

Modelação e Previsão do Consumo Mensal de Eletricidade em Espanha

Hugo Veríssimo. hugoverissimo@ua.pt, 124348.

Resumo—Este estudo analisa e prevê o consumo mensal de eletricidade em Espanha com base em dados do Eurostat (2008–2025). Foram comparados vários modelos estatísticos de séries temporais, incluindo SARIMA, GARCH, ETS e STLM. Os modelos SARIMA e STLM apresentaram um bom desempenho preditivo, sendo que o STLM se destacou por gerar resíduos mais regulares e intervalos de previsão mais estreitos, mantendo a cobertura. Este modelo foi, assim, selecionado como a melhor opção para previsão.

I. INTRODUÇÃO

A monitorização e previsão do consumo de eletricidade desempenham um papel central na gestão energética de qualquer país desenvolvido. Em contextos de transição energética, planeamento de capacidade, definição de políticas públicas e avaliação de impacto ambiental, a compreensão detalhada dos padrões de consumo é essencial. Espanha, sendo uma das principais economias da Europa, apresenta uma dinâmica energética particularmente relevante, com variações sazonais marcadas e um peso crescente das fontes renováveis [1].

Assim, o presente trabalho tem como objetivo modelar e prever o consumo mensal de eletricidade em Espanha, utilizando dados oficiais do Eurostat [2]. Através da análise de séries temporais, pretende-se identificar padrões estruturais e comportamentos sazonais, com vista à construção de modelos preditivos que possam apoiar a tomada de decisão estratégica no domínio da energia.

Com este objetivo, serão aplicados diversos modelos estatísticos de séries temporais, nomeadamente SARIMA, GARCH, ETS e STLM. A seleção destes modelos justifica-se pela sua reconhecida eficácia na captura de componentes sazonais, tendências estruturais e padrões residuais em séries univariadas, propriedades que caracterizam a série em análise.

A comparação entre os modelos será efetuada com base no desempenho preditivo em dados de teste, bem como na análise dos resíduos obtidos durante o treino. A avaliação conjunta destes critérios permite não apenas identificar o modelo com melhor capacidade de previsão, mas também assegurar que os pressupostos estatísticos subjacentes à modelação são razoavelmente cumpridos.

II. ANÁLISE EXPLORATÓRIA DOS DADOS

Como já referido anteriormente, os dados utilizados neste estudo estão disponíveis na plataforma do Eurostat [2]. Esta base de dados inclui, para além dos valores mensais de consumo de eletricidade em Espanha, informações equivalentes para um total de 40 países europeus. Tal abrangência torna esta fonte

particularmente robusta e adequada para estudos comparativos e análises aprofundadas no domínio da energia.

Neste estudo, foi selecionado exclusivamente o caso de Espanha, dado tratar-se de uma análise de séries temporais univariadas. Os dados correspondem ao consumo mensal de eletricidade neste país, medido em Gigawatt-hora (GWh), abrangendo o período de janeiro de 2008 a fevereiro de 2025. A série é completa, sem quaisquer valores em falta, totalizando 206 observações consecutivas.



Figura 1: Série temporal univariada do consumo mensal de eletricidade em Espanha (2008–2025).

Na Fig. 1 podem observar-se as 206 observações relativas ao caso em estudo, demonstrando a presença de sazonalidade nos dados, juntamente com uma possível tendência decrescente e um valor discrepante no início de 2020.

Este valor, após uma breve investigação, aparenta estar relacionado com o impacto da pandemia de COVID-19, que provocou alterações significativas nos padrões de consumo energético devido à redução da atividade económica, confinamentos e mudanças no comportamento da população [3]. Por não se tratar de um erro de registo ou medição, o valor não será removido nem ajustado. No entanto, será considerado com especial atenção na análise do desempenho dos diferentes modelos, caso se revele relevante.

A. Separação dos Dados em Treino e Teste

Para avaliar a capacidade preditiva dos modelos, a série temporal foi dividida em dois conjuntos: treino e teste. O conjunto de treino inclui as observações de janeiro de 2008 a outubro de 2021 (166 observações), enquanto que o conjunto de teste inclui as observações de novembro de 2021 a fevereiro de 2025 (40 observações).

Esta divisão resulta numa proporção aproximada de 80% para treino e 20% para teste, respeitando a ordem cronológica da série, o que é essencial em problemas de previsão temporal. Os modelos serão ajustados exclusivamente com os dados de treino, sendo a sua capacidade de generalização avaliada com base no desempenho sobre os dados de teste.

