o desenvolvimento de cada passo:

1. Definir diretório de trabalho: É nesse local onde ficarão os arquivos necessários para realizar o ajuste sazonal. O comando a seguir, direciona o diretório para a pasta *work* contida na unidade *C*.

```
> setwd("C:/work")
```

2. Baixar X-13-ARIMA-SEATS: O software pode ser encontrado no site do US Census Bureau. Deve ser baixado<sup>4</sup> e colocado na pasta de diretório de trabalho. Mas, para facilitar, seguindo os comandos<sup>5</sup> abaixo, o usuário pode obter o arquivo sem precisar sair do R. O arquivo será armazenado no diretório de trabalho especificado anteriormente com o nome de **x13.zip**, e em seguida será descompactado.

```
> download.file("https://www.census.gov/ts/x13as/pc/x13as_V1.1_B19.zip",  
+               destfile = "./x13.zip")  
> unzip("x13.zip")
```

3. Informar o local do arquivo executável baixado em (2): O arquivo *x13\_as.exe*, desenvolvido pelo US Census Bureau, é que de fato faz o ajuste sazonal. É necessário informar

---

<sup>4</sup>Vale lembrar que não é preciso fazer o download do arquivo sempre que precisar fazer um ajuste sazonal. Apenas deixe-o em seu diretório de trabalho para futuras utilizações e passe para etapa seguinte, ignorando esta.

<sup>5</sup>Ocasionalmente, novas versões do programa podem ser criadas causando mudanças no link de download e, consequentemente, o não funcionamento da linha de comando.



o seu caminho ao R.

```
> local <- paste(getwd(), "/x13as", sep = "")  
> Sys.setenv(X13_PATH = local)
```

4. Instalar e carregar pacote **seasonal**: Christoph Sax (2015a) desenvolveu o pacote **seasonal** que executa o X-13ARIMA-SEATS no R. As funções do pacote permitem ao usuário, além do ajuste sazonal, obter diagnósticos, tais como os citados na seção 3.1.

```
> install.packages("seasonal")  
> library(seasonal)
```

5. Verificar se todas as etapas foram executadas corretamente: A função `checkX13()` do pacote **seasonal** permite verificar se será possível a realização do ajuste sazonal. Se os 4 procedimentos anteriores foram executados de maneira correta, a função retorna uma mensagem, vista a seguir, de confirmação e é possível dessazonalizar uma série temporal com o X-13ARIMA-SEATS no R.

```
> checkX13()
```

```
X-13 installation test:  
- X13_PATH correctly specified  
- binary executable file found  
- command line test run successful  
- seasonal test run successful  
Congratulations! 'seasonal' should work fine!
```

seasonal now supports the HTML version of X13, which offers a more accessible output via the `out()` function. For best user experience, download the HTML version from:

[http://www.census.gov/srd/www/x13as/x13down\\_pc.html](http://www.census.gov/srd/www/x13as/x13down_pc.html)

and copy `x13ashtml.exe` to:

`C:/work/x13as`

## 4 Estudo de Caso: Índice de Produção Industrial

A produção industrial de uma região, geralmente é afetada pelas épocas do ano. Por volta de outubro, por exemplo, é esperado um aumento no nível de produção devido às comemorações natalinas em dezembro. Nos meses seguintes, no entanto, é esperado uma queda nesse nível. Com ajuste sazonal, a série de produção industrial poderá ser interpretada sem os efeitos do calendário, permitindo realizar comparações entre os meses de forma adequada. Dessa forma, o algoritmo de ajuste sazonal apresentado na seção 3 será aplicado na série temporal do índice de produção industrial geral do Brasil, estimado mensalmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2015c) pela Pesquisa Industrial Mensal - Produção Física (PIM-PF) desde a década de 1970.

Os dados, que podem ser extraídos do sistema SIDRA do IBGE mas também estão disponíveis em <https://github.com/pedrocostaferreira>, representam um índice sem ajuste sazonal com base fixa em 2012 (média de 2012 = 100), e compreendem o espaço de tempo de janeiro de 2002 a dezembro de 2014, totalizando 156 observações.

### 4.1 Single Seasonal Adjustment (SSA)

Após download da série histórica do índice de produção industrial em formato .csv, o arquivo foi lido no R com a função `read.csv2()` e, em seguida, transformado em um objeto de séries temporais utilizando a função `ts()` como se pode ver nos próximos comandos.

```
> pim <- read.csv2("pimpf.csv")  
> pim.ts <- ts(pim, start = c(2002,1), freq = 12)
```

A seguir, a execução dos quatro passos do algoritmo da seção 3.1.

#### 4.1.1 Análise Gráfica

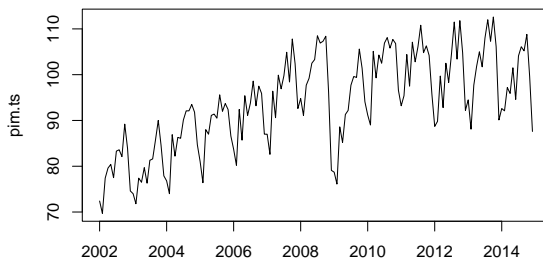
Para esboçar, de forma simples, o gráfico de uma série temporal, utiliza-se a função `plot()`. Outro gráfico que pode ajudar a entender o comportamento de séries temporais é dado pela

função `monthplot()`. Nele, é possível comparar a série histórica de cada mês do ano. A análise gráfica da série temporal (figura 1) permite supor que o índice de produção industrial

