

交通信息检测与处理第二次实验报告

1854084 徐晨龙 交通信息

一 数据来源

设备布设位置为上海中环外圈，数据采集周期为 20s，数据种类包括流量、平均速度和时间占有率，线圈编号和车道序号如图所示：

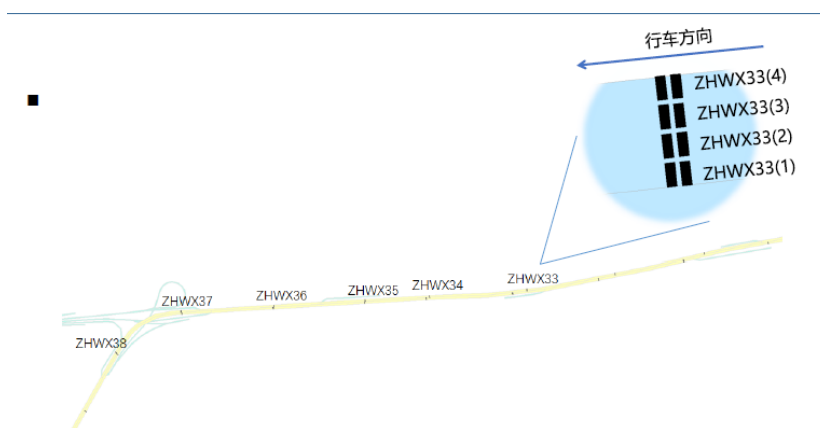


图 1：检测线圈摆放位置

二 数据结构

数据种类包括线圈编号、采集时间以及是否有效，空间上为 ZHWX33（1）~ZHWX38（1）数据，时间范围为 11 月 4 日~11 月 5 日。

	A	B	C	D	E	F
1	FSTR_LOOPGROUPID	FDT_TIME	FSTR_VALIDITY	FINT_VOLUME	FINT_SPEED	FINT_OCCUPY
2	ZHWX33(1)	2019/11/4 16:58	valid	7	77	8
3	ZHWX33(1)	2019/11/4 17:01	valid	8	70	10
4	ZHWX33(1)	2019/11/4 16:44	valid	8	74	9
5	ZHWX33(1)	2019/11/4 16:55	valid	6	81	7
6	ZHWX33(1)	2019/11/4 16:47	valid	9	74	10

图 2：数据示例

三 数据处理要求

（1）故障数据检测

a.以 ZHWX33（1）-1105 交通数据为例，采用 20s 采样间隔的流量、平均速度和时间占有率，分别利用独立判断法、联合判断法、平均有效长度法进行故障数据检测，并分别给出故障数据示例

b.以 ZHWX33（1）_1105 交通数据计算 20s 数据的小时有效性，并给出小时为颗粒度的有效性时效图，计算出不同方法检测出的故障数据比例

c.使用同样的故障数据检测方法，计算出六个检测器 4 日和 5 日的日有效性

(2) 数据修复

a.从 ZHWX33(1)-1105 中随机选取 100 条数据,利用原始数据对上述 100 条缺失数据进行修复,比较估计值和真实值,并量化估计误差,给出方法的修复精度

b.利用得到的估计模型,给出 ZHWX33(1)-1105 中故障数据的估计值,使用同样的故障数据修复方法,修复其他检测器中的故障数据

(3) 数据平滑处理

a.对修复后的 ZHWX33(1)-1105 交通流数据针对流量、平均速度、时间占有率进行多个时间窗口宽度的移动平均平滑处理和对比

b.对修复后的 ZHWX33(1)-1105 交通流数据针对流量、平均速度、时间占有率进行多个 α 值的指数平均平滑处理和对比

(4) 绘制交通态势图

a.利用 Windows=30 移动平均后的数据针分别绘制出 4 日和 5 日的交通态势图,各检测器的位置如下:

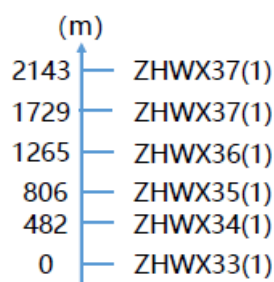


图 3: 各检测器之间的距离

b.看图说话,简单分析拥堵现象和原因

四 数据处理分析

(1) 故障数据检测

a.独立判断,即根据流量、时间平均速度和占有率的取值范围来判断异常值,本实验取值范围如下:

$$\begin{aligned}0 &\leq q \leq 17(pcu/20s) \\ 0 &\leq v \leq 80 * 1.4(km/h) \\ 0 &\leq o \leq 95\end{aligned}$$

由此可以检测出对应的故障数据 Error_independent_3305:

