CHAPTER 4

Analysis of solar radiation absorption

This chapter describes the data concerning solar radiation observations. In this study, many statistical models are used to analyze this data. This section also presents the analysis results from these models and discusses the spatial and temporal patterns of solar radiation and estimated cloud cover. The results in the chapter also appeared in Cheung *et al*. (2014).

The amount of solar radiation actually incident at the top of the atmosphere, daily-extra terrestrial radiation (ETR), depends on time of year, time of day and latitude which controls the obliquity of the solar beam. Only 25% of solar radiation, which represents the maximum possible short-wave radiation receipt at any point in the Earth-atmosphere system. Components of global solar radiation, surface solar radiation comprises direct and diffuse radiation. These combine into global radiation using formula $R_G = R_D + R_B\cos(\zeta)$ where R_D is diffuse and R_B is direct radiation, and ζ is the Sun's Zenith. Differences between levels of ETR and Global radiation (R_G) indicate the amount of solar energy absorbed by the atmosphere. In this study, R_G was the observation values at the stations. Assuming that the Earth is a sphere rotation around its axis with constant declination to its elliptical orbit around the Sun, the formula for extra-terrestrial solar radiation (ETR) above the Earth's atmosphere is the following (Klein 1977);

$$\mathrm{ETR} = (24/\pi)K(1+0.033\cos(2\pi d/365))(\cos\phi\cos\delta\sin\omega + \omega\sin\phi\sin\delta)$$

In this formula, ϕ is latitude angle, d is the day of the year, δ is the Earth's declination $23.45(\pi/360)\sin(2\pi(284+d)/365)$, ω is sunset hour angle given by $\omega = \mathrm{acos}(-\tan\phi\tan\delta)$, and K is the solar constant equal 4.925 MJ/hr.m^2, about 1,360 watts per square meter, according to measurements made by NASA (The National Aeronautics and Space Administration) satellite missions (Lindsey, 2009). The solar constant is the average energy radiated by the Sun per time unit on a unitary surface situated outside the Earth's atmosphere and perpendicular to the Sun's rays (Lorenzini *et al*., 2010).

4.1 Preliminary data analysis

The extra-terrestrial solar radiation (ETR) is reduced first by the upper atmosphere about 20% of ETR. The remaining 80% of ETR is the maximum solar energy level and the observations for each day of the year during 1990-2012 for 8 selected stations are shown in Figure 4.1

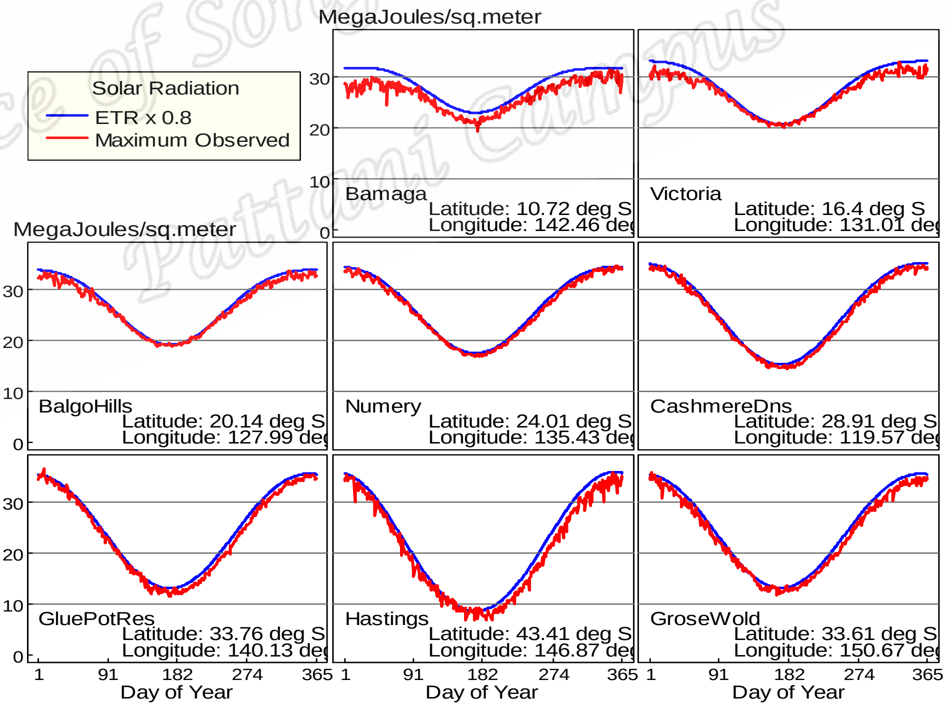

Figure 4.1 Plots of the observed maximum solar radiation and 0.8 times the extra-terrestrial solar radiation for eight selected stations

The eight example stations in Figure 4.1 represent 144 stations which are located in different latitudes and longitudes covering all parts in Australia. These graphs show the values of observations at ground stations are not higher than 0.8 times the extra-terrestrial solar radiation (above line).

