香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

Forecasting the Volume of Patient Arrival to an Emergency Department in Hong Kong using Time Series Regression Method and Artificial Intelligence Method

JIANG Shancheng and CHIN, K.S.

Dept. of Systems Engineering and Engineering Management

Email:scjiang2-c@my.cityu.edu.hk

Outline

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

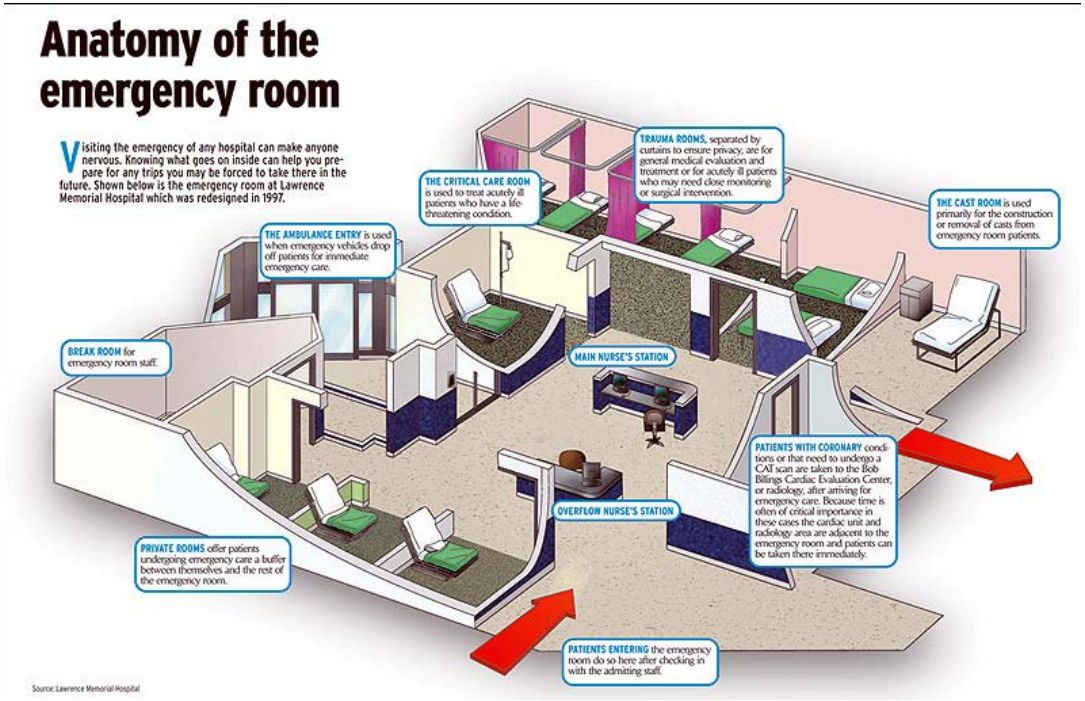
- 1.Introduction
- 2.Literature Review
- 3.Model Analysis and Comparison
- 4.Case Study
- 5.Conclusion

Outline

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

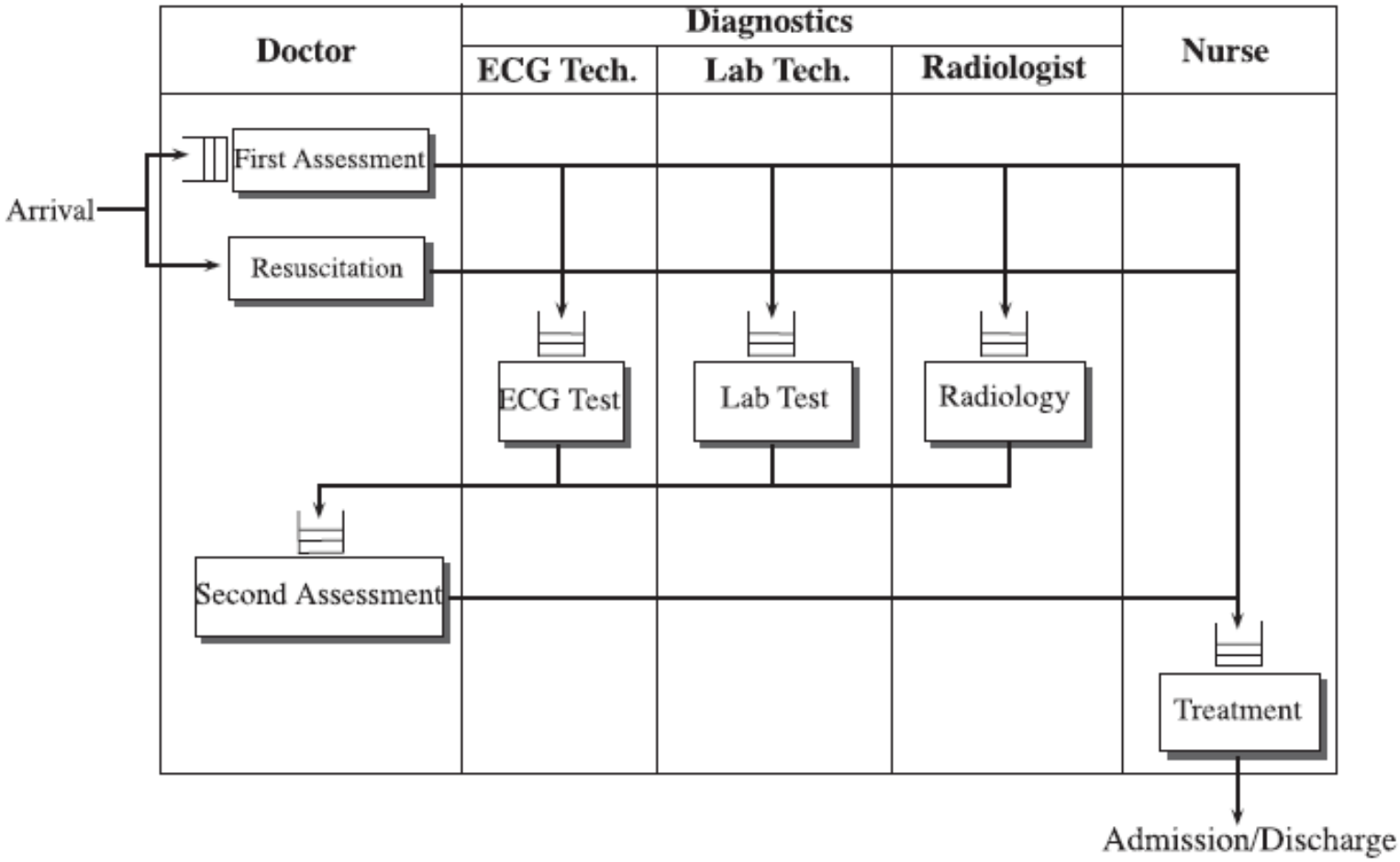
- 1.Introduction
- 2.Literature Review
- 3.Model Analysis and Comparison
- 4.Case Study
- 5.Conclusion