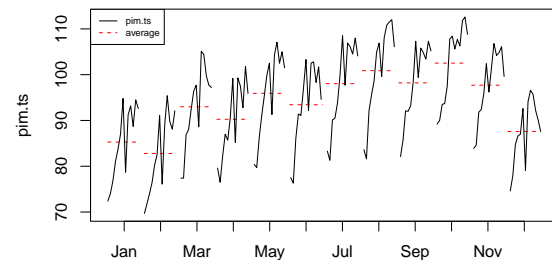
- tem característica sazonal, uma vez que de janeiro a outubro o índice tem comportamento crescente e, nos outros dois meses, decrescente. Esse comportamento se repete em todos os anos;
- tinha uma tendência crescente antes da crise econômica (final de 2008) e parece estar estável (sem crescimento ou quedas expressivas) após tal acontecimento;
- parece ter variação constante, não necessitando que os dados sejam transformados para estabilizá-la;
- foi extremamente afetado pela crise econômica no final de 2008. Nota-se que o índice atingiu um valor discrepante em relação ao comportamento habitual.

Para criar a figura 1, executou-se os comandos abaixo.

```
> plot(pim.ts)
> monthplot(pim.ts, col.base = 2, lty.base = 2, labels = month.abb)
> legend("topleft", legend = c("pim.ts", "average"),
+       cex = 0.6, lty = c(1,2), col = c(1,2))
```



(a) Série histórica (`plot`)



(b) Série histórica por mês (`monthplot`)

Figura 1: Análise gráfica do índice de produção industrial (IBGE,2015).

### 4.1.2 Execução do X-13ARIMA-SEATS no modo automático

Após visualizar graficamente o comportamento histórico da série temporal de interesse, o próximo passo é efetuar o ajuste sazonal no modo automático. A função `seas()` do pacote `seasonal`<sup>6</sup> desempenhará esse papel. Os principais argumentos da função `seas()` são:

- `x`: série temporal de interesse;
- `arma.model`: permite especificar do modelo SARIMA para a série de interesse;
- `outlier`: permite definir se o programa deve ou não detectar automaticamente outliers;
- `regression.variables`: permite especificar outliers e variáveis de calendário como Páscoa, trading days, ano bissexto, entre outras variáveis<sup>7</sup>;
- `regression.aictest`: permite definir se o programa deve ou não detectar automaticamente variáveis de regressão.
- `transform.function`: permite especificar a transformação que deve ser aplicada na série de interesse. As opções disponíveis são `log` (transformação logarítmica), `none` (nenhuma transformação) ou `auto` (o programa define se deve ser ou não aplicada a transformação log);

A função `seas()`, no entanto, requer apenas a série temporal em que se pretende fazer o ajuste sazonal para o seu funcionamento. Os outros argumentos, nem todos exemplificados anteriormente, funcionarão no modo automático. Logo, para executar o ajuste sazonal em uma série temporal no modo automático, utilize `seas(x)`. Para o estudo de caso desse artigo, utilizou-se

```
> (ajuste <- seas(pim.ts))
```

```
Call:
seas(x = pim.ts)
```

```
Coefficients:
```

---

<sup>6</sup>Mais detalhes sobre o pacote ver Sax (2015b).

<sup>7</sup>Outras variáveis pré-definidas podem ser encontradas em X13-ARIMA-SEATS Reference Manual Accessible HTML Output Version (U.S. Census Bureau, 2015, chapter 7, pp. 144-147)

Mon	Tue	Wed	Thu
0.0055494	0.0053673	0.0022694	0.0052843
Fri	Sat	Easter[1]	LS2008.Dec
0.0004402	-0.0002930	-0.0242646	-0.1334385
A02011.Feb	A02014.Feb	MA-Seasonal-12	
0.0612578	0.0629949	0.6797172	

No modelo ajustado automaticamente foram detectados efeitos da Páscoa, de dias da semana e também de outliers. Foi detectado um outlier level shift no mês de dezembro de 2008 (LS2008.Dec), mês extremamente afetado pela crise econômica. Outros dois outliers, não esperados visualmente pela análise da figura 1, foram detectados em fevereiro de 2011 (A02011.Feb) e 2014 (A02014.Feb). Esses são do tipo aditivo. O próximo passo é avaliar esse ajuste.

#### 4.1.3 Avaliação do ajuste sazonal

Para avaliar o ajuste sazonal feito na seção 4.1.2, precisa-se dos resultados das ferramentas apresentadas no tópico 3 da seção 3.1. Um breve resumo desses resultados são obtidos pela função `summary()`.

```
> summary(ajuste)
```