Our problem is reducing the serial correlation between 365 daily observations in each year. One way to partially rectify this problem is to take the five-day averages for each station. These five-day averages are called the period solar radiation absorption data in the sequel. The number of the period data is 73 in each year. Missing data will also be reduced substantially as a result.

4.2 Studies completed of solar radiation absorption

To consider for data observed from a given weather station, linear model was taken in the form Rc = factor(year) + factor(season). For station k at year t, let y_{tsk} be the percentage of solar absorption observation for period s. The first linear regression model is, for $t = 1, \dots, 23$ and $s = 1, \dots, 73$,

$$y_{tsk} = \alpha_{0k} + \sum_{i=1}^{22} \beta_{ik} x_{ti} + \sum_{j=1}^{27} \gamma_{jk} z_{sj} + \epsilon_{tsk} \quad (4.1)$$

where ϵ_{tsk} are independent error terms following a normal distribution with mean 0 and variance σ^2, x_{ti} are indicator covariates for comparing years t and i and z_{sj} are indicator covariates for comparing periods s and j. Therefore, x_{ti} takes value 1 if $t = i$ and 0 otherwise and z_{sj} takes value 1 if $s = j$ and 0 otherwise. In (4.1), year 1990 and period 1 are used as the baseline year and the baseline period respectively. For station k, α_{0k} represents the average period solar radiation absorption corresponding to the baseline year and the baseline period, β_{ik} is the year i effect and γ_{jk} the period j effect. For each station, the five-day averages of the percentage of solar radiation absorbed by clouds (Rc) were fitted by using model (4.1) as shown in Figure 4.2.

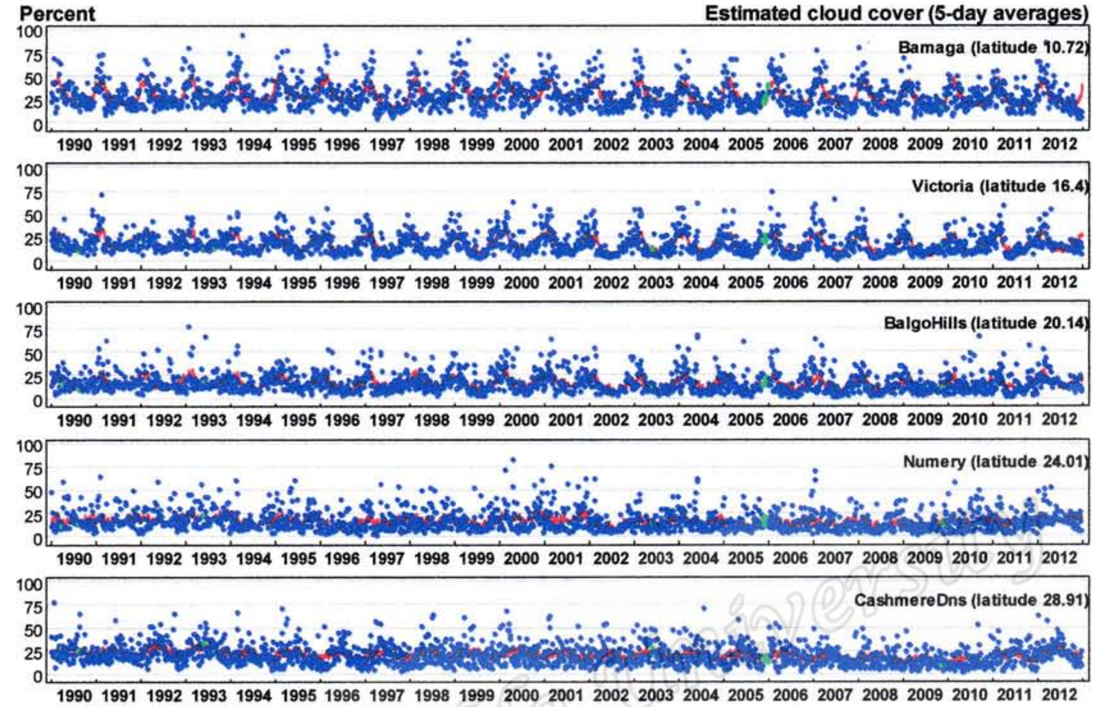

Figure 4.2 Five patterns of solar radiation absorption and their mean estimation for 5 stations

Figure 4.2 presents the estimated cloud cover of five stations over the year 1990 – 2012. The amount of solar energy reaching the Earth depends on latitude angle and the day of the year.

