

FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS (FPCA): ANALYSIS OF HOURLY TRAFFIC DATA ON INTERSTATE-5, BETWEEN ROUTES 14 AND 99

JAN DE LEEUW, IRINA KUKUYEVA

ABSTRACT. The authors examine traffic on a section of the Interstate-5 using the Functional Principal Component Analysis. The results show that if we only use the first three principal components, our results model the original data closely.

1. DATA OVERVIEW

The traffic data is obtained from loop counts on the Interstate-5 (I-5) freeway, between Routes 14 and 99. It is hourly data over 11 years, with a total of 96,432 observations for each direction.

1.1. North Direction. Missing Percentages:

- Route 126 - 8.82%
- Hungry Valley - 30.1%
- Wheeler Ridge - 29.9%
- Route 14 - 70.1%

1.2. South Direction. Missing Percentages:

- Route 126 - 8.81%
- Hungry Valley - 33%
- Wheeler Ridge - 30.2%
- Route 14 - 61.1%

To use FPCA, we have organized our data to have each *Hour* as a column and each *Day* as a row of the matrix. As a result, our matrix is 4018x24. Then, we limit ourselves to only those days that have 5 or fewer observations. This is done to ensure convergence within a reasonable amount of time.

Date: March 31, 2008.

Key words and phrases. Functional Principal Component Analysis, Traffic, Interstate-5.

2. NORTH DIRECTION

2.1. Fitted Data after FPCA. The fitted data after applying the FPCA is presented below for each intersection. (Please see previous research reports for plots of original data.)

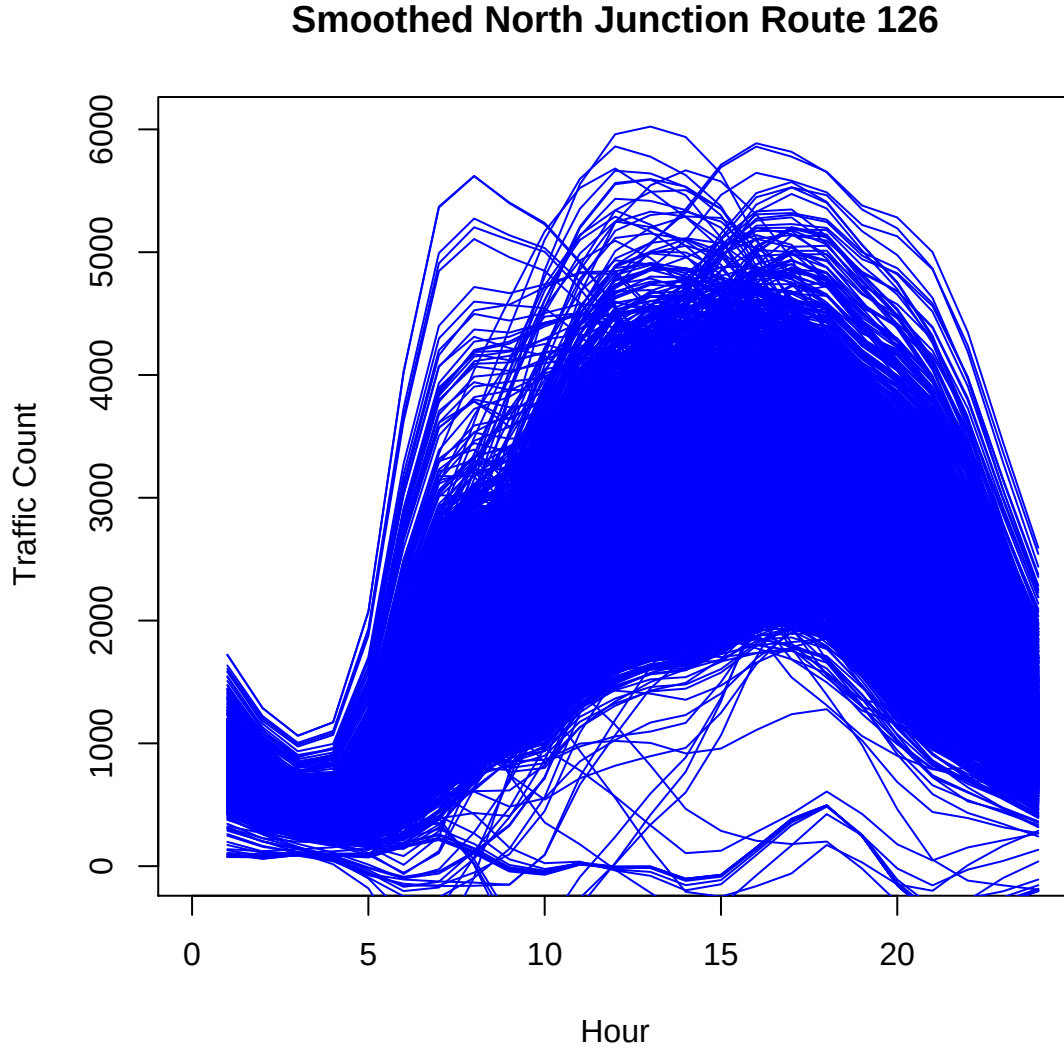
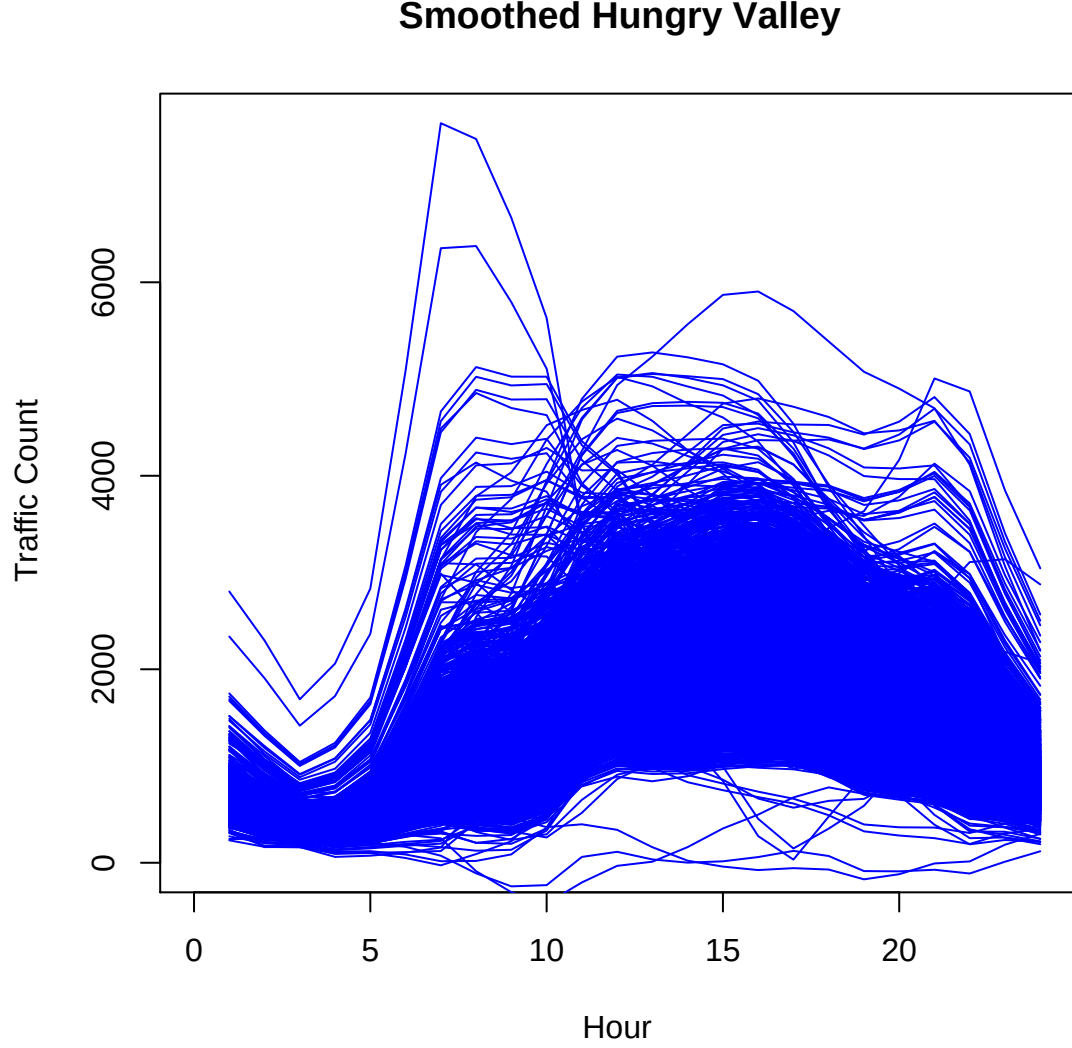


FIGURE 1. Route 126 Fitted Data

2.1.1. *Comparison of the Fitted Data.* Legend: Red is the fitted data and the original data is in blue.

