

Relatório do Mini Projeto

Índice

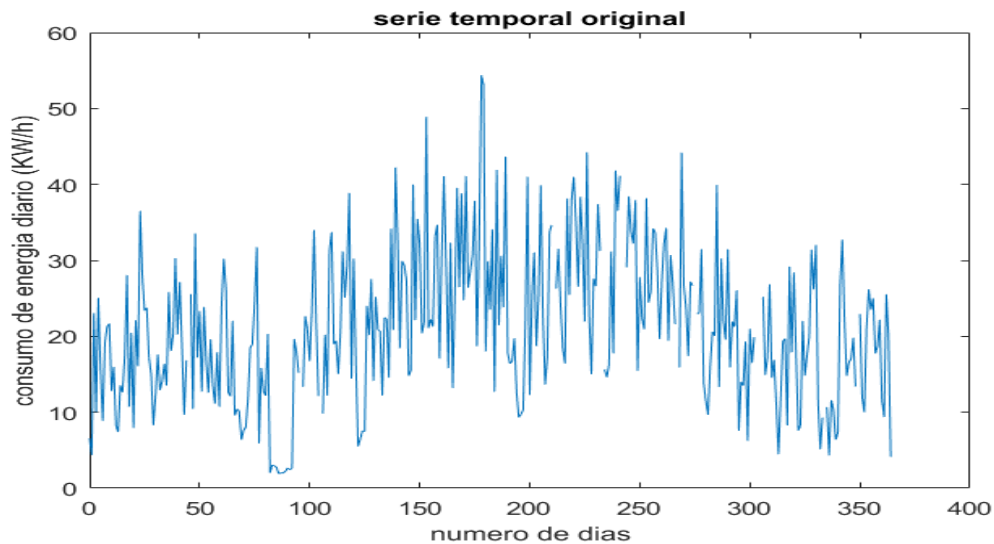
Introdução	3
Pré-processamento dos dados do dataset “dataset_ATD_PL8.csv”	3
Análise da série temporal.....	4
Estimação do Modelo da Série Temporal	6
Conclusão.....	12

Introdução

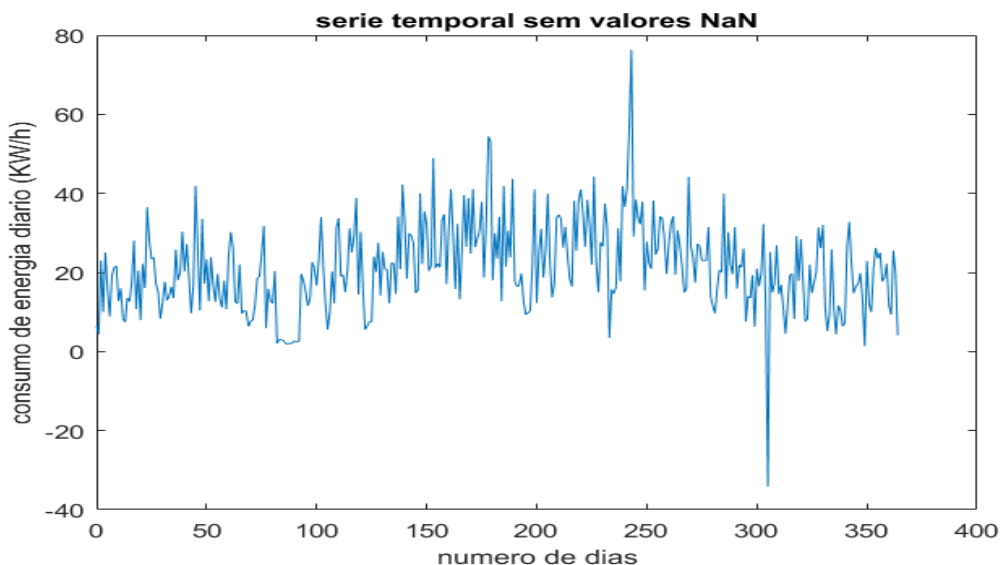
Neste trabalho, foi-nos proposta a análise de uma serie temporal dada por um dataset de valores de consumos totais diários de energia, tendo como objetivo estimar um modelo adequado, tendo em vista, a previsão de valores futuros.

Pré-processamento dos dados do dataset “dataset_ATD_PL8.csv”

