Portland Traffic Analysis: Final Report

Xu, Suru
suru@pdx.edu

Nelson, Austen
ajn6@pdx.edu

February 21, 2025

Contents

1 Executive Summary

2 Introduction

3 Background

4 Data

The data we analyzed consists of vehicle counts over hourly time periods at 12 different highway locations, each consisting of 2 opposite directions (either NB/SB or EB/WB). Some locations have many more years available than others but in total there are about 3.2 million hours of data recorded over the 24 location and direction combinations. For a visual of the data coverage for each location / direction combination, see figure 1.

Many trends and area specific patterns can be easily identified by simply looking at the median values in the data. Table 1 presents an approximation to weekday peak-time and peak-volume based on cubic periodic spline interpolation of hourly medians.

General visible trends seem to be a slightly later peak traffic time in the mornings and location dependent changes in afternoon peak times. The peak volume is lower post-Covid all across the board, with very few exceptions outside of data availability issues. Lastly, looking at the occasional near-hour shift for some peak times (see Vista Ridge EB 2018, Wilsonville SB 2024, and Interstate Bridge NB 2024) indicates that there are potentially issues with Day Light Savings inconsistencies in the dataset. Duplicated hours (UTC time / no timezone) in the raw data additionally supports this hypothesis; so large changes in peak times may not be real in some cases.

5 Methods

5.1 Generalized Additive Model (GAM)

We have a dataset $D := \{x_i, y_i\}_{i=1}^n$, where x_i is an hour in the day and y_i is the observed number of cars in that hour. As expected, the correlation between the traffic volume and the time of the day is strong, with clear distinct patterns emerging for a different locations/directions and days of the week. To isolate different variables to make conclusions about changes in hourly traffic dynamics, we seek to create a simple model to summarize characteristics of interest. In particular, we are interested in how the rush hour has changed before and after the 2019 pandemic.

One of the most simple but powerful models in data science is linear regression, but as the name suggests isn't suitable for capturing non-linear relationships in data on its own. A common technique to address this constraint is to choose a suitable non-linear mapping such that the mapped inputs can be modeled with linear regression. GAMs are a class of Generalized Linear Models which provides a framework for selecting such a mapping as the sum over *smooth* functions. The smoothness of these functions is desirable for our application because traffic volume as a function of time is "smooth" and we may be interested in rate characteristics of this function which can be obtained through differentiation (since smooth functions are differentiable).

The GAM literature and software packaging calls these smooth function 'smoothers' and many different functions are provided, but we are mostly interested in splines. Spline fitting is similar to fitting polynomials to the data but are generally easier to work with for a variety of reasons (see Runge effect, sklearn example).

5.2 Maximum Mean Discrepancy (MMD)

6 Results

6.1 Generalized Additive Model (GAM)

6.2 Maximum Mean Discrepancy (MMD)

7 Conclusions

A Median Peak Time and Volumes

Table 1: Weekday Peak Time Table

Location Info			Median Peak Time				Median Peak Volume			
ID	**Name**	**Dir**	**2018**	**2019**	**2023**	**2024**	**2018**	**2019**	**2023**	**2024**
26002	Vista Ridge Tunne	EB	06:41	07:35	07:21	07:30	5062	4961	4735	4623
		WB	07:12	07:18	07:27	07:24	5669	5623	5244	5239
3016	Stafford	NB	14:02	14:05	14:08	14:11	3288	3424	3195	3315
		SB	06:20	06:18	06:52	06:52	3370	3390	3271	3342
26001	Troutdale	WB	14:34	14:48	14:48	14:54	1188	1252	1219	1246
		EB	10:34	10:26	10:34	10:03	1224	1313	1213	1282
26024	Glenn Jackson Bridge	NB	15:11	15:11	15:17	14:45	7184	7177	6781	6664
		SB	06:15	06:18	06:49	06:20	7511	7430	6334	6328
26014	Hoyt	EB	13:56	13:53	13:48	13:48	5597	5596	5307	5255
		WB	05:40	05:34	06:00	06:00	5870	5831	5443	5411
3011	Wilsonville	SB	15:03	14:51	14:40	15:40	3749	3756	3602	3621
		NB	05:57	05:54	06:06	06:03	3885	3895	3823	3924
26016	Iowa Street	NB	06:29	06:29	07:09	07:07	5707	5678	5356	5319
		SB	16:35	16:35	16:32	16:32	5391	5380	4838	4793
26022	Lents	NB	06:20	06:20	07:01	06:58	5066	5071	5198	5155
		SB	16:29	16:35	16:32	16:32	5524	5578	5287	5214
34010	Beaverton-Bethany	WB	07:15	07:15	16:24	16:24	5040	5196	4509	4552
		EB	16:03	16:03	06:49	06:52	4574	4498	4255	4283
26028	Fairview	EB	16:26	16:35	16:18	16:32	4468	4522	4375	4367
		WB	06:29	06:32	06:41	06:35	4069	4106	3939	3834
26004	Interstate Bridge	NB	16:26	16:12	16:03	16:52	4636	4807	4637	4383
		SB	06:29	05:34	06:09	06:00	1906	5647	5262	5021
34007	North Plains	WB	16:03	16:03	15:52	15:58	1102	1135	1086	1137
		EB	06:35	06:38	06:49	06:52	927	950	809	823

Hourly median traffic volume for each entry (location, direction, year) for weekdays is interpolated with a cubic spline to find approximate peak traffic time.