We can see that when we filter out for missing data, our fit is not perfect. We will need to smooth the proposed function.

2.1.2. *Eigenvalues for Each Intersection.* Table 1 contains eigenvalues for each intersection, based on complete data only.



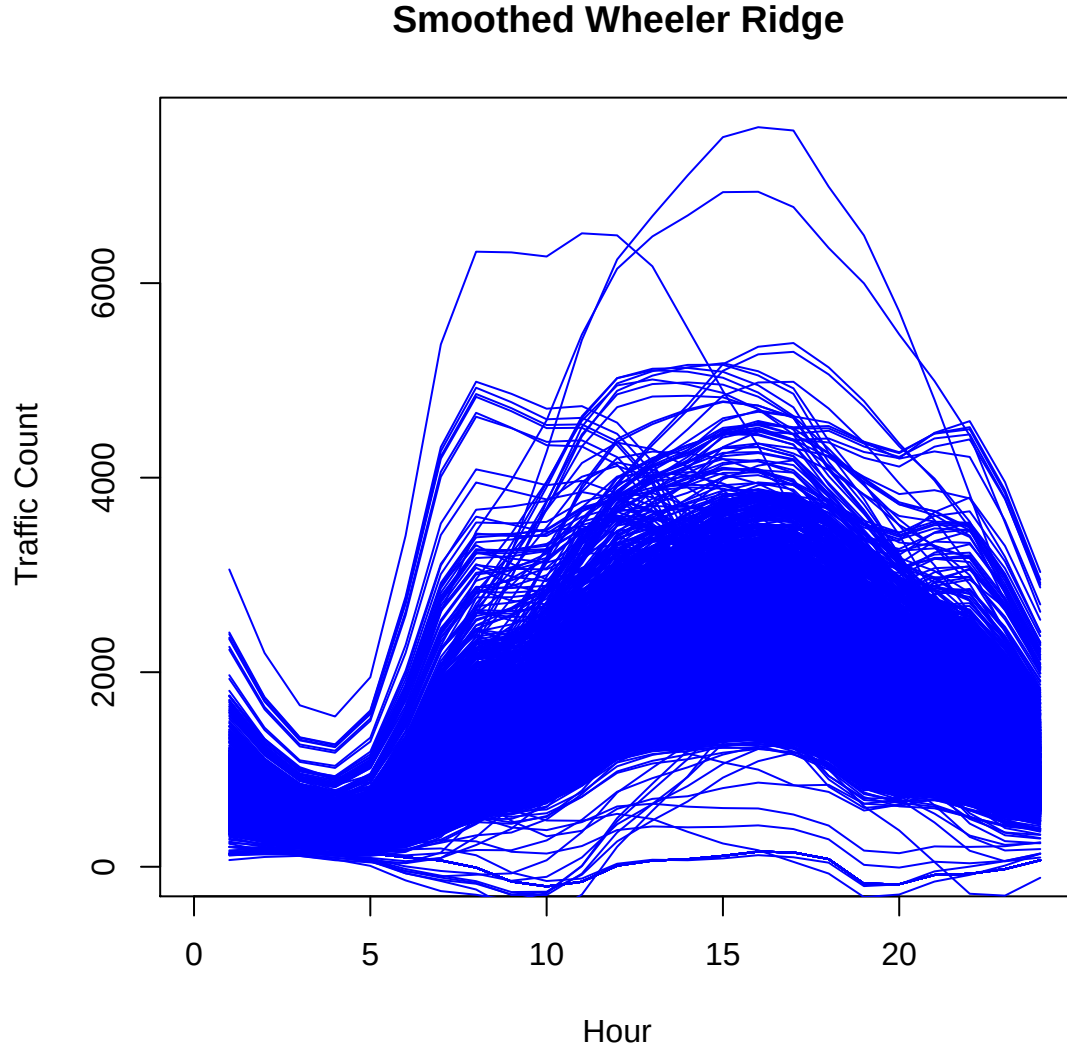
n_{fitted}.pdf

FIGURE 2. Hungry Valley Fitted Data

Table 2 contains percentage of total eigenvalues for each intersection, based on complete data only.

As a result, we can see that if we are to use the first three principal components, we would explain 91% of the variability observed in Route 126 data, 87% in Hungry Valley data, 93% in Wheller Ridge data and 83% in Route 14 data.

2.1.3. Plots of the Eigenvalues for Each Intersection. To verify that our choice of using the first three principal components, we present the following plots of the eigenvalues for each intersection. Since each eigenvalue corresponds to the variance explained by that component, we can further see that our choice is valid.



n_ffitted.pdf

FIGURE 3. Wheeler Ridge Fitted Data

3. SOUTH DIRECTION

3.1. Fitted Data after FPCA. The fitted data after applying the FPCA is presented below for each intersection. (Please see previous research reports for plots of original data.)

3.1.1. *Comparison of the Fitted Data.* Legend: Red is the fitted data and the original data is in blue.

We can see that when we filter out for missing data, our fit is not perfect. We will need to smooth the proposed function.

3.1.2. *Eigenvalues for Each Intersection.* Table 1 contains eigenvalues for each intersection, based on complete data only.

Smoothed North Junction Route 14

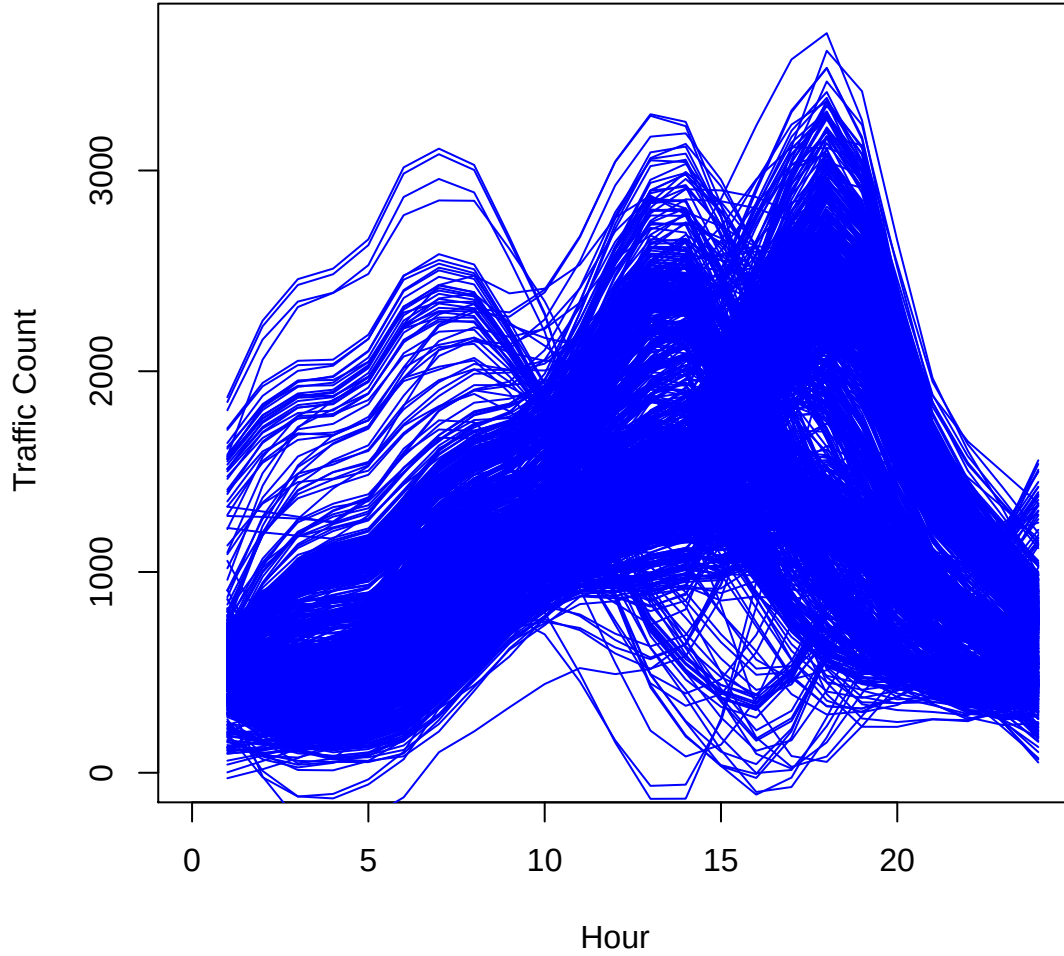


FIGURE 4. Route 14 Fitted Data

Table 2 contains percentage of total eigenvalues for each intersection, based on complete data only.

