CHAPTER 5

TYPICAL WEATHER FOR BUILDING ENERGY ANALYSIS

he selection of climatic time-series collections, to provide meaningful simulation boundary conditions, will require great care. Of fundamental importance is the typicality of the collection in relation to the elements of the design problem in hand." – (Clarke, 1985, pp. 215)

The commonly used forms of weather data, such as outdoor design temperatures and degree days, are not adequate for building energy analysis. A set of detailed weather data is required for modelling of building thermal characteristics over the long term. This chapter depicts the concepts of typical weather and investigates two common methods for selecting typical years: ASHRAE Test Reference Year (TRY) method and Typical Meteorological Year (TMY) method. Typical years for Hong Kong are determined and weather files for building energy simulation are developed. The weather files are then evaluated by studying their simulated performance and the effectiveness of the typical year approach is discussed.

5.1 Basic Concepts

The weather input for building energy simulation is described and the basic properties of typical years are depicted.

5.1.1 Weather data for simulation

Weather data are required for building energy simulation and analysis to serve as input for driving the thermal models within the simulation tool.

The type and format of weather data required depends on the objective of the study and the type of simulation tool.

Using the right data

The purpose and end use of the simulation must be considered when choosing the weather data (ASHRAE, 1991c, pp. 36.5). For an analysis to check the design performance in a specific year (such as an energy audit), the weather data of the year concerned should be used. For comparative studies and long-term energy estimation, a year representative of the average climatic conditions is often used. Keeble (1990) has classified three types of hourly weather data for use in building energy simulation:

- *Multi-year datasets* – They are the fundamentals and include substantial amount of information for a number of years.

- *Typical years* – A typical or reference year is a single year of hourly data selected to represent the range of weather patterns that would typically be found in a multi-year dataset *. Definition of a typical year depends on its satisfying a set of statistical tests relating it to the multi-year parent dataset.

- *Representative days* – They are hourly data for some average days developed to represent the typical climatic conditions. Representative days are economical for small-scale analysis and are often found in simplified simulation and design tools.

For detailed simulation, typical years are most commonly used since analysis using a multi-year dataset is often not feasible and economical for the common problems whereas representative days are too limited and not accurate enough. Typical year has two major advantages:

- Computational efforts in simulation and weather data handling are reduced by using one year instead of multiple years.

* A ‘typical year’ is also known in many other names, such as reference year, test year, test reference year, short reference year, design reference year, typical meteorological year, weather year, typical weather year, example weather year, standard year and standard weather year.

- A consistent form of weather data is ensured so that results from different studies can be compared.

Weather data generation

The simplest way to obtain the weather data for simulation is to take the typical data in the weather library of the simulation program. However, most simulation programs used nowadays only provide weather data for limited number of locations, such as in North America and Europe. There is very little information for developing countries like Hong Kong *. When the weather data files are not available, climatic records may be used to generate them using raw weather data. Three major steps are involved for weather data generation for detailed energy simulation:

- Determine and establish the multi-year dataset (to represent the long term).
- Select typical year(s) from the multi-year dataset.
- Prepare the required weather files (for typical years and multi-years).

When preparing the weather files, the data format and requirements of the simulation program must be considered. The typical year data may have to be adjusted and reconstructed (Skartveit, Lund and Olseth, 1992); missing and unavailable parameters have to be supplied through other means (such as empirical models). Provided that the long-term climatic data exist and are of sufficient quality (see Section 4.2), typical year weather data can be established from the multi-year dataset using suitable method.

5.1.2 Typical year studies

The concept of typical year comes from solar energy simulation (Gansler, Klein and Beckman, 1994; Petrie and McClintock, 1978; Benseman and Cook, 1969). It is applied to building energy simulation since the two applications have similar requirements for weather input. Basically, a typical

---

* The simplified simulation programs Carrier HAP (Carrier Corporation, 1990), TRACE 600 (Trane Company, 1992b) and BUNYIP (Moller and Wooldridge, 1985), have provided weather data files for Hong Kong, but the files are not full 8,760 hours.

year can be considered as a compromise between using the multi-year dataset and portraying accurately the statistics of the long-term weather data, such as using a synthetic year (Knight, Klein and Duffie, 1991). To ensure that the sequence of weather occurrence in nature is preserved, a typical year is usually generated by selecting the entire year from the past weather data or by linking together selected months or seasons. The major research studies on typical years have been examined and a list of them is given in Table 5.1.

The methodology and basis of the typical year studies are different. Some require analysis of full hourly data, such as WYEC (Crow, 1981); some only use daily or monthly averages, such as CIBSE Example Year (Holmes and Hitchin, 1978; Hitchin, *et al.*, 1983) and ASHRAE TRY (NCC, 1976). The period of weather data used usually ranges from 10 to 40 years, depending on the availability of weather data. Although the selection procedure may vary across these studies, the basic concepts of the more well-established early studies are often referred to and adopted. The most common typical year methods for locations with warm climate include:

- ASHRAE Test Reference Year (TRY) (NCC, 1976).
- Typical Meteorological Year (TMY) (NCC, 1981).

The information about typical years in Hong Kong was very limited and the author has tried earlier to investigate and develop some basic data for Hong Kong (Hui and Lam, 1992; Lam, Hui and Yuen, 1992). The approach and data have been expanded in this thesis.

Table 5.1 Major Research Studies on Typical Years

Typical year study (Reference sources)	Period used	Number of years	Select by
USA			
ASHRAE Test Reference Year (TRY) (NCC, 1976)	1948-75	28	Year
Typical Meteorological Year (TMY) by NCC (NCC, 1981)	1952-75	15 - 24 (mostly 23)	Month
New TMY by NREL (Marion, 1994)	1961-90	30	Month
Weather Year for Energy Calculations (WYEC) (Crow, 1981 & 1984)	1941-80	30 to 40	Month
Europe			
Belgium TRY (Lund and Eidroff, 1980)	1958-75	15	Month
CIBSE Example Year (Holmes and Hitchin, 1978)	1956-75	20	Year
Danish TRY (Lund and Eidroff, 1980)	1959-73	15	Month
Finnish Test Year (Gabrielsson and Wiljanen, 1994)	1968-87	20	Month
Italy TRY (Lund and Eidroff, 1980)	1958-70	13	Month
Short Reference Year and EC-TRY (Lund, 1985)	1958-75	13 to 18	Month
Other countries			
Athens' TMY (Greece) (Pissimanis, *et al.*, 1988)	1966-82	17	Month
Canadian TMY (Siurna, D ndrea and Hollands, 1984)	1967-76	10	Month
China Standard Year (Baizhen and Shengyuan, 1989)	1974-83	10	Month
Ibadan Test Reference Year (Nigeria) (Fagbenle, 1995)	1979-88	10	Month
Japan Standard Year (Matsuo, *et al.*, 1972)	1960-69	10	Month
Saudi Arabia Typical Weather Year (Said and Kadry, 1994)	1970-91	22	Month
Standard Solar Year (New Zealand) (Benseman and Cook, 1969)	1954-64	11	Month
Hong Kong			
HK Example Weather Year (Wong and Ngan, 1993)	1967-91	25	Year
HK Test Reference Year (Hui and Lam, 1992)	1948-90	42	Year
HK Typical Weather Year (Lam, Hui and Yuen, 1992)	1980-89	10	Month
Present study in this thesis			
Analysis using ASHRAE TRY method (exclude years 1940-46)	1884-1994	104	Year
Analysis using TMY method	1979-94	16	Month

Characteristics

Generally, the typical year data are valid only for a limited geographical area and should not be considered as a climatological description of a specific town or region (Lund, 1991). The importance of the typical weather is based on the assumption that the weather characteristics over a suitable time period will ecur’ in the future. Therefore, a typical year which can represent the multi-year dataset can be used to assess the future performance of buildings. However, typical years are not necessarily good indicators of weather conditions over the next year, the next five years, or even the next 10 years (Marion, 1994). Rather, they are expected to represent average conditions over a long period of time (the number of years in the multi-year datasets they are drawn from). If the typical year is formulated from weather data of the past 30 years, it should remain useful for the next 30 years. The effectiveness of a typical year relies on how it is generated and what sort of data it is based on. The representativeness of a set of weather data depends on how ypicality’ is defined and measured. A year considered ypical’ for a situation might not be so for another application since the weather variables and their effects on building performance are varying. It is impossible to establish a year which is ypical’ for all situations.

Two approaches to typical years

There are two approaches to define suitable criteria for selecting and analysing typical years:

- *Statistic-based approach* – It is based solely on statistical analysis of the weather attributes.

- *Simulation-based approach* – It is based on comparison of simulation results obtained using the weather data.

