BUS TRAVEL TIME PREDICTION USING STATE SPACE MODELS

Presentation for National conference on
Urban mobility – Challenges, Solutions, Prospects
during July 13-14, 2012 at IIT Madras

S.Vasantha Kumar
Ph.D Research Scholar, Dept. of Civil Engg., IIT Madras

Lelitha Vanajakshi
Assistant Professor, Dept. of Civil Engg., IIT Madras

Shankar C.Subramanian
Assistant Professor, Department of Engineering Design., IIT Madras

Introduction

- Real time bus arrival time information
- Encourage public transport and reduce congestion
- The expected modal share is 70:30 (Public transit: Personal vehicles) in Chennai in 2026 against the modal share of 35:65 in 2009 (Source: Chennai Metropolitan Development Authority, 2008)

(Source: Muenster Planning Office, Germany)

Bus Arrival Time Prediction Models and Algorithms

- Historical Data based Models
- Time Series Models → (Based on ARIMA models)
- Regression Models
- Machine Learning Models
- Artificial Neural Network Models
- Support Vector Machines
- Kalman Filtering Technique (KFT) based Models

Disadvantages with ARIMA

- Requires specialized software for parameter estimation and forecasting
  - Suwardo *et al.* (2010) used STATISTICA software for forecasting of bus arrival times in Malaysia.
- Difficult to explain to others
- Site-specific

Alternate to ARIMA

- Exponential smoothing
- Most preferred method for short term forecasting in business and industry
- Simple to understand, easy to implement, data storage and computing requirements are very minimal

Basic equation of exponential smoothing

$$S_t = \alpha y_{t-1} + (1-\alpha) S_{t-1}$$

S_t ⟶ Forecast at time period 't'

y_{t-1} ⟶ Observed value at time 't-1'

S_{t-1} ⟶ Forecasted value at time 't-1'

α ⟶ Smoothing constant $(0 \le \alpha \le 1)$

State space formulation of exponential smoothing

- A state space model for a time series consists of two equations.

- The first is called the state equation and determines the state X_{t+1} at time $t+1$ in terms of the previous state X_t and a noise term. The state equation is

$$X_{t+1} = F_t X_t + W_t, \quad t = 1,2,...,$$

- The second equation, called the observation equation, expresses the observation Z_t as a function of a X_t plus noise.

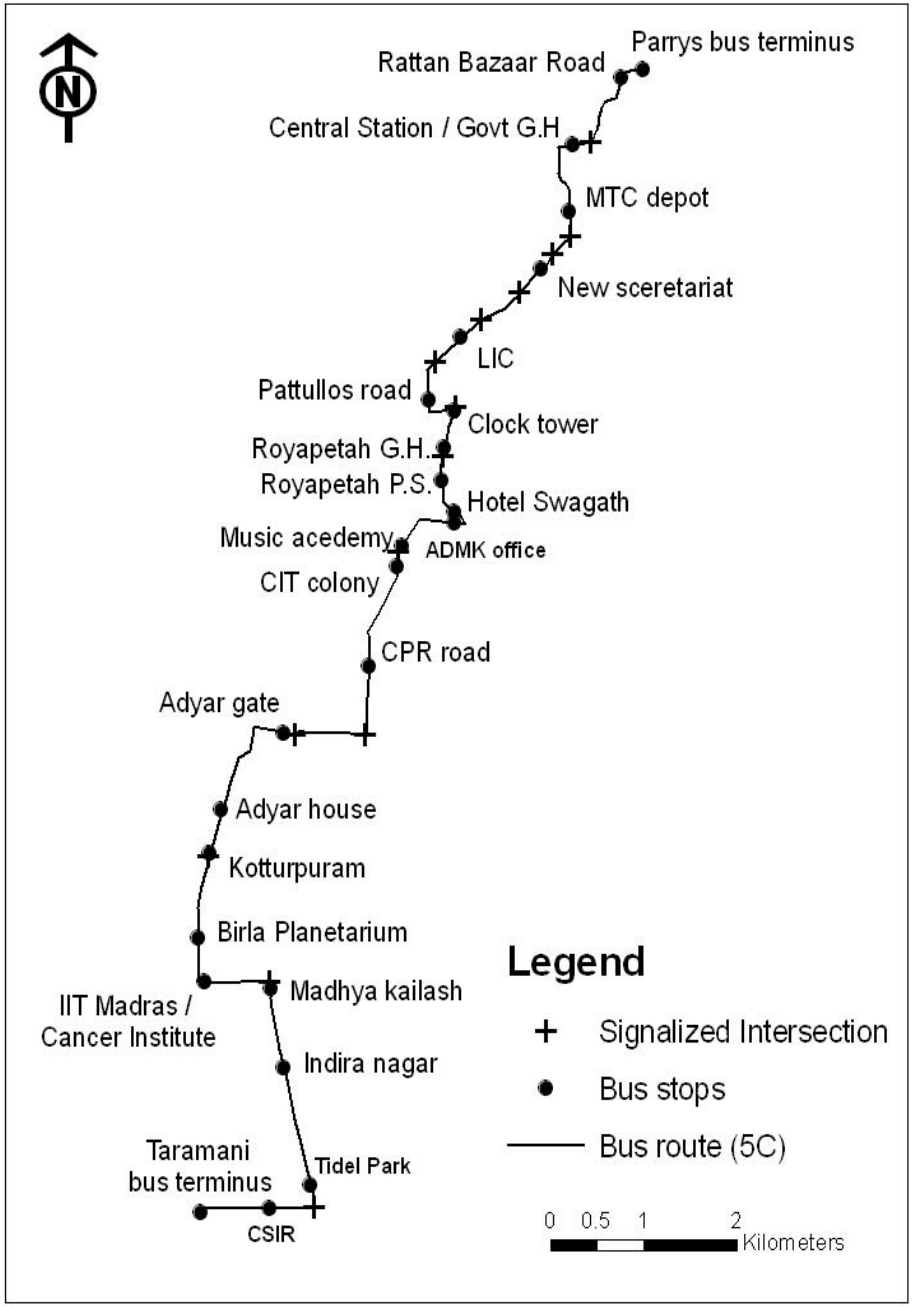
$$Z_t = G_t X_t + V_t, \quad t = 1,2,...,$$