As a result, we can see that if we are to use the first three principal components, we would explain 91% of the variability observed in Route 126 data, 93% in Hungry Valley data, 96% in Wheller Ridge data and 87% in Route 14 data.

3.1.3. Plots of the Eigenvalues for Each Intersection. To verify that our choice of using the first three principal components, we present the following plots of the eigenvalues for each intersection. Since each eigenvalue corresponds to the variance explained by that component, we can further see that our choice is valid.

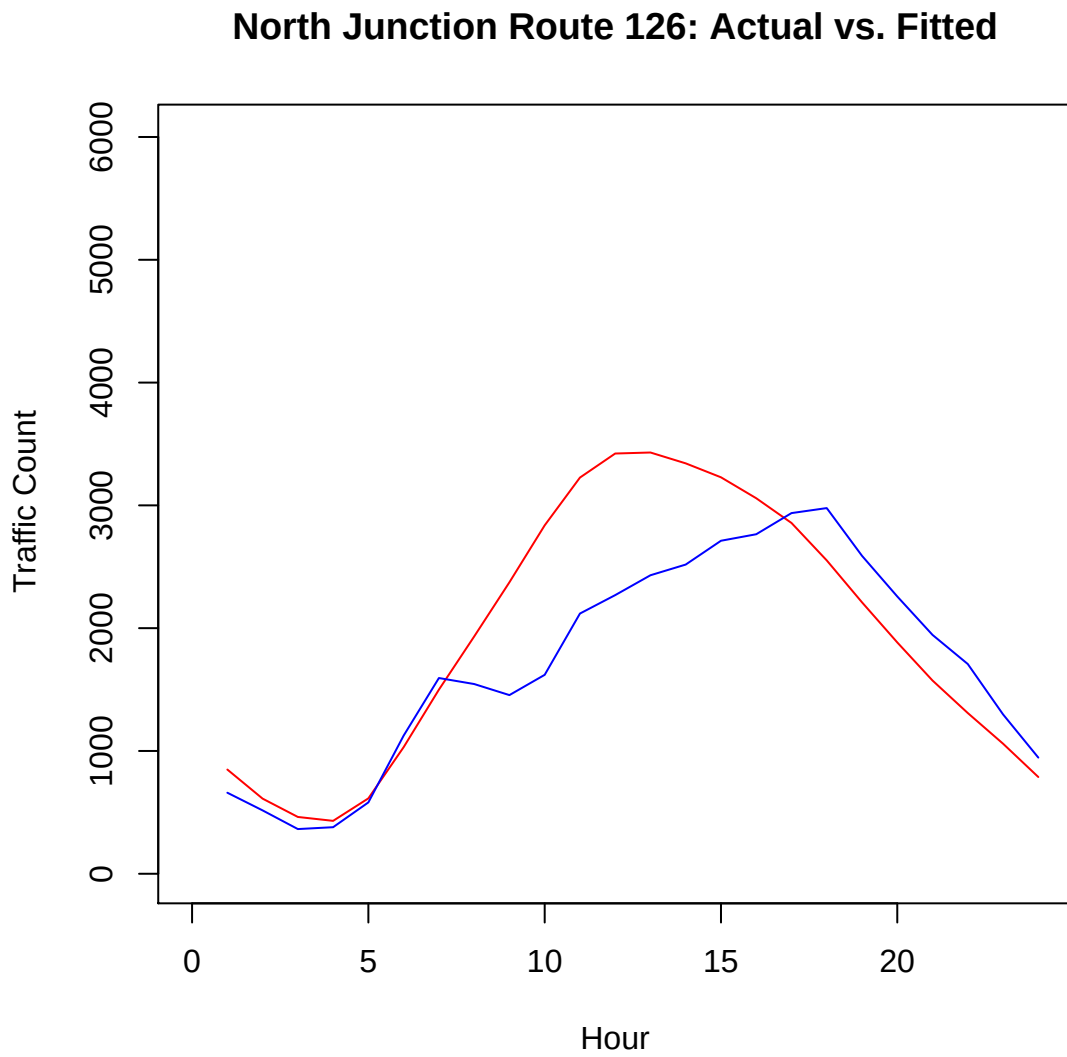


FIGURE 5. Route 126 Fitted Data

ú-

E-mail address: deleeuw@stat.ucla.edu, ikukuyeva@stat.ucla.edu

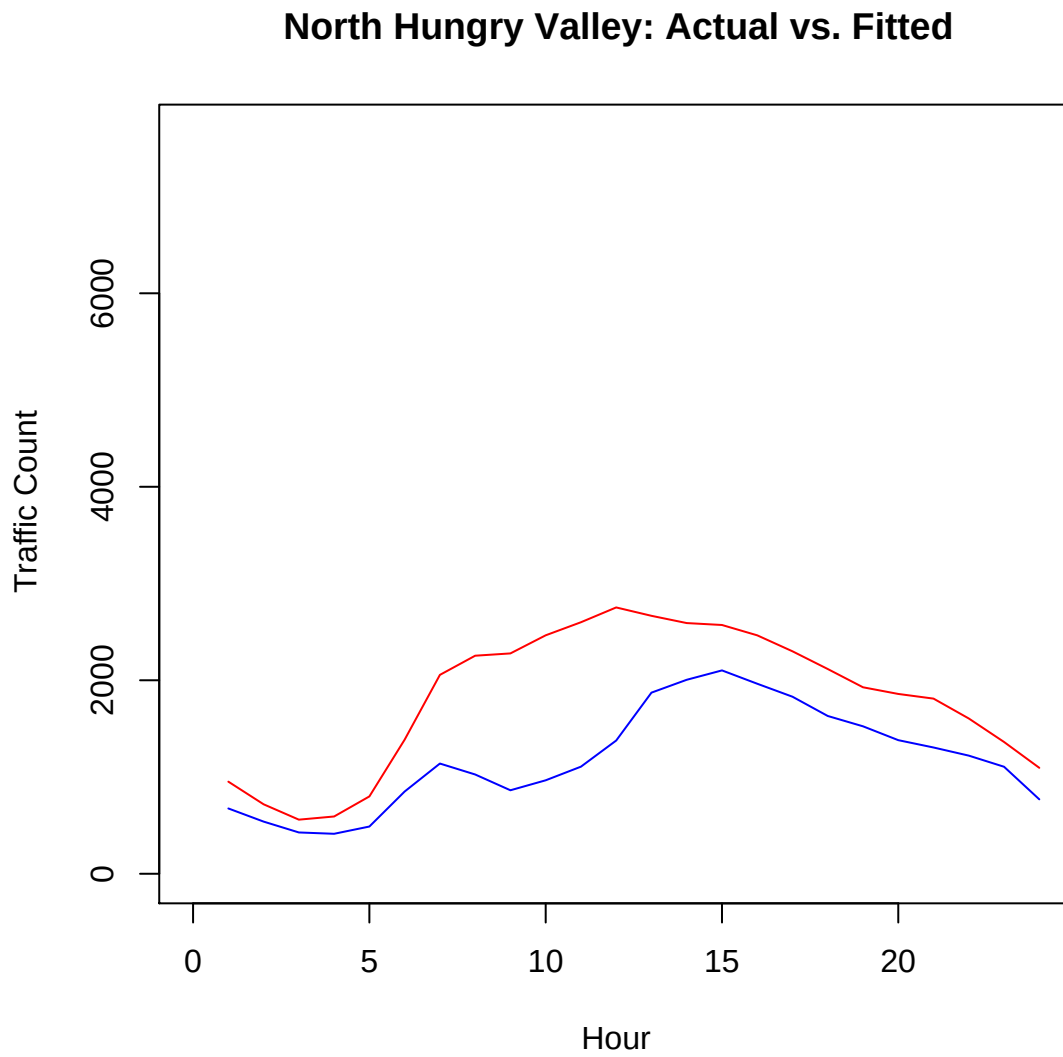


FIGURE 6. Hungry Valley Fitted Data

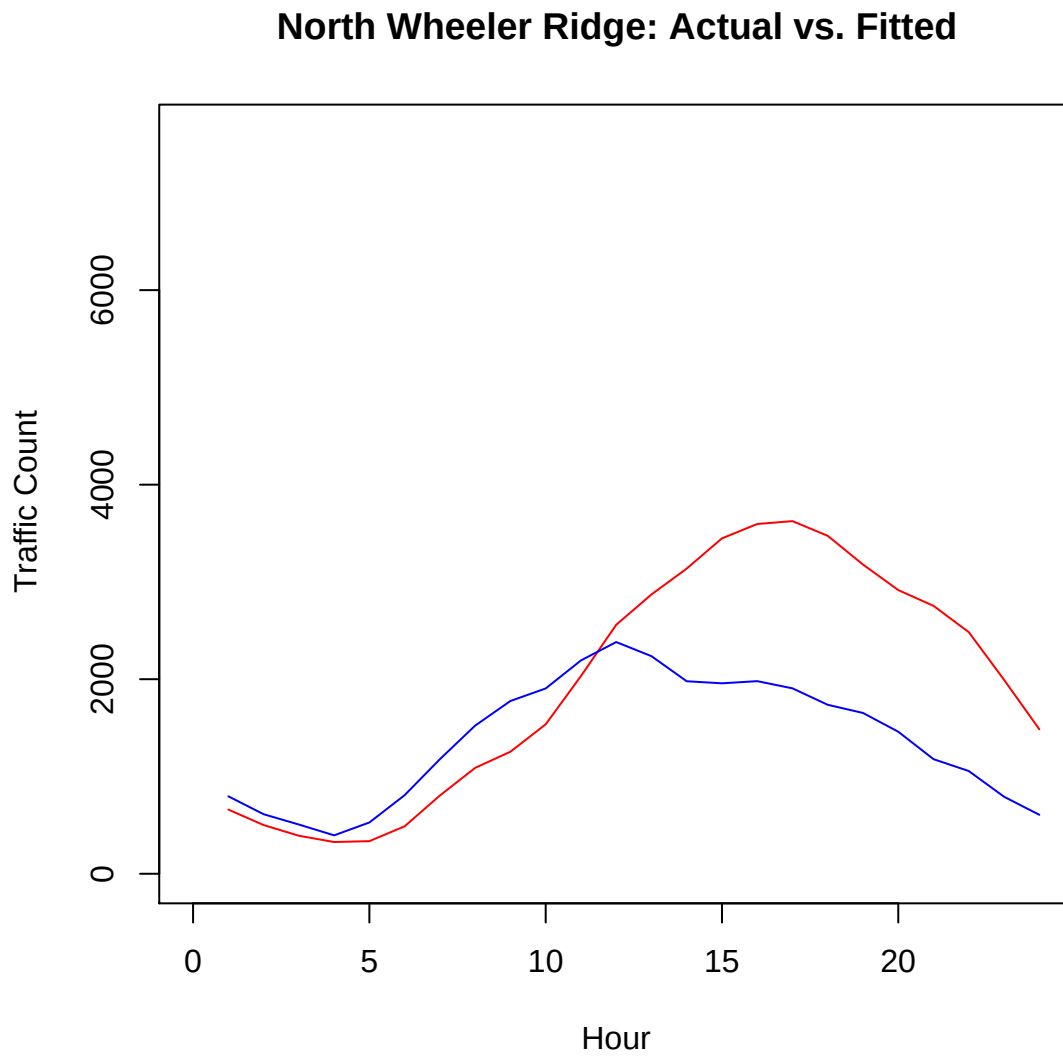


FIGURE 7. Wheeler Ridge Fitted Data

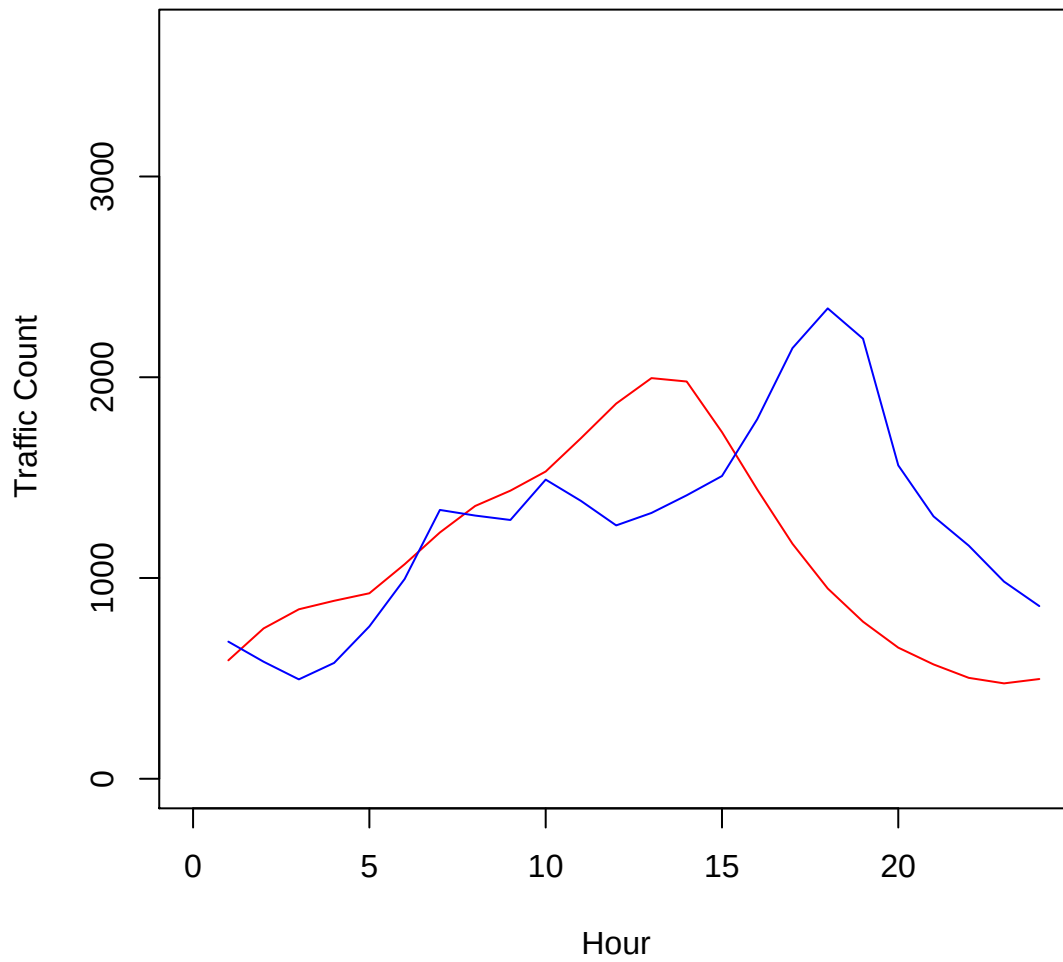
North Junction Route 14: Actual vs. Fitted

FIGURE 8. Route 14 Fitted Data

	Route 126	Hungry Valley	Wheeler Ridge	Route 14
1	18461236565.27	11575235988.71	14350344480.16	2863211000.16
2	4322870279.36	2367339431.83	2598717936.67	1280948278.88
3	2186878681.71	891683574.29	758308794.75	859718208.53
4	749647567.51	440387711.21	464364658.44	364024254.60
5	350870459.32	276439817.75	200538317.80	225324325.44
6	217516499.02	179185827.09	125133174.12	91209059.58
7	200704585.61	159136084.77	120383247.02	63631651.66
