# Chapter 1

# Introduction

## 1.1 Statement and Significance of the Problem

Outdoor air pollution is a serious problem threatening the health around the world. From the data from nearly 1100 cities across 91 countries, World Health Organization (WHO) found that, in many cities air pollution is reaching levels that threaten people's health, the great majority of urban populations have an average annual exposure to PM10 particles in excess of the WHO Air Quality guideline recommended maximum level of 20 μg/m$^3$. For 2008, the estimated mortality attributable to outdoor air pollution in cities amounts to 1.34 million premature deaths.

Air pollution has become a big concern of China, especially in urban areas. According to the World Bank, 16 of the world’s 20 cities with the worst air are in China. According to Chinese governmental sources, about a fifth of urban Chinese breath heavily polluted air. Air pollution problems in cities and their immediate vicinities have been and will continue to be one of the environmental concerns in the next decade in China (Chak K. Chan, Xiaohong Yao, 2008).

In the past over 30 years, China has experienced rapid economic growth with over 8% GDP growth rate annually. This unprecedented growth increase people's wealth at the cost of environmental degradation. Among which atmospheric pollution is a important one. Increased coal burning by power plants and heavy industries, increased commuting caused by urbanization, the changing of the travel mode from the cycling to private motorized modes et al. all contributed to the air pollution. In turn, serious air pollution has become threaten to health of Chinese people as well as Chinese economy. By measured the harmful effects of two air pollutants: ozone and particulates, which can lead to respiratory and cardiovascular diseases, and taking into account people's lost leisure time because of illness or death, .Kira Matus et al. (2012) found that, the estimated marginal welfare impact to the Chinese economy of air pollution levels above background levels increased from $22 billion in 1975 to $112 billion in 2005 (1997 US$).

With increasing concern about environment problems and human health, the Chinese government has begun to struggle with its major urban air quality problem. Many efforts have been made by the authorities. Since June 2000, as required by the State Environment Protection Agency of China, 42 cities report to the public the average daily air pollution index (API), a referential parameter describing air pollution levels as alerts. More and more cities have followed, and by August 2012, 115 cities have been reporting their APIs to the public. Some cities began to restrict the car ownership, policies encouraging desulfurization and low coal economy is established and implemented. Moreover, the inter-region cooperation to cut air pollution and improve air quality has been required. The three key regions required to carry out inter-region cooperation are northern China which includes Beijing, Tianjin and Hebei; the Yangtze River Delta, and Pearl River Delta. These three regions are the most developed and populated ones. Their cooperation seems more urgent, and their cooperation will provide experiences for other regions and national cooperation. As a big country with wide territory, investigating more sample cities is necessary, as it will provide more information for further research and policy decision.

Since it can be released to the public so than they can decide and adjust their activities accordingly, API forecasting is important, but existence of uncertainty impacts the accuracy of forecasting. API is affected by a series of factors like energy use (V. Kimmel, 2002), transportation (Xie, 2006), topographic features (Chu et al. 2008), wind speed and temperature (Euro Cogliani, 2001.), pressure (Chen et al. 2008, Jiang et al. 2004), much of the uncertainty may arise from data used, which again maybe based on sub-data from the observation stations, simple parameterized representation of atmospheric processes, and so on. There may also be factors omitted or ignored in the modeling, either because they are not recognized as significant, or because of incomplete knowledge. Volatility modeling can improve the forecasting by controlling and forecasting uncertainty in API alerts.

Seasonality has been one focus of previous studies, but they are mostly limited in one or two pollutants, or in dust storms impact on air quality, studies seldom focus about the API; spatially, they are limited in one region or city. National wide studies have yet to be taken, to my knowledge.

Air quality is a public good that generates trans-boundary spillovers. Good understanding of spatial contagion underpins better API modeling, forecasting and air pollution policy decision, cooperation and implementation. Spatial contagion in this study means the transmission of air pollution between city air pollution and regional/national air pollution. Even though international cooperation in air pollution control is important, as has been documented by previous studies, regional and national cooperation within one region, one country is more feasible. Considering the homogeneity within one country and one region, and heterogeneity between regions, understanding the spatial contagions of regional and national APIs is very important. But it seems that study in this field has not been carried out.