The statistic-based approach, which is most common nowadays, aims at finding a set of weather data that is ypical’ for the important weather parameters from statistical points of view. The selection issue is isolated from the building-related factors (such as building systems and plant) and can be assessed independently. On the other hand, the simulation-based approach targets on a set of weather data that has load and energy performance close to the long term. The effects of climate are assessed for the particular circumstance with the present of all other factors. The key criterion is: the simulation output with the typical year as input should deviate as little as possible from the mean output obtained by using the complete multi-year dataset (Skartveit, Lund and Olseth, 1992). Both approaches have some limitations. The statistic-based approach is indirect in nature and may not yield close correlation with actual load and energy performance; the simulation-based approach is specific for the building, system, simulation tool and method used in the selection process. A common way to study typical years is to select a year first by a statistic-based approach and then carry out simulation exercises to see how far the selected data resemble the long term.

5.2 Typical Year Methods

The ASHRAE Test Reference Year method and the Typical Meteorological Year method are investigated. The basic concepts of these two methods are extended to develop typical years for Hong Kong.

5.2.1 ASHRAE Test Reference Year method

In the 1970 ASHRAE has established a simple procedure for selecting a TRY for HVAC applications. The selection procedure was applied to sixties cities in USA to determine TRY for these locations (Stamper, 1977). A standard format, known as the TRY format, was designed for presenting hourly weather data of the weather tapes (NCC, 1976) *.

Selection procedure

The principle for determining TRY is to eliminate the years which contain extremely high or low monthly mean air temperatures until only one year remains. The months are arranged in order of importance for energy comparisons; hot Julys and cold Januarys are considered the most important. All months are ranked by alternating between the warm half (May to October) and the cold half (November to April) of the year, as shown in Table 5.2.

* The TRY and TMY formats should not be mixed up with the selection of 'TRY and TMY'. Data for a selected TRY can be compiled using the TRY format as well as the TMY format provided that there are enough weather data.

Table 5.2 Rankings in the ASHRAE Test Reference Year Procedure

1. Hottest July	9. Hottest May	17. Coldest June
2. Coldest January	10. Coldest November	18. Hottest December
3. Hottest August	11. Hottest October	19. Coldest September
4. Coldest February	12. Coldest April	20. Hottest March
5. Hottest June	13. Coldest July	21. Coldest May
6. Coldest December	14. Hottest January	22. Hottest November
7. Hottest September	15. Coldest August	23. Coldest October
8. Coldest March	16. Hottest February	24. Hottest April

The selection procedure begins by marking the extreme months according to the rankings. If more than one year remain without any marked months, elimination will continue with the next-to-hottest July, the next-to-coldest January, and so on, until one year is left without being marked (i.e. the TRY). To facilitate the analysis process, the author has developed a supporting computer program (in BASIC language), called RYSEL’ (TRY selection), to perform the selection process efficiently using the monthly data as inputs (such as DBT, WBT and GSR). Figure 5.1 shows a brief overview of the RYSEL’ program. Further details about the program can be found in Lam and Hui (1995b).

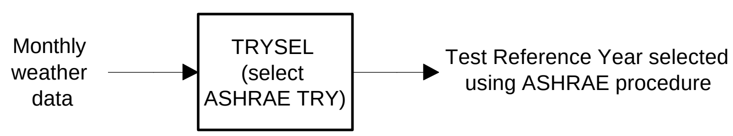

Figure 5.1 *Overview of the RYSEL’ Program*

TRY selection for Hong Kong

The ASHRAE TRY method has been performed by Hui and Lam (1992) to determine TRY for Hong Kong. The original TRY procedure only looks at the monthly mean DBT. The concept has been extended to include four important weather parameters in the analysis: DBT, WBT, DPT and GSR. To study the properties of the selection method, TRY for Hong Kong based on different parameters and periods have been determined using the monthly average weather data in Appendix II. The results are summarised in Table 5.3. Four periods of weather data have been studied: (a) 1979-94 which has full hourly GSR data, (b) 1960-94 which is the length of the hourly weather database developed in this research (see Section 4.3.1), (c) 1947-94 which is the period after the World War II and (d) 1884-1994 which is the maximum period of climatic records in Hong Kong (exclude 1940-46).

Table 5.3 Selection of Test Reference Years for Hong Kong Using ASHRAE Method

		Selected Test Reference Year (TRY) [1]			
Period used [2]	Nos. of year	DBT	WBT	DPT [3]	GSR
1979 - 1994	16	1989	1991	1989	1993
1960 - 1994	35	1989	1961	1961	1976
1947 - 1994	48	1989	1961	1961	–
1884 - 1994	104	1949	1920	1913 / 1903	–
Recommended TRY:					
· If measured hourly GSR is req		1989	1991	1989	1993
· If measured hourly GSR is not req		1989	1961	1961	1976
Suggested applications		General HVAC	Humidity-sensitive		Solar-sensitive

Notes: 1. DBT is dry-bulb temperature, WBT is wet-bulb temperature, DPT is dew-point temperature and GSR is the global solar radiation.

2. The period 1884 - 1994 does not include the seven years 1940 - 1946 during the World War II.

3. The years 1903 and 1913 are selected for DPT based on the period 1884 - 1994 because the two years are marked at the same time at the end of the selection.

It is found that the selected TRY may be different if different parameters or periods of years are used. For DBT, the year 1989 was selected for the periods 1979-1994, 1960-1994 and 1947-1994 while the year 1949 was selected for 1884-1994. The results for WBT, DPT and GSR are more varying and the years 1993, 1991, 1989, 1976, 1961, 1920, 1913 and 1903 have been suggested in different cases. One limitation of the use of TRY is the availability of GSR data. Since hourly GSR in Hong Kong is only measured starting from December 1978 (see Section 4.3.1), there will be a problem composing an hourly weather file if a year before 1979 is selected. The recommended TRY for the four parameters and the suggested applications are given in Table 5.3. For general HVAC applications, the year 1989 is a suitable candidate. Incidentally, a study by Wong and Ngan (1993) using the CIBSE method (Holmes and Hitchin, 1978) and the weather data for the years 1967-1991, also suggested 1989 as an example weather year for Hong Kong. For specific studies where outdoor humidity is important, the year 1989 should be used if DPT data are required and the year 1991 should be used if WBT are required. If both DPT and WBT are required, either 1989 or 1991 may be used. For studies on solar loads, the year 1993 is more appropriate.

Characteristics of TRY

TRY can provide a simple approach for determining weather data for building energy analysis. Reeves, Robart and Stamper (1976) showed that the energy use predicted using TRY is sufficient for comparative energy studies; when testing the representativeness of TRY, Arens, Nall and Carroll (1979) found that the TRY tapes in USA contain cooler temperatures than the long-term average climate record. TRY is often considered useful for comparative analysis but not typical enough for estimating the average long-term energy requirements (NCC, 1976). The general view is that selection based on linked months (such as TMY) has a better chance of forming a year close to the long term. However, there is no definite and quantitative answer to how typical a TRY is or is not. Generally, the consequence of picking a less typical set of weather data will not affect the overall interpretations and analysis of most simulation studies which are used for comparative evaluations.

5.2.2 Typical Meteorological Year method

The TMY method, developed by Sandia National Laboratories in USA (NCC, 1981; Hall, *et al.*, 1978), is the most widely accepted method for determining typical years and its concept has been used in many places including Canada (Siurna, D ndrea and Hollands, 1984), Greece (Pissimanis, *et al.*, 1988), Nigeria (Fagbenle, 1995), Saudi Arabia (Said and Kadry, 1994) and USA (Hall, *et al.*, 1978; Marion, 1994); TMY weather tapes for locations in USA have been created using data supplied from the National Climatic Data Center (NCDC) *. Descriptions of the TMY method and the TMY format for presenting hourly weather data can be found in the *TMY User Manual* (NCC, 1981). When adopted in countries outside USA, the original TMY method has been modified to suit particular needs and emphases of that location; different weather parameters and selection criteria may have been used. An overview of the TMY method is provided below to help understand the basic concept.

Basic approach

A TMY consists of twelve ypical Meteorological Months' (TMM) selected from the calendar months in a multi-year weather database. For example, the January of 1980 may be selected as the first TMM, the February of 1989 as the second TMM, and so on. All the twelve selected months will then combine to form the TMY. Smoothing of data for discontinuities may be needed to avoid abrupt changes at the boundary between two adjacent months coming from different years †. Selection of a TMM is based on statistical analysis and evaluation of four weather parameters: GSR, DBT, DPT and WSP. A set of thirteen indices were initially considered, including the daily total GSR and the daily maxima, minima, means and ranges of DBT, DPT and WSP. But the three daily ranges and the daily minimum WSP were assigned zero weightings, thus leaving only nine indices in the actual

---

* The National Climatic Data Center (NCDC) is previously known as the National Climatic Center (NCC).

† Smoothing for pressure, temperatures and wind speed data is recommended in the *TMY User's Manual* (NCC, 1981).

selection. The nine daily indices and their respective weightings are shown in Table 5.4. Factors for the original TMY method (NCC, 1976) and two modified schemes by WSL (1992) and NREL (Marion, 1994) are compared in the Table *. Principally, TMY was designed for solar system simulation; the weather parameters and indices influential to solar system performance are given high priority (40% to 50% as shown in Table 5.4).