8	140660283.33	138853987.12	103546060.74	42346667.59
9	124097716.09	127188117.63	69829097.55	36194223.10
10	103016042.65	105159268.43	69393701.97	28457225.65
11	92304273.58	95009760.31	56092128.50	24339106.59
12	72078032.01	90616024.37	48204085.20	21441819.54
13	62224169.29	80924139.62	39014381.82	19427878.77
14	58307176.08	70170477.84	37685206.99	17132303.65
15	53344694.05	58314564.86	31627198.41	16233180.68
16	49804556.11	54510323.88	26342594.84	14743502.42
17	45197745.71	54184581.13	24644440.65	12276837.06
18	40896477.60	50644558.25	21533853.32	11223030.20
19	29198664.80	46601893.53	19099000.51	7066087.72
20	25841682.26	43142064.87	14881591.64	6256882.28
21	19924355.34	33564660.72	8734192.93	5330893.08
22	13930419.25	29987081.81	7256748.30	4605682.39
23	6018216.06	23936821.47	4438598.84	2181105.76
24	3141323.96	13395269.00	2394262.04	1858078.02

	Route 126	Hungry Valley	Wheeler Ridge	Route 14
1	0.67	0.68	0.75	0.48
2	0.16	0.14	0.14	0.21
3	0.08	0.05	0.04	0.14
4	0.03	0.03	0.02	0.06
5	0.01	0.02	0.01	0.04
6	0.01	0.01	0.01	0.02
7	0.01	0.01	0.01	0.01
8	0.01	0.01	0.01	0.01
9	0.00	0.01	0.00	0.01
10	0.00	0.01	0.00	0.00
11	0.00	0.01	0.00	0.00
12	0.00	0.01	0.00	0.00
13	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00
15	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00
20	0.00	0.00	0.00	0.00
21	0.00	0.00	0.00	0.00
22	0.00	0.00	0.00	0.00
23	0.00	0.00	0.00	0.00
24	0.00	0.00	0.00	0.00