Study stretch

Route number 5C
from Parrys to Taramani

Route length 15 km,
24 bus stops and
14 signalized intersections

Average time headway
between the buses was
15-30 minutes.



Data collection and extraction

Sl.No	Test Date	Number of trips	Corresponding Week 1	Number of trips	Corresponding Week 2	Number of trips
1	Sep 30, 2010	23	Sep 23, 2010	19	Sep 16, 2010	21
2	Nov 02, 2010	27	Oct 26, 2010	18	Oct 19, 2010	23
3	Nov 16, 2010	15	Nov 09, 2010	21	Nov 02, 2010	29
4	Dec 17, 2010	22	Dec 10, 2010	20	Dec 03, 2010	30
5	Dec 22, 2010	18	Dec 15, 2010	16	Dec 08, 2010	17
Total		105		94		120

Data extraction involved the extraction of 100m section travel time of 105 trips (test period), 94 trips (week 1) and 120 trips (week 2).

Evaluation

- Comparison of observed section travel time with the predicted travel time for all the 105 trips by both exponential smoothing method and exponential smoothing integrated with Kalman filtering method

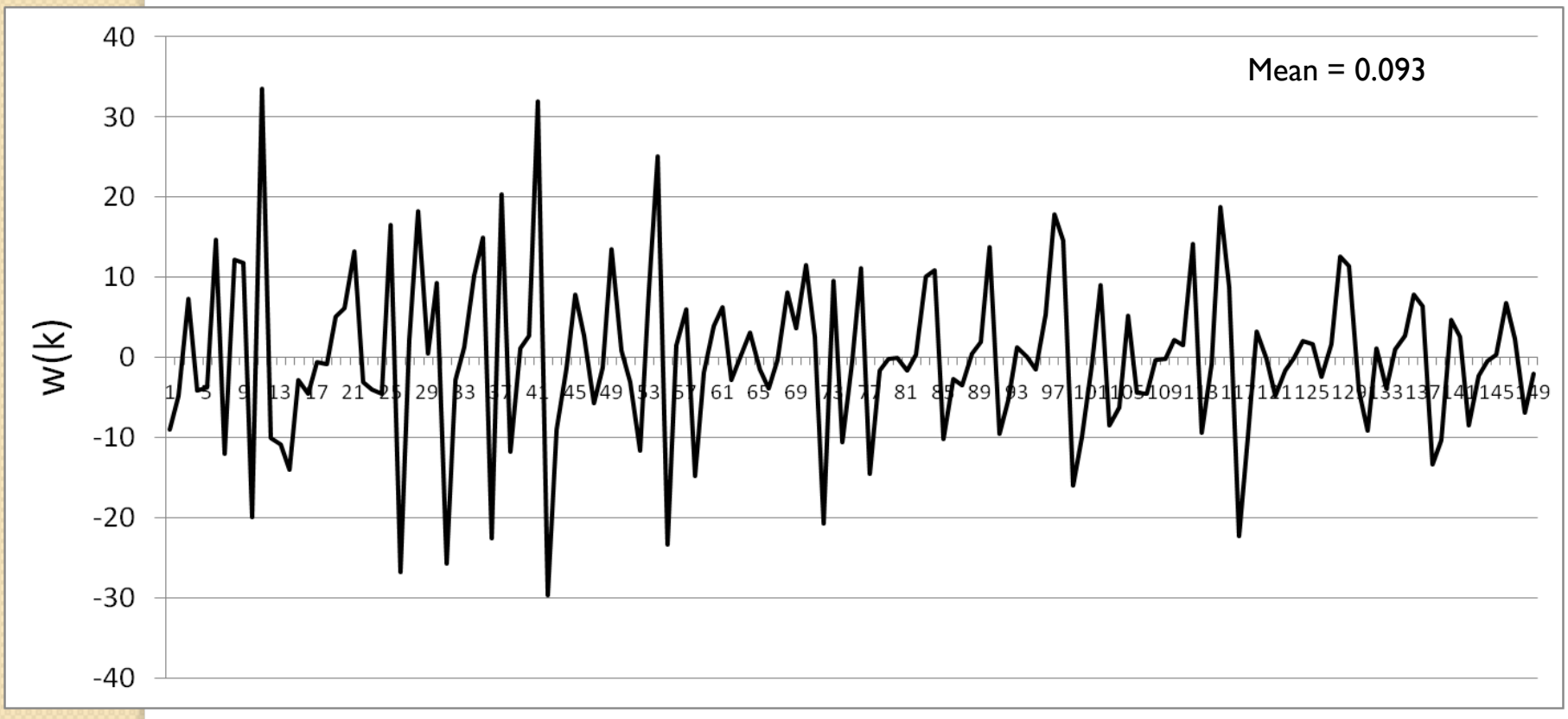
$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{tt_{pre} - tt_{obs}}{tt_{obs}}\right| \times 100$$