This dissertation will add to research uncertainty, seasonality and spatial contagion of air pollution in China by: applying new econometric models, and specifically, examining and comparing the predictive ability of different GARCH models with various volatility specifications in API, then focus on the tri-variate series to investigate the spatial contagion between city APIs and corresponding regional, national APIs. Further examination will research whether the seasonal effect exists in both first and second moment and how the introduction of seasonal effect affect the dynamic correlations of the local, regional and national APIs. Finally, copula methods will be used to examine the dynamic dependence between city API and corresponding regional and national levels, comparison between the results implied by DCC-GARCH model and copula methods will be made.

## 1.2 Literature Review

Air pollution problems are complicated because there are so many pollutants, and each has its source of origin, movement feature, et al. Air pollution indices of different countries are not comparable because of the difference of the base pollutant and the calculation method. The following review of selected studies in the field is organized according to the most heated field of study, air pollution forecasting, volatility study, the regional nature of air pollution, the impact of emission or dust from one region to another and seasonality of air pollution.

### 1.2.1 Air pollution studies

Early studies of air pollution problems focus on the relationship of some pollutant and input factors, and mostly focus on one single city or region(e.g. S.M. et al, 2007; Peter C. et al. 2008; Edussuriya, P., 2009). Some researchers studied the spillover of dust from one region to another (e.g. Chung-Ming Liu et al, 2006; Y.C. Lee et al, 2010).Tracy et al (2008) studied the impact of global emissions on regional air quality in Asia.

### 1.2.2 Air pollution forecasting

Many research papers on air pollution forecasting have been published. Most widely used models in this field of study are artificial neutral network based models (Uwe et al, 2006, 2003). Some tried Linear multiple partial correlation statistical method (Euro Cogliani, 2001). Other researchers used time series models. Xie and Wei used the auto-Regressive moving average (ARMA) method to forecast the API time series in different seasonal specifications and found that the ARMA model can provide reliable, satisfactory predictions for the problem interested (Xie, Wei, 2006).

### 1.2.3 Volatility studies

Instrumental in most of volatility studies has been the Generalized Autoregressive Conditional Heteroskedasticity(GARCH) family models which are widely used in finance. Although volatility clustering was documented earlier, it was not until Engle (1982) and the advent of the ARCH and GARCH (Bollerslev, 1986) models that financial econometricians started to seriously model this phenomenon. It then became a popular tool for volatility modeling and forecasting. However, despite the success of the GARCH model, it has been criticized for failing to capture asymmetric volatility. This limitation has been overcome by introducing more flexible volatility treatments by accommodating the asymmetric responses of volatility to positive and negative shocks. This more recent class of asymmetric GARCH models includes the Exponential GARCH (EGARCH) of Nelson (1991) and the threshold GARCH by Glosten, Jagannathan, and Runkle (1993) (GJR-GARCH) (Hung, Jui, 2010).

### 1.2.4 Regional nature of the air pollution

Some previous studies have surveyed the regional nature of the air pollution. Paul J. Miller et al. (1998) provides an overview of the scientific basis that supports the actions of the United States and Canada to implement regional ozone pollution control measures. F. Cousin et al. (2005) studied the synoptic and local meteorological situation in Marseille region Europe and found that ozone levels are due both to regional and local factors. By using a time series, cross-sectional panel of pollution data from Hong Kong and southern China, along with weather variables from Hong Kong and employing an econometrically-based structural model, Feng Xiao et al. (2006) conducted Granger causality tests and found that pollutants PM10 and NO2 causality runs in both directions, which confirmed the regional nature of the air pollution problem.