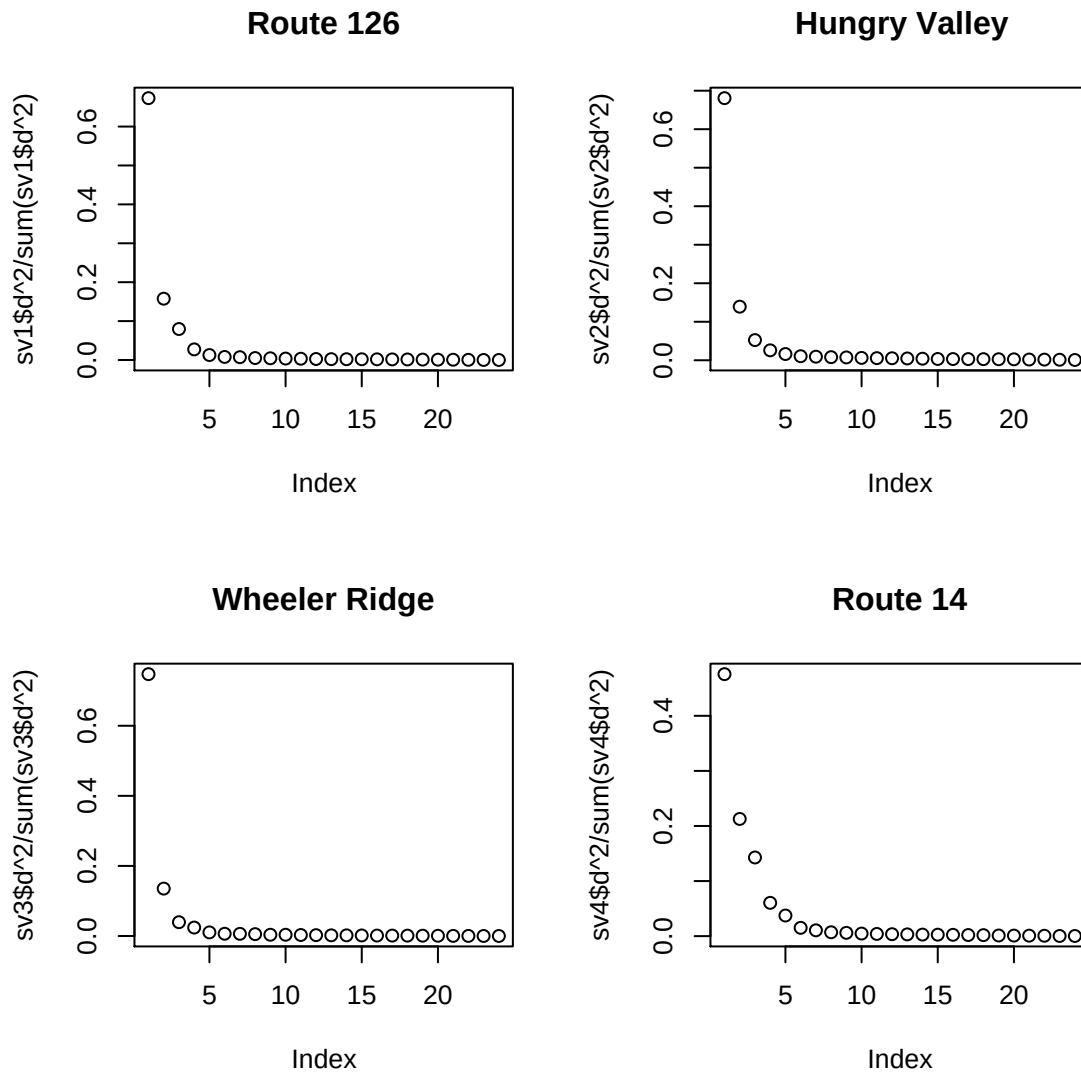


FIGURE 9. Eigenvalue Contributions for Complete Data

