Y:/Sage/LRT/Articles/LRT074483/LRT074483.3d[x] Friday, 1st December 2006 19:26:45

Lighting Res. Technol. 39,1 (2007) pp. 1–12

Predicting daylight availability dynamically based on forecasts of a weather observatory

E Ng PhD RIBA FRSA, **A Gadi** BSc MSc, **F Wong** BSc MPhil, **J Mu** BA MPhil, **A Gadi** BSc and **M Lee** BSc Meng

AQ2 Department of Architecture, The Chinese University of Hong Kong, Hong Kong

Received 18 October 2006

By using measured data of the research class IDMP station, a subset of three sky types and their probability of occurrence has previously been established to represent sky conditions of Hong Kong. This was dubbed the Hong Kong Representative Sky (HKRS). The HKRS allows a better prediction of daylight on the vertical surfaces of buildings than the CIE Overcast Sky. The HKRS was still a static representation. This paper investigates the possibility of using the weather predictions of the observatory to establish dynamically the HKRS. A methodology is proposed, and the results indicate that the approach is feasible. It is established that the weather prediction information issued by the Hong Kong Observatory (HKO) provides a reliable means for estimating probabilities of the sky types at a particular moment. By using this information, a better prediction of vertical sky component (VSC) on building surfaces, with a reduction of 20–30% mean absolute error, can be achieved. Engineers and architects may use this more accurate information to design more dynamically.

1. Introduction

The Hong Kong Representative Sky (HKRS) consists of three sky types – CIE General Sky type 1, 8 and 13[1] – representing overcast, intermediate and clear sky, respectively, could characterise sky conditions of Hong Kong.[2] The HKRS is represented by:

$$\text{HKRS} = \rho_1 \ \text{sky} \ 1 + \rho_8 \ \text{sky} \ 8 + \rho_{13} \ \text{sky} \ 13 \qquad (1)$$

where ρ_1 = probability of sky type 1 = 0.42; ρ_8 = probability of sky type 8 = 0.36; and ρ_{13} = probability of sky type 13 = 022.

It was concluded that a better estimation of vertical sky component (VSC) on building facades, which is a major component of calculating the daylight factor (DF) of interiors, could be obtained by using the HKRS as compared to the standard CIE Overcast Sky model. In an earlier paper,[2] the usefulness of using HKRS in daylight prediction and modelling was established.

In addition, it was found that although the HKRS provides a better basis for estimating VSC compared to the CIE Overcast Sky model, the overall error, measured by using the root mean squared error (RMSE), is still large for practical purposes. For instance, the RMSE was found to be 42% for a surface facing East (Table 1).

The generic rigidity of the HKRS – the sky type probabilities are assumed to be constant – is the reason for the poor performance. Also, a ‘comprehensive description of a daylight climate’ that can allow dynamic simulation of building performance,[3,4] requires the probabilities of the different sky types at a particular moment, hence the ‘static’ probabilities in Equation 1 do not suffice for this purpose. Without resorting to measured data, how could a dynamic prediction be achieved?

 10.1177/1477153507074483

Table 1 Error analysis based on the representative sky and sky 1

	East	West	North	South
HKRS				
RMSE (%)	42	32	21	28
Sky 1 (CIE overcast sky)				
RMSE (%)	51	52	29	44
Reduction in error (%)	18	38	27	36

This paper attempts to establish a relationship between HKRS sky types (sky type 1, 8 and 13) and the local weather forecast issued by the Hong Kong Observatory (HKO). Since the forecast is issued several times a day, it might provide means for an instantaneous and accurate estimate of the HKRS sky types probabilities, which could be used more 'dynamically' to:

- Obtain a better estimate of VSC, which, in turn, will enable a better DF prediction;
- Better predict solar illuminance (irradiance) on an inclined surface – this can be done following the method outlined by Li and Lam,[5] where the sky radiance data is used to predict solar irradiance on inclined surface.

2. The HKO key words

The HKO,[6] issues the local weather forecast on its website at regular intervals during the day. On a typical day at a given time, the forecast may take the following form:

Cloudy with sunny periods. There will also be a few showers. Moderate to fresh southerly winds. (29 April 2005: 13:45)

The underlined words in the forecast above are the key words, which have been described very precisely and unambiguously on the HKO's website.[6]

Special care was given to the order in which the terms appear in the forecast. Therefore, 'cloudy with sunny periods' is treated as not the same as 'sunny periods with clouds'. The weather forecast generally follows the format in Table 2.