The errors from the fitted models were then assessed. The normality assumption requires that errors be normally distributed and this assumption was assessed by plotting residuals against normal quantiles after a model is fitted.

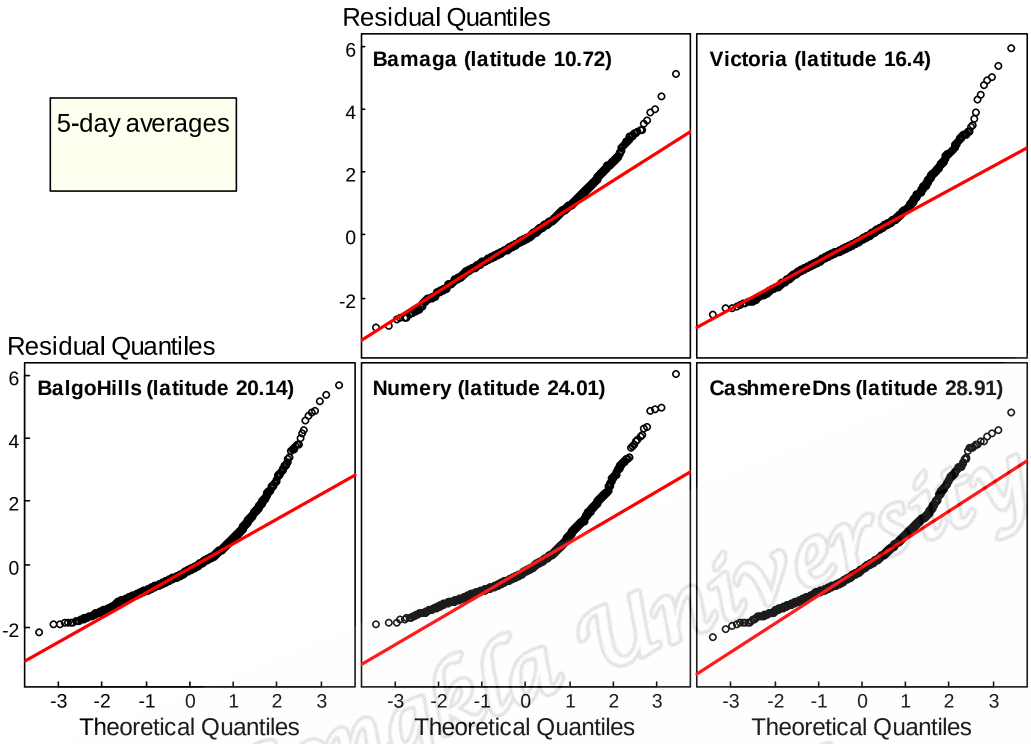

Figure 4.3 Residual quantile plots for five selected stations, where the response of the additive model is a five-day average of the daily percentage solar radiation absorption

To assess independence assumption in the errors from the fitted models, the residuals were non- normally distributed. The straight line corresponds to an exact fit of the data to a normal distribution but the residuals stretched up and down from the normal line. Thus Rc was transformed to Rc*, $Rc^* = 10\sqrt{|Rc|}$, and removed auto-correlations (filtered data) by the first order autoregressive process, the result was shown in Figure 4.4.