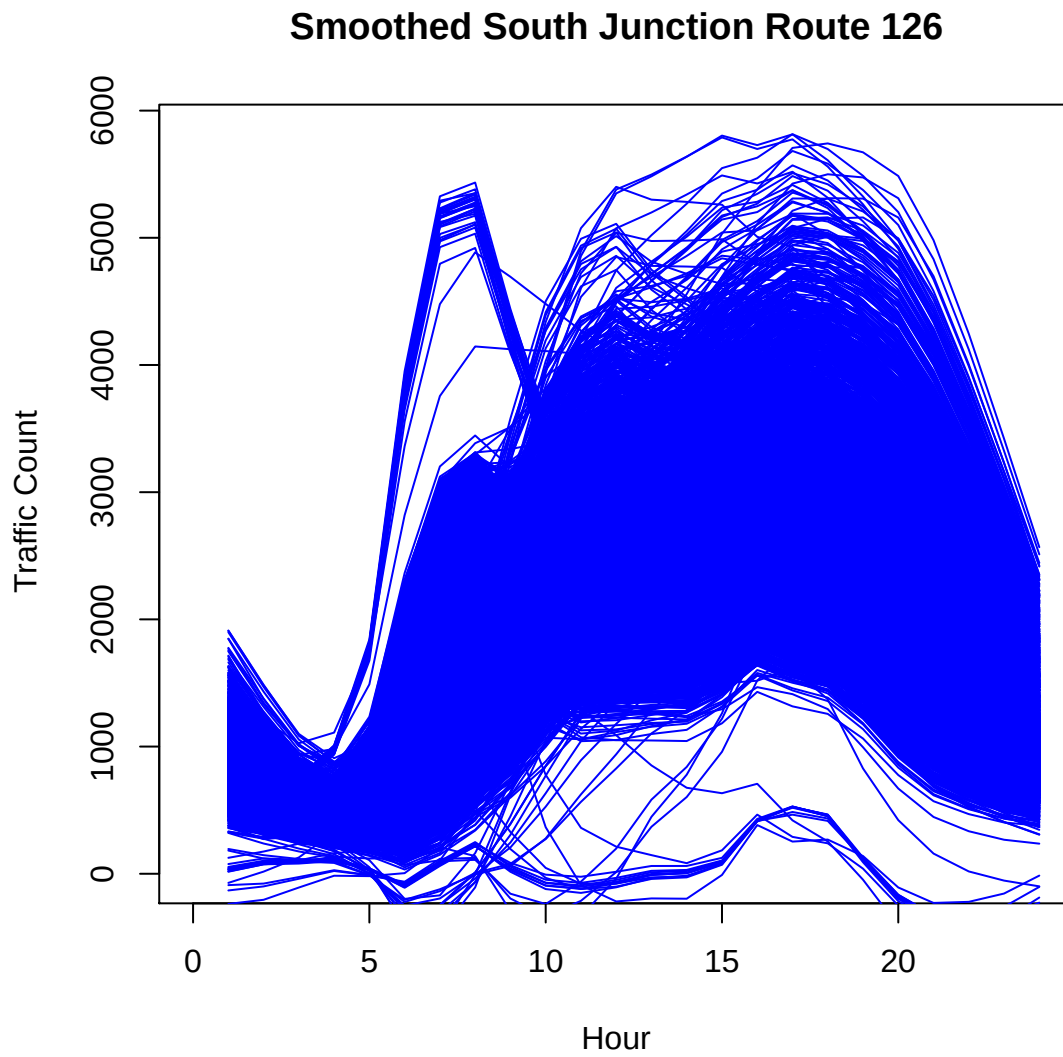
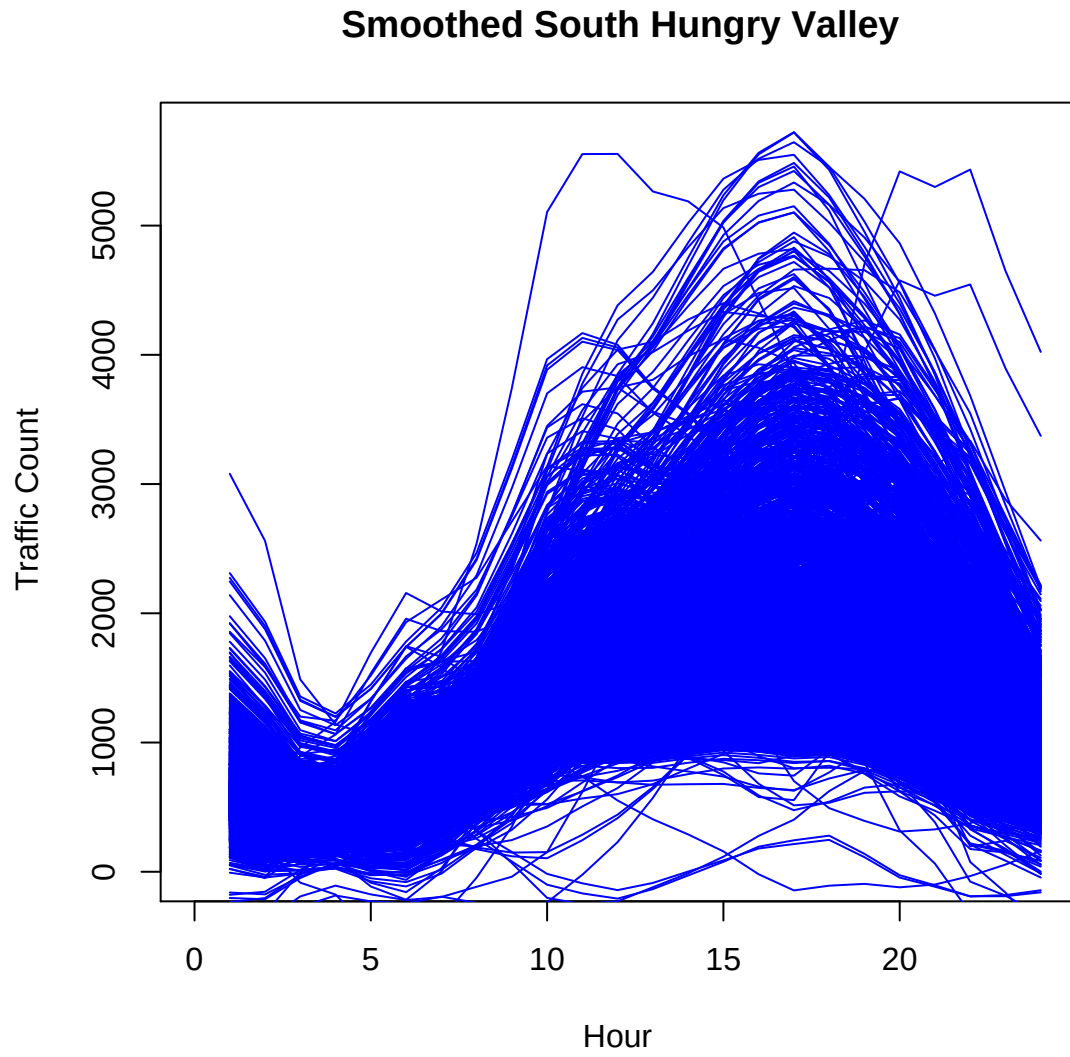
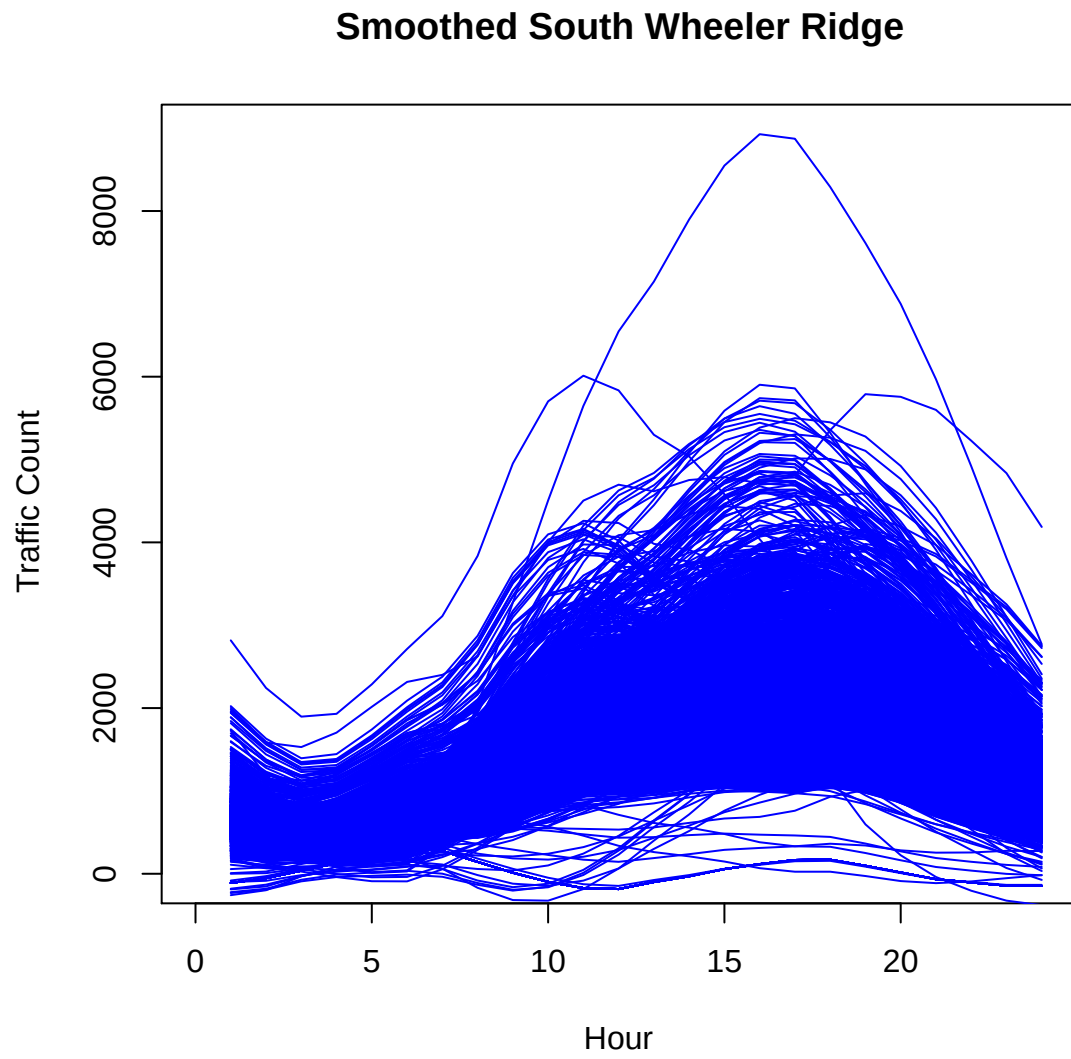