The local weather forecast data in the period April 2003 to June 2005 was combined with the sky type analysis to construct a table in the following form shown in Table 3. The sky type information in this table was obtained from reference 2.

A few assumptions were made to construct a table with a similar format to that used for Table 3. These assumptions were:

1) A weather forecast for a given time is assumed to be valid for one hour. For example, if a weather forecast is available for 07:45 hours, it applies to all times between 07:45 and 08:45 hours.
2) On many occasions, no data were available because the weather forecast was not recorded. Therefore, the following steps were adopted to solve the problem.
    a) If no data were available for a particular day then it was completely ignored, although a weather forecast from the night before might have been available.

Table 2 Typical format for local weather forecast

HKO's typical format for local weather forecast			
Key word	1 (HKO I)	2 (HKO II)	3 (HKO III)
Weather forecast	Describes immediate weather	Adds more precision to key word 1	Generally provides a more distant outlook, very rarely adds precision to key word 1 and key word 2
Wind information	Describes the wind type, two key words are used rarely	Describes wind direction	

Table 3 Sample table showing a summary of HKO key words and sky types

Date (dd/mm/yy)	Time	Sky type	HKO I	HKO II	HKO III	Wind type	Wind direction
06/04/2004	07:45	Sky 1	C	B	R	Moderate	E
06/04/2004	08:05	Sky 1	C	B	R	Moderate	E
06/04/2004	08:15	Sky 1	C	B	R	Moderate	E
06/04/2004	08:25	Sky 1	C	B	R	Moderate	E
06/04/2004	08:35	Sky 1	C	B	R	Moderate	E

C, cloudy; B, bright intervals; R, rain; E, east.

b) If no data were available for a given period, then the forecast of adjacent time periods was used provided the adjacent forecast was almost the same. For example, if no data were available for the period 08:45–09:45 hours, then provided that the data for 07:45–8:45 hours were the same as for 09:45–10:45 hours, then the forecast for the period was assumed to be the same as for the adjacent periods.

3) A weather forecast issued around 17:00 hours describes the night time weather, which is not relevant to the purpose of this paper, and, therefore, there are almost no data available for the period after 17:00 hours.

The above assumptions allowed approximately 8600 data points (sky scans) to become available for analysis out of the potential 25 444 data points available for the period from 24 April 2003 to June 2005.

3. Methodology

3.1 Identify dominant key words

As mentioned earlier, the first two key words are the most important key words in the weather forecast issued by the HKO. This is because, generally speaking, the first two words almost always describe the weather in the immediate future, whereas the third key word describes weather in the more distant future. Therefore, only the first two key words were considered for this analysis.

Table 4 presents an analysis that was conducted to determine the most frequently used key words. It becomes clear that there are five dominant key words for HKO I, namely, cloudy, fine, showers, sunny intervals and sunny periods, and together they account for almost 97% of the total number of observations. Therefore, for the purpose of establishing sky type probabilities, only the most dominant key words and their combinations were considered.

3.2 Combination of key words and estimating sky type probability

As mentioned earlier, HKO I is usually accompanied by HKO II to provide more precision to the weather forecast. For example, the key word 'fine' may appear as 'fine with cloudy' or just 'fine'. Therefore, one can estimate sky type probability with any two combinations of HKO I and HKO II provided they are observed in the given data period. For example, consider the case of 'fine with haze' (Table 5).

Table 4 Frequency of occurrence of the most frequently used key words for HKO I in the weather forecasts issued by the Hong Kong Observatory

HKO I Frequency of occurrence	
Key word	Frequency (%)
Cloudy	37.1
Fine	44.5
Showers	1.3
Sunny intervals	1.4
Sunny periods	12.7
Total	97

Table 5 Estimating sky type probability for key words 'fine with haze'

Fine	Sky 1	Sky 8	Sky 13	Total no. of observations
Haze	60	692	422	1174

The above information can be used to estimate the probability associated with 'fine with haze', which was found to be 0.05, 0.59 and 0.36 for sky 1, sky 8 and sky 13, respectively.

