

Métodos de previsão para a taxa de desemprego mensal: uma análise de séries temporais

Dieison Lenon Casagrande¹

Felipe Resende Oliveira²

Guilherme Studart³

Inaldo Bezerra da Silva⁴

Paulo Henrique Monteiro Guimarães⁵

RESUMO

O presente trabalho tem como objetivo prever a taxa de desemprego mensal no Brasil. Utilizam-se algoritmos de alisamento exponencial e modelos estocásticos para tal. O algoritmo que melhor representa o processo gerador de dados é um Holt-Winters com tendência multiplicativa atenuada e com sazonalidade multiplicativa. Os modelos estocásticos que se destacaram foram o SARIMA (3, 0, 1)(2, 0, 1) e o SARIMAX (4, 0, 2)(2, 0, 2) que inclui a taxa de inflação como regressor.

Palavras-Chaves: Desemprego; Algoritmo de alisamento exponencial; Série Temporal.

Classificação JEL: C01, E24.

ABSTRACT

This work aims to predict the monthly unemployment rate in Brazil. It uses exponential smoothing models and stochastic algorithms. The algorithm that best represents the data generating process is a Holt-Winters with trend multiplicative and attenuated and multiplicative seasonal. Stochastic models that stood out were the SARIMA (3, 0, 1) (2, 0, 1) and the SARIMAX (4, 0, 2) (2, 0, 2) which includes the inflation rate as a regressor.

Key-words: Unemployment; Exponential smoothing algorithm; Time Series.

JEL Classification: C01, E24.

1 Doutorando em Economia pelo Programa de Pós-graduação em Economia da Universidade Federal de Pernambuco - PIMES/UFPE - e-mail: dieisonlenon@yahoo.com.br

2 Doutorando em Economia pelo Programa de Pós-graduação em Economia da Universidade Federal de Pernambuco - PIMES/UFPE - e-mail: felipexresende@gmail.com

3 Doutorando em Economia pelo Programa de Pós-graduação em Economia da Universidade Federal de Pernambuco - PIMES/UFPE - e-mail: guistudart@gmail.com

4 Doutorando em Economia pelo Programa de Pós-graduação em Economia da Universidade Federal de Pernambuco - PIMES/UFPE - e-mail: inaldobs@gmail.com

5 Doutorando em Economia pelo Programa de Pós-graduação em Economia da Universidade Federal de Pernambuco - PIMES/UFPE - e-mail: gmpaulohenrique@hotmail.com.

1. INTRODUÇÃO

Um estudo mais detalhado a respeito da atividade econômica de um país exige uma análise da dinâmica dos seus mais diversos componentes. Alguns desses componentes nos fornecem diagnósticos acerca da situação econômica de uma nação em termos de distribuição de renda, desenvolvimento e estabilidade econômica, além de outras características.

O nível de desemprego é uma das principais variáveis econômicas estudadas na tentativa de compreender e diagnosticar os principais fenômenos econômicos e sociais. Pode-se explicar o grande interesse por tal variável pelo fato de se usar informações sobre desemprego para formulação de políticas públicas que afetem diretamente famílias e firmas. Por muito tempo o nível de desemprego de uma economia foi tema controverso. Os defensores da chamada escola clássica classificam o desemprego como sendo apenas voluntário ou transitório, já para a escola keynesiana acredita que desemprego pode ser persistente e não apenas voluntário. Essa teoria ganhou força com a grande depressão da década de 1930.

Existem inúmeros estudos sobre desemprego. Há estudiosos que analisam o desemprego de forma a atribuir como principal causador de tal fenômeno a escolaridade dos indivíduos (Fraga e Dias, 2007), outros autores como Mattos (2001) defendem que o desemprego carrega componentes estruturais, características demográficas e importância do setor público como grandes determinantes do nível de emprego e portanto da taxa de desemprego (Mattos, 2001).

Por outro lado, há ainda autores que caracterizam a taxa de desemprego como sendo um componente que pode se explicar ao longo do tempo, Becker (2010) trabalha o desemprego com modelos de séries temporais a fim de obter previsões para suas taxas futuras. Tal autor usa diversos métodos de séries temporais para tentar encontrar o melhor ajustamento ao conjunto de dados. Ele usa modelos do tipo SARIMA, além de vários algoritmos de alisamento exponencial.

Neste trabalho seguiremos a lógica proposta por Becker em que usa a taxa de desemprego passada como subsídio para previsões de taxa de desemprego futura. Para tanto utilizaremos modelos de previsões da família SARIMA, além de usar algoritmos de alisamento exponencial visando dar melhor ajustamento aos dados e poder fazer previsões consistentes acerca do desemprego dos próximos períodos.

O método para modelar com o SARIMA será a modelagem BOX-JENKINS, seguindo as etapas de identificação onde serão determinados os valores dos parâmetros do SARIMA $(p, d, q) \times (P, D, Q)_s$, começando pela determinação da ordem de integração através da análise dos testes de Dickey e Fuller e o teste Dickey e Fuller aumentado e ainda pelo correlograma, depois serão definidos os valores das defasagens dos processos autoregressivos e de médias móveis envolvidos no SARIMA, diagnóstico em que será analisado o ajustamento do modelo e por fim as previsões.

À utilização de algoritmos de alisamento exponencial, serão estimados vários deles, como o algoritmo de Holt e Winters aditivo, multiplicativo, com atenuação de tendência e outros que estão descritos na seção de metodologia.

Desse modo este trabalho objetiva fazer uma análise dos métodos de estimação de séries temporais, com foco em uma verificação de qual dos métodos melhor se ajusta à série histórica de desemprego, e ainda, dado o melhor modelo, fazer previsões consistentes que possibilitem uma melhor tomada de decisão quando se trata de políticas públicas de combate ao desemprego e manutenção da renda. Onde prever os níveis futuros de desemprego pode permitir a antecipação de cenários recessivos, e com isso combater antecipadamente tendências de queda da renda.

2. Revisão de Literatura

No presente trabalho, utilizaremos a definição de desemprego e de taxa de desemprego comumente utilizada pelo IBGE (Instituto Brasileiro de Geografia e Estatística), onde a taxa de desemprego é formada pela razão entre a população desocupada e a população economicamente ativa multiplicada por 100.

A população economicamente ativa (PEA) é uma das partes que compõem a população em idade ativa (PIA). A PIA é composta pela população economicamente ativa e a população economicamente inativa, essa classificação compreende o conjunto de todas as pessoas teoricamente aptas a exercer uma atividade econômica. Para o IBGE, a população economicamente ativa é composta pela população ocupada e a população desocupada, ou seja, a PEA compreende o potencial de mão de obra com que o setor produtivo pode contar.

Ainda segundo o IBGE, constituem a população ocupada: os empregados (aquelas pessoas que trabalham para um empregador ou mais, cumprindo uma jornada

de trabalho, recebendo em contrapartida uma remuneração em dinheiro ou outra forma de pagamento), os conta própria (aquelas pessoas que exploram uma atividade econômica ou exercem uma profissão ou ofício, sem empregados), os empregadores (aquelas pessoas que exploram uma atividade econômica ou exercem uma profissão ou ofício, com auxílio de um ou mais empregados) e os trabalhadores não remunerados (aquelas pessoas que exercem uma ocupação econômica, sem remuneração, pelo menos 15 horas na semana). Já a população desocupada é composta por pessoas que não têm trabalho, mas estão dispostas a trabalhar, e que, para isso, tomam alguma providência efetiva.

Por fim, o IBGE usa a definição de quem forma a população economicamente inativa, são as pessoas não classificadas como ocupadas ou desocupadas da PIA, ou seja, tal grupo é formado por pessoas incapacitadas para o trabalho ou que desistiram de buscar trabalho ou não querem trabalhar. Inclui os incapacitados, os estudantes e as pessoas que cuidam de afazeres domésticos. Inclui também os desalentados, pessoas em idade ativa que já não buscam trabalho, uma vez que já o fizeram e não obtiveram sucesso. O IBGE considera desalentado aquele que está desempregado e há mais de um mês não busca emprego.