1 FSTR_LOOPGROUPID	2 FDT_TIME	3 FSTR_VALIDITY	4 FINT_VOLUME	5 FINT_SPEED	6 FINT_OCCUPY
ZHWX33(1)	2019/11/05 00:59	valid	1	116	1
ZHWX33(1)	2019/11/05 01:25	valid	1	121	1
ZHWX33(1)	2019/11/05 04:45	valid	1	121	1
ZHWX33(1)	2019/11/05 09:47	valid	0	0	100
ZHWX33(1)	2019/11/05 22:34	valid	1	127	1
ZHWX33(1)	2019/11/05 23:54	valid	1	116	1

图 4: ZHWX33(1)-1105 独立判断后的故障数据

联合判断,即根据流密速之间的逻辑关系进行判别,主要判断方式如下

$$((q == 0) + (v == 0) + (o \geq 95) + (0 == 0) == 3) \text{ or } ((q == 0) + (v == 0) + (o \geq 95) == 0)$$

若该式子成立则数据正确

1 FSTR_LOOPGROUPID	2 FDT_TIME	3 FSTR_VALIDITY	4 FINT_VOLUME	5 FINT_SPEED	6 FINT_OCCUPY
ZHWX33(1)	2019/11/05 00:44	valid	0	0	1
ZHWX33(1)	2019/11/05 04:12	valid	0	0	1
ZHWX33(1)	2019/11/05 09:53	valid	0	0	86

图 5: ZHWX33 (1) -1105 联合判断后的故障数据

有效车长判断:

$$Avel = v * o / q$$

$$3 \leq AveL \leq 12$$

由此可得到故障数据:

	1 FSTR_LOOPGROUPID	2 FDT_TIME	3 FSTR_VALIDITY	4 FINT_VOLUME	5 FINT_SPEED	6 FINT_OCCUPY
1	ZHWX33(1)	2019/11/05 00:47	valid	2	76	1
2	ZHWX33(1)	2019/11/05 05:39	valid	2	78	1
3	ZHWX33(1)	2019/11/05 05:51	valid	2	78	1

图 6: ZHWX33 (1) --1105 有效车长判断后的故障数据 (部分)

b.综合上述三种方法,可以得到应该错误值矩阵 Error, 绘制出 ZHWX33 (1) -1105 的 20s 数据的小时有效性图:

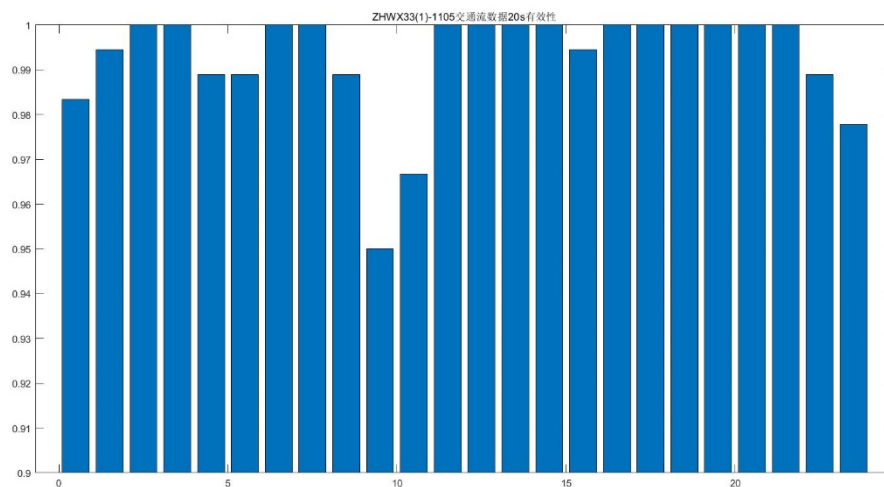


图 7: ZHWX33 (1) -1105 数据小时有效性

并由此得到不同方法检测出的故障数据比例:

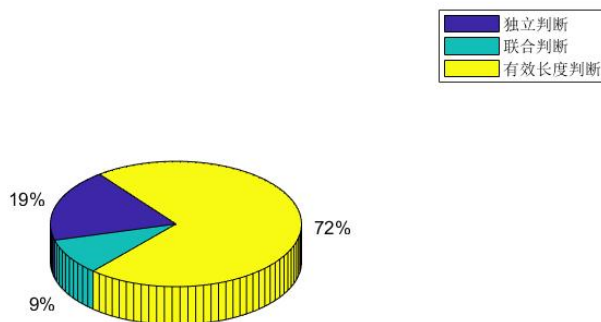


图 8: 三种故障检测数据比例

c. 编写 matlab 的 function:

`[rate, Error_correct, data_bug] = data_check(data)`

其中 `rate` 是输入数据的日有效性, `Error_correct` 是输入数据中的错误值, `data_bug` 即为修改之后的数据 (与第 (2) 题有关)

