

**AÇÃO 2.1 - AVALIAÇÃO DOS RESULTADOS DO
NOVO MODELO DE PREVISÃO E
MONITORAMENTO DE INDICADORES DE
RECEITAS MUNICIPAIS**

**PRODUTO 2 – AVALIAÇÃO PRELIMINAR DO
MODELO DE PREVISÃO**

AGOSTO/2024

Marcelo Eduardo Alves da Silva

Professor Associado do Departamento de Economia da UFPE e PIMES/UFPE

Ph.D. pela The University of North Carolina at Chapel Hill

1.	Apresentação	2
2.	Metodologia.....	3
2.1.	Problema de previsão de séries temporais.....	3
2.2.	Modelos ARMA	4
2.3.	Metodologia Box-Jenkins	5
2.4.	Uso dos modelos para previsão.....	7
3.	O modelo de previsão da Sefin/PCR	7
3.1.	Modelo para prever as receitas do ISSQN.....	8
3.1.1.	Análise dos dados.....	8
3.1.2.	Estimação dos modelos	12
3.1.3.	Ajuste do modelo dentro da amostra	15
3.1.4.	Ajuste do modelo fora da amostra	16
3.1.5.	Previsões para julho a dezembro de 2024.....	19
3.2.	Modelo para prever as receitas do IPTU.....	22
3.2.1.	Análise dos dados.....	22
3.2.2.	Estimação dos modelos	25
3.2.3.	Ajuste do modelo dentro da amostra	27
3.2.4.	Ajuste do modelo fora da amostra	29
3.2.5.	Previsões para julho a dezembro de 2024.....	31
4.	Conclusões	34

1. APRESENTAÇÃO

Com o objetivo de melhorar a classificação na CAPAG, passando de “C” para “B”, o município buscou o apoio do BID para melhorar a gestão fiscal em duas dimensões: (i) melhoria da gestão da receita pública; e (ii) melhoria da gestão dos gastos públicos. Este apoio possibilitou ao município aumentar seu espaço fiscal por meio da melhoria dos indicadores de liquidez de caixa e poupança corrente e assim tornou-o apto a realizar novas operações de empréstimo e novos investimentos na cidade.

As políticas envolvidas na operação foram:

Componente 1 - Estabilidade macroeconômica. Apoiar a melhoria dos indicadores fiscais para que o Município do Recife alcance o CAPAG “B”, conforme acordado com o Governo Federal, através das seguintes condições políticas:

- a) Melhorar os indicadores fiscais para contribuir no alcance da estabilidade de suas finanças públicas da forma prevista no PEF.

Componente 2 - Efetividade da arrecadação tributária. O objetivo do componente é fortalecer a gestão da cobrança, através das seguintes condições políticas:

- a) Melhorar a capacidade institucional de prever e monitorar as receitas municipais com o uso de dados fiscais eletrônicos e novas tecnologias digitais.
- b) Reforçar o controle fiscal, através da depuração contínua do registo dos contribuintes imobiliários e comerciais.
- c) Promover o cumprimento voluntário das obrigações fiscais através do acompanhamento de grandes Contribuintes do ISSQN.
- d) Fortalecer o atendimento ao cidadão por meio da digitalização, automação e maior transparência dos serviços.

Componente 3 - Eficiência dos gastos públicos. O objetivo do componente é fortalecer a gestão dos gastos públicos, através das seguintes condições políticas:

- a) Melhorar as regras de execução orçamentária e financeira do município.
- b) Melhorar o controle das despesas de pessoal da Prefeitura do Recife.
- c) Fortalecer as compras públicas para melhorar a qualidade do gasto público.
- d) Melhorar a qualidade dos gastos fortalecendo o sistema de investimento público.

Este documento apresenta o segundo produto de consultoria dentro do componente 2 - **Efetividade da arrecadação tributária**, em particular no suporte à implementação do Tranche IV – (a) Avaliação dos resultados do novo modelo de previsão e monitoramento dos indicadores da receita municipal, na qual se apresentem os benefícios gerados. Em particular, apresenta a avaliação preliminar do modelo de previsão utilizado pela equipe da Secretaria de Finanças da PCR para realizar projeções de receita tributárias relacionadas ao ISSQN e IPTU.

2. METODOLOGIA

O objetivo do trabalho consiste na avaliação dos resultados do novo modelo de previsão e monitoramento de indicadores de receitas municipais, onde são apresentados os benefícios gerados. Para alcançar esse objetivo estão sendo executadas três metodologias principais: i. Análise econométrica do modelo desenvolvido pela Secretaria de Finanças; ii. Análise da documentação e iii. Reuniões com a equipe responsável pela modelagem e uso do modelo de previsão.

O primeiro passo consistiu na análise dos dados (séries temporais) e documentação relacionada aos modelos (códigos computacionais, avaliações internas, relatórios etc.). A partir dos dados, documentação e conversas com os responsáveis pela modelagem e uso do modelo de previsão, foi realizada uma análise da adequabilidade dos modelos ao problema de previsão de receitas tributárias.

Nesta primeira etapa, a ideia é avaliar os modelos e os seus resultados em termos de ajuste aos dados e em termos de acurácia de previsão. Na etapa subsequente, será realizada uma análise comparativa do desempenho do modelo da SEFIN/PCR com modelos alternativos. Também será avaliado o uso dos resultados do modelo por parte da PCR.

2.1. Problema de previsão de séries temporais

O problema de previsão consiste em estimar os valores futuros de uma ou mais séries temporais com base numa amostra de dados. Para isso pode ser utilizado apenas os valores passados da série temporal ou ainda valores de outras variáveis consideradas importantes na previsão da série de interesse.

Seja Y_t uma série de tempo com valores conhecidos até o momento T . O problema de previsão consiste em prever o comportamento futuro da série temporal no horizonte h utilizando as informações até T . Formalmente:

$$E_t y_{t+h} = E_t(y_{t+h} | y_T, y_{T-1}, y_{T-2}, \dots, y_1, e_T, e_{T-1}, e_{T-2}, \dots, e_1)$$

Diferentes modelagens podem ser utilizadas para alcançar o objetivo de previsão. No entanto, este documento focará nos modelos Univariados, em particular, nos modelos ARMA ou ARIMA.

2.2. Modelos ARMA

Resumidamente, os modelos ARMA – Autoregressivos de Médias Móveis consiste numa representação de uma série temporal y_t na forma:

$$y_t = a_0 + \sum_{j=1}^p a_j y_{t-j} + \sum_{j=0}^q b_j e_{t-j}$$

Onde a_j para $j = 1, 2, \dots, p$ e b_j para $j = 0, 1, 2, \dots, q$, são constantes, e $b_0 = 1$. E e_t é um processo estocástico com média zero, variância constante e não é correlacionado com todas as outras realizações (ausência de autocorrelação serial). Esse processo é chamado de **ruído branco** e formalmente $\{e_t\}$ é tal que: $E[e_t] = 0$ para todo t , $\text{Var}[e_t] = \sigma^2$ e $E[e_t, e_{t-j}] = 0$ para todo $j \in \mathbb{N}$.

Se as raízes características de $\{y_t\}$ estiverem todas no círculo unitário, $\{y_t\}$ é chamado de modelo ARMA para y_t . A parte autorregressiva do modelo é a equação de diferença dada pela porção homogênea e a parte da média móvel é a sequência $\{e_t\}$. Se a parte homogênea da equação de diferença contiver p defasagens e a parte de médias móveis contiver q defasagens, o modelo é chamado de modelo ARMA(p, q).

Entretanto, se uma ou mais raízes características do modelo forem maiores ou iguais à unidade, a sequência $\{y_t\}$ é considerada um processo integrado e o modelo é chamado de modelo autorregressivo de média móvel integrado de ordem d ou ARIMA(p, d, q). Onde d consiste no número de diferenças para tornar a sequência $\{y_t\}$ estacionária.

O problema da estimação consiste em identificar as ordens p e q do modelo, quando se tratar de uma série temporal estacionária. Quando se tratar de uma variável não-estacionária (e essa identificação pode ser realizada através de testes estatísticos – testes de raiz unitária), aplica-se d diferenciações à série para

então proceder na identificação das ordens p e q do modelo. É possível ainda incluir variáveis exógenas na estimação. Neste caso o modelo é modificado da seguinte maneira:

$$y_t = a_0 + \sum_{j=1}^p a_j y_{t-j} + \sum_{j=0}^q b_j e_{t-j} + \sum_{j=1}^n \sum_{i=0}^m w_{ij} x_{i,t-j}$$

Onde w_{ij} para $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$ são constantes associadas com cada uma das variáveis exógenas x_i na defasagem j . Esses modelos são denominados ARIMAX e incluem o modelo ARIMA como caso particular.

Quando a série possuir um comportamento sazonal e esse comportamento for importante para o objetivo de modelagem, é possível estimar um modelo que capture tanto a sazonalidade quanto a possibilidade de não-estacionariedade sazonal. Esses modelos são chamados de SARIMA ou ainda SARIMAX. Neste último caso, admitindo a possibilidade de variáveis exógenas na estimação. A estimação de modelos da classe ARIMA é tipicamente realizada pelo método de maximização da verossimilhança.

2.3. Metodologia Box-Jenkins

O método de Box–Jenkins (1976) para seleção do modelo ARIMA mais apropriado consiste em três etapas: identificação, estimação e diagnóstico. Na etapa de **identificação**, o objetivo é identificar a ordem (p, d, q) do modelo. Dentro dessa etapa, procede-se uma inspeção visual do gráfico temporal da série, da função de autocorrelação (FAC) e da função de correlação parcial (FACP) amostrais. Esse passo é importante porque fornece informações úteis sobre tendências temporais (determinísticas ou estocásticas), outliers, valores ausentes e quebras estruturais nos dados. Variáveis não-estacionárias tendem a apresentar uma tendência pronunciada ou “serpentejar” sem qualquer tendência de reversão à média ou variância constante de longo prazo. Valores ausentes e outliers podem ser corrigidos neste ponto. Nesta etapa, é realizada uma comparação da FAC e FACP amostrais com aquelas de vários processos ARMA teóricos o que pode sugerir vários modelos plausíveis. Aqui surge um desafio, já que vários possíveis modelos podem ser sugeridos na comparação entre funções amostrais e teóricas. Neste caso, a recomendação típica é seguir com dois ou três modelos mais plausíveis. Exceto quando há uma clara identificação do modelo, o que tipicamente não ocorre.

No estágio de **estimação**, cada um dos modelos plausíveis é estimado, e os vários coeficientes a_j e b_j são examinados. Neste segundo estágio, o objetivo é selecionar um modelo estacionário e parcimonioso que tenha um bom ajuste. Ajuste e parcimônia são objetivos conflitantes de certa forma, ao aumentar o número de regressores, por um lado, melhora-se o ajuste do modelo aos dados. Por outro, corre-se o risco de *overfitting* ou sobreparametrização do modelo (muitos parâmetros a serem estimados), aumentando a incerteza das estimativas. Daí comumente se utiliza um processo de minimização de uma função de perda que ao mesmo atribua valor ao ajuste do modelo aos dados (minimizando os erros) e penalize a sobreparametrização. Dois desses critérios são o Critério de Informação de Akaike (AIC) ou ainda sua versão com correção de viés para pequenas amostras (AICc), e o Critério de Informação Bayesiano (BIC). Como há diferenças entre esses critérios, nem sempre eles concordam em escolher o mesmo modelo. Neste caso, a recomendação é seguir com os modelos sugeridos por esses critérios para o próximo estágio.