B Data Availability

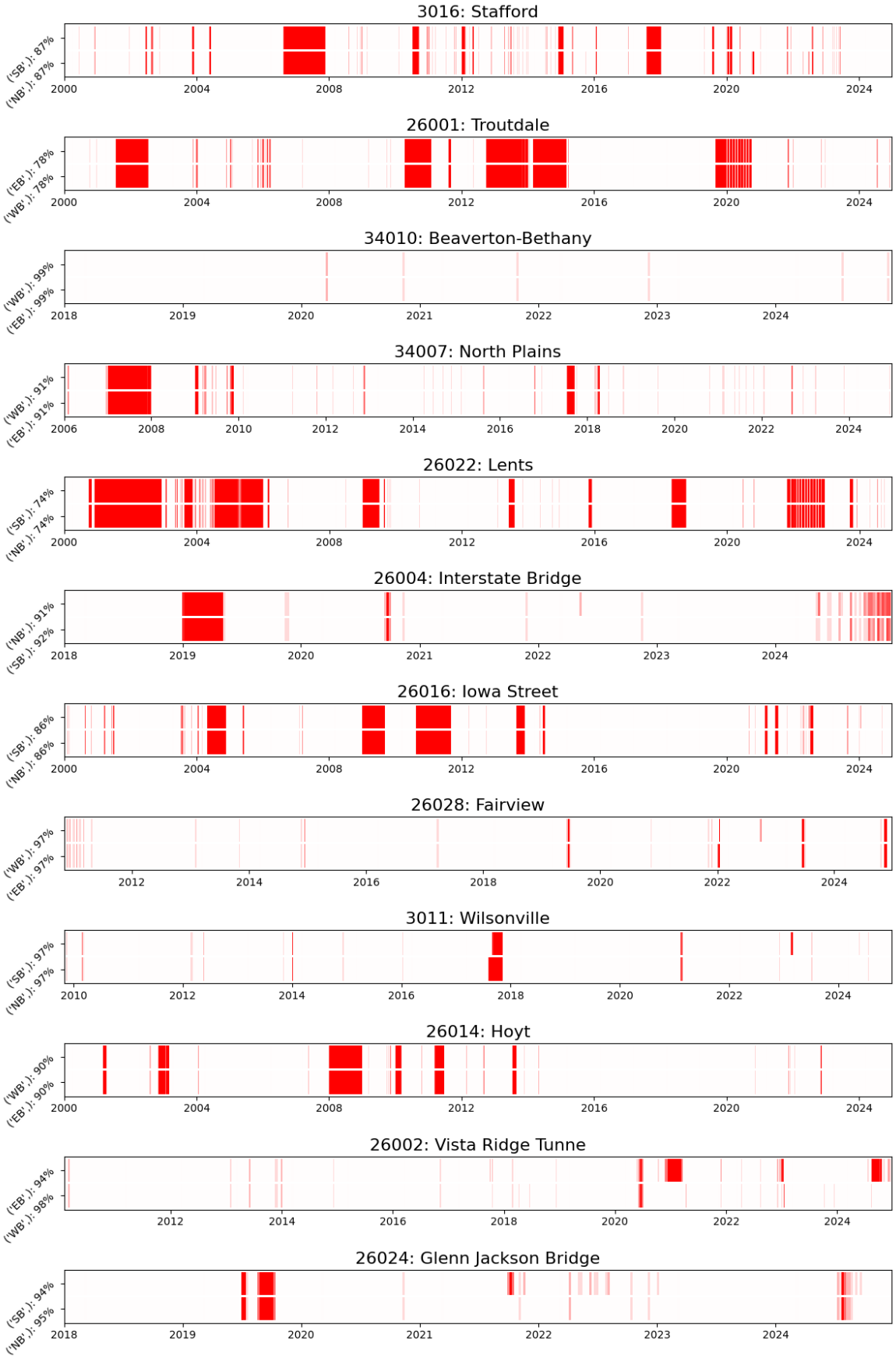
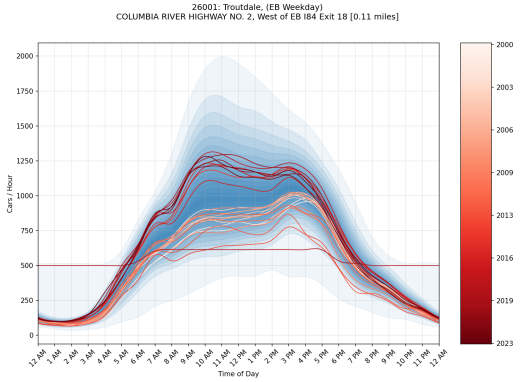