Frame & Resources of ED



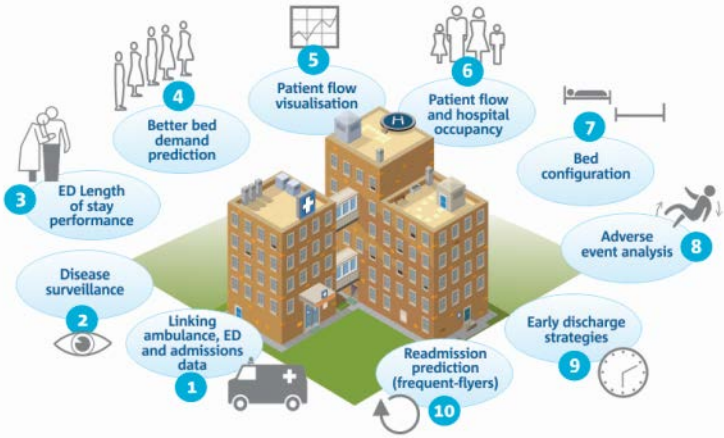
- Nurse
- Physician
- Other staff
- Trauma room
- Waiting room
- Medical equipment

Process Chart of Patients in ED



Challenges in ED

- Timeliness → Scheduling
  - Waiting time target of patients is less than hospital ward.
- Shortage → Assignment
  - Room, personnel, equipment…
- Uncertainty → Forecasting
  - Patients' arrivals without appoint
  - Need among different patient



Motivation

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

- **A series of deterministic models were proposed to describe rostering, scheduling or staffing problems in ED**
  - Valouxis, C., Gogos, C., Goulas, G., Alefragis, P., & Housos, E. (2012). A systematic two phase approach for the nurse rostering problem. European Journal of Operational Research, 219(2), 425-433.
  - Komarudin, Guerry, M., De Feyter, T., & Vanden Berghe, G. (2013). The roster quality staffing problem – A methodology for improving the roster quality by modifying the personnel structure. European Journal of Operational Research, 230(3), 551-562
- **In reality, there exist some uncertain variables, leading ED to a stochastic system.**
- **Among those uncertain variables, the volume of patient arrivals is an important prerequisite of conducting the rostering and scheduling, either in the simulation-based approach, or in mathematical programming approach.**
  - Sinreich, D., Jabali, O., & Dellaert, N. P. (2012). Reducing emergency department waiting times by adjusting work shifts considering patient visits to multiple care providers. IIE Transactions, 44(3), 163-180.
  - Erdem, E., Qu, X., & Shi, J. (2012). Rescheduling of elective patients upon the arrival of emergency patients. Decision Support Systems, 54(1), 551-563.

Objective

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

- To acquire a promising contributing variable set(e.g. meterologic variable, calendar-based variable…) based on literature review and experts' opinion.