Table 5.4 Weather Indices and Weighting Factors for TMY method

Weather parameter	Daily index	Weighting factors		
		Old TMY	New TMY	CWEC
Dry-bulb temperature (DBT)	Daily maximum	1/24 (4.2%)	1/20 (5%)	5%
	Daily minimum	1/24 (4.2%)	1/20 (5%)	5%
	Daily mean	2/24 (8.3%)	2/20 (10%)	30%
Dew-point temperature (DPT)	Daily maximum	1/24 (4.2%)	1/20 (5%)	2.5%
	Daily minimum	1/24 (4.2%)	1/20 (5%)	2.5%
	Daily mean	2/24 (8.3%)	2/20 (10%)	5%
Wind speed (WSP)	Daily maximum	2/24 (8.3%)	1/20 (5%)	5%
	Daily mean	2/24 (8.3%)	1/20 (5%)	5%
Global solar radiation (GSR)	Daily total	12/24 (50%)	5/20 (25%)	40%
Direct normal solar radiation	Daily total	Not used	5/20 (25%)	Not used

Notes: 1. Old TMY refers to the original TMY method from the National Climatic Center (NCC) (NCC, 1981); new TMY refers to the method proposed by the National Renewable Energy Laboratory (NREL) (Marion, 1994); CWEC is the Canadian Weather for Energy Calculations (WSL, 1992).

2. The percentage of each index is given in parenthesis next to the weighting factor.

Selection criteria

TMY selection involves minimising the difference from long-term distributions, means and daily persistence of the weather indices. To determine a TMM, three basic properties are considered (NCC, 1981):

* The Watsun Simulation Laboratory (WSL) in Canada has adopted the TMY method, with some changes in the weighting factors, for creating the Canadian Weather for Energy Calculations (CWEC) (WSL, 1992). In a recent research study carried out by the National Renewable Energy Laboratory (NREL), an index for the direct normal solar radiation was

- *Frequency distributions* – Climatic elements should have frequency distributions close to the long term.

- *Sequences* – The sequences of daily measurements should be ike’ the sequence often registered at that location.

- *Correlations* – The relationships among different climatic elements should be ike’ the relationships observed in nature.

Hall, *et al.* (1978) and Lund and Eidorff (1980) believed that by using real weather data, the requirements on sequences and correlations could be satisfied. It was hoped that a TMY meeting all these criteria would have weather patterns and system performance close to the long term. However, there is no clear indications on how these criteria are achieved.

Screening process

For each TMM, a screening process is first performed to select five candidate years. A nonparametric method, known as Finkelstein-Schafer (FS) statistic (Finkelstein and Schafer, 1971) is used to determine the candidates by comparing the yearly cumulative distribution function (CDF) with the long-term CDF in the month concerned (the long-term CDF is the aggregate sum of all the yearly CDF) *. An empirical CDF, which is a monotonic increasing function, is defined using the order statistic (Conover, 1980), like this:

$$S_n(x) = \begin{cases} 0 & \textit{for } x < x_{(1)} \\ (k - 0.5)\ /\ n & \textit{for } x_{(k)} \le x \le x_{(k+1)} \\ 1 & \textit{for } x \ge x_{(n)} \end{cases} \quad \text{where } k = 1, \ldots, n - 1 \qquad (5.1)$$

where $Sn\,(x)$ = value of the cumulative distribution function at x

n = total number of elements

k = rank order number

(the subscript number in parenthesis is the anked order’)

---

added to be one of the criteria for TMY selection. The weighting factors for wind speed were also reduced.

Values of the FS statistic are calculated for each of the daily indices using the following equation (NCC, 1981; Hall, *et al.*, 1978):

$$FS = \frac{1}{n}\sum_{i=1}^{n} d_i \tag{5.2}$$

where FS = value of FS test statistic

d_i = absolute difference between the long-term CDF and the yearly CDF at $x_{(i)}$ value

n = the number of daily readings for that month (such as for January, n = 31, for April, n = 30, and so on)

A weighted-sum average (WS) or composite index is then computed for each year and the five years with the smallest WS values are selected as candidates to be taken to the final selection. The WS values is calculated by:

$$WS = \sum_{i=1}^{9} WF_i \times FS_i \tag{5.3}$$

where WS = weighted sum average

WF_i = weighting factor for the i^{th} parameter

FS_i = FS test statistic calculated for the i^{th} parameter

Final selection

The final selection encompasses a two-step process (Hall, *et al.*, 1978). The first step checks on the statistics associated with the mean daily DBT and daily total GSR, including the FS statistic and the deviations of the monthly mean and median from the long-term mean and median. The second step looks at the persistence in the mean daily DBT and daily total GSR by examining their run structure. Persistence is considered important for solar systems since in some cases the distribution of a given year can be quite close to that of the long term, yet there still can be atypically long runs of cloudy or warm or cool days. Unfortunately, the final selection procedure has not been

---

* The names of the authors of the statistical test, 'Finkelstein and Schafer', was taken as the

clearly defined in the *TMY User Manual* (NCC, 1981). It was only mentioned that an attempt should be made to select years with small FS values, small deviations, and typical run structures. But the exact procedure and criteria have not been provided. Various methods for the final selection have been proposed and used by different researchers following the TMY concept. Some of them look at the root mean square difference of GSR (Pissimanis, *et al.*, 1988); some assess the persistence and the monthly mean and median (Marion, 1994); some take the year with the lowest WS value as the TMM (Said and Kadry, 1994). It is difficult to compare the results of different TMY studies.

Shortcomings of current TMY method

Although TMY method is commonly used nowadays, it suffers from a number of drawbacks:

- *Only daily indices are considered* – There is no reason why hourly data cannot be used. By using hourly data, the daily variations can be preserved and the power of the statistical tests can be increased.

- *Selection criteria are vague* – The basic criteria (frequency distributions, sequences and correlations) are only qualitatively described.

- *Selection procedure has not been clearly defined* – The lack of a clear, standard procedure for final selection is a serious problem of TMY.

- *Statistical basis is ambiguous* – The statistical theory and inference used for CDF comparison are ambiguous.

- *Formulation of the test statistic is not clear* – The formulation of the FS statistic and the procedure for statistical calculation are not clear.

Performing TMY selection in practice is problematic because of the ambiguity and vagueness of the definition and procedure. It is believed that an unambiguous procedure is needed to define the steps and properties of the

---

name of the test.

selection. The TMY concept has been clarified and expanded in this research to determine a set of TMY for Hong Kong.

5.2.3 TMY selection for Hong Kong

The basis of the onparametric statistic' used in the TMY method has been studied carefully *. Efforts have been made to clarify the principle of statistical inference and the formulation of test statistic in TMY so that an unambiguous method can be established for TMY selection. The statistical basis and the calculation procedure are explained in Appendix III. To make the best use of the information in the weather data, the statistical calculations are performed not only on daily indices but also on hourly data. Different sets of TMY using daily and hourly indices have been determined. Apart from the FS statistic, another nonparametric test statistic, known as the Kolmogrov-Smirnov (KS) two-sample statistic, is proposed for the analysis (Gibbons and Chakraborti, 1992, Chp. 15). While the FS statistic is based on the magnitude of the CDF difference, the KS statistic is based on the maximum deviation like this:

$$KS = \max \left| d_i \right| \qquad \text{for } i = 1, 2, \ldots, p \tag{5.4}$$

The KS statistic is chosen because it is more well-known (Festa and Ratto, 1993; Lohrding, 1973) and its critical values can be determined from relevant statistical references. In order to select TMY for Hong Kong, the basic process on frequency distributions is performed to determine the values of the nonparametric statistic. However, the assessment on persistence and run structure as suggested in the *TMY User Manual* is not followed since there is no evidence that it can provide useful information for relating building energy performance. To avoid ambiguity and to simplify the process in the final selection, the year with the lowest weighted-sum average of the test statistic (FS and KS statistics) is selected as the TMM (this approach has also

---

* 'Nonparametric' or 'distribution-free' methods are statistical methods with the property that no assumption is being made concerning the specific distribution from which the sample is drawn (Gibbons and Chakraborti, 1992; Neave and Worthington, 1988). The studies and mathematical theory of nonparametric inference are still embryonic (Gibbons and Chakraborti, 1992) and understandings of its properties are limited nowadays.

been used by Said and Kadry (1994)). To facilitate comparison with other typical year methods (such as TRY), the TMY concept has been extended to consider selection of the entire year as well. The actual selection process begins with the determination of weather parameters.

Weather parameters

Since cooling requirements are the most important energy use in commercial buildings in Hong Kong, the weather parameters essential for building cooling loads are considered. DBT, DPT, WBT, GSR and WSP are the five common climatic variables taken for the analysis of summer air conditioning (Cuncliffe, 1965). Both daily and hourly indices are used respectively in this study to determine different sets of TMY. Table 5.5 shows the indices and their relative weightings adopted for the TMY selection in the present study. As hourly GSR data of Hong Kong are not available before December 1978, to ensure a complete hourly dataset, the years 1979-1994 (16 years) is employed for the TMY analysis. For daily indices, the weightings of the original TMY method (NCC, 1981) is followed. For hourly indices, a set of weighting factors is proposed for the five parameters (DBT, WBT, DPT, GSR and WSP): WSP is considered less important to building energy analysis whereas WBT and DPT are grouped together to represent humidity.