之后可以计算出各检测器 4, 5 日的日有效性

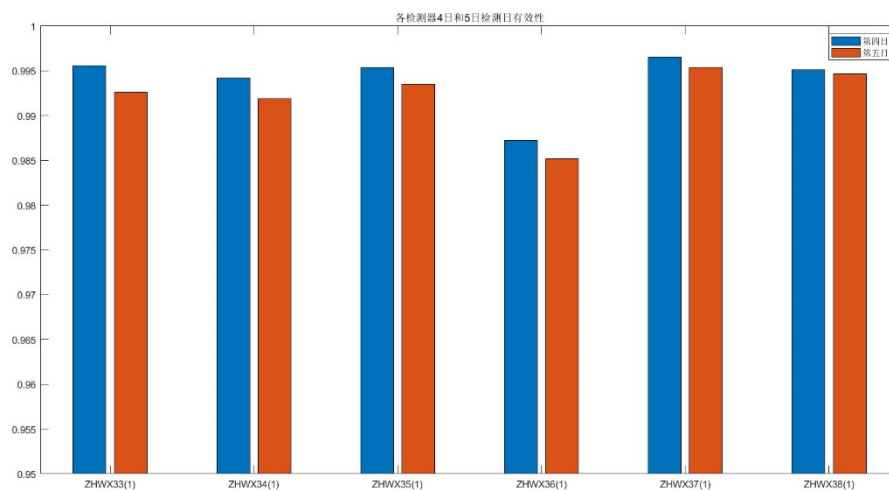


图 9 各检测器 4 日和 5 日数据日有效性

(2) 数据修复

a. 利用 matlab 的随机数生成函数 `randperm` 生成 100 个不相等的随机数 `delete_Index`, 由于实验中仅有两天的数据, 因此利用相邻时段中离线处理的方法:

$$y(k, t) = (y(k, t - 1) + y(k, t + 1)) / 2$$

可以得到估计值和真实值之间的差别

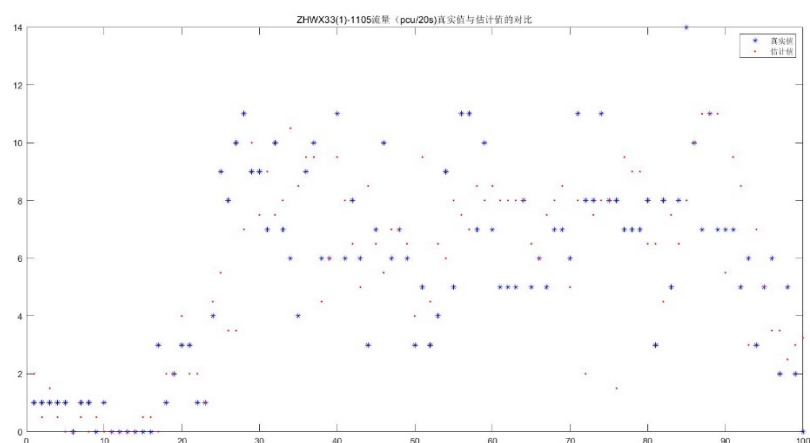


图 10: ZH WX33 (1) --1105 流量 (pcu/20s) 真实值和估计值对比

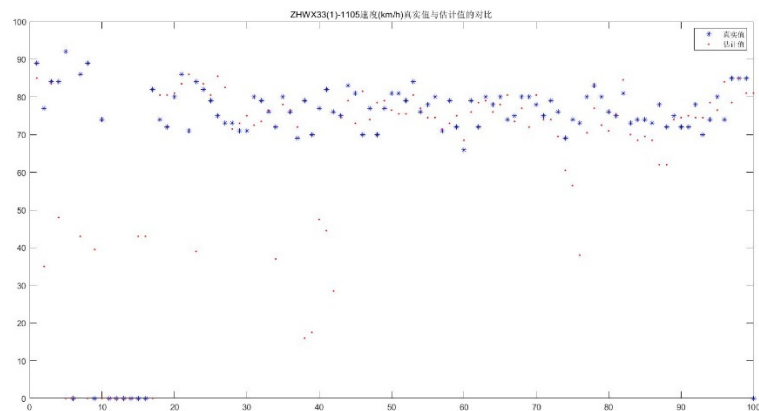


图 11: ZHWX33 (1) --3305 速度 (km/h) 真实值和估计值对比

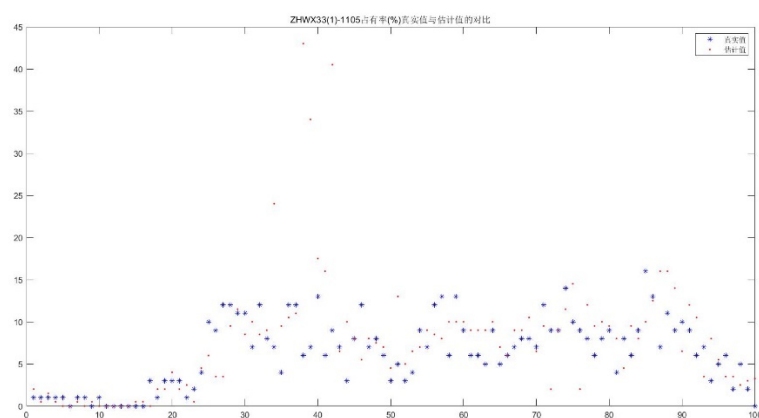


图 12: ZHWX33 (1) --3305 占有率 (%) 真实值和估计值对比

计算量化误差:

$$D\text{elta} = \text{abs}(\text{treatment} - \text{raw}) / \text{sum}(\text{raw})$$