B. Análise da Estacionariedade

A estacionariedade é uma propriedade fundamental em muitas técnicas de modelação de séries temporais, exigindo que as características estatísticas da série, como a média, a variância e a autocovariância, permaneçam constantes ao longo do tempo [4].

Antes de proceder a uma avaliação mais formal da estacionariedade, analisou-se a necessidade de aplicar uma transformação de Box-Cox, com o objetivo de estabilizar a variância ao longo do tempo.

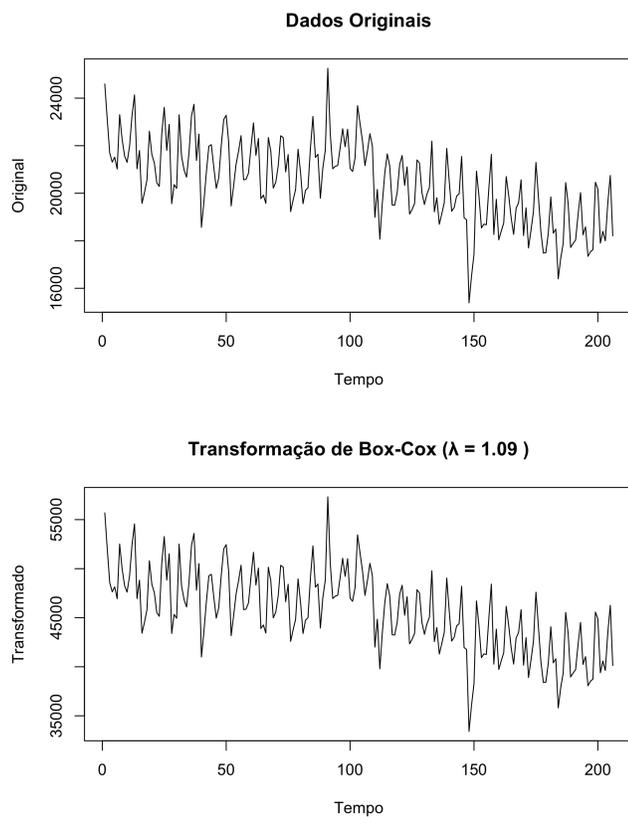


Figura 2: Avaliação da transformação Box-Cox.

Através da análise da Fig. 2, e tendo em conta o valor de λ ajustado automaticamente ($\lambda = 1.09$), bem como a representação gráfica da série transformada, conclui-se que não há qualquer benefício relevante em aplicar a transformação. Não se observam indícios visuais de heterocedasticidade na série original.

Para confirmar esta impressão, foi aplicado o teste de Breusch-Pagan à série original. Este teste avalia a presença de heterocedasticidade, sendo a hipótese nula de que a série tem variância constante (homocedasticidade), contra a alternativa

de que a variância não é constante. O p-valor obtido foi 0.5541, valor claramente superior a qualquer nível de significância usual, pelo que não se rejeita a hipótese nula. Assim, conclui-se que não há evidência estatística de heterocedasticidade na série, o que reforça a decisão de não aplicar qualquer transformação estabilizadora da variância e suporta a suposição de variância constante exigida pela estacionariedade.