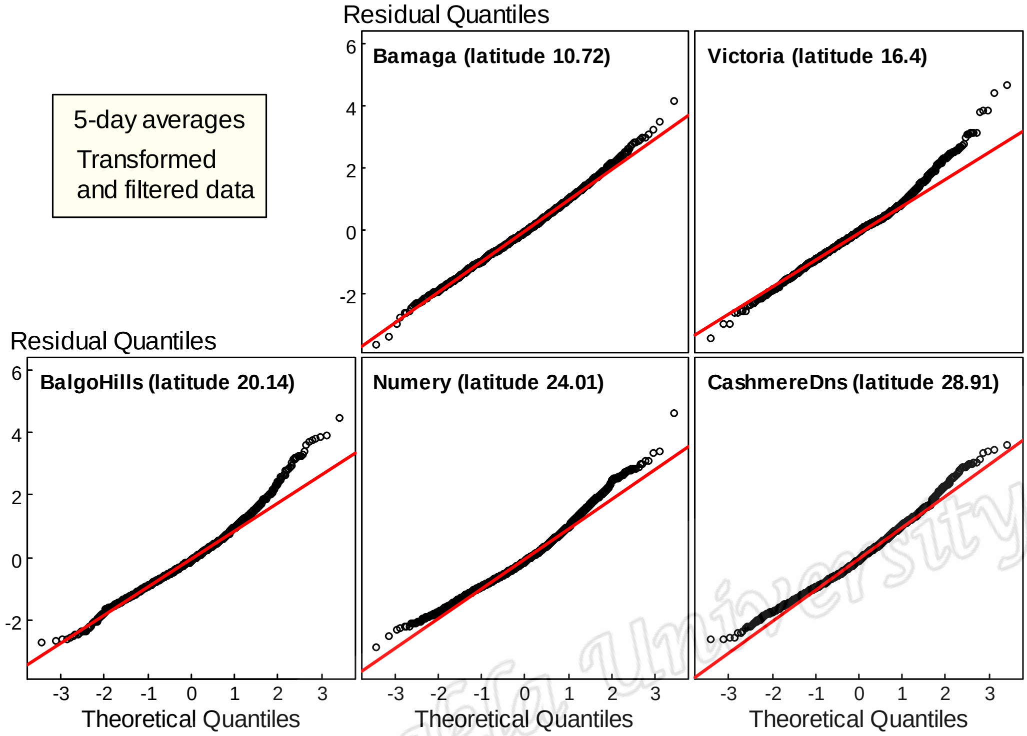

Figure 4.4 Residual quantile plots for the same five selected stations, where the response is location dependent square root transformed five-day average of the daily percentage solar radiation absorption, and then filtered with a lag 1 term

Figure 4.4 shows that the normality assumption is reasonable, then statistical model (4.2) is used to analyse the period solar radiation absorption.

$$g(y_{tsk)} = \alpha_{0k} + \sum_{i=1}^{22} \beta_{ik} x_{ti} + \sum_{j=1}^{72} \gamma_{jk} z_{sj} + g(y_{t(s-1)k}) + \epsilon_{tsk} \qquad (4.2)$$

Where $g(y_{tsk})$ denotes the transformation function of y_{tsk}, $\epsilon_{tsk}$are independent error terms following a normal distribution with mean 0 and variance σ^2, x_{ti} are indicator covariates for comparing years t and i and z_{sj} are indicator covariates for comparing periods s and j. Therefore, x_{ti} takes value 1 if $t = i$ and 0 otherwise and z_{sj} takes value 1 if $s = j$ and 0 otherwise.