3.3 The accuracy of the estimate and the margin of error analysis

The sky type probability in Step 2 was estimated by using the data comprising a sky type and HKO key words (as shown in Table 3). However, from a practitioners perspective, it is important to know the margin of error associated with an estimate. For this purpose, 'z-interval procedure' was used. The steps for z-interval are:[7]

1) For a confidence interval (CI) of $1-\alpha$, use the standard normal table to find $z_{\alpha/2}$.
2) The margin of error or the estimate of ρ, a sky type probability, is

$$E = z_{\alpha/2} \cdot \sqrt{\hat{\rho}\ \ (1-\hat{\rho})/n} \quad (2)$$

where $z_{\alpha/2}$ is found in the first step, n is the sample size, and $\hat{\rho}$ is the estimated value of a sky type probability.

For example, consider the margin of error analysis for the estimated sky type probabilities associated with 'fine with haze'. First consider the analysis for $\hat{\rho}=0.59$, ie, the probability of sky 8.

1) If 95% confidence interval (CI) is desired then $\alpha=0.05$, and for 99% confidence level, $\alpha=0.01$. From the standard normal table: $z_{\alpha/2}=1.96$ and 2.58 for 95% and 99% confidence levels, respectively.
2) Equation 2 gives $E=0.03$ and $E=0.04$ for 95% and 99% confidence level, respectively.

Hence, one could be 95% (99%) confident that the error in estimating the probability, $\hat{\rho}=0.59$, for sky type 8 for 'fine with haze' key words is at most 0.03 (0.04), ie, plus or minus three (or four) percentage points.

However, there are three estimates, and, hence, one must calculate the margin of error for all three probability estimates. It is logical to select the largest value of the margin of error, as this approach will guarantee that the confidence level specification will be either met or bettered. For 95% confidence level specification, $E=0.01$ for sky type 1 and $E=0.03$ for sky type 13. Therefore, one could be at least 95% confident that the error in estimating sky type probabilities associated with key words 'fine with haze' is at most 0.03, ie, plus or minus three percentage points.

Table 6 The sky type probabilities and the HKO's key words (HKO 1 is fine)

HKO I	HKO II	Sky 1	Sky 8	Sky 13	No. of observations	*E* for 95% CI	*E* for 99% CI
Fine	Only	0.08	0.46	0.46	1862	0.02	0.03
	Cloudy	0.03	0.53	0.43	120	0.09	0.12
	(Fog)	0.23	0.77	0	13	0.23	0.30
	Haze	0.05	0.59	0.36	1174	0.03	0.04
	Isolated showers	0.19	0.47	0.35	397	0.05	0.06
	Mist	0.09	0.68	0.23	141	0.08	0.1
	Showers	0.12	0.42	0.46	152	0.08	0.1
Overall		0.07	0.52	0.41	3859	0.02	0.02

CI, confidence level.

Table 7 The sky type probabilities and the HKO's key words (HKO 1 is cloudy)

HKO I	HKO II	Sky 1	Sky 8	Sky 13	No. of observations	*E* for 95% CI	*E* for 99% CI
Cloudy	Only	0.72	0.28	0.01	250	0.06	0.07
	Bright	0.84	0.16	0	68	0.09	0.12
	(Drizzle)	1	0	0	10	n/a	n/a
	(Fine)	0.44	0.32	0.24	25	0.19	0.26
	Fog	0.89	0.11	0	101	0.06	0.08
	Haze	0.88	0.12	0	115	0.06	0.08
	(Isolated showers)	0.5	0.05	0	26	0.19	0.25
	Mist	0.86	0.14	0	204	0.05	0.06
	Overcast	0.93	0.06	0.01	99	0.05	0.07
	(Occasional showers)	0.25	0.75	0	4	0.42	0.56
	Rain	0.85	0.14	0.01	824	0.02	0.03
	Showers	0.75	0.22	0.03	676	0.03	0.04
	Sunny intervals	0.67	0.27	0.06	279	0.06	0.07
	Sunny periods	0.45	0.46	0.09	341	0.05	0.07
	(Squally showers)	0.74	0.23	0.03	31	0.15	0.2
Overall		0.76	0.21	0.03	3217	0.01	0.02

CI, confidence level.