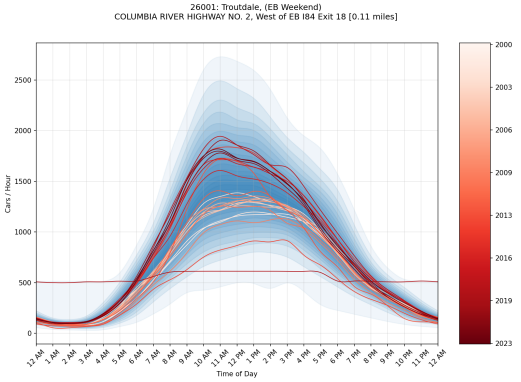
Figure 1: Red indicates missing data, opacity indicates how much of that week is missing. This doesn't account for "0" data which is technically missing and misrepresented.

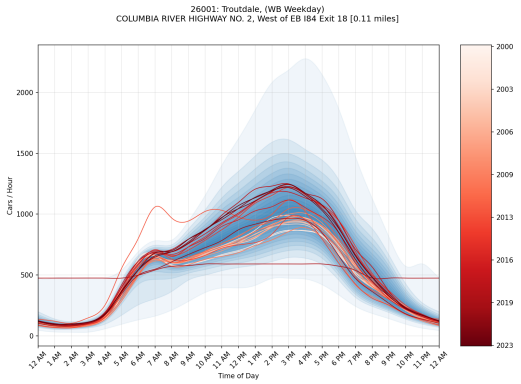
C Smooth Hourly Trends

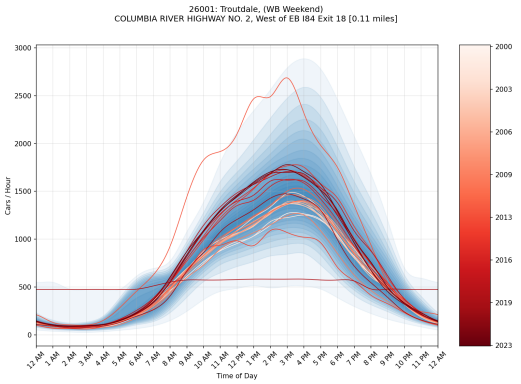
These plots provide a visual for similar data presented in table 1 but includes every year available in the dataset with quantile shading. Each line is the periodic cubic spline interpolation of the hourly medians/quantiles for each (location, direction, weekday/weekend) tuple. Some strange data issues become apparent in these plots, but they are almost entirely only in years we aren't concerned with in this analysis.