### 1.2.5 Inter region contagion of air pollution

Most studies focus on the impact of emission or dust from one region to another. Yongxin Zhang et al. (2010) examined the impact of tripled anthropogenic emissions from China and India over the base level (gaseous species and carbonaceous aerosols for 2000) on air quality over the U.S. The simulations indicate an extensive area of elevated pollutant concentrations spanning from the Arabian Sea to the Northern Pacific and to the Northern Atlantic. Tracey Holloway et al. (2008) studied the geographic variability and seasonality of inflow of global pollutants to Asia. They found that imported $o_3$ contributes significantly throughout Asia. In detail, both North America and Europe contribute to ground-level $o_3$ concentrations throughout the region though the seasonality of these two sources varies. Based on Aluminium and Calcium concentrations in PM10 Y.C. Lee et al. (2010) studied the transport of dusts from East Asian and non-East Asian sources to Hong Kong, and found that dust events (96%) involve non-East Asian sources. Chung-MingLiu et al. (2006) studied the influence of Asian dust storms on air quality in Taiwan. The high PM10 concentration on the east coast is associated with long-range transported dust, but the high PM10 concentration at the western side is due to a mixture of dust particles and local pollutants.

### 1.2.6 Seasonality of air pollution

Seasonality is another focus of former research. X.X. Zhang et al. (2010) found that total suspended particle (TSP) was the chief pollutant influencing Air Pollution Index (API) in northern China in spring and winter seasons. Sand and dust storm might be a major factor affecting the temporal variability and spatial distribution of TSP and dust fall in China. G. Grivas et al. (2008) analyzed PM10 concentration data collected at 8 sites over the Greater Athens. Statistical examination showed that the seasonal variation of PM10 levels was not found to be uniform across the eight sites, and significant inter-site correlations were observed. KazuyoYamaji et al. (2006) analyzed the seasonal variation of surface O3 over Japan. They concluded that surface O3 distribution over East Asia varies dynamically from season to season according to the meteorological condition, with O3 concentrations characterized by two peaks in spring and autumn and summer minimum. Guor-Cheng Fang (2010) analyzed atmospheric particulate (PM10 and PM2.5) mass concentration using hourly data collected from seven air quality monitoring stations in Taiwan, and found that the highest mean PM2.5 particle concentration was detected during spring whereas the highest mean PM2.5 particle concentration was recorded during winter. By selecting four sites covering urban, suburban, rural and coastal areas as representatives for detailed analysis, Junyu Zheng et al. (2010) analyzed Pearl River Delta (PRD) regional air quality monitoring network and investigated the characteristics of ground-level ozone in the region. Their results showed that there are distinct seasonal and diurnal cycles in ground-level ozone across the PRD region. L.Y. Chan et al. (2001) investigated the seasonal variation of different types of particulates in a fixed roadside station in heavily trafficked urban area of Hong Kong. They found that large-size particles had an apparent seasonal variation, with higher concentration level in winter and lower in summer. The dry continental winter monsoon and the wet oceanic summer monsoon are the dominating factors. On the other hand, variation of PM2.5 is much smaller since they are more affected by local traffic emission. K.L. So et al. (2007) studied long term trend and spatial variations of PM2.5 mass and chemical composition in Hong Kong covering three sites with different land-use characteristics, namely roadside, urban, and rural environments. They found that seasonal variations of PM2.5 mass concentration at the three sites were similar: higher in autumn/winter and lower in summer. SuhejlaHoti et al. (2005) analyzed the trends and volatility in atmospheric carbon dioxide concentrations levels using daily data from 1 January 1991 to 31 December 2002 collected at Ryori, Japan and Mauna Loa, USA. Their results implied that the two regions, namely Asia and the South-West Pacific, were independent in terms of the shocks to the ACDC levels. Using hourly data of PM10 concentration, Kuang-Ling Yang et al. (2002) evaluated the spatial and seasonal variations in the four regions in Taiwan, they found that spatial and seasonal variations of PM10 concentrations are rather large over Taiwan; seasonal variation is characterized by high concentrations in winter and low in summer. A.K. Gupta et al. (2008) investigated the air quality of urban region of Kolkata, India consisting of residential, commercial and industrial sites having high population density and pollution. Winter concentrations of ambient SO2, NO2, NH3 and PM10 were observed to be higher irrespective of the monitoring sites.