Ao longo da história do pensamento econômico, o desemprego sempre foi tema de muita polêmica e discordância, as mais diversas opiniões e correntes de pensamento surgiram, tendo como ponto de conflito o mercado de trabalho. Por exemplo, os economistas marxistas viam na persistência do desemprego, através da manutenção do “exército industrial de reserva”, a manifestação concreta das contradições do sistema capitalista e da inerente luta de classes que acabaria por conduzir o capitalismo a derrocada.

Já para a tradicional escola clássica, o desemprego decorre do descompasso de curto prazo entre oferta e demanda de trabalho, em caso de elevado desemprego, a redução dos salários nominais restabeleceria o equilíbrio, tendo o desemprego caráter puramente transitório. Ao analisar, a economia clássica no livro *A Teoria Geral do Emprego, do Juros e da Moeda*, Keynes classificou as formas de desemprego existentes na teoria clássica do emprego em duas: o desemprego friccional e o desemprego voluntário, não admitindo, a teoria clássica, pelos postulados a ela impostos, outra modalidade de desemprego.

Posteriormente, na sua Teoria Geral, Keynes observa que o desemprego assume uma terceira modalidade, chamada de desemprego involuntário, oriunda de um fenômeno conjuntural, decorrente da insuficiência da Demanda Agregada. Os trabalhadores, para os quais a produção não há demanda, permanecerão involuntariamente desempregados, independentemente de aceitarem, inclusive, alguma redução de salários. Este desemprego é involuntário, independe do salário comumente pago e, para Keynes, pode ser evitado ou mitigado através de políticas ativas, fiscais e monetárias que elevem a demanda efetiva.

O desemprego voluntário ocorre segundo Keynes, “oriundo da recusa ou incapacidade de determinada unidade de mão de obra em aceitar uma remuneração equivalente à sua produtividade marginal, em decorrência da legislação ou dos costumes sociais, ou de um entendimento para contrato coletivo de trabalho, ou ainda da lentidão em adaptar-se às mudanças, ou simplesmente, em consequência da obstinação humana.” Já o desemprego friccional decorre do fato que as transferências de um emprego para outro não se realizam sem certa demora, de modo que, numa sociedade não estática, sempre existe certa proporção de recursos não empregados. Por fim, o terceiro tipo de desemprego que existe, o desemprego conjuntural, cíclico ou involuntário ocorre nas fases descendentes do ciclo econômico, quando o produto fica abaixo do nível de pleno emprego (Demanda Agregada inferior à Oferta Agregada).

O florescimento da teoria keynesiana representou a ascensão do capitalismo de estado na economia, pois justificava o intervencionismo governamental como fonte estimuladora do mercado de trabalho e geradora de emprego. Porém, nos anos 50, o economista Willian Phillips, relacionou uma suposta relação estável de trade-off entre inflação e desemprego, aonde uma menor taxa de desemprego levaria a uma maior taxa de inflação, tal trade-off ficou conhecido como curva de Phillips, que estabeleceu uma nova conceituação o que justificava certo grau de desemprego em troca da estabilidade dos preços.

Posteriormente, Milton Friedman e Edmund Phelps, apresentaram a curva de Phillips aceleracionista, que acrescenta à equação original a análise das expectativas adaptativas, tal curva indica que, para mantermos a taxa de desemprego a níveis inferiores ao da taxa de desemprego natural, o que importa não é a taxa de inflação, mas sim sua variação, necessitando-se assim de taxas de inflação cada vez maiores para manter as taxas de desemprego abaixo da taxa natural. Segundo Friedman (1968) e

Phelps (1968), a taxa natural de desemprego é a taxa que reflete a estrutura do mercado de trabalho e a natureza das relações entre firmas e trabalhadores, e que vigora no equilíbrio, após todos os ajustes serem realizados.

Friedman e Phelps concluíram que a inflação e o desemprego não estão relacionados no longo prazo, assim sendo, a Curva de Phillips é perfeitamente inelástica no ponto da Taxa Natural de Desemprego que é a taxa de desemprego quando a inflação corrente é igual a inflação esperada. Eles constroem esta nova versão da curva de Phillips para tentar provar que as políticas de emprego keynesianas, em vez de reduzir o desemprego, apenas provocam inflação.

Este termo taxa natural de desemprego, surgiu quando Milton Friedman o utilizou em seu discurso de posse na presidência da American Economic Association, a fim de expressar a ideia de que o nível de desemprego, em uma sociedade, não poderia ser determinado pela política monetária, uma vez que a taxa de desemprego seria resultado apenas de forças reais atuantes na economia. Para Phelps (1995), a taxa natural de desemprego serve como um atrator para a taxa corrente e choques monetários causam apenas desvios temporários dela.

O presente estudo visa prever o comportamento futuro da taxa de desemprego, obviamente, o eventual comportamento intervencionista do governo em situações futuras ou circunstâncias conjunturais de curto prazo adversas, como crises econômicas, poderão gerar desvios no mercado de trabalho, levando-o a um superaquecimento ou resfriamento temporário que irá afastar nossas previsões dos dados que efetivamente irão se verificar, porém acreditamos que, no longo prazo, esses efeitos se anulem e que a taxa de desemprego, converge para a taxa natural de desemprego esperada.

3. Metodologia

3.1. Algoritmos de Alisamentos Exponencial

Alguns algoritmos de alisamento exponencial são usados na tentativa de obter previsões de séries de maneira ad hoc, ou seja, sem grande justificativa probabilística, mas que tendem a funcionar bem. Tratam-se de algoritmos baseado em suposições sobre os componentes não observáveis que determinam a dinâmica da série, sendo estimados a cada instante de tempo. Tais componentes não precisam ser necessariamente fixos, porém suas alterações devem ser imprevisíveis.

Uma das classes desses algoritmos é o Simplex que possui a seguinte ideia: a série evolui em torno de um nível e uma componente não observável. A estimativa do nível da equação (1) é uma média ponderada entre a observação recente (y_t) e a estimativa anterior do nível. A equação (1) é representada a seguir:

$$N_t = \alpha y_t + (1 - \alpha)N_{t-1} \quad (1)$$

Observe que $0 < \alpha < 1$. Quanto maior α , maior o peso dado à observação mais recente e, conseqüentemente, menor o peso dado à estimativa anterior N_{t-1} . Seja $\hat{y}_t(h)$ a previsão de y_{t+h} feita no tempo t por:

$$\hat{y}_t(h) = N_t, \quad h = 1, 2, 3, \dots \quad (2)$$

A forma de correção de erros é dada por:

$$N_t = N_{t-1} + \alpha \epsilon_t \quad (3)$$

em que ϵ_t é o erro de previsão um passo à frente. Escolhe-se o α de forma a minimizar a soma dos erros de previsão um passo a frente ao quadrado. Essa forma de correção de erros é um mecanismo de autoajuste em cada ponto do tempo. E a intensidade é controlada pelo α .

Outra classe de algoritmo é chamada de *Holt* cuja suposição é a seguinte: y_t se desenvolve ao longo do tempo de forma determinada por duas componentes não observáveis, a saber, nível e tendência. As estimativas dessas componentes são dadas por:

$$N_t = \alpha y_t + (1 - \alpha)(t - 1 + N_{t-1}); 0 < \alpha < 1 \quad (4)$$

$$T_t = \beta(t - N_{t-1}) + (1 - \beta)T_{t-1}; 0 < \beta < 1 \quad (5)$$

Observe que as estimativas, em cada instante de tempo, sempre serão uma média ponderada entre o estabelecido e a surpresa (y_t). As previsões agora refletirão o fato de que a tendência está sendo considerada.