Call:

```
seas(x = pim.ts)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
Mon	0.0055494	0.0023373	2.374	0.0176 *
Tue	0.0053673	0.0023757	2.259	0.0239 *
Wed	0.0022694	0.0022926	0.990	0.3222
Thu	0.0052843	0.0023300	2.268	0.0233 *
Fri	0.0004402	0.0023069	0.191	0.8487
Sat	-0.0002930	0.0023159	-0.127	0.8993
Easter[1]	-0.0242646	0.0043734	-5.548	2.89e-08 ***
LS2008.Dec	-0.1334385	0.0182778	-7.301	2.87e-13 ***
A02011.Feb	0.0612578	0.0129029	4.748	2.06e-06 ***
A02014.Feb	0.0629949	0.0140520	4.483	7.36e-06 ***
MA-Seasonal-12	0.6797172	0.0694161	9.792	< 2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

SEATS adj. ARIMA: (0 1 0)(0 1 1) Obs.: 156 Transform: log

AICc: 613.8, BIC: 646.9 QS (no seasonality in final): 0  
Box-Ljung (no autocorr.): 23.4 Shapiro (normality): 0.9817 \*

Foi ajustado um modelo SARIMA(0 1 0)(0 1 1). O parâmetro MA-Seasonal-12 foi significativo ao considerar nível de significância de 5%. O mesmo pode ser dito para o efeito dos três outliers e da variável que reflete a Páscoa (**Easter[1]**). Embora nem todos os dias da semana sejam significativos considerando nível de 5% de significância, três dias foram (**Mon, Tue and Thu**), e isso é suficiente para mantê-los no modelo e concluir que há indícios de que a produção industrial seja afetada pelos dias da semana. Observa-se que a hipótese de normalidade dos resíduos foi rejeitada com 95% de confiança e a transformação log foi aplicada na série original, embora acreditou-se que isso não fosse necessário como observado na seção 4.1.1. E o teste de Ljung & Box, sugere não haver evidências de autocorrelação residual até o lag 24.

A criação de um gráfico spectral para análise de efeitos sazonais e de trading days pode ser feita no R após extrair o "spectral output" utilizando a função **series()** do pacote **seasonal** e transformá-lo em um objeto da classe **data.frame**. O código, a seguir, mostra como extrair o spectral output para a série original<sup>8</sup> (**sp0**).

```
> spec.orig <- data.frame(series(ajuste, "sp0"))
```

O gráfico pode ser feito utilizando o pacote **ggplot2** com as configurações abaixo.

```
> library(ggplot2)
> ggplot(aes(x=Pos,y = X10.Log.Spectrum_AdjOri.), data = spec.orig, colour = "black") +
+   geom_line() +
+   geom_vline(colour = "red", xintercept = c(10, 20, 30, 40, 50), linetype = 5) +
+   geom_vline(colour = "blue", xintercept = c(42, 52), linetype = 3) +
+   ylab(" ") + xlab(" ") + theme_bw() +
+   ggtitle("Spectral plot of the first-differenced original series") +
+   theme(plot.title = element_text(lineheight=2, face="bold",size = 16))
```

O gráfico spectral (figura 2), mostra indícios de efeitos de sazonalidade na série original (2a), visto que nas frequências sazonais (10, 20, 30, 40, 50) a série temporal toma forma de picos bem

---

<sup>8</sup>Para ver a codificação para as outras séries além da série original ver (U.S. Census Bureau, 2015, chapter 7, pp. 194)

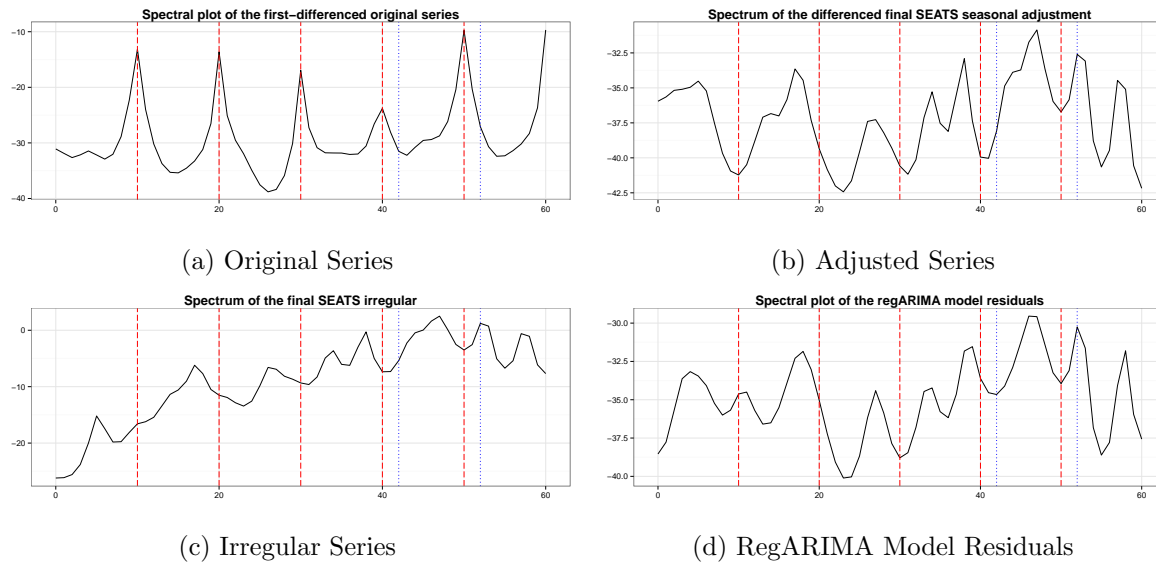


Figura 2: Análise espectral do ajuste sazonal.

definidos. Já nas frequências de trading days (linha pontilhada em azul, aproximadamente 42 e 52), não se pode concluir o mesmo pois não há forma de picos. Para as outras três séries não foi detectado picos de sazonalidade, porém, há leves indícios de efeitos de trading days, o que é estranho pois foi incluído o efeito de trading days no modelo do ajuste.

As suposições de sazonalidade na série original não são rejeitadas com a análise da estatística QS. Os testes de sazonalidade nas demais séries apresentadas na tabela 1 da seção 3.1, podem ser vistos utilizando a função `qs()` do pacote **seasonal**:

