

# Towards Improving Visibility Forecasts in Taiwan: A Statistical Approach

John Chun-Han Lin<sup>1,\*</sup>, Jyh-Huei Tai<sup>2</sup>, Chin-Hui Feng<sup>2</sup>, and De-En Lin<sup>2</sup>

<sup>1</sup> Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, Ontario, Canada

<sup>2</sup> Air Force Weather Wing, Taipei, Taiwan, ROC

Received 7 October 2008, accepted 13 May 2009

---

## ABSTRACT

Reduced visibility is a major safety concern at airports leading to flight delays or diversions. The primary motivation for this study is to enhance understanding of visibility and to build a simple, practical visibility forecasting method across Taiwan in an operational setting by using readily available, ground-based observations.

This paper presents for the first time a systematic, quantitative examination of the controls on visibility over the entire Taiwan region by adopting a statistical approach relating visibility to various physical variables. A multiple linear regression is carried out for the early morning hours during the months of November ~ April, when visibility is especially low. The regression reveals that on the west coast of Taiwan, the concentration of fine particles ( $PM_{2.5}$ ) and relative humidity ( $RH$ ) are most related to visibility, and to a lesser extent, coarse particles ( $PM_{10-2.5}$ ) and windspeed. The significantly elevated PM concentrations would, therefore, cause a marked reduction in visibility on Taiwan's west coast - where most of Taiwan's population and anthropogenic PM emissions are found. Visibility on the east coast appears to be controlled by somewhat different mechanisms, with rainfall playing a larger role. The probability of occurrence of especially low visibility ( $< 1600$  m) was revealed by logistic regression to be the most statistically related to  $RH$ , and, to a less extent, to PM concentrations.

An uncertainty analysis to understand current limitations in predicting visibility indicated that 24-hour visibility forecasts were dominated by a) errors in forecasting  $RH$  and b) inadequacies in the adopted statistical model, followed by c) errors in PM forecasts. Hence to minimize the considerable uncertainties in visibility forecasts, which could reach standard deviations of several thousand meters, future work needs to adopt a more sophisticated statistical model as well as reduce the considerable errors in predicting  $RH$ . Finally, the uncertainties associated with PM can be reduced by improving PM emission estimates through an inverse analysis method.

Key words: Visibility, Weather forecasting, Statistical analyses, Particulate matter, Airport operations

Citation: Lin, J. C. H., J. H. Tai, C. H. Feng, and D. E. Lin, 2010: Towards improving visibility forecasts in Taiwan: A statistical approach. *Terr. Atmos. Ocean. Sci.*, 21, 359-374, doi: 10.3319/TAO.2009.05.13.01(A)

---

## 1. INTRODUCTION/MOTIVATION

Visibility is defined as the “greatest distance in a given direction at which an object can be visually identified with unaided eyes (Yuan et al. 2006).” Reduced visibility has long been recognized as a major concern at airports (George 1960; Martin and Suckling 1987) by causing flight delays or diversions, with significant attendant financial costs (Teixeira and Miranda 2001). Given the large impact of visibility on aviation, accurate forecasts are important. Imperfect visibility forecasts adversely affect airport operation, resulting in diverted flights when visibility is lower than expected and

lost capacity when visibility is higher than expected (Müller et al. 2007). The motivation of this research project is to improve visibility forecasts at various airports around Taiwan. A related objective is to gain insight and scientific understanding of factors controlling visibility.

We adopt an empirical, statistical approach to characterize visibility with a goal towards employing the results in an operational forecasting setting. A simple multiple linear regression method is adopted, as this work is envisioned to provide an initial, overall picture of the factors controlling visibility around the entire Taiwan region. The dataset derives from readily available, ground-based observations at airports as part of routine aviation weather reports (METARs).