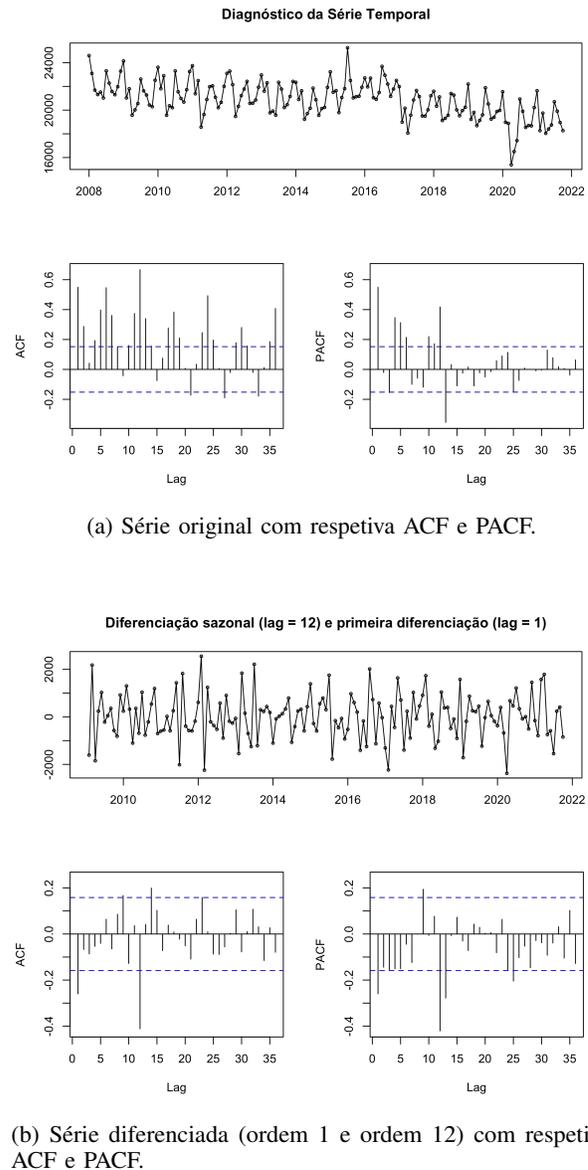


Figura 3: Avaliação visual da estacionariedade: série temporal original (a) e após diferenciações (b).

De seguida, procedeu-se à restante avaliação da estacionariedade da série, começando por uma inspeção visual (Fig. 3). Esta análise foi complementada pelas funções de autocorrelação (ACF) e autocorrelação parcial (PACF), que permitem identificar padrões de dependência temporal e indicar a presença de componentes sazonais ou tendências persistentes. Estas ferramentas oferecem uma primeira aproximação ao comportamento estrutural da série, orientando eventuais

transformações necessárias para satisfazer os pressupostos dos modelos a aplicar.

A partir da análise gráfica e das funções ACF e PACF, torna-se evidente que a série original não é estacionária, apresentando tanto sazonalidade, como uma leve tendência. No entanto, após a aplicação de diferenciações de ordem 1 (para remover a tendência) e 12 (para eliminar a componente sazonal), a série transformada revela um comportamento estacionário, com médias e variâncias aproximadamente constantes ao longo do tempo e autocorrelações significativamente reduzidas.

Esta transformação é essencial para garantir a validade dos pressupostos dos modelos lineares clássicos, como o SARIMA, que exigem estacionariedade da série de entrada.

Para confirmar formalmente esta avaliação, foram aplicados dois testes estatísticos complementares: o teste de Dickey-Fuller aumentado (ADF), cuja hipótese nula assume a presença de raiz unitária (não estacionariedade), e o teste KPSS, cuja hipótese nula assume estacionariedade.

A aplicação dos testes formais confirmou a avaliação visual inicial. No caso da série original, esta é indicada não estacionária. Após a aplicação de diferenciações de ordem 1 (para eliminar a tendência) e 12 (para remover a sazonalidade), ambos os testes conduzem à conclusão oposta: o teste ADF rejeita a presença de raiz unitária e o teste KPSS não rejeita a hipótese de estacionariedade.

Assim, a série transformada é considerada estacionária para qualquer nível de confiança típico, incluindo 5%.

III. PROPOSTAS DE MODELOS

Com o objetivo de modelar e prever o consumo mensal de eletricidade em Espanha, serão considerados vários modelos clássicos de séries temporais, nomeadamente modelos SARIMA, GARCH, ETS e STLM. Cada abordagem será avaliada prioritariamente com base na validade dos pressupostos estatísticos subjacentes. Apenas no caso dessa validação ser bem-sucedida, será então considerado o seu desempenho preditivo com base nos dados de teste, assegurando que a capacidade de previsão não compromete a consistência do modelo.

A. SARIMA

Dado que o modelo SARIMA permite incorporar explicitamente as diferenciações de primeira ordem e sazonal ($d = 1 \wedge D = 1$), será seguida uma abordagem dual para a sugestão dos parâmetros: por um lado, a análise da ACF e da PACF da série diferenciada orientará a escolha inicial das ordens dos componentes autorregressivos e de médias móveis, tanto sazonais como não sazonais; por outro lado, será realizada uma seleção sistemática de modelos SARIMA através da comparação do critério de informação corrigido (AICc), testando modelos com e sem imposição prévia das ordens de diferenciação definidas pela análise exploratória. O modelo final será escolhido com base no menor valor de AICc, desde que respeite os pressupostos de adequação estatística.

