



A systematical approach for upgrade count time series models for DOC

Dong Wang

2019-11-01

Abstract

This article reports **the progress** how count time series models has been improved for Department of Conservation. This includes:

- 1: Identify the flaws of the current models (Pauls's)
- 2: Explore possible modelling approaches
- 3: Model selections
- 4: Models comparsion between the new and current models on 700+ counters (sensors)
- 5: Recommendations and future works

Keywords: GAM models, Time Series, visitor counter, sensor data

Contents

| | |
|--|-----------|
| 1 Identify the flaws of the current models (Pauls's) | 2 |
| 1.1 GAM models in the current weekly counter reports | 2 |
| 2 Explore possible modelling approaches | 3 |
| 2.1 Endless possibilities of cadidate modelling approaches and paradigms, which one? | 3 |
| 3 Model selections | 4 |
| 4 Models comparsion between the new and current models on 700+ counters (sensors) | 5 |
| 4.1 Models comparison: a full view on 700+ counters | 5 |
| 4.2 Top 4 counter sites where Dong's outperformed Paul's | 7 |
| 5 Recommendations and future works | 9 |
| 6 Reference | 11 |
| 7 Acknowledgement | 12 |

1 Identify the flaws of the current models (Pauls's)

1.1 GAM models in the current weekly counter reports

The flaws are blablabla... but there are hopes.

2 Explore possible modelling approaches

2.1 Endless possibilities of candidate modelling approaches and paradigms, which one?

3 Model selections

4 Models comparsion between the new and current models on 700+ counters (sensors)

4.1 Models comparison: a full view on 700+ counters

For 635 visitor counter sites (excludes flow meter sensors), Dong's and Paul's models were compared head to head for each site in terms of Root Mean Squared Error (RMSE). Dong's models outperformed Paul's for 623 sites. Table 1 is the summary statistics for model performance.

Table 1: Models comparison Statistics in terms of RMSE for all counter sites

| Model | min | q25 | median | q75 | max | mean | sd |
|-------|--------|----------|----------|----------|----------|----------|----------|
| Paul | 0.7133 | 7.523775 | 16.38805 | 35.83470 | 900.5800 | 41.47800 | 82.54003 |
| Dong | 0.7049 | 7.226825 | 15.52360 | 34.33105 | 721.5588 | 37.27302 | 70.10260 |