FIGURE 10. Route 126 Fitted Data



s_{fitted}.pdf

FIGURE 11. Hungry Valley Fitted Data



s_{fitted}.pdf

FIGURE 12. Wheeler Ridge Fitted Data

Smoothed South Junction Route 14

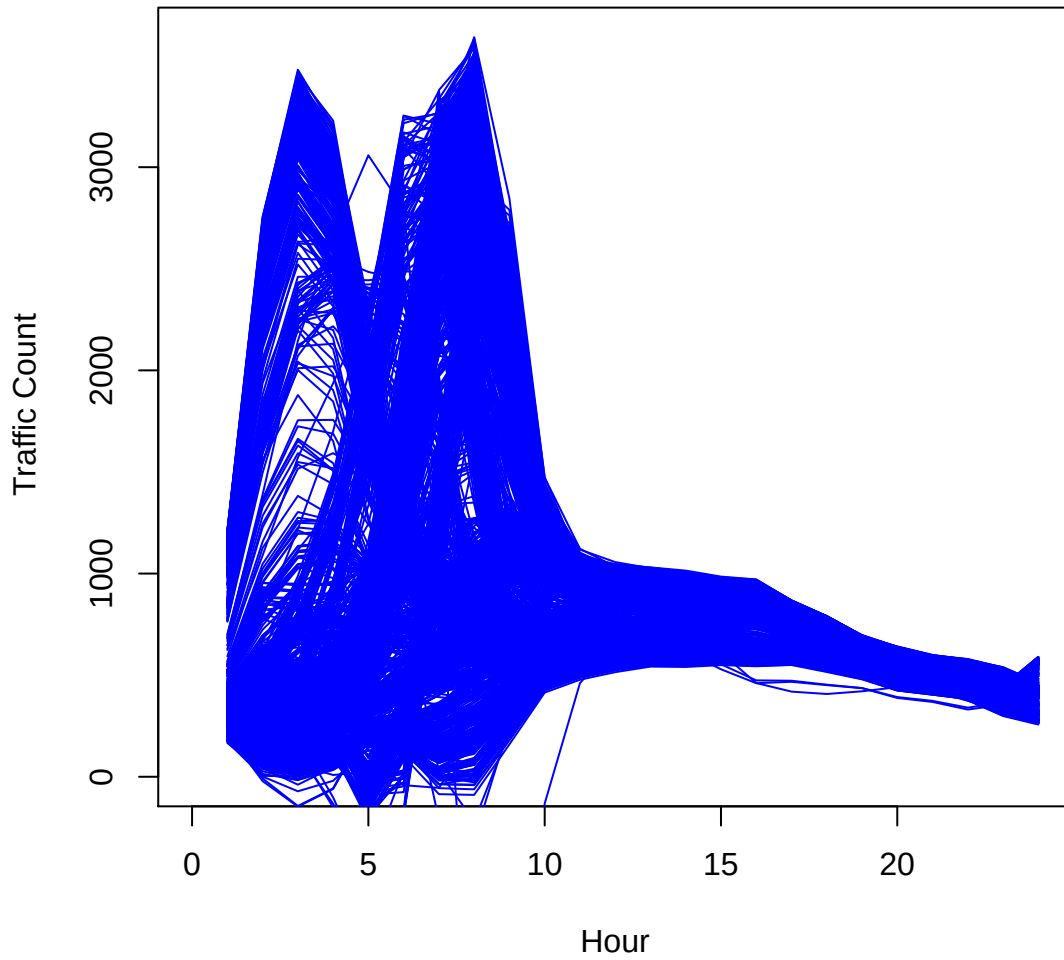


FIGURE 13. Route 14 Fitted Data

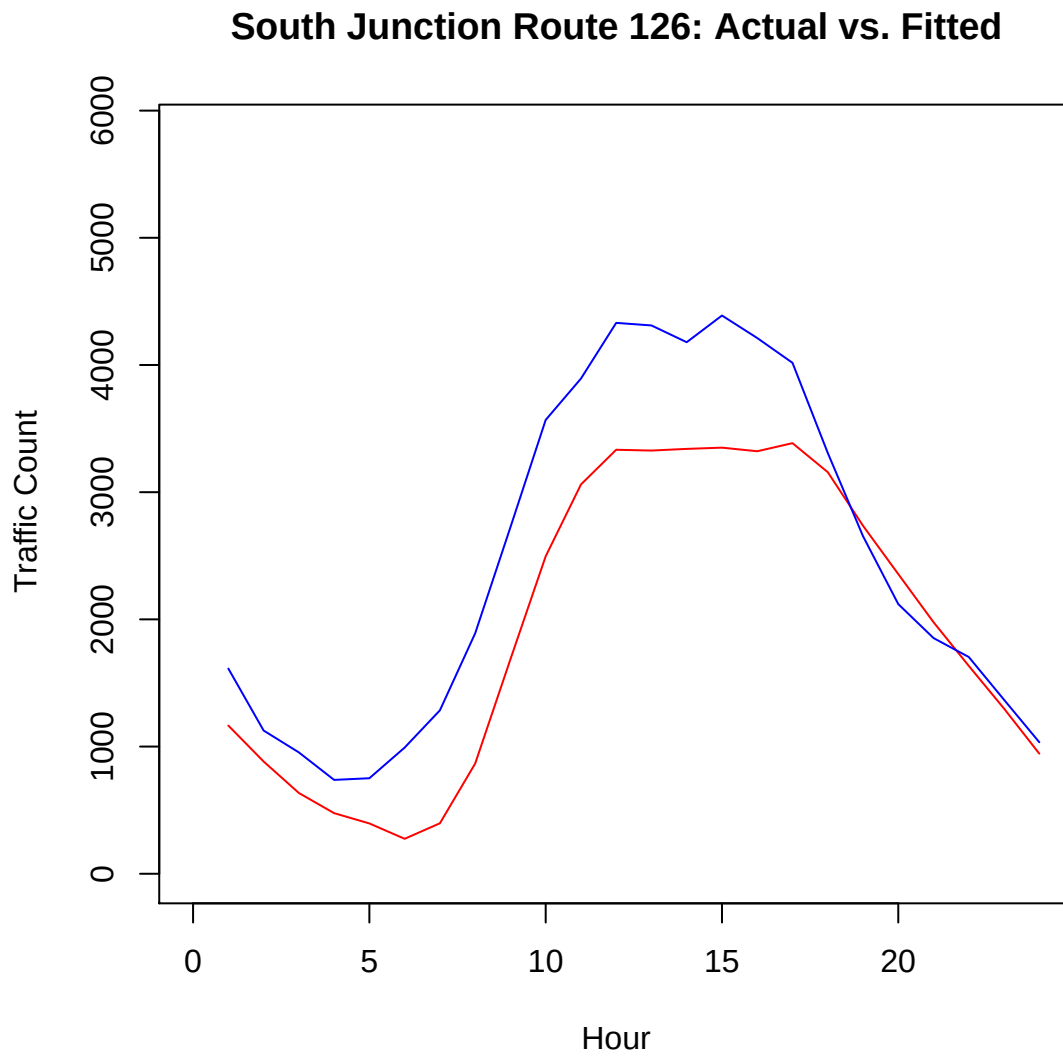


FIGURE 14. Route 126 Fitted Data

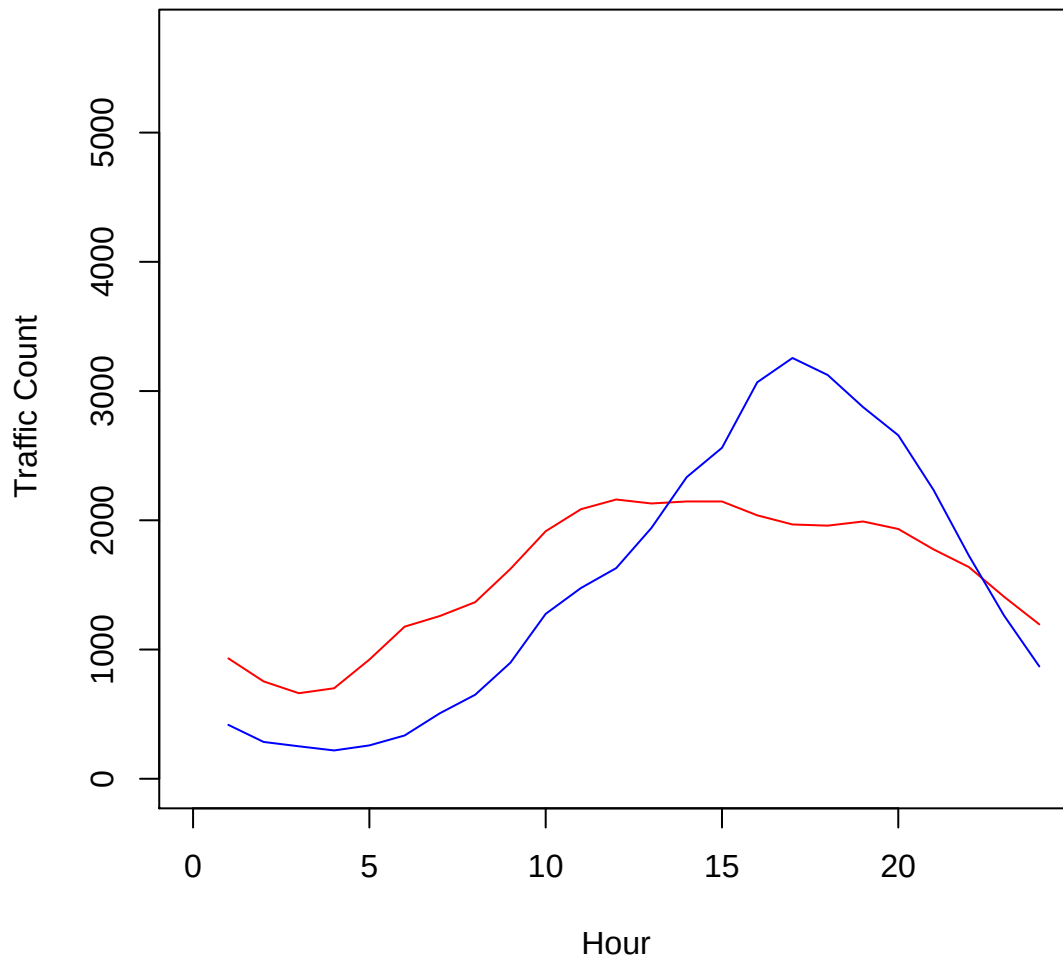
South Hungry Valley: Actual vs. Fitted

FIGURE 15. Hungry Valley Fitted Data

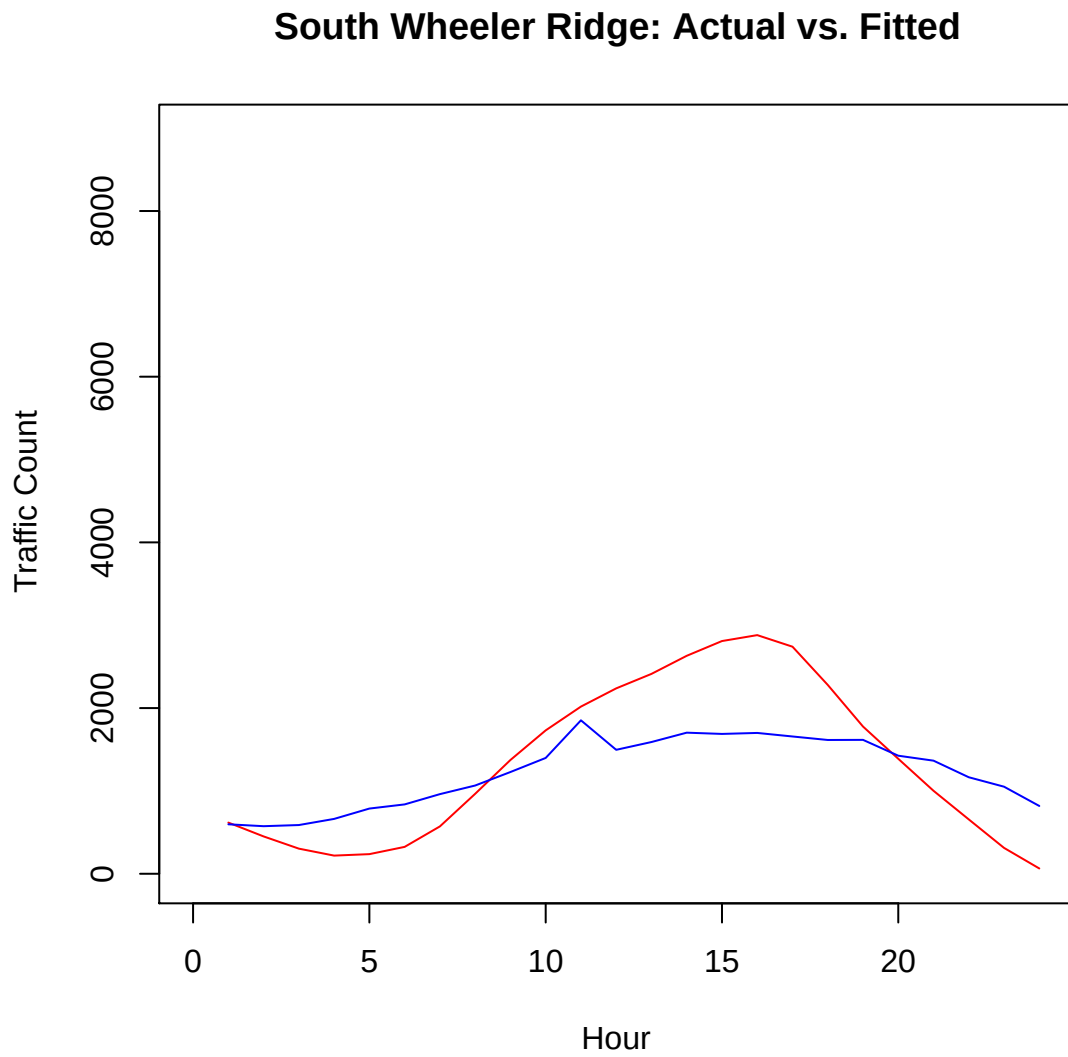


FIGURE 16. Wheeler Ridge Fitted Data

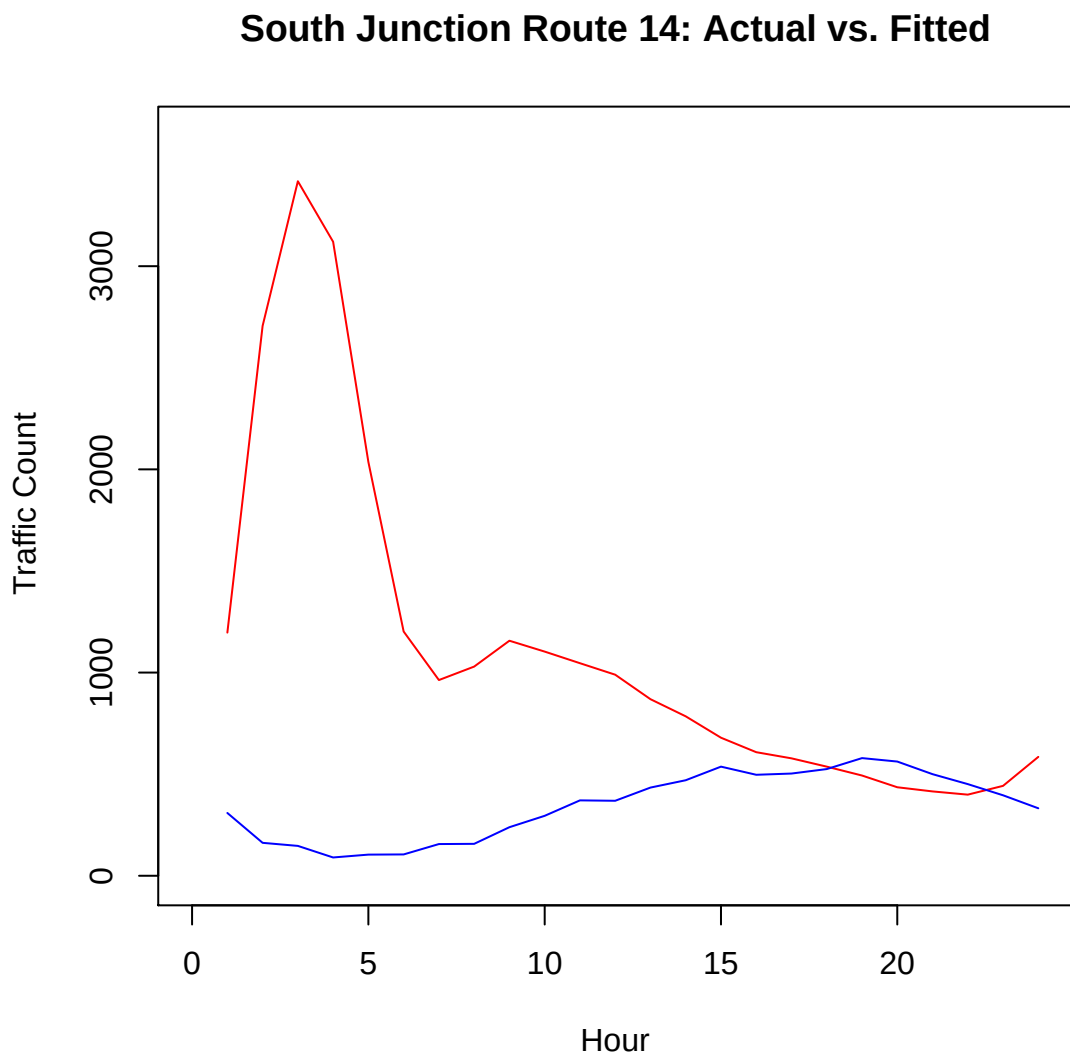


FIGURE 17. Route 14 Fitted Data

	Route 126	Hungry Valley	Wheeler Ridge	Route 14
1	21769014818.69	12949035665.07	16021220389.57	5722224351.74
2	7826839616.89	2212508385.82	2837888007.13	4163096415.29
3	2444256736.87	985412651.99	904936208.95	878760905.00
4	880924407.49	259985553.22	241367959.19	475585796.82
5	594928866.62	205780122.20	165708210.59	400463400.07
6	246506137.55	168943552.07	104291024.02	193546251.05
7	206864862.58	149617423.11	69510554.35	115167012.85
8	135804249.87	124869016.47	40667848.71	75789041.90
9	97599722.51	82286988.84	30167239.05	70551171.87
10	82772549.27	66239798.62	25828232.21	52976438.27
11	71455543.59	49012588.05	24718273.92	31719663.48
12	67915502.06	45929110.46	19926958.12	29925949.38
13	60848771.58	41775483.21	18546713.06	22989101.41
14	51769056.60	34325733.54	17313628.59	21165948.88
15	51270361.06	28343981.39	15522154.14	18393186.19
16	46011889.02	28027049.08	15464391.11	14174419.04
17	40198955.61	25091557.49	11981548.96	11557334.32
18	39567716.92	21283378.70	10973878.17	10741271.12
19	36638597.05	16421334.76	10489563.34	7578192.63
20	29357740.65	12937444.01	8562439.99	7324513.54
21	22601680.66	11601338.41	7481023.53	6697957.21
22	15519853.46	9486960.53	5603789.54	5532986.65
23	8234210.59	8402833.79	4973577.03	4107143.86
24	4413888.49	4420501.82	3307918.29	2511278.00