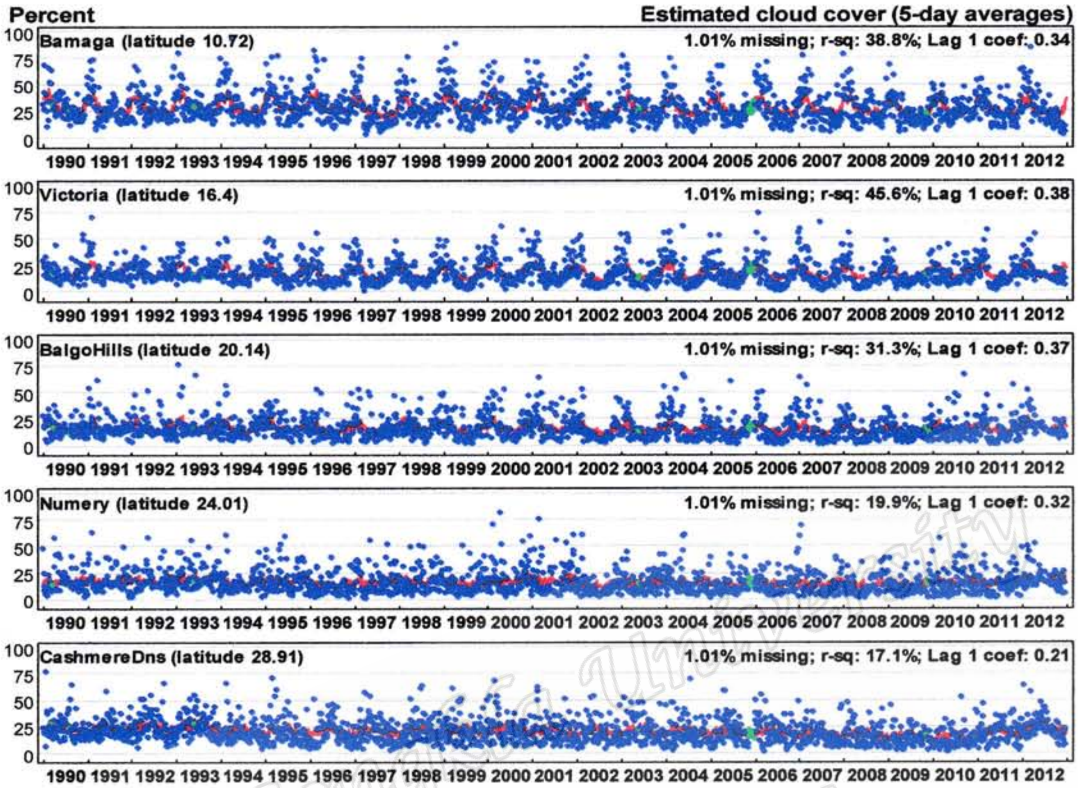

Figure 4.5 Solar radiation absorption dots and their mean estimation for 5 stations with latitudes between 10 °S and 30°S based on model (4.2)

The transformation function $g(\cdot)$ in (4.2) is defined as: $g(y_{tsk}) = 10 \times \sqrt{y_{tsk}}$ if station k has latitude less than 30 degrees; otherwise $g(y_{tsk}) = y_{tsk}$. Another issue for this type of data is possible dependence among the response observations, which violates the independence assumption. Here number 10 is used to multiply square root of y_{tsk} so that the scale of $g(y_{tsk})$ is approximately the same as y_{tsk}, and therefore different plots involving $g(y_{tsk})$ and y_{tsk} are directly comparable.

The third model is also a linear model, but its response observations are generated from a factor analysis and the explanatory variables are the year and period factors. The criterion we use to classify these factor groupings is based on the values of the factor loadings greater than 0.333 with each factor.

Let u_{tsk} be the average period radiation absorption for factor analysis based region k at year t and month s, then the model is

$$h(u_{tsk}) = \delta_{0k} + \sum_{i=1}^{22} \eta_{ik} x_{ti} + \sum_{j=1}^{72} \xi_{jk} w_{sj} + h\big(u_{t(s-1)k}\big) + \epsilon_{tsk} \qquad (4.3)$$

where x_{ti} are indicator covariates for comparing year t with year i, w_{sj} are indicator covariates for comparing period s with period j, region k (k= 1,2…7), In (4.3), function $h(\cdot)$ is defined as $h(u_{tsk}) = 10 \times \sqrt{u_{tsk}}$ for $k \neq 2$ and $\epsilon_{tsk}$are independent errors following normal a distribution with mean 0 and variance σ^2. Due to the stations in the region k = 2 (South-East Australia) located in latitude more than 30 degrees, the type of data is possible dependence among the response observations. We successfully remove dependence by including a lag 1 term (equivalent to 5-day period) of the response variable in the model, evidenced by the corresponding auto-correlation function plots displayed in Figure 4.7.

Figure 4.6-4.7 show auto-correlation function plots before and after transformation. The normality and independence assumptions on ϵ_{tsk} are verified.

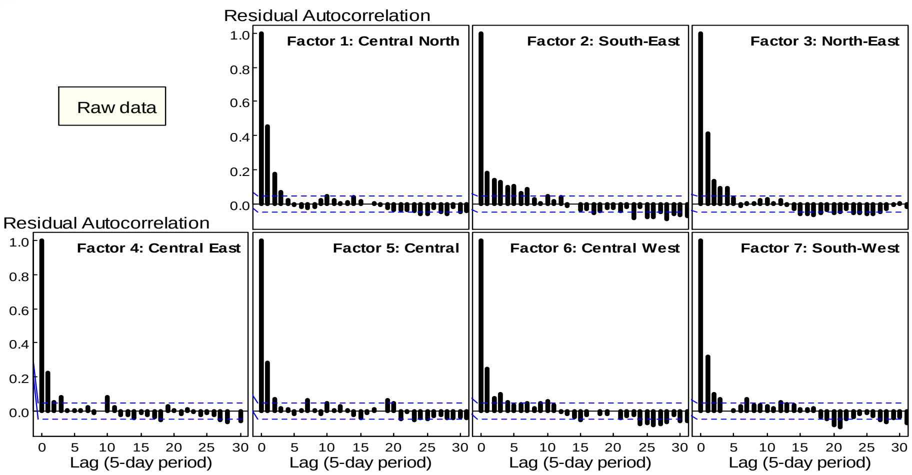

Figure 4.6 Auto-correlation function plots of residual from raw data and no filtering