O estágio de **diagnóstico** envolve verificação dos resíduos para garantir que os resíduos do modelo estimado imitem um processo de ruído branco. Neste estágio são realizadas inspeções visuais da série temporal dos resíduos, da Kernel (distribuição empírica) e da FAC dos resíduos. São implementados ainda testes estatísticos para verificar se os resíduos do modelo exibem autocorrelação significativa (teste Ljung-Box) e ainda aqueles que avaliam se a distribuição empírica dos resíduos possui assimetria e curtose tais quais aqueles de uma distribuição normal (teste Jarque-Bera). É importante notar duas coisas. A primeira é que diferentes modelos podem produzir resíduos que se comportam como ruído branco. A segunda é que a não-normalidade dos resíduos não é um impeditivo para o uso dos modelos já que a presença de valores discrepantes pode viesar os resultados do teste. Além disso, os testes são baseados em grandes amostras (teoria assintótica), o que tipicamente não é o caso da maioria das séries temporais da vida real. O mais importante é ter a sinalização da ausência de autocorrelação dos resíduos. Quando este não for o caso, deve-se reiniciar o processo de identificação até que a presença de autocorrelação dos resíduos seja eliminada. Passados esses estágios, procede-se com o uso do modelo para realizar previsões fora-da-amostra para a série temporal.

2.4. Uso dos modelos para previsão

Uma vez selecionados os modelos mais plausíveis, é possível usá-los para previsão fora-da-amostra. A questão importante a ser respondida nessa etapa é: qual modelo? Uma possibilidade natural é usar o modelo que melhor se ajustou aos dados observados e que tenha passado nos testes de diagnóstico. Como discutido anteriormente, essa identificação não é tão simples assim, uma vez que diferentes modelos podem ser apontados como plausíveis e podem inclusive gerar resíduos que se comportam como ruído branco.

Além disso, como o objetivo final é o uso dos modelos para previsão, nada garante que o melhor modelo que se ajustou aos dados passados produzirá as melhores previsões sobre o comportamento futuro da série temporal. No entanto, essa avaliação é impossível de ser feita, uma vez que os dados futuros não se realizaram ainda. Uma prática usual é deixar parte da amostra de fora das etapas de identificação, estimação e diagnóstico, e implementar essas etapas com a outra parte da amostra. Suponha uma amostra de tamanho total T . Dividindo-se essa amostra em duas partes $T=N+h$, por exemplo, é possível estimar os melhores modelos sobre a subamostra N e usar a parte restante h para avaliar os modelos em sua capacidade preditiva. Isso permitiria conhecer o modelo de melhor desempenho fora-da-amostra. A prática é deixar de 1/3 a 1/4 da amostra como período de teste, no entanto, outros tamanhos de h podem ser usados, mas, no mínimo, deve-se usar o mesmo período para o qual deseja-se realizar as previsões.

A avaliação pode ser feita a partir do cálculo de estatísticas como RMSE - Erro Quadrático Médio, o MAE - Erro Absoluto Médio, o MPE - Erro Percentual Médio, o MAPE - Erro Percentual Absoluto Médio, dentre outros. Com essas estatísticas em mãos é possível selecionar o melhor modelo. Contudo, não necessariamente um único modelo será apontado por esses critérios, deixando para o econometrista a decisão sobre qual modelo escolher. É possível ainda combinar previsões de modelos diferentes e com isso há a possibilidade de obter previsões melhores. Além disso, pode-se ainda implementar testes estatísticos como o Diebold-Mariano para avaliar se há evidências estatísticas de que as previsões de um modelo são mais precisas do que um modelo alternativo.

3. O MODELO DE PREVISÃO DA SEFIN/PCR

Em linhas gerais, a equipe responsável pelas previsões segue a metodologia de Box-Jenkins na estimação de modelos ARIMA e no uso posterior do modelo escolhido na realização das previsões. Em particular, a

equipe faz uso de um algoritmo desenvolvido por Hyndman e Khandakar (2018) “para reduzir o tempo dedicado à previsão de todas as receitas do município”, denominado **auto.arima**, o qual implementa as etapas de identificação, estimação e diagnóstico de maneira automática. O procedimento adotado pela equipe da PCR é selecionar o modelo com melhor ajuste aos dados de acordo com o critério de informação de Akaike (AIC). Escolhido o modelo, ele é utilizado nas previsões.

A qualidade das previsões é realizada com base na comparação entre os valores previstos e os valores efetivamente realizados. De acordo com informações da equipe da PCR, o critério para se julgar a qualidade das previsões é que estas se situem em torno de +/- 5% do valor realizado. Os resultados das previsões são comparados com previsões *ad hoc* da secretária de finanças com base em seus conhecimentos e análises técnicas. Se os resultados do modelo escolhido forem considerados satisfatórios, este é mantido na realização de novas previsões. Caso contrário, são sugeridas alternativas. No que se refere às revisões dos modelos, elas são implementadas bimestralmente, quando os modelos são reestimados usando as informações adicionais (dados realizados das receitas de impostos).

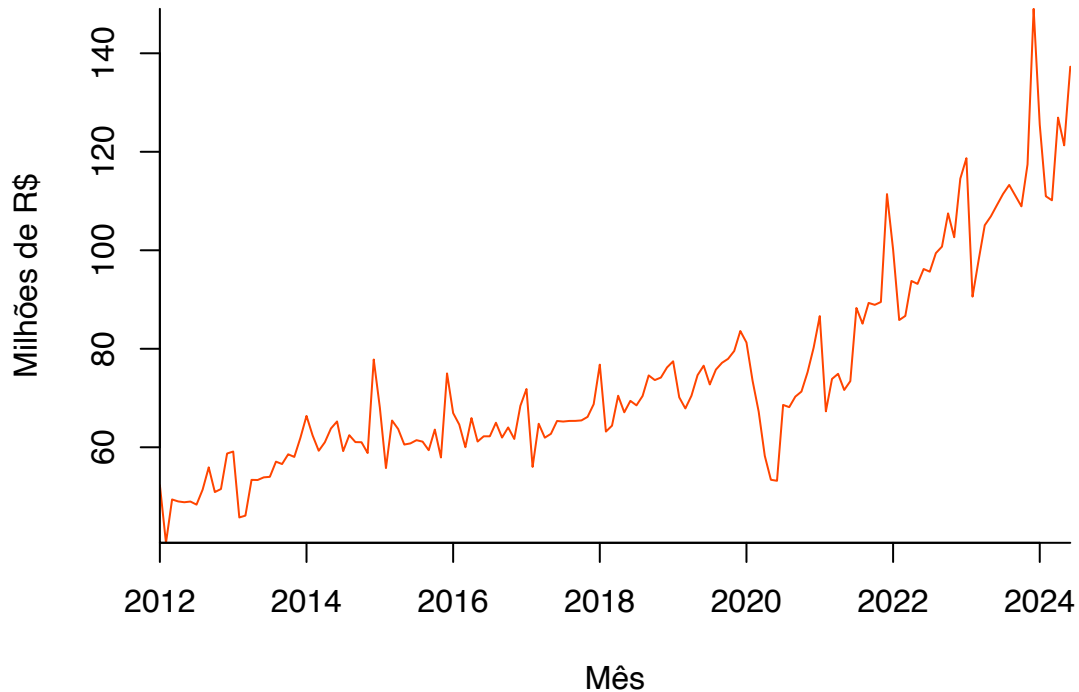
A seguir é implementado um exercício semelhante ao da equipe da PCR na realização das previsões das receitas de ISS e IPTU. Alguns ajustes na metodologia são apresentados ao longo do exercício e são deixados como sugestão para futuras implementações da metodologia.

3.1. Modelo para prever as receitas do ISSQN

3.1.1. Análise dos dados

A Figura 1 apresenta a receita mensal (em milhões de R\$ correntes) com o ISS de janeiro de 2012 a junho de 2024. Dado o objetivo de identificar padrões relevantes, o primeiro elemento a se notar é que a série apresenta um comportamento de crescimento ao longo do tempo (tendência) além de sugerir uma mudança de inclinação nos meses iniciais de 2020.

Figura 1
Receita mensal com o Imposto sobre Serviços
(Em milhões correntes)

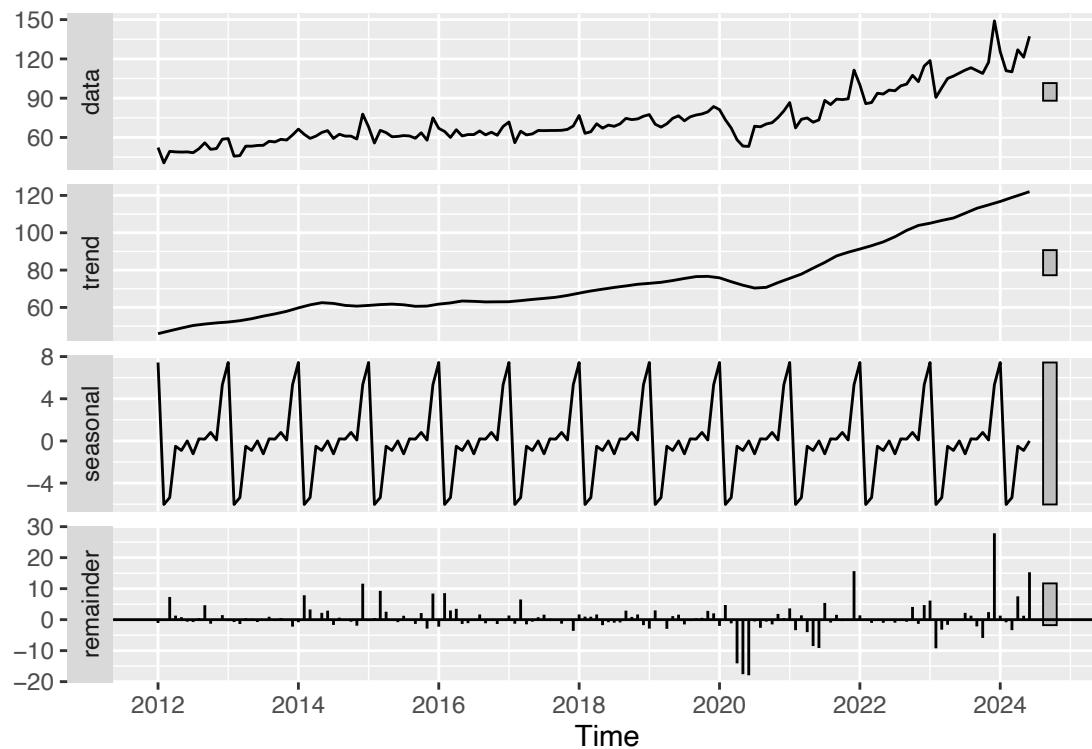


Fonte: SFIN/PCR

É possível ter uma indicação melhor a partir do uso de uma decomposição clássica da série em três componentes: tendência, sazonalidade e componente aleatório. A figura a seguir apresenta tal decomposição. Como sugerido anteriormente, a série temporal apresenta tanto uma tendência como uma aparente mudança de inclinação nos meses iniciais de 2020. Além disso, apresenta comportamento sazonal, o que na etapa de identificação pode sugerir modelos SARIMA.

Figura 2

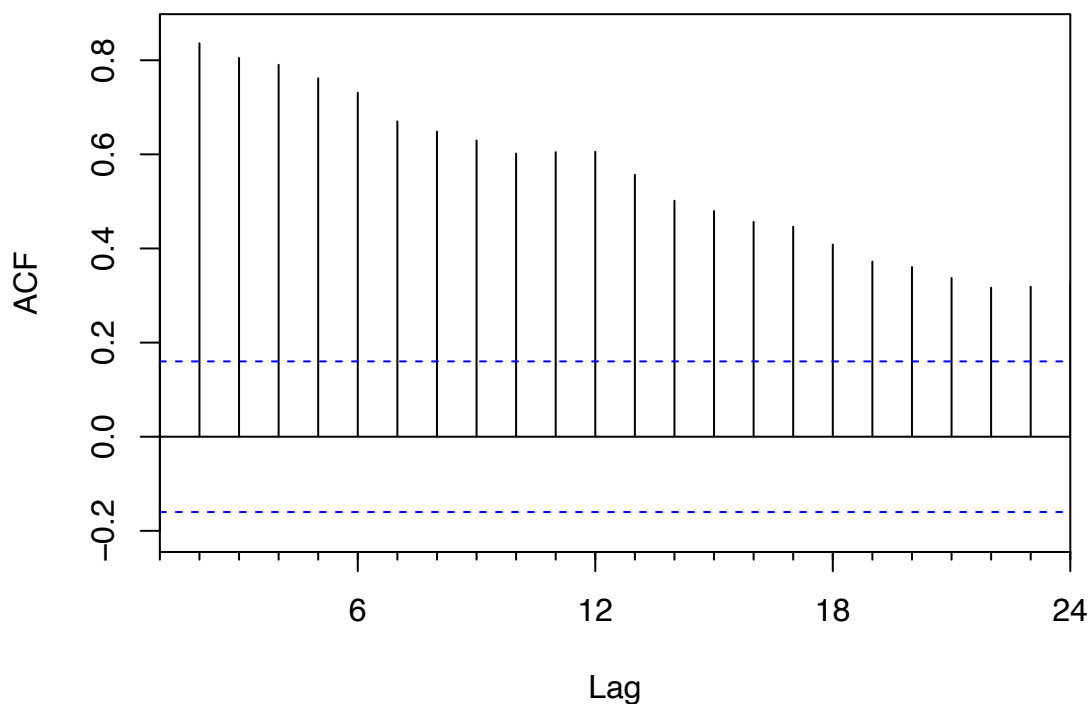
Decomposição clássica da série da receita mensal com o Imposto sobre Serviços



Fonte: BID

Uma inspeção da função de autocorrelação sugere um decaimento das autocorrelações indicando possivelmente a presença de elementos autorregressivos. Além disso, percebem-se certos picos a cada 12 defasagens aproximadamente, possivelmente revelando a presença de sazonalidade na série.