Table 5.5 Proposed Weighting Factors and Indices for TMY Selection

Weather parameter	Index	Weighting factors	
		For daily indices	For hourly indices
Dry-bulb temperature (DBT)	Daily maximum	1/24 (4.2%)	–
	Daily minimum	1/24 (4.2%)	–
	Daily mean	2/24 (8.3%)	–
	Hourly mean	–	2/10 (20%)
Dew-point temperature (DPT)	Daily maximum	1/24 (4.2%)	–
	Daily minimum	1/24 (4.2%)	–
	Daily mean	2/24 (8.3%)	–
	Hourly mean	–	1/10 (10%)
Wind speed (WSP)	Daily maximum	2/24 (8.3%)	–
	Daily mean	2/24 (8.3%)	–
	Hourly mean	–	1/10 (10%)
Global solar radiation (GSR)	Daily total	12/24 (50%)	–
	Hourly mean	–	5/10 (50%)
Wet-bulb temperature (WBT)	Hourly mean	–	1/10 (10%)

Notes: 1. The percentage of each index is given in parenthesis next to the weighting factor.

The relative importance of the weather parameters is difficult to determine since different building designs and applications may result in some weather parameters becoming more influential (for example, extensive use of windows and skylights may make solar radiation a dominating factor for building energy consumption). The individual values of the test statistic calculated for each index can be used to work out a different set of weighted-sum values for typical year selection if a particular application requires. For example, a ypical wind year' (NCC, 1981), a ypical solar year' and a ypical humidity year' may be constructed by putting more emphasis on the wind, solar and humidity data, respectively.

Supporting program for nonparametric calculations

To facilitate the calculations of nonparametric statistics for the TMY analysis, the author has developed a supporting computer program, called PARM' (nonparametric statistics), to perform the statistical calculations on the weather data. This program performs nonparametric calculations for hourly or daily data, including the FS and KS statistics. Further details of the PARM' program can be found in Lam and Hui (1995b). Figure 5.2 gives a brief overview of the PARM' program. The procedure described in Appendix III forms the basis of the computer algorithm for PARM'.

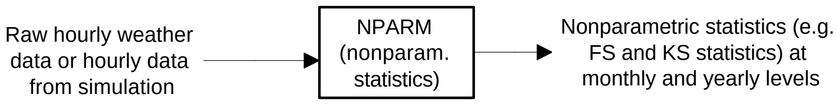

Figure 5.2 *Overview of the PARM' Program*

TMY selection using daily and hourly indices

Table 5.6 shows the results of the TMY selection using daily indices (five candidate years are shown to study the selection patterns). It can be seen that TMY selections using the FS and KS statistics often give similar results. Eight out of the twelve TMM are the same and the candidate years suggested from the two methods are similar. The year 1989 stands out to be quite distinctive since it was suggested by both FS and KS statistics as the ypical year'. It is interesting since 1989 was also selected as the TRY in Section 5.2.1. When the FS and KS statistics for each individual index are examined, it is found that the behaviours often differ from one index to another, and from month to month. Years considered ypical' for a certain index may not necessarily be good at the others. The temperatures (i.e. DBT, WBT and DPT) tend to agree more often with each others whereas GSR and WSP exhibit very different patterns.

Table 5.6 TMY Selection for Hong Kong Using Daily Indices

Ranks	Year selected (such as 80 = 1980)												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
	Five candidate years selected from FS statistic using daily indices												
1st	80	90	93	80	89=	86	86	86=	82	84	89	93	89=
2nd	94	88	84	93	93=	87	90	93=	93	86	84	85	92=
3rd	85	82	80	79	86	90	92	89	89	93	91	89	90
4th	91	86	86	86	82	91	80	82	90	89	86	82	93
5th	82	81	92	92	84	92	83	85	87	91	85	86	81/85
	Five candidate years selected from KS statistic using daily indices												
1st	85	90	93	93	93	86	89	86	82	84	89	93	89
2nd	80	88	92	79	89	87	86	93	93	86	86	89	93
3rd	94	82	86	80	84	91=	92	89	87=	93	84	85	92
4th	82	86	84	87	86	92=	80	82	90=	89	91	82	90
5th	93	80	80	89/92	87	90	83	91	89	81	85	90	85/88

Note: 1. If two figures are given or the figure is followed by '=', there are equal scores (ties) at three decimal places during the selection.

Table 5.7 TMY Selection for Hong Kong Using Hourly Indices

Ranks	Year selected (such as 80 = 1980)												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
	Five candidate years selected from FS statistic using hourly indices												
1st	80	88	93	80	82	90	86=	86=	82	84	89	93=	89
2nd	94	90	84	93	86=	86	90=	93=	87=	86	84	85=	92
3rd	89	82	82=	87=	93=	87=	89	89	93=	89	91	86	90
4th	85	91	92=	79=	89	89=	80	82=	90=	80	86	80	81=
5th	82	92	80	92	87	91	91	91=	91=	91	90	82	93=
	Five candidate years selected from KS statistic using hourly indices												
1st	80	88	93	87=	86	90	89	86	87	84	89	85	89
2nd	89	82	86	93=	82	86	86	93	82	80	84	86	93
3rd	82	90	92	80	89	87	90	91	93	86	90	93	86
4th	94	89	82	79	87	92	91	89	90	89	86	89	92
5th	87	91	90	92	93	91	80	82	91	90	83	79	90

Note: 1. If two figures are given or the figure is followed by '=', there are equal scores (ties) at three decimal places during the selection.

By extending the original TMY concept to hourly data, TMY selection for Hong Kong has been performed based on full hourly indices as shown in Table 5.7. Many of the years selected from hourly indices are the same as those from daily indices (eight TMM in the FS statistic and five TMM in the KS statistic are the same). When comparing the results from the FS and KS statistics for hourly indices, seven TMM selected are the same. One thing worth mentioning is that the year 1989 is again chosen to be the ypical year' in all these cases.

The patterns found in Table 5.6 and 5.7 indicate that the CDF difference as detected by the two nonparametric methods have similar properties. If the two distributions have no significant difference statistically at hourly level, they should also be close at the daily level. Since the hourly data contain more information, TMY selection based on hourly indices is a more reasonable choice for the analysis if hourly data are required for the end use. It should be noted that comparison of CDF and the results of the nonparametric statistic depends only on the ank order' of the variables, not on their values. The relative positions, not the magnitude, are important for determining the closeness of the distributions. This is an essential properties of nonparametric methods. In general, the larger the sample size for the test statistic, the higher the probability that any real differences will be detected. In statistical terms, the critical region within which the result of the test statistic is considered significant will be larger if more data are analysed. If only a limited amount of data is used for testing the oodness-of-fit', then it is likely that the test will give non-significant results. The uncertainty band within which the true population lies will be wider in the case of daily indices (fewer data) than hourly ones (24 times more data).

Critical values of the KS statistic

Estimation of the critical values for a test statistic can provide a conditional measure of uncertainty. However, the critical values for many nonparametric methods are difficult to compute even for a small sample since recursive formulae and complicated probability derivations are involved

(Gibbons and Chakraborti, 1992). The KS two-sample statistic is the one that has its critical values defined in statistical references, such as Gibbons and Chakraborti (1992), Neave and Worthington (1988) and Conover (1980). The two-sided critical region of the KS statistic at a significant level ♋ is:

$$KS_{m,n} \geq c_a \tag{5.5}$$

where $KS_{m,n}$ = KS test statistic calculated for the two samples of sizes m and n

$c_{♋}$ = critical value of KS statistic at significant level ♋

If the calculated statistic is larger than the critical value, then the null hypothesis will be rejected (i.e. the two CDF are different). For large sample sizes, the two-sided critical value of the KS statistic can be approximated by an asymptotic distribution of the form (Gibbons and Chakraborti, 1992; Neave and Worthington, 1988), like this:

$$c_a = K\sqrt{\frac{m+n}{m \times n}} \tag{5.6}$$

where K = coefficient relating to the significant level (see Table 5.8)

In the TMY selection, the KS statistic is used for comparing the long-term and yearly CDF . The size of the long-term CDF will be equal to the number of years multiplied by the total number of data in each yearly CDF. Let N_y be the number of years and n be the size of each yearly CDF, then the critical value of the KS statistic can be written as:

$$c_a = K\sqrt{\frac{N_y \cdot n + n}{N_y \cdot n \times n}} = K\sqrt{\frac{N_y + 1}{N_y \cdot n}} \tag{5.7}$$

Equation (5.7) has been used to calculate the critical values of the KS statistic for significant levels from 1% to 20% for 16 years of weather data, as shown in Table 5.8.