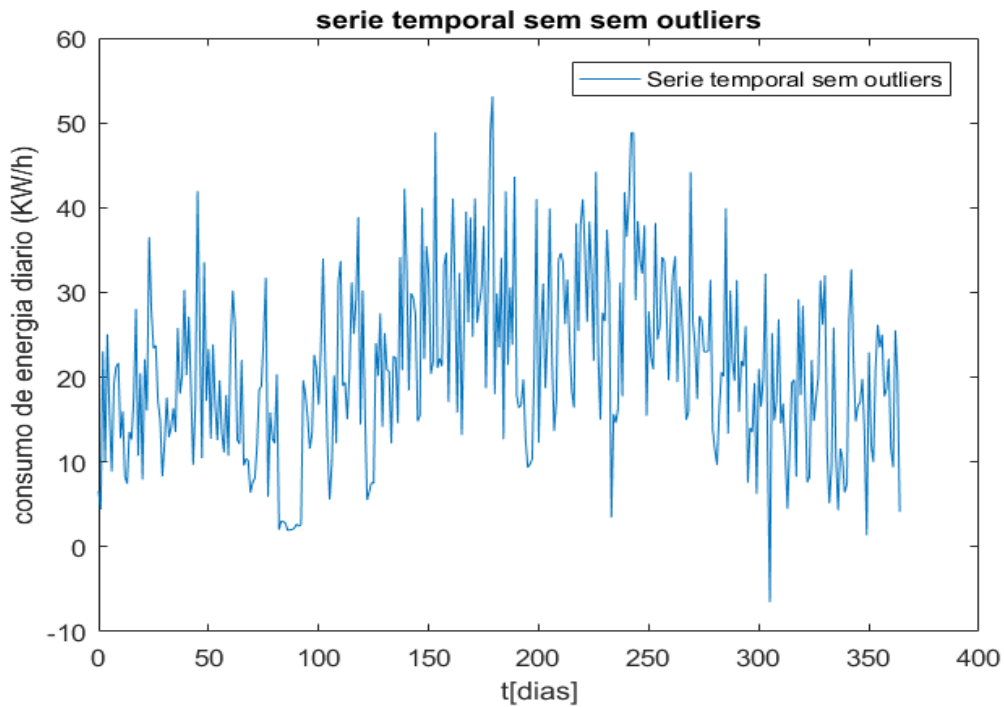
Ao carregar o dataset, verifica-se que o mesmo não se verifica completo e encontra-se com valores NaN (Not a Number), como a imagem abaixo o comprova.



Neste caso, devemos identificar quais os índices com valores em falta e substituindo por valores estimados por interpolação.