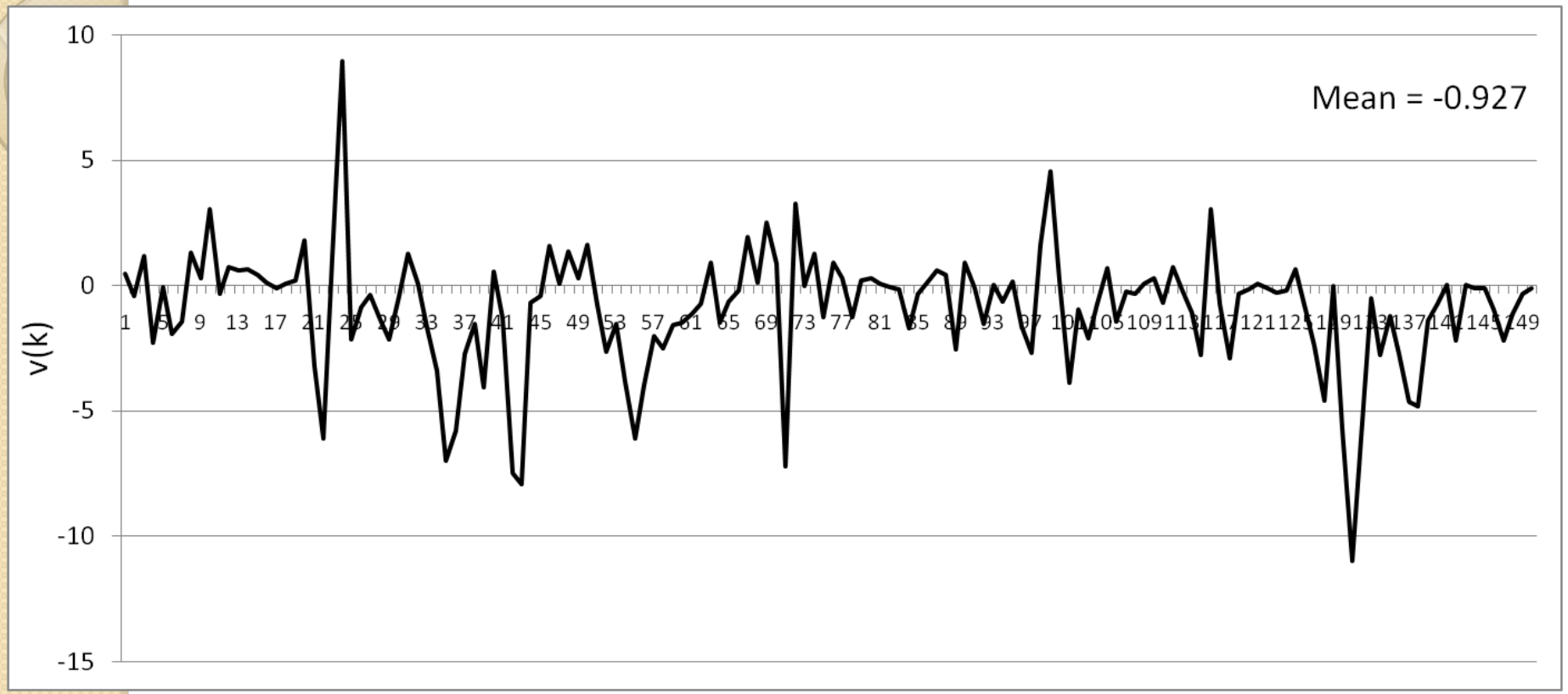
Evaluation (contd.)

- Comparison of observed arrival time of the bus at 21 bus stops along the route with the predicted arrival time, for all the 105 trips of the five days considered by both the methods