Also see Figure 1 and Figure 2

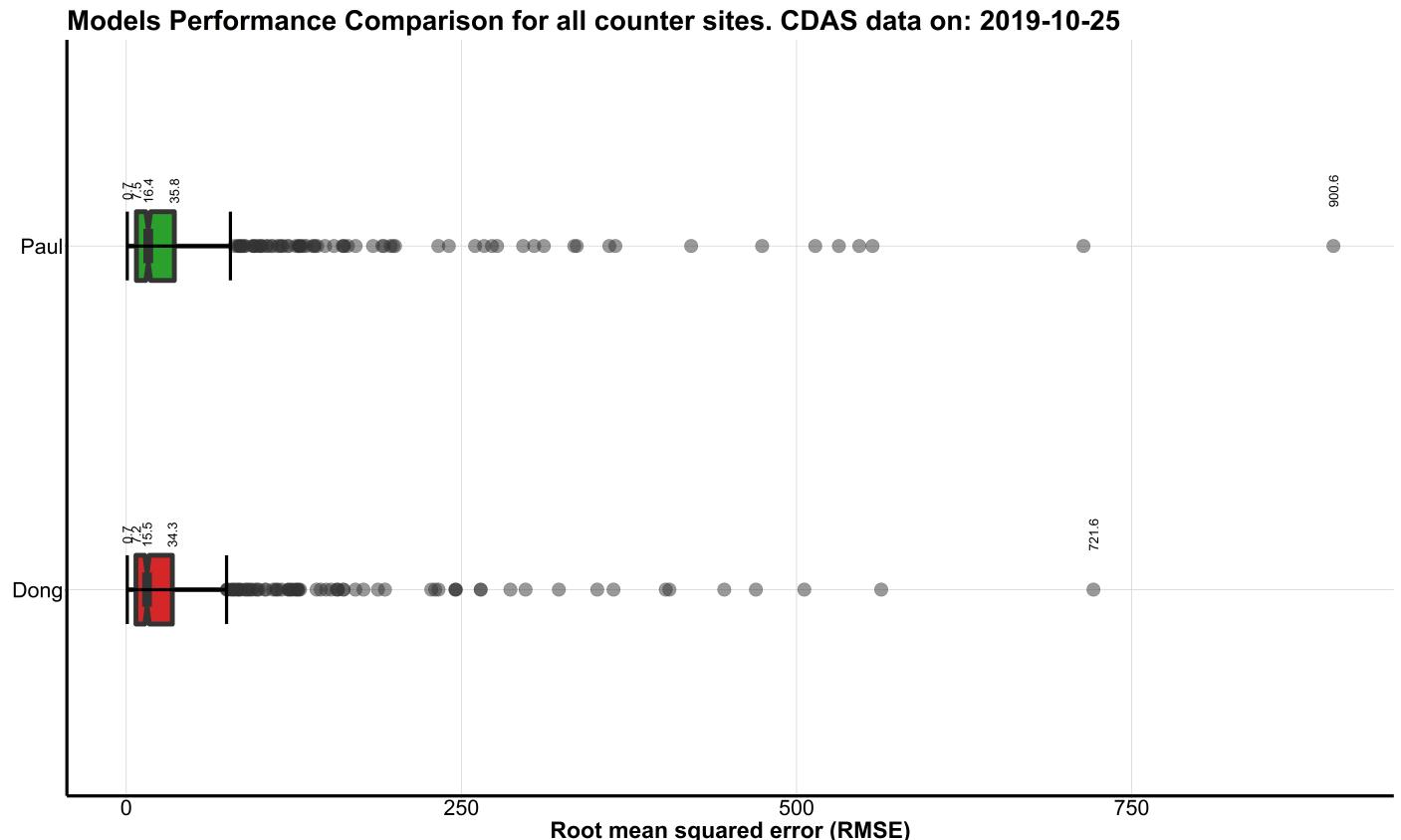


Figure 1: Models comparison in boxplots for all visitor counters

Models Performance Comparison. CDAS data on: 2019-10-25

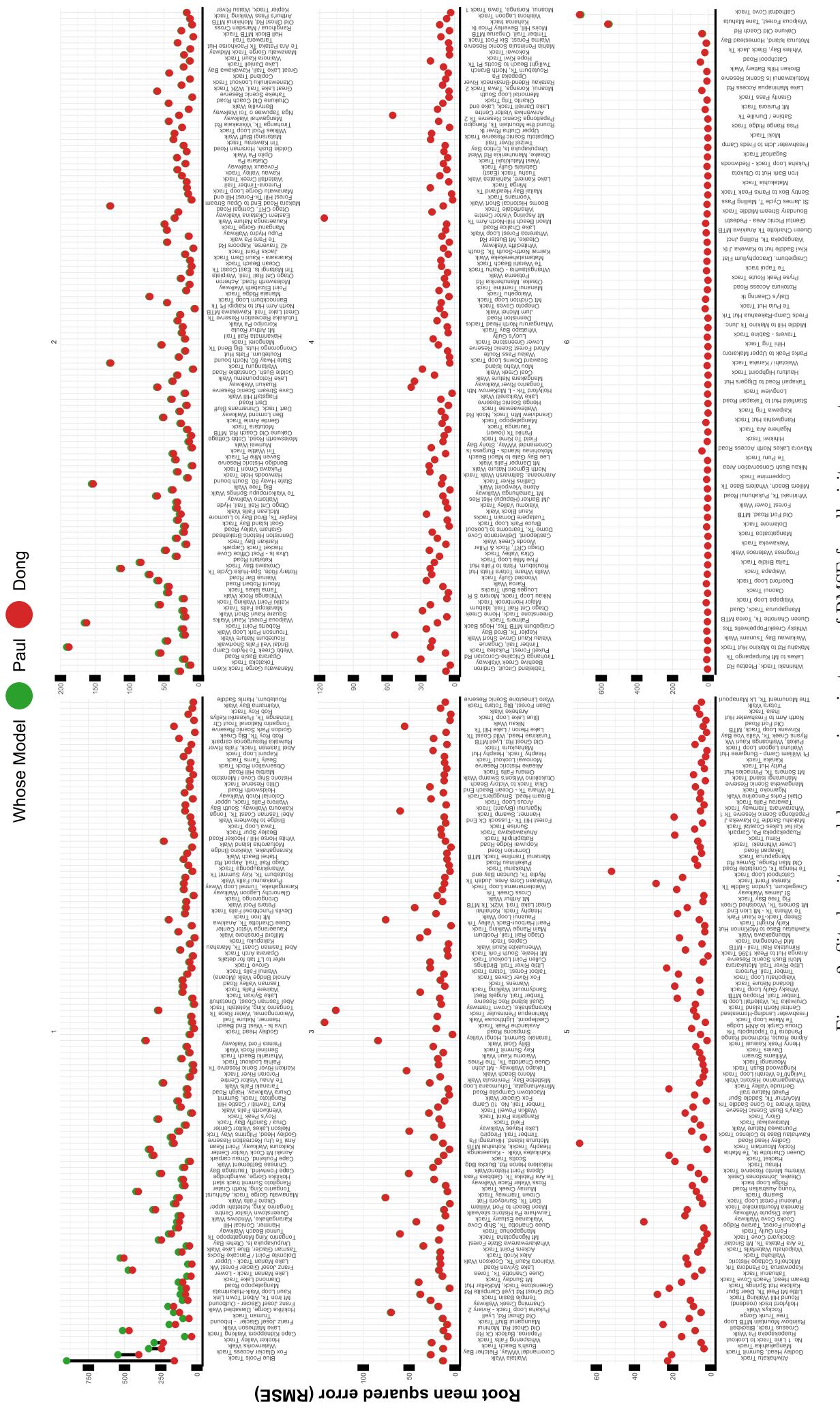


Figure 2: Site-by-site models comparison in terms of RMSE for all visitor counters

4.2 Top 4 counter sites where Dong's outperformed Paul's

Nex page is Figure 3, showing predictive performance gain

Top 4 sites: Dong' models are better than Paul's

Models Performance evaluated by Root mean squared errors. CDAS data on: 2019-10-25