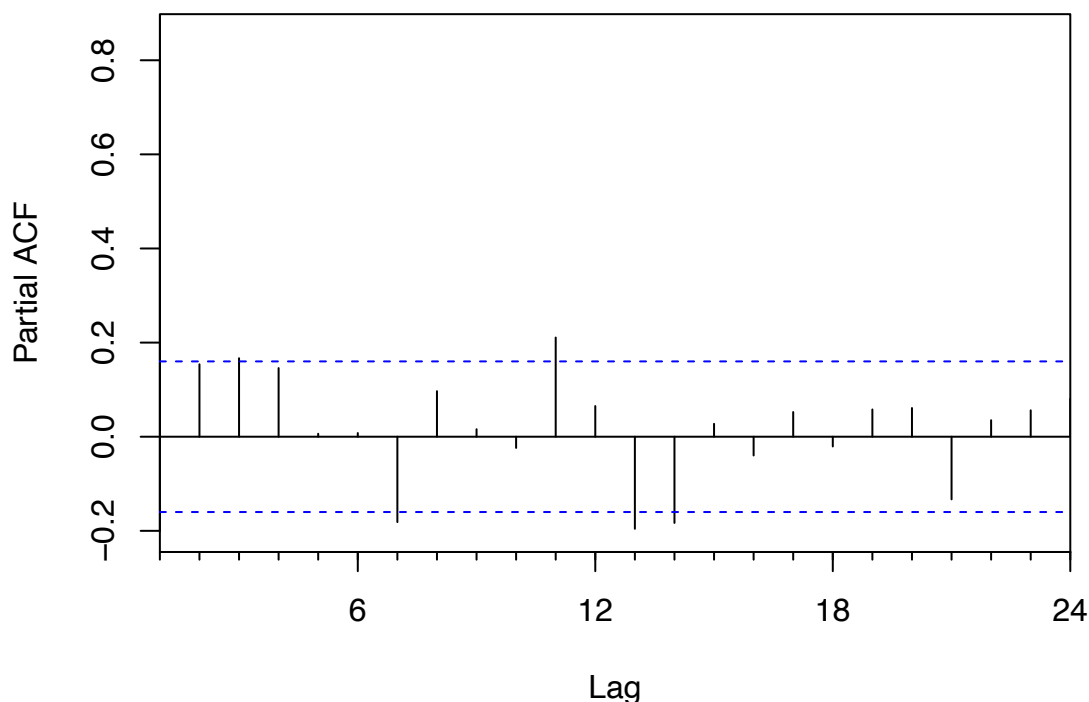
Figura 3
FAC da receita mensal com o Imposto sobre Serviços



Fonte: BID

A inspeção da FACP indica um modelo SARIMA(3,0,0)(1,0,0) pelo fato das três primeiras autocorrelações parciais aparentarem significância e pelos “saltos” na FAC a cada 12 defasagens. No entanto, diferentes possibilidades de modelos ARIMA são possíveis. A opção será utilizar critérios de informação como o AIC para definir o melhor modelo ou ainda o AICc ou o SBC.

Figura 4
FACP da receita mensal com o Imposto sobre Serviços _



Fonte: BID

3.1.2. Estimação dos modelos

Para fins de seleção do modelo mais apropriado, a amostra será dividida em duas partes: treinamento e teste. O período de treinamento será usado para estimar diferentes alternativas, avaliando o ajuste do modelo aos dados dentro deste período. Enquanto a fase de teste será usada para avaliar o desempenho do modelo em termos de sua capacidade preditiva. O período de treinamento foi definido como o período que vai de janeiro de 2012 a junho de 2023, deixando os doze meses seguintes como período de teste. Note que escolhas diferentes poderiam ter sido feitas, o que poderia indicar possivelmente diferentes modelos.

A escolha do “melhor” modelo no período de treinamento será realizado utilizando o AICc. A tabela a seguir apresenta os resultados dos critérios de informação para os três melhores modelos de acordo com o AICc. É possível notar que o “melhor” modelo indicado pelo menor AICc foi o ARIMA(3,1,0)(2,0,0) ou SARIMA(3,1,0)(2,0,0). A tabela apresenta ainda os dois outros modelos com melhor ajuste dentre as várias combinações testadas. Note que enquanto os critérios AICc e AIC concordam em escolher o mesmo modelo, o BIC indica um modelo mais parcimonioso. Esse conflito é comum de ocorrer em trabalhos

empíricos, cabendo ao econometrista a escolha do critério a ser usado e, por seguinte, do modelo a ser empregado no exercício de construção das previsões. Trabalharemos com os três para comparar a capacidade preditiva deles.

Tabela 1
Critérios de informação

Modelo	ARIMA	AICc	AIC	BIC
1	ARIMA(3,1,0)(2,0,0)	847.04*	846.39*	863.91
2	ARIMA(2,1,0)(2,0,0)	848.02	847.56	862.16*
3	ARIMA(3,1,0)(2,0,0) com constante	848.16	847.29	867.73

Nota: Um "*" indica o melhor modelo de acordo com o critério. Fonte: BID

O melhor modelo foi um ARIMA(3,1,0)(2,0,0) e a tabela a seguir apresenta os coeficientes estimados.

Tabela 2
Modelo 1: ARIMA (3,1,0)(2,0,0)

	ar1	ar2	ar3	sar1	sar2
coeficiente	-0.5256	-0.2973	-0.1525	0.4461	0.2568
Desvio padrão	0.0861	0.0927	0.0851	0.0853	0.0939

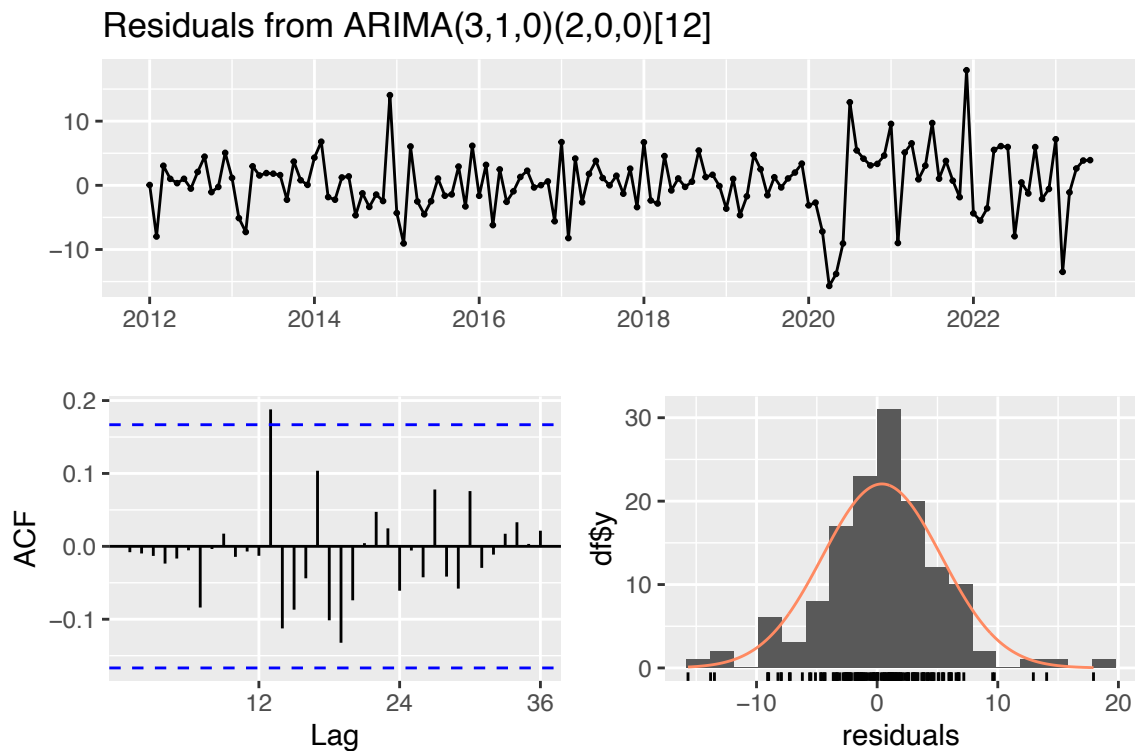
$\sigma^2 = 25.43$: log likelihood = -417.2

AIC=846.39 AICc=847.04 BIC=863.91

Fonte: BID

A Figura 5 a seguir apresenta o diagnóstico dos resíduos do modelo 1. Em particular apresenta a evolução dos resíduos ao longo do tempo, sua distribuição empírica (kernel) e a FAC dos resíduos. Aparentemente, os resíduos se assemelham a um processo de ruído branco.

Figura 5
Diagnóstico dos resíduos do modelo 1



Fonte: BID

Duas verificações adicionais consistem na implementação do teste de Ljung-Box de autocorrelação serial e do teste Jarque-Bera para normalidade. Os resultados do teste de Ljung-Box indicam ausência de autocorrelação serial dos resíduos do modelo, enquanto os resultados do teste de Jarque-Bera indicam que a assimetria e curtose dos resíduos não se assemelham aos de uma distribuição normal. Os quadros a seguir apresentam esses resultados.