For this paper, the threshold for the margin of error (E) at 0.1 for 95% confidence level. If the E value is >0.1, the estimated probability is not considered accurate for practical purposes. This is indicated later in Tables 6–10 by the enclosed HK II keywords.

4. Results

Tables 6–10 provide sky type probabilities associated with HKO I of 'fine', 'cloudy', 'sunny periods', 'showers' and 'sunny intervals'.

Each table provides the probabilities of sky types associated with the observed combinations of HKO I and HKO II. Also provided are the overall sky type probabilities associated with HKO I, which can be used if the user observes a combination not reported in the table. The overall sky type probabilities are particularly useful for 'sunny periods', 'showers' and 'sunny intervals' key words, since the number of observations is not enough to provide a detailed insight.

5. Discussion

Based on the Observatory's weather prediction key words, the probabilities of the three sky

Table 8 The sky type probabilities and the HKO's key words (HKO 1 is sunny periods)

HKO I	HKO II	Sky 1	Sky 8	Sky 13	No. of observations	*E* for 95% CI	*E* for 99% CI
Sunny periods	Only	0.28	0.49	0.23	300	0.07	0.07
	Cloudy	0.41	0.54	0.06	145	0.08	0.11
	(Fine)	0.15	0.85	0	26	0.14	0.18
	Haze	0.20	0.62	0.19	183	0.07	0.09
	(Isolated-showers)	0.30	0.55	0.15	89	0.10	0.14
	(Mist)	0.27	0.55	0.18	77	0.11	0.15
	(Rain)	1.00	0	0	3	n/a	n/a
	Showers	0.31	0.53	0.16	275	0.06	0.08
	(Scattered showers)	0	1	0	4	n/a	n/a
Overall		0.29	0.55	0.17	1102	0.03	0.04

CI, confidence level.

Table 9 The sky type probabilities and the HKO's key words (HKO 1 is showers)

HKO I	HKO II	Sky 1	Sky 8	Sky 13	No. of observations	*E* for 95% CI	*E* for 99% CI
Showers	(Only)	0.33	0.67	0	6	0.38	0.50
	(Fog)	0	0.86	0.14	7	0.26	0.34
	(Hazy)	0	1	0	7	n/a	n/a
	(Sunny periods)	0.39	0.45	0.16	56	0.13	0.17
	(Squally showers)	0.62	0.35	0.04	26	0.19	0.25
	(Thunder storms)	0.69	0	0.31	13	0.25	0.33
Overall		0.43	0.44	0.13	115	0.09	0.12

CI, confidence level.

types could be 'dynamically' estimated. The approach should be able to provide a better estimate of VSC compared to the 'static' probabilities in Equation 1. This argument can be validated by comparing the accuracy of VSC estimated using the two methods. A small sample of approximately 100 data points was selected for this comparative study. To ensure that this sample correctly represents the total population, the following steps were taken:

1) The sample size contains data from throughout the day.
2) For each hour, 8–12 data points were randomly selected. This ensured that data were well spread throughout the period under investigation (April 2003 to June 2005).
3) Table 11 was generated to check that the sample correctly represented the population. It confirmed that the sample was a good representation of the population, as the frequency of 'key words' is close to that of the total population (Table 4).

The sample was used to study the usefulness of the sky type probability approach. The VSC, based on both the 'dynamic' and the 'static' method, was compared to the observed values of VSC, which were calculated using the sky scanner data. Table 12 shows the mean values of the absolute error in estimating VSC in four cardinal directions based on the two methods. It is clear from Table 12 that the 'dynamic' method has resulted in a useful improvement in estimating VSC. Moreover, the standard deviations of the error based on the 'dynamic' method in the four orientations are constantly lower than the 'static' method, which implies the performance of the 'dynamic' method should be more stable.

A Student's *t*-test was conducted to confirm the superiority of the new method statistically. The test hypothesis is stated as follows.[7]

Table 10 The sky type probabilities and the HKO's key words (HKO 1 is sunny intervals)

HKO I	HKO II	Sky 1	Sky 8	Sky 13	No. of observations	*E* for 95% CI	*E* for 99% CI
Sunny intervals	(Only)	0.7	0.3	0	30	0.16	0.22
	(Cloudy)	0.62	0.38	0	13	0.26	0.35
	(Haze)	0.67	0.33	0	24	0.19	0.25
	(Showers)	0.68	0.26	0.06	50	0.13	0.17
	(Squally showers)	1	0	0	2	n/a	n/a
Overall		0.68	0.29	0.03	119	0.08	0.11

CI, confidence level.