### 1.2.7 Copula Model

In dependence studies, DCC-GARCH model, as a conventional linear-based correlation method is somewhat restrictive due to its requirements of normality for the joint distribution and of linear relationships among variables. More flexible copula-based models have become a common practice to cope with dependence between random variables. Moreover, methods have emerged for dealing with non-normality and dependence dynamics with asymmetry over time using copula-based GARCH models(Patton 2006; Jondeauand Rockinger 2006), but were mostly focused on the financial market.

None of the studies have explored the uncertainty of city API for such a big sample. None studied the dynamic correlations of Chinese city API and regional, national levels. Previous seasonality studies focus mostly on the first moment only. The proposed study is aimed to partially fill this gap by Approximation of regional and national APIs; combining uncertainty, seasonality and spatial contagion in air pollution study based on APIs of 42 sample cities; examining the seasonality of both first and second moment of APIs; examining the dependence structure of the city air pollution and corresponding regional, national levels.

## 1.3 Principles, Models, Rationale or Hypothesis

1) ARCH coefficients of sample cities' air pollution index equal to zero;

2) GARCH coefficients of sample cities' air pollution index equal to zero;

3) The impact of bad news and good news on the conditional variance is the same;

4) The conditional correlation is constant;

5) The conditional correlation with regional, national shocks are the same for the cities in the same region;

6) The conditional correlation with regional is higher than the national levels

7) Summer, Autumn and Winter are not significantly different from spring in both mean and variance;

8) Dependence and dependence structure implied by DCC-GARCH model and GARCH-Copula method are the same.

## 1.4 Objectives of study

The objective of this thesis is to study the volatility of Beijing API, and to find a prime parsimonious model which gives the best forecast, Then, based on the first stage of study, examine time varying correlations of the city APIs with the regional and national levels of China through 42 sample cities, so as to check: the impact of shock from region and the whole nation to the local cities; to examine whether the impact is mainly from the region or the nation; the existence of region homogeneity and/or heterogeneity. Further, study the seasonality of the 42 sample cities and their influence on the dynamic correlations of the sample cities with the regional, national levels. Examine the dependence structure of city API and its corresponding regional, national level, and compare the results with what implied by the DCC-GARCH model.

## 1.5 Scope of Study

The data series for this study comprises of 42 groups of daily average Air Pollution Index (APIs) of 42 Chinese sample cities. Each group consists of three series: city API, their corresponding regional level and national API during the period from June 5th, 2000 to March 04th, 2010, that is, the data consists of 126 series, with 3560 observations each.

Data on urban air pollution index come from the data base of Ministry of Environmental Protection of the People's Republic of China (http://www.zhb.gov.cn//) (MEPPRC). The reason why we choose the 42 sample cities is that even though APIs of more cities were reported later on, say 115 cities till August 2012, the data of theses 42 cities were available since June 5th, 2000, so we have more observation; moreover, the 42 cities were scattered in the 9 regions of China, that can meet the need of the research objectives.

The data of regional and national level are integrated from APIs of the other cities within the region and nation respectively, by calculating weighted average of city APIs for all other cities in the region and in the nation based on direct distance. The regional APIs are integrated as:

$$\mathrm{API}_{ri} = \sum_{j=1}^{n} \mathrm{w}_j \mathrm{API}_j$$

$$\mathrm{w}_j = \frac{{}^{1}\!/\!_{\mathrm{D}_j^2}}{\sum_{j=1}^{n} {}^{1}\!/\!_{\mathrm{D}_j^2}} \tag{1.1.}$$

where n is the number of the sample cities in the region, which is different in different regions. $\mathrm{API}_{ri}$ is the regional API corresponding to city i, $\mathrm{API}_j$ is the API of city j in the same region of city i. $\mathrm{w}_j$ is the weight of city j, $\mathrm{D}_j^2$ is the squared distance from city j to city i.