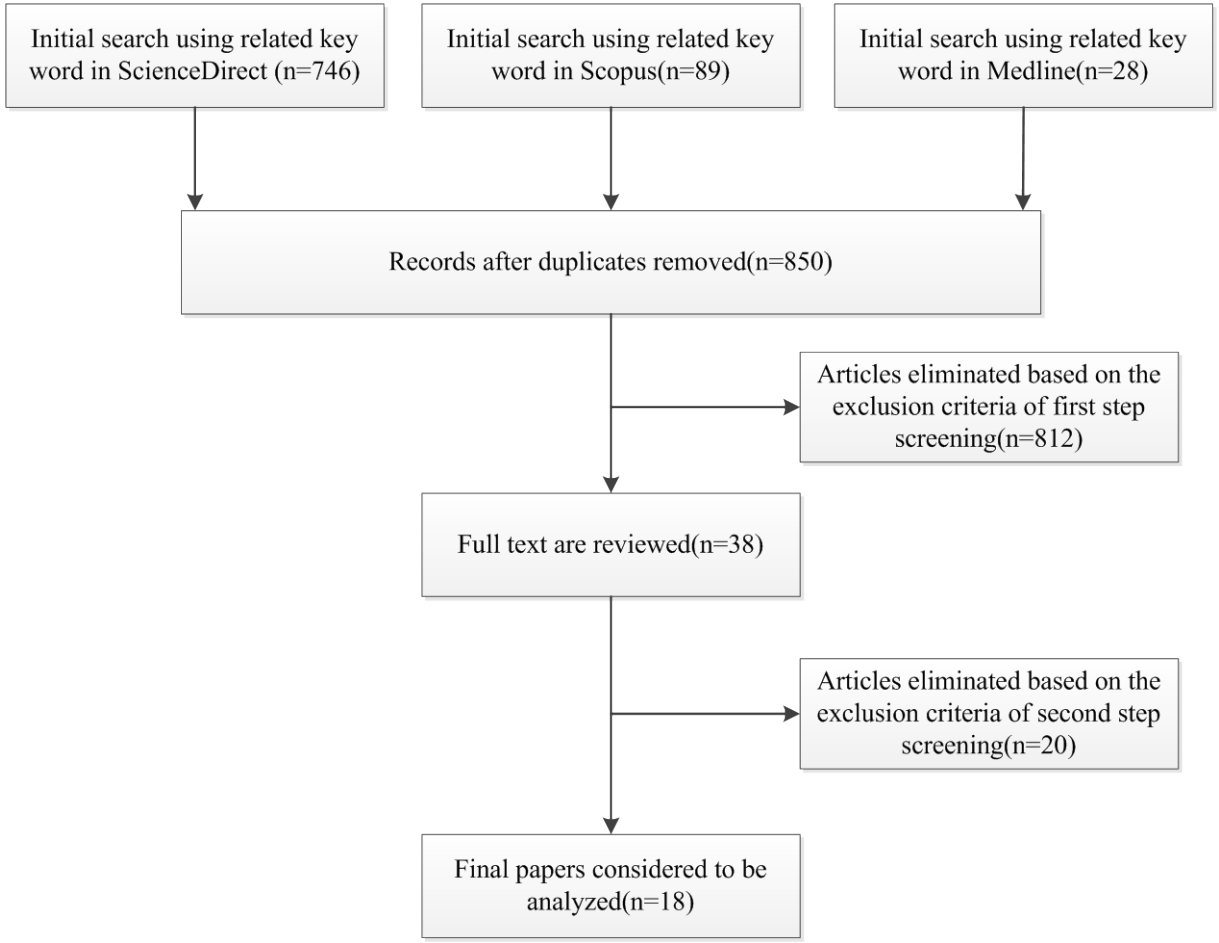
- To design a comparative case study to test the predictive power of proposed models on this particular problem.

- To find out the advantage & weakness of proposed models based on the result of case study

Outline

- 1.Introduction
- 2.Literature Review
- 3.Model Analysis and Comparison
- 4.Case Study
- 5.Conclusion

Literature Searching Strategy



Proposed Prediction Model

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

Category	Type	Number of studies	References
Regression (causal)	Multiple linear regression (MLR)	6	Jones, 2008[12]; Tai, 2007[11]; Xu, 2013[13]; Wargon, 2010[14]; Boyle, 2012[15]; Boyle, 2008[16]
	Poisson regression	2	McCarthy, 2008[17]; Marcilio, 2013[10]
	Quadratic	1	Boyle, 2008[16]
	Nonlinear least square regression	1	Xu, 2013[13]
Time series	Autoregressive integrated moving average (ARIMA)	8	Boyle, 2008[16]; Sun, 2009[9]; Jones, 2008[12]; Schweihler,2009[18]; Xu, 2011[19]; Boyle, 2012[15]; Marcilio, 2013[10]; Kadri, 2014[20]
	Exponential smoothing (ES)	4	Jones, 2008[12]; Xu, 2011[19]; Boyle, 2012[15]; Bergs, 2014[21]
	Other smoothing methods	3	Schweihler,2009[18]; Boyle, 2008[16];; Boyle, 2012[15]
Artificial intelligence	Artificial neural networks (ANN)	5	Jones, 2008[12]; Aladag, 2012[22]; Xu, 2013[13]; Moustris, 2012[23]; Menke, 2014[24]
Time series regression	Generalised estimating equations (GEE)	1	Marcilio, 2013[10]
	Multiple linear regression (MLR) with time series terms	2	Xu, 2011[19]; Jones, 2008[12]