Table 11 Frequency of 'key words' in the randomly generated sample of 100 data points

	Fine	Cloudy	Sunny periods	Showers	Sunny intervals
Frequency (%)	46	43	10	1	1

The null hypothesis is 'the difference between the population mean of the absolute error produced by the dynamic method and the static method is zero'. The alternative hypothesis is 'the difference between the population mean of the absolute error produced by the dynamic method and the static method is greater than zero'. Mathematically, the test hypothesis can be written as

$$H_o{:}\mu_{|\mathrm{Error(oldmethod)}|} - \mu_{|\mathrm{Error(newmethod)}|} = 0$$

$$H_a{:}\mu_{|\mathrm{Error(oldmethod)}|} - \mu_{|\mathrm{Error(newmethod)}|} > 0$$

where $\mu_{|\mathrm{Error(oldmethod)}|}$ is the population mean of the absolute the static method and $\mu_{|\mathrm{Error(newmethod)}|}$, the population mean of the absolute the dynamic method.

Before applying the Student's t-test, it is necessary to check if our samples meet the test requirement, which is the sample distribution has to be normal or the sample size is large enough (normally 30) given that the sample distribution is not too far from normal. After using the normal probability plot for testing the normality of the four set of samples, which are shown in Figures 1–4, the samples roughly fall on the straight lines, which means the sample distributions are not too far from normal. In addition, as one already has >100 samples of VSC error for each orientation, the test requirement is met and the Student's t-test could be applied reliably.

According to Table 13, which is the result of the Student's t-test, the P-value calculated by the samples of the absolute error of East, South, West and North VSC are 0.018, 0.001,

Table 12 Mean error is four cardinal directions in estimating VSC for the two methods

Orientation	Method	Mean absolute error	Standard deviation
East	Static	0.15	0.18
	Dynamic	0.11	0.13
South	Static	0.11	0.12
	Dynamic	0.07	0.08
West	Static	0.11	0.13
	Dynamic	0.07	0.08
North	Static	0.07	0.06
	Dynamic	0.05	0.04

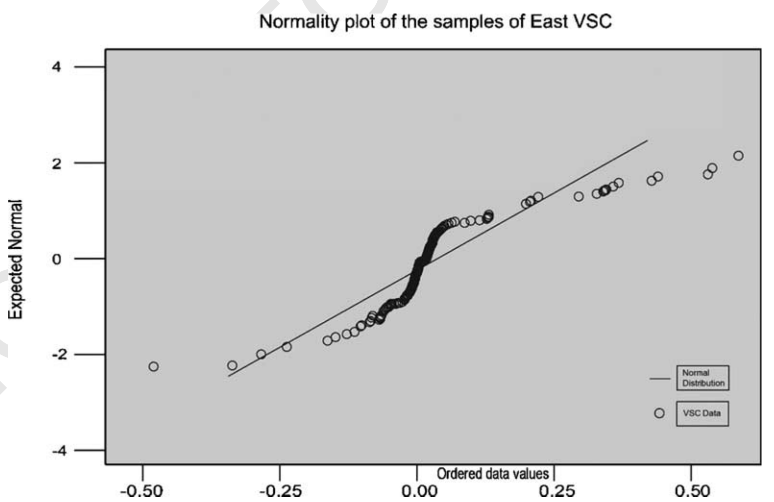

Figure 1 Normality plot of the samples of the difference of absolute error between the old and new methods for East VSC