```
> qs(ajuste)
```

	qs	p-val
qsori	162.66893	0.00000
qsorievadj	236.53944	0.00000
qsrsd	0.02387	0.98814
qssadj	0.00000	1.00000
qssadjevadj	0.00000	1.00000
qsirr	0.00000	1.00000
qsirrevadj	0.00000	1.00000
qssori	81.40339	0.00000
qssorievadj	136.59753	0.00000
qssrsd	0.00000	1.00000
qssadj	0.00000	1.00000

```

qsssadjevadj  0.00000 1.00000
qssirr        0.00000 1.00000
qssirrevadj   0.00000 1.00000

```

Uma vez que a série `pim.ts` apresenta mais de 96 observações, o teste de sazonalidade foi calculado para a série completa e para os 8 anos mais recentes. Nota-se que o p-value é pequeno tanto para série original (`qsori`) como para a série original corrigida por outliers (`qsorievadj`), isto é, não há evidências de que a série do índice de produção industrial não seja sazonal. Nas demais séries, pelo p-value ser próximo de 1, conclui-se o contrário: há evidências de não sazonalidade nas séries temporais. Assim, se tratando de sazonalidade, o X-13ARIMA-SEATS cumpriu bem o seu dever de removê-la.

Mais uma ferramenta para avaliar a qualidade do ajuste sazonal é dado pelo gráfico *SI ratio*. Para esboçá-lo, utiliza-se a função `monthplot()`.

```

> monthplot(ajuste, col.base = 1, lty.base = 2,
+           labels = month.abb, lwd.base = 2)
> legend("topleft", legend = c("SI", "FS", "Mean FS"),
+       cex = 0.7, lty = c(1,1,2), col = c(4,2,1), lwd = c(1,2,2))

```

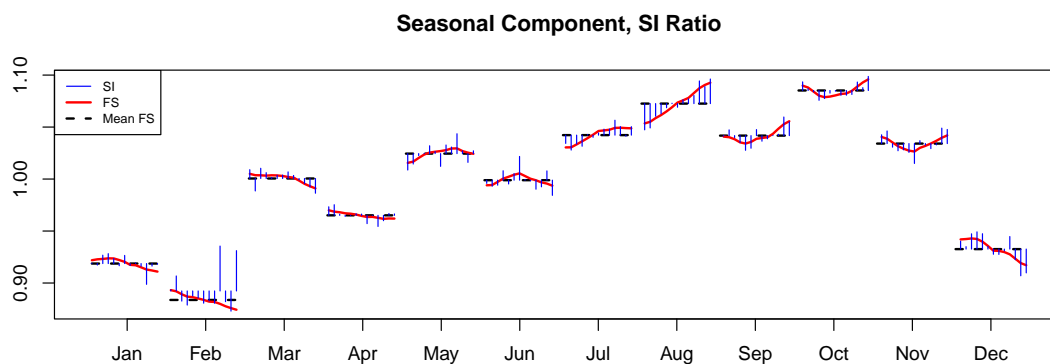


Figura 3: SI Ratio and Seasonal Factors.

Na figura 3, as linhas azuis (verticais) referem-se a componente SI ratio (componentes sazonal e irregular agregadas). As linhas vermelhas (realçadas) representam os fatores sazonais. E a linha tracejada é a média dos fatores sazonais naquele mês. Pode-se observar que os fatores



sazonais tendem a acompanhar os SI ratio. Isso significa que a componente SI ratio não é dominada pela componente irregular, isto é, os erros tem um comportamento estável em torno de zero e a decomposição das componentes não observáveis da série temporal foi feita adequadamente. No entanto, note que para o mês de fevereiro (em que dois outliers foram encontrados), o SI ratio é dominado pela componente irregular.

Pode-se concluir então que o ajuste sazonal automático já forneceu bons resultados, porém, como alguns pressupostos necessários (normalidade dos resíduos) não foram confirmados estatisticamente, o modelo precisa ser especificado com mais detalhes.

#### 4.1.4 Correção do ajuste sazonal

Após a análise do ajuste sazonal automático na seção 4.1.3, verificou-se que o modelo precisava ser corrigido. O IBGE, utilizando o método X-12-ARIMA, adiciona ao modelo de ajuste sazonal, além de efeitos de trading days e Páscoa, o efeito do Carnaval IBGE (2015a). Esse efeito também será acrescentado e espera-se que as alterações corrijam a normalidade dos resíduos e o novo modelo tenha um critério de informação inferior ao do modelo automático. Para criar a variável de Carnaval, utiliza-se a função `genhol()` do **seasonal**:

`genhol(x, start, end, frequency)`

- **x**: um vetor da classe "Date", contendo as datas de ocorrência do feriado. Pode ser criado pela função `as.Date`;
- **start**: inteiro, desloca o ponto inicial do feriado. Use valores negativos se o efeito começa antes da data específica;
- **end**: inteiro, desloca o ponto final do feriado. Use valores negativos se o efeito termina antes da data específica;
- **frequency**: inteiro, frequência da série temporal resultante.

Para os argumentos **start** e **end** escolheu-se colocar uma janela de 3 dias antes e um dia depois do feriado, uma vez que no Brasil o feriado dura quase uma semana. Ao argumento **frequency** foi atribuído 12 pois os dados são mensais.