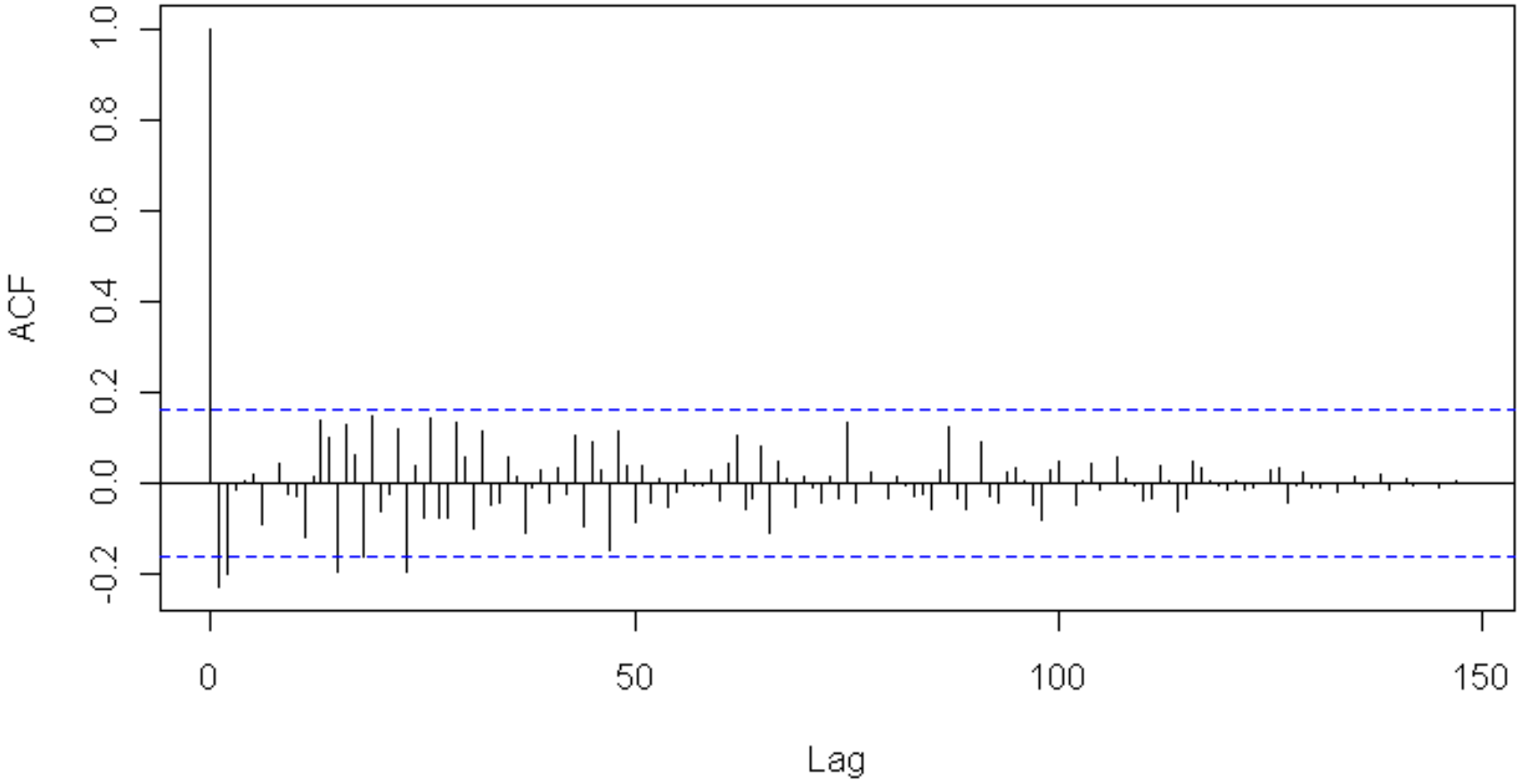
$$\mathrm{MAPE}_{bs} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{b_{pre} - b_{obs}}{b_{obs}}\right| \times 100$$

Assumptions about process disturbance (w(k)) and measurement noise (v(k))