National APIs is integrated similarly:

$$\mathrm{API}_{ni} = \sum_{k=1}^{m} \mathrm{w}_k \mathrm{API}_k$$

$$\mathrm{w}_k = \frac{{}^{1}\!/\!_{\mathrm{D}_j^2}}{\sum_{k=1}^{m} {}^{1}\!/\!_{\mathrm{D}_k^2}} \tag{1.2.}$$

where m is the number of the cities in the sample except i, which is 41 in this study. $\mathrm{API}_{ni}$ is the national API corresponding to city i, $\mathrm{API}_k$ is the API of city k in the sample. $\mathrm{w}_k$ is the weight of city k, $\mathrm{D}_k^2$ is the squared distance from city k to city i.

After descriptive statistic analysis of the data, this study will start with univariate modeling. Daily API data exhibit the same volatility clustering feature as financial data. GARCH type models have advantages in this field of reaches, but to the best of our knowledge, it has not been used in the study of urban API. Before entering the multivariate study of time varying correlations between urban

APIs and the corresponding regional, national levels, this study first will examine and compare the predictive ability of different GARCH models with various volatility specifications in APIs.

By employing the Dynamic Conditional Correlation (DCC) model introduced by Engle (2002) as a generalization of the Constant Conditional Correlation (CCC) model of Bollerslev(1990), the dissertation will study dynamic correlations between local cities, regional and national APIs. Following previous research, by introducing seasonal dummy in DCC-GARCH model this paper attempt to investigate whether the seasonal effect exists in both first and second moment and how the introduction of seasonal effect affect the dynamic correlations of data series.

Further, by employ the GARCH copula, the dissertation will examine the dependence structure of a special city(based on the previous results, among 42 sample cities) API and its corresponding regional, nation levels, comparison with the results revealed by DCC-GARCH model will be made.

## 1.6 Definitions

### 1.6.1 Air Pollution Index

Air Pollution Index (API) is a referential parameter describing air pollution levels, provides information to enhance the public awareness of air pollution. The API reporting in China requires to convert monitored daily average air quality data into integer values, and then to report to the public. In China, Shanghai is the first to report APIs dating back to June 1997. Before June2000, three major pollutants, including total suspended particulates (TSPs), sulfur dioxide (SO2) and nitrogen oxides (NOx) , were selected for API reporting. After June 2000, required by the State Environment Protection Agency of China, these pollutants were switched to respirable particulate matter (PM10), SO2 and nitrogen dioxide (NO2) (Kai, et al,2008).

### 1.6.2 Region

The spatial distribution of air pollutants is determined by meteorology, terrain, source and pollutant characteristics, factors that can vary significantly in different applications (Huang, Batterman, 2000). So the type and size of regions should depend on meteorology, terrain and weather characterization.

Based on the climate type, region geographical features, etc, there are many ways of Chinese regional divisions. Lin Chao, LuoKaifu, Huang Bingwei, RenMeie, Zhao Xueyu, Zhao Songqiao, Xi Chengfanetc suggested different schemes(Zhen Du, et al, 2005). This dissertation will employ the scheme which is based on the division of RenMeie and Yang Renzhang. There are nine major geographical regions which overlap somewhat with climate and culture region, and administrative divisions.The reason why the study start with this notion of region division is that the scale of the region is smaller,so the author expect that region homogeneity and city, region correlation are significant; moreover, cities from the same province are included in a same region so that intra-region and regional cooperation is more feasible.

Central China consists of five cities: Changsha, Hefei, Nanchang, Wuhan and Zhenzhou, which are the capital cities of Hunan, Anhui, Jiangxi, Hubei and Henan province respectively.

East China consists of nine cities. Hangzhou, Wenzhou Zhejiang province; Jinan Yantai and Qingdao in Shandong province; Nanjing, Nantong and Suzhou in Jiangsu province and Shanghai.

North China consists of five cities: Beijing; Huhehaote, Inner-Mongolia; Shijiazhuang and Tianjin, Hebei province and Taiyuan, Shanxi Province.