Assim, foram escolhidos três modelos para uma análise mais detalhada:

- SARIMA(1,1,1)(1,1,1)[12]

- SARIMA(2,0,2)(2,1,0)[12]
- SARIMA(0,1,2)(3,1,0)[12]

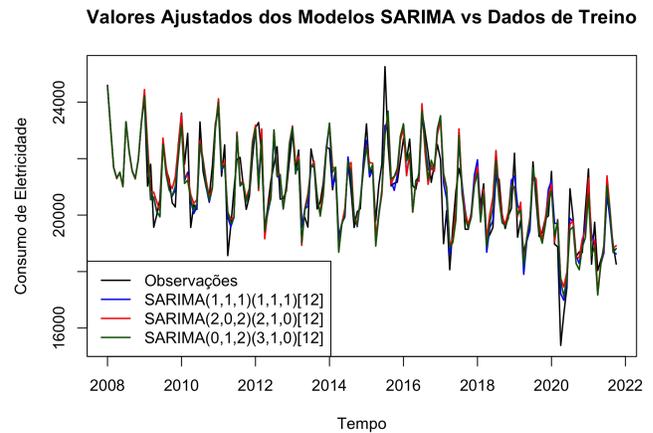


Figura 4: Ajuste dos diferentes SARIMA aos dados de treino.

Apesar de todos os modelos aparentarem um bom ajustamento aos dados de treino, como se observa na Fig. 4, a análise dos resíduos revelou falhas estruturais relevantes. Com base na função de autocorrelação dos resíduos e na aplicação do teste de Ljung-Box, constatou-se a presença de autocorrelação em todos os modelos testados. Tal indício compromete um dos pressupostos fundamentais da modelação SARIMA, a ausência de estrutura nos resíduos, colocando em causa a adequação estatística dos modelos ajustados.

Ainda assim, ao testar diferentes combinações de parâmetros em modelos SARIMA, o modelo SARIMA(1,1,1)(1,1,1)[12] foi o que apresentou resultados mais razoáveis em termos de ajustamento e simplicidade estrutural.

Apesar de o teste de Ljung-Box aplicado aos seus resíduos ter produzido um p-valor de 0.0513, ligeiramente acima do limiar convencional de 5%, tal resultado encontra-se num intervalo ambíguo. Embora tecnicamente não se rejeite a hipótese nula de ausência de autocorrelação, a proximidade ao valor crítico sugere cautela na aceitação da adequação do modelo.

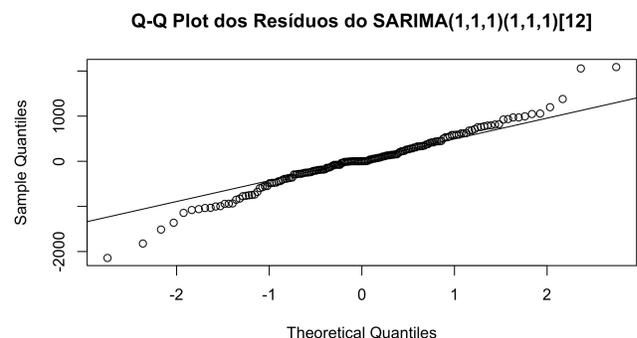


Figura 5: Gráfico QQ dos resíduos do modelo SARIMA.

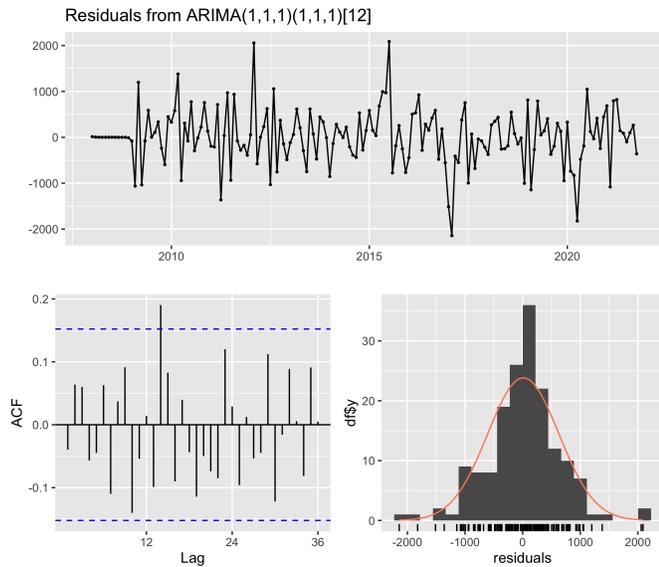


Figura 6: Análise dos resíduos do modelo SARIMA.