Table 5.8 Critical Values of KS Statistic for 16 Years of Weather Data

Signif. level ♋	’ coeff.	KS critical value for daily indices				KS critical value for hourly indices			
		28-day (28)	30-day (30)	31-day (31)	Year (365)	28-day (672)	30-day (720)	31-day (744)	Year (8760)
0.01	1.63	0.318	0.307	0.302	0.088	0.065	0.063	0.062	0.018
0.02	1.52	0.296	0.286	0.281	0.082	0.060	0.058	0.057	0.017
0.025	1.50	0.292	0.282	0.278	0.081	0.060	0.058	0.057	0.017
0.05	1.36	0.265	0.256	0.252	0.073	0.054	0.052	0.051	0.015
0.10	1.22	0.238	0.230	0.226	0.066	0.049	0.047	0.046	0.013
0.20	1.07	0.208	0.201	0.198	0.058	0.043	0.041	0.040	0.012

Notes: 1. The size of each yearly sample used for calculating the critical value is given in parenthesis: 28-day for February, 30-day for April, June, September and November, 31-day for January, March, May, July, August, October and December.

2. The above critical values are calculated based on 16 years (see Equation (5.7) if different number is used.

Implications

To check the selected TMY of Hong Kong, the critical values for the KS statistic have been compared with the values of the KS statistic calculated in the TMY selection. A comparison has been made with a critical value of 2.5% (a significant level of 0.025 from Table 5.8) and the results are given in Table 5.9. The 2.5% level is chosen since the same significant level has been used in Section 4.3 for the outdoor design temperatures and it is a common significant level for general HVAC applications. It can be seen that all the weighted-sum averages of the KS statistics calculated for the TMM using daily indices are smaller than the 2.5% critical values. This implies that there is not enough evidence that the TMM are ifferent’ from the long term under this test. But, for the hourly indices, five TMM and the whole-year selection have the weighted-sum averages above the 2.5% critical values. This means that although they are selected, these candidates are ifferent’ from the long term. Users of the typical weather data should be aware that the selected data are uncertain from statistical point of view and the simulated system performance determined using the hourly data may not be close to the long term under the 2.5% criteria.

Table 5.9 Comparison of KS Statistic and Their 2.5% Critical Values

	Year selected (such as 80 = 1980)												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
	TMM selected from KS statistic using daily indices												
TMM	85	90	93	93	93	86	89	86	82	84	89	93	89
WS	0.193	0.165	0.156	0.116	0.146	0.136	0.144	0.119	0.125	0.116	0.115	0.132	0.040
c_{α} (2.5%)	0.278	0.292	0.278	0.282	0.278	0.282	0.278	0.278	0.282	0.278	0.282	0.278	0.081
Reject? [2]	No	No	No	No	No	No	No	No	No	No	No	No	No
	TMM selected from KS statistic using hourly indices												
TMM	80	88	93	87 / 93	86	90	89	86	87	84	89	85	89
WS	0.060	0.065	0.064	0.048	0.065	0.058	0.048	0.040	0.058	0.036	0.050	0.068	0.019
c_{α} (2.5%)	0.057	0.060	0.057	0.058	0.057	0.058	0.057	0.057	0.058	0.057	0.058	0.057	0.017
Reject? [2]	Yes	Yes	Yes	No	Yes	No	No	No	No	No	No	Yes	Yes

Note:
1. TMM is the typical meteorological month, WS is the weighted-sum average of the KS statistic and c_{α} (2.5%) is the critical value at 0.025 significant level.
2. No = null hypothesis cannot be rejected; Yes = null hypothesis is rejected (i.e. not identical distributions)

The above example illustrates that the selected typical years / months have high degree of uncertainty. Even the best candidate is chosen from the dataset, it may differ from the long term under the usual significant criteria. The critical values proposed here can be used to quantify and assess the reliability of TMY selection in a way similar to how the outdoor design temperatures are defined using the significant frequency level (see Chapter 4). As weather files for the typical years will be used for building energy simulation, the degree of uncertainty of the weather input in building energy analysis can be better understood through a rational approach to weather data.

5.3 Weather Files for Building Energy Simulation

The process of weather file generation is described and the essential weather files for building energy simulation in Hong Kong are developed. The solar conditions and local holidays are considered when generating the weather files.

5.3.1 Standard formats

The weather data collected from the weather station cannot be used directly for building energy simulation. They must be converted into a suitable format acceptable by the simulation tool (Degelman, 1991). The author has developed three types of weather files for Hong Kong:

- *Raw weather data files* – They contain one parameter in each file and are prepared from the hourly data obtained from the weather station (ROHK).
- *ASCII weather text files* – They are in a standard format with each line containing the weather data for one hour in fixed column positions [*].
- *Binary weather files* – They are the weather files used directly by the simulation program (such as DOE-2 and BLAST).

To prepare and analyse the weather data of Hong Kong, the basic hourly weather data collected from ROHK have been checked thoroughly for missing data and compiled into the raw data files suitable for processing by the supporting analysis programs (such as STAT', REQ' and PARM'). Figure 5.3 shows how the weather files are generated. The process for weather file generation takes two steps: raw weather data are converted to ASCII file and then the ASCII file is converted to binary file by the weather processing program.

---

[*] 'ASCII' is the American Standard Codes for Information Interchange.

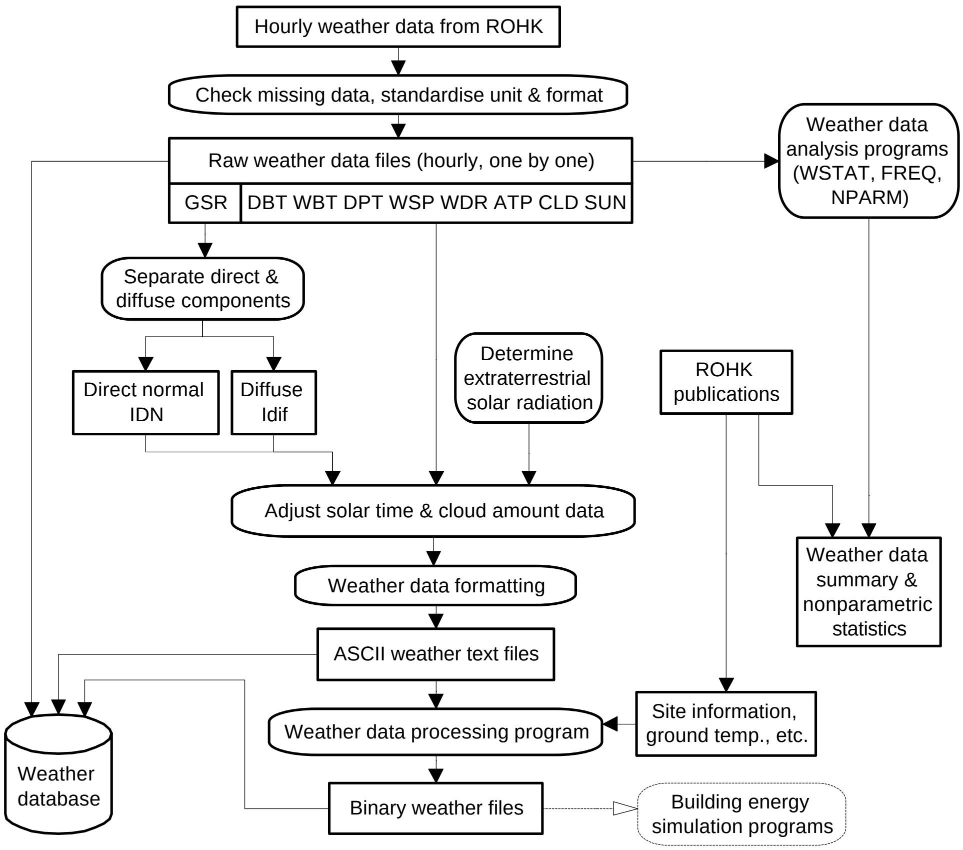

Figure 5.3 *Generation of Weather Files for Building Energy Simulation*

ASCII weather text file

Common formats for presentation of hourly weather data are the TRY and TMY formats mentioned earlier. Other formats include:

- Australian format developed by the Commonwealth Science and Industry Research Organisation (CSIRO) (NSW Public Works, 1993, pp. 309).
- Formats for CIBSE Example Weather Year and the SERC meteorological database (Keeble, 1990).
- European Community Test Reference Year (EC-TRY) and Short Reference Year (SRY) format (Lund, 1985, pp. 18).
- Weather Year for Energy Calculations (WYEC) format developed by ASHRAE (Crow, 1981 & 1984).

- WYEC2 and Canadian Weather for Energy Calculations (CWEC) formats (WSL, 1992).

The differences between these different formats are the tabular layout, the types of climatic elements and the units used for the weather parameters. To facilitate generation of the ASCII files, the author has developed some supporting computer programs to carry out the conversion job. Currently, there are three programs, called MYF', RYF' and USF', to generate weather text files for the TMY, TRY and Australia formats, respectively. Figure 5.4 gives a general overview of the three supporting programs. Further details of the programs can be found in Lam and Hui (1995b).