$$\hat{y}_t(h) = N_t + hT_t, \quad h = 1,2,3, \dots \quad (6)$$

Então, a previsão vai crescer ou decrescer, dependendo se a tendência naquele instante de tempo está crescente ou decrescente. A forma de correção de erros para *Holt* é a seguinte:

$$N_t = N_{t-1} + T_{t-1} + \alpha\epsilon_t \quad (7)$$

$$T_t = T_{t-1} + \alpha\beta\epsilon_t \quad (8)$$

Assim, escolhem-se α e β de forma a minimizar a soma dos erros de previsão um passo à frente ao quadrado. Perceba nas equações (7) e (8) que a magnitude do nível é controlada por α e que a magnitude da tendência é controlada por α e β . Uma alternativa é o algoritmo de Pegels (1969), ou seja, o de *Holt* com tendência multiplicativa. As estimativas desse algoritmo são dadas por:

$$N_t = \alpha y_t + (1 - \alpha)(N_{t-1}T_{t-1}) \quad (9)$$

$$T_t = \beta \frac{N_t}{N_{t-1}} + (1 - \beta)T_{t-1} \quad (10)$$

A equação de previsão de $y_t + h$ para esse tipo de algoritmo é dado por:

$$\hat{y}_t(h) = N_t T_t^h, \quad h = 1,2,3, \dots \quad (11)$$

A forma de correção de erros para esse tipo de algoritmo é dado por:

$$N_t = N_{t-1}T_{t-1} + \epsilon_t \quad (12)$$

$$T_t = T_{t-1} + \alpha\beta \frac{\epsilon_t}{N_{t-1}} \quad (13)$$

Além disso, há o algoritmo de Holt-Winters aditivo que possui três componentes não observáveis: nível, tendência e sazonalidade. Esses componentes, respectivamente, são dados por:

$$N_t = \alpha(t - F_{t-s}) + (1 - \alpha)(N_{t-1} + T_{t-1}) \quad (14)$$

$$T_t = \beta(t - N_{t-1}) + (1 - \beta)T_{t-1} \quad (15)$$

$$F_t = \gamma(y_t - N_t) + (1 - \gamma)F_{t-s}, 0 < \gamma < 1 \quad (16)$$

Em geral, $s = 12$ para dados mensais e $s = 4$ para dados trimestrais. A previsão de y_{t+h} para esse tipo de algoritmo é dada pela seguinte estrutura:

$$\hat{y} = \begin{cases} N_t + hT_t + F_{t+h-s}p/h = 1,2,3 \dots \\ N_t + hT_t + F_{t+h-2s}p/h = s + 1, \dots, 2s \\ N_t + hT_t + F_{t+h-3s}p/h = 2s + 1, \dots, 3s \\ \dots \dots \end{cases} \quad (17)$$

Aqui, tem-se que tomar cuidado com a divisão de $\hat{y}_t(h)$ em infinitas equações. Para prever um mês à frente, dois meses à frente e assim por diante, usa-se a primeira equação. Agora, para prever 13 meses à frente e assim por diante, usa-se a segunda equação. E assim se segue. A forma de correção de erros para esse tipo de algoritmo é dado por:

$$N_t = N_{t-1} + T_{t-1} + \alpha \epsilon_t \quad (18)$$

$$T_t = T_{t-1} + \alpha \beta \epsilon_t \quad (19)$$

$$F_t = F_{t-s} + \gamma(1 - \alpha)\epsilon_t \quad (20)$$

Assim, escolhem-se α , β e γ de forma a minimizar a soma dos erros de previsão um passo à frente ao quadrado. Além disso, será executado o algoritmo de alisamento exponencial de Holt-Winters multiplicativo. Seguem abaixo as estimativas das componentes não observáveis:

$$\begin{aligned} & N \\ (1 - \alpha)(t - 1 + T_{t-1}) & \quad (21) \\ N_t = \alpha \frac{y_t}{F_{t-s}} + & \end{aligned}$$

$$T_t = \beta(N_t - N_{t-1}) + (1 - \beta)T_{t-1} \quad (22)$$

$$F_t = \gamma \frac{y_t}{N_t} + (1 - \gamma)F_{t-s} \quad (23)$$

A previsão de y_{t+h} para esse tipo de algoritmo é dada pelo seguinte sistema:

$$\begin{aligned}
 & \frac{N}{(t + hT_t)F_{t+h-s}p/h} = 1,2,3 \dots \\
 & \frac{N}{(t + hT_t)F_{t+h-2s}p/h} = s + 1, \dots, 2s \quad (24) \\
 & \frac{N}{(t + hT_t)F_{t+h-3s}p/h} = 2s + 1, \dots, 3s \\
 & \hat{y} = \{ \dots \dots
 \end{aligned}$$

Para o algoritmo de Holt-Winters multiplicativo, a forma de correção de erros é dada pelas seguintes equações:

$$N_t = N_{t-1} + T_{t-1} + \alpha \frac{\epsilon_t}{F_{t-s}} \quad (25)$$

$$T_t = T_{t-1} + \alpha\beta \frac{\epsilon_t}{F_{t-s}} \quad (26)$$

$$F_t = F_{t-s} + \gamma(1 - \alpha) \frac{\epsilon_t}{N_t} \quad (27)$$

Esses algoritmos são bem úteis quando se quer prever poucos passos à frente. Mas às vezes, é desejável prever muitos passos à frente. Conseqüentemente, ocorrem superprevisões. Recentemente na literatura, vem surgindo modificações desses algoritmos. São os chamados algoritmos com atenuação de tendência. Neste trabalho, será utilizada a atenuação de tendência sobre o algoritmo de Pegels (Taylor, 2003). As estimativas seguirão da seguinte forma:

$$N_t = \alpha y_t + (1 - \alpha)N_{t-1}T_{t-1}^\phi \quad (28)$$

$$T_t = \beta \left(\frac{N_t}{N_{t-1}} \right) + (1 - \beta)T_{t-1}^\phi \quad (29)$$

Se $\phi \in (0, 1)$, tem-se atenuação de tendência. Se $\phi = 1$ tem-se o algoritmo de Pegels. A equação de y_{t+h} para esse tipo de algoritmo é dado por:

$$\hat{y}_t(h) = N_t T_t \sum_{j=1}^h \phi^j \quad (30)$$

Quando $h \rightarrow \infty$, as previsões convergem a para $N_t T_t \frac{\phi}{1-\phi}$.

3.2. Modelos estocásticos

São modelos que possuem uma representação da realidade. Então, a partir da inferência dos dados, podemos imitar a aleatoriedade do sistema real através de distribuições probabilísticas. Toda modelagem estocástica é baseada em dois pilares. Um pilar diz que as previsões são funções lineares das observações passadas, pois funções lineares, em muitos casos, aproximam bem funções não-lineares. O outro pilar assegura que há uma preferência por modelos simples, compactos. É mais fácil encontrar modelos enormes que explicam bem o passado, porém tais modelos tendem a ser muito ruins para prever o futuro. A seguir, serão apresentadas algumas classes de modelos de previsão.

3.3. Modelos Autorregressivos (AR)

De forma geral, o modelo autorregressivo de ordem p ($AR(p)$) é dado por:

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + u_t \quad (31)$$

em que c é uma constante, ϕ_p é um parâmetro autoregressivo (p) e u_t segue um processo ruído branco que é o processo bem mais comportado que existe. É um processo puramente aleatório, sem nenhuma correlação temporal.

O modelo da equação (31) possui a dependência serial porque é uma ligação entre diferentes instantes de tempo. Essa ligação é chamada de Dinâmica. Além disso, um processo é estacionário quando possui média, variância e covariância constantes. Nessa classe de modelos, a estacionaridade depende dos valores dos parâmetros. Então, pode-se escrever a equação (31) da seguinte maneira:

$$\phi(B)y_t = c + u_t, \quad (32)$$

em que o polinômio autorregressivo $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ e B é um operador de defasagem. Com isso, y_t é estacionário se todas as raízes de $\phi(B) = 0$ estiverem fora do círculo unitário.

3.4. Modelos de Médias Móveis (MA)

De forma geral, o processo de médias móveis de ordem q ($MA(q)$) é dado por:

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \dots + \theta_q u_{t-q} \quad (33)$$

em μ é um constante, e u_t segue um processo ruído branco. Em resumo, um processo MA é uma combinação linear dos termos de erro de um ruído branco. Então, aqui a dinâmica está sobre o ruído branco que é algo não observável. No processo AR, a dinâmica está sobre y que é algo observável. Além disso, um processo MA é sempre estacionário, mas nem todo processo MA está associado a uma única estrutura de autocorrelação. Para que ele seja invertível, todas as raízes $\theta(u) = 0$ têm que estar fora do círculo unitário.

