Semester One 2025
Examination Period

Faculty of Business & Economics

UNIT CODES: **ETF3231/ETF5231**

TITLE OF PAPER: **Business Forecasting**

EXAM DURATION: 2 hours 10 minutes

AUTHORISED MATERIALS

This is a closed book exam, with the following permitted items.

- A physical calculator of any type or Virtual Calculator:
  - Inbuilt Mac/Windows calculator
  - Website https://www.educalc.net/2336211.page
  - 10bii Financial Calculator for Mac by K2 Cashflow, https://apps.apple.com/au/app/10bii-financial-calculator/id473144920
- 5 blank pages for use as working sheets
- 2 pre-printed answer sheets

RULES

During your eExam, you must not have in your possession any item/material that has not been authorised for your exam. This includes books, notes, paper, electronic device/s, smart watch/device, or writing on any part of your body. Authorised items are listed above. Items/materials on your device, desk, chair, in your clothing or otherwise on your person will be deemed to be in your possession. Mobile phones must be switched off and placed face-down on your desk during your exam attempt.

You must not retain, copy, memorise or note down any exam content for personal use or to share with any other person by any means during or following your exam. You are not allowed to copy/paste text to or from external sources unless this has been authorised by your Chief Examiner.

You must comply with any instructions given to you by Monash exam staff.

As a student, and under Monash University's Student Academic Integrity procedure, you must undertake all your assessments with honesty and integrity. You must not allow anyone else to do work for you and you must not do any work for others. You must not contact, or attempt to contact, another person in an attempt to gain unfair advantage during your assessment. Assessors may take reasonable steps to check that your work displays the expected standards of academic integrity.

Failure to comply with the above instructions, or attempting to cheat or cheating in an assessment may constitute a breach of instructions under regulation 23 of the Monash University (Academic Board) Regulations or may constitute an act of academic misconduct under Part 7 of the Monash University (Council) Regulations.

The exam contains FIVE sections. ALL sections must be completed. The exam is worth 100 marks in total.

Below are the State Space equations for each of the models in the ETS framework.

ADDITIVE ERROR MODELS

Trend	Seasonal N	Seasonal A	Seasonal M
N	$y_t = \ell_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$	$y_t = \ell_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/\ell_{t-1}$
A	$y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + b_{t-1})$
A$_d$	$y_t = \ell_{t-1} + \phi b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + \phi b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = \phi b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + \phi b_{t-1})$

MULTIPLICATIVE ERROR MODELS

Trend	Seasonal N	Seasonal A	Seasonal M
N	$y_t = \ell_{t-1}(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$	$y_t = (\ell_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + \alpha(\ell_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + s_{t-m})\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m}(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$
A	$y_t = (\ell_{t-1} + b_{t-1})(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$
A$_d$	$y_t = (\ell_{t-1} + \phi b_{t-1})(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$

SECTION A

Respond to any **four** of the following topics. For each, demonstrate your understanding by presenting at least five key insights or explanations.

1. An STL decomposition into trend, seasonal and remainder components is only useful for forecasting when there are no cycles in the data.

   5 marks

2. Forecasting assumes that the patterns in the past will continue in the future.

   5 marks

3. Benchmark forecasting methods are too simplistic to be used in practice.

   5 marks

4. Box-Cox transformations are useful for making a time series stationary.

   5 marks

5. Time series cross-validation is not really necessary because we can just use the AICc.

   5 marks

6. Never produce forecasts using a model that has correlated residuals.

   5 marks

Total: 20 marks

SECTION B

Figures 1–3 relate to electricity demand for the whole state of Queensland in Australia. The data are recorded in 5-minute intervals, spanning the period 1 January 2022 to 31 March 2025.

```
qld_data |>
  autoplot(Demand) +
  labs(
    title = "5-min electricity demand for the state of Queensland",
    y = "Demand (MW)"
  )
```