```
> dates <- c("02/12/2002", "03/04/2003", "02/24/2004", "02/08/2005",
+           "02/28/2006", "02/20/2007", "02/05/2008", "02/24/2009",
+           "02/16/2010", "03/08/2011", "02/21/2012", "02/12/2013",
+           "03/04/2014", "02/17/2015", "02/09/2016", "02/28/2017")
> carnaval.date <- as.Date(dates, "%m/%d/%Y")
> carnaval <- genhol(carnaval.date, start = -3, end = 1, frequency = 12)
```

Para acrescentar a variável `carnaval` ao ajuste sazonal, especificou-se o argumento `xreg = carnaval`.

Os trading days poderiam ser específicos para cada dia da semana como foi feito no ajuste automático, ou apenas uma variável que combinassem essas informações específicas (ver nota de rodapé 7). Ambos os tipos foram testados com e sem o efeito do ano bissexto. A opção que melhor caracterizou o modelo foi apenas uma variável que indicasse efeitos do dia da semana mais o efeito de anos bissextos e esta é indicada por `td1coef`.

```
> ajuste_novo <- seas(pim.ts, transform.function = "none",
+                    xreg = carnaval, regression.variables = "td1coef")
> summary(ajuste_novo)
```

Call:

```
seas(x = pim.ts, xreg = carnaval, transform.function = "none",
     regression.variables = "td1coef")
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
carnaval	-2.99294	0.48858	-6.126	9.02e-10	***
Leap Year	2.41881	0.72837	3.321	0.000897	***
Weekday	0.35307	0.03042	11.605	< 2e-16	***
Easter[1]	-2.98161	0.41291	-7.221	5.16e-13	***
A02008.Nov	-6.80666	1.49598	-4.550	5.37e-06	***
LS2008.Dec	-17.09895	1.79260	-9.539	< 2e-16	***
A02011.Dec	5.05730	1.18157	4.280	1.87e-05	***
AR-Nonseasonal-01	-0.28065	0.07945	-3.533	0.000411	***
MA-Seasonal-12	0.52321	0.07747	6.754	1.44e-11	***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

SEATS adj. ARIMA: (1 1 0)(0 1 1) Obs.: 156 Transform: none  
AICc: 581.9, BIC: 609.8 QS (no seasonality in final): 0.3735  
Box-Ljung (no autocorr.): 20.46 Shapiro (normality): 0.9908

Com as mudanças inseridas na função `seas()`, o modelo SARIMA também foi modificado de  $(0\ 1\ 0)(0\ 1\ 1)$  para  $(1\ 1\ 0)(0\ 1\ 1)$  com todos os parâmetros significativos com 95% de confiança. As variáveis de regressão também foram significativas considerando o mesmo nível de confiança e, assim, pode-se concluir que a quantidade de dias da semana e o ano bissexto influenciam na produção industrial brasileira. Além disso, as alterações permitiram concluir que os resíduos seguem distribuição normal, e os critérios de informação AICc e BIC são consideravelmente inferiores ao do ajuste automático. O programa encontrou mais um outlier além dos relacionados a crise econômica de 2008 (A02011.Dec). O teste de Ljung & Box não mostrou autocorrelação residual nos lags de 1 a 24.

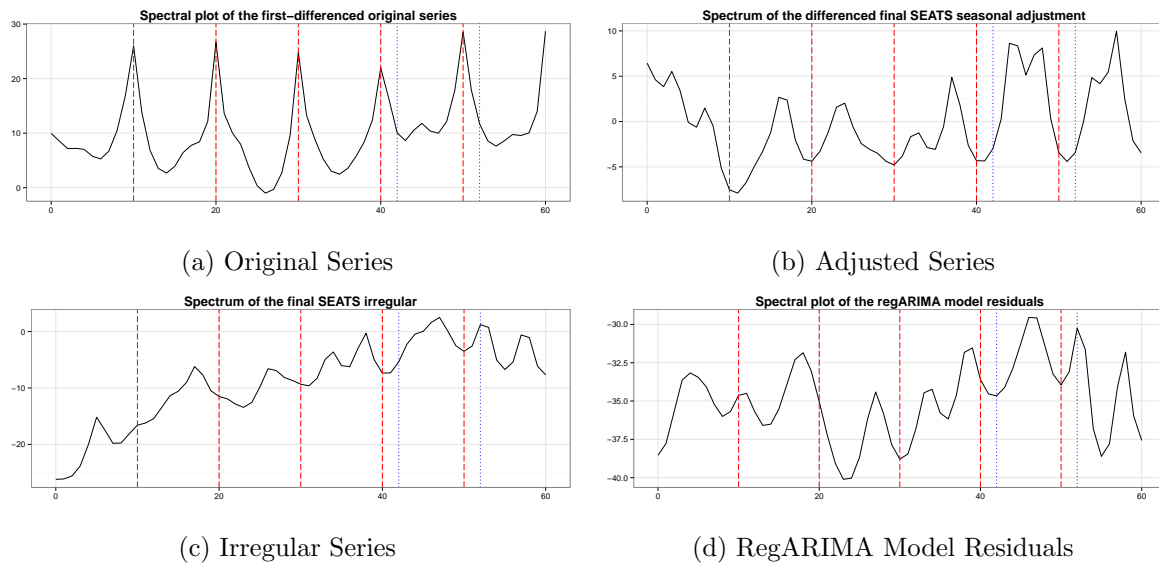


Figura 4: Análise spectral do ajuste sazonal corrigido.