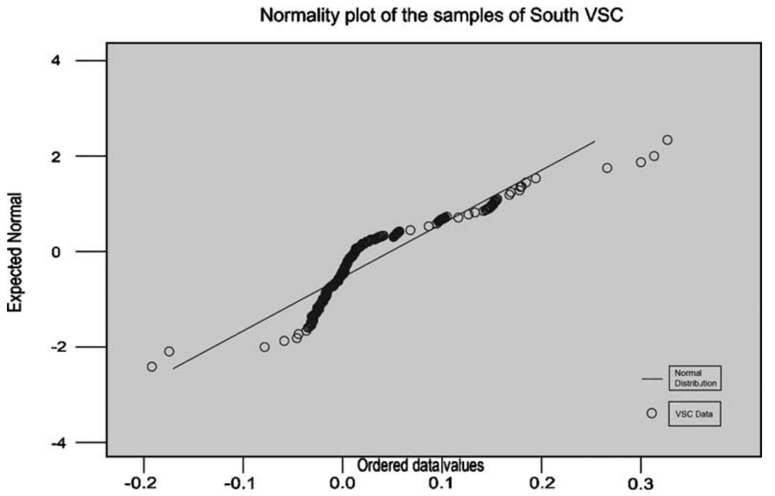

Figure 2 Normality plot of the samples of the difference of absolute error between the old and new methods for South VSC

0.001 and 0.014, respectively. As all four values are less than the critical value 0.05, we should reject the null hypothesis and conclude that the new method is superior in all orientations with 95% confidence level.

6. Conclusion

This paper introduces briefly, a previous attempt to define a HKRS.[2] Its 'static' nature has been a concern. A dynamic model is desired. In lieu of relying on measured data, which could be expensive, as well as not normally readily available, an attempt to reference the HKO's weather prediction information is made. This demonstrates that the dynamic model performs better. For example, if the following is noted from the Observatory:

Cloudy with sunny periods. There will also be a few showers. Moderate to fresh southerly winds. (29 April 2005; 13:45)

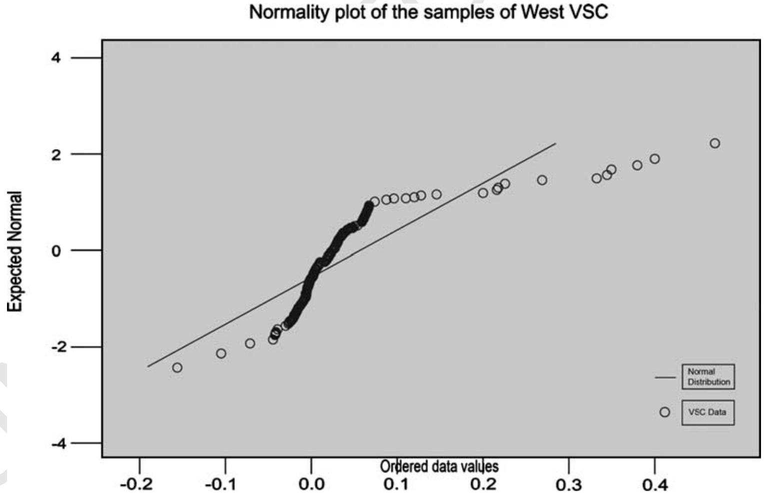

Figure 3 Normality plot of the samples of the difference of absolute error between the old and new methods for West VSC

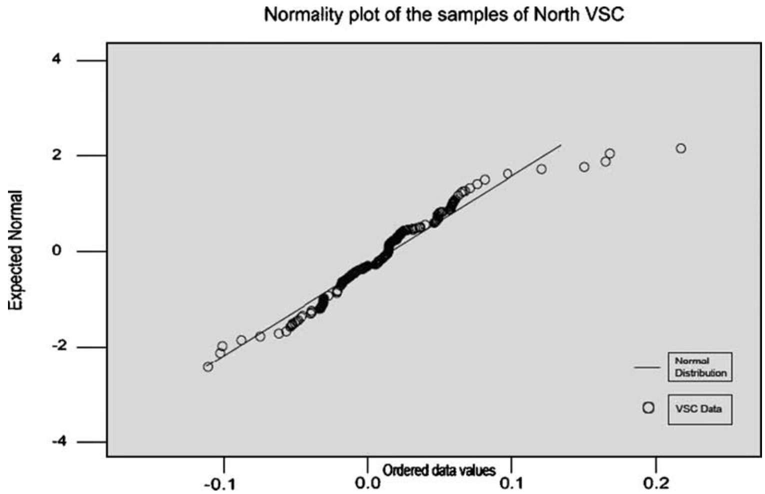

Figure 4 Normality plot of the samples of the difference of absolute error between the old and new methods for North VSC

Then Equation 1 will become

$$\text{HKRS} = 0.45 \ \text{sky} \ 1 + 0.46 \ \text{sky} \ 8 + 0.09 \ \text{sky} \ 13.$$

Using the Observatory information, which is on-line and is regularly updated, engineers can estimate, in almost real time and conveniently, the daylight availability and perhaps the irradiation load of the building surfaces. This information, together with probability of sunshine, can be used to dynamically control the buildings for daylight and solar radiation considerations. Table 14 provides an easy reference.