---

\* Corresponding author  
E-mail: jcl@uwaterloo.ca

The reasons for using such “simple” observations are two-fold: 1) wide availability of continuous observations; 2) ease of use when formulating visibility forecasts. “Ease of use” refers to the availability of timely, up-to-date observations that can be readily used for forecasting purposes. This means that quantities such as chemical composition of atmospheric particulates (Hornig and Cheng 2008), although demonstrated to have important consequences for influencing visibility (Tsai and Cheng 1999; Yuan et al. 2006), will not be considered, since such observations are available only on a limited, campaign basis. In addition, we make use of particulate matter (PM) measurements being carried out by automated instrumentation on a continuous basis around Taiwan by the Environmental Protection Administration (EPA) (Chang and Lee 2007a).

To our knowledge from the published literature, a statistical modeling approach to predict visibility has yet to be attempted over the entire Taiwan region. This is the first time a large number of stations are analyzed to reveal spatial patterns in visibility around Taiwan and the factors controlling visibility. Previous studies generally focused on visibility at a single location (e.g., Tsai and Cheng 1999; Yuan et al. 2006) or only a few, generally urban sites (Chen et al. 1999). While the statistical approach may lack direct connection to physical mechanisms, it is nonetheless a first step towards physical understanding, as the statistical results would demonstrate exactly which variables are most related to visibility; also, the regression coefficients themselves may hint at physical mechanisms at play.

After a discussion of the observational dataset and statistical methodology (section 2), this paper proceeds as follows. First, we present figures summarizing the visibility data and identify dominant features in the data (section 3). Secondly, we examine which physical variables are the most important in determining visibility, as a guide towards building the statistical model (section 4.1). Thirdly, we carry out multiple regression to relate visibility to physical variables in a quantitative manner. The multiple regression consists of two parts. The first part (section 4.2) is based upon a linear model that predicts visibility, while the second part (section 4.4) hones in on the *probability of occurrence* of extremely low visibility events ( $< 1600$  m) by using a logistic regression model. A quantitative analysis of uncertainties in forecasts of visibility by using the regression model follows next (section 4.3). Finally, our summary and conclusions are presented in section 5, along with a discussion of how the uncertainties currently limiting accuracy of visibility forecasts can be minimized.

## 2. MEASUREMENTS/METHODOLOGY

### 2.1 Visibility and Meteorological Measurements

A large atmospheric observational database has been collected over numerous years by the Taiwanese Air Force at

13 stations around Taiwan (Table 1). Two of the stations were situated on the small off-shore islands of Magong (QC) and Lyudao (LT). Except for MQ and LT, all of the other stations were close to sea level, with elevations  $< 50$  m above sea level. The geographical locations of the 13 stations can be seen in Fig. 1.

At these stations measurements of visibility and standard meteorological variables such as temperature, dew point, rainfall, wind speed, and wind direction were carried out at least every hour. All of these variables were measured as part of an automated weather observation system (AWOS). These stations were outfitted with the AWOS 2000 (Artis Weather Check, Ohio, USA) in the early 1990s and are serviced and calibrated every 3 months. A description of the AWOS 2000 can be found in Arpino (1994).

The visibility measurement is referred to as “horizon visibility” and is defined as the visibility in the horizontal direction, in all 360 degrees. These observations derive primarily from human observers looking out towards the horizon, at objects of a known distance. The observations were supplemented with measurements from the AWOS-based visibility sensor, particularly under conditions of extremely low visibility, when distant objects can no longer be distinguished.

The data were reported following the format of routine aviation weather reports (METARs). Temperature and dewpoint were reported as whole numbers, thus introducing errors in the derived relative humidity (see section 4.3 for details). Visibility observations were reported at discrete levels following International Civil Aviation Organization (ICAO) guidelines (ICAO Annex 3 2007), at every 1000 m above 5000 m and with the maximum value set to 9999 m.

This study makes use of data from the years of 2005 ~ 2007. 2005 was the first year in which  $PM_{2.5}$  was monitored widely around Taiwan on a continuous basis. The 2005 and 2006 data were used in fitting the regression model, while 2007 was used as an independent dataset to evaluate the statistical fit and to quantify uncertainties in extrapolating the fitted statistical coefficients to other years.