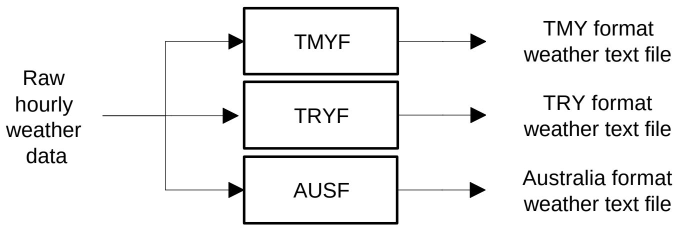

Figure 5.4 *Overview of the MYF', RYF' and USF' Programs*

Binary weather file

The next step is to produce the binary weather file from the ASCII file. The simulation package usually contains a weather processing program to convert ASCII files into binary weather files which are used by the main simulation module. For example, OE2WTH' in DOE-2 (LBL, 1981) and IFE' in BLAST are the weather processing programs (BLAST, 1991). Preparation of weather files for the current simulation programs is a complex task since the amount of data is large and little information is provided about this by the program developers (the program developers usually expect the users to get the weather files from the weather library or send raw weather data to them for preparing specific weather files). If the program users are to generate the weather files by themselves, they must understand clearly the specific operation and requirements of the program.

5.3.2 Generate weather files for Hong Kong

Essential weather files for Hong Kong have been generated in this research to provide a weather database for simulation and climatic analysis. If the data required for the weather files are not present (such as components of solar radiation data), they are determined by other means and supplied to the weather file. The climatic records from ROHK do not contain all the weather parameters which are necessary for compiling weather files for building energy simulation. Values of those unavailable parameters are estimated from relevant models and empirical equations. Three parameters are calculated when generating the weather text files for Hong Kong:

- Extraterrestrial solar radiation.
- True solar time.
- Direct and diffuse components of solar radiation.

The methods for estimating the above parameters are described in Appendix IV. Generation of weather files should consider the final application of building simulation programs so that the correct amount and format of data are taken.

Weather files for DOE-2 and BLAST

The DOE-2 and BLAST simulation programs have been selected in this research as the simulation tools (see Section 6.2). Weather files for Hong Kong have been prepared for these two programs for the years 1979 to 1994 (this period has complete hourly GSR data). Generally, two types of binary weather files are used by DOE-2 (LBL, 1982, pp. III.17):

- *Solar files* – They contain solar radiation data in the form of GSR and direct normal solar radiation.

- *Normal files* – The solar radiation data needed are estimated in the load or weather sub-program using cloud cover and cloud type data.

DOE-2 has two options for generating solar radiation data for normal files: (a) *Boeing cloud cover modifier model* (the default option) and (b) *Kimura-Stephenson cloud cover modifier model* (by specifying the RYKS’ option in OE2WTH’, a solar packed weather file will be generated). By processing the ASCII files created previously (i.e. TRY and TMY formats), three types of binary weather files have been prepared for DOE-2:

- *TMY format weather files* – GSR and direct normal solar radiation data are contained in the file.

- *TRY format weather files* – The solar radiation data will be estimated by the load sub-program of DOE-2 using the Boeing cloud cover model.

- *TRYKS format weather files (or simply TKS files)* – The solar radiation data will be estimated by OE2WTH’ using the Kimura-Stephenspn cloud cover model, and then supplied to the binary weather file.

In the DOE-2 terminology, TMY and TRYKS format files are solar files whereas TRY format files are normal files. Although the input raw weather data are identical, simulation using weather files of different formats will produce different results, since the values of solar radiation data are not necessarily the same. The case for BLAST is simpler since there is only one model used for the TRY format data. Two types of binary weather files have been prepared for BLAST (TMY and TRY format files). Weather files for both the typical years (ASHRAE TRY and TMY) and the multi-year dataset (16 years from 1979 to 1994) have been prepared for Hong Kong.

Holidays adjustments

One important thing often overlooked by users of simulation tools is the holidays assumed by the simulation program. The holiday list has influences on the simulation results since the internal loads (occupancy, lighting and equipment) and system operation are affected by the changes in day schedule. The same set of holidays should be employed for comparative studies and the variations of holidays should be checked when comparing the simulation results of different calendar years and at different locations. By defaults, the U.S. holidays are often used in the U.S. simulation programs, such as DOE-2 and BLAST. When these programs are applied in other countries, the default holiday list becomes inappropriate.

To ensure that the local conditions are considered when doing building energy simulations, the schedule of public holidays in Hong Kong (Hong Kong Government, 1983; *ong Kong $ Directory*", 1979-94) has been studied and the holiday lists in DOE-2 and BLAST has been adjusted. The holiday list of DOE-2 is changed by writing macro functions in the building input (Hui, 1995) and that of BLAST is changed by using the holiday command in the weather processing program (BLAST, 1991). It is found by Hui (1995) that the difference in annual building energy consumption may vary from 0.8% to 2.6% when the default U.S. holidays are replaced by the Hong Kong local holidays. If a typical year comprises of calendar months from different years, the assignment of holiday list for the weather file may present a problem since the day schedules will be affected by the holidays in each month. To avoid this, the holiday schedule of the typical year weather files in this research have assumed the respective holiday schedule of the year they come from.

5.4 Evaluation of Weather Files

The simulation results of the weather files have been analysed at yearly and monthly levels. Energy performance are compared against the climatic variables. The effectiveness of the typical year approach is discussed.

5.4.1 Compare different programs and formats

The weather files generated for Hong Kong have been evaluated using the simulation tools and the base case model developed in this research (see Chapter 6 for detail of the simulation tools and model). Simulated performance of the multi-year dataset and the typical year data (TRY and TMY) have been studied and analysed at yearly and monthly levels [*].

Yearly variations

Figure 5.5 gives a comparison of the annual building electricity consumption in Megawatt-hour (MWh) calculated by DOE-2 and BLAST. It can be seen that the three types of weather files of DOE-2 produce similar patterns along the years, with TRY files being the highest and TMY files being the lowest. Simulation results of BLAST are quite close to each other and they are below the DOE-2 predictions in most cases. Table 5.10 gives the deviations of annual MWh for the different weather files (the standard, maximum and minimum deviations and the deviations of the typical year (1989) are shown). The standard deviations range from 1.2% (BLAST TMY and TRY files) to 1.9% (DOE-2 TMY) while the maximum deviations are from 1.9% (BLAST TMY files) to 3.7% (DOE-2 TMY files). The typical year (1989) is seldom the one with minimum deviation. The results show that annual MWh is not sensitive to the change of weather year and the year closest to the long-term mean may vary in different weather files and programs.

---

[*] The TRY is an entire year in the multi-year dataset (year 1989 as determined in Section 5.2.1) and the TMY consists of TMM's selected using hourly KS statistics (see Table 5.7 for detail). It is found that the effect of smoothing of weather parameters between adjacent months coming from different years is small. Therefore, the energy performance of the TMY is studied by picking the simulation results of the multi-year data month by month.

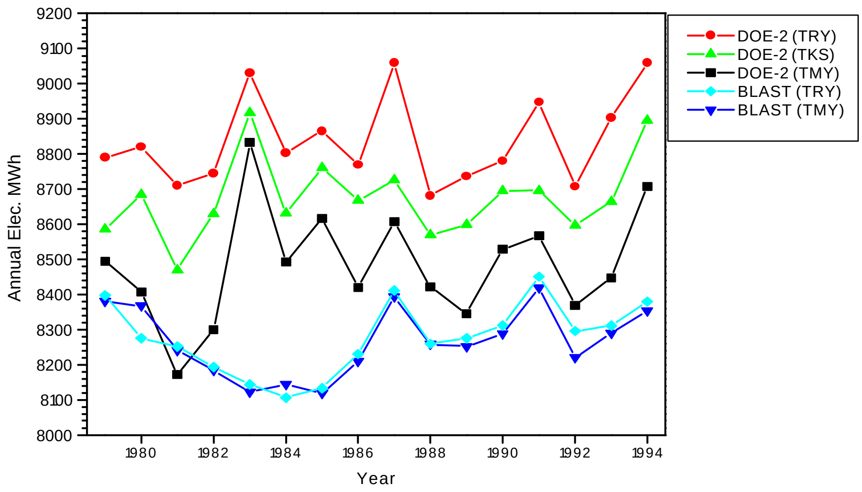

Figure 5.5 *Comparison of Annual Electricity MWh for Different Weather Files*

Table 5.10 Deviations of Annual MWh for Different Weather Files

Program	DOE-2			BLAST	
File format	TMY	TRY	TKS	TMY	TRY
Long-term average (MWh)	8483	8838	8674	8265	8277
Standard deviation (MWh)	162 (1.9%)	127 (1.4%)	114 (1.3%)	97 (1.2%)	101 (1.2%)
Max. deviation (MWh) [Year]	311 (3.7%) [1981]	221 (2.5%) [1991]	243 (2.8%) [1983]	154 (1.9%) [1991]	174 (2.1%) [1991]
Min. deviation (MWh) [Year]	9 (0.11%) [1984]	18 (0.21%) [1980]	7.6 (0.08%) [1986]	7 (0.08%) [1988]	0.8 (0.01%) [1989/80]
Typical year deviation (i.e. year 1989)	138 (1.6%)	101 (1.1%)	76 (0.9%)	13 (0.16%)	0.8 (0.01%)

Notes: 1. The percentage given in parenthesis is obtained by dividing the deviation by the long-term average.