Quanto aos restantes pressupostos, através das Fig. 6 e 5, e dos testes de Shapiro–Wilk e Kolmogorov–Smirnov, foi possível avaliar a normalidade dos resíduos. As visualizações sugerem alguma aproximação à distribuição normal, mas com desvios ligeiros nas caudas. Esta ambiguidade visual reflete-se nos testes formais: o teste de Shapiro–Wilk rejeita a hipótese de normalidade, enquanto o teste de Kolmogorov–Smirnov não apresenta evidência suficiente para a rejeitar. Assim, a normalidade dos resíduos permanece uma questão debatível, com sinais contraditórios entre os métodos utilizados.

Quanto aos parâmetros do modelo SARIMA, os mesmos foram analisados tendo em conta a sua significância estatística e a correlação entre coeficientes. A análise revelou a insignificância do parâmetro sazonal autorregressivo (*sar1*), o que levanta dúvidas sobre a sua contribuição efetiva para o modelo. Adicionalmente, observou-se uma correlação elevada entre os parâmetros autorregressivo e de médias móveis não sazonais (*ar1* e *ma1*), com um coeficiente aproximado de -0.81 . Esta correlação forte pode indicar redundância estrutural no modelo e potenciais problemas de identificação.

Ainda assim, a remoção do parâmetro não significativo e de um dos parâmetros fortemente correlacionados resultou no agravamento da autocorrelação nos resíduos, comprometendo ainda mais a adequação do modelo. Pelo que, apesar das limitações identificadas, optou-se por manter a estrutura completa do modelo SARIMA(1,1,1)(1,1,1)[12], dado ser o compromisso mais aceitável entre qualidade do ajustamento, simplicidade estrutural e cumprimento dos pressupostos essenciais.

B. GARCH

Tendo em conta a possibilidade de existência de heterocedasticidade condicional, foi testada a presença de efeitos ARCH nos resíduos do modelo SARIMA ajustado (SARIMA(1,1,1)(1,1,1)[12]), recorrendo ao teste de Engle (ARCH-LM) com 12 lags.

O p-valor obtido foi 0.6337, valor claramente superior aos níveis de significância típicos. Assim, não se rejeita a hipótese

nula de ausência de efeitos ARCH, não havendo, portanto, evidência de heterocedasticidade condicional nos resíduos. Por esta razão, considera-se que a modelação GARCH não se justifica neste contexto, sendo descartada como proposta válida para a série em análise.

C. ETS

Procedeu-se também à aplicação de um modelo ETS (Error, Trend, Seasonal), uma abordagem adequada para séries temporais com componentes de tendência e sazonalidade bem definidas. Ao contrário dos modelos SARIMA, os modelos ETS não requerem estacionariedade, o que os torna particularmente úteis em contextos onde a estabilização da série pode distorcer a estrutura subjacente.

O modelo ajustado foi selecionado automaticamente com base no critério AICc, tendo sido escolhido o modelo ETS(M,N,A), ou seja, com erro multiplicativo, sem tendência e com sazonalidade aditiva.

No entanto, a análise dos resíduos revelou autocorrelação significativa. O teste de Ljung-Box aplicado aos resíduos produziu um p-valor de 0.002727, rejeitando claramente a hipótese nula de ausência de autocorrelação. Dada esta violação de um pressuposto fundamental, conclui-se que os modelos desta família não são adequados para a série em análise e são, por conseguinte, descartados.

D. STLM

Foi também considerado um modelo STLM (Seasonal and Trend decomposition using Loess + ARIMA), que combina a decomposição da série temporal em componentes sazonais e de tendência através de suavização local (Loess), seguida da modelação do componente restante com um modelo ARIMA [5].

Esta abordagem tem como vantagem permitir um tratamento explícito da sazonalidade, antes de ajustar a estrutura dinâmica residual, o que pode ser útil em séries com sazonalidade complexa ou não linear.

Após a decomposição e ajuste do ARIMA residual (selecionado automaticamente via AICc), a análise dos resíduos indicou ausência de autocorrelação significativa, tanto através do teste Ljung-Box (p-valor de 0.2003) como da análise gráfica dos resíduos (Fig. 8), e um comportamento globalmente aceitável do modelo.