### 2.2 Particulate Matter (PM) Measurements

The visibility dataset was combined with hourly measurements of particulate matter (PM) concentrations (in  $g\ m^{-3}$ ) carried out by Taiwan’s Environmental Protection Administration (EPA). Particulate matter refers to “any material, except uncombined water, that exists in the solid or liquid state in the atmosphere or gas stream at standard condition” (Malm 1999).

PM concentration measurements were carried out at stations in the Taiwan Air Quality Monitoring Network (TAQMN) (Chang and Lee 2007a). Observations of both  $PM_{2.5}$  and  $PM_{10}$  were available, where  $PM_{2.5}$  and  $PM_{10}$  refer to PM below aerodynamic diameters of 2.5 micrometers

Table 1. Site information for the 13 observational stations used in this study and the average values of observed quantities during the years 2005 ~ 2007. “sub” refers to the subset of data used for the multiple linear regression (section 2.3.1): only the hours 0400 ~ 0800 LT, and months of November ~ April, years 2005 ~ 2006.

Station Name	Latitude	Longitude	Elevation [m asl]	PM station [km]	% of Hrs < 1600 m (sub)	Visibility		PM <sub>2.5</sub>		RH		Windspeed			
						[m]	(sub) [m]	(sub) [ $\mu\text{g m}^{-3}$ ]	(sub) [ $\mu\text{g m}^{-3}$ ]	(sub) [%]	(sub) [%]	(sub) [ $\text{m s}^{-1}$ ]	(sub) [ $\text{m s}^{-1}$ ]		
PO	24.83	120.93	11	3.9	5.97	6827	33.87	18.97	83.95	8.67	7984	31.02	20.19	78.61	9.38
GM	25.05	121.22	48	10.6	5.90	6988	32.21	28.14	83.71	8.654	8082	33.95	29.14	78.61	9.02
MQ	Chingchuan Kang	24.23	120.60	205	5.2	5678	36.52	24.05	87.18	8.425	7199	32.72	27.65	82.28	8.73
KU	Chiayi	23.47	120.37	29	6.2	5148	53.18	34.84	91.89	4.542	7188	46.10	35.85	85.22	5.63
NN	Tainan	22.95	120.20	24	4.8	5577	50.01	36.18	86.70	5.877	7666	41.43	31.79	80.64	6.05
AY	Gangshan	22.83	120.30	13	7.9	5065	59.63	32.86	88.55	3.752	7324	46.71	33.79	81.39	5.37
SQ	Pingtung (North)	22.70	120.47	32	3.2	4841	55.07	32.68	86.06	0.6317	6794	47.81	32.17	78.95	2.48
DC	Pingtung (South)	22.67	120.45	27	3.2	4856	55.07	32.68	85.80	1.134	6818	47.81	32.17	78.60	2.97
QC	Magong	23.57	119.62	35	19.1	6676	27.28	23.21	89.54	13.23	7606	23.57	23.15	87.44	11.82
YU	Hualien	24.17	121.62	19	22.1	8885	18.50	12.27	86.89	4.781	8853	20.15	17.08	82.82	5.66
CS	Chiashan	24.29	121.59	22	35.6	8845	18.50	12.27	88.26	2.205	8819	20.15	17.08	83.89	3.57
ZN	Taitung	22.77	121.20	16	6.1	9227	15.76	13.02	78.64	4.244	8962	15.29	19.38	77.72	6.00
LT	Lutao	22.66	121.48	281	36.3	8103	15.76	13.02	85.42	11.96	8422	15.29	19.38	82.67	10.41