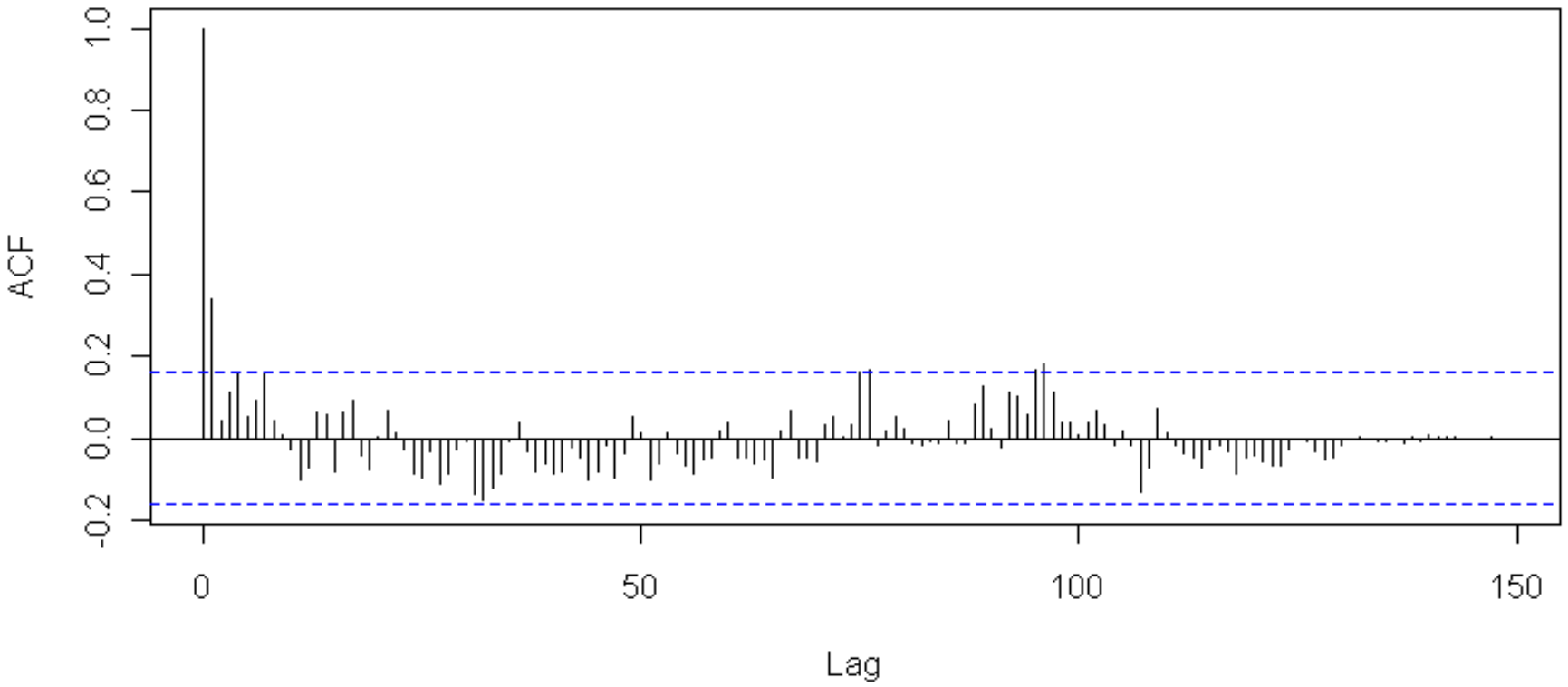


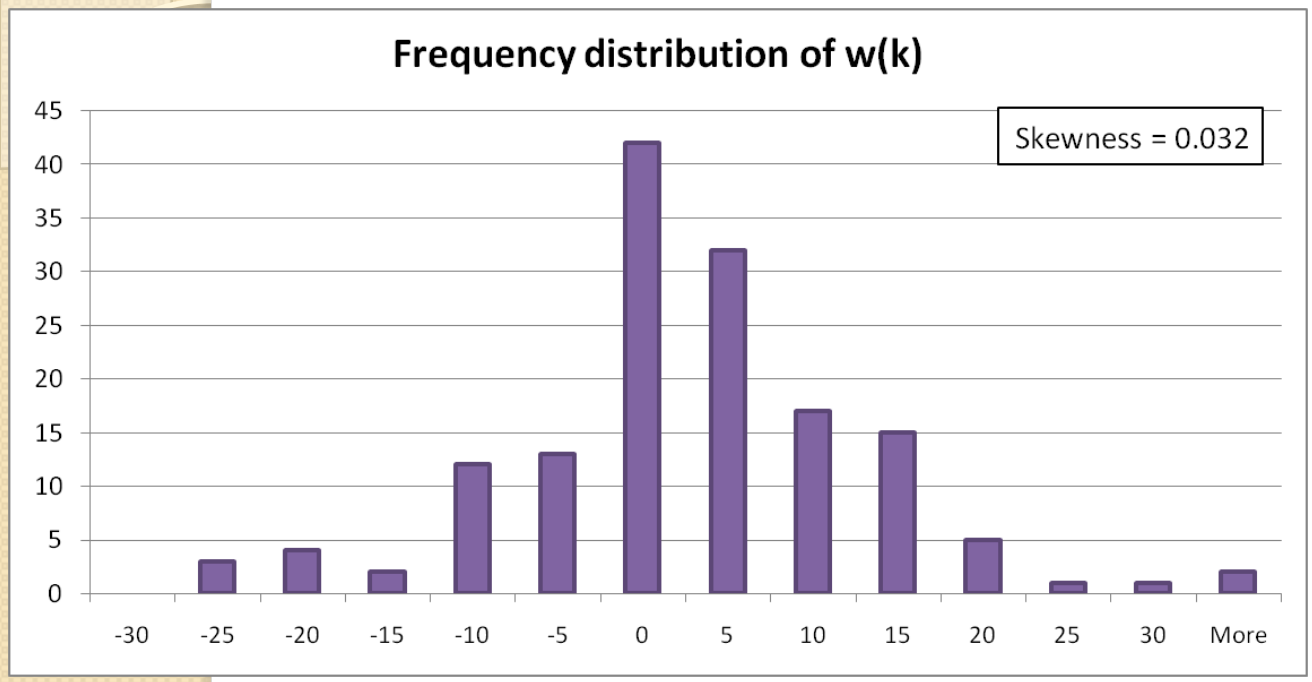

Mean = -0.927
v(k)
10
5
0
-5
-10
-15
1
5
9
13
17
21
25
29
33
37
41
45
49
53
57
61