Table 13 Result of the Student's *t*-test for the difference of absolute error of VSC produced by the old and new methods in four orientations

	t-Value	Degrees of freedom	*P*-value (one-tailed)
East	2.679	105	0.018
South	4.649	105	0.001
West	4.676	105	0.001
North	2.748	105	0.014

7. Further work

This study makes use of a relatively short-term measured data. Comparison of Table 7

Table 14 A summary of the sky type probabilities and the HKO's key words

HKO 1	HKO II	Sky 1	Sky 8	Sky 13
Fine	Only	0.08	0.46	0.46
	Cloudy	0.03	0.53	0.43
	Haze	0.05	0.59	0.36
	Isolated showers	0.19	0.47	0.35
	Mist	0.09	0.68	0.23
	Showers	0.12	0.42	0.46
Cloudy	Only	0.72	0.28	0.01
	Bright	0.84	0.16	0
	Fog	0.89	0.11	0
	Haze	0.88	0.12	0
	Mist	0.86	0.14	0
	Overcast	0.93	0.06	0.01
	Rain	0.85	0.14	0.01
	Showers	0.75	0.22	0.03
	Sunny intervals	0.67	0.27	0.06
	Sunny periods	0.45	0.46	0.09
	Scattered showers	0.87	0.10	0.03
Sunny periods	Only	0.28	0.49	0.23
	Cloudy	0.41	0.54	0.06
	Haze	0.20	0.62	0.19
	Showers	0.31	0.53	0.16

and Table 14 shows that some keywords are not featured. To address that, a larger number of observations are needed so that, ultimately, a more comprehensive Table 14 can be formulated.

A possible further investigation to follow up this work would be to examine closely the reasons how and why the weather conditions of the HKO key words characteristically affect the sky and daylight conditions. This may be useful later to develop a physical and theoretical model.

Another possible follow-up investigation may be to apply the Dynamic HKRS to attempt a Building Management System (BMS) optimisation procedure, and to evaluate its usefulness and reliability under real building conditions.

References

1 CIE, Commission Internationale de I'Eclairage. CIE Standard (CIE S 011/E:2003) spatial distribution of daylight – CIE standard general sky. CIE Vienna, 2003.
2 Edward N, Vicky C, Ankur G, Mu J, Max L, Ankit G. Defining standard skies for Hong Kong, *Buildings Environ* 2007; 42(2): 866–76.
3 Tregenza PR. Standard skies for maritime climates. *Lighting Res Technol* 1999; 31(3): 97–106.
4 Tregenza PR. Analysing sky luminance scans to obtain frequency distributions of CIE Standard General Skies. *Lighting Res Technol* 2004; 36(4): 271–81.
5 Li DHW, Lam J. Predicting solar irradiance on inclined surfaces using sky radiance data. *Energy Conservation Manage* 2004; 45: 1771–84.
6 The Hong Kong Observatory. www.hko.gov.hk.
7 Weiss NA. *Introductory statistics*, sixth edition. Addison-Wesley, 2002.

Discussion

Comment 1 on 'Predicting Daylight Availability Dynamically Based on Forecasts of Weather Observatory' by E Ng, Ankur Gadi, F Wong, Jun Mu, Ankit Gadi, and M Lee

D H W Li (City University of Hong Kong, Dept of Building and Construction, Tat Chee Avenue, Kowloon, HONG KONG SAR, CHINA)

This is a very interesting paper which investigates the possibility of using weather forecasts from the local meteorological station to predict the standard sky types. There are a few points that I would like to ask the authors.

The prediction approach relies on the identification of the dominant key words issued by the Hong Kong Observatory. Is it possible to standardise all weather forecast key words used by the Hong Kong Observatory?

Hong Kong is a small city but contains many residential and business districts. Some of them may be far away from the city centre (over 20km) such that the sky conditions observed from these places may be different from the weather forecast. Do such variations severely affect the proposed approach?