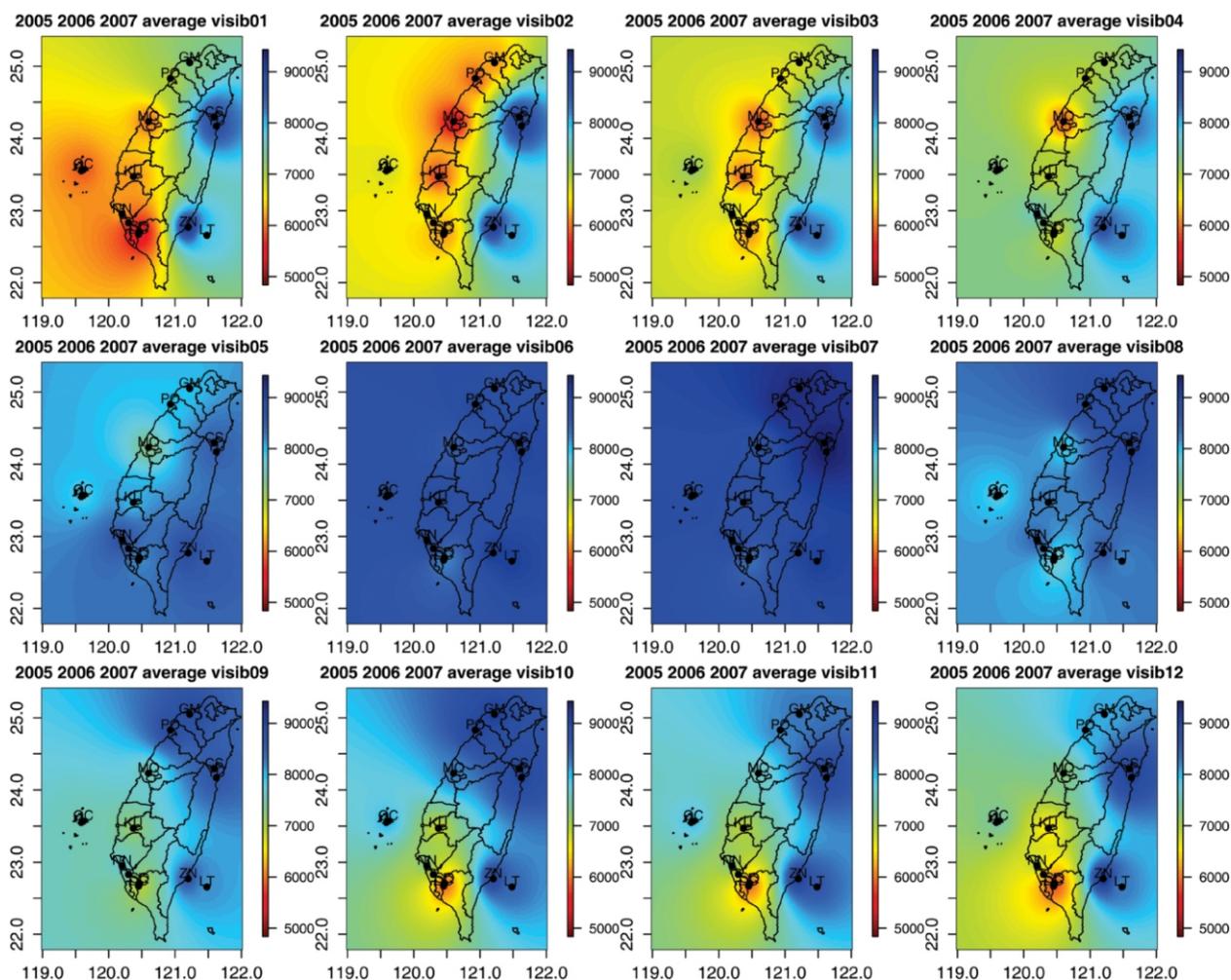


Fig. 1. Monthly averaged visibility (years 2005 ~ 2007) based on the 13 observational stations around Taiwan and outlying islands. The continuous fields were constructed by interpolating the station data using a Kriging method (Venables and Ripley 2002). Note the color scale increases from red (low visibility) to blue (high).