可以得到 ZHWX33 (1) -1105 的流量、速度、占有率的量化误差:

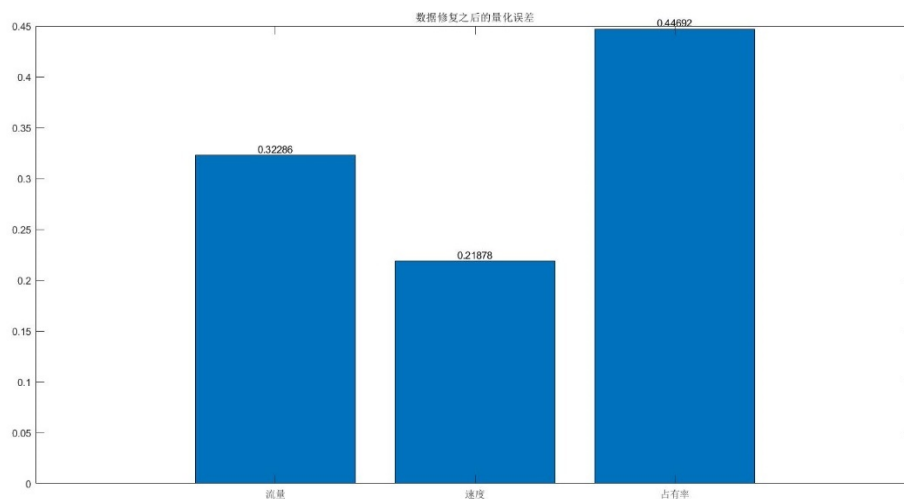


图 13: 数据修复的量化误差

b. ZHWX33 (1) --1105 处理前后的流量、速度和占有率（部分）为：

4	5	6	1	2	3
FINT_VOLUME	FINT_SPEED	FINT_OCCUPY			
0	0	1	0.5000	43	0.5000
1	116	1	0.5000	38.5000	0.5000
1	121	1	1	74	1
0	0	1	1	42.5000	1
1	121	1	0	0	0
0	0	100	4	11.5000	62.5000
0	0	86	4.5000	10	44
1	127	1	6	80	6
1	116	1	1.5000	84	1.5000

图 14：ZHWX33 (1) --1105 故障数据 q、v、o（部分）修复前后对比
利用函数 data_check 对其他数据进行修复。

(3) 数据平滑处理

a. 移动平均的计算即为，对于原始数据 $(y_1,y_2,y_3 \dots ,y_n)$
平滑后的数据 $(z_1,z_2,z_3, \dots ,z_n)$

$$\begin{aligned} z_1 &= y_1 \\ z_2 &= (y_1 + \dots + y_k)/k \\ &\dots \\ z_{(n-k+2)} &= (y_{(n-k+1)} + \dots + y_k)/k \\ z_{(n-k+3)} &= y_{(n-k+3)} \\ &\dots \end{aligned}$$