Contributing Variable

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

- Calendar-based variable
  - All approaches demonstrate a consistent conclusion that calendar variables are significant correlated with patient arrival
  - Patient arrival shows evident seasonal pattern
- Meteorologic factors
  - Daily max and min temperature, humidity, wind speed, etc.
  - The most controversial issues for whether they impact the patient arrival and what degree of influence is.
  - **The change in temperature**, rather than the absolute value of temperature, provides a more sensitive marker for total ED patient arrival(Tai, Lee, Shih, & Chen, 2007)

Contributing Variable(Cont'd)

- Other factors
  - **Patient triage level**-Used to divided the raw data into several subsets with similar character(arrival pattern and motivation to see a doctor)
  - **Signal of influenza**-influenza epidemic can produce an abnormal peak in ED visits (Marcilio et al., 2013; Xu et al., 2011; Xu et al., 2013)
  - **Socioeconomic indicators**
    - Fluctuation in the stock market index (Chen et al., 2011)
    - Maybe produce side-effect of the accuracy of prediction(Moustris et al., 2012)

Outline

- 1.Introduction
- 2.Literature Review
- 3.Model Analysis and Comparison
- 4.Case Study
- 5.Conclusion

Selected Models to Conduct Case Study

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

- Multiple linear regression(MLR)
- Poisson Regression
- Seasonal Autoregressive Integrated Moving Average(S-ARIMA)
- Regression with ARMA Errors
- Artificial Neural Network(ANN)

Model Analysis (1)

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

- MLR
  - **Linearity** -The relationship between the input variable and the response is linear
  - **Independence** -The errors are independent with each other and have no serial correlation
  - **Homoscedasticity** -The variance of errors is constant
  - **Normality** -The errors is distributed normally
- Poisson Regression
  - **Assumption:** the response variable(i.e. arrival volume) has a Poisson distribution, and assumes the logarithm of its expected value can be modeled by a linear combination of unknown parameters

Model Analysis (2)

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

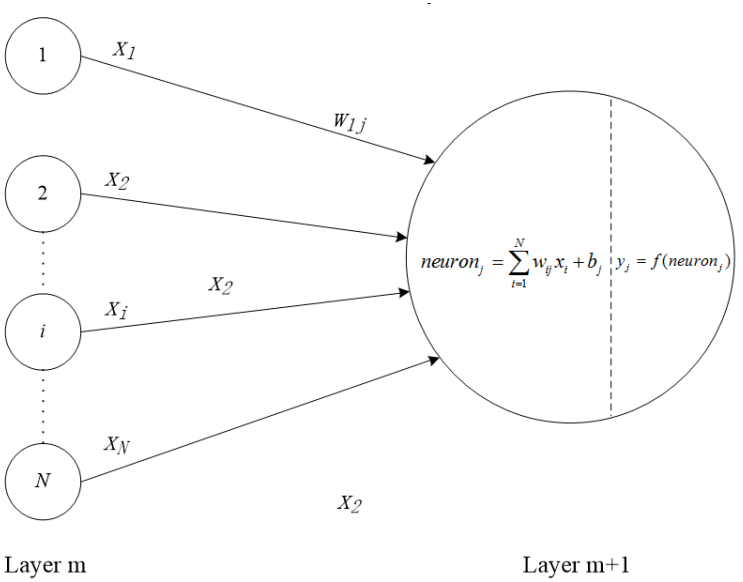
- S-ARIMA
  - **Premise:** introducing autocorrelation as a factor, and it is generated through lagged linear relations, i.e., historical trend of the time series itself.
  - **General form:** $\Phi_P(B^s)\phi(B)(1-B)_s^D(1-B)^d x_t = \Theta_Q(B^s)\theta(B)\omega_t$
    - Autoregressive term: $\phi(B) = 1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p$
    - Moving average term: $\theta(B) = 1+\theta_1 B+\theta_2 B^2+\ldots+\theta_q B^q$
- Regression with ARMA Errors
  - Combine the linear regression model and the ARMA model

$$y_t = \beta' z_t + x_t,\ where\ \Phi(B)x_t = \theta(B)\omega_t$$

Model Analysis (3)

- ANN
  - The ANN model simulated a machine that is designed to model the way in which the brain performs a particular task.

$$y_k = \tanh(\sum_k w_{kj} \tanh(\sum_i w_{ki} \tanh(\sum_j w_{ji} x_j - \theta_j) - \theta_i) - \theta_k)$$