Monthly variations

Figure 5.6 shows the correlation of the monthly MWh values for different weather files. The DOE-2 TMY files are chosen as the base (x-axis) and the results for other files are compared with them. It can be seen that the results of the three DOE-2 formats are closely related (points tend to cluster near a straight line) while the BLAST results are related quite well.

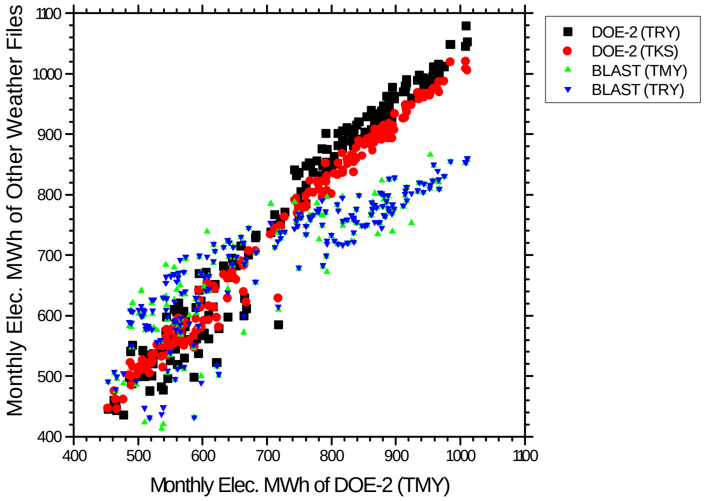

Figure 5.6 *Correlation of Monthly Electricity MWh for Different Weather Files*

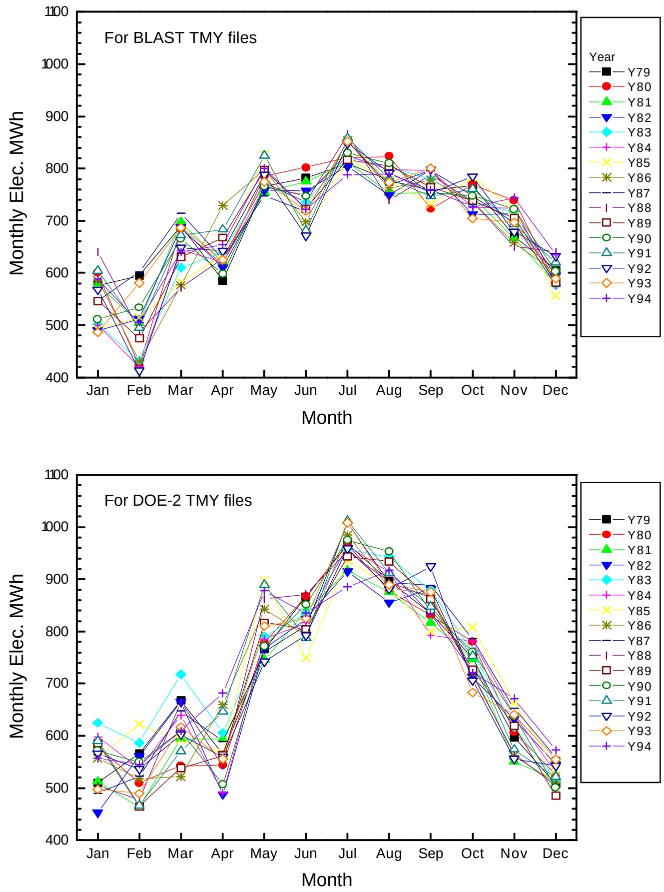

Figure 5.7 *Monthly Electricity MWh from DOE-2 and BLAST*

To study the seasonal changes of energy consumption, the monthly MWh profiles of the TMY format files of DOE-2 and BLAST are plotted and compared in Figure 5.7. It is found that the BLAST profiles are slightly flattened as compared with the DOE-2 and both programs indicate a peak consumption in July. The general shape of the MWh profiles resembles the monthly temperature and GSR profiles of Hong Kong as shown in Section 4.4.1 (Figures 4.2 to 4.4).

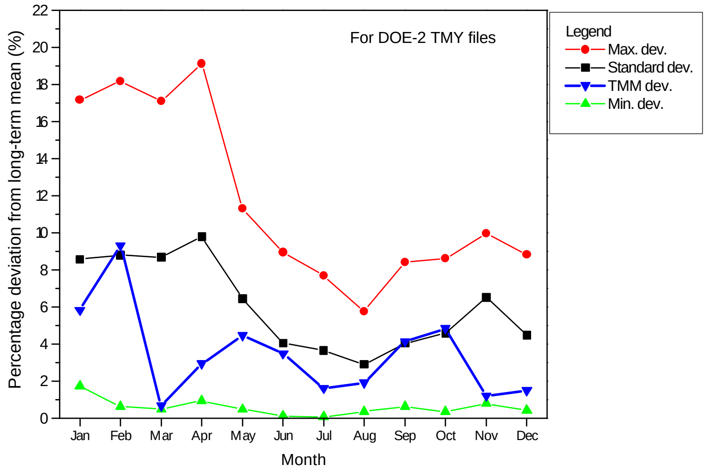

Figure 5.8 *Deviations of Monthly MWh for DOE-2 TMY Files*

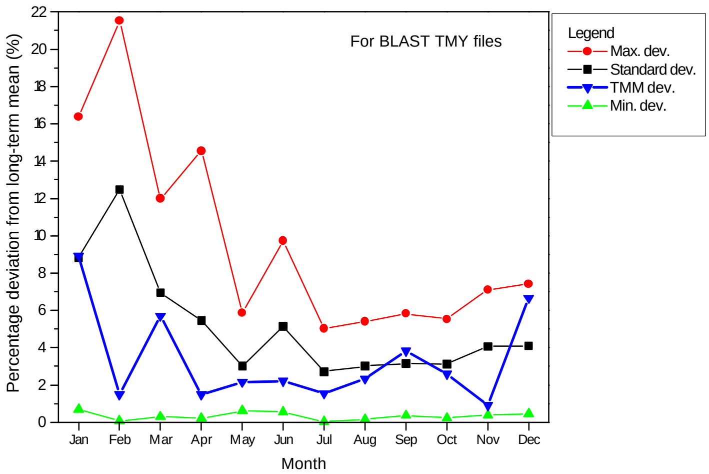

Figure 5.9 *Deviations of Monthly MWh for BLAST TMY Files*

Figures 5.8 and 5.9 gives the percentage deviations of the monthly MWh for the DOE-2 TMY and BLAST TMY files, respectively. The maximum deviations range from 5.8% to 19.1% in DOE-2 and from 5% to 21.5% in

BLAST. Generally, the variations are larger at the monthly level than at the yearly level and the monthly MWh is sensitive to the selection of typical month. In most cases, the deviations of the selected TMM (shown as thick lines in the figures) lie between the minimum and standard deviations. This implies that many of the selected months have performance quite near to the long term, but not necessarily the closest ones.

5.4.2 Correlation with climatic variables

To examine how the climatic conditions relate to the performance of the typical years / months, major weather parameters and their effects on building electricity MWh and peak kW have been studied. The monthly means, medians and standard deviations of six weather parameters (DBT, WBT, DPT, GSR, WSP and RHM) for the years 1979-1994 have been determined and the cooling degree-days (base temperature 18.3^{oC}) for this period have been calculated on a monthly basis. These monthly parameters are correlated with the simulated monthly MWh and peak kW. Table 5.11 gives a summary of the correlation coefficients (*R*) of the monthly parameters (a negative value of R means a negative slope of the relationship). It is found that the cooling degree-days have the highest correlation (*R* = 0.947 for MWh and 0.925 for peak kW) while the monthly means and medians of temperatures show quite good results (*R* = 0.88 to 0.91). WSP and RHM are not sensitive to the energy performance (*R* = -0.55 to 0.57) whereas GSR gives moderate results (*R* = 0.66 to 0.76).

Table 5.11 Correlation Coefficients of Weather Parameters

	Correlation coefficient for monthly means for 1979-94						
	DBT	WBT	DPT	GSR	WSP	RHM	CDD183
MWh	0.912	0.907	0.890	0.756	-0.120	0.315	0.947
Peak kW	0.907	0.918	0.913	0.687	-0.117	0.419	0.925
	Correlation coefficient for monthly medians for 1979-94						
	DBT	WBT	DPT	GSR	WSP	RHM	
MWh	0.909	0.901	0.883	0.724	-0.137	0.254	
Peak kW	0.900	0.910	0.906	0.663	-0.132	0.363	
	Correlation coefficient for monthly standard deviations for 1979-94						
	DBT	WBT	DPT	GSR	WSP	RHM	
MWh	-0.630	-0.770	-0.762	0.735	0.021	-0.549	
Peak kW	-0.608	-0.753	-0.769	0.690	0.035	-0.573	

Notes: 1. DBT = dry-bulb temperature; WBT = wet-bulb temperature; DPT = dew-point temperature; GSR = global solar radiation; WSP = wind speed; RHM = relative humidity; CDD183 = cooling degree-days (base 18.3 ^{o}C); MWh = monthly electricity MWh; Peak kW = peak electricity kW (monthly).