65
69
73
77
81
85
89
93
97
101
105
109
113
117
121
125
129
133
137
141
145
149

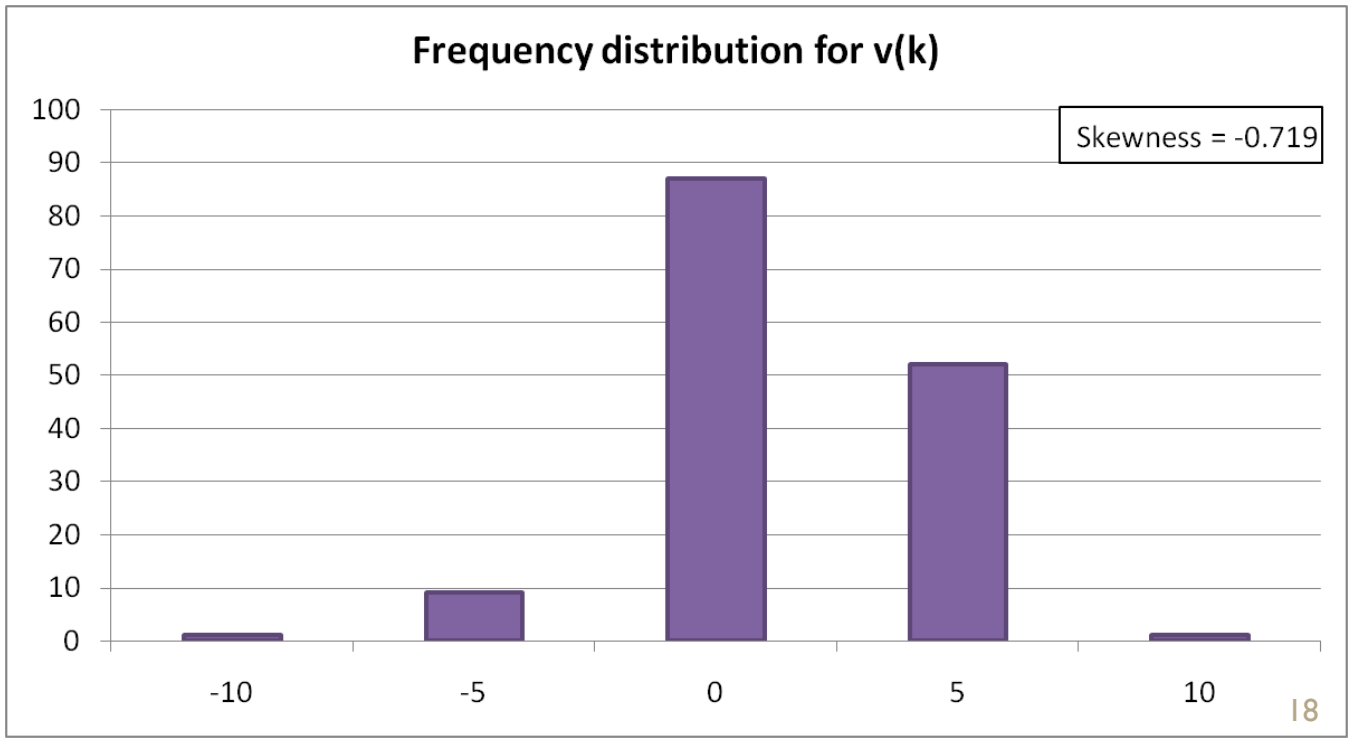
AUTO CORRELATION FUNCTION CALCULATED FOR w(k)