编写 matlab 函数

$$function output = yidongpingjun(input,a)$$

Output 即为移动平均平滑后的数据，input 即为原始数据，a 为移动平均的窗口大小，

ZHWX33(1)-1105线圈流量（veh/20s）移动平均平滑后结果

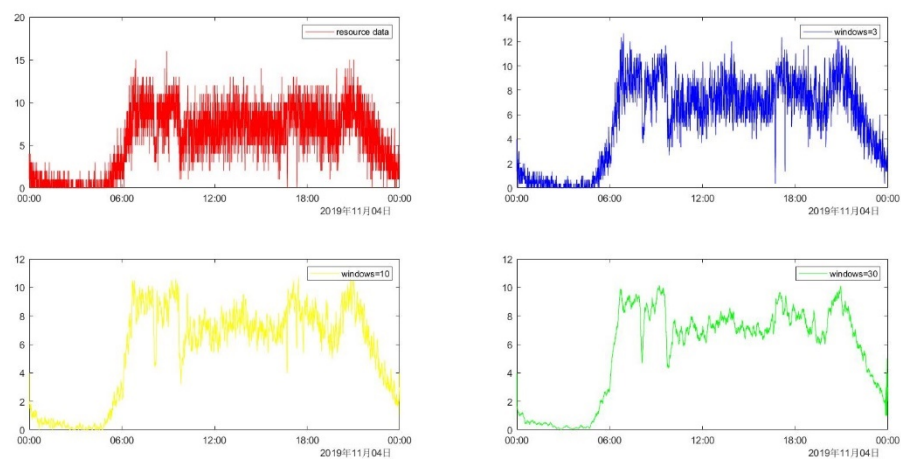


图 15:ZHWX33(1)—1105 流量（pcu/20s）不同窗口移动平滑

ZHWX33(1)-1105线速度 (km/h)移动平均平滑后结果

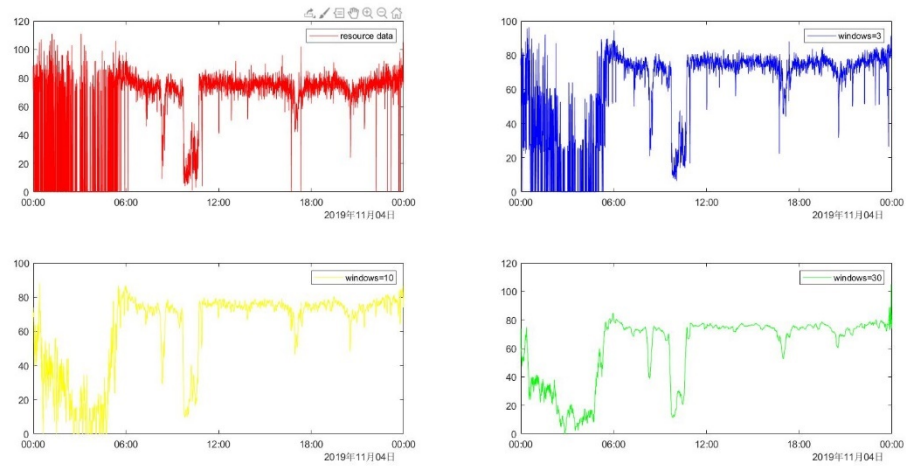


图 16: ZHWX33(1)—1105 速度 (km/h) 不同窗口移动平滑

ZHWX33(1)-1105线图占有率 (%)移动平均平滑后结果

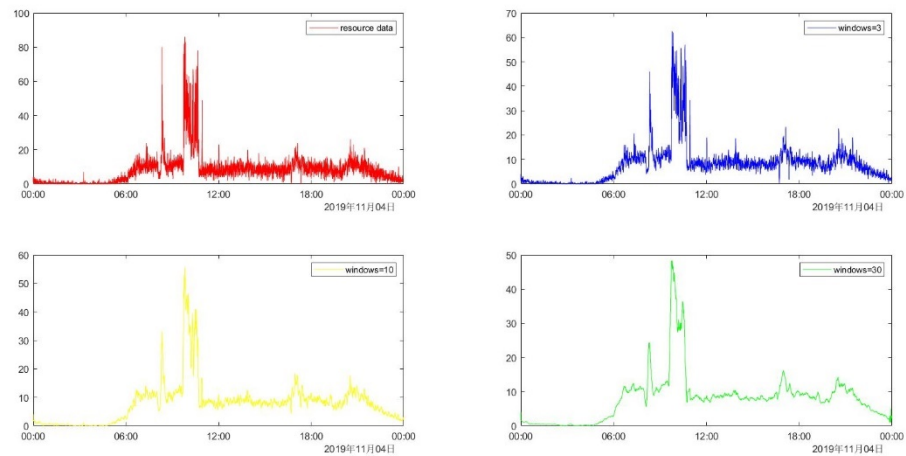


图 17: ZHWX33(1)—1105 占有率 (%) 不同窗口移动平滑

b. 指数平滑计算过程即为: (y_1, y_2, \dots, y_n)
平滑后的数据 (z_1, z_2, \dots, z_n)

$$z_1 = y_1$$

...

$$z(k) = a * y_k + (1 - a) * z_{(k-1)}$$

编写 matlab 函数 zhishupingjun

function output = zhishupingjun(input,a)