To assess the relative importance of the weather parameters, an attempt has been made to find out a mathematical expression relating the monthly values of simulated MWh and peak kW to the monthly values of the weather parameters using multiple linear regression method (Norusis, 1993a). Table 5.12 gives the regression coefficients determined for the climatic variables. It can be seen that the monthly MWh is most sensitive to the standard deviation of GSR and the mean WBT, DBT and DPT; the peak kW is most sensitive to the standard deviation of WBT and the median and standard deviation of DPT. The multiple coefficient of determination (R^2) for MWh and peak kW are 0.9335 and 0.9423, respectively. This indicates that the monthly descriptive statistics (mean, median and standard deviation) of the weather parameters are useful for locating the years which have energy performance close to the long term.

Table 5.12 Multiple Linear Regression of Weather Parameters

Weather parameter	Unit	Regression coefficients	
		For MWh	For peak kW
CDD183	oC	2.67	6.69
Mean DBT	oC	37.3	-6.80
Mean WBT	oC	-40.05	-163.22
Mean DPT	oC	-35.91	153.29
Mean GSR	MJ/m²/day	28.77	-159.85
Mean WSP	m/s	-16.98	57.24
Mean RHM	%	15.82	-23.60
Median DBT	oC	-15.61	209.16
Median WBT	oC	14.48	96.46
Median DPT	oC	-1.65	-367.77
Median GSR	MJ/m²/day	-21.95	-4.34
Median WSP	m/s	9.07	-132.02
Median RHM	%	-3.12	88.53
Standard Dev. DBT	oC	-10.79	-176.63
Standard Dev. WBT	oC	31.95	459.88
Standard Dev. DPT	oC	-18.59	-315.42
Standard Dev. GSR	MJ/m²/day	-619.25	17.38
Standard Dev. WSP	m/s	9.76	30.41
Standard Dev. RHM	%	3.23	41.98
Regression constant	–	130 MWh	-1335 kW
Multiple R^2	–	0.9335	0.9423
Standard error	–	42.6 MWh	194.6 kW

Notes: 1. DBT = dry-bulb temperature; WBT = wet-bulb temperature; DPT = dew-point temperature; GSR = global solar radiation; WSP = wind speed; RHM = relative humidity; CDD183 = cooling degree-days (base 18.3 oC); MWh = monthly electricity MWh; Peak kW = peak electricity kW (monthly).

Effects of FS and KS statistics

To assess how well the FS and KS statistics relate to the building energy performance, the monthly values of FS and KS statistics for both daily and hourly indices have been correlated with the monthly MWh and peak kW. Tables 5.13 and 5.14 summarise the correlation coefficients of the hourly and daily indices, respectively. The coefficients of the weighted-sum averages (WS) and the individual components are given.

Table 5.13 Correlation Coefficients of Hourly FS and KS Parameters

	Correlation coefficient of hourly FS for					
	DBT	WBT	DPT	WSP	GSR	WS
MWh	-0.196	-0.373	-0.388	-0.1558	-0.091	-0.330
Peak kW	-0.180	-0.390	-0.416	-0.115	-0.153	-0.358
	Correlation coefficient of hourly KS for					
	DBT	WBT	DPT	WSP	GSR	WS
MWh	-0.270	-0.225	-0.256	0.026	-0.130	-0.278
Peak kW	-0.267	-0.274	-0.302	0.059	-0.151	-0.301

Notes: 1. DBT = dry-bulb temperature; WBT = wet-bulb temperature; DPT = dew-point temperature; GSR = global solar radiation; WSP = wind speed; WS = weighted-sum average; MWh = monthly electricity MWh; Peak kW = peak electricity kW.

Table 5.14 Correlation Coefficients of Daily FS and KS Parameters

	Correlation coefficient if daily FS for									
	Max DBT	Min DBT	Mean DBT	Max DPT	Min DPT	Mean DPT	Max WSP	Mean WSP	Total GSR	WS
MWh	-0.127	-0.171	-0.011	-0.253	-0.307	-0.097	-0.130	-0.009	-0.152	-0.196
kW	-0.121	-0.198	-0.029	-0.285	-0.341	-0.153	-0.124	0.017	-0.215	-0.257
	Correlation coefficient if daily KS for									
	Max DBT	Min DBT	Mean DBT	Max DPT	Min DPT	Mean DPT	Max WSP	Mean WSP	Total GSR	WS
MWh	-0.048	-0.179	-0.015	-0.103	-0.230	-0.053	-0.043	0.055	-0.165	-0.169
kW	-0.056	-0.228	-0.049	-0.156	-0.262	-0.120	-0.030	0.101	-0.226	0.229

Notes: 1. DBT = dry-bulb temperature; WBT = wet-bulb temperature; DPT = dew-point temperature; GSR = global solar radiation; WSP = wind speed; WS = weighted-sum average; MWh = monthly electricity MWh; kW = peak electricity kW.

It can be seen from Tables 5.13 and 5.14 that the hourly indices give better correlation than the daily ones. However, the correlation coefficients are low in all these cases (R is all below 0.4) which means the FS and KS statistics do not relate well to the simulated performance. The result implies that for the case in this research, the current TMY approach cannot ensure that the simulated performance is close to the long term at the monthly level.

5.4.3 Rethink about typical years

The study of the simulated performance in the previous sections has revealed some deficiencies of the ASHRAE TRY method and the TMY method. As Klein (1976, pp. 155-156) pointed out, the adequacy of using an average year with a simulation model to provide an estimate of the long-term system performance depends on the sensitivity of system performance to the weather sequences. It is found in this research that the monthly performance is more sensitive than the yearly performance to the change of weather years. If only the annual MWh is concerned, then the error of choosing a less typical (entire) year is not so critical (maximum 1.9% to 3.5% deviation). But, if the monthly MWh is concerned, the effect of choosing a less typical month can be large (maximum 5% to 21.5% deviation).

To select a typical year which can provide a good estimate of the long-term system performance (both annual and monthly), suitable criteria which relate well to the intended application and performance are required. However, it is found that the nonparametric statistics (FS and KS statistics) of the TMY method do not have good correlation with the simulated performance. It seems that the TMY approach is limited when applied to the particular climate and conditions in this research for Hong Kong. The descriptive statistics of the weather parameters are found useful for indicating the simulated performance and this may be a good potential for revising the typical year so that the performance can be better reflected. To help quantify the degree of uncertainty of the typical years, the percentage deviations of the TRY and TMY for Hong Kong are given in Table 5.15 (the minimum deviations achieved by selecting the est’ months are also shown). In term of deviations for the whole year, the TMY is slightly better than the TRY. But when the summer months are concerned (i.e. June to September), the TRY is considered better. For the comparative studies in Chapter 6, the TRY (1989) is recommended since it can give better results for summer months which is important for Hong Kong. Although this comparison is rather crude, it can give us an indication of how well the typical years perform.

Table 5.15 Percentage Deviations of Selected TRY and TMY

	Percentage deviation from long-term mean (%)												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
TRY [1989]	7.1	11.6	12.1	1.8	1.3	2.2	1.6	3.7	1.1	2.3	1.2	7.7	1.6
TMY (hourly KS)	5.8 [80]	9.3 [88]	0.7 [93]	3.0 [87]	4.5 [86]	3.5 [90]	1.6 [89]	1.9 [86]	4.2 [87]	4.9 [84]	1.2 [89]	1.5 [85]	1.5
Min. dev.	1.8 [94]	0.7 [87]	0.5 [90]	1.0 [85]	0.5 [93]	0.1 [82]	0.1 [92]	0.4 [79]	0.6 [91]	0.4 [81]	0.8 [80]	0.4 [80]	0.1

Note: 1. The figure in the square bracket is the year selected for that month or year.

2. The percentage deviation for TMY and Min. dev. (minimum deviations) is determined from the difference between the sum of all monthly consumption selected and the multi-year annual average of the consumption.

Simulation using multi-year data

A direct but cumbersome method to avoid the typical year problems is to use the multi-year dataset in the simulation. Running building energy simulations for multiple years might be a formidable job in the past due to the enormous amount of computations involved. But, with fast development of computing power, it is not a very difficult task now and will be easier with better and faster number crunching machines. It is time to rethink about the approach to weather data for building energy analysis. It is found in the evaluation of weather files that simulation using the multi-year weather data can provide more information for analysing the long-term climatic conditions and energy properties. Provided that the simulation process has been fully automated (see also Section 6.2.2), it is feasible to carry out ulti-year simulation’ to develop the full perspective of the climate. The use of multi-year dataset can also offer solutions to the problems of design weather and typical weather since the long term can be represented by one dataset which is used simultaneously for determining the design loads and energy performance. The author believes that more research is needed to explore this option. Of prime importance is the progress of the simulation methods which make use of the weather data.