Quadro 1 Teste Ljung-Box

data: Residuals from ARIMA(3,1,0)(2,0,0)[12]

$Q^* = 18.466$, $df = 19$, $p\text{-value} = 0.4916$

Model $df: 5$. Total lags used: 24

Fonte: BID

Quadro 2 Teste Jarque-Bera

data: residuals(model1)

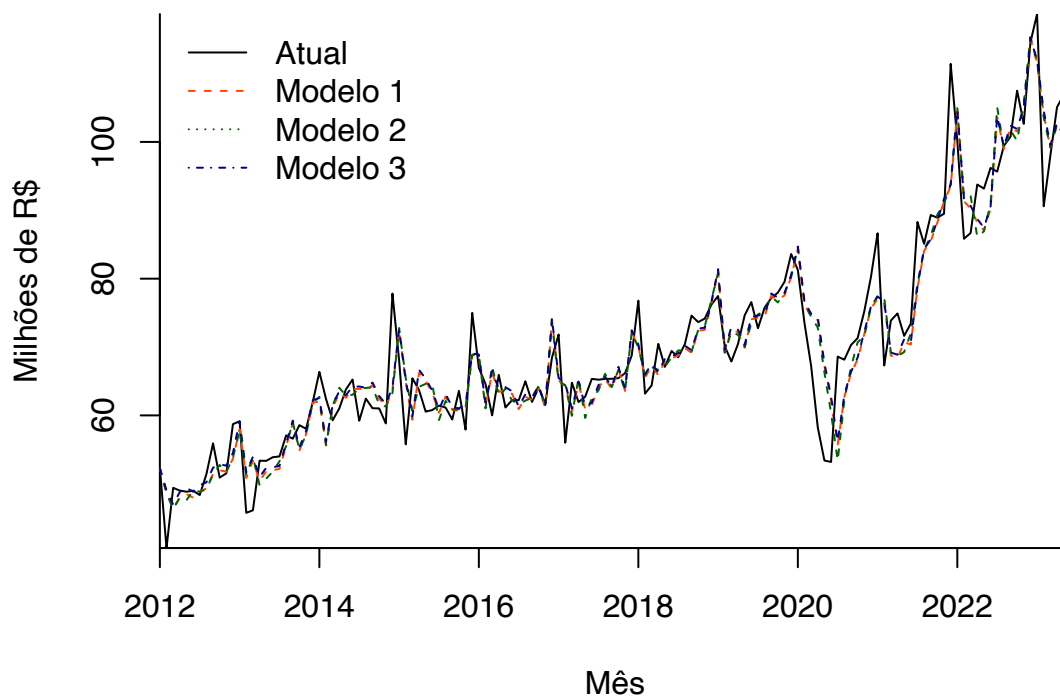
X-squared = 19.551, df = 2, p-value = 5.683e-05

Fonte: BID

3.1.3. Ajuste do modelo dentro da amostra

Uma forma de verificar o ajuste do modelo aos dados é comparar os valores estimados (usando os coeficientes estimados) e os dados utilizados na estimação. A Figura 6 a seguir apresenta a evolução da receita mensal com o ISS de janeiro de 2012 a julho de 2023 (período de treinamento) e os valores previstos pelos três modelos em questão (para dentro da amostra). Visualmente, os três modelos apresentam ajustes semelhantes aos dados observados. Todos eles capturam razoavelmente bem a dinâmica da arrecadação mensal com o ISS em Recife.

Figura 6
Receita mensal ISS: Valores atuais e previstos Jan/2012 a Jun/2023



Nota: A figura apresenta o ajuste dos modelos aos dados efetivamente realizados de janeiro de 2012 a junho de 2023. Fonte: BID

No entanto, a simples inspeção visual não nos ajuda muito em definir o modelo que melhor se ajustou aos dados durante o período de treinamento. Neste sentido, é possível obter algumas estatísticas que nos ajudam nessa decisão. A tabela a seguir apresenta algumas dessas.

Tabela 3
Critérios de desempenho preditivo dentro-da-amostra

Modelo	ARIMA	RMSE	MAE	MAPE	MASE
1	ARIMA(3,1,0)(2,0,0)	4.932	3.648	5.298	0.491
2	ARIMA(2,1,0)(2,0,0)	4.995	3.690	5.351	0.496
3	ARIMA(3,1,0)(2,0,0) com constante	4.917*	3.598*	5.244*	0.484*

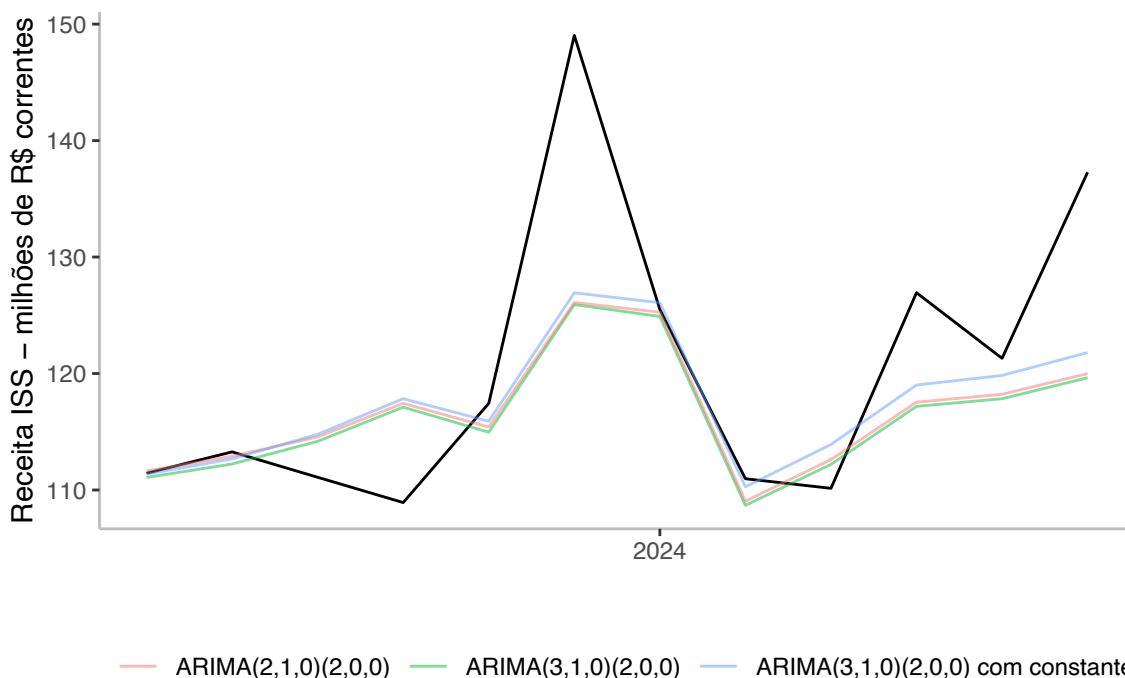
Nota: Um "*" indica o melhor modelo de acordo com o critério. Fonte: BID

O interessante a ser notado é que o modelo com melhor ajuste aos dados foi o modelo ARIMA(3,1,0)(2,0,0) com constante, modelo este que apresentou o pior AICc (critério utilizado para a escolha do modelo). Esse resultado é importante porque revela que os critérios de informação e as estatísticas tipicamente usadas para avaliar a capacidade preditiva dos modelos nem sempre concordarão na escolha do "melhor" modelo. A seguir os três modelos serão avaliados em suas capacidades preditivas fora-da-amostra.

3.1.4. Ajuste do modelo fora da amostra

Como indicado anteriormente, a amostra total foi dividida em duas partes: treinamento e teste. Na fase de treinamento, utilizando os critérios de informação, em particular o AICc, o melhor modelo, no conjunto parcimônia e ajuste, foi o ARIMA(3,1,0)(2,0,0) (Modelo 1). A partir da estimação dos três modelos e obtenção dos valores estimados dos parâmetros, eles podem ser aplicados aos dados do período de teste e a partir disso, os valores previstos da série temporal, \hat{y}_t , podem ser comparados aos valores efetivamente observados, y_t , e conseqüentemente tem-se a possibilidade de avaliar o desempenho dos modelos em sua capacidade preditiva. A figura a seguir apresenta as previsões para a receita mensal do ISS com base nos três modelos (linhas coloridas) e compara-as com os valores efetivamente realizados no período de teste dos modelos (julho de 2023 a junho de 2024).

Figura 7
Desempenho preditivo fora-da-amostra



Nota: A figura apresenta o ajuste dos modelos aos dados efetivamente realizados de julho de 2023 a junho de 2024. A linha sólida preta representa os dados realizados, enquanto as linhas coloridas representam os valores pontuais previstos por cada modelo indicado na legenda. Fonte: BID

A inspeção visual indica que os três modelos produziram previsões semelhantes e que, de certa maneira, acompanham a dinâmica dos dados. Fica claro ainda que nos dois primeiros meses as previsões são mais precisas. A partir dessas previsões e dos dados efetivos é possível calcular estatísticas que nos ajudam a sintetizar o desempenho preditivo dos modelos. A tabela abaixo apresenta essas estatísticas.

Tabela 4
Critérios de desempenho preditivo fora-da-amostra

Modelo	ARIMA	RMSE	MAE	MAPE	MASE
1	ARIMA(3,1,0)(2,0,0)	9.338	6.173	4.747	0.830
2	ARIMA(2,1,0)(2,0,0)	9.225	5.992	4.605	0.806
3	ARIMA(3,1,0)(2,0,0) com constante	8.679*	5.570*	4.291*	0.749*

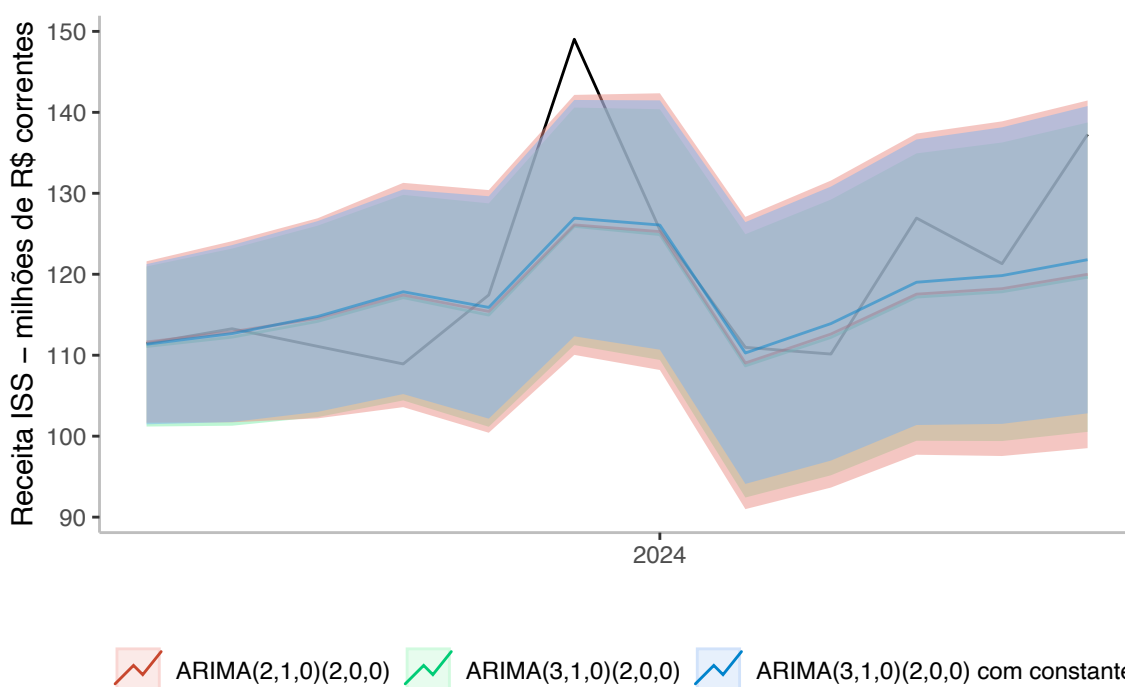
Nota: Um “*” indica o melhor modelo de acordo com o critério. Fonte: BID

Uma análise da tabela acima nos permite indicar que o modelo com o melhor desempenho no período de teste foi o modelo ARIMA(3,1,0)(2,0,0) com constante (modelo 3). Interessante notar que este não foi o modelo escolhido como o “melhor” modelo de acordo com o AICc ou ainda o AIC. Portanto, mais uma vez,

ressaltando a importância de julgar os modelos de acordo com sua capacidade preditiva e não necessariamente no seu ajuste aos dados passados.

Um ponto importante a ser considerado é que previsões são sujeitas a incertezas e, portanto, o quão boas foram (ou serão) as previsões devem ser julgadas com base no intervalo de confiança das previsões, não necessariamente na distância entre as estimativas pontuais e os valores efetivamente realizados. A figura a seguir apresenta as previsões pontuais dos três modelos e seus respectivos intervalos de confiança.

Figura 8
Desempenho preditivo fora-da-amostra (com intervalo de confiança)



Nota: A figura mostra os dados realizados e as previsões para a receita mensal com o ISS com base nos três modelos testados assim com os intervalos de confiança de 95% das previsões para o período de julho de 2023 e junho de 2024. Fonte: BID.

Exceto por um período apenas, todos os valores realizados *ex-post* das receitas do ISS ficaram dentro dos intervalos de confiança dos três modelos estimados. Portanto, é seguro afirmar que os modelos apresentaram bom desempenho preditivo e podem ser utilizados com o objetivo de prever as receitas de ISS para além da amostra total dos dados (de julho de 2024 em diante). Um ponto a ser explorado no relatório final será a implementação de testes estatísticos comparando as previsões de modelos alternativos.