A análise do gráfico spectral (figura 4), de forma diferente da análise feita na seção 4.1.3, não mostra indícios de efeitos de trading days na série com ajuste sazonal. Os fatores sazonais (figura 5) aparentam ter um comportamento mais suave do que o do ajuste automático. Repare que o mês de fevereiro foi melhor captado depois da correção. Note também o impacto dos outliers nos fatores sazonal de novembro de 2008 e de dezembro de 2011: a componente SI assume um valor discrepante comparado aos outros valores do mesmo mês.

A conclusão do teste de sazonalidade é semelhante à conclusão do teste de sazonalidade para o ajuste automático, também qualificando este ajuste em adequado.

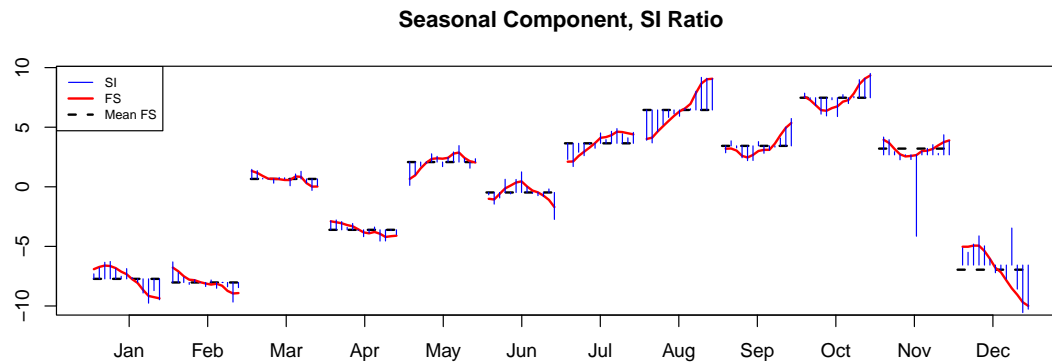


Figura 5: SI Ratio and Seasonal Factors para o modelo corrigido.

```
> qs(ajuste_novo)
```

	qs	p-val
qsori	162.66893	0.00000
qsorievadj	230.29764	0.00000
qsrsd	0.00000	1.00000
qssadj	0.37351	0.82965
qssadjevadj	0.00000	1.00000
qsirr	0.00000	1.00000
qsirrevadj	0.00000	1.00000
qssori	81.40339	0.00000
qssorievadj	131.02081	0.00000
qssrsd	0.00000	1.00000
qsssdadj	0.00186	0.99907
qsssdjevadj	0.00000	1.00000
qssirr	0.00000	1.00000
qssirrevadj	0.00000	1.00000

Por fim, tem-se o gráfico do índice de produção industrial com ajuste sazonal pelo X-13ARIMA-SEATS em que nota-se seu comportamento decrescente nos últimos meses.

```
> plot(ajuste_novo)
> legend("topleft", legend = c("Observada", "Com ajuste sazonal"),
+       cex = 0.7, lty = 1, col = c(1,2), lwd = c(1,2))
```

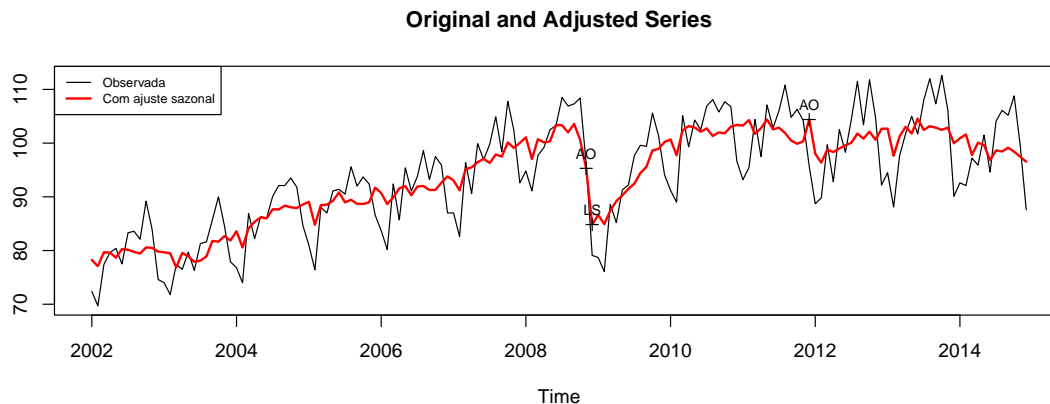


Figura 6: Índice de produção industrial geral do Brasil com ajuste sazonal.

## 4.2 Multiple Seasonal Adjustment (MSA)

Para mostrar sua utilidade, a rotina `MSA()` apresentada na seção 3.2 será aplicada às séries temporais que compõem o índice de produção industrial. O índice de produção industrial geral pode ser desagregado em indústria extrativa e indústria de transformação, sendo este último composto por outros 25 grupos. Dessa forma, o ajuste sazonal será aplicado nos 25 grupos e na indústria extrativa. Ao final, as séries com ajuste sazonal serão agrupadas. Esse método de ajuste sazonal é conhecido como método indireto (Hungarian Central Statistical Office, 2007), em que no lugar de o ajuste sazonal ser aplicado na série final (como foi feito na seção 4.1), faz-se nas séries que a compõem. Assim, o ajuste sazonal será feito em 26 séries temporais.