Output 为指数平滑后的数据, input 为原始数据, a 即为平滑系数

ZHWX33(1)-1105线圈流量(veh/20s)指数平均平滑后结果

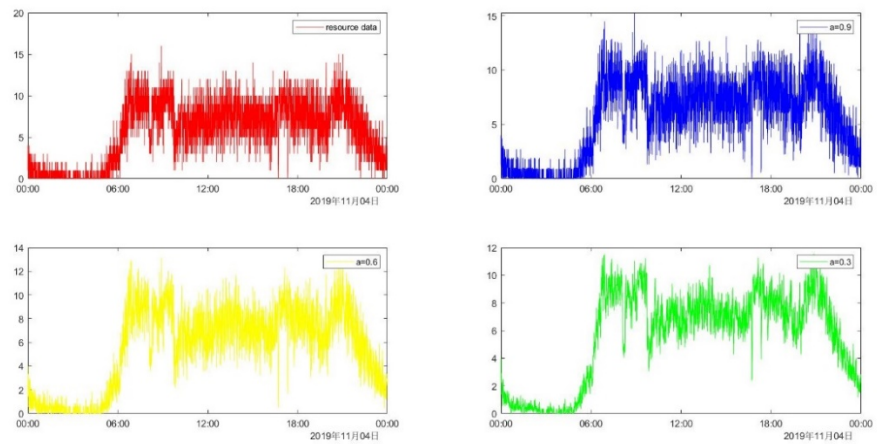


图 18: ZHWX33(1)—1105 流量 (pcu/20s) 不同 a 值指数平滑

ZHWX33(1)线圈速度 (km/h)指数平均平滑后结果

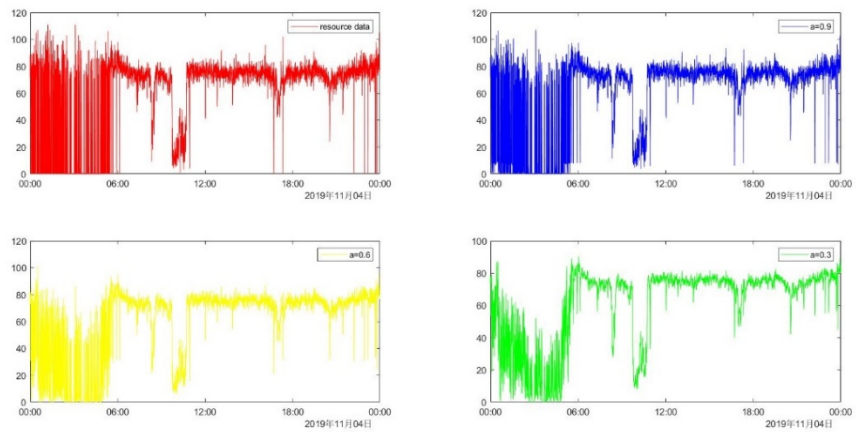


图 19: ZHWX33(1)—1105 速度 (km/h) 不同 a 值指数平滑

ZHWX33(1)线圈速度 (km/h)指数平均平滑后结果

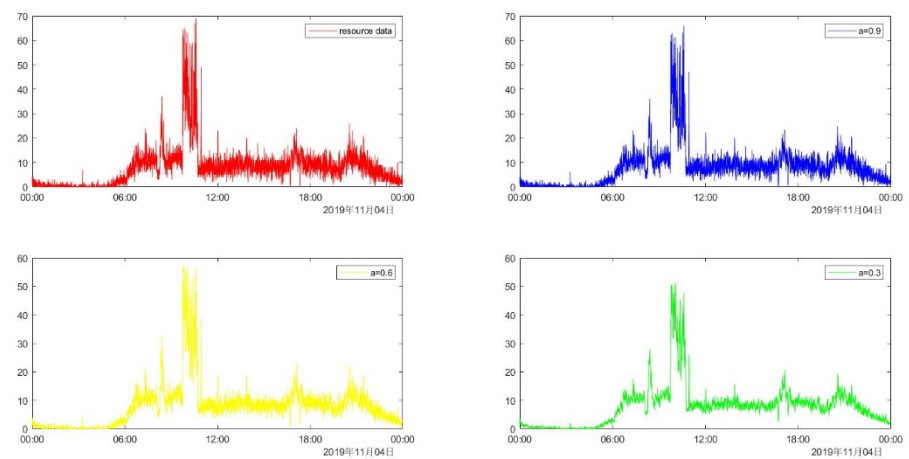


图 20 ZHWX33(1)—1105 占有率 (%) 不同 a 值指数平滑

c. 分析

1. 移动平均和指数平均都可以从交通流波动较大的原始数据中提取出数据的总体趋势，有利于数据趋势揭示。
2. 移动平均随着窗口值逐渐增加，平滑效果越来越好
3. 指数平均随着 a 值增加，平滑效果并无明显改善
4. 指数平均平滑效果没有移动平均好

(4) 绘制交通态势图

a. 首先将六个检测器两天一共 12 个数据进行移动平均 (window=30) 进行平滑处理，之后编写 matlab 的函数

$$function Q = data_imagesc(y, data1, data2, data3, data4, data5, data6)$$

Y 即为各线圈的相对位置，data1~data6 为对应线圈检测点的 20s 的共 24h 的平滑数据，检测器之间的数据采用线性拟合的方式进行计算：

$$(data_{(k+1)} - data_k) * [1:1: y(k+1) - y(k)] / (y(k+1) - y(k)) + data_k$$

输出的 Q 即为检测器所在路段的 24 小时数据，利用 matlab 的 `imagesc` 进行绘制可以得到 4 日和 5 日的交通(20s 流量、速度和占有率)态势图：