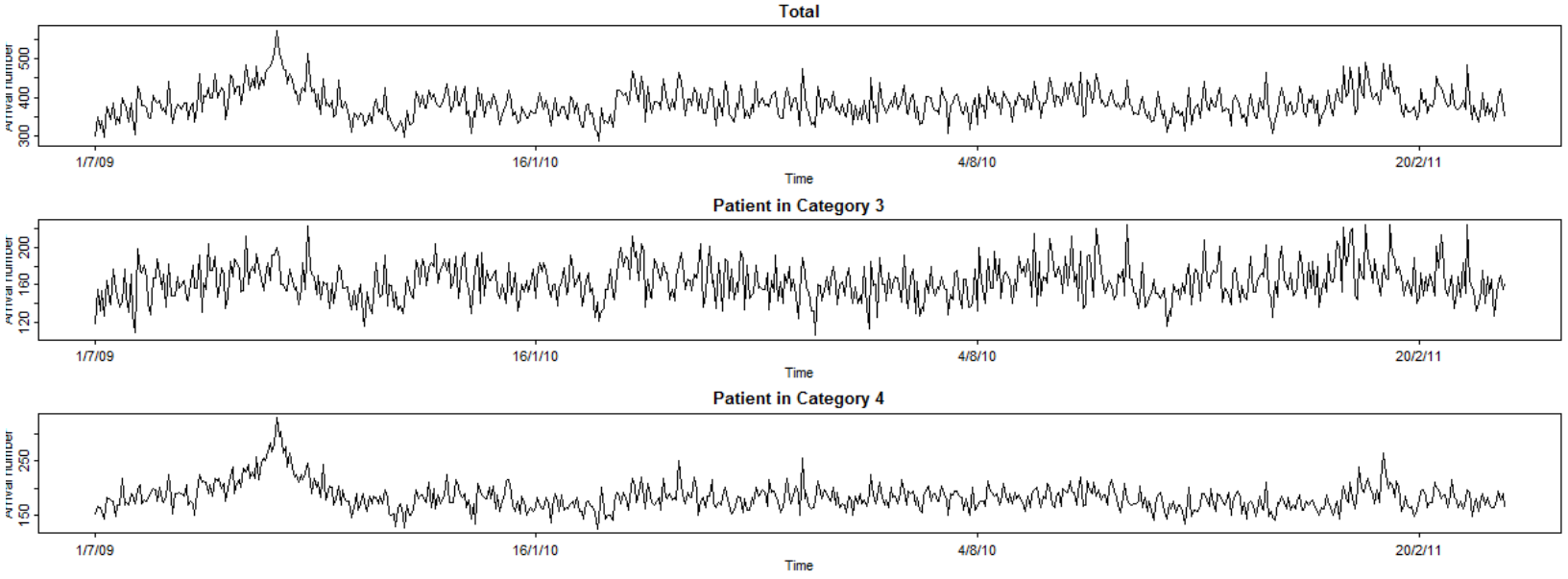
Outline

- 1.Introduction
- 2.Literature Review
- 3.Model Analysis and Comparison
- 4.Case Study
- 5.Conclusion

Data Pre-processing

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

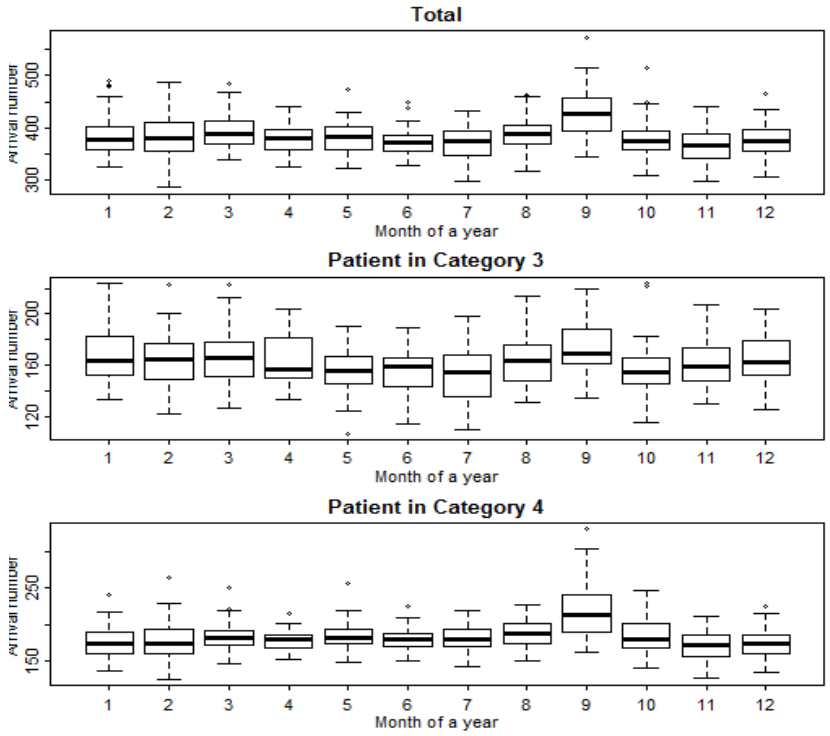
- **Administrative data**-from 1st July 2009 to 31th March 2011 in a certain ED center in Hong Kong.
- **Triage System** in Hong Kong Accident and Emergency Departments