Após o download das 26 séries temporais<sup>9</sup>, estas foram armazenadas em um único arquivo .csv e lidas no R utilizando a função `read.csv2()`. Em seguida, o objeto foi transformado em um de classe `mts` utilizando a função `ts()`.

```
> pim_multiple <- read.csv2("pim_multiple.csv")
> pim_multiple.ts <- ts(pim_multiple[, -1], start = c(2002, 1), freq = 12)
> dim(pim_multiple.ts)
```

```
[1] 156 26
```

<sup>9</sup>Disponíveis em <https://github.com/pedrocostaferreira>.

Para fazer o ajuste sazonal<sup>10</sup> em todas séries simultaneamente utilizando a função `MSA()`, é necessário carregá-la antes utilizando a função `source()`.

```
> source("./MSA.R")
> ajuste_msa <- MSA(pim_multiple.ts)
```

O objeto retornado pela função `MSA()` é uma lista que contém um objeto da classe `seas` para cada série temporal se o ajuste sazonal pôde ser executado, ou então contém um objeto com a própria série temporal caso não tenha sido possível dessazonalizar tal série. A função demorou em torno de 15 segundos para tentar fazer o ajuste sazonal nas 26 séries temporais. Para auxiliar na análise do ajuste sazonal feito pela função `MSA()`, também foram desenvolvidas as rotinas `MSA.GA()` (Graphical Analysis) e `MSA.EV()` (Evaluation of seasonal adjustment) que permitem a elaboração das mesmas análises das seções anteriores.

```
> # Graphical Analysis
> MSA.GA(pim_multiple.ts)
> # Evaluation of seasonal adjustment
> MSA.EV(ajuste_msa)
```

Se após a análise de cada ajuste, algum parâmetro (por exemplo, modelo SARIMA) necessitar de correção, o ajuste individual pode ser refeito utilizando a função `seas()`.

Para extrair as séries com ajuste sazonal do objeto `ajuste_msa`, utiliza-se a função `final()` do pacote **seasonal**. Por exemplo, para o primeiro ajuste execute:

```
> final(ajuste_msa[[1]])
```

Após agregar<sup>11</sup> todas as séries com ajuste sazonal, obteve-se o índice de produção industrial com ajuste sazonal (figura 7).

---

<sup>10</sup>Lembre-se que os procedimentos informados na seção 3.3 ainda precisam ser executados para que `MSA()` funcione.

<sup>11</sup>Os pesos utilizados na agregação foram calculados baseados na correlação das 26 séries com a série do índice de produção industrial (todas sem ajuste sazonal).

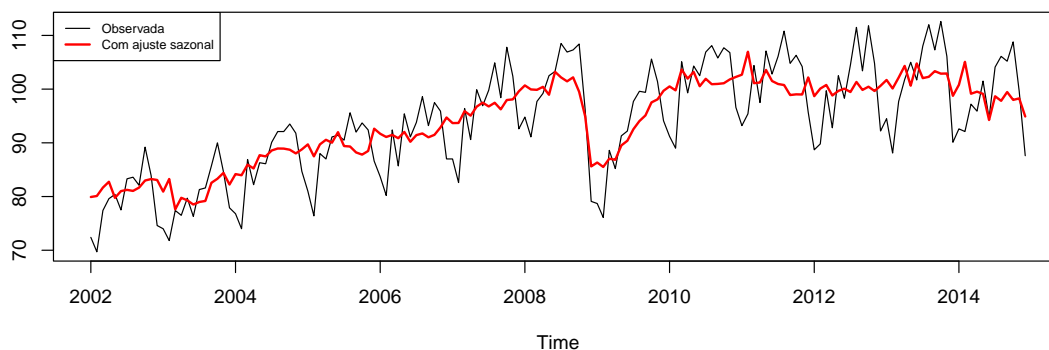


Figura 7: Índice de produção industrial geral do Brasil com ajuste sazonal (método indireto).

## 5 Final Remarks

O artigo apresentou como dessazonalizar séries temporais utilizando X-13ARIMA-SEATS com interface no software R fazendo uso do pacote **seasonal** desenvolvido por Christoph Sax. Além de apresentar o passo-a-passo do ajuste sazonal, mostrou-se também como analisar cada output retornado pelo programa. O método foi aplicado a série temporal do índice de produção industrial geral do Brasil e sua análise permitiu concluir que, além de o índice ter minimizado seu crescimento após a crise econômica de 2008 e mostrar queda nos últimos meses, nem sempre o ajuste automático fornecido pelo programa possui a melhor qualidade possível, assim recomenda-se que o responsável por fazer o ajuste sazonal avalie a saída automática.

O X-13ARIMA-SEATS fornece diversas ferramentas para diagnosticar um ajuste sazonal e meios de aprimorá-lo. Para o índice de produção industrial, por exemplo, verificou-se que inserir o efeito do carnaval e efeitos de calendário melhoram a qualidade do ajuste. A maneira como a variável de carnaval foi construída foi baseada no número de dias que o feriado ocorre no Brasil, porém, o efeito de um feriado pode afetar a produção industrial dias, semanas ou até meses antes de sua ocorrência. Logo, essa variável poderia ter sido construída de diversas outras maneiras afetando o resultado final do ajuste sazonal e isso não foi levado em conta neste artigo. Destaca-se também que apesar de o programa oferecer diversos recursos para executar um ajuste sazonal, nem todos foram explorados neste artigo.