3.4.1. Processo Autorregressivo e de Médias Móveis (ARMA)

De forma geral, um processo ARMA (p,q) é dado por:

$$\phi(B)y_t = c + \theta(B)u_t, \quad (34)$$

em que c é uma constante, u_t segue um processo ruído branco, $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ e $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$. Nessa classe de modelos, a condição de estacionaridade e a condição de invertibilidade são as mesmas que nas classes citadas anteriormente. Além disso, o processo ruído branco ($p = 0$ e $q = 0$), o processo AR ($q = 0$) e o processo MA ($p = 0$) são casos particulares da classe ARMA. Mais ainda, com a inclusão de regressores, esse processo passa a ser chamado de modelo ARMAX(p,q). Esses regressores podem ser do tipo dummy. Se os regressores não forem tipo dummy, haverá acumulação de erros de previsão no modelo dinâmico. Então, recomenda-se apenas a inclusão de regressores que realmente forem muito importantes, portanto, tem que ter muito cuidado quando a previsão for realizada.

3.4.2. Processo Autorregressivo Integrado de Médias Móveis (ARIMA)

Foram vistos, até então, os processos que trabalham com séries estacionárias. Todavia, a maioria das séries são não estacionárias. Dessa forma, por meio de um processo de diferenciação, parte dessas séries não estacionárias pode ser transformada em séries estacionárias. Esse tipo de série são as chamadas séries integradas. Então, se

y_t é estacionário, esse processo será integrado de ordem zero, $I(0)$. Se y_t não for estacionária, mas Δy_t (a primeira diferença de y_t) for, então, $y_t \sim I(1)$.

Diz-se que y_t é integrado de ordem d : $y_t \sim I(d)$. Em geral, $d = 0$ ou $d = 1$; raramente $d = 2$.

De forma geral, um processo ARIMA (p,d,q) é dado por:

$$\phi(B)(\Delta^d y_t - \mu) = \theta(B)u_t, \quad (35)$$

em que μ é uma constante, u_t segue um processo de ruído branco, $\Delta^d y_t$ é o número de diferenças do processo y_t , $\phi(B)$ e $\theta(B)$ são os polinômios autorregressivos e de médias móveis, respectivamente. Neste modelo, pode-se, também, incluir regressores e obter a classe ARIMAX (p,d,q) .

3.5. Modelos SARIMA

Aqui, pode-se levar em consideração que os dados possam ter algum comportamento sazonal que são movimentos recorrentes de frequência intra-anual. De forma geral, o modelo SARIMA $(p,d,q) \times (PDQ)$ é representado da seguinte maneira:

$$\frac{1 - \phi_1 B^s - \dots - \phi_p B^{ps}}{\phi(B)} (1 - B)^d (1 - B^s)^D y_t = \theta(B) (1 + \theta_1 B + \dots + \theta_q B^{qs}) u_t \quad (36)$$

em que u_t é processo de ruído branco, B é o operador de defasagem e s é o período de sazonalidade. Observe que esse modelo apresenta duas dinâmicas, uma delas é a dinâmica ordinária que foi a que se tinha visto até agora, a outra é a dinâmica sazonal. Também se pode incluir regressores ao modelo e essa classe passa a ser chamada de SARIMAX.

3.5.1. Estratégia de Modelagem

Dada uma determinada série temporal, o que será apresentado agora é como modelar e fazer previsões de tal série. O método a ser seguido será o Box-Jenkins. Inicialmente, deve-se postular uma classe de modelos. Se seus dados foram sazonais, a

classe deverá ser SARIMA. Se seus dados não forem sazonais, postula-se um ARIMA. No caso de que seus dados sejam estacionários, temos então uma classe ARMA a ser trabalhada.

Depois de postular uma dessas classes de modelo, se identifica o modelo particular que melhor representa a dinâmica da série. Nessa fase, devem-se encontrar os valores de p , d e q .

Para encontrar o valor de d , utiliza-se o teste de Dickey-Fuller aumentado (ADF), que é um teste mais amplo do teste DF de Dickey-Fuller (1979). Então, no teste ADF vai ser testada a seguinte hipótese:

$$H_0: y_t I(1), \text{ ou seja, } y_t \text{ possui raiz unitária.}$$

$$H_1: y_t I(0), \text{ ou seja, } y_t \text{ não possui raiz unitária.}$$

A regressão utilizada no teste de DF era dada por $\Delta y_t = \varphi y_{t-1} + u_t$, em que u_t segue um ruído branco. Said e Dickey (1984) modificaram essa regressão, ampliando de três formas distintas:

$$\Delta y_t = \varphi y_{t-1} + \sum_{j=1}^m \gamma_j \Delta y_{t-j} + u_t \quad (37)$$

$$\Delta y_t = \mu + \varphi y_{t-1} + \sum_{j=1}^m \gamma_j \Delta y_{t-j} + u_t \quad (38)$$

$$\Delta y_t = \mu + \sigma + \varphi y_{t-1} + \sum_{j=1}^m \gamma_j \Delta y_{t-j} + u_t \quad (39)$$

A seguir, estima-se o modelo ampliado escolhido e calcula-se a estatística de teste. Existe uma regra básica para escolher o modelo ampliado. Se a série possuir forte tendência, usa-se a equação (39). Caso contrário, aplica-se a equação (38). A equação (37) quase nunca é usada. A estatística de teste é dada por:

$$\tau_{ADF} = \frac{\hat{\phi}}{\sqrt{\text{var}(\hat{\phi})}} \quad (40)$$

Said e Dickey (1984) mostraram que a distribuição limite, que não é uma distribuição normal, de τ_{ADF} é a mesma de τ_{DF} . A implicação disso é que pode-se usar os mesmos valores críticos do teste de DF. Então, rejeita-se a hipótese nula se τ_{ADF} for mais negativa do que o valor crítico tabelado por Dickey e Fuller. Outra forma de

determinar o valor de d é através da seguinte regra de bolso: se o correlograma cair lentamente, $d = 1$. Se o correlograma da série cair rapidamente $d = 0$.

Em relação aos valores de p , q , P e Q , utilizam-se os critérios de seleção de modelos AIC, BIC e AICc. Para simplificação serão determinados apenas p e q . O critério de informação de Akaike (1974) é calculado da seguinte forma:

$$AIC = -2\log\hat{L} + 2(p + q), \quad (41)$$

em que \hat{L} é a função de verossimilhança maximizada. Então, estimam-se vários modelos e calcula-se o AIC. O primeiro termo do lado direito da equação (41) representa o ganho da ampliação do modelo. O segundo termo representa a penalização por ampliar o modelo.

O critério de informação bayesiano ou Schwarz (1978) é dado pela seguinte equação:

$$BIC = -2\log\hat{L} + (p + q)\log T, \quad (42)$$

em que T é número de observações. A ideia do BIC consiste em, quanto maior for a série, maior será o desafio. O desafio no AIC é constante, a saber, $(p + q)$. Então, há necessidade de fazer esse desafio aumentar. Entretanto, esse aumento não pode ser rápido. Esse aumento dar-se-á de forma lenta. Então, Schwarz encontrou que a velocidade ótima é $\log T$, que aumenta lentamente com T .

Por fim, o critério de informação de Akaike corrigido (Hurvich, 1989) é calculado da seguinte forma:

$$AIC_c = \frac{-2\log\hat{L} + 2(p+q)T}{T-p-q-1} \quad (43)$$

Comparando esses três critérios acima, observa-se que o AIC tende a superdimensionar o modelo (até assintoticamente). Significa que o AIC tende a selecionar modelos um pouco mais parametrizados. O AICc é uma forma corrigida do AIC, com fator de correção que desaparece quando $T \rightarrow \infty$. Nesse caso, o AIC fica perto do AICc. Mas, com T pequeno, essa condição melhora o desempenho do AIC. E como esse fator de correção pode desaparecer, o AICc possui as mesmas propriedades assintóticas do AIC. Além disso, a seleção de modelos é consistente quando feita apenas

via BIC. Então, isso é uma grande vantagem do BIC sobre o AIC. Também, o BIC tende a selecionar modelos mais compactos que o AIC, pois $\log T > 2$, para $T > 8$ observações. Com Isso, o fator de penalidade do BIC é maior do que o fator de penalidade do AIC.

