**Using Mobility Data for Meteorological Applications – Report Summary**

**Mohamed Imthiyas Abdul Rasheeth [08/09/2024]**

**Introduction:** This study examines the influence of weather variables on mobility patterns across the UK, with the goal of understanding how different meteorological conditions affect human movement. By analysing these mobility patterns, the research aims to improve early warning systems, optimize urban planning, and enhance emergency response strategies, ultimately bolstering resilience against weather-related disruptions.

**Data Sources:** The research utilizes mobility data from Vodafone, meteorological data from the ERA5 Reanalysis dataset, shapefiles for boundary mapping, and sunrise and sunset data generated using the ephem library in Python to estimate daylight duration.

**Methodology:** The analysis involved generating correlation matrices and assessing feature importance using Random Forest to identify the key weather variables that influence mobility. These top variables were then used to create derived weather metrics, which were analysed using a range of statistical and machine learning models, including Linear Regression, Log Transformation, Polynomial Regression, Generalized Additive Models (GAM), Random Forest, and Gradient Boosting. Confidence intervals were estimated to determine which weather variables had a significant impact on mobility.

**Note:** While the ERA5 dataset provides hourly meteorological data, the mobility data is time-aggregated, and the date ranges differ between the two datasets. To address this, the ERA5 data was processed to align with the aggregated mobility data. The analysis focuses on time-aggregated data for both weekends and weekdays across five diverse UK regions, as this provides a more balanced and representative trend compared to including all locations and groups.

**Results:** Random Forest and Gradient Boosting models showed the strongest performance, explaining approximately 38% of the variability, followed closely by GAM at 37%. The table below summarizes the confidence intervals (CI) for key weather variables:

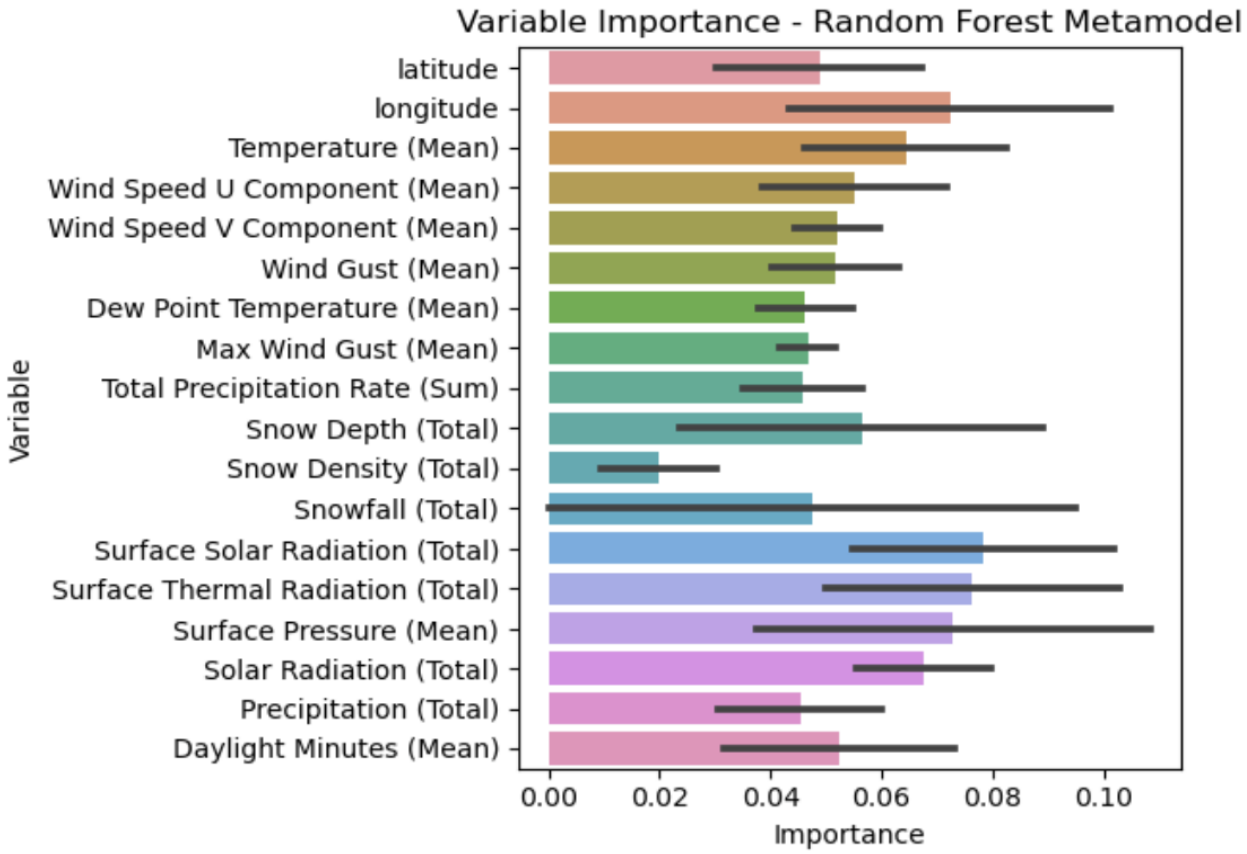
| Variable | Parameters | Weekday (CI) | Weekend (CI) |
|---|---|---|---|
| Latitude | - | [0.0117, 0.1220] | [0.0868, 0.1748] |
| Longitude | - | [0.9751, 1.14877] | [0.8556, 1.0341] |
| Clear Days | Solar Radiation > 10.2M J/m2 | [-0.1331, 0.6617] | [0.1520, 0.2356] |
| Cloudy Days | Solar Radiation < 10.2M J/m2 | [-0.1851, 0.4395] | [0.1948, 0.2903] |
| Dry Days | Daily Precipitation < 0.1mm | [0.1925, 0.2639] | [0.1856, 0.2317] |
| Wet Days | Daily Precipitation > 25mm | [0.1882, 0.2557] | [0.2145, 0.2375] |
| Comfortable Days | Temperature: 18C to 22C | [0.0426, 0.4480] | [0.2122, 0.2640] |
| Uncomfortable Days | Temperature: Outside 18°C to 22°C | [0.0306, 0.6047] | [0.1742, 0.2210] |

The results indicate that weather variables significantly influence mobility, with differences observed between weekdays and weekends. Notably, clear and cloudy days do not significantly impact mobility on weekdays, though they do on weekends

**Recommendations:** The study suggests that incorporating more diverse locations and additional variables could improve the model's performance. Expanding beyond simple wet and dry day classifications increased the model's variability from 3% to 38%. Introducing more time granularity could provide deeper insights into mobility patterns. Additionally, amending the way unique visitor counts are measured—by capturing visit frequency—could further enhance the accuracy of mobility analyses.

Figures and relevant plots are included on the following page for further review.

**Feature Importance using Random Forest:**




**Comparative Plots for December:**