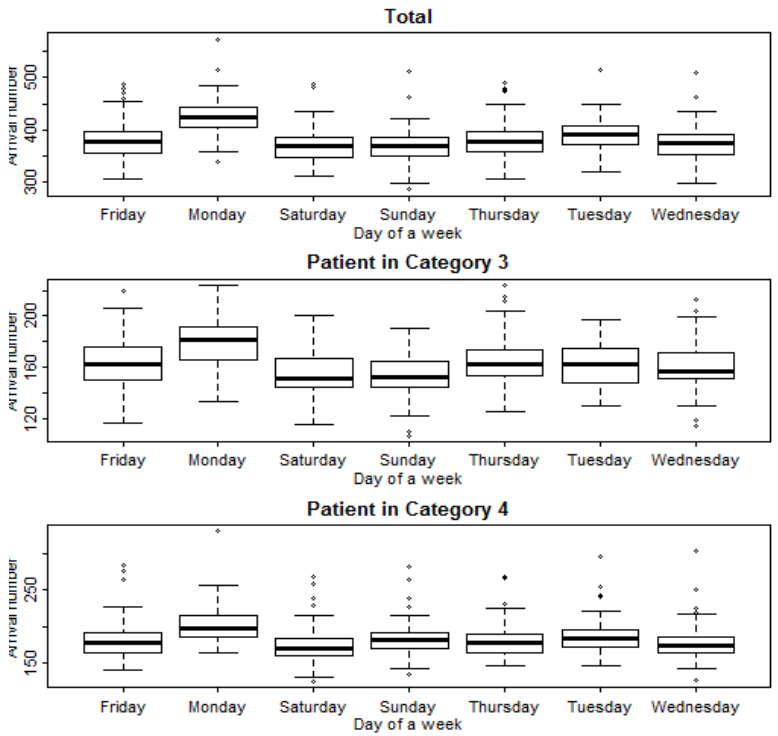
Data Pre-processing(Cont'd)

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

- Boxplot



Boxplot on month of a year



Boxplot on day of a week

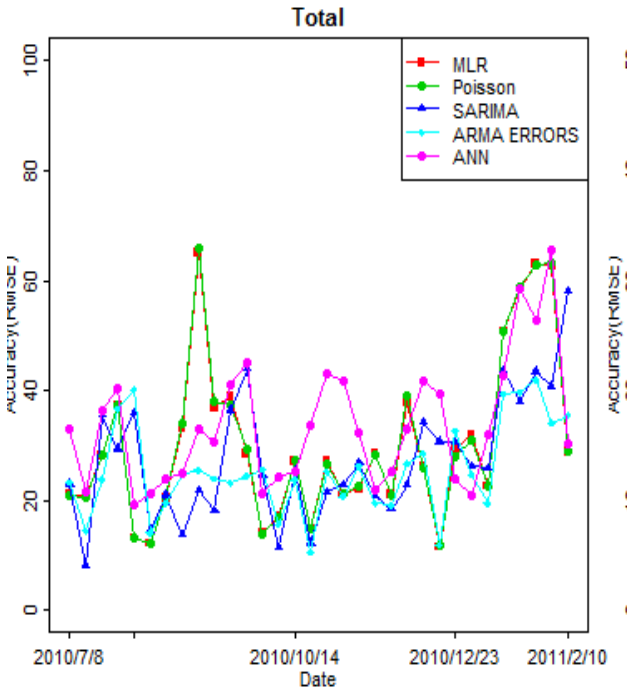
Case Study Design

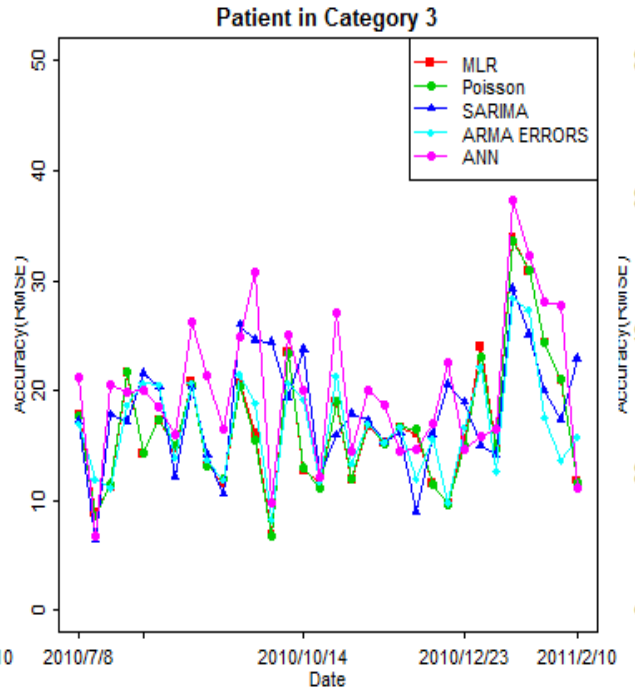
香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