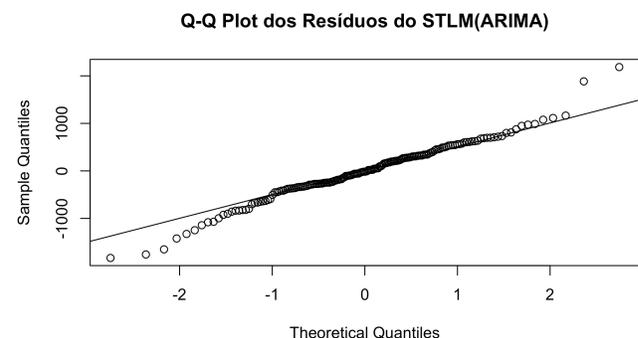


Figura 7: Gráfico QQ dos resíduos do modelo STLM.

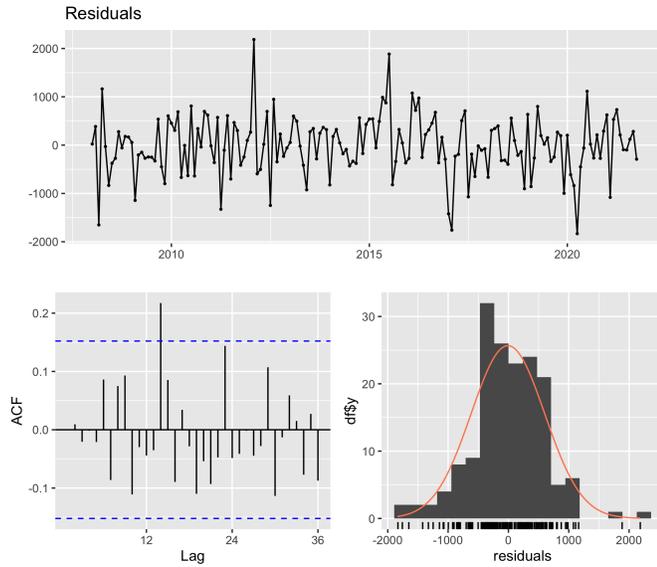


Figura 8: Análise dos resíduos do modelo STLM.

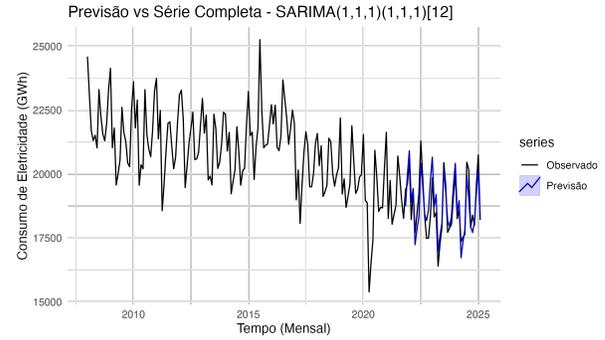
Este comportamento globalmente aceitável do modelo inclui, nomeadamente, a normalidade dos resíduos. A análise gráfica (Fig. 8 e 7) revela uma aproximação satisfatória à distribuição normal, sem desvios ou assimetrias relevantes. Esta impressão foi apoiada por testes formais. O teste de Shapiro–Wilk produziu um p-valor de 0.01422 que, embora sugira rejeição da hipótese de normalidade ao nível de significância de 5%, situa-se perto do limiar e pode ser interpretado com cautela, tendo em conta a elevada sensibilidade do teste a pequenas imperfeições. Por outro lado, o teste de Kolmogorov–Smirnov devolveu um p-valor de 0.3263, não permitindo rejeitar a hipótese de normalidade. Apesar da ligeira inconsistência entre os testes, o comportamento global dos resíduos pode ser considerado compatível com a normalidade, não comprometendo os pressupostos do modelo STLM nem a sua validade inferencial.

Dado o cumprimento dos pressupostos e a estabilidade do comportamento residual, o modelo STLM será considerado como um possível modelo para previsão, a ser testado posteriormente no conjunto de dados de teste. A sua seleção baseia-se na solidez estrutural e na adequação estatística observada durante o ajuste aos dados de treino.