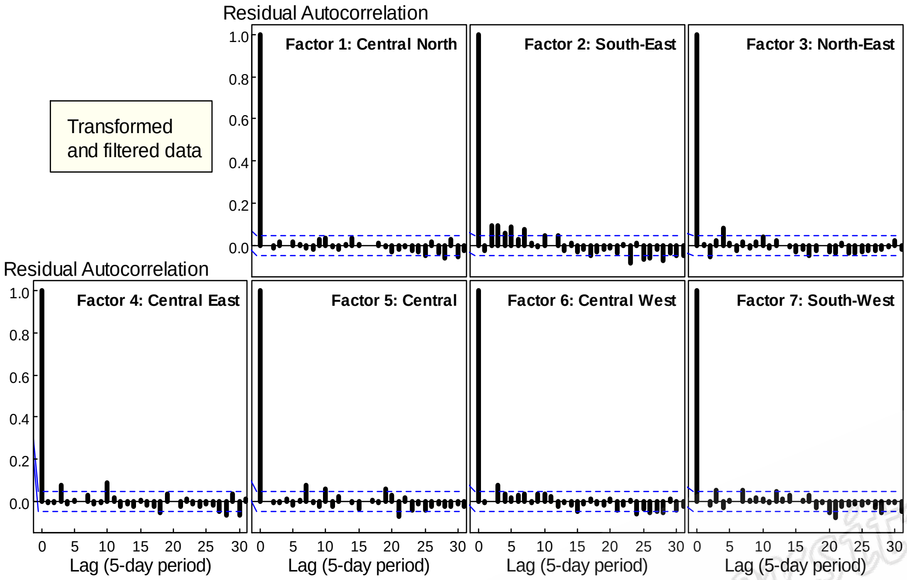

Figure 4.7 Auto-correlation function plots of residual from transformed and filtered data

Factor analysis was applied to identify correlations between solar radiation absorption, which is related to the estimation of the monthly variation of the percentage of solar radiation energy absorbed by clouds and other bodies in the lower atmosphere. This result divides 143 (of 144) stations into 7 regions geographically : Central North (CN, Factor 1), South East (SE, factor 2), North East (NE, factor 3), Central East (CE, factor 4), Central (C, factor 5), Central West (CW, factor 6) and North West (NW, factor 7). There is one station which is not correlated to the rest of the other stations.

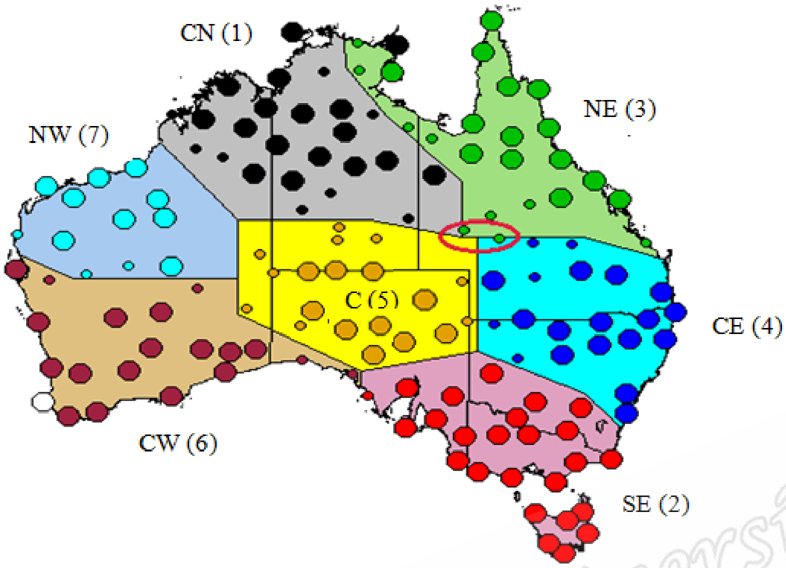

Figure 4.8 Factor analysis divides the stations into 7 geographical groups

The model (4.3) originally gave eight factors but factor 8 only contributed in two stations in NE (Springvale and Westerton represented as the circled stations in Figure 4.8) with smaller loadings than other factors. Therefore, factor 8 is not a dominating factor in any region. There is a station (Warner Glen) in CW (empty circle) with all loadings of factors smaller than 0.333 with uniqueness 0.691, and thus there is no identified dominating factor for this station. In addition, there are small circles indicating mixed (two) identified dominating factors.