C.0.1 Troutdale

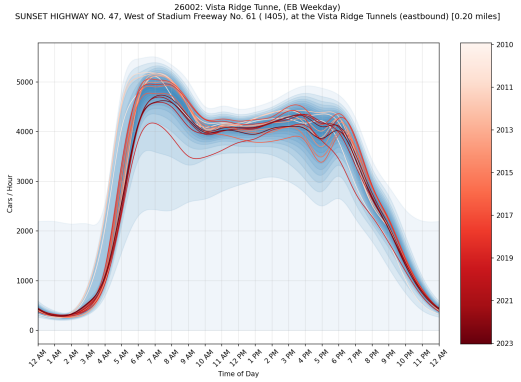


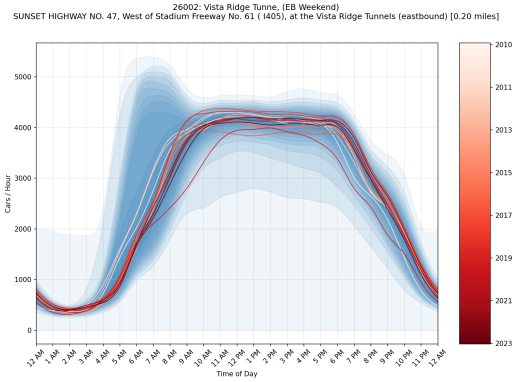


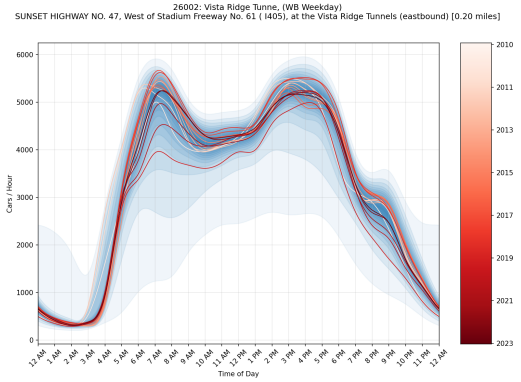


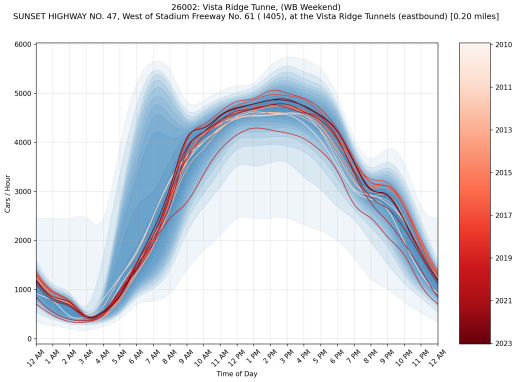


C.0.2 Vista Ridge Tunnel

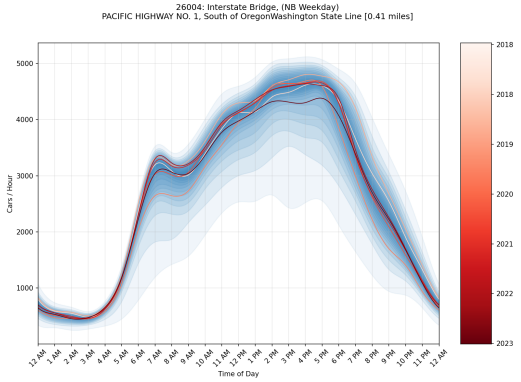


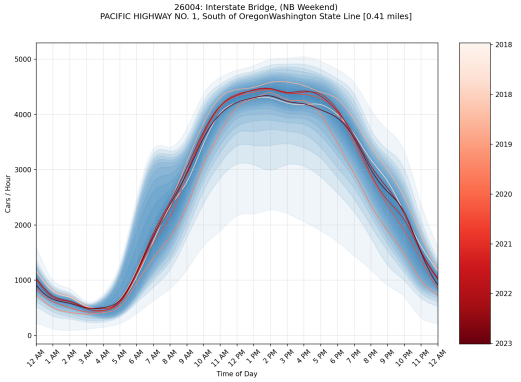


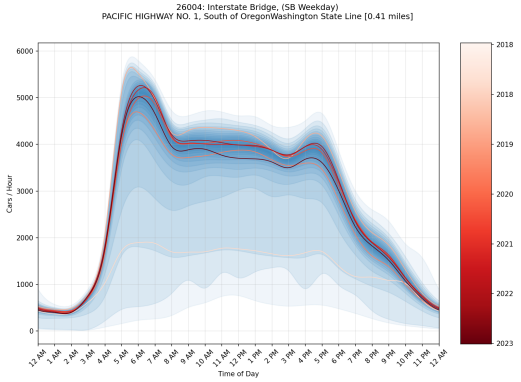


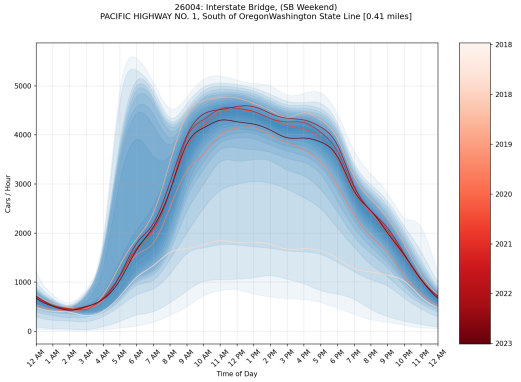