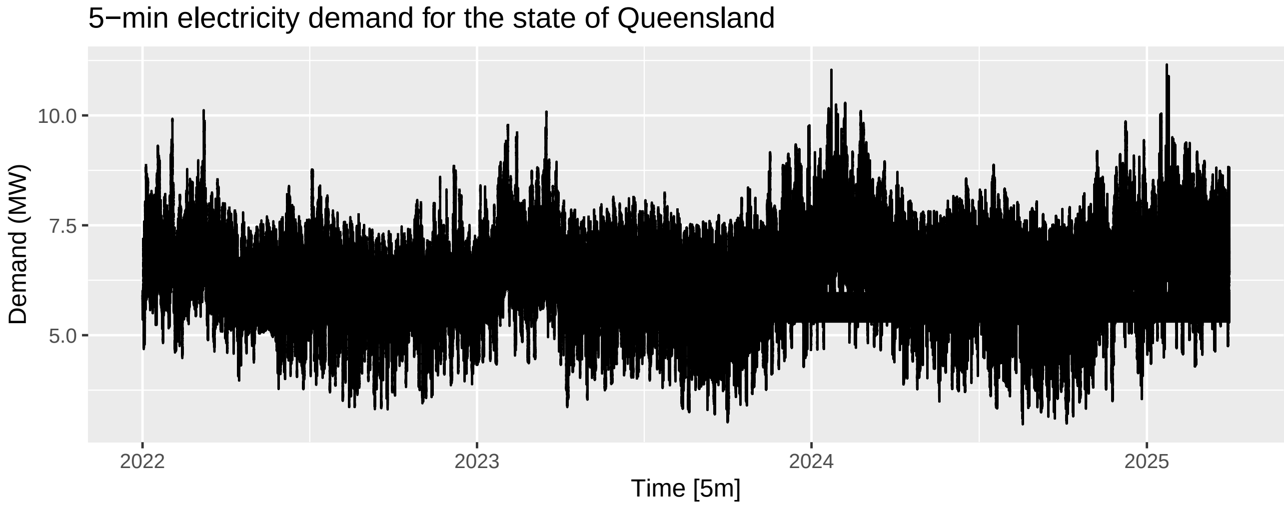

Figure 1

```
qld_data |>
  gg_season(Demand, period = "day") +
  labs(
    title = "5-min electricity demand for the state of Queensland",
    y = "Demand (MW)",
    x = "Time of day"
  )
```

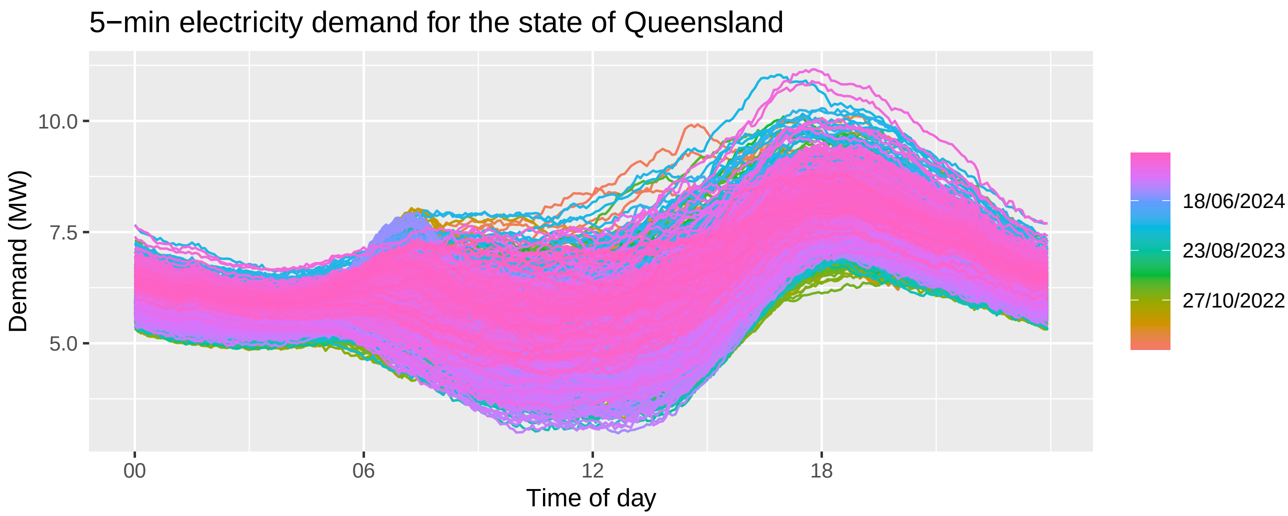

Figure 2

```r
qld_data |>
  gg_season(Demand, period = "week") +
  labs(
    title = "5-min electricity demand for the state of NSW",
    y = "Demand (MW)",
    x = "Time of week"
  )
```