Em relação à estimação, os principais métodos utilizados são os mínimos quadrados (MQ) condicional, os MQ não-condicional, máxima verossimilhança (MV) aproximada e MV exata. De acordo com Ansley e Newbold (1980), simulações foram realizadas com processos AR, MA e ARMA. Eles concluíram que, em geral, MV funciona melhor do que MQ. Além disso, eles concluíram que a estimação pontual funciona bem, mas a estimação intervalar não. A implicação disso é que testes tipo t ou z também não são confiáveis. É importante lembrar de que existe uma relação muito próxima entre teste de hipótese e intervalo de confiança. Quando se está testando um parâmetro, pode-se fazer isso em intervalos. Então, se o intervalo não funcionar bem, o teste de hipótese também não funcionará bem. É por isso que utiliza-se dos critérios AIC, AICc e BIC.

Há duas razões por trás dessa segunda conclusão de Ansley e Newbold (1980). Uma delas é que quando se está perto da região de não estacionaridade e/ou da região de invertibilidade, as variâncias assintóticas dos estimadores ficam muito pequenas e qualquer viés, por menor que seja, influencia bastante a cobertura do intervalo desejado. Outra razão é que quando se faz esses testes e intervalos, é utilizado uma aproximação normal para a distribuição do estimador. Se T não for grande, essa aproximação pode ser ruim.

Há uma má notícia além dessas duas conclusões de Anseley e Newbold (1980). É que poucos softwares como o SPSS e o R usam MV porque, numericamente, é muito complicado. O default do R é o seguinte. Ele estima os parâmetros por MQ condicional, pega essas estimativas e vai usar como valores iniciais por MV exata.

A terceira etapa do método de Box-Jenkins refere-se à verificação do diagnóstico. Essa etapa procura analisar se o modelo estimado descreve adequadamente a dinâmica presente nos dados. Para essa análise, procura-se ver se há correlação serial dos resíduos. Caso não exista tal correlação, use o modelo para fazer previsões, pois esses resíduos serão ruído branco. Caso contrário, busque um novo modelo.

Uma forma de verificação de diagnóstico é bastante conhecida como testes de Portmanteau que consistem em testar conjuntamente a significância das primeiras

autocorrelações residuais. O teste que será utilizado neste trabalho é o teste mais usado na prática. É o teste de Ljung e Box (1978). A estatística de teste é dada por:

$$Q_{LB} = T(T + 2) \sum_{j=1}^m (T - j)^{-1} \hat{\rho}_j^2, \quad (44)$$

em que $\hat{\rho}_j$ é a j-ésima correlação serial, com $j=1,2,3,\dots$.

O valor crítico é obtido da distribuição da estatística de teste quando H_0 é verdadeira. Só que a distribuição nula dessa estatística, assim como para as demais estatísticas de Portmanteau, é desconhecida. Então, propõe-se uma coisa parecida com o ideal. Sabe-se que para modelos ARMA, se esses modelos estiverem corretos, $Q_{LB} \underset{d}{\sim} \chi^2_{m-p-q}$. Isso serve para os demais testes de Portmanteau. Com isso, tem-se um teste aproximado, Rejeite H_0 se $Q_{LB} > \chi^2_{1-\alpha; m-p-q}$.

3.5.2. Análise descritiva da série temporal

Os dados da Pesquisa Mensal de Emprego (PME) utilizado no artigo foram divulgados pelo Instituto de Pesquisa Econômica Aplicada (IPEA) no período de março de 2002 até setembro de 2014 de 6 regiões metropolitanas do Brasil (São Paulo, Belo Horizonte, Porto Alegre, Salvador, Rio de Janeiro e Recife). A tabela 1 abaixo mostra algumas medidas de tendência central e de dispersão.

Tabela 1: Estatística descritiva

| | Mínimo | Mediana | Média | Máximo | Desvio Padrão |
|----------------|--------|---------|-------|--------|---------------|
| Série completa | 4.30 | 8.20 | 8.41 | 13.10 | 2.52 |
| Tendência | 4.59 | 8.10 | 8.40 | 12.47 | 2.46 |

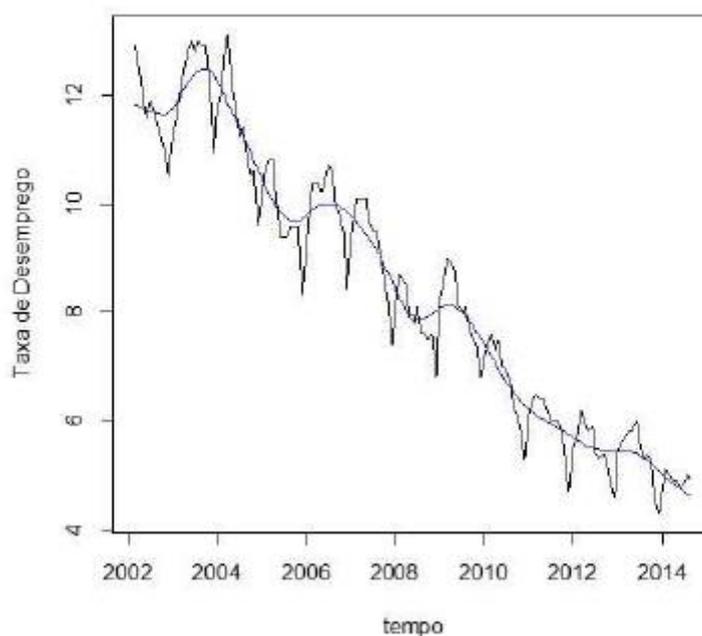
Fonte: elaboração própria com base nos da PME.

Vemos que a taxa de desemprego médio é cerca de 8.41% com um mínimo de 4.3% e máximo de 13.10%, o desvio padrão é em torno de 2.52%. Vale lembrar que essas medidas levam em consideração todos os componentes da série, sendo eles, a sazonalidade, o componente aleatório ou irregular (choques de produtividade) e a

tendência de longo prazo da série. Uma análise realista da taxa de desemprego no Brasil levaria em consideração apenas a tendência de longo prazo da série. Tendo em vista esse objetivo, a figura 1A do apêndice decompõe a série em 3 componentes: Sazonalidade, Tendência e componente irregular. A linha 2 mostra as estatísticas descritivas com base na tendência de longo prazo. Note que as medidas são bastante semelhantes, dando um forte indício de que choques aleatórios e ou de produtividade não impactam de maneira significativa no índice mensal de emprego, reforçando a ideia de Phelps (1995) de que a taxa natural de desemprego funciona como um atrator para a taxa corrente e choques monetários causam apenas desvios temporários dela.

A figura 1 mostra o comportamento da taxa de desemprego e seu componente de tendência de longo prazo.

Figura 1: Taxa de desemprego no Brasil



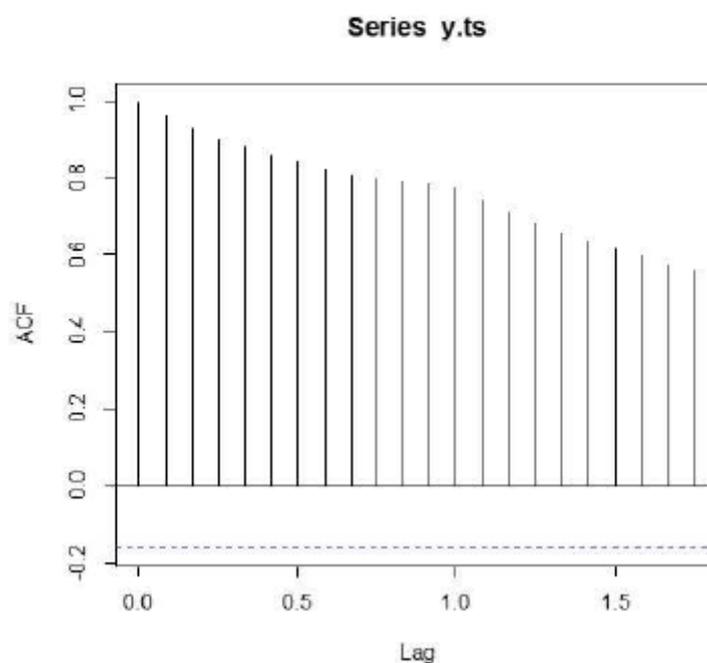
É possível perceber que a taxa de desemprego no Brasil vem caindo significativamente durante os últimos doze anos, convergindo para uma taxa em torno de 4,5%, refletindo a taxa natural de desemprego e equilíbrio de mercado proposta por Friedman (1968). A estabilidade macroeconômica dos últimos dez anos, os ganhos em termos de troca (estimulado em grande parte pelo crescimento Chinês) devido ao

elevado preços das commodities, queda na taxa de juros internacional e uma economia brasileira mais sólida são razões que justificam essa queda na taxa de desemprego.