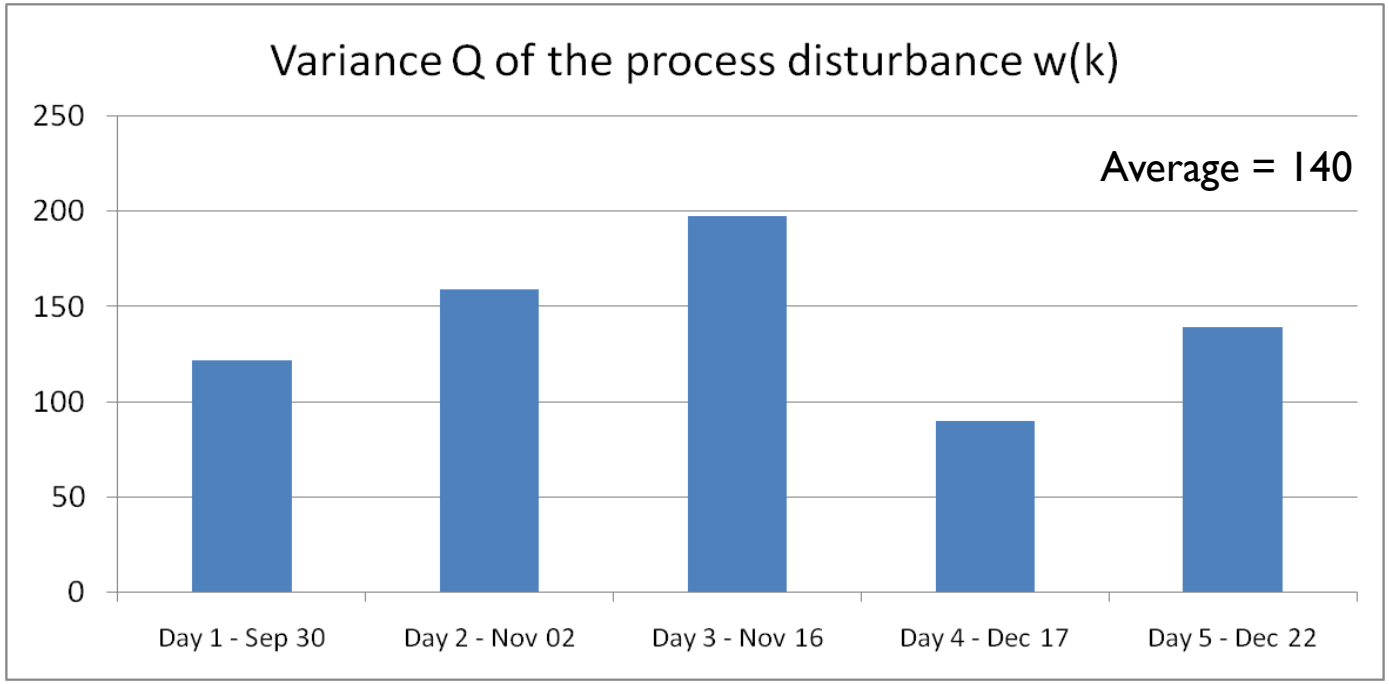


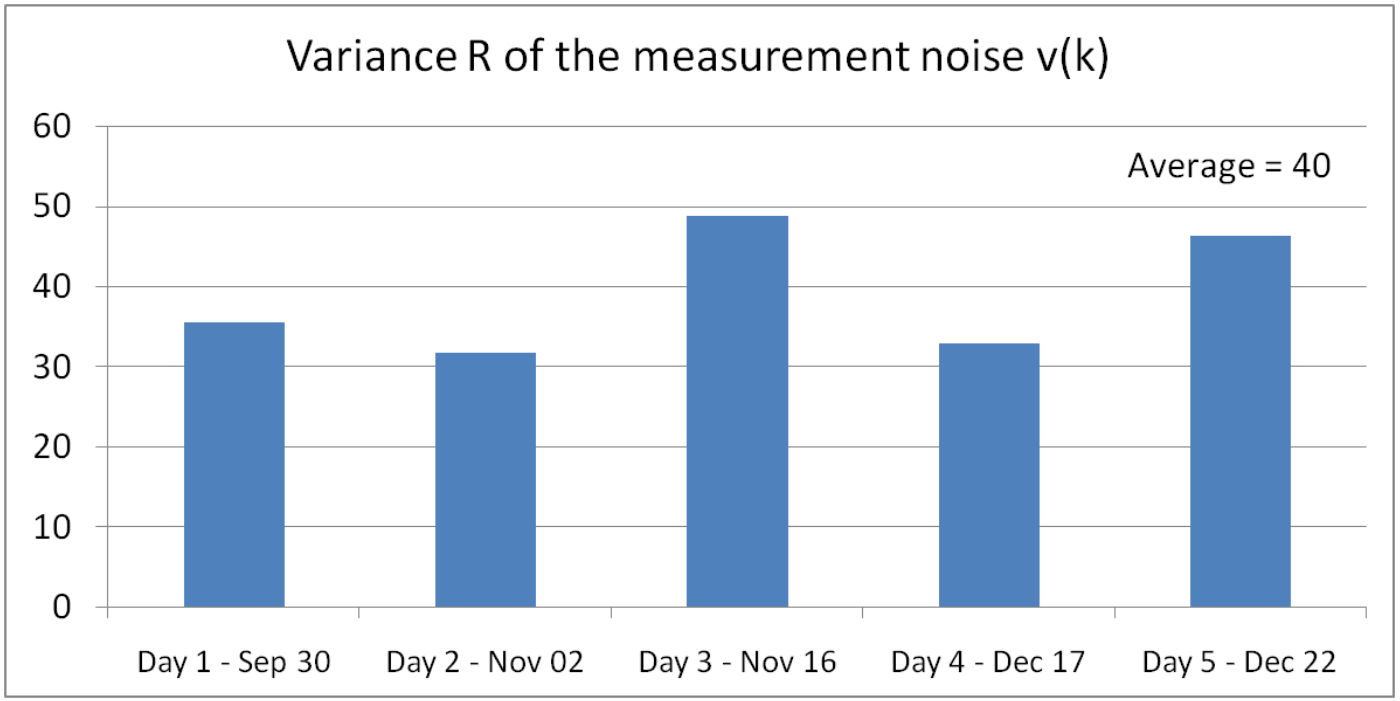
AUTO CORRELATION FUNCTION CALCULATED FOR v(k)



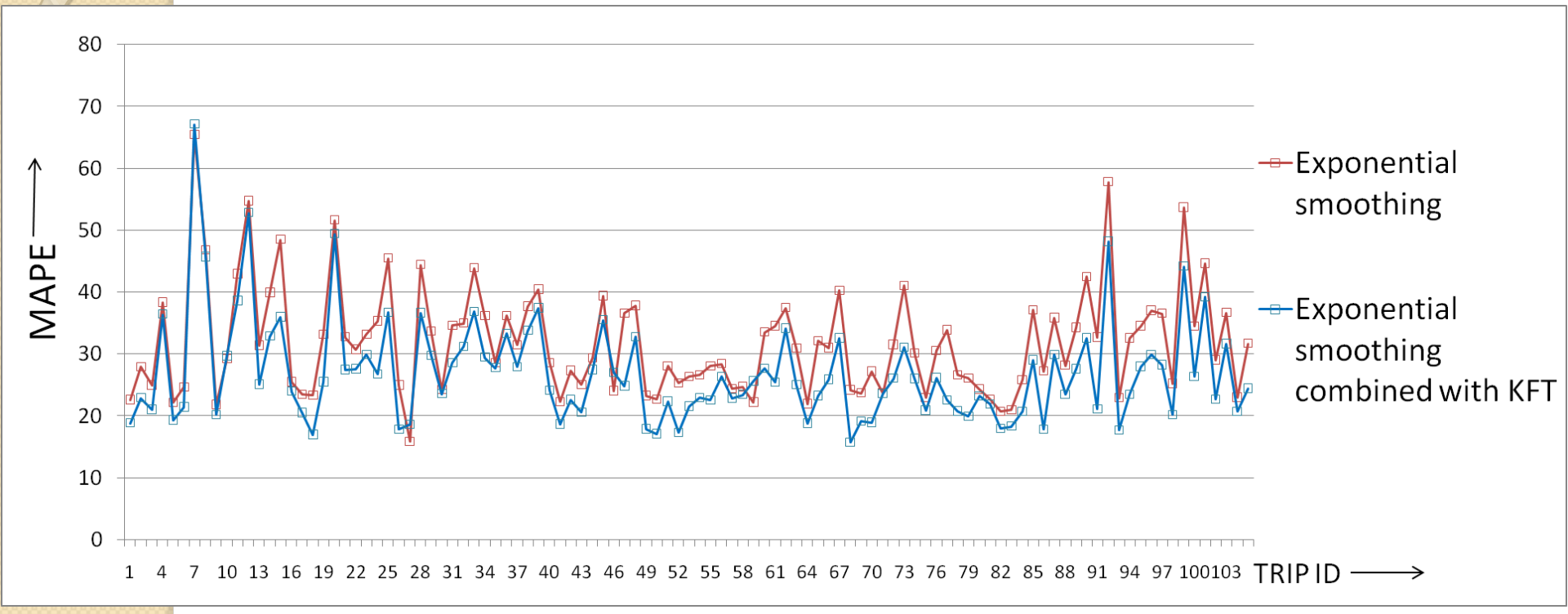

Frequency distribution of w(k)
Skewness = 0.032
45
40
35
30
25
20
15
10
5
0
-30
-25
-20
-15
-10
-5
0
5
10
15
20
25
30
More