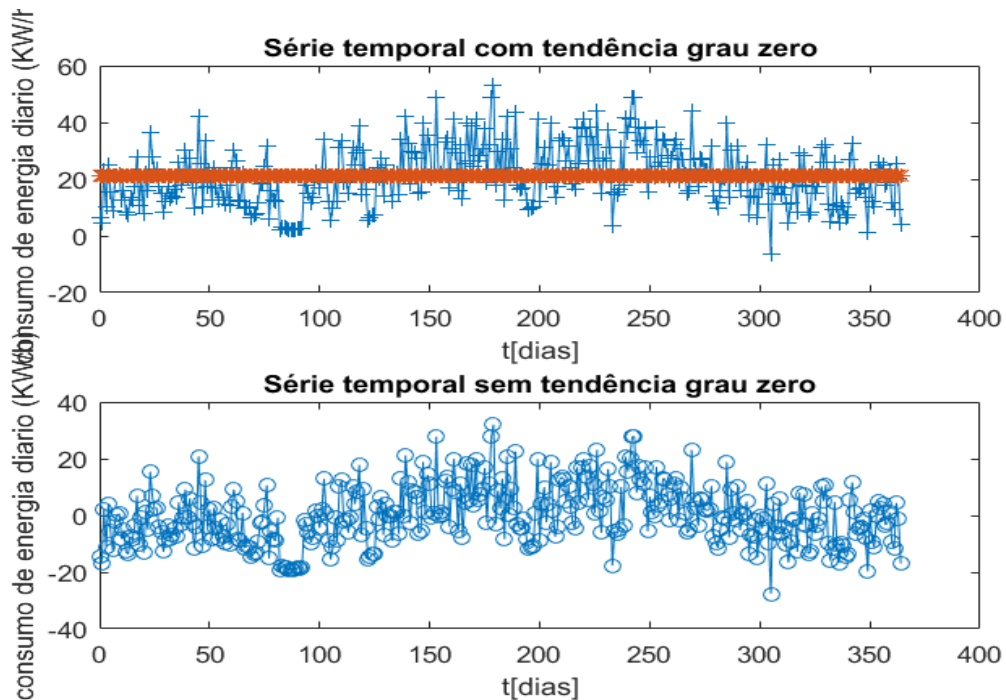


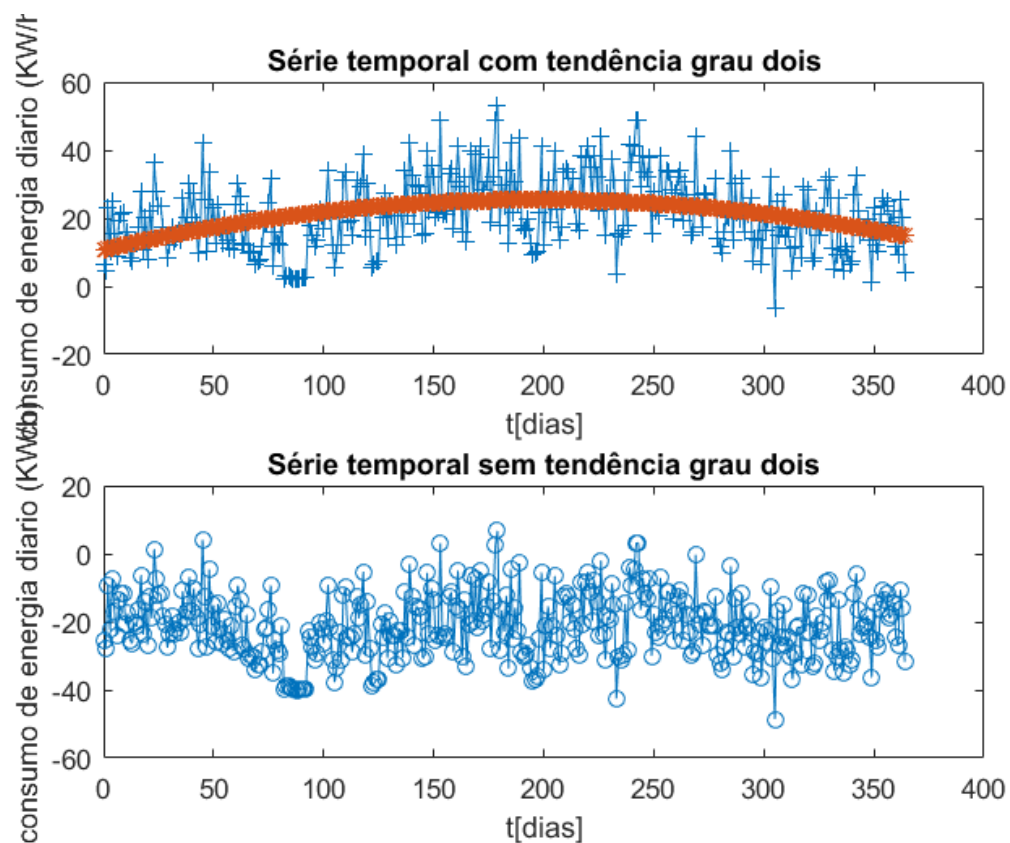
Como podemos verificar existiam outliers na série em estudo. No seu tratamento usou-se o critério $|x_i - \mu| > 3\sigma$.



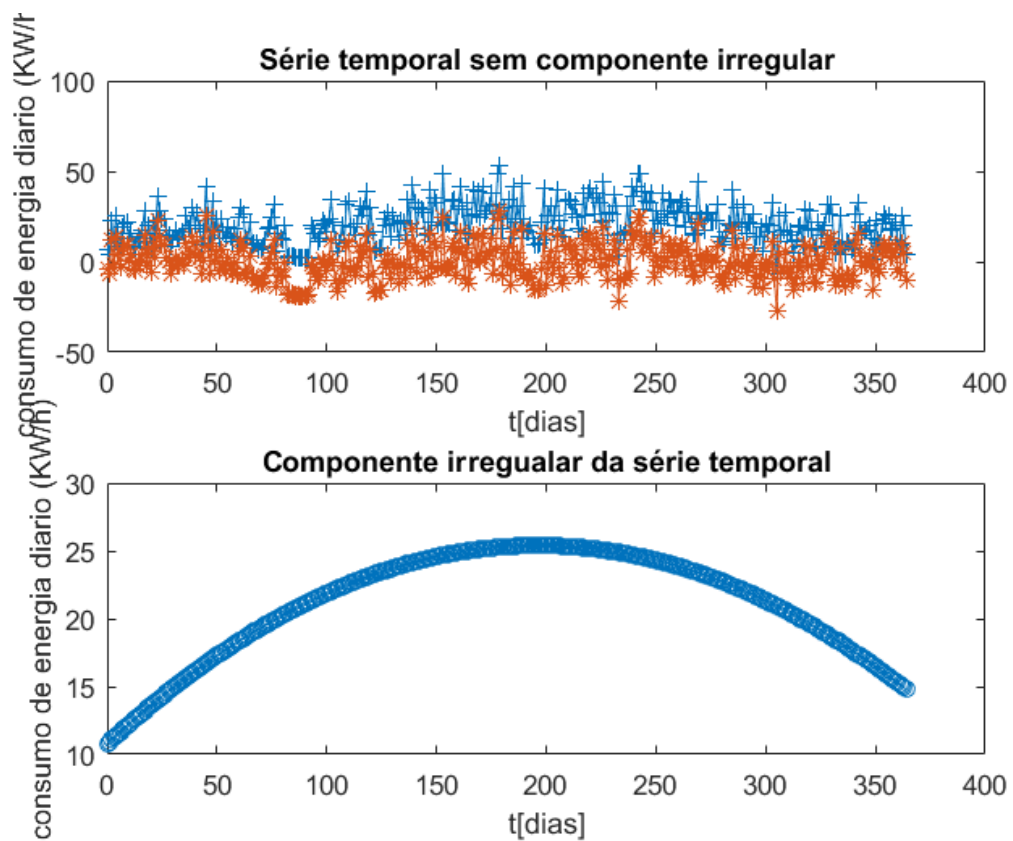
Análise da série temporal

Depois da série regularizada procedeu-se a identificação da tendência da série.