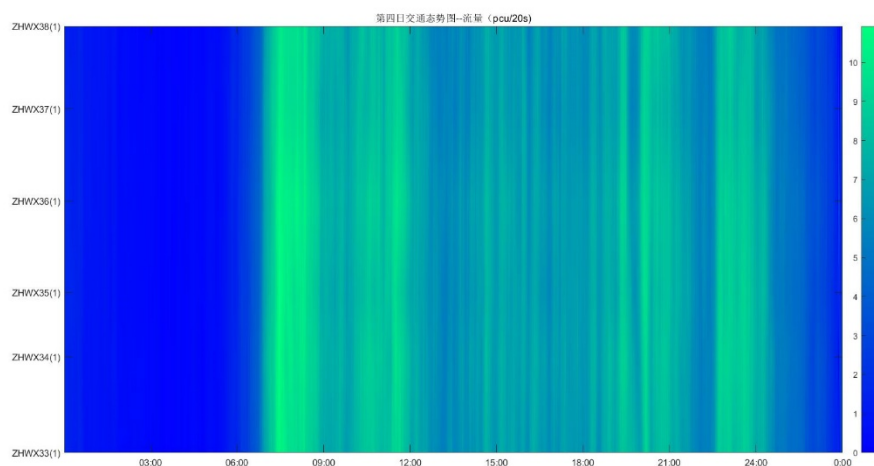


图 21：第四日交通态势图—流量 (pcu/20s)

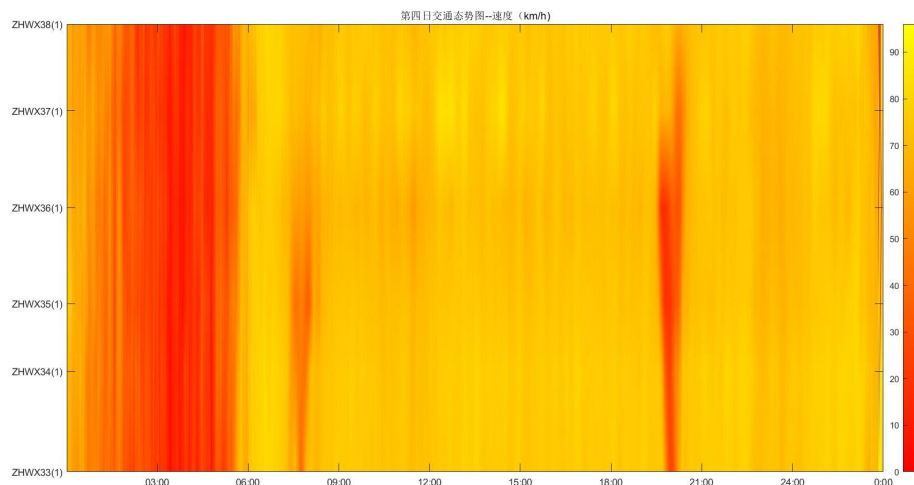


图 22：第四日交通态势图—速度 (km/h)

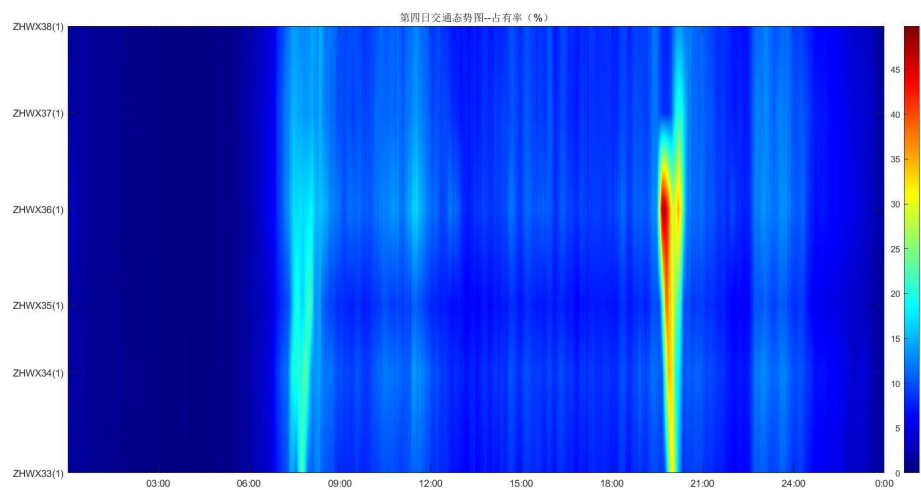


图 23:第四日交通态势图—占有率 (%)