C.0.3 I5 Bridge

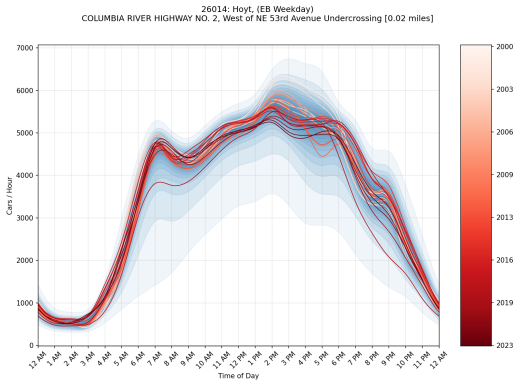


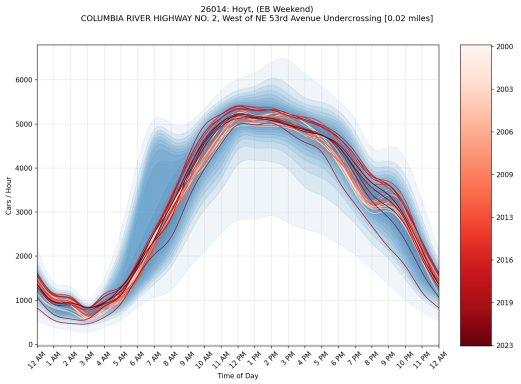


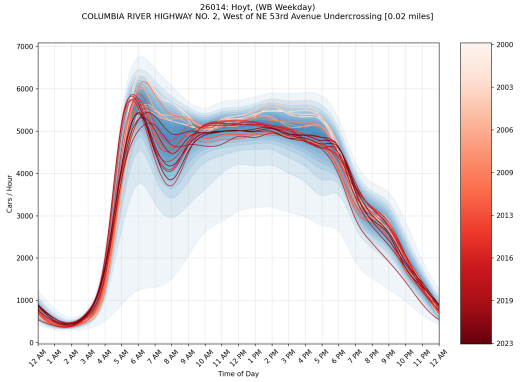


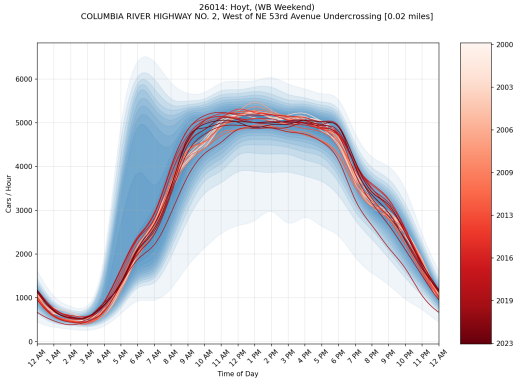


C.0.4 Hoyt

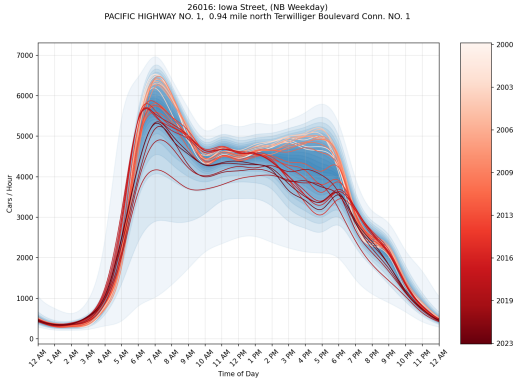


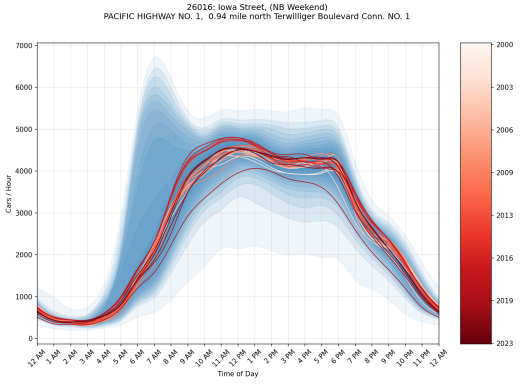


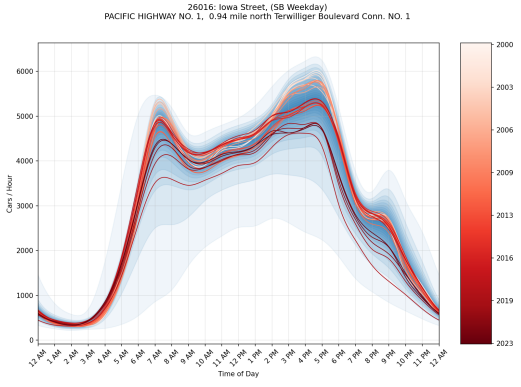


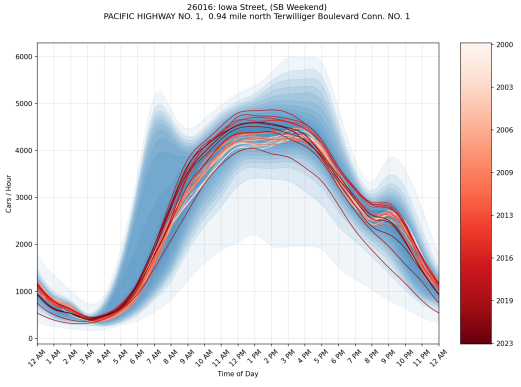


C.0.5 Iowa Street