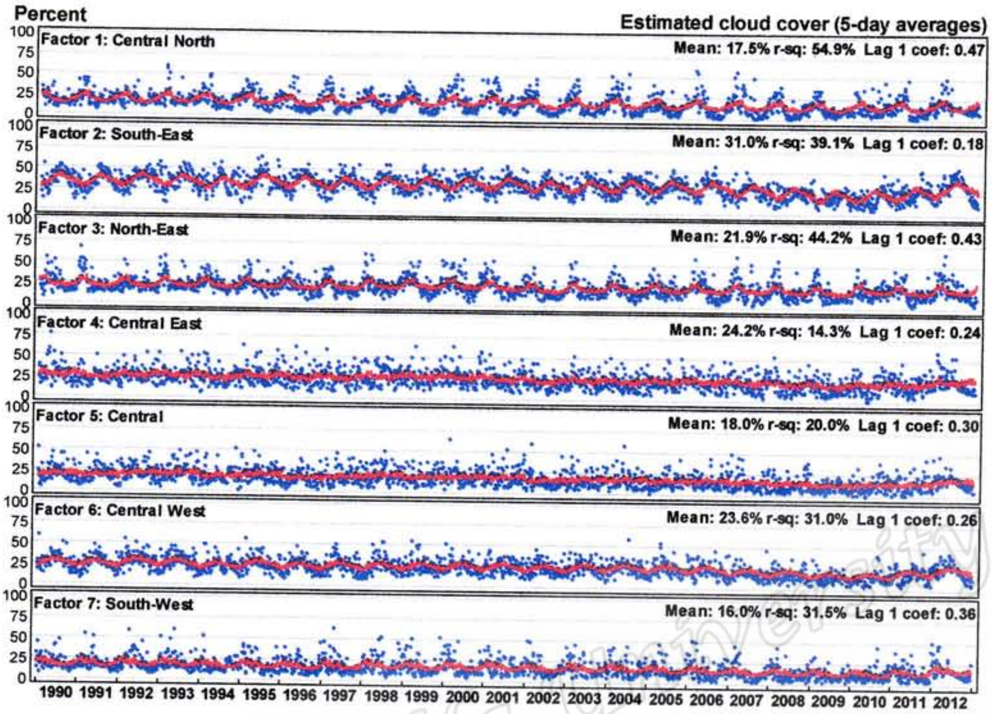

Figure 4.9 The seven dominating factors in the trend estimation based on model (4.3)

Seven patterns of cloud cover from factor analysis model are shown in Figure 4.9. These patterns were classified into three types of estimated cloud cover; factors 1, 3, 7 which represent the tropical regions of CN, NE and NW respectively, have distinct seasonal variation (downward then upward). Factor 2 represents the SE Australian region with the highest latitudes among all the regions (upward then downward). Comparatively, solar absorption in the regions of CE, C and CW with dominating factors 4, 5 and 6 respectively has large temporal fluctuation in the sub-seasonal timescale.

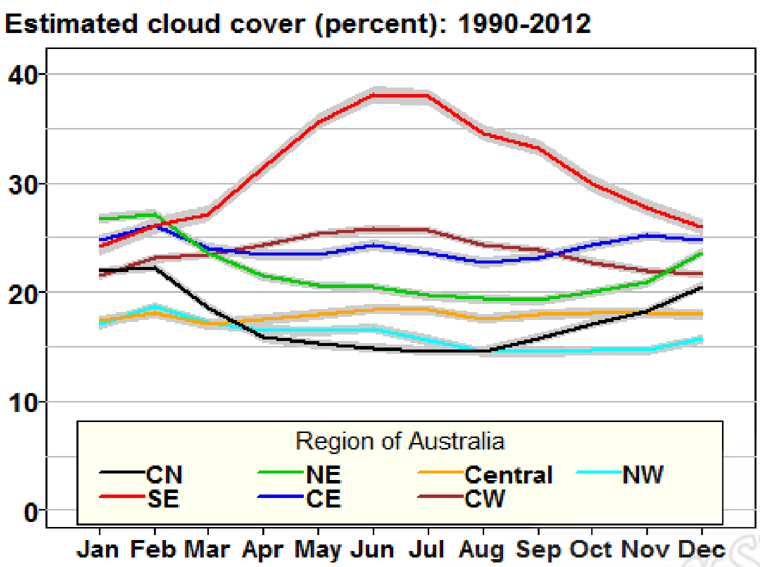

Figure 4.10 Estimated percentage cloud cover (line) with 68% confidence interval (shaded) in each month of a year.

