Quick summary

Pablo Paulsen

06/12/2021

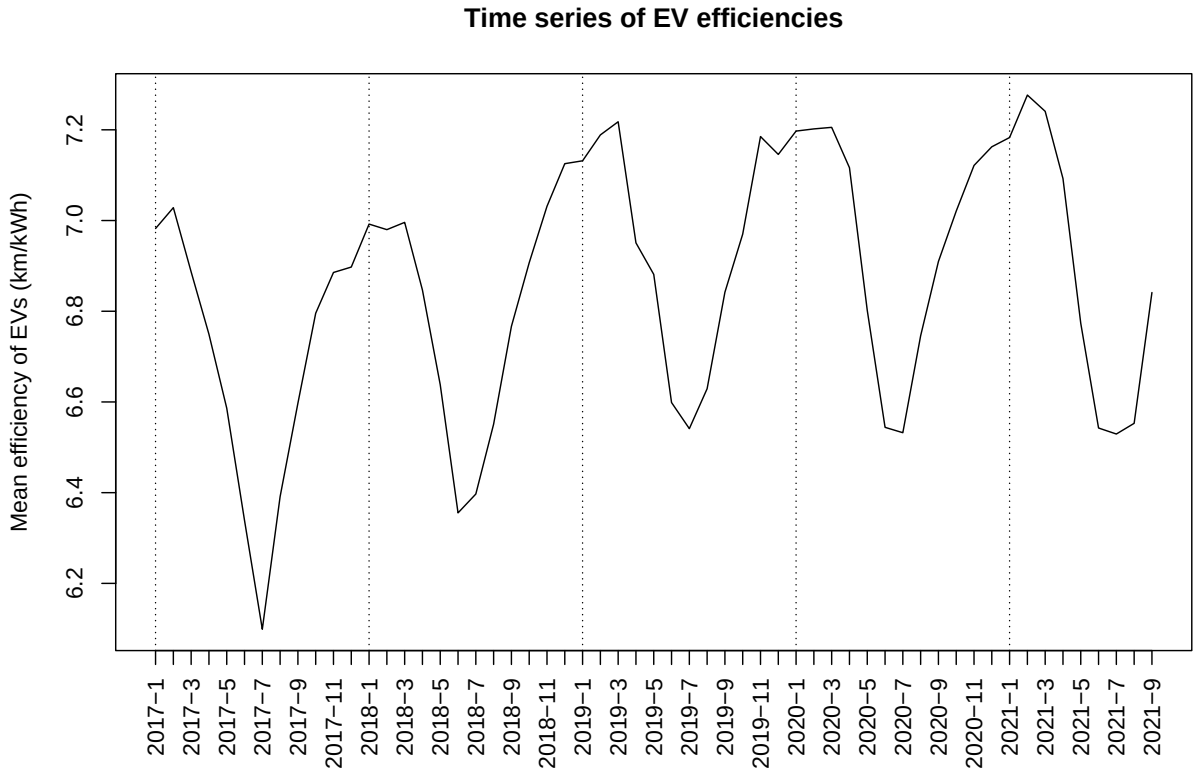
Load in and set up all of the data. Data loaded in includes

- Flip the fleet EV monitoring data with PHEVs removed. Includes monthly stats on the EV including distance traveled, efficiency and region of the vehicle. also only using data from 2017 as 2015 only included data from 5 cars and by the end of 2016 only included 23 cars.
- Weather data from
  - Auckland
  - Upper Hutt
  - Christchurch
  - Dunedin
  - Hamilton
  - Rotorua
  - Nelson
  - Palmerston North
  - Stratford
  - Napier
  - Invercargill

Using the weather data I calculated the regional average temperature for the month, heating degree days and cooling degree days using base temperature of 16 and 22 respectively. Base temperature is based around what is comfortable for most people as (https://www.geotab.com/blog/ev-range/) research shows that a majority of the seasonal variation in EV efficiency is due to cabin temperature control. The base temps could be changed slightly or possibly even use cross validation to find the ideal for best model fit.

I then added the calculated monthly weather statistics by region to the monthly EV data based on the regions of vehicle. Assuming that vehicle stays in it own region for a majority of the time.

I also created a quick monthly average for all of NZ and also by region of the EV statistics.