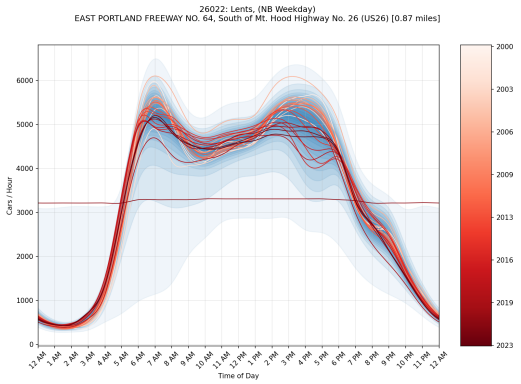


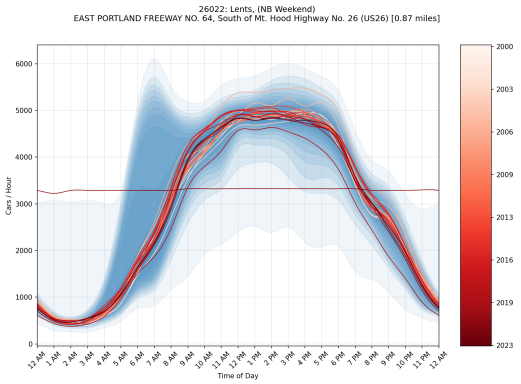


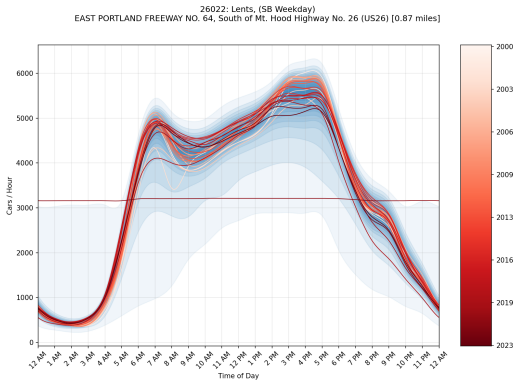


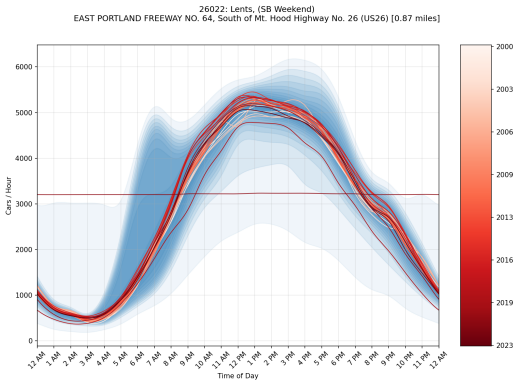


C.0.6 Lents

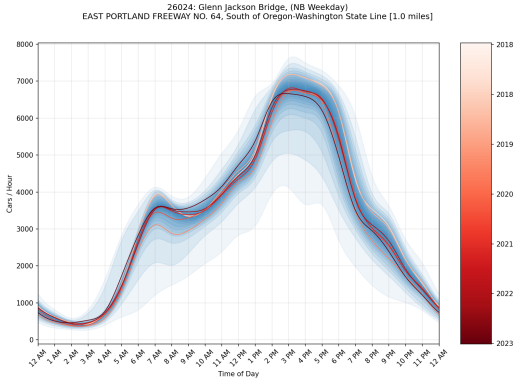


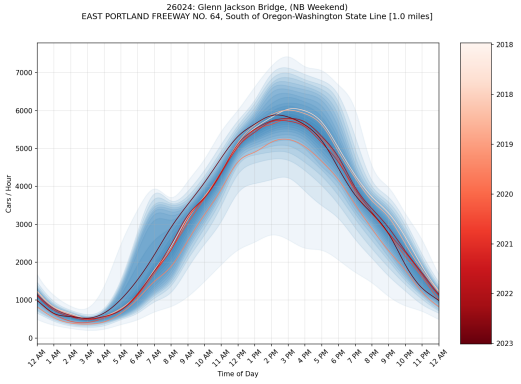


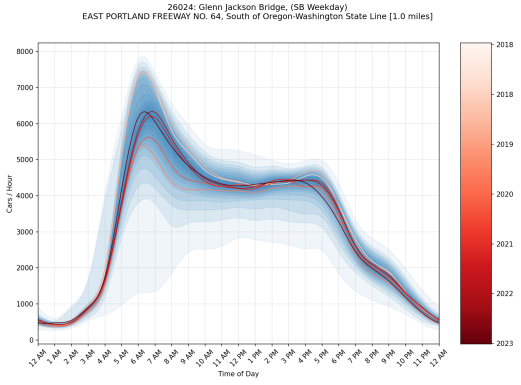


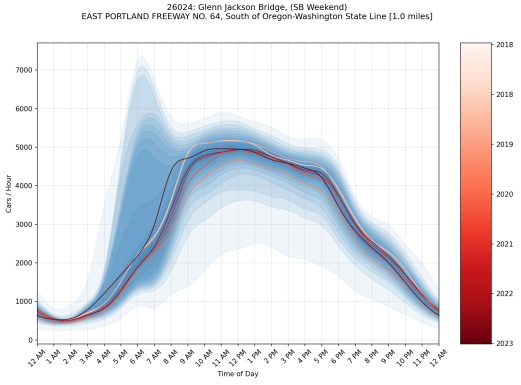


C.0.7 Glenn Jackson Bridge