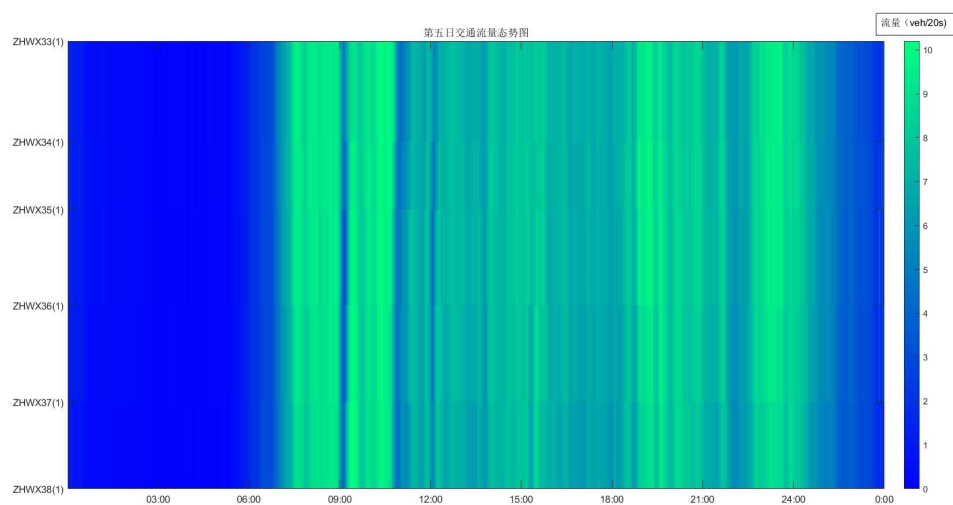


图 24:第五日交通态势图—流量 (pcu/20s)

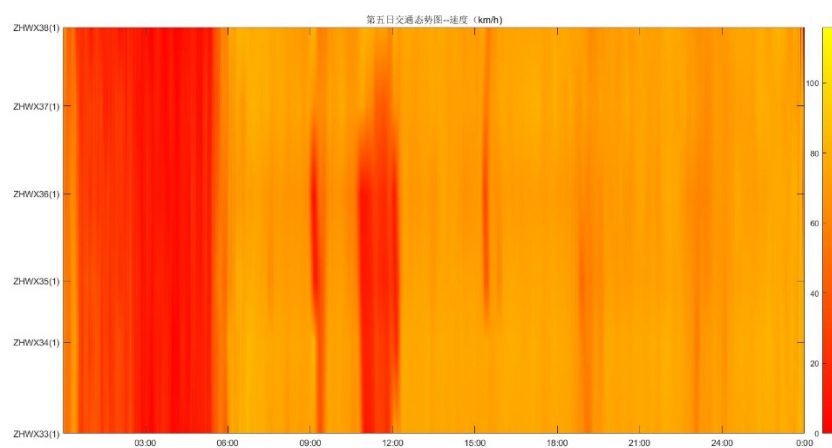


图 25:第五日交通态势图—速度 (km/h)

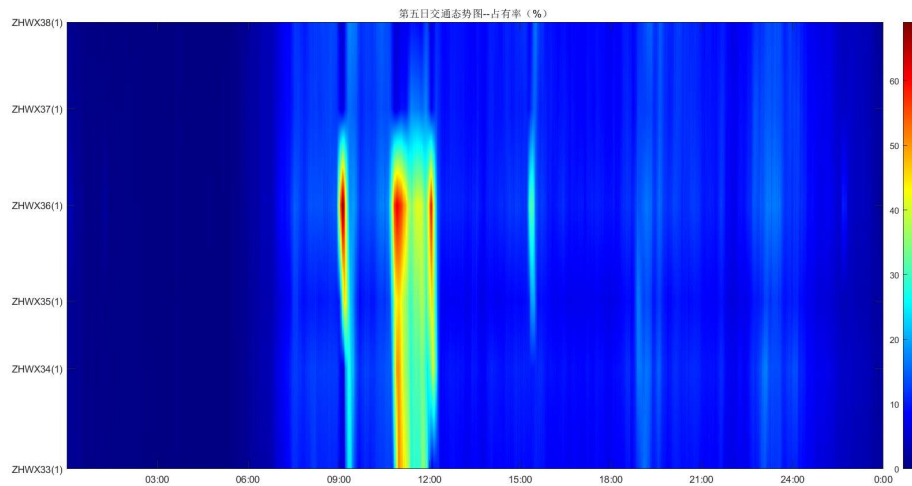


图 26:第五日交通态势图—占有率 (%)

b. 分析拥堵

第四日交通拥堵发生在晚上 20:00 左右,主要集中在 ZHWX38 和 ZHWX37 之间,有图 1 的检测线圈的排列图可以看出 ZHWX38~ZHWX36 之间有交叉口,因此造成拥堵,预测为晚高峰造成拥堵,属于常发性交通拥堵。

第五日交通拥堵发生在 09:00~10:00 和 11:00~12:00,前者结合第四日可以看出因为早高峰通行需求和交叉口通信能力过低造成的常发性拥堵,后者持续时间较长并且第四日并没有对应趋势,因此可以判断为发生交通事故或其他原因,经过新闻查找,为 2019.11.05~2019.11.10 日上海开展进博会,可能因此导致需求异常增多,属于突发性交通拥堵。

附件:

Resource_data.mat

Experiment2.m

Data_check.m

Yidongpingjun.m

Zhishupingjun.m

Data_imagesc.m