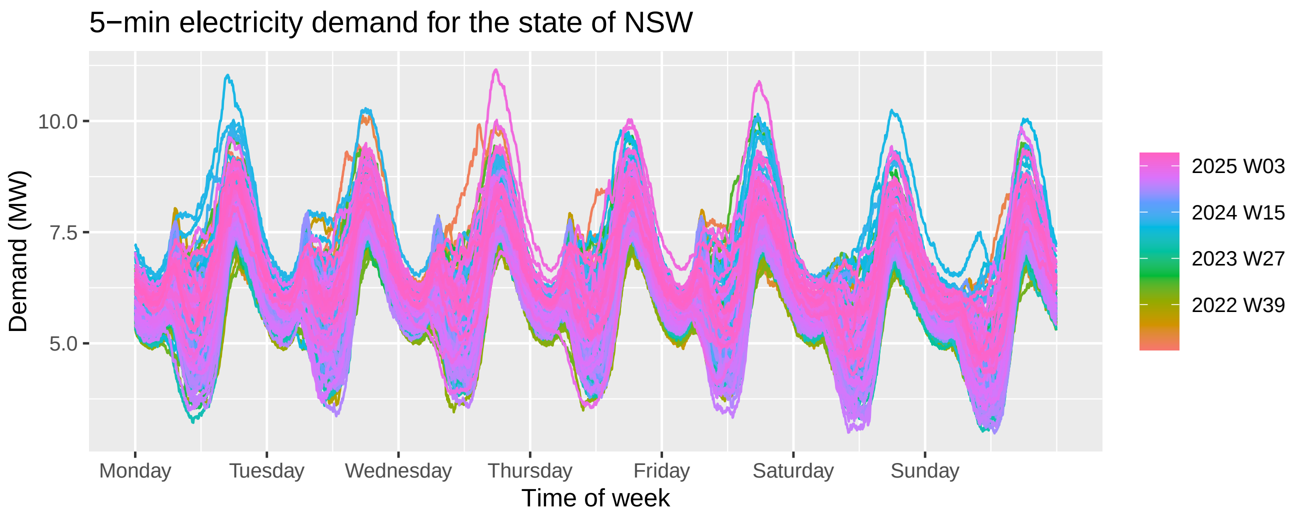

Figure 3

1. Describe Figures 1–3. Carefully comment on the interesting features of the plots.

7 marks

2. The data has been aggregated to daily. For the STL decomposition in Figure 4, briefly discuss what is shown in each panel. How has the annual seasonality been handled?

3 marks

```
qld_daily |>
  model(
    stl = STL(Demand ~ trend(window = 101) +
                season(period = "week", window = "periodic"))
  ) |>
  components() |>
  autoplot()
```