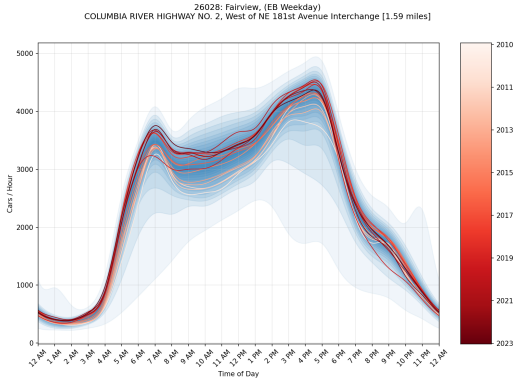


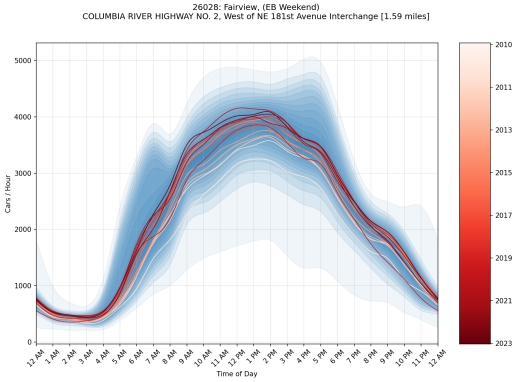


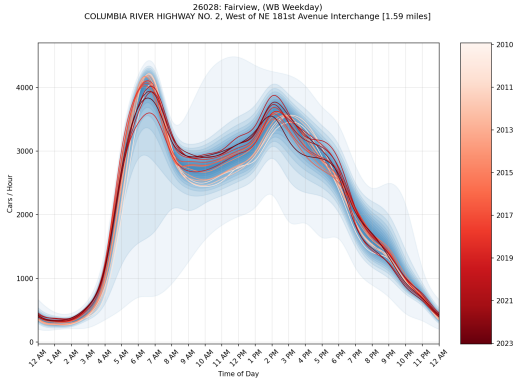


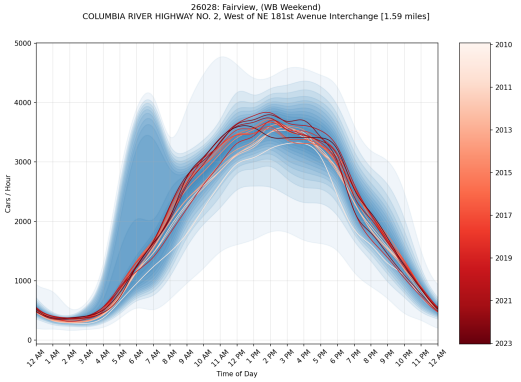


C.0.8 Fairview

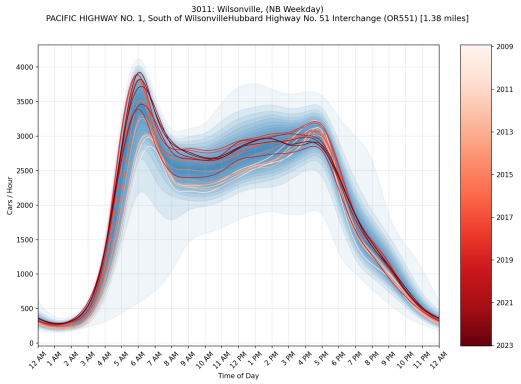


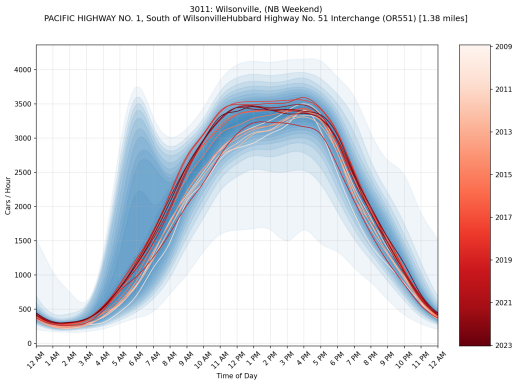


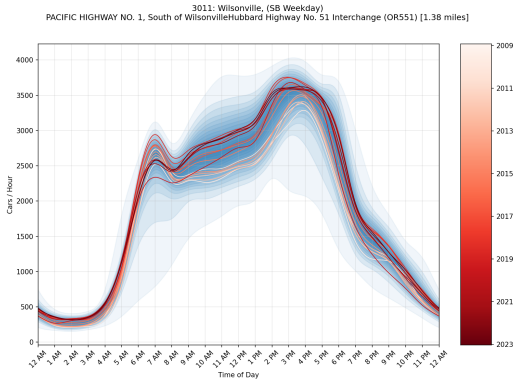


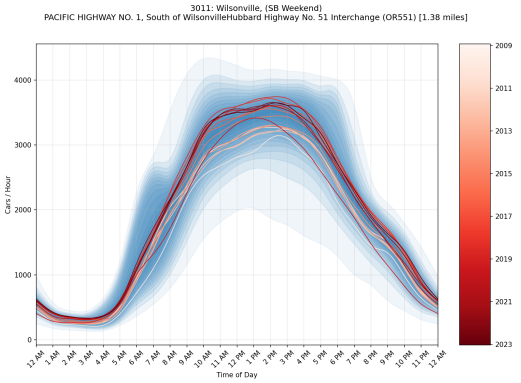


C.0.9 Wilsonville

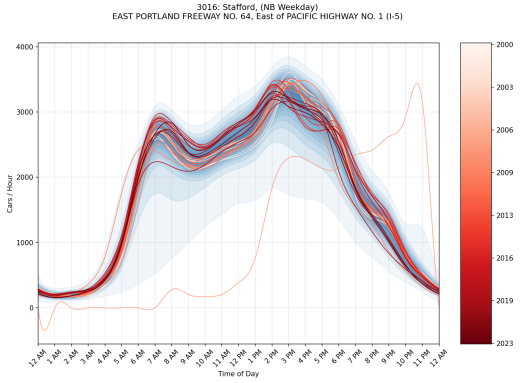