As taxas de desemprego costumam ser menores nos períodos de Dezembro (indicando sazonalidade da série), isso acontece devido a um aquecimento na demanda por bens e serviços, ocasionando um aumento na procura por trabalhadores por parte das empresas e conseqüentemente um aumento na taxa de emprego. Também pode-se observar que houve alguns períodos em que a taxa de desemprego aumentou, como por exemplo, no início de 2008, devido à crise norte-americana.

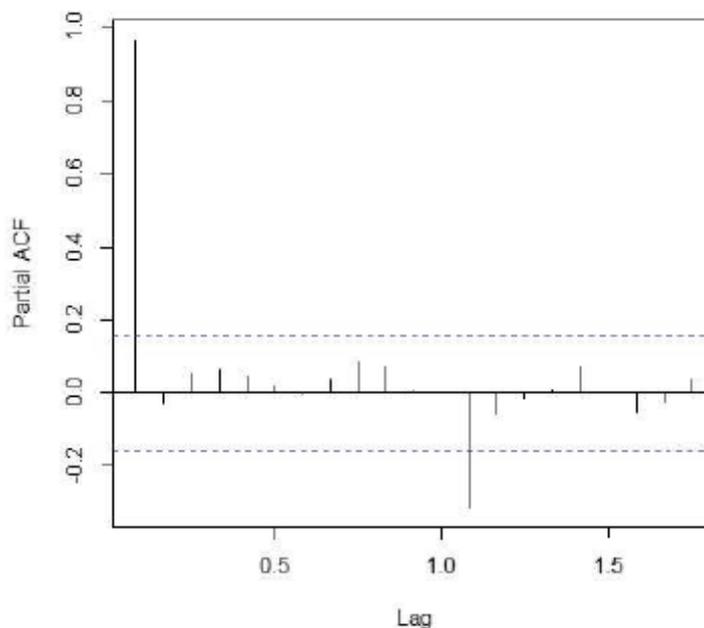
Outro aspecto importante de uma série é sua estacionaridade⁶, observa-se um indício de não estacionaridade na figura 1, pois a série não aparenta se situar em torno de uma média (vem caindo ao longo do tempo). Além disso, a função de autocorrelação decai lentamente, suportando a ideia de não estacionaridade. A Figura 2 e 3 descrevem o correlograma e o correlograma parcial, respectivamente:

Figura 2: Correlograma (acf)



⁶ Uma série temporal $(y_t; t \in T)$ é estacionária se possui média e variância constantes e se sua função de covariância não depende de "t".

Figura 3: Correlograma Parcial (pacf)
Series y.ts



É claro que esses indícios de não-estacionariedade não são suficientes, eles servem apenas para indicar um possível comportamento da série. Testes de raiz unitária são mais precisos e mais formais para a detecção de estacionariedade. O artigo utilizou o teste Dickey-Fuller Aumentado (ADF), para implementar o teste utilizamos uma regressão linear que incorpora uma constante e uma variável de tendência linear e para a escolha do número de defasagem foi utilizado um critério ad-hoc dado por:

$$\text{Número de defasagens} = \text{trunc}(T - 1)^{\frac{1}{5}},$$

onde T é o número de observações, e “trunc” é uma função que remove a parte fracionária do contradomínio. Como foi visto na seção de metodologia, caso ocorra não rejeição da hipótese nula, há não estacionariedade. A tabela 2 mostra os resultados do teste:

Tabela 2: Teste Dickey- Fuller Aumentado

| | |
|---------------------------|---------|
| Estatística Dickey-Fuller | -3,9053 |
| Número de defasagens | 5 |

p-valor 0,01577

Fonte: Elaboração própria

Como a estatística do teste é suficientemente pequena e o p-valor é menor que 5% rejeita-se a hipótese nula de não-estacionaridade a um nível de 5% de significância. Portanto, conclui-se que a série da taxa mensal de desemprego é estacionária.

3.5.3. Algoritmos de alisamento exponencial: seleção e resultados

Como foi visto na metodologia, os algoritmos ad hoc são técnicas largamente utilizadas pelo setor industrial com objetivo de fazer previsões. Esta seção fará uso desses algoritmos para a previsão das taxas de desemprego mensal. O primeiro passo para uma análise desse tipo é encontrar o algoritmo mais adequado para descrever a série temporal. Segundo Hyndman e Khandakar (2008), buscou-se selecionar o algoritmo com menor AIC (Akaike Information Criterion) e compará-lo com outros algoritmos alternativos. Outra maneira de comparar algoritmos é verificar sua capacidade de prever bem o futuro, os erros de previsão são medidas que resumizam essa capacidade, para isso, comparamos a previsão dos últimos 3 meses com os valores que realmente ocorreram nesses últimos 3 meses. A tabela 3 a seguir mostra os resultados para alguns algoritmos selecionados:

Tabela 3: Seleção dos Algoritmos

| | A1(M,N,A) | A2(M,Ad,A) | A3(M,Md,M) | A4(M,Ad,M) | A5(M,Md,N) |
|-------------------|-----------|------------|------------|------------|------------|
| AIC | 371,38 | 367,74 | 371,38 | 371,62 | 524,44 |
| AICc | 376,09 | 371,9 | 376,08 | 376,32 | 524,87 |
| BIC | 422,33 | 415,7 | 422,33 | 422,57 | 539,43 |
| Previsão 07/14 | 4,74 | 4,70 | 4,74 | 4,73 | 4,8 |
| Previsão 08/14 | 4,70 | 4,64 | 4,70 | 4,68 | 4,8 |
| Previsão 09/14 | 4,61 | 4,54 | 4,61 | 4,59 | 4,8 |
| Erro 07/14 | 3,26% | 4,08% | 3,26% | 3,47% | 2,04% |
| Erro 08/14 | 6,00% | 7,20% | 6,00% | 6,40% | 4,0% |
| Erro 09/14 | 5,91% | 7,34% | 5,91% | 6,33% | 2,4% |

Fonte: Elaboração própria

A primeira coluna da tabela 3 mostra o algoritmo com menores critérios de comparação (AIC, AICc e BIC), o algoritmo selecionado não possui tendência e possui sazonalidade aditiva, os erros de previsão parecem aumentar com o tempo. Os resultados do estudo de Taylor (2003) mostram que algoritmos com atenuação de tendência costumam ser mais robustos em termos de previsão do que os algoritmos sem atenuação. Logo, consideram-se esses algoritmos para fins de comparação. A segunda coluna da tabela 3 mostra o algoritmo com atenuação de tendência e com menor critério de comparação, os erros de previsão, porém os erros de previsão altos e aumentam conforme o período da previsão aumenta. A terceira e a quarta coluna mostram algoritmos mais complexos, ambos possuem sazonalidade multiplicativa, sendo que o terceiro possui tendência multiplicativa enquanto o quarto possui tendência aditiva. A coluna 5 mostra o algoritmo proposto por Taylor (2003), um Pegels com atenuação de tendência. Assim, como foi encontrado no estudo de Taylor (2003), esse é o algoritmo com menor erro de previsão.