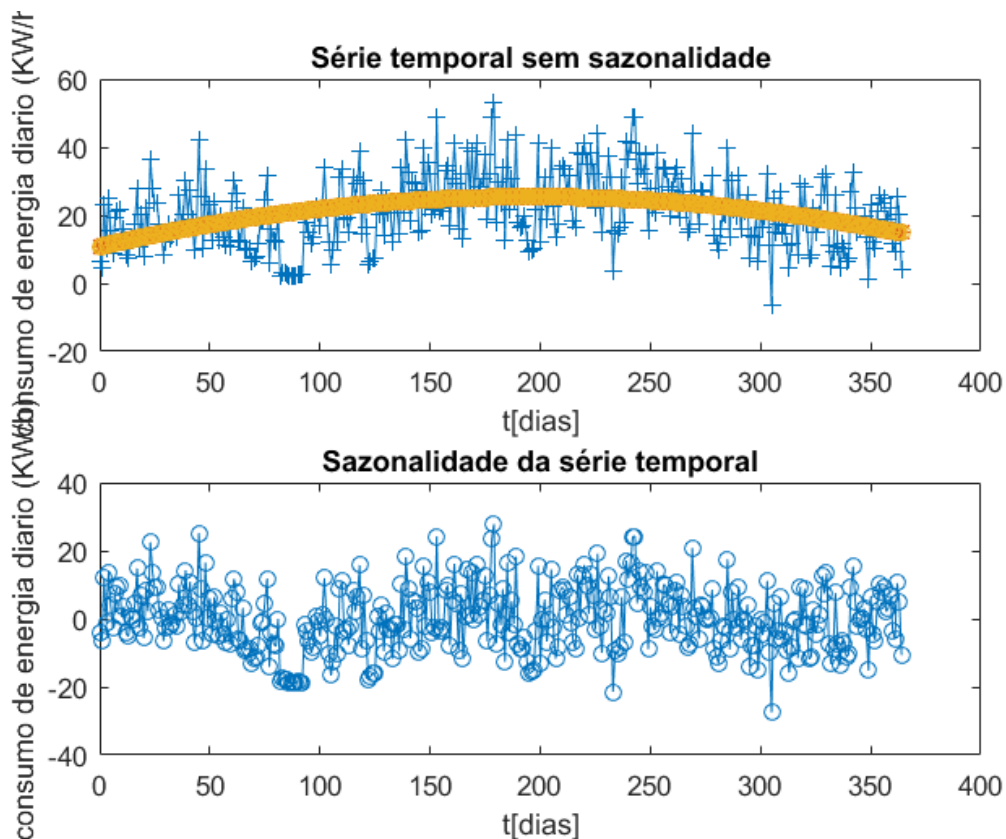




Imagens relativas à identificação da componente irregular da série.



Imagens relativas à identificação da componente sazonal da série.