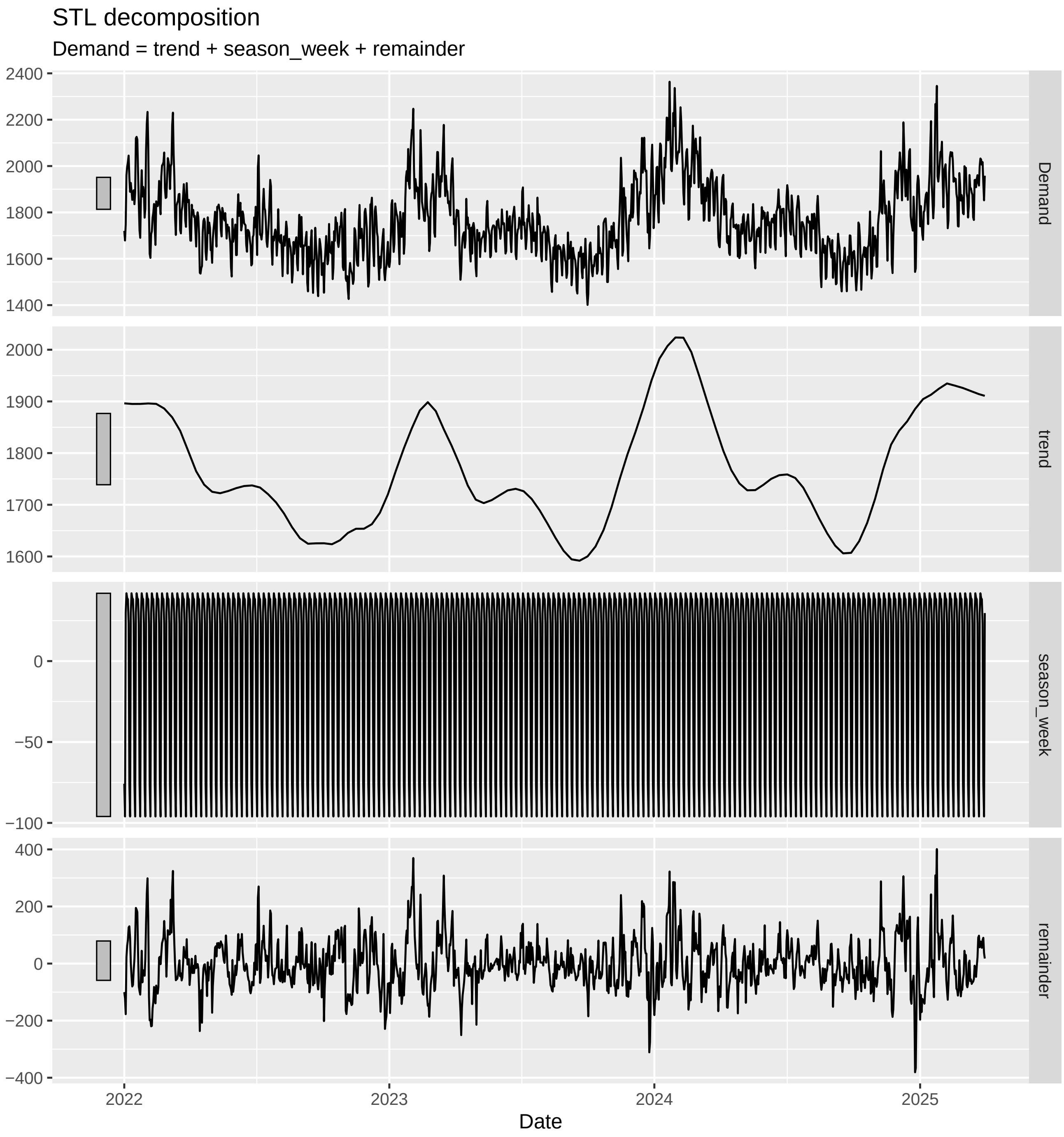

Figure 4

3. You have been asked to provide forecasts for the next two weeks using the daily data. Consider applying each of the methods and models below. Comment, in a few words each, on whether each one is appropriate for forecasting the next two weeks of data. No marks will be given for simply guessing whether a method or a model is appropriate without justifying your choice.

10 marks

Start your response by stating: **suitable** or **not suitable**.

(a) Seasonal naïve method using weekly seasonality.
(b) Naïve method.
(c) The STL decomposition shown in Figure 4, combined with an ARIMA to forecast the seasonally adjusted component, and seasonal naïve method for the seasonal component.
(d) Holt-Winters method with no trend and additive weekly seasonality.
(e) ETS(A,N,A) with annual seasonality.
(f) ETS(M,A,M) with weekly seasonality.
(g) ARIMA(2,0,1)(1,0,0)$_{52}$.
(h) ARIMA(1,1,2)(2,0,0)$_{7}$.
(i) Dynamic regression with Fourier terms for the annual seasonality and a seasonal ARIMA model to handle the weekly seasonality and other dynamics.
(j) Dynamic regression with Fourier terms for the weekly seasonality and a seasonal ARIMA model to handle the annual seasonality and other dynamics.