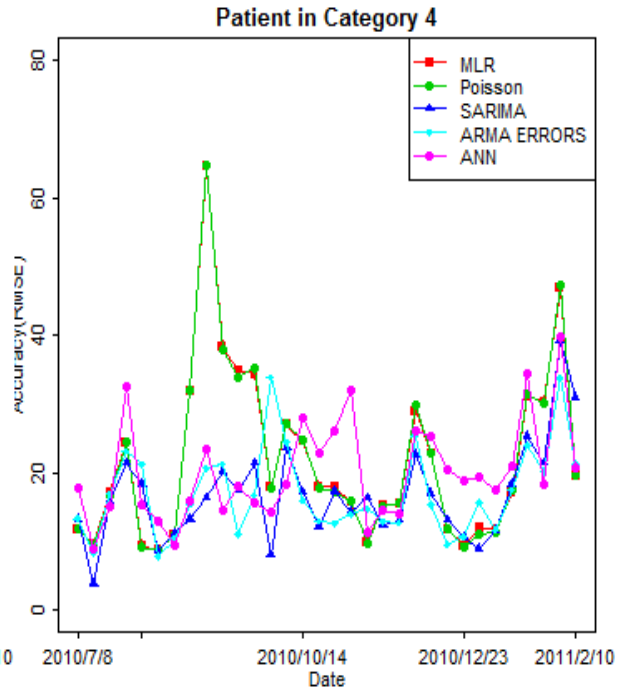
- **Training Set -** First year's data(from 1st July 2009 to 30th June 2010)
- **Testing Set** - Rest data(from 1st July 2009 to 30th June 2010)
- 3 different types of prediction horizon
  - 7days ahead, 14 days ahead and 28 days ahead
  - To test the stability of different prediction models
- Contributing Variable
  - Calendar based dummy variable
  - Air temperature, mean dew point, mean relative humidity, mean amount of cloud, total rainfall, mean wind speed
- Metric: Root Mean Square Deviation

Performance of 7 Days ahead Prediction

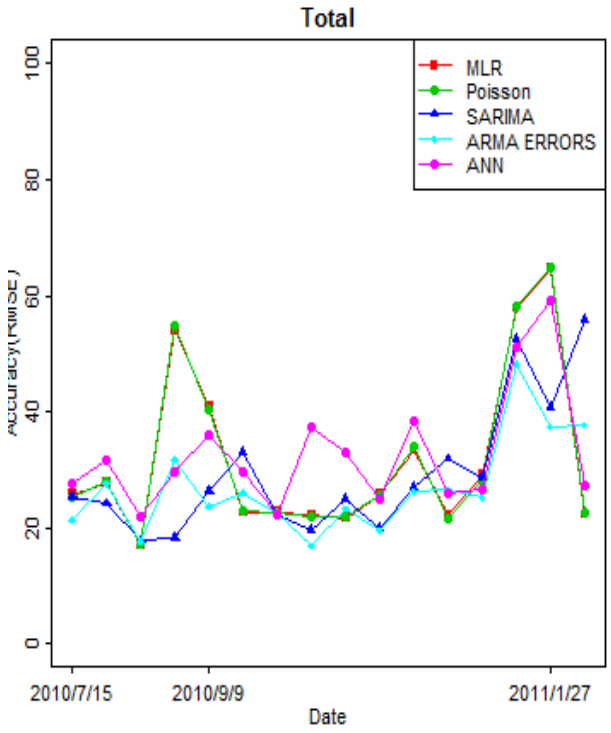
香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