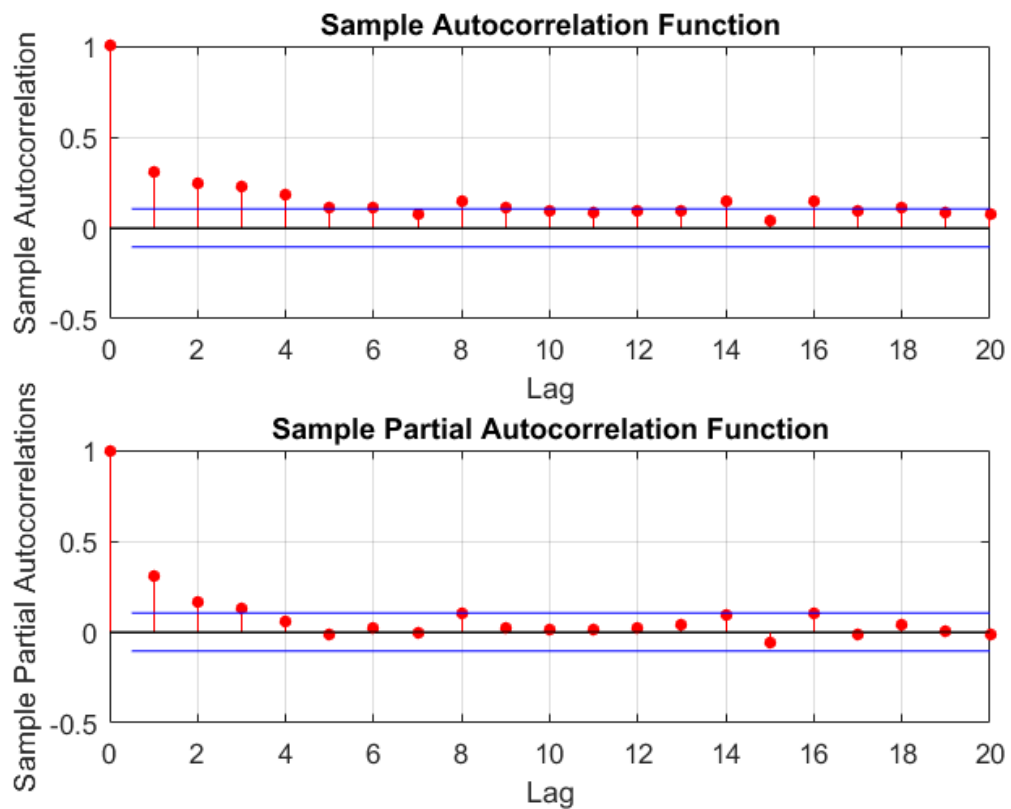


Estimação do Modelo da Série Temporal

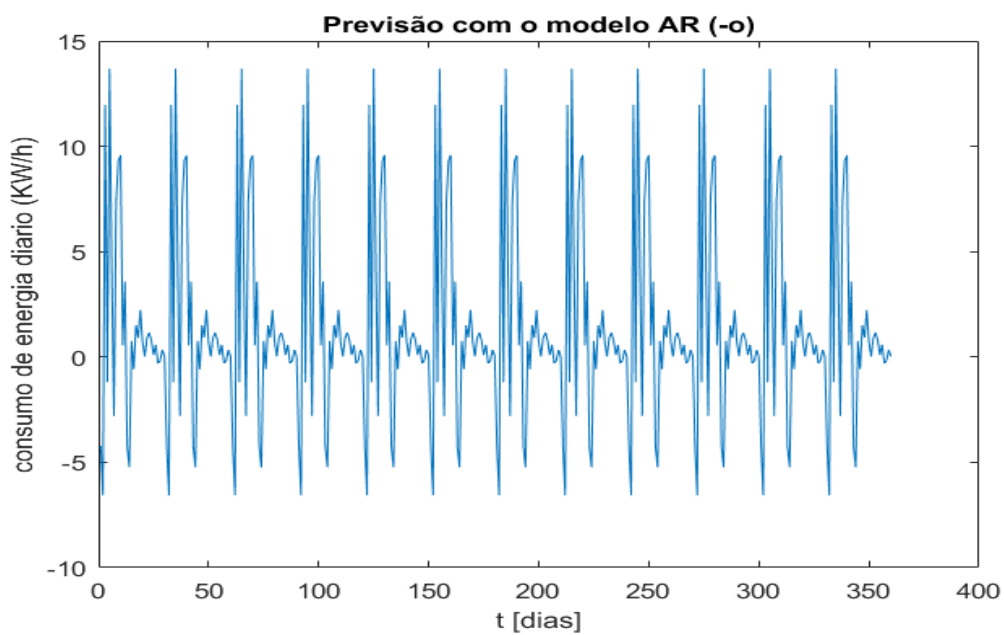
Na estimação do modelo desta série, procedeu-se à realização de três modelos, para se encontrar o que melhor se aproxima do comportamento da mesma.

O primeiro modelo de estudo foi o AR (Auto-regressivo), antes de se proceder ao desenvolvimento deste modelo, houve tarefas *à priori* em estudo.

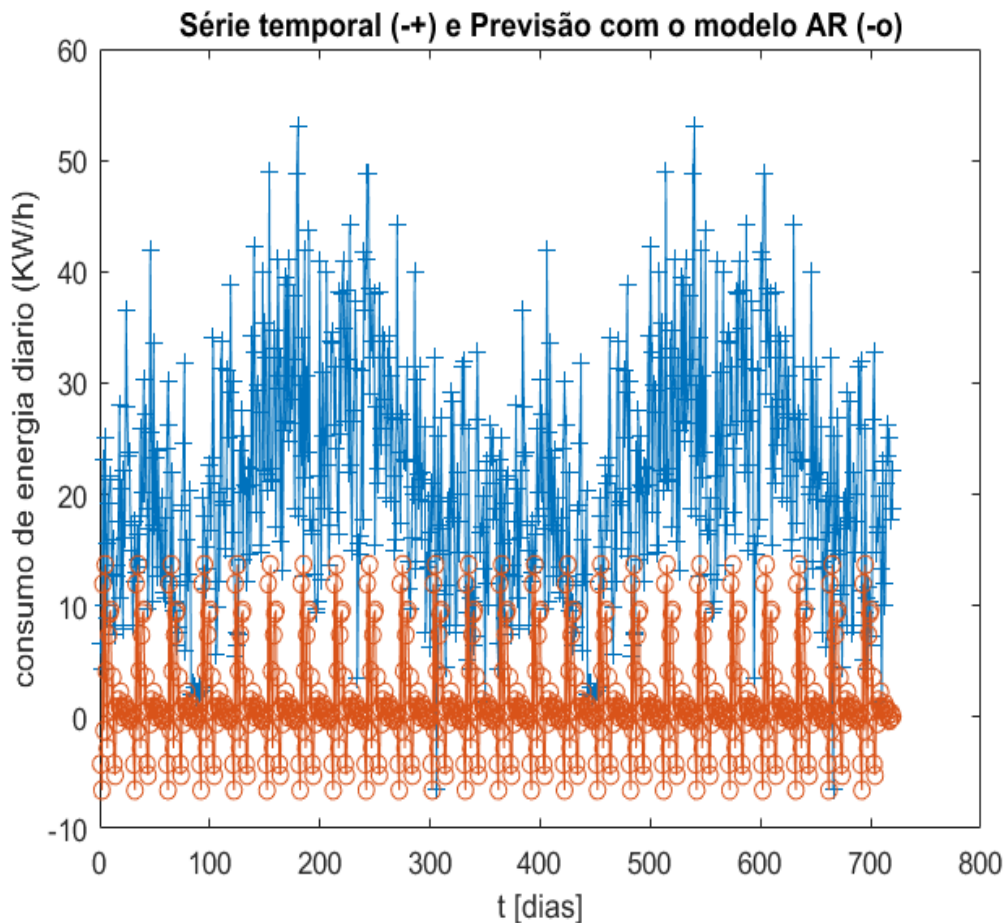
A primeira tarefa teve que ver com a estacionaridade da série era 1, o que se acabou por verificar. De seguida, terá sido necessário obter a auto correlação parcial, assim como a auto correlação da série.