Total: 20 marks

SECTION C

You build a model to forecast the daily electricity demand from Figure 4 with an ETS model. The estimated model is shown below.

```
fit_ETS <- qld_daily |>
  model(ets = ETS(Demand))
report(fit_ETS)
```

```
Series: Demand
Model: ETS(M,Ad,M)
  Smoothing parameters:
    alpha = 0.979
    beta  = 0.000581
    gamma = 0.000208
    phi   = 0.971

  Initial states:
 l[0] b[0] s[0] s[-1] s[-2] s[-3] s[-4] s[-5] s[-6]
 1870 3.33 1.01  1.02  1.03  1.03  1.02 0.945 0.955

  sigma^2:  0.0015

  AIC  AICc   BIC
18415 18415 18481
```

1. Write down the equations for the fitted model, and explain how the output above relates to the model parameters.

6 marks

2. Plots of the residuals are shown in Figure 5. Discuss what these tell you about the model.

4 marks

```r
gg_tsresiduals(fit_ETS)
```

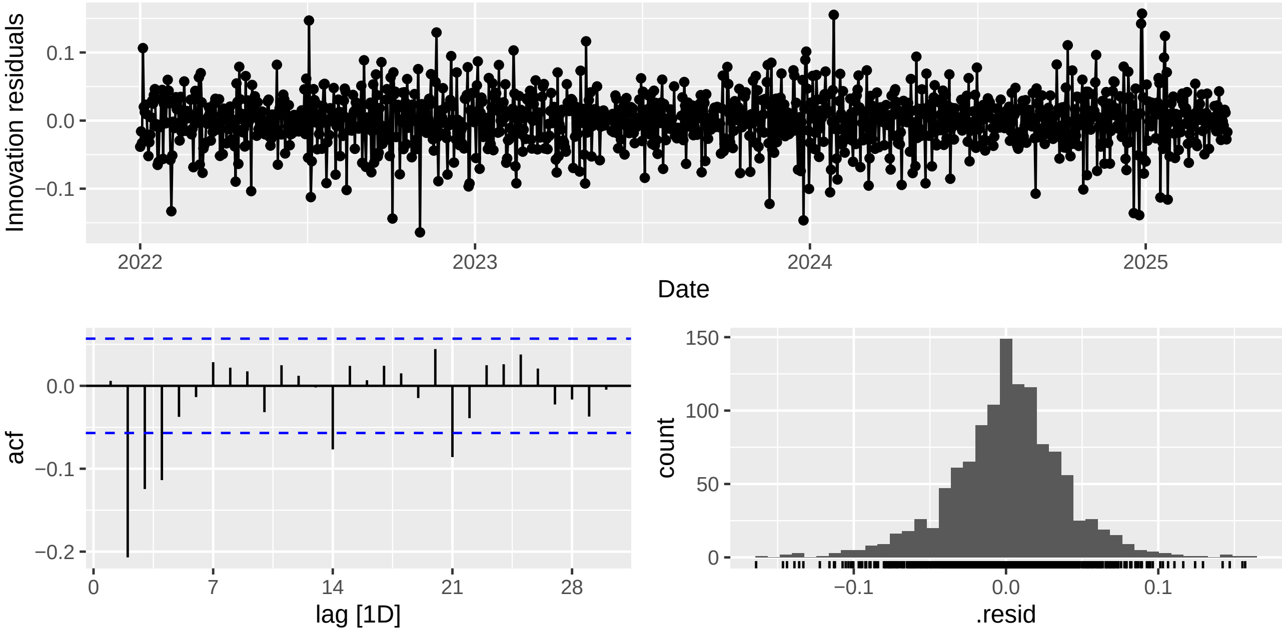

Figure 5:

3. If a Ljung-Box test was conducted on the residuals from this model, what would you expect the p-value to be? Justify your answer.

2 marks

4. The last few values of the model states are shown below. Use these, along with the model output shown earlier, to compute the point forecast for the next day.

6 marks

```r
fit_ETS |> components() |> tail(10)
```

```
# A dable: 10 x 7 [1D]
# Key:     .model [1]
# :        Demand = (lag(level, 1) + 0.971097463322547 * lag(slope, 1)) *
#   lag(season, 7) * (1 + remainder)
   .model Date       Demand level   slope season remainder
   <chr>  <date>      <dbl> <dbl>   <dbl>  <dbl>     <dbl>
 1 ets    2025-03-22  1930. 2019. 0.0716   0.955   0.0429
 2 ets    2025-03-23  1916. 2028. 0.0747   0.945   0.00441
 3 ets    2025-03-24  2000. 1971. 0.0390   1.02   -0.0284
 4 ets    2025-03-25  2032. 1980. 0.0429   1.03    0.00436
 5 ets    2025-03-26  2028. 1977. 0.0401   1.03   -0.00134
 6 ets    2025-03-27  2006. 1963. 0.0305   1.02   -0.00735
 7 ets    2025-03-28  2017. 1994. 0.0476   1.01    0.0158
 8 ets    2025-03-29  1926. 2016. 0.0597   0.955   0.0116
 9 ets    2025-03-30  1852. 1962. 0.0257   0.945  -0.0276
10 ets    2025-03-31  1958. 1930. 0.00577  1.02   -0.0168
```

5. Forecasts and prediction intervals for the next two weeks are shown in Figure 6. How many of the 80% intervals do you expect to contain the actual demand values?