3.1.5. Previsões para julho a dezembro de 2024

Uma vez escolhido o modelo, ele pode então ser reestimado sobre a amostra completa (jan de 2012 a junho de 2024). A tabela a seguir apresenta os resultados da estimação dos parâmetros do modelo ARIMA(3,1,0)(2,0,0) com constante.

Tabela 5
Modelo escolhido: ARIMA(3,1,0)(2,0,0) com constante

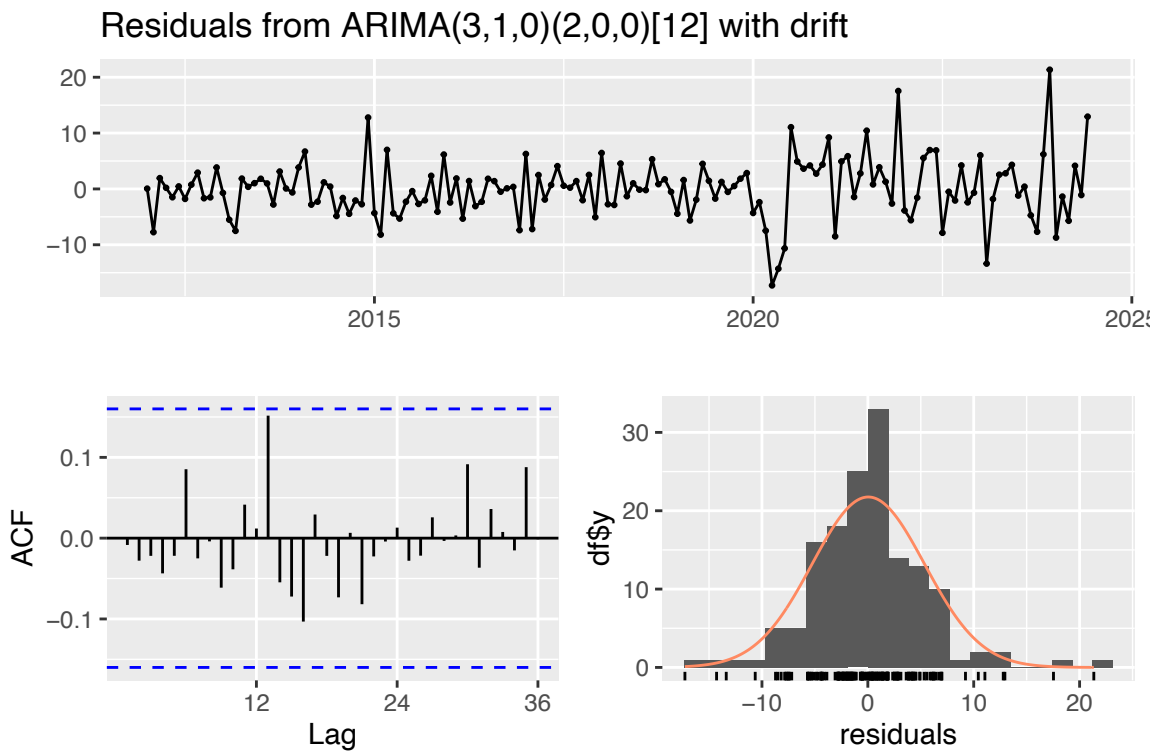
	ar1	ar2	ar3	sar1	sar2	constante
Coefficientes	-0.5536	-0.3324	-0.2488	0.3908	0.3761	0.7702
Desvio padrão	0.0831	0.0884	0.0819	0.0836	0.0873	0.6577

$\sigma^2 = 29.4$: log likelihood = -465.31

AIC=944.62 AICc=945.42 BIC=965.65

Após a estimação os resíduos do modelo são avaliados e há a indicação de que imitam um ruído branco, ao menos não apresentam autocorrelação como indicam a FACP e os resultados do teste de Ljung-Box abaixo. Mais uma vez, pode-se rejeitar a hipótese de que a curtose e a assimetria da distribuição empírica dos resíduos se assemelham ao de uma distribuição normal.

Figura 9
Diagnóstico dos resíduos



Fonte: BID

Quadro 3
Teste Ljung-Box

data: Residuals from ARIMA(3,1,0)(2,0,0)[12] with drift
 $Q^* = 12.49$, $df = 19$, $p\text{-value} = 0.8636$

Model df: 5. Total lags used: 24

Fonte: BID

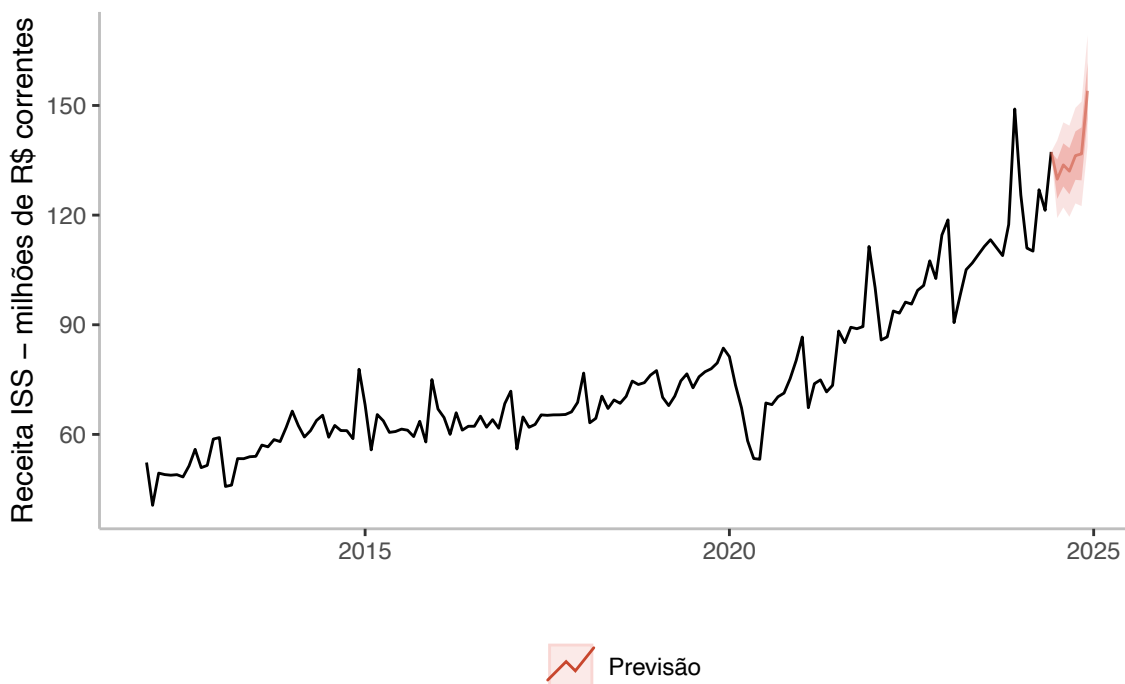
Quadro 4
Teste Jarque-Bera

data: residuals(model3)
 $X\text{-squared} = 39.624$, $df = 2$, $p\text{-value} = 2.487e-09$

Fonte: BID

A figura a seguir apresenta as previsões para a receita do ISS para o período de julho a dezembro de 2024 com base no modelo selecionado. Esses valores são resumidos na tabela abaixo, que apresenta ainda os intervalos de confiança de 68% e 95% das previsões.

Figura 10
Valores previstos para julho a dezembro de 2024



Nota: A figura mostra os dados realizados e as previsões para a receita mensal com o ISS com base no modelo selecionado na etapa anterior assim com os intervalos de confiança de 68% e 95% das previsões para o período de julho de 2024 a dezembro de 2024. Fonte: BID.

Tabela 6
Valores previstos para julho a dezembro de 2024

Data	ISS	LI_68	LS_68	LI_95	LS_95
jul-24	129.83	124.44	135.23	119.20	140.46
ago-24	133.73	127.83	139.64	122.09	145.37
set-24	132.02	125.69	138.35	119.55	144.49
out-24	136.28	129.65	142.91	123.21	149.35
nov-24	136.76	129.50	144.01	122.46	151.05
dez-24	154.01	146.33	161.70	138.86	169.16

Nota: A tabela apresenta as previsões para a receita mensal com o ISS com base no modelo selecionado na etapa anterior assim com os intervalos de confiança de 68% e 95% das previsões para o período de julho de 2024 a dezembro de 2024. Fonte: BID.

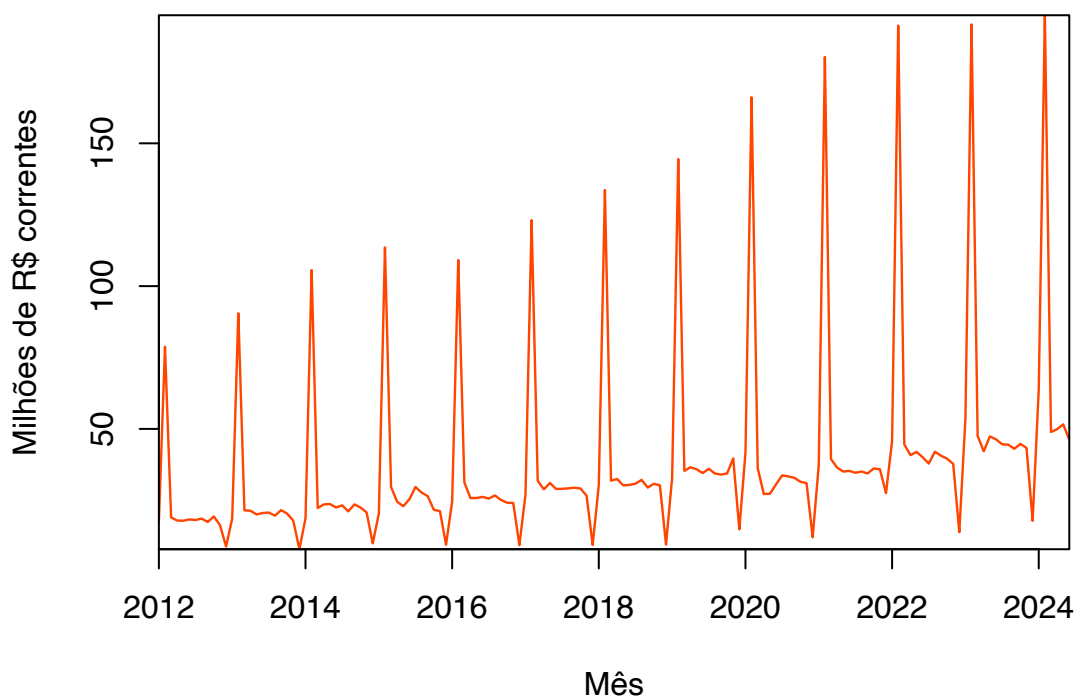
A previsão pontual é de que a receita com o ISS em dezembro de 2024 cresça em aproximadamente 3%, em termos nominais, em relação a dezembro de 2023.

3.2. Modelo para prever as receitas do IPTU

3.2.1. Análise dos dados

A Figura 11 apresenta a receita mensal (em milhões de R\$ correntes) com o IPTU de janeiro de 2012 a junho de 2024. Dado o objetivo de identificar padrões relevantes, o primeiro elemento a se notar é que a série apresenta um comportamento de crescimento ao longo do tempo (tendência) e um forte componente sazonal com picos no mês de fevereiro de cada ano.