Depois de obter estes dados, é possível calcular o modelo da série para este modelo. Para isso os valores de $na - 15$, devido ao índice 15 estar mais próximo do limite inferior. Depois do modelo calculado foi só obter o gráfico final.

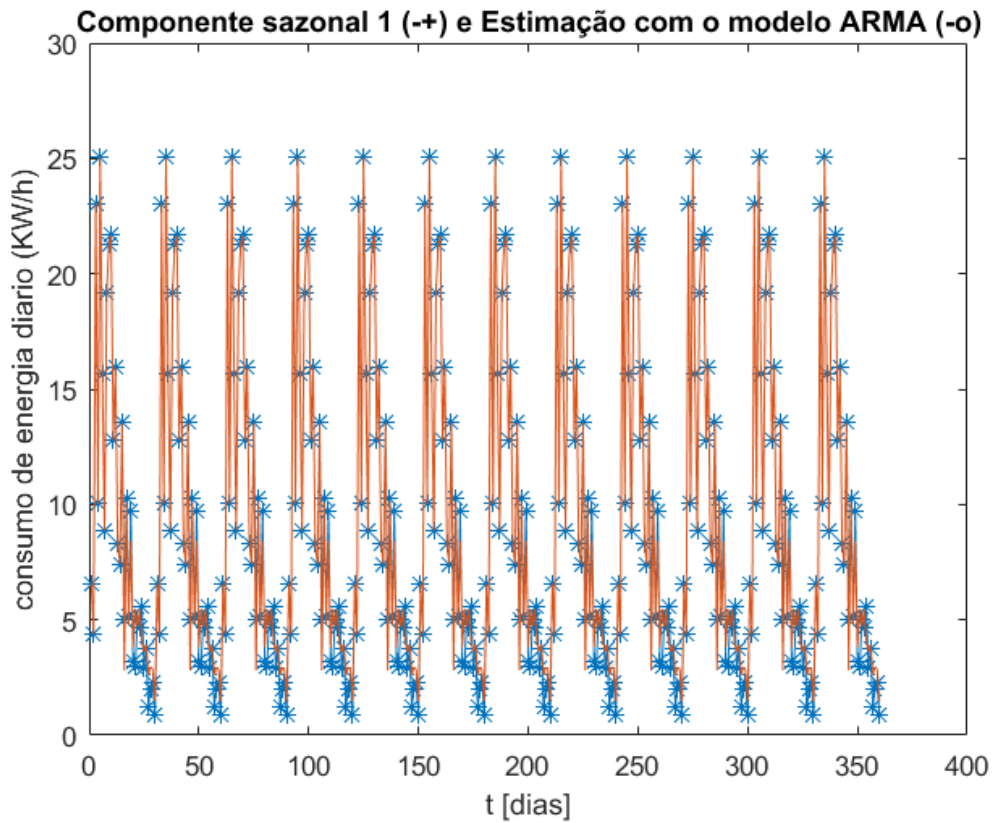


Ao observar o gráfico do modelo AR, nota-se que a série é periódica no período de estudo de 30 dias, e replicado para os restantes meses. O mesmo se verifica na previsão para o dobro do tamanho da série.