2 marks

```
fit_ETS |>
  forecast(h = "2 weeks") |>
  autoplot(qld_daily |> filter(year(Date) >= 2025))
```

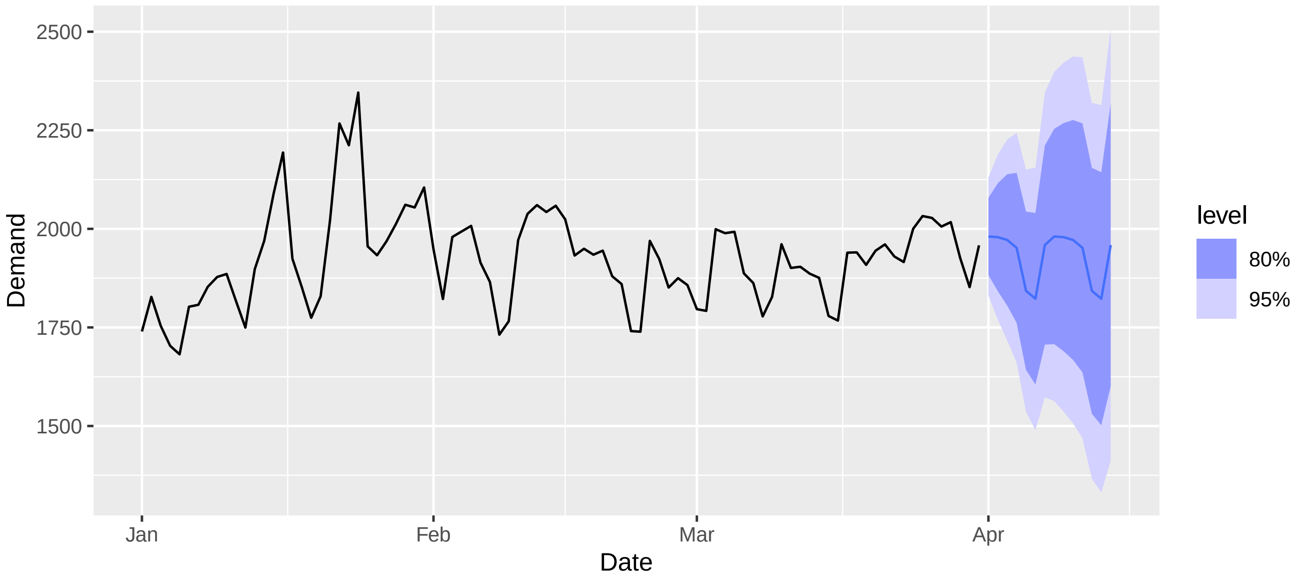

Figure 6:

Total: 20 marks

SECTION D

An ARIMA model is fitted to the daily electricity demand shown in Figure 4.

```
fit_ARIMA <- qld_daily |>
  model(ARIMA(Demand))
report(fit_ARIMA)
```

```
Series: Demand
Model: ARIMA(1,1,2)(2,0,0)[7]

Coefficients:
         ar1      ma1      ma2    sar1    sar2
      0.5458  -0.6655  -0.2877  0.3508  0.2071
s.e.  0.0411   0.0462   0.0399  0.0295  0.0293

sigma^2 estimated as 5442:  log likelihood=-6777
AIC=13566   AICc=13566   BIC=13596
```

1. Write down the equations for the model using backshift notation, including specifying the values for all model parameters.