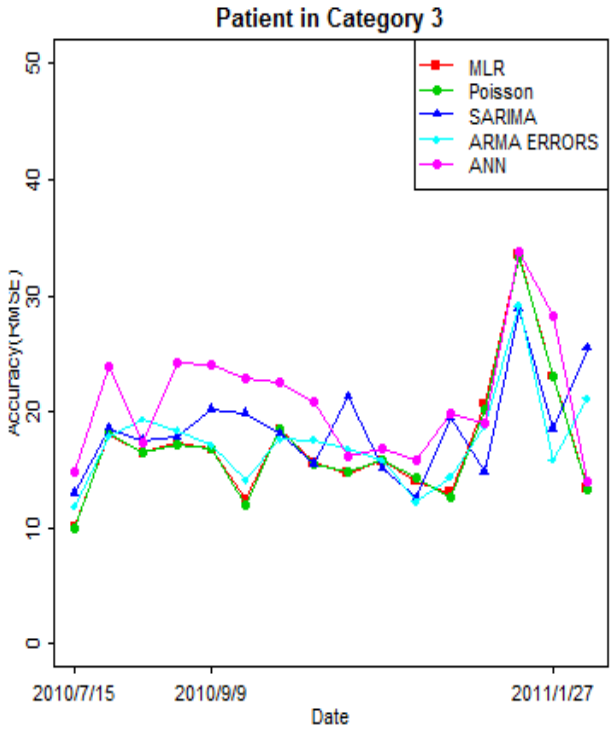


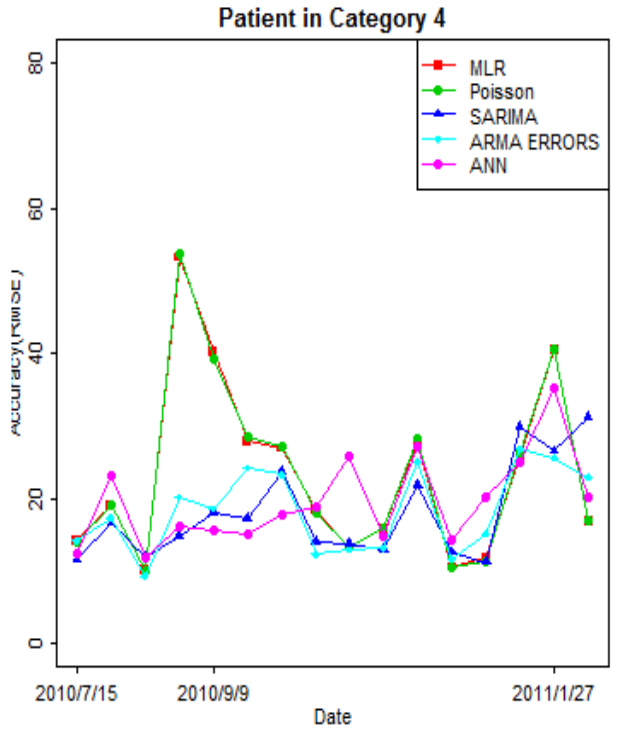




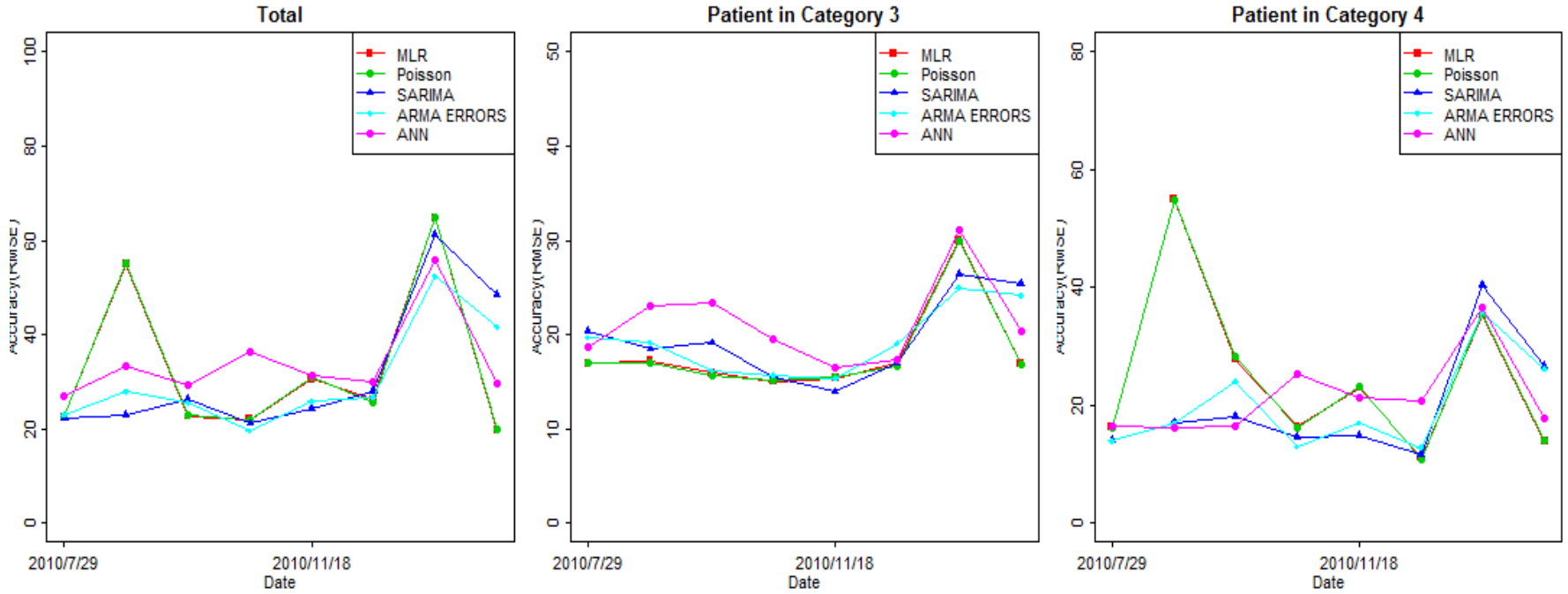
Performance of 14 Days ahead Prediction







Performance of 28 Days ahead Prediction



Discussion

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

- The S-ARIMA model and the Regression with ARMA error model perform better than other models.
  - *Patient arrival patterns are depend much on the latest historical trend.*
- The ANN model shows weaker predictive power than other models though it does have the complex and flexible structure.
  - *Over-fitting phenomenon.*
- The results of MLR and Poisson Regression are almost the same in different time periods.

Outline

- 1.Introduction
- 2.Literature Review
- 3.Model Analysis and Comparison
- 4.Case Study
- 5.Conclusion

Conclusion

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

- Although a specific prediction model does not exist that always performs well than any other kinds of models, adding ARMA terms distinctly enhances model's predictive power and adding too much nonlinear terms can bring side-effect.

- There still exists plenty of room for improvement in the topic of predicting the number of patient arrivals.
  - The meteorologic factor does affect the number of patient arrivals, while it is difficult to modeling this relationship.
  - **Complexity** and **flexibility** of the prediction model need to be improved.

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

Thanks

Q & A