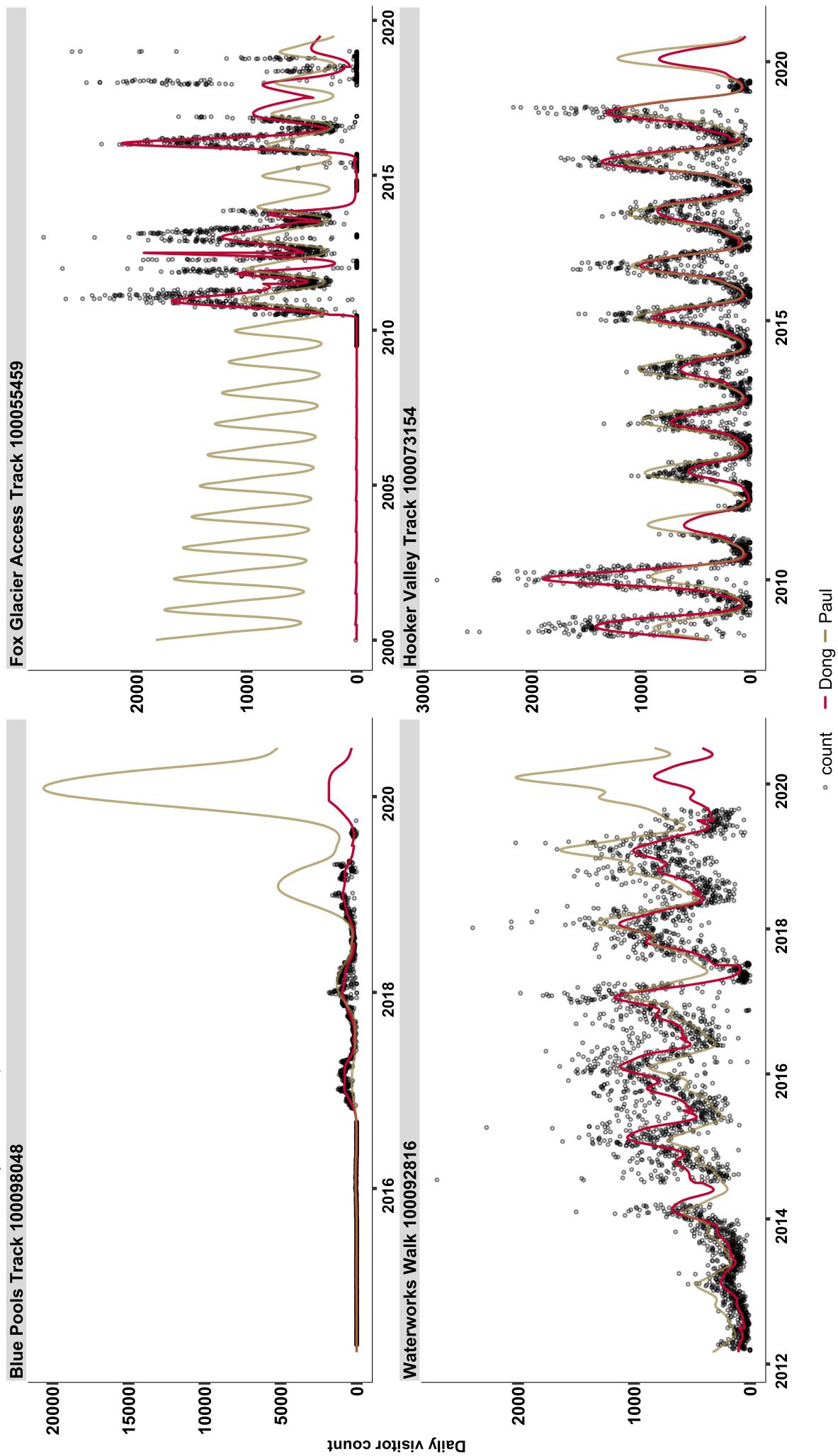


Figure 3: Top 4 counter sites where Dong's outperformed Paul's in terms of RMSE

5 Recommendations and future works

Below is a trivial dummy plot Figure 4

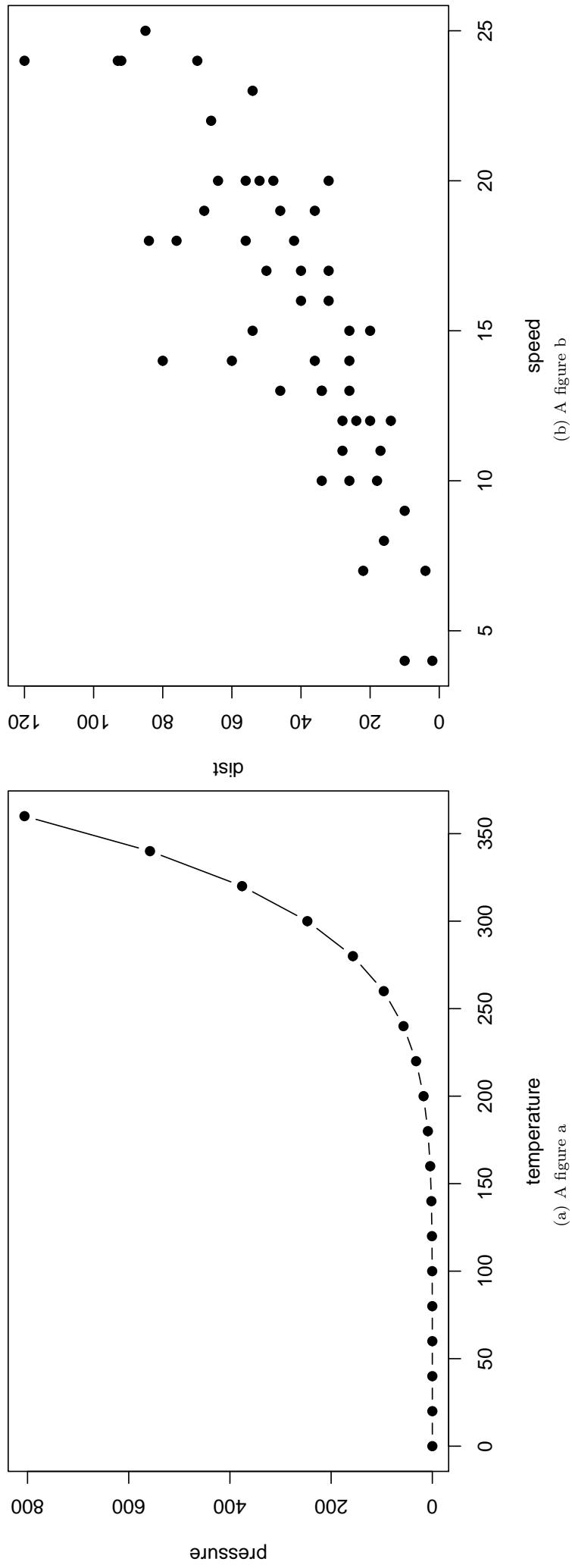


Figure 4: dummy plots

6 Reference

7 Acknowledgement