Northeast includes four cities: Changchun, Jilin province; Dalian and Shenyang, Liaoning province; Haerbin, Heilongjiang province.

Northwest includes four cities: Lanzhou, Gansu province; Wulumuqi, Xinjiang province; Xian, Shannxi province and Yinchuan, Ningxia province.

Qinghai-Tibet consists of Lasa and Xining.

South China consists of five cities: Guangzhou, Shenzhen, Zhuhai and Zhanjiang, Guangdong province and Haikou, Hainan province.

Southeast consists of three cities, which are in the same province, Fujian province: Fuzhou, Shantou and Xiamen.

Southwest consists of five cities: Chengdu, Sichuan province; Chongqing; Guiyang, Guizhou province; Kunming, Yunnan province and

Nanning, Guangxi province.

### 1.6.3 Season

There are two notions of season. First, a calendar year is divided into two seasons: dust season and non-dust season. According to the dust storm record of Beijing, dust season consists of three months: March, April and May, the left nine months are included in non-dust season. Another notion is to divide one calendar year into four seasons: spring includes March, April and May; summer includes June, July and August; autumn includes September, October and November; winter includes December, January and February.

## 1.7 Overview

Since air pollution can transport across regions, cross boundary cooperation in air pollution monitor and control is very important for air pollution elimination and health protection. Some countries have signed agreements of actual cooperation. For example, Canada and the United States signed Air Quality Agreement on 13 March 1991; on July 6, 2011, the US Environmental Protection Agency (EPA) finalized Cross-State Air Pollution Rule. Understanding air pollution spatial contagion and correlation is important for air pollution forecasting, control and cooperation.

China, the most rapidly developing economy, with the second largest population, faces serious air pollution problem and health threaten. According to the World Bank 16 of the world's 20 cities with the worst air are in China. According to Chinese government sources, about a fifth of urban Chinese breath heavily polluted air. Only a third of the 340 Chinese cities that are monitored meet China's own pollution standards. Many cross boundary air pollution cooperation are under implementation, most are within one province. The very first regions implementing inter region cooperation are Beijing-Tianjin-Hebei region, the Yangzi River Delta, the Pearl River Delta and Hongkong-Guangdong region. There are some studies examining the cross boundary of air pollution in China, but API based spatial volatility spillover studies covering China are not noticed.

The object of this study is to understand the uncertainty, seasonality and the dynamic correlations of Chinese city API and regional, national levels. Specifically, the dissertation tries to answer these questions: are GARCH type models feasible in studying air pollution uncertainty? Does asymmetry effect exist in air pollution index volatility? Which model gives the best forecast? Does the air pollution in China exhibit region homogeneity, nation homogeneity, inter-region heterogeneity? Whether the impact of new information on a city comes mostly from regional or national level? What's the seasonality of APIs of the cities in different regions? Does seasonality consideration change the feature of spatial contagion of city APIs of China? Do results from GARCH copula tell a story different from what DCC-GARCH?

The dissertation is organized as follows: chapter 2 introduces the methodology of this study. This study starts by univariate modeling, Chapter 3 models the uncertainty of Beijing API by introducing the econometric models widely used in financial econometrics in this field. In particular, this chapter focus on three aspects: comparing the estimation and forecast performance of GARCH, GJR-GARCH, EGARCH and GARCH-M models; examine the seasonal dust effect of data, and the existence of asymmetry in the data. With model diagnostic criteria, EGARCH out performs other models, while out-of-data static forecast performance does not.

In order to understand the spatial relationship of urban air pollution in China, the study goes further into the multivariate analysis. Chapter 4 focus on the seasonality of the first and second moment of the daily APIs of 42 Chinese sample cities over 10 years, from June 5th, 2000 to March 04th, 2010, and investigate the dynamic correlation of air pollution Indices (APIs) between 42 Chinese cities and their corresponding regional and national levels, comparison with the model without seasonal consideration is made.