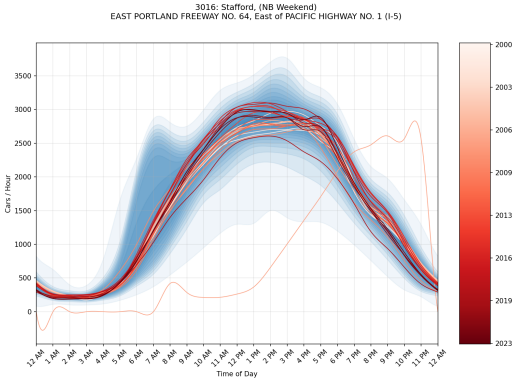


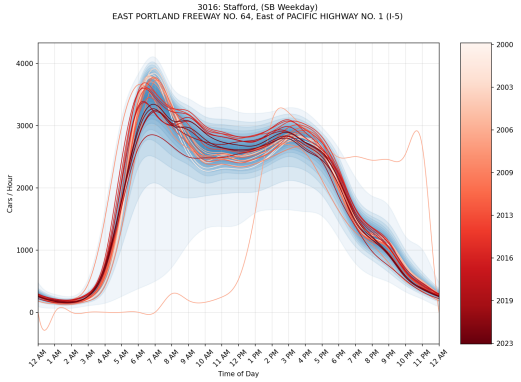


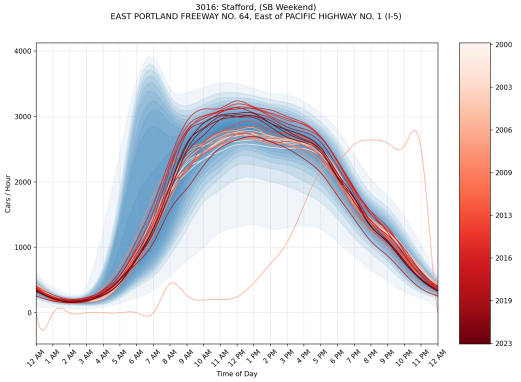


C.0.10 Stafford

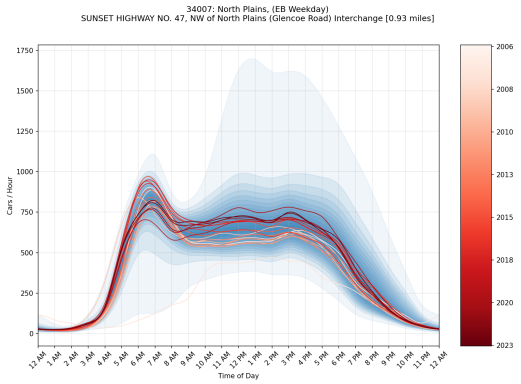


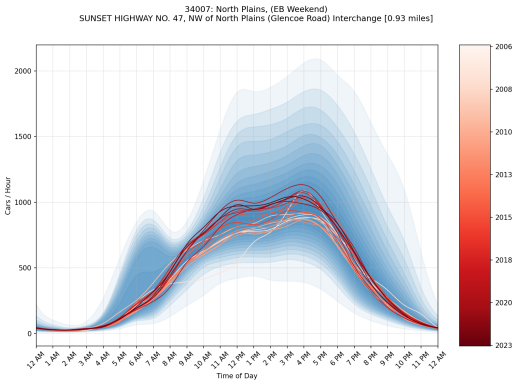


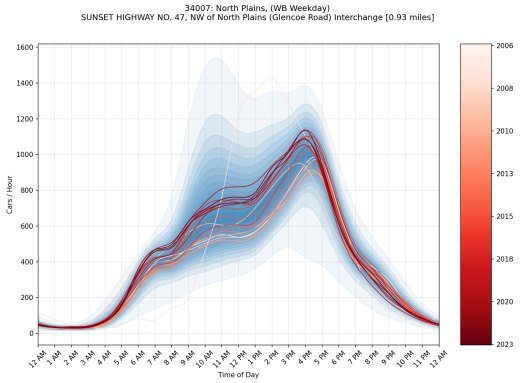


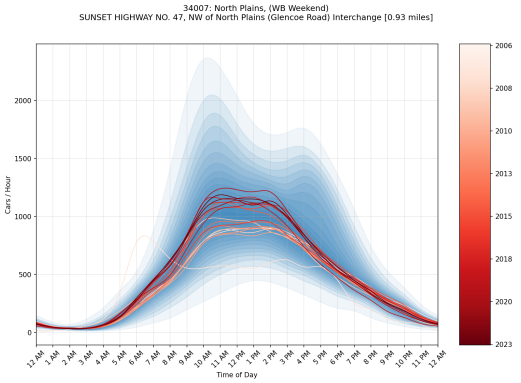


C.0.11 North Plains









C.0.12 Beaverton Bethany