Por saber que geralmente o ajuste sazonal é feito para mais de uma série temporal e visando a otimizar esse processo, desenvolveu-se uma rotina (Multiple Seasonal Adjustment) no R para o ajuste de várias séries temporais simultaneamente. Sua utilização mostrou-se razoável no que se refere ao tempo de execução: aproximadamente 15 segundos para 26 séries temporais de 156 meses (em sua maioria). No entanto, a análise individual de cada ajuste não foi realizada. Para continuar o propósito desse artigo, futuramente pretende-se comparar, considerando os métodos direto e indireto, o ajuste sazonal do índice de produção industrial feito pelo X-13ARIMA-SEATS e o ajuste feito pela instituição oficial do cálculo do índice, o IBGE, que ainda utiliza o programa X12-ARIMA.

## Referências

- JJ Allaire, Jeff Allen, Aron Atkins, Bill Carney, Winston Chang, Joe Cheng, Jim Clemens, Garrett Grolmund, Fereshteh Karimeddini, Tareef Kawaf, Andrew Kipp, Pete Knast, Jonathan McPherson, Roger Oberg, Stephen Peters, Kevin Ushey, Hadley Wickham, and Yihui Xie. Rstudio, 2015. URL <http://www.rstudio.com/>.
- BLS. Bureau labor of statistics, 2015. URL <http://www.bls.gov/>.
- Eurostat. Eurostat, 2015. URL <http://ec.europa.eu/eurostat>.
- D. F. Findley, B. C. Monsell, W. R. Bell, M. C. Otto, and Bor-Chung Chen. New capabilities and method of the x-12-arima seasonal adjustment program. *Journal of Business and Economic Statistics*, 16, 1998.
- D. Fok, P. H. Franses, and R. Paap. Performance of seasonal adjustment procedures: Simulation and empirical results. *Econometric Institute Report*, 2005.
- Víctor Gómez and Augustín Maravall. Guide for using the programs tramo and seats (beta version). *Banco de España*, 1997. URL <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/98/Fic/dt9805e.pdf>.



- A. Harvey and N. Shepard. Structural time series models. *Elsevier Science Publishers B.V.*, 1993.
- Hungarian Central Statistical Office. Seasonal adjustment methods and practices. *European Commission Grant*, 2007.
- IBGE. Pesquisa industrial mensal produção física - brasil - notas metodológicas, 2015a. URL [http://www.ibge.gov.br/home/estatistica/indicadores/industria/pimpfbr/notas\\_metodologicas.shtm](http://www.ibge.gov.br/home/estatistica/indicadores/industria/pimpfbr/notas_metodologicas.shtm).
- IBGE. Instituto brasileiro de geografia e estatística, 2015b. URL [www.ibge.gov.br](http://www.ibge.gov.br).
- IBGE. Pesquisa industrial mensal - produção física, 2015c. URL <http://www.ibge.gov.br/home/estatistica/indicadores/industria/pimpfbr/>.
- S. J. Koopman, A. Harvey, J. Doornik, and N. Shepard. Structural time series analyser, modeler and predictor. *Timberlake Consultants*, 2009.
- Alexander Kowarik, Angelika Meraner, Matthias Templ, and Daniel Schopfhauser. Seasonal adjustment with R packages x12 and x12gui. *Journal of Statistical Software*, 62, 2014. URL <http://www.jstatsoft.org/v62/i02>.
- James Livsey, Osbert Pang, and Tucker McElroy. Effect of trading day regressors on seasonal adjustment of growth rates. *RESEARCH REPORT SERIES, US Census Bureau*, 2014. URL <https://www.census.gov/srd/papers/pdf/rrs2014-09.pdf>.
- G. M. Ljung and G. E. P. Box. On a measure of a lack of fit in time series models. *Biometrika*, 65, 1978.
- C. I. Plosser. A time series analysis of seasonality in econometric models. *The National Bureau of Economic Research*, 1979.
- R. Rasmussen. On time series data and optimal parameters. *The International Journal of Management Science*, 2004.

- R Core Team. The R project for statistical computing, 2015. URL <http://www.r-project.org/>.
- Christoph Sax. Github christoph sax, 2015a. URL <https://github.com/christophsax>.
- Christoph Sax. **seasonal**: *R Interface to X-13-ARIMA-SEATS*, 2015b. URL <http://cran.r-project.org/package=seasonal>.
- S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52, 1965.
- J. Shiskin, A. H. Young, and J. C. Musgrave. The x-11 variant of the census method ii seasonal adjustment program. *Bureau of the Census*, 52, 1967.
- U.S. Census Bureau. X13-arima-seats reference manual accessible html output version. 2015. URL <https://www.census.gov/ts/x13as/docX13AS.pdf>.
- Hadley Wickham and Winston Chang. **ggplot2**: *An Implementation of the Grammar of Graphics*, 2015. URL <http://cran.r-project.org/package=ggplot2>.
- Arnold Zellner. Front matter to ‘seasonal analysis of economic time series. *The National Bureau of Economic Research*, 1979.



[www.fgv.br/ibre](http://www.fgv.br/ibre)

**Rio de Janeiro**

Rua Barão de Itambi, 60  
22231-000 - Rio de Janeiro – RJ

**São Paulo**

Av. Paulista, 548 - 6º andar  
01310-000 - São Paulo – SP