The study allows engineers to estimate the real time daylight availability and the irradiation load of the building surfaces. Would the authors further suggest some practical applications from which the users can benefit?

Apart from reference 5, we have published a paper that uses sky luminance data to predict daylight illuminance on inclined surfaces. Both the two papers are based on the same approach.

Reference

Li DHW, Lau CCS, Lam JC. Predicting daylight illuminance on inclined surfaces using sky luminance data. *Energy – The International Journal* 2005; 30(9): 1649–65.

Comment 2 on 'Predicting Daylight Availability Dynamically Based on Forecasts of Weather Observatory' by E Ng, Ankur Gadi, F Wong, Jun Mu, Ankit Gadi, and M Lee

PR Tregenza (School of Architecture, The University of Sheffield, SHEFFIELD, S10 2TN UK)

Measurements of daylight, such as the data collected in the International Daylight Measurement Project, are of limited value in themselves. Long-term records from a well-run measurement station, may give a statistical description of the available light at a particular place. A set of measurements made over a few days or months offers only a description of what happened at that place at that time.

Like any other scientific data, the general value of daylight measurements lies in the extent to which they provide points of reference within a predictive model. Daylight has been measured at few points on the globe: we need to be able to forecast its availability everywhere.

The first step towards this is to examine the relationship between daylight and other, more commonly measured, values, such as solar irradiation. The next is to link daylight quantitively with established categories of climate. We want to be able to predict the standard sky types and the illuminances that are likely to occur if, for example, all we know about a place is its latitude, its probability of sunshine through the year, and that it has a temperate maritime climate.

This is a realistic objective, and by associating daylight measurements with standard weather descriptions the authors have made an important step towards generalisation of IPMP data. It is potentially of considerable value, and I would like the authors to describe ways in which they see the work being continued.

Authors' response to DHW Li and PR Tregenza

E Ng, Ankur Gadi, F Wong, Jun Mu, Ankit Gadi, and M Lee

We thank Dr Li and Professor Tregenza their observations.

To address Dr Li's queries, the Hong Kong Observatory has a standardised list of key words. This follows the World Meteorological Office's guidelines. The "weather terms" the observatory uses can be accessed from their website. Unfortunately, the Hong Kong Observatory only issue one weather forecast for the entire territory of Hong Kong, there is at present no way for one to check if variations exist. Apart from the possibility of predicting daylight availability and irradiation load, the information could be used to predict the 'near real time' solar renewable energy potentials.

The authors thank Professor Tregenza for laying out the scenario upon which our paper is making a small contribution. The queries raised at the end of the discussion have actually been mostly addressed in the paragraphs preceding it. The authors are grateful for the hints.

So far we have only made a statistical correlation between the two sets of data. The initial investigation confirms the possibility and the practical value of using one to predict the other. Of course, our current two year IDMP sky scan data are not enough to capture all the combinations of weather terms used by the Observatory to describe Hong Kong's skies. Further funding permitted; the data set needs to be extended.

The next question to ask is whether or not by looking further into the definition of the keywords, one could understand the theoretical basis of the correlation. For example, the Hong Kong Observatory's definition of the weather term "cloudy" is 'The sky is covered with a total cloud amount of between six eighths and eight eighths'; for "bright", it is 'The sky is covered by a large amount of thin cloud with sunshine occasionally'; for "isolated showers", it is 'The rain-bearing cloud amount is small and isolated, resulting in showers occurring only in small parts of the territory at a time'; and for "fine", it is 'The sky is covered by a total cloud amount of less than six eighths. However, it can still be described as fine even though the total cloud amount is greater than six eighths if the cloud layer is thin enough to let plenty of sunshine to penetrate'. The four keywords are related to the nature of cloud in the sky. How would the characteristics of the cloud structure in a typical tropical maritime sky be further understood as the basis for knowing the sky types?

Currently, the authors are examining the possibility of using satellite images, which are, like weather data, becoming increasingly available, to correlate the observations and the sky types. One of the key objectives is precisely this need for using the images to give an indication of the cloud structures. Initially, it was found that the satellite image pixel on its own might not be enough. A generic correlation might not be possible. There might be a need to be weather or location specific; that is to say, there might be a need to combine other ground weather observation data to formulate a more holistic understanding. Suffice to say that our current speculations are only preliminary, and a lot more work is needed.