   5 marks

2. Explain how the model above has been selected.

   2 marks

3. Using the plots in Figure 7 choose an alternative model? Justify your choices.

   6 marks

```
qld_daily |>
  gg_tsdisplay(difference(Demand, 7), plot_type = "partial")
```

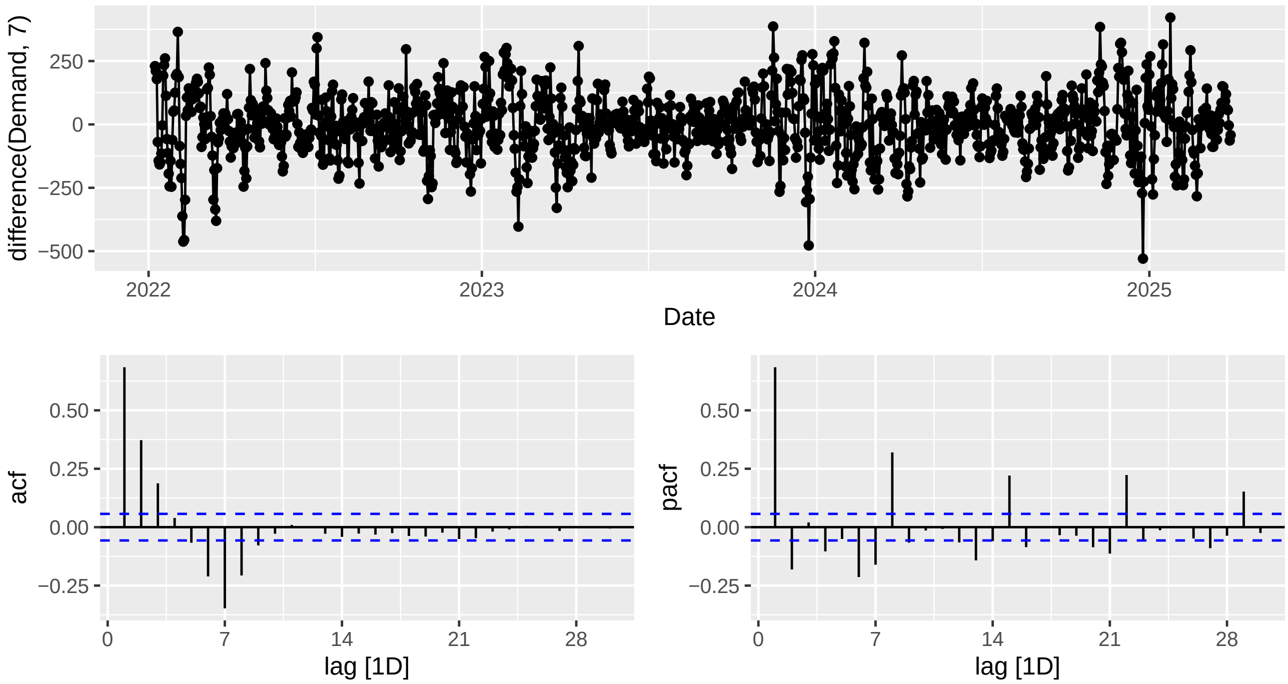

Figure 7:

4. Comment on the residuals shown in Figure 8. What do you conclude in terms of the fit of the model. Would forecasts generated from the model be reliable based on the residuals?