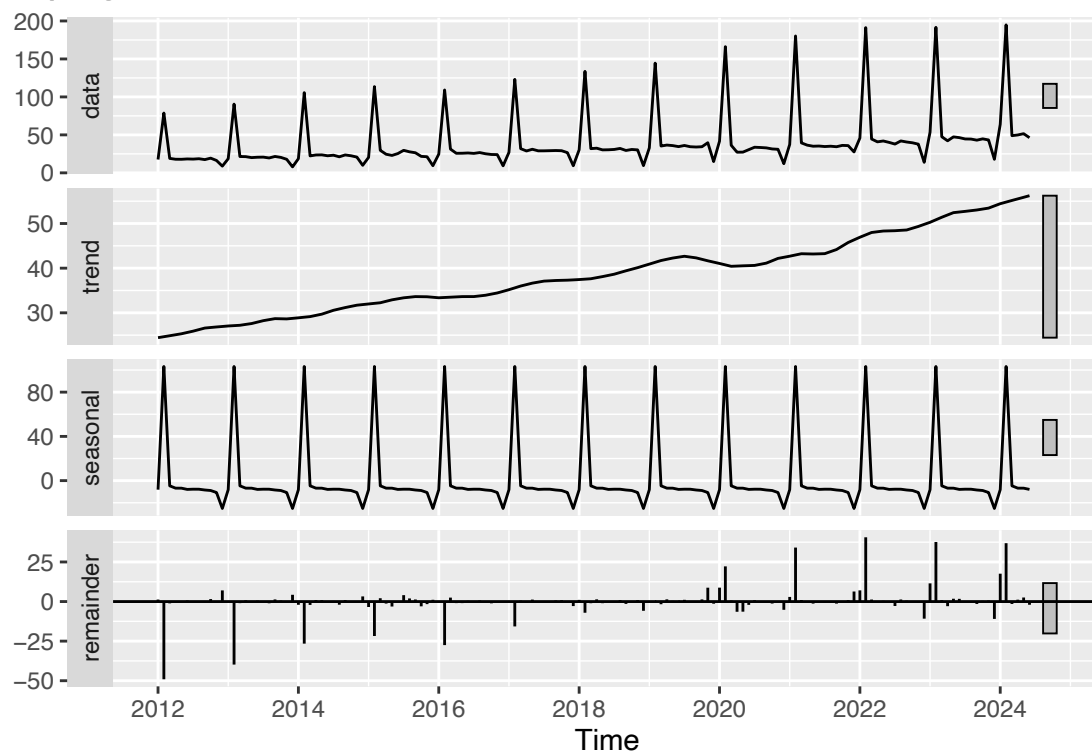
Figura 11
Receita mensal com o IPTU
(Em milhões correntes)



Fonte: SFIN/PCR

Assim como no caso do ISS, é possível aplicar uma decomposição clássica da série em três componentes: tendência, sazonalidade e componente aleatório. Como sugerido anteriormente, a série temporal apresenta tanto uma tendência como um comportamento sazonal, o que na etapa de identificação pode sugerir modelos SARIMA.

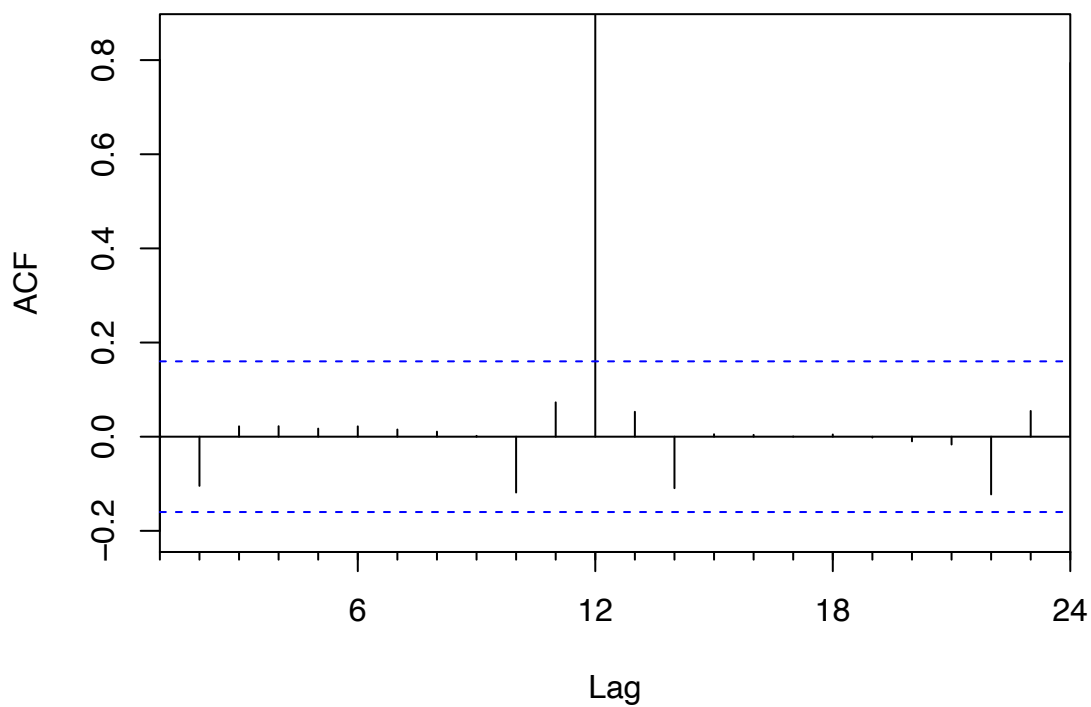
Figura 12
Decomposição clássica da série da receita mensal com o IPTU



Fonte: BID

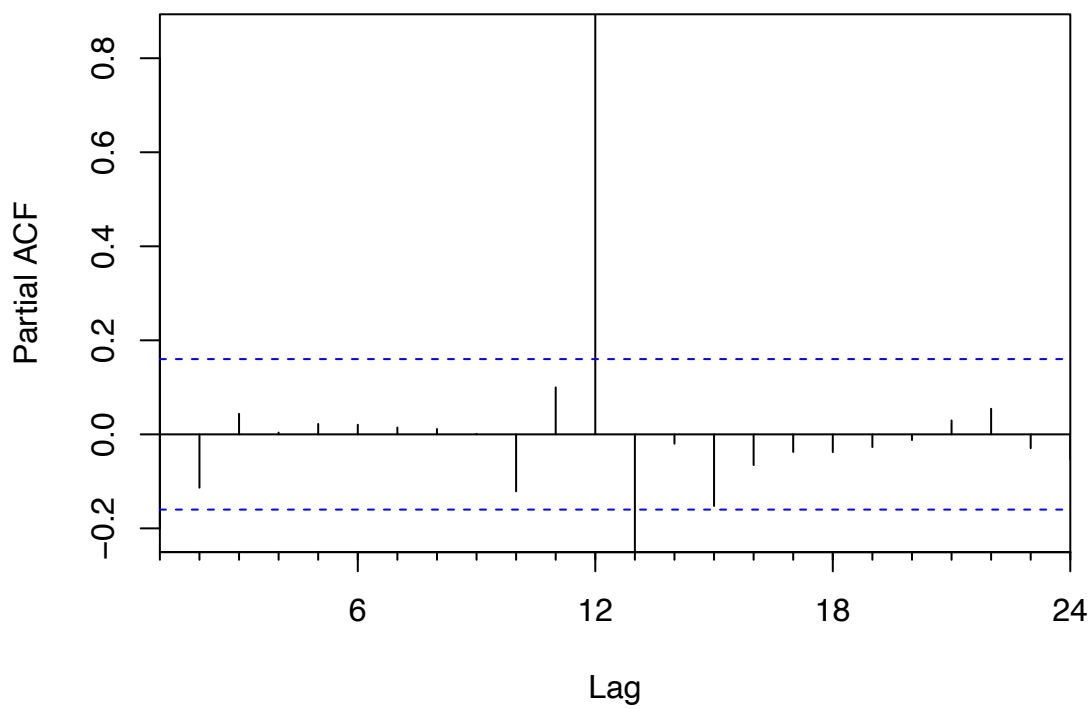
Uma inspeção das funções de autocorrelação e autocorrelação parcial sugere a presença de sazonalidade na série, o que havia sido indicado anteriormente.

Figura 13
FAC da receita mensal com o IPTU



Fonte: BID

Figura 14
FACP da receita mensal com o IPTU



Fonte: BID

A opção será utilizar critérios de informação como o AIC para definir o melhor modelo ou ainda o AICc ou o SBC.

3.2.2. Estimação dos modelos

Para fins de seleção do modelo mais apropriado, assim como no caso do ISS, a amostra será dividida em duas partes: treinamento e teste. O período de treinamento será usado para estimar diferentes alternativas, avaliando o ajuste do modelo aos dados dentro deste período. Enquanto a fase de teste será usada para avaliar o desempenho do modelo em termos de sua capacidade preditiva. O período de treinamento foi definido como o período que vai de janeiro de 2012 a junho de 2023, deixando os doze meses seguintes como período de teste.

A escolha do “melhor” modelo no período de treinamento será realizado utilizando o AICc. A tabela a seguir apresenta os resultados dos critérios de informação para os três melhores modelos de acordo com o AICc. É possível notar que o “melhor” modelo indicado pelo menor AICc foi o ARIMA(1,0,0)(0,1,0) com constante ou SARIMA(1,0,0)(0,1,0) com constante. A tabela apresenta ainda os dois outros modelos com melhor ajuste dentre as várias combinações testadas. Note que os três critérios concordam em escolher o mesmo modelo. Trabalharemos com os três para comparar a capacidade preditiva deles.

Tabela 7
Critérios de informação

Modelo	ARIMA	AICc	AIC	BIC
1	ARIMA(1,0,0)(0,1,0) com constante	746.30*	746.11*	754.62*
2	ARIMA(2,0,0)(0,1,0) com constante	747.36	747.03	758.37
3	ARIMA(1,0,0)(1,1,0) com constante	747.39	747.06	758.40

Nota: Um “*” indica o melhor modelo de acordo com o critério. Fonte: BID

A tabela a seguir apresenta os coeficientes estimados para o modelo 1.

Tabela 8
Modelo 1: ARIMA(1,0,0)(0,1,0) com constante

	ar1	constante
coeficiente	0.253	0.2466
Desvio padrão	0.086	0.0452

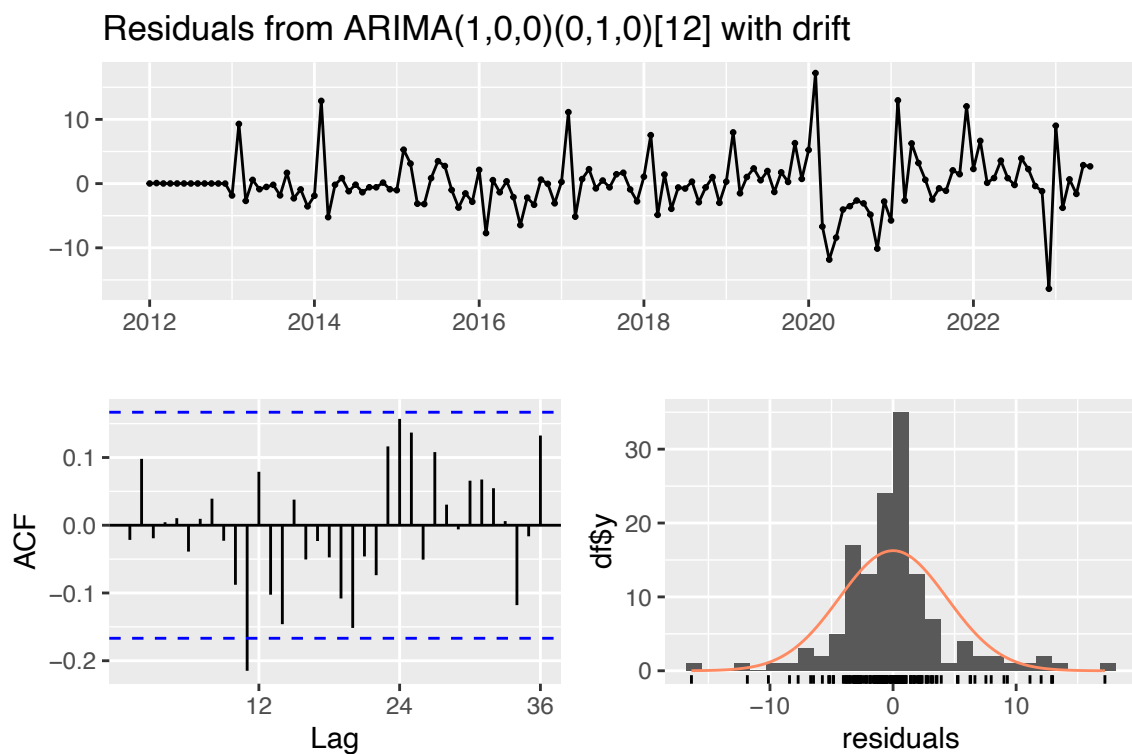
$\sigma^2 = 21.15$: log likelihood = -370.05

AIC=746.11 AICc=746.3 BIC=754.62

Fonte: BID

A figura a seguir apresenta o diagnóstico dos resíduos do modelo 1. Em particular apresenta a evolução dos resíduos ao longo do tempo, sua distribuição empírica (kernel) e a FAC dos resíduos. Exceto pela 11^o defasagem, não há evidência de autocorrelação serial. Fora isso, aparentemente os resíduos se assemelham a um processo de ruído branco. Note que uma forma de eliminar a aparente autocorrelação serial seria aumentar o número de componentes autorregressivos, mas isso ocorreria ao custo de menor parcimônia do modelo.