and 10 micrometers, respectively. The PM measurements were carried out by the Beta attenuator method (Chang and Lee 2007a) using the beta gauge automated particle monitor (EQPM-0391-081) for  $PM_{10}$  and model EQPM-0990-076 with a 2.5  $\mu$ m cut size inlet for  $PM_{2.5}$  (Chang and Lee 2007b). TAQMN instruments underwent regular daily zero and span checks, as well as calibration (Chang and Lee 2007a, b). Following Yuan et al. (2006), we refer subsequently to  $PM_{2.5}$  and  $PM_{10-2.5}$  as the concentration of “fine” and “coarse” particles, respectively.  $PM_{10-2.5}$  was calculated by simply taking the difference between  $PM_{10}$  and  $PM_{2.5}$ .

Since PM measurements were not coincident with those from the Air Force, we merged the two datasets by simply choosing PM measurement sites closest to each Air Force airport. Distances between airport and PM measurement sites ranged from 1 to 36 km (Table 1), resulting in uncertainties due to spatial displacement in visibility and PM. The

displacement at LT was as much as 36 km, because no PM observations were available on Lyudao Island; the closest one was on the main island of Taiwan. In consideration of the ocean and large distance separating the visibility and PM measurements, we removed PM from consideration at LT in subsequent analyses.

## 2.3 Statistical Modeling

### 2.3.1 Multiple Linear Regression

The degree to which different variables were “useful” in predicting visibility was first established by fitting a linear, multiple regression model and then calculating the change in Akaike Information Criterion (AIC; Akaike 1974) when each explanatory variable was dropped from the regression (Venables and Ripley 2002). AIC is a measure of the goodness of fit of a statistical model and often used for model selection. AIC is defined as follows (Akaike 1974):

$$\text{AIC} = -2\text{maximized log-likelihood} + 2 \# \text{ parameters} \quad (1)$$

The coefficient of  $-2$  means that *lower* AIC is associated with *higher likelihood* (statistically “superior”). However, a cost is imposed when the tighter fit is achieved by additional parameters to the model (second term on the RHS). If dropping a particular variable resulted in a lower AIC, we removed it from consideration in the subsequent multiple regression. Detailed results from the AIC analysis will be shown in section 4.1.

From the AIC analysis the following linear model was established:

$$\begin{aligned} \text{Visib} = & M_i + B_{(100 - RH)} * [100 - RH] + B_{PM_{2.5}} * [PM_{2.5}] \\ & + B_{PM_{10-2.5}} * [PM_{10-2.5}] + B_{rain} * [Rain] \\ & + B_{wind} * [Windspd] \end{aligned} \quad (2)$$

where *Visib* indicates visibility, *RH* is relative humidity,  $PM_{2.5}$  and  $PM_{10-2.5}$  are the particulate matter concentration of fine and coarse particles, respectively, and *Windspd* is the windspeed. *Rain* indicates the presence of rain: 1 when rain is present, 0 when otherwise.  $M_i$  and  $B_{var}$  (where *var* indicates any explanatory variable) are the parameters to be solved in the regression.  $M_i$  denotes a “month” factor with *i* denoting the various months of the year.  $M_i$  can be also thought of as the “base-line visibility” with  $RH = 100\%$  and in the absence of PM, rain, or wind.  $(100 - RH)$  can be regarded as a measure of the deviation from saturation - i.e., a measure of the atmosphere’s “dryness.”  $(100 - RH)$  was chosen rather than *RH*, because the mean values of *RH* were high over Taiwan (Table 1), closer to 100% than to 0%. This meant that selecting 100% as the “origin” - i.e., the point of departure - was a natural choice. In any case, due to the linear relationship between *RH* and  $(100 - RH)$  values of  $B_{var}$  for other explanatory variables would not be changed with the choice of  $(100 - RH)$ .

Equation (2) is simple and is by no means meant to be exhaustive in its consideration of variables that might control visibility (synoptic weather systems, strength of turbulence, mixed-layer height are just a few of the other possible variables). It also does not attempt to fit a separate  $B_{var}$  for each month. The relative simplicity of the statistical model follows from the overarching goal of this study, which is meant as an initial step towards gaining an overall understanding of visibility around the entire Taiwan region rather than focusing on details at individual sites.