Como foi visto na seção anterior, a série de taxa de desemprego mensal possui sazonalidade e tendência, sendo assim optou-se pela escolha do algoritmo 3 (que possui essas características). Além dessa vantagem, o algoritmo 3 possui critérios de comparação semelhantes aos algoritmo 1 e 2. O algoritmo 3 pode ser descrito pelas seguintes equações:

$$N_t = \alpha \left(\frac{y_t}{F_{t-m}} \right) + (1 - \alpha) N_{t-1} T_{t-1}^\phi \quad (45)$$

$$T_t = \beta \left(\frac{N_t}{N_{t-1}} \right) + (1 - \beta) T_{t-1}^\phi \quad (46)$$

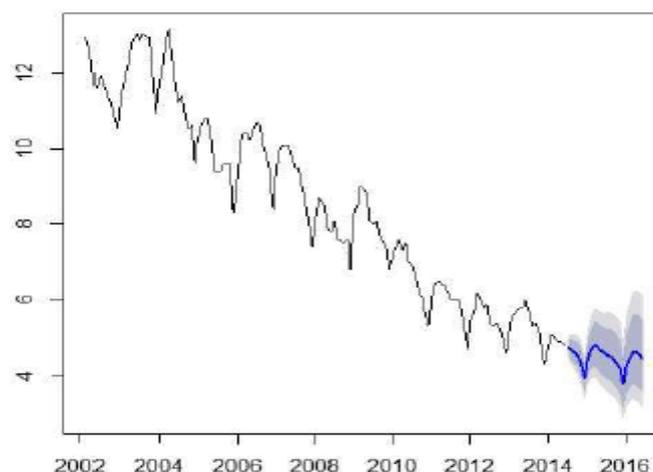
$$F_t = \gamma \left(\frac{y_t}{N_{t-1} T_{t-1}^\phi} \right) + (1 - \gamma) F_{t-1} \quad (47)$$

onde N_t é nível, T_t é a tendência e F_t é a sazonalidade. Tem-se que $\alpha = 0,88, \beta = 0,018, \gamma = 0,0001 e \phi = 0,98$. A previsão é dada por:

$$y_{t+h} = N_t T_{t-1}^{\phi h} F_{t-m+h} \quad (48)$$

A tabela 6 no apêndice mostra previsões de longo prazo e a figura 4 representa essas previsões graficamente:

Figura 4: Previsões de longo prazo com o modelo(M,Md,M)
Forecasts from ETS(M,Md,M)



A figura 4 mostra as previsões da taxa de desemprego utilizando o algoritmo de alisamento exponencial do tipo (M,Md,M). Esse algoritmo aponta para uma variabilidade baixa da taxa de desemprego, com máximo em Março de 2015 (4,81%) e mínimo em Dezembro de 2015 (3,83%). Observe que a taxa de desemprego começa a variar em torno de uma média aproximada de 4,5%, dando indicio de que esse é valor do chamado desemprego natural.

4. Modelos estocásticos de previsão – Seleção e resultados

Outra forma de obter previsões da taxa de desemprego mensal é utilizar modelos estocásticos de séries temporais. Este artigo utilizou-se da metodologia Box-Jenkins. O primeiro passo da modelagem é postular uma classe de modelos estocásticos. Como foi visto na seção 5, a serie temporal da taxa de desemprego é sazonal e estacionária, logo temos que postular um modelo do tipo SARMA(p,q,P,Q). Em seguida, o processo de modelagem exige encontrar os valores ó de p,q,P e Q, utilizou-se o algoritmo de Hyndman e Khandakar (2008), que identifica o modelo com menor Akaike Criterion Information (AIC), os valores de de p, q, P e Q são encontrados com base nesse modelo. A tabela 4 mostra o modelo encontrado com tal critério:

Tabela 4: Modelo SARIMA (3,0,1)(2,0,1)

| Coeficientes | | Erros de previsão | |
|-----------------|--------|-------------------|-------|
| AR ₁ | 1,1960 | Jul/14 | 7.14% |

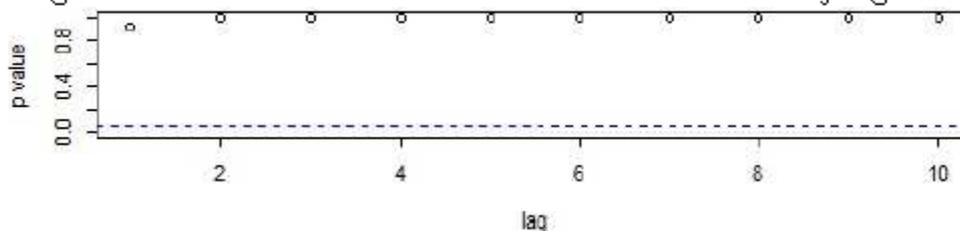
| | | | |
|------------------|---------|--------------------------------|--------|
| AR ₂ | -0,2349 | Set/14 | 11.20% |
| AR ₃ | 0,0303 | Ago/14 | 10.20% |
| MA ₁ | -0,1994 | Crítérios de comparação | |
| Sar ₁ | 0,8287 | AIC | 93,91 |
| Sar ₂ | 0,1323 | AICc | 94,95 |
| Sma ₁ | -0,6191 | BIC | 117,89 |
| σ^2 | 0,08571 | log likelihood | -38,96 |

Fonte: Resultados da pesquisa

O critério de estimação utilizado foi à máxima verossimilhança. O modelo identificado é um SARIMA (3,0,1)(2,0,1). Este modelo possui três coeficientes autorregressivos ordinários e dois sazonais, possui um coeficiente de média móvel ordinal e um sazonal. Note que o modelo analisado não possui nenhuma diferenciação, uma vez que a série taxa de desemprego é estacionária. Os erros de previsão são menores se compararmos com outros modelos estocásticos.

A última etapa da metodologia Box-Jenkins é verificar se o modelo consegue captar bem a dinâmica contida nos dados, caso seja válido, podemos fazer previsões. Para isso, foi utilizado o teste de Ljung-Box (1978), que é uma pequena alteração do teste de BoxPierce, este teste é utilizado porque apresenta melhores resultados que o segundo. Como foi visto na metodologia, ele testa a hipótese nula de que os resíduos não têm correlação serial (indício de um bom modelo) contra a hipótese nula de que os resíduos possuem correlação serial. A figura 5 mostra o p-valor associado a cada grau de defasagem da autocorrelação, como todo p-valor é superior a 5%, não se rejeita a hipótese nula. Sendo assim, o modelo SARIMA (3,0,1)(2,0,1) descreve bem a dinâmica dos dados e portanto podemos fazer previsões.

Figura 5: P-valor associados com a estatística de Ljung-Box



A tabela 6 e a figura 2, no apêndice, mostram o comportamento apresentado pelo modelo estocástico. A taxa de desemprego máxima prevista é de 4.57% em março de 2015 e mínima de 3.18% em dezembro de 2015. Assim como foi visto na seção anterior, nota-se que a taxa de desemprego fica em torno de 4,5%.

Tendo em vista a discussão feita na revisão da literatura sobre uma possível existência de um *trade-off* entre taxa de desemprego e taxa de inflação, optou-se pela estimação de um modelo que incluísse a taxa de inflação como regressora, essa classe de modelo é denominada SARIMAX. A tabela 5 retrata o modelo SARIMAX (4,0,2)(2,0,2) que é o modelo com menor AIC.

Tabela 5: Modelo SARIMAX (4,0,2)(2,0,2)

| Coeficientes | | Erros de previsão | |
|------------------|---------|--------------------------------|--------|
| AR ₁ | 0,0532 | Jul/14 | 6,32% |
| AR ₂ | 0,3462 | | |
| AR ₃ | 0,6454 | set/14 | 11% |
| AR ₄ | -0,0746 | | |
| MA ₁ | 1,0488 | ago/14 | 9,80% |
| MA ₂ | 0,6926 | Crítérios de comparação | |
| Sar ₁ | 0,1374 | AIC | 104,55 |
| Sar ₂ | 0,7704 | | |
| Sma ₁ | 0,1623 | AICc | 107,27 |
| Sma ₂ | -0,4869 | | |
| Intercepto | 85.376 | BIC | 143,51 |
| Inflação | -0,1029 | | |
| σ^2 | 0,08819 | log likelihood | -39,28 |

Fonte: Resultados da pesquisa

Note que os critérios de comparação são maiores que o modelo sem o regressor, taxa de inflação. Os erros de previsão são menores do que o modelo proposto anteriormente, uma vez que este novo modelo apresenta um melhor ajuste. Um resultado interessante deste modelo é o sinal negativo do coeficiente da inflação. Este fato empírico confirma o *trade-off* existente entre inflação e desemprego. A ideia associada a isso é o fato de que uma inflação elevada está associada ao aquecimento da economia, que por sua vez, pode aumentar a procura por emprego, fazendo com que a taxa de desemprego seja menor. Este novo modelo também passou no teste de Ljung-Box, ou seja, a hipótese nula não foi rejeitada (resíduos não possuem correlação serial), portanto o modelo está apto para descrever a dinâmica dos dados e com isso pode ser utilizado para fazer previsões.