Figura 15
Diagnóstico dos resíduos do modelo 1



Fonte: BID

Duas verificações adicionais consistem na implementação do teste de Ljung-Box de autocorrelação serial e do teste Jarque-Bera para normalidade. Os resultados do teste de Ljung-Box indicam ausência de autocorrelação serial dos resíduos do modelo, enquanto os resultados do teste de Jarque-Bera indicam que a assimetria e curtose dos resíduos não se assemelham aos de uma distribuição normal.

Quadro 5 Teste Ljung-Box

data: Residuals from ARIMA(1,0,0)(0,1,0)[12] with drift Q* = 30.577, df = 23, p-value = 0.1335 Model df: 1. Total lags used: 24

Fonte: BID

Quadro 6 Teste Jarque-Bera

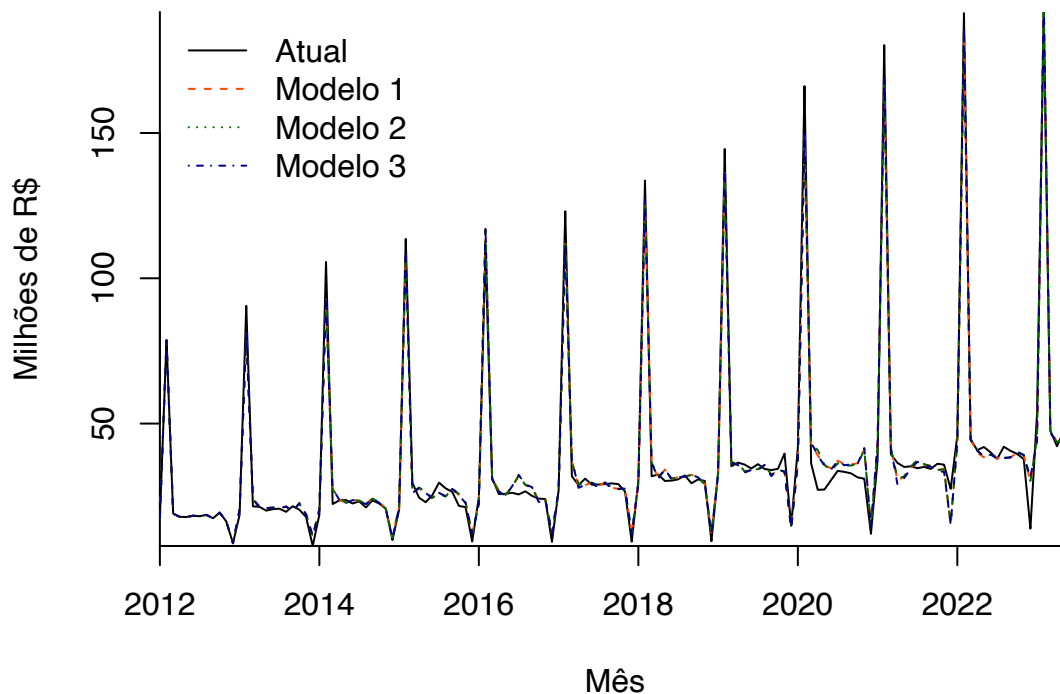
data: residuals(model1) X-squared = 81.677, df = 2, p-value < 2.2e-16
--

Fonte: BID

3.1.3. Ajuste do modelo dentro da amostra

A figura a seguir apresenta a evolução da receita mensal com o IPTU de janeiro de 2012 a julho de 2023 (período de treinamento) e os valores previstos pelos três modelos em questão. Visualmente, os três modelos apresentam ajustes semelhantes aos dados observados. Todos eles capturam razoavelmente bem a dinâmica da arrecadação mensal com o IPTU em Recife.

Figura 16
Receita mensal IPTU: Valores atuais e previstos Jan/2012 a Jun/2023



Nota: A figura apresenta o ajuste dos modelos aos dados efetivamente realizados de janeiro de 2012 a junho de 2023. Fonte: BID

No entanto, a simples inspeção visual não nos ajuda muito em definir o modelo que melhor se ajustou aos dados durante o período de treinamento. Neste sentido, é possível obter algumas estatísticas que nos ajudam nessa decisão. A tabela a seguir apresenta algumas dessas.

Tabela 9
Critérios de desempenho preditivo dentro-da-amostra

Modelo	ARIMA	RMSE	MAE	MAPE	MASE
1	ARIMA(1,0,0)(0,1,0) com constante	4.359	2.807	8.898	0.681
2	ARIMA(2,0,0)(0,1,0) com constante	4.340	2.738*	8.724*	0.665*
3	ARIMA(1,0,0)(1,1,0) com constante	4.338*	2.794	8.898	0.678

Nota: Um "*" indica o melhor modelo de acordo com o critério. Fonte: BID

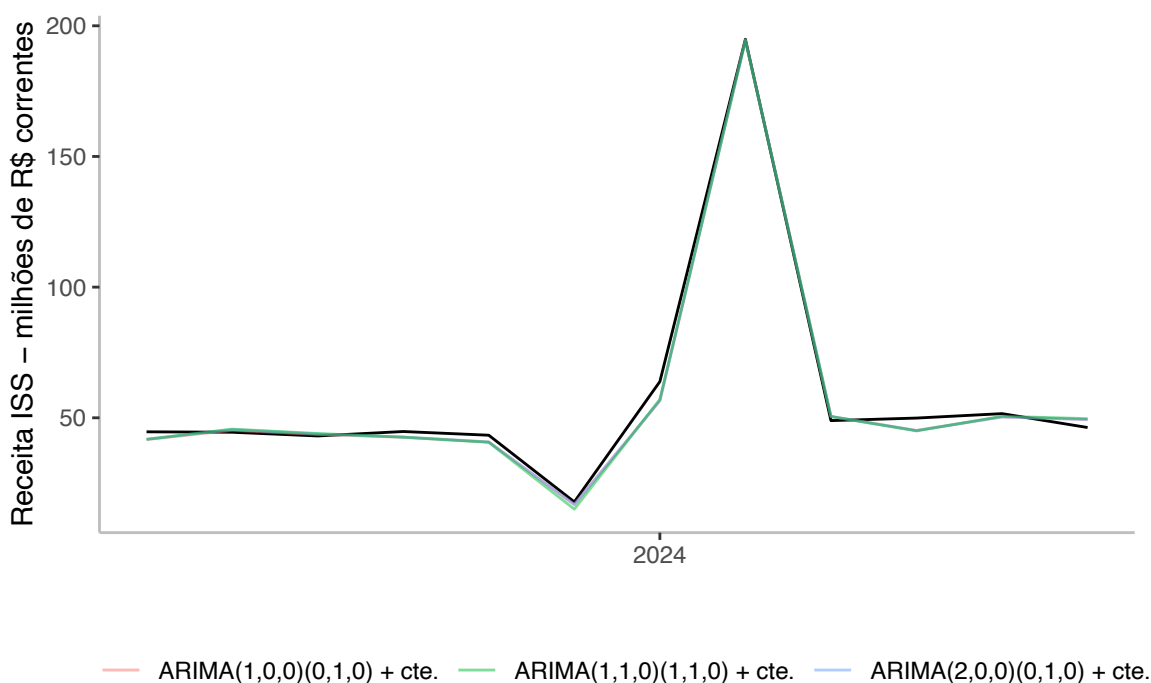
O interessante a ser notado é que o modelo com melhor ajuste aos dados foi o modelo ARIMA(2,0,0)(0,1,0) com constante, diferente do que havia sido escolhido pelos critérios de informação. Ou seja, assim como no caso do ISS, na escolha do modelo para o IPTU houve divergência entre os critérios e as estatísticas

tipicamente usadas para avaliar a capacidade preditiva dos modelos. A seguir os três modelos serão avaliados em suas capacidades preditivas fora-da-amostra.

3.1.4. Ajuste do modelo fora da amostra

Como indicado anteriormente, a amostra total foi dividida em duas partes: treinamento e teste. Na fase de treinamento, utilizando os critérios de informação, em particular o AICc, o melhor modelo, no conjunto parcimônia e ajuste, foi o ARIMA(1,0,0)(0,1,0) com constante (Modelo 1). A partir da estimação dos três modelos e obtenção dos valores estimados dos parâmetros, eles podem ser aplicados aos dados do período de teste e a partir disso, os valores previstos da série temporal, \hat{y}_t , podem ser comparados aos valores efetivamente observados, y_t , e conseqüentemente tem-se a possibilidade de avaliar o desempenho dos modelos em sua capacidade preditiva. A Figura 17 apresenta as previsões para a receita mensal do IPTU com base nos três modelos (linhas coloridas) e compara-as com os valores efetivamente realizados no período de teste dos modelos (julho de 2023 a junho de 2024).

Figura 17
Desempenho preditivo fora-da-amostra



Nota: A figura apresenta o ajuste dos modelos aos dados efetivamente realizados de julho de 2023 a junho de 2024. A linha sólida preta representa os dados realizados, enquanto as linhas coloridas representam os valores pontuais previstos por cada modelo indicado na legenda. Fonte: BID

A inspeção visual indica que os três modelos produziram previsões muito boas e que acompanham bem a dinâmica dos dados. A partir dessas previsões e dos dados efetivos é possível calcular estatísticas que nos ajudam a sintetizar o desempenho preditivo dos modelos. A tabela abaixo apresenta essas estatísticas.

Tabela 10
Critérios de desempenho preditivo fora-da-amostra

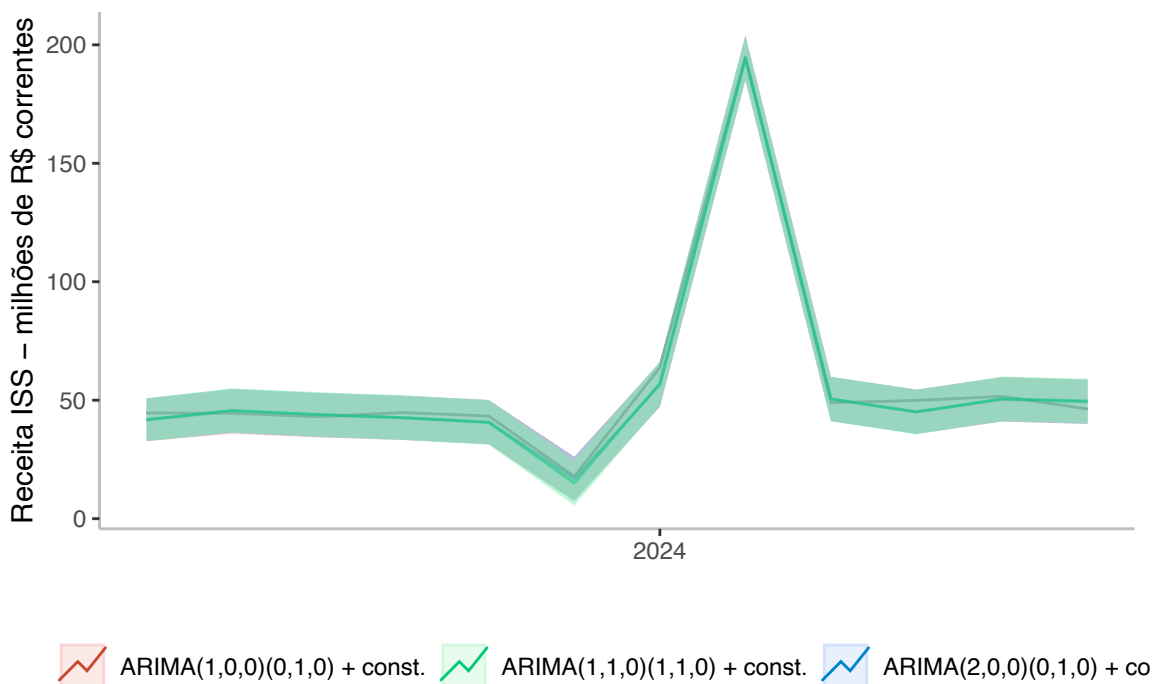
Modelo	ARIMA	RMSE	MAE	MAPE	MASE
1	ARIMA(1,0,0)(0,1,0) com constante	3.056	2.362*	4.977*	0.573*
2	ARIMA(2,0,0)(0,1,0) com constante	3.039*	2.372	4.997	0.576
3	ARIMA(1,0,0)(1,1,0) com constante	3.115	2.576	5.940	0.625

Nota: Um “*” indica o melhor modelo de acordo com o critério. Fonte: BID

Uma análise da tabela acima nos permite indicar que o modelo com o melhor desempenho no período de teste foi o modelo ARIMA(1,0,0)(0,1,0) com constante (modelo 1). Interessante notar que este foi o modelo escolhido como o “melhor” modelo de acordo com critérios de informação. Neste caso, portanto, eles concordaram na escolha do modelo.