Figure 4.10 shows estimated percentage cloud cover with ± 1standard error (68% confidence interval) in each factor. With the monthly explanatory variable, in the SE region the percentage increases sharply from January to June, then decreases from July to December. The largest percentage in the CN, CE and NW is in February, with a substantial decline to the minimum in August. However, the percentages in the CE and NW regions increase slightly in June. The largest percentage in the NE region is in February, with a sharp decrease to the minimum in September. Overall, the percentage of cloud cover is different in each region: the tropical regions have the highest cloud cover in the Austral summer during which the monsoon is active, while the mid-latitude region of SE has the highest cloud cover in the Austral winter. In contrast, the other central regions do not have clear seasonal cycles.

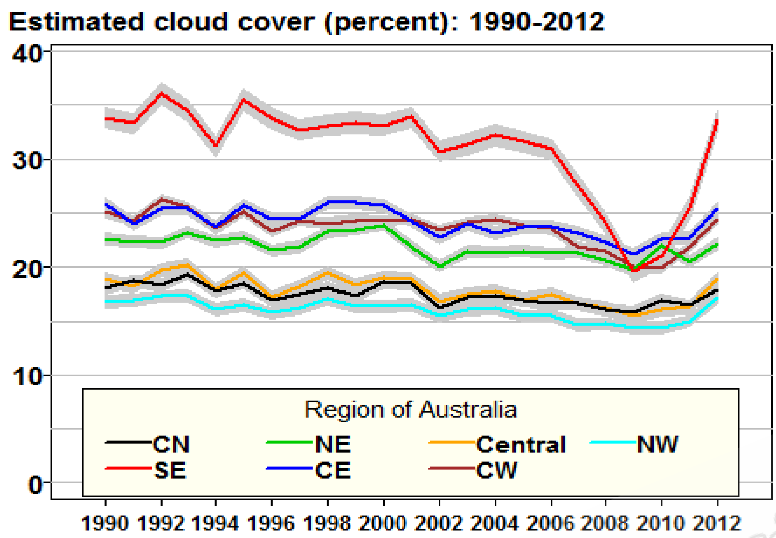

Figure 4.11 Estimated annual average percentage cloud cover (line) with 68% confidence interval (shaded) during 1990-2012

Examination of the seasonal cycle and inter annual variation of solar radiation based on the factor analysis model indicates three major regional groups in term of the estimated annual cloud cover as shown in Figure 4.11. The SE region forms a group with the highest annual cloud cover percentage above 30%, the group consisting of CE, NE and CW has medium percentage of about 25%, and that with C, CN and NW as members has the lowest percentage of less than 20%.

4.3 Conclusions

The daily solar irradiance during 1990-2012 in Australia was investigated. For our study we selected an area-based sample of 144 surface stations across Australia. For each station, a linear regression model was used to analyze the variables of year and pentad factors. Data from the weather stations are also highly correlated spatially. To accommodate these correlations, we use exploratory factor analysis to create larger regions within which data are correlated, but between which correlations are substantially reduced. From factor analysis, it is found that the remaining seven factors have distinct geographic distribution. In each of the identified geographic regions (CN, NE, CE, SE, CW, NW and C) a unique factor is the dominant component.

Our results show that the estimated percentage of solar radiation energy absorbed by clouds and other bodies in the lower atmosphere varies with the month of the year, with patterns for different geographical locations in Australia. In addition, the decreasing trend of solar radiation and the estimated cloud cover in the Australian region starting from the early twenty-first century and reaching a minimum in 2009 indicates that the availability of solar radiation is much influenced by climate variability and change. Under such influences, the major factors affecting surface solar irradiance such as cloud cover, cloud types and water vapour will be modified spatially and temporally. However, while annual trends showed variation, no overall increase or decrease over the 23-year period was evident. The fraction of radiation absorbed by the upper atmosphere (P_0) is not constant value. The rational are able to calculate by observing maximum solar radiation from stations. The maximum is $R_E(1-P_0)$. In this study $P_0 = 0.2$.