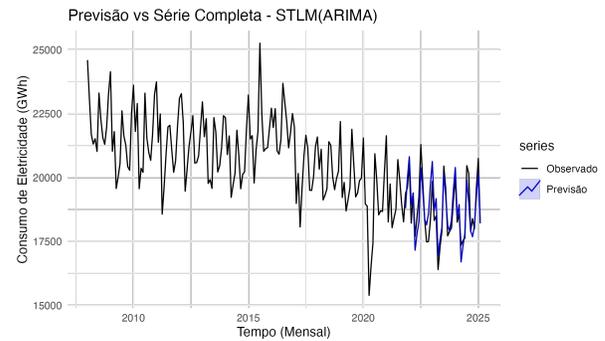
IV. PREVISÃO DE OBSERVAÇÕES FUTURAS

Tendo em conta a análise dos modelos propostos, os modelos selecionados para avaliação da capacidade preditiva serão o SARIMA(1,1,1)(1,1,1)[12] e o STLM com componente residual modelado por ARIMA. Ambos demonstraram ajustamentos estatisticamente aceitáveis aos dados de treino, cumprindo os principais pressupostos. Estes serão agora comparados com base no seu desempenho sobre o conjunto de teste.

Através da comparação das métricas de previsão obtidas sobre os dados de teste com as métricas de ajustamento sobre os dados de treino, observou-se uma semelhança significativa entre ambas. Esta estabilidade entre fases indica que não há indícios de *overfitting*, o que reforça a robustez dos modelos selecionados para fins preditivos.



(a) Modelo SARIMA(1,1,1)(1,1,1)[12].



(b) Modelo STLM(ARIMA).

Figura 9: Comparação entre valores reais e previsões para o modelo SARIMA (a) e STLM (b).

A Fig. 9 apresenta a comparação gráfica entre os valores observados e as previsões geradas pelos modelos, ao longo do período de teste. Visualmente, ambos os modelos acompanham com precisão a estrutura da série, sem desvios sistemáticos aparentes, reforçando a consistência do ajustamento estatístico já verificado anteriormente.

Tabela I: Comparação das métricas de previsão no conjunto de teste: SARIMA vs STLM

Métrica	SARIMA(1,1,1)(1,1,1)[12]	STLM(ARIMA)
ME	-52.21608	-26.13118
RMSE	562.8728	563.2452
MAE	465.7346	460.1395
MPE	-0.35018	-0.21075
MAPE	2.48337	2.45174
MASE	0.56503	0.55824
ACF1	0.58261	0.58367
Theil's U	0.42074	0.42183

A Tabela I apresenta a comparação direta entre os modelos SARIMA e STLM com base no seu desempenho preditivo sobre o conjunto de teste. As métricas obtidas revelam valores bastante semelhantes entre os dois modelos, não existindo diferenças expressivas que justifiquem a escolha inequívoca de um em detrimento do outro apenas com base na precisão da previsão.

Dado este equilíbrio em termos de erro preditivo pontual, torna-se particularmente relevante a análise dos intervalos de previsão. Esta dimensão permite avaliar a incerteza associada às

previsões de cada modelo, bem como a amplitude e a cobertura dos respetivos intervalos.

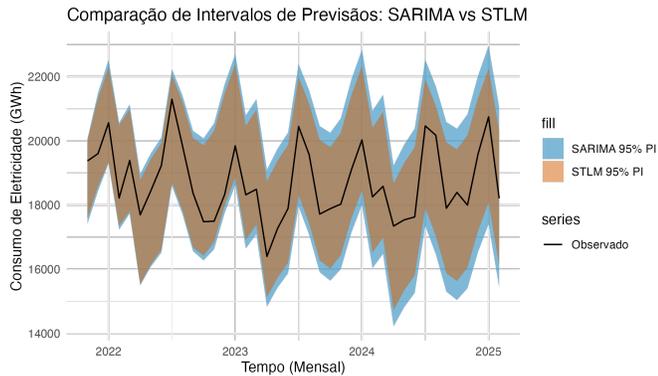


Figura 10: Intervalos de previsão a 95% para os modelos SARIMA e STLM.

A Fig. 10 ilustra os intervalos de previsão a 95% dos dois modelos. Ambos os modelos apresentam cobertura total dos valores reais dentro dos seus intervalos de confiança, ou seja, 100% das observações do período de teste encontram-se contidas nas previsões com 95% de confiança, tanto para o modelo SARIMA como para o STLM.

No entanto, a largura média dos intervalos difere entre os modelos: o SARIMA apresenta uma amplitude média de 4378.183, enquanto o modelo STLM regista uma largura média inferior, de 3639.196. O modelo STLM produz, assim, intervalos de previsão a 95% mais estreitos em média, quando comparado com o SARIMA. Isto significa que o STLM gera previsões com menor incerteza e idêntica precisão. Intervalos mais estreitos indicam maior confiança do modelo nas suas previsões, o que é particularmente desejável quando a cobertura não é sacrificada, como é o caso presente, visto que ambos os modelos mantêm 100% de cobertura.