Um ponto importante a ser considerado é que previsões são sujeitas a incertezas e, portanto, o quão boas foram (ou serão) as previsões devem ser julgadas com base no intervalo de confiança das previsões, não necessariamente na distância entre as estimativas pontuais e os valores efetivamente realizados. A figura a seguir apresenta as previsões pontuais dos três modelos e seus respectivos intervalos de confiança.

Figura 18
Desempenho preditivo fora-da-amostra (com intervalo de confiança)



Nota: A figura mostra os dados realizados e as previsões para a receita mensal com o ISS com base nos três modelos testados assim com os intervalos de confiança de 95% das previsões para o período de julho de 2023 e junho de 2024. Fonte: BID.

Todos os valores realizados *ex-post* das receitas do IPTU ficaram dentro dos intervalos de confiança dos três modelos estimados. Portanto, é seguro afirmar que os modelos apresentaram bom desempenho preditivo e podem ser utilizados com o objetivo de prever as receitas de IPTU para além da amostra total dos dados (de julho de 2024 em diante).

3.2.5. Previsões para julho a dezembro de 2024

Uma vez escolhido o modelo, ele pode então ser reestimado sobre a amostra completa (jan de 2012 a junho de 2024). A tabela a seguir apresenta os resultados da estimação dos parâmetros do modelo ARIMA(1,0,0)(0,1,0) com constante.

Tabela 11
Modelo 1: ARIMA(1,0,0)(0,1,0) com constante

	ar1	constante
coeficiente	0.2497	0.2559
Desvio padrão	0.0822	0.0420

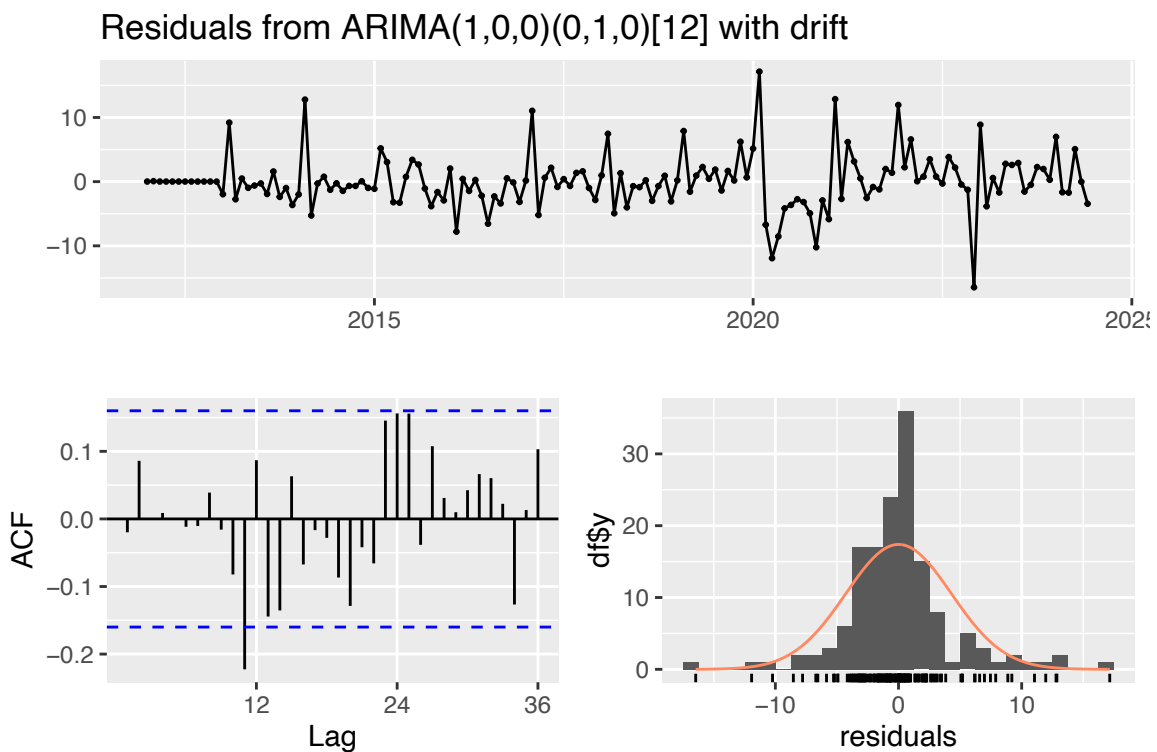
$\sigma^2 = 20.11$: log likelihood = -401.93

AIC=809.86 AICc=810.04 BIC=818.64

Fonte: BID

Após a estimação, os resíduos do modelo são avaliados. A indicação de que imitam um ruído branco é mais fraca no caso do IPTU. Há evidência de autocorrelação possivelmente na 11^o defasagem como indica a FACP, mas os resultados do teste de Ljung-Box rejeitam a 5% a hipótese nula de autocorrelação serial. Idealmente, prosseguiríamos com uma nova etapa de identificação, estimação e diagnóstico. Como indicado anteriormente, incluir mais elementos autoregressivos resolveria o problema, mas ao custo da parcimônia do modelo. Mais uma vez, pode-se rejeitar a hipótese de que a curtose e a assimetria da distribuição empírica dos resíduos se assemelham ao de uma distribuição normal.

Figura 19
Análise dos resíduos do modelo 1



Fonte: BID

Quadro 7 Teste Ljung-Box

data: Residuals from ARIMA(1,0,0)(0,1,0)[12] with drift

$Q^* = 33.677$, $df = 23$, $p\text{-value} = 0.06997$

Model $df: 1$. Total lags used: 24

Fonte: BID

Quadro 8 Teste Jarque-Bera

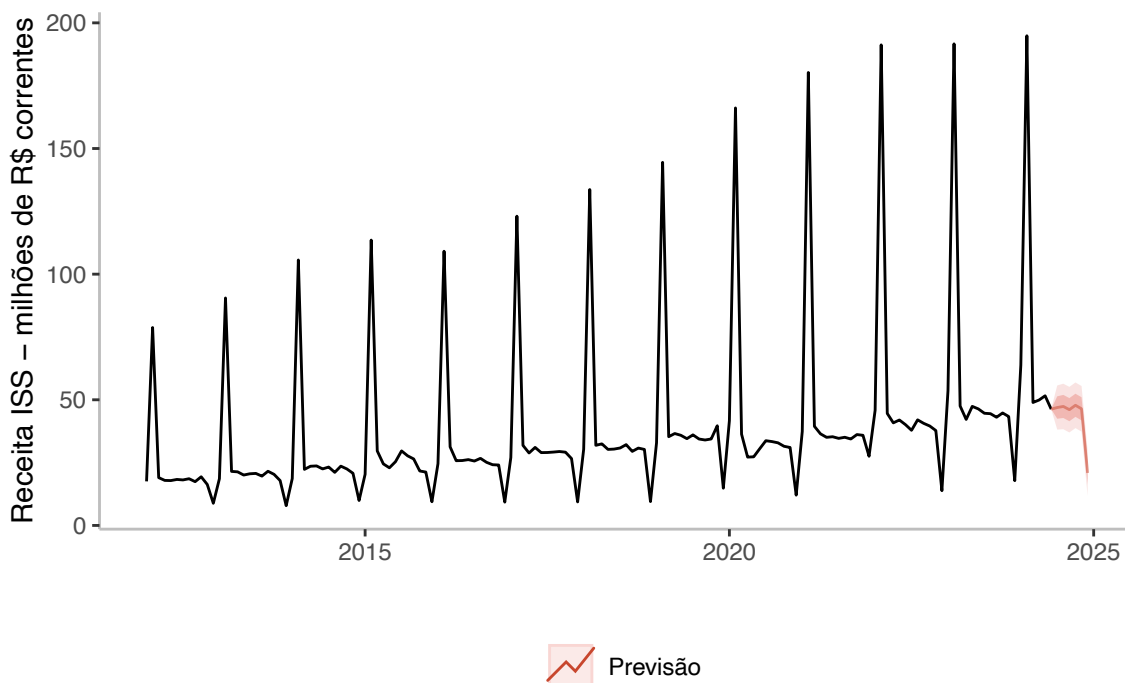
data: residuals(model)

$X\text{-squared} = 89.021$, $df = 2$, $p\text{-value} < 2.2e-16$

Fonte: BID

A figura a seguir apresenta as previsões para a receita do IPTU para o período de julho a dezembro de 2024 com base no modelo selecionado. Esses valores são resumidos na tabela abaixo, que apresenta ainda os intervalos de confiança de 68% e 95% das previsões.

Figura 20
Valores previstos IPTU para julho a dezembro de 2024



Nota: A figura mostra os dados realizados e as previsões para a receita mensal com o IPTU com base no modelo selecionado na etapa anterior assim com os intervalos de confiança de 68% e 95% das previsões para o período de julho de 2024 a dezembro de 2024. Fonte: BID.

Tabela 12
Valores previstos IPTU para julho a dezembro de 2024

Data	IPTU	LI_68	LS_68	LI_95	LS_95
jul-24	46.92	42.46	51.38	38.13	55.71
ago-24	47.35	42.75	51.95	38.29	56.41
set-24	46.08	41.47	50.68	37.00	55.15
out-24	47.82	43.21	52.42	38.74	56.89
nov-24	46.41	41.80	51.01	37.33	55.48
dez-24	20.89	16.29	25.50	11.82	29.97

Nota: A tabela apresenta as previsões para a receita mensal com o IPTU com base no modelo selecionado na etapa anterior assim com os intervalos de confiança de 68% e 95% das previsões para o período de julho de 2024 a dezembro de 2024. Fonte: BID.

4. CONCLUSÕES

Este produto avaliou de forma preliminar a metodologia utilizada para a previsão das receitas municipais com o ISS e IPTU. Em particular, o estudo aplicou a mesma metodologia e mostrou a utilidade e aplicabilidade dos modelos (metodologia) ao problema em questão. Na versão final, o trabalho será investigar se é possível melhorar a metodologia e com isso obter estimativas mais precisas.

Duas recomendações iniciais são: i. dividir a amostra em período de treinamento e teste, utilizando o período de teste para selecionar o modelo com base em sua capacidade preditiva e não exclusivamente com base em critérios de informação; ii. Seguir com o uso de dois ou três melhores modelos para a etapa de previsão na escolha do melhor modelo. Isso envolverá reavaliar o uso das previsões automáticas utilizando o algoritmo auto.arima, atualmente utilizado pela equipe da PCR.