Time Series Analysis 2024

Part 1 – Project 1 **100 Marks** **Due**: Sunday, 15 September 2024 @ 23:55

Data: You have been provided with hourly electricity demand data for South Africa for the period 01/04/2019 to 13/09/2023 as obtained from Eskom. The data can be found in the attached CSV file named 'ESK6816.csv'. You can also find descriptions of the variables in the dataset at https://www.eskom.co.za/dataportal/glossary/.

Task: Your assignment is to forecast the *peak daily demand series* **and** *unplanned generation outages (UCLF + OCLF)* utilising the techniques we have covered in Part 1 of the course.

Your report: Please structure your analysis in a report format, including sections such as an abstract, introduction, body, conclusion, references, and any relevant appendices.

- The write-up must be your work.
- Make sure you show all your work and include enough technical detail so that someone else will be able to precisely replicate your work.
- The main body of your report should include only the relevant calculations, tables and plots and should not exceed 25 pages in length.
    - All relevant supplementary information including your R codes should be included in the appendix or as a separate file. Where necessary, set a seed and report it, so that your work can be reproduced.

At a minimum, your report should comprise of the following components:

- An introduction and a description of your data.
- A summary of the main question(s)/hypotheses to be addressed.
    - Approach this as if it were a consulting project. Consider what the recipient of your work would be interested in knowing about the problem and what interesting questions might arise. Please make it explicit what type of forecast you intend to conduct (short-term, medium-term, long-term, or a blend of these) and why.
- Exploratory data analysis including a discussion of relevant data cleaning/transformation tasks you have implemented.
- Results from fitting different models to your data, including at least one model from each of the following: AR, MA, ARMA, ARIMA, SARIMA, ARCH and GARCH models.
    - In general, discuss all outputs (tables and plots), and give good interpretations and explanations of results.
    - You should state clearly what you have done and assumed in your own words: justify your choice(s) and state any assumptions you make. Choices related to data partitioning, choice of forecast horizon, accuracy assessment and others should be duly considered.
    - Compare your models (including the forecasts from the models) and carry out residual analysis.
    - Please be aware that Eskom's predictions for the timeframe spanning from 01/04/2019 to 31/03/2024 can be located in the dataset under the column labelled "RSA Contracted Forecast." make sure you consider this information when conducting your comparisons.
- Discussion and Conclusions: Revisit your main question(s)/hypotheses and demonstrate that you have chosen a satisfactory model.
    - Discuss the limitations of your chosen model and how the model might be improved.
- A few marks are reserved for proper referencing, grammatical correctness, and spell checks.

Please hand in an R Markdown (RMD) file and the rendered PDF file.