5. CONSIDERAÇÕES FINAIS

Muitas decisões econômicas são baseadas em previsões sobre o caminho que a economia pode trilhar. Este artigo investigou o comportamento da taxa de desemprego com diferentes critérios de previsão. Para tanto, foram utilizados algoritmos de alisamento exponencial e modelos estocásticos. O algoritmo selecionado foi Holt-Winter com sazonalidade e tendência multiplicativa atenuada. Os modelos com fundamentação probabilística selecionados foram o SARIMA (3,0,1)(2,0,1) e um SARIMAX (4,0,2)(2,0,2) com o regressor taxa mensal de inflação.

Os resultados encontrados no artigo mostram que tanto os modelos estocásticos como o algoritmo ad hoc representam bem a dinâmica dos dados e portanto apresentam boas previsões para o comportamento da série. Um resultado bastante interessante e que corrobora com a teoria econômica é a evidência do trade-off entre a taxa de desemprego e a inflação, uma vez que foi encontrado uma correlação negativa entre essas variáveis. Temos também o fato da taxa de desemprego ser estacionária, com uma média de 4,5%, sugerindo a existência de uma taxa natural de desemprego, que é a ideia proposta por Friedman (1968).

6. REFERÊNCIAS

AKAIKE, Hirotugu. A new look at the statistical model identification. **Automatic Control**, IEEE Transactions on, v. 19, n. 6, p. 716-723, 1974.

C.F e P. Newbold Ansley. Finite sample properties of estimators for autoregressive moving average models. *journal of econometrics*, 13, 159-183. 1980.

Marcel Henrique Becker. Modelos para previsão em séries temporais: uma aplicação para a taxa de desemprego na região metropolitana de porto alegre.

IBGE. Instituto Brasileiro de Geografia e Estatística. Acessado em 12/12/2014.

D.A e W.A. Fuller Dickey. Distribution of the estimators for autoregressive time series with a unit root. *journal of the american statistical association*, 427-431. 1979.

Said E. e David A. Dickey. Testing for unit roots in autoregressive moving average models of unknown order. *biometrika*, 71, 599-607. 1984.

Cribari-Neto F. Notas de aula de econometria. pimes/ufpe. 2014.

Joilson Fraga, Gilberto J. Dias. Taxa de desemprego e a escolaridade dos desempregados nos estados brasileiros: estimativas dinâmicas de dados em painéis.

M. Friedman. The role of monetary police. the american economic review, vol. 58, pp. 1-17. 1968.

C.M. e C.L. Tsai Hurvich. Regression and time series model selection in small samples. biometrika, 76, 297-307. 1989.

R.J. e Y. Khandakar Hyndman. Automatic time series forecasting: the forecast package for r. journal of statistical software, 27. 2008.

J. M. Keynes. A teoria geral do emprego, do juro e da moeda. 1aed. atlas. São Paulo.1982.

G.M. e G.E.P. Box Ljung. On a measure of lack off it in time series models. biometrika, 65, 297-303. 1978.

MATTOS, Fernando Augusto et al. Decomposição da taxa de crescimento do nível de emprego dos Estados Unidos e da União Européia desde o pós-guerra. 2008.

C. C. Pegels. Exponential forecasting: some new variations. management science, 15, 311-315. 1969.

E.S. Phelps. The structuralist theory of unemployment. the american economic review, vol. 85, pp. 226-231.

E.S. Phelps. Money-wage dynamics and labor-market equilibrium. the journal of political economy, vol. 76, pp. 678-711. 1968.

G. Schwarz. Estimating the dimension of a model. the annals of statistics,6, 461-464. 1978.

J. W. Taylor. Exponential smoothing with a damped multiplicative trend. international journal of forecasting, 19, 715-725. 2003.

Anexos

Apêndice A – Gráficos e Tabelas

Figura 6: Série decomposta em sazonalidade, tendência e componente irregular

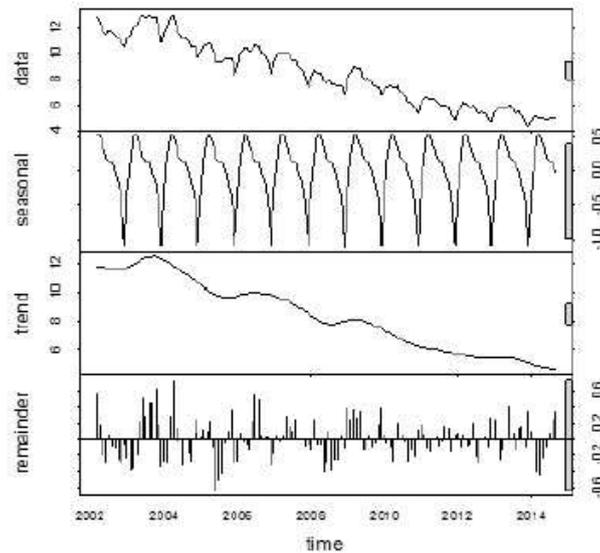


Figura 7: Previsões de longo prazo da taxa de desemprego segundo modelo SARIMA(3,0,1)(2,0,1)

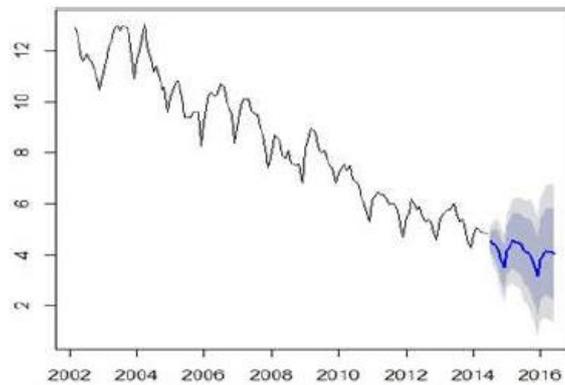


Tabela 6: Previsões com o algoritmo e com o modelo estocástico

| | (M,Md,M) | SARIMA(3,0,1)(2,0,1) |
|--------|----------|----------------------|
| jul/14 | 4,743636 | 4,554365 |
| Aug/14 | 4,696523 | 4,443921 |
| Sep/14 | 4,612725 | 4,403983 |
| Oct/14 | 4,52748 | 4,274452 |
| nov/14 | 4,347366 | 3,895028 |
| Dec/14 | 3,971734 | 3,51386 |
| jan/15 | 4,431088 | 4,173648 |
| Feb/15 | 4,66275 | 4,425339 |
| mar/15 | 4,813249 | 4,573595 |
| Apr/15 | 4,787398 | 4,522915 |
| May/15 | 4,704654 | 4,486047 |
| jun/15 | 4,598524 | 4,452558 |
| jul/15 | 4,550488 | 4,200077 |
| Aug/15 | 4,509041 | 4,071457 |
| Sep/15 | 4,432198 | 4,054457 |
| Oct/15 | 4,353763 | 3,923593 |
| nov/15 | 4,183832 | 3,532706 |
| Dec/15 | 3,825264 | 3,18004 |
| jan/16 | 4,270886 | 3,795814 |
| Feb/16 | 4,497483 | 4,046917 |
| mar/16 | 4,645999 | 4,159352 |
| Apr/16 | 4,624316 | 4,106908 |
| May/16 | 4,547543 | 4,079115 |
| jun/16 | 4,447977 | 4,040863 |