V. DISCUSSÃO DOS RESULTADOS

A análise desenvolvida permitiu avaliar comparativamente múltiplas abordagens de modelação para a previsão do consumo mensal de eletricidade em Espanha. Embora tanto o modelo SARIMA(1,1,1)(1,1,1)[12] como o STLM com componente ARIMA tenham revelado desempenho preditivo semelhante, a avaliação aprofundada dos resíduos e dos intervalos de previsão revelou diferenças estruturais relevantes.

O modelo STLM demonstrou menor autocorrelação residual, variância inferior e maior conformidade com a suposição de normalidade dos erros. Estas características sugerem uma modelação mais robusta e estatisticamente fiável. Além disso, a capacidade de gerar intervalos de previsão mais estreitos, mantendo 100% de cobertura no período de teste, evidencia maior confiança associada às suas previsões, um fator crítico na comunicação de resultados em contextos operacionais e de apoio à decisão.

Apesar do bom desempenho do modelo SARIMA, a sua estrutura residual mais ruidosa e a dependência de transformações diferenciais podem limitar a sua interpretabilidade e estabilidade em cenários futuros. Já o STLM, ao

separar explicitamente os componentes estruturais da série antes da modelação dinâmica, oferece maior transparência no ajustamento e potencial adaptabilidade a alterações na sazonalidade.

Em síntese, a escolha do modelo STLM(ARIMA) surge como uma decisão natural e justificada, não se baseando exclusivamente em métricas pontuais de erro, mas numa análise integrada dos pressupostos, da estabilidade dos resíduos e da qualidade informativa das previsões. Estes fatores convergem para uma decisão fundamentada e alinhada com os princípios da modelação estatística em séries temporais.

VI. CONCLUSÃO

Este estudo teve como objetivo modelar e prever o consumo mensal de eletricidade em Espanha, com base em dados oficiais do Eurostat. A análise envolveu a aplicação e comparação de vários modelos estatísticos de séries temporais, incluindo SARIMA, ETS, GARCH e STLM, avaliados segundo critérios de ajustamento, validação dos pressupostos e desempenho preditivo.

Apesar do modelo SARIMA(1,1,1)(1,1,1)[12] ter apresentado um bom desempenho geral, foi o modelo STLM(ARIMA) que demonstrou maior robustez estrutural: resíduos mais regulares, menor autocorrelação, maior aproximação à normalidade e intervalos de previsão mais estreitos, mantendo a cobertura total. Estes resultados sustentam a sua escolha como modelo final para previsão.

Conclui-se que o modelo STLM, ao conjugar decomposição estrutural com modelação ARIMA, representa uma solução eficaz para séries com forte sazonalidade, como é o caso do consumo de eletricidade. A abordagem seguida neste estudo fornece um enquadramento sólido para aplicações semelhantes noutras séries energéticas ou económicas com características análogas.

REFERÊNCIAS

- [1] Red Eléctrica de España, "Informe del sistema eléctrico 2023," March 2024, https://www.sistemaelectrico-ree.es/sites/default/files/2024-03/ISE_2023.pdf, acesso em 28 maio 2025.
- [2] Eurostat, "Supply, transformation and consumption of electricity - monthly data," 2025, https://doi.org/10.2908/NRG_CB_EM, acesso em 18 maio 2025.
- [3] I. Santiago, A. Moreno-Munoz, P. Quintero-Jiménez, F. Garcia-Torres, and M. González-Redondo, "Electricity demand during pandemic times: The case of the covid-19 in Spain," *Energy Policy*, vol. 148, p. 111964, January 2021, 10.1016/j.enpol.2020.111964.
- [4] A. Petrică, S. Stancu, and V. Ghițulescu, "Stationarity – the central concept in time series analysis," *International Journal of Emerging Research in Management and Technology*, vol. 6, pp. 6–16, 01 2017.
- [5] R. Hyndman, G. Athanasopoulos, C. Bergmeir, G. Caceres, L. Chhay, K. Kuroptev, M. O'Hara-Wild, F. Petropoulos, S. Razbash, E. Wang, and F. Yasmeeen, *forecast: Forecasting functions for time series and linear models*, 2024, r package version 8.24, <https://pkg.robjhyndman.com/forecast/reference/forecast.stl.html>.