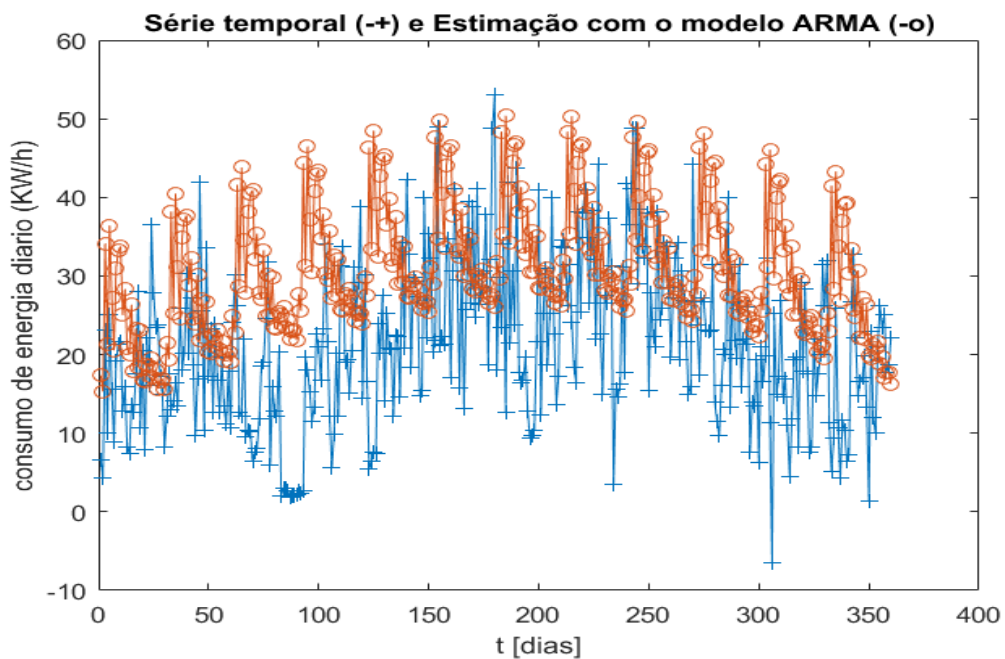


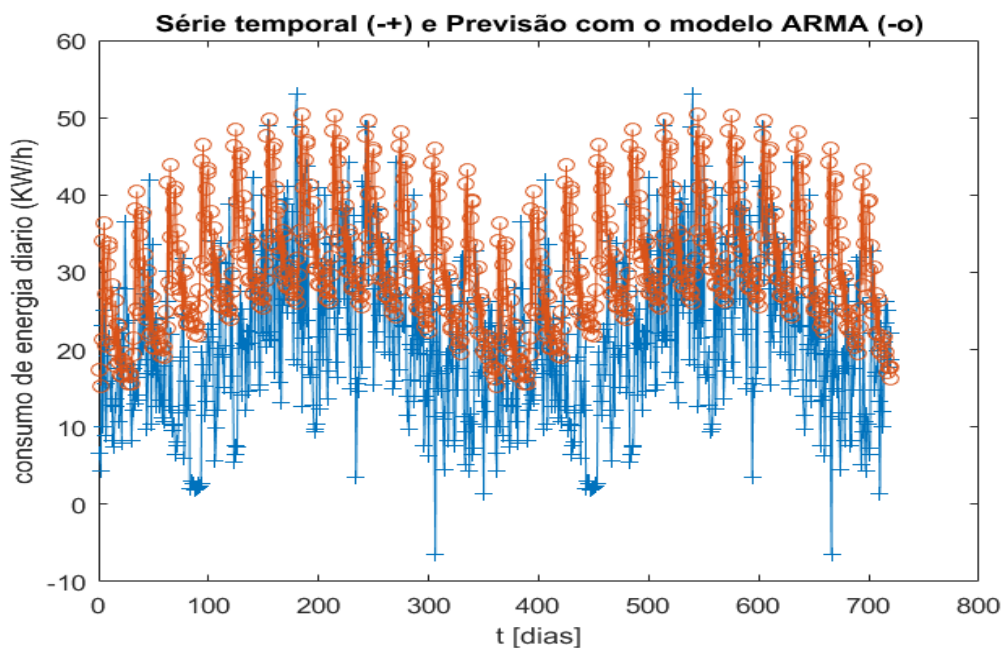
O próximo modelo estudado foi o ARMA (Auto-Regressivo de Médias Móveis). Este modelo difere do anterior, porque usa o ruído branco. Para tal, usou-se o mesmo valor de n_a , ou seja 15 e n_c igual a 1. Para o ruído branco, usou-se um vetor de valores pseudoaleatórios de tamanho 30, estudo em média para 30 dias por mês. Feito a previsão do modelo de 30 dias, replicou-se por 12 meses, fazendo um total de 360 dias, não afetando a série no tamanho integral dos 365 dias.

Antes de analisar o modelo obteve-se a sazonalidade do modelo, em que se verifica que é periódica.

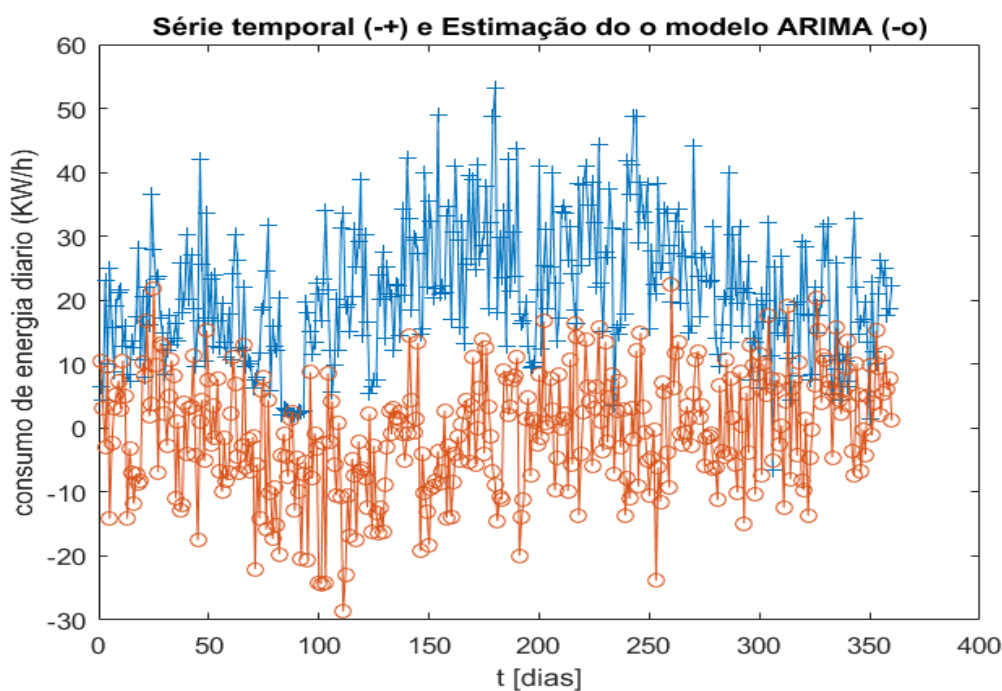


Podemos verificar que a série, neste modelo, tende a ser periódica no caso de estudo, seguidamente fez-se a previsão para um ano, ou seja para o dobro do tamanho da série, em que podemos também verificar que o comportamento tende a ser igual.