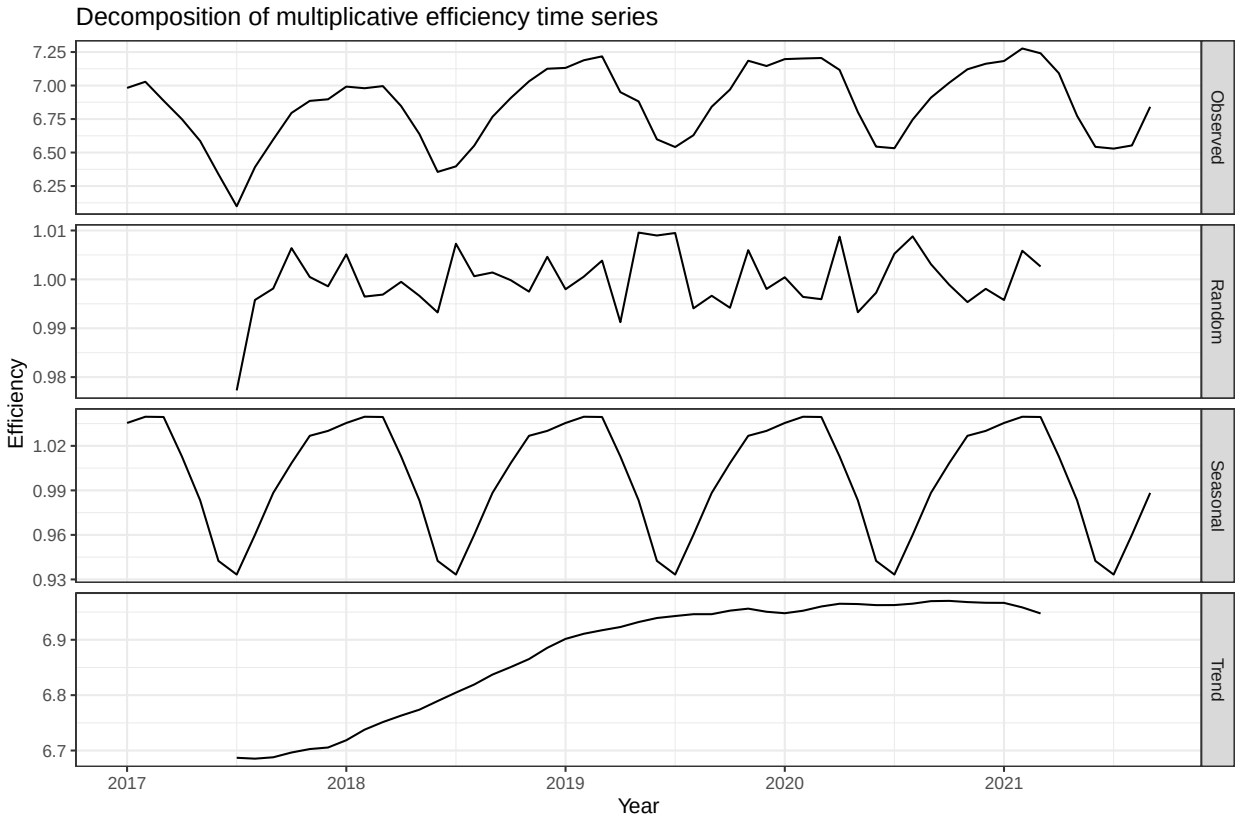
Plotting monthly average efficiency for all of NZ we can see that there is a very clear seasonal trend.

Used 2 different methods or decomposition of the seasonal trend of efficiency.

- Linear model with each month as an independent factor
  - offers more control and flexibility (could add vehicle type etc in further analysis)
  - shows confidence interval
  - requires to define an arbitrary function that can fit to the overall trend to separate from seasonal trend
  - least squares is sensitive to single large deviation that could just be outlier (such as lockdown)
- Time series Decomposition
  - designed for time series
  - automatically finds a overall trend based on the period to isolate the seasonal trend from
  - less sensitive to a large deviation (such as lockdown) as attributed to noise compared to linear model
  - no confidence interval

In the end seems better to use Time series Decomposition for overall efficiency trend but is still useful to see from the linear model without assuming any correlation between the months it still has very strong confidence intervals (p-value $< 2^{-16}$). Could be worth doing some more in depth using linear model and modeling by car.

The decomposition shows that the seasonal trend goes from 0.93 times the efficiency in June to 1.025 the average efficiency in February in March


Decomposition of multiplicative efficiency time series
Observed
Random
Seasonal
Trend
Efficiency
Year
7.25
7.00
6.75
6.50
6.25
1.01
1.00
0.99
0.98
1.02
0.99
0.96
0.93
6.9
6.8
6.7
2017
2018
2019
2020
2021