Frequency distribution for v(k)
Skewness = -0.719
100
90
80
70
60
50
40
30
20
10
0
-10
-5
0
5
10


Variance Q of the process disturbance w(k)
250
200
150
100
50
0
Average = 140
Day 1 - Sep 30
Day 2 - Nov 02
Day 3 - Nov 16
Day 4 - Dec 17
Day 5 - Dec 22

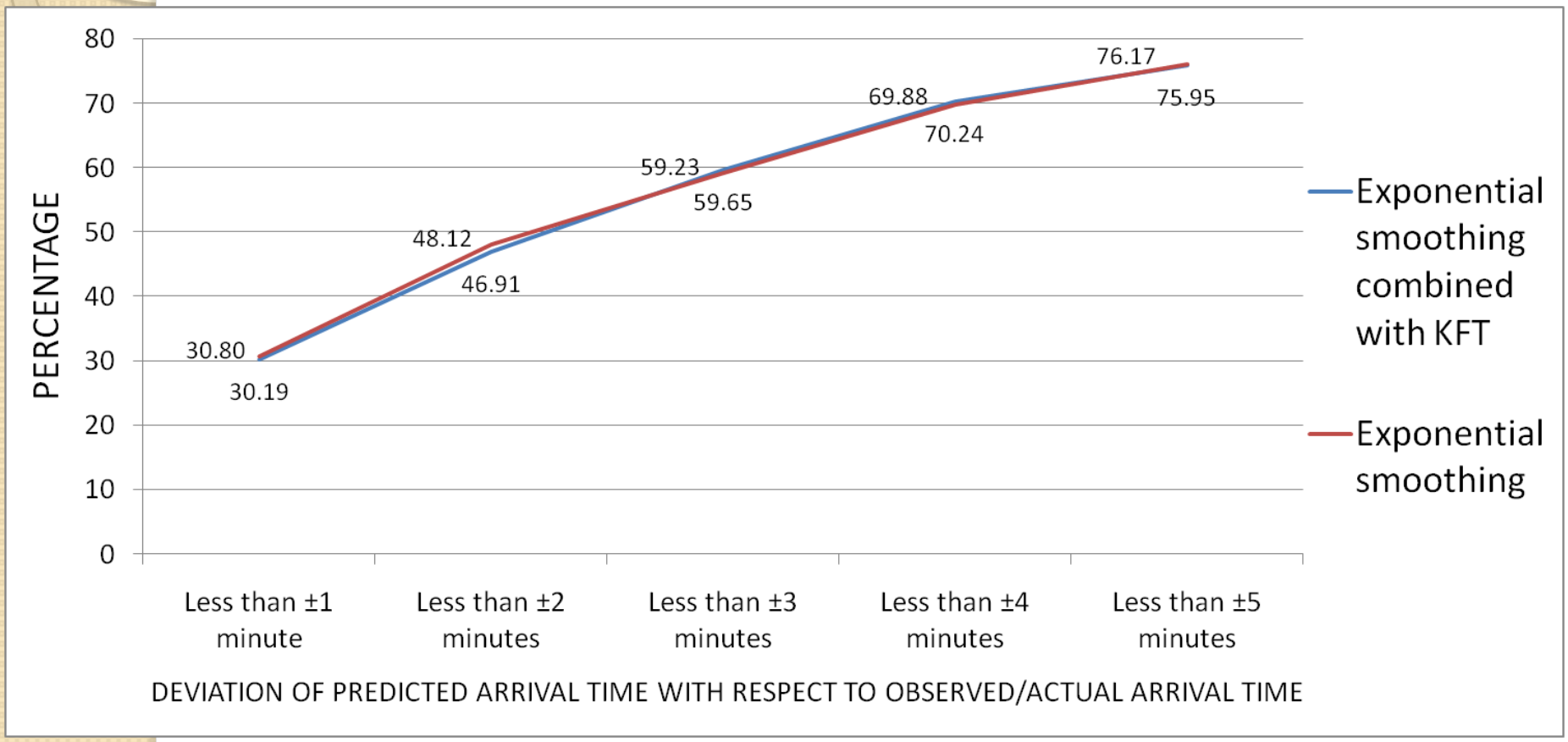

Variance R of the measurement noise v(k)
60
50
40
30
20
10
0
Average = 40
Day 1 - Sep 30
Day 2 - Nov 02
Day 3 - Nov 16
Day 4 - Dec 17
Day 5 - Dec 22

MAPE BETWEEN OBSERVED AND PREDICTED SECTION TRAVEL TIMES FOR VARIOUS TRIPS



- In 100 out of 105 trips, exponential smoothing combined with KFT method shows lesser MAPE

Number of times the predicted arrival time deviated with respect to observed/actual arrival time for various time intervals (expressed in percentage)



Concluding remarks

- Use of exponential smoothing for bus travel time/arrival time prediction
- State space form – integration with KFT
- Exponential smoothing combined with KFT shows better performance when compared to exponential smoothing

When compared to complex ARIMA models for bus travel time prediction, the proposed model provides the advantages of

- ✓ Real-time data handling,
- ✓ Easy understanding,
- ✓ Flexibility of transferring to other sites as there are no model coefficients involved,
- ✓ No dependence on specialized software,
- ✓ Limited input data.

ACKNOWLEDGEMENT

The authors acknowledge the support for this study as part of the project CIE/10-11/168/IITM/LELI by Ministry of Urban Development, Government of India, through letter No. N-11025/30/2008-UCD.

THANK YOU