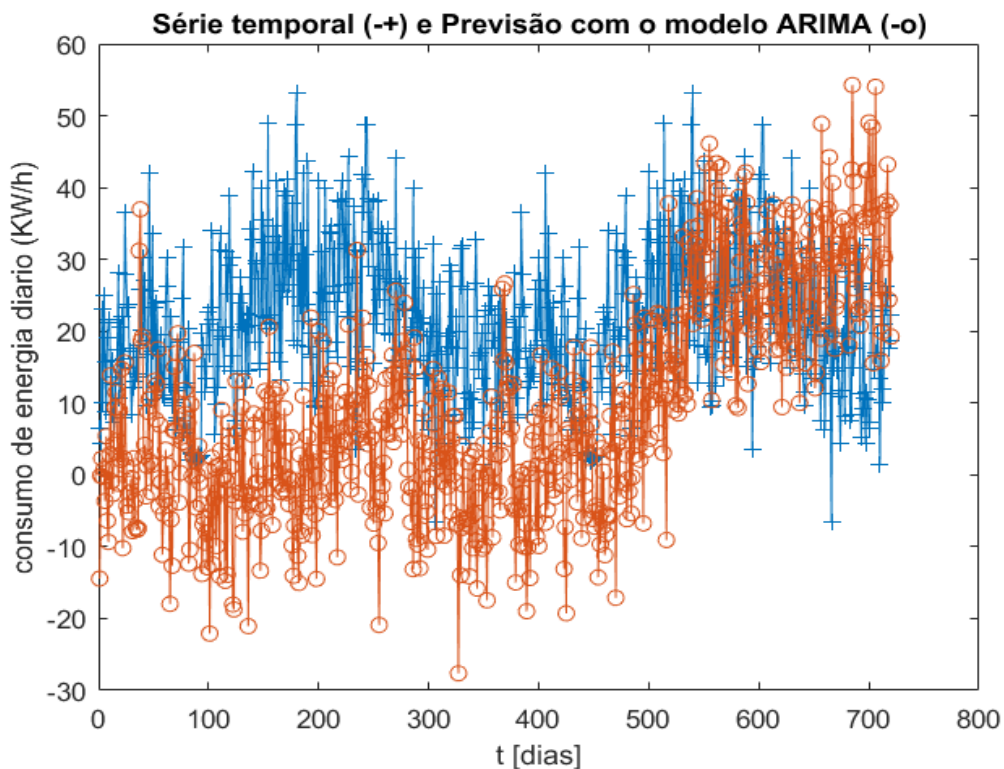




Por fim estudou-se o modelo ARIMA (Auto-Regressivo Integrado de Médias Móveis). Para o cálculo do modelo foram usados os seguintes valores para a criação do modelo; $P=30$, $D=1$ e $Q=1$. Designa-se o valor de P (30) de um modo elevado, por ter sido baseado no mesmo princípio dos outros anteriormente descritos (estudo para um mês, e de seguida, replicar para os restantes meses). A estimação do modelo foi calculada com um tamanho de 360 dias. A grande diferença deste modelo para os anteriores, é que este usa os valores originais da série, com a tendência, enquanto os anteriores usam a componente sazonal da série para fins de estudo.



A primeira conclusão que retiramos da estimação do modelo, como usa os valores originais, a previsão, dá-nos valores negativos, que até faz sentido, visto que a cada valor calculado, usa a média dos anteriores, e nos leva a concluir que num dado dia o consumo tenha sido menor, podemos concluir que nem possa ter existido consumo de energia nesse instante da série. Estas conclusões voltaremos a retirá-las ao fazer a previsão futura para 1 ano desta série.



Para um horizonte futuro podemos verificar que a série, tende a convergir para os valores originais ao fim de um ano, com isto podemos concluir que este modelo se adequa a série que se estudou.

Conclusão

Em conclusão, na análise final deste trabalho, podemos retirar vários aspetos. Em primeiro lugar, conclui-se que nem todos os modelos para estudo se encaixam num *dataset*, tendo em conta este caso em específico.

Os modelos AR e ARMA, vão dar resultados periódicos, fazendo sentido *à priori*, visto que se estudou para o período temporal de 30 dias, e resultando para os restantes meses de igual forma. Em último ponto, o Modelo ARIMA é o único que se comporta como os dados originais, mesmo tendo um tempo de convergência tardio, isto dá-nos uma previsão teórica do comportamento da série em função do tempo.