5 marks

```
fit_ARIMA |> gg_tsresiduals()
```

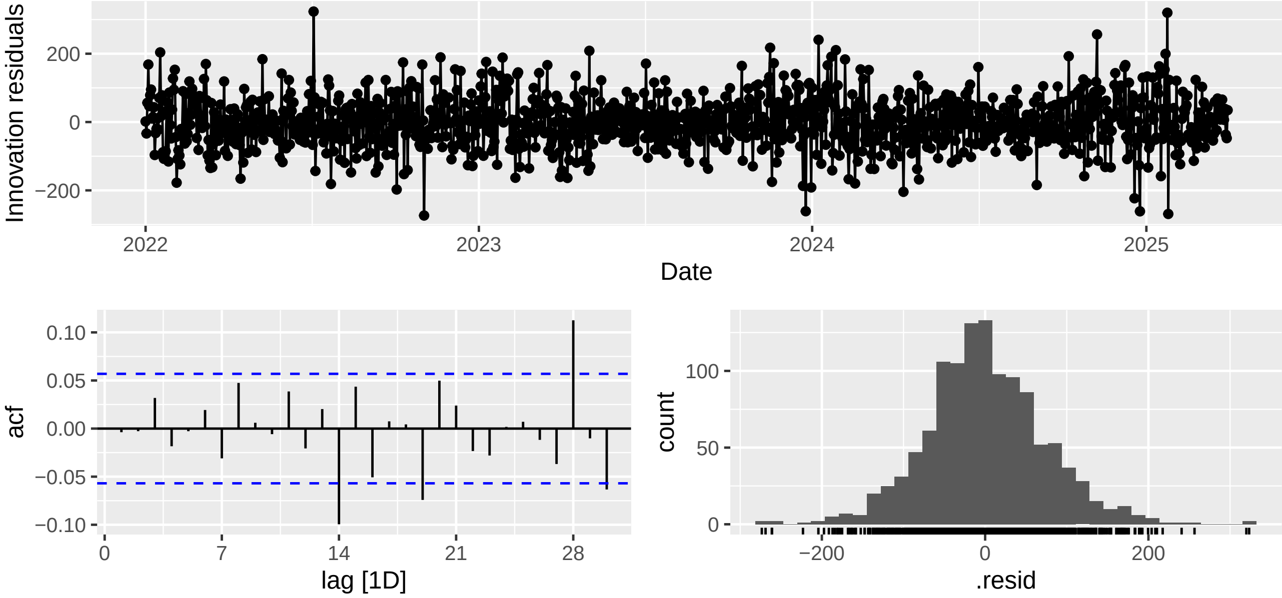

Figure 8:

5. The point forecast for the next day is 1982 MW. Compute the corresponding 95% prediction interval.

2 marks

Total: 20 marks

SECTION E

The time series shown in Figure 4 is split into a training set and a test set, where the test set comprises the last 2 weeks of available data. A dynamic regression model is fitted to training data, along with an ETS and an ARIMA model.

```
fit <- qld_daily |>
  filter(Date <= max(Date) - 14) |>
  model(
    ets = ETS(Demand),
    arima = ARIMA(Demand),
    dr = ARIMA(Demand ~ fourier("week", K = 3) + PDQ(0, 0, 0))
  )
```

1. The model details for the dynamic regression are shown below. Write down the equations for the model. There is no need to give numerical values for the coefficients.

4 marks

```
fit |>
  select(dr) |>
  report()
```

```
Series: Demand
Model: LM w/ ARIMA(1,1,2) errors

Coefficients:
        ar1      ma1      ma2  fourier("week", K = 3)C1_7
      0.508  -0.6090  -0.2463                       41.21
s.e.  0.053   0.0546   0.0369                        3.89
      fourier("week", K = 3)S1_7  fourier("week", K = 3)C2_7
                          -52.27                        3.98
s.e.                        3.89                        2.05
      fourier("week", K = 3)S2_7  fourier("week", K = 3)C3_7
                           38.19                       -6.85
s.e.                        2.05                        1.32
      fourier("week", K = 3)S3_7
                          -10.74
s.e.                        1.32

sigma^2 estimated as 4335:  log likelihood=-6561
AIC=13141   AICc=13141   BIC=13192
```

2. Which coefficients in the dynamic regression model relate to weekly seasonality? What do the remaining coefficients handle?

3 marks

3. We use all three models to forecast the next 2 weeks of data, and compare the results against the actual values.

```
fc <- fit |> forecast(h = "2 weeks")
accuracy(fc, qld_daily)
```

```
# A tibble: 3 x 10
  .model .type    ME  RMSE   MAE   MPE  MAPE  MASE RMSSE  ACF1
  <chr>  <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 arima  Test   73.4  86.6  75.8  3.70  3.83 0.817 0.707 0.592
2 dr     Test   68.6  83.5  74.1  3.48  3.77 0.799 0.681 0.657
3 ets    Test   49.0  67.9  59.2  2.49  3.02 0.639 0.554 0.618
```

Which model produces the most accurate forecasts for the test period? Justify your answer using the forecast accuracy metrics. Then, compare this result with the residual diagnostics presented in Sections C and D. How do you explain the difference between forecast accuracy and model fit?

3 marks

4. How would you extend each of the three models to also handle the annual seasonality?

4 marks

5. Energy experts tell you that electricity demand in Queensland increases on hot and humid days due to air-conditioning. Explain how you could use this information to improve your model. What complications would this add to the forecasting process?

6 marks

Total: 20 marks