DCC-GARCH estimation results reveal that i) the correlations of local APIs between regional and national levels are time varying; ii) most cities exhibit positive conditional correlations with both regional and national shocks; iii) the conditional correlations of most cities with regional and national shocks are only slightly different; iv) cities in East China, North China and Southeast exhibit high correlation with both regional and national shocks, while the cities in vast territory and special terrain are not very sensitive to shocks of regional and national APIs; v) transformed DCC-GARCH model including seasonality dummies improve the estimation result in this study; vi) seasonality feature of the second moment follow that of the first moment, with the condition mean and variance of the second and autumn significantly lower than spring, whereas that of winter is higher than spring; vii) compare with DCC-GARCH model estimation, the transformed model does not change the feature of the dynamic correlations very much.

Based on the previous study of previous section, chapter 5 investigates the dependence structure between the Air Pollution Index (API) of Shenzhen and corresponding regional, national levels based on copula based GARCH models. In particular, time varying normal copula and time varying SJC copula are compared and employed to model the dependence structure. Comparison with the results of DCC-GARCH model is made. We find that there exists significant asymmetric upper and lower tail dependence between Shenzhen and regional, national levels; tail dependence captures the change in dependence better; dependence structure change across time. Followed by conclusions in chapter 6.

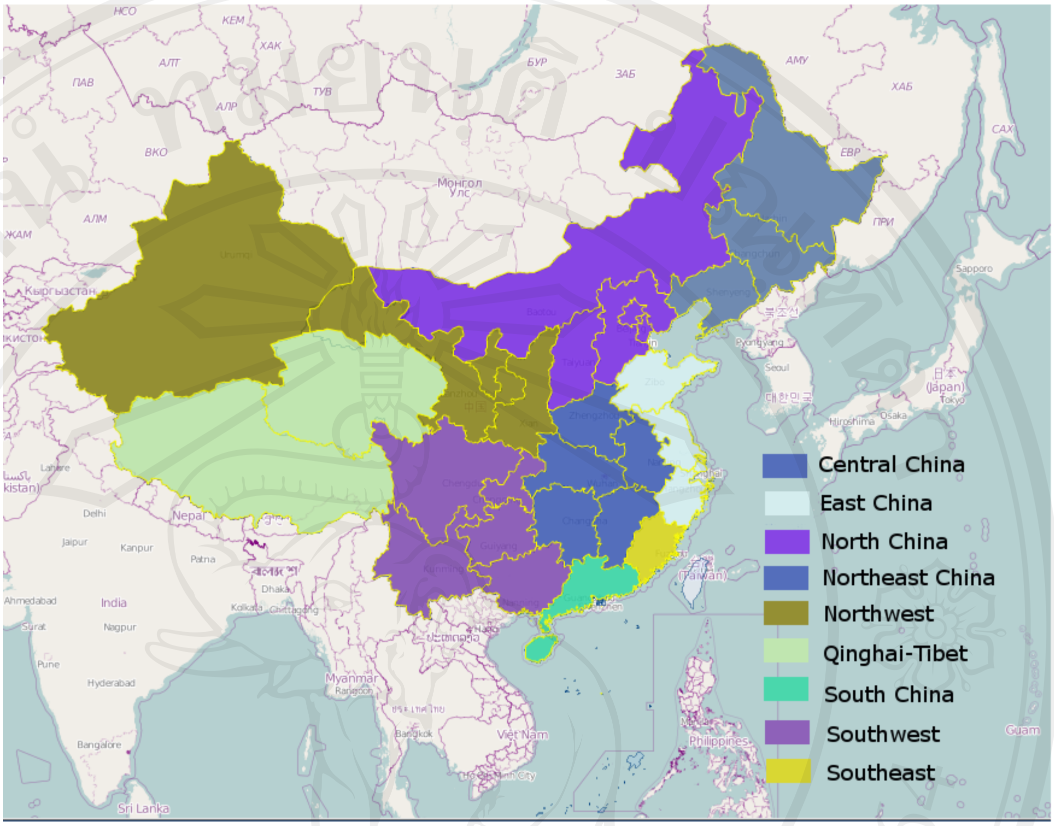

Figure 1-1 Regions of China Mainland