Values of  $M_i$  and  $B_{var}$  were established through a least-square method that minimizes the squared differences between the observed and calculated visibilities. We carried out the AIC and multiple regression calculations using the function “lm” within “R”, an open source data analysis software (R Development Core Team 2005).

## 2.3.2 Logistic Regression

We further zeroed in on the variables controlling occurrence of particularly low visibility events, defined as when visibility dipped to values  $\leq 1600$  m. The threshold of 1600 m was chosen following personal communication with personnel in the Air Force familiar with airport operations, who pointed out this as the value below which airport operations are significantly affected by the low visibility.

Because in this case the variable to be explained takes on only two possible outcomes (presence/absence of low visibility), a statistical model is used to predict the probability of occurrence of extremely low visibility events ( $p$ ).

We adopt the logistic regression method to predict  $p$  (Dalgaard 2002):

$$\begin{aligned} \text{logit}(p) = \log_e[p/(1 - p)] = & m_i + b_{(100 - RH)} * [100 - RH] \\ & + b_{PM_{2.5}} * [PM_{2.5}] + b_{PM_{10-2.5}} * [PM_{10-2.5}] \end{aligned} \quad (3)$$

The logit function transforms the probability such that the response variable is no longer limited to values between 0 and 1.  $p/(1 - p)$  is known as the “odds”. Thus positive (negative) values of  $m_i$  and  $b_{var}$  mean that the  $\log_e$  (odds) of the event occurring is enhanced (reduced). For instance, if  $b_{var} = 0.693$ , then the odds increase by a factor of  $\exp(0.693) = 2$  for a unit increase in *var*.

$m_i$  and  $b_{var}$  were fitted through the method of maximum likelihood. Variables such as rainfall, windspeed, and wind direction were not sufficiently related to low visibility to be considered in Eq. (3). Similar to the multiple linear regression, regression variables were established by examining AIC. Hours when visibility was  $\leq 3200$  m, close to the threshold of 1600 m, were selected for analysis. Thus the AIC selects for variables that are statistically varying with visibility at low values, identifying the variables that may control the onset of extremely low visibility events.

## 3. OBSERVED PATTERNS

We first show general temporal and spatial features identified in the data.

### 3.1 Observed Patterns in Visibility

#### 3.1.1 Spatial Patterns in Visibility

The spatial distribution of mean visibility in each month (average of years 2005 ~ 2007) is shown in Fig. 1. In most months visibility along the western coast of Taiwan is noticeably lower than that on the east. During the winter months of November ~ January visibility is particularly low in the southwestern region of the island, with average values dipping to 6000 m or so. In contrast, visibility is significantly higher during the summer months of June ~ August, with mean values of approximately 9000 m throughout Taiwan.

Note that these averages are biased toward lower values, since the reported visibility had a ceiling of 9999 m.

### 3.1.2 Temporal Patterns in Visibility: Diurnal + Monthly

As an example of the mean diurnal variation of visibility, Fig. 2 shows the average diurnal cycle at various months for AY, close to the port city of Kaohsiung in southwest Taiwan. A distinct diurnal pattern can be observed in the non-summer months, in which the visibility reaches the diurnal maximum at 1500 LT and decreases afterwards throughout

the night to especially low values at dawn, reaching the lowest levels at 0700 LT. The same time of day was also reported by Yuan et al. (2006) to be when the lowest visibility at Kaohsiung was observed. The impact of rainfall on visibility is particularly evident in the summer as witnessed in the higher (gray points) visibility when hours with rainfall are removed. The aforementioned pattern in AY is widely repeated in other west coast stations (not shown).

In contrast, the diurnal pattern for YU on the east coast (Fig. 3) shows no distinct morning drop in visibility. There appears instead to be a decrease in the afternoon, which is not seen in the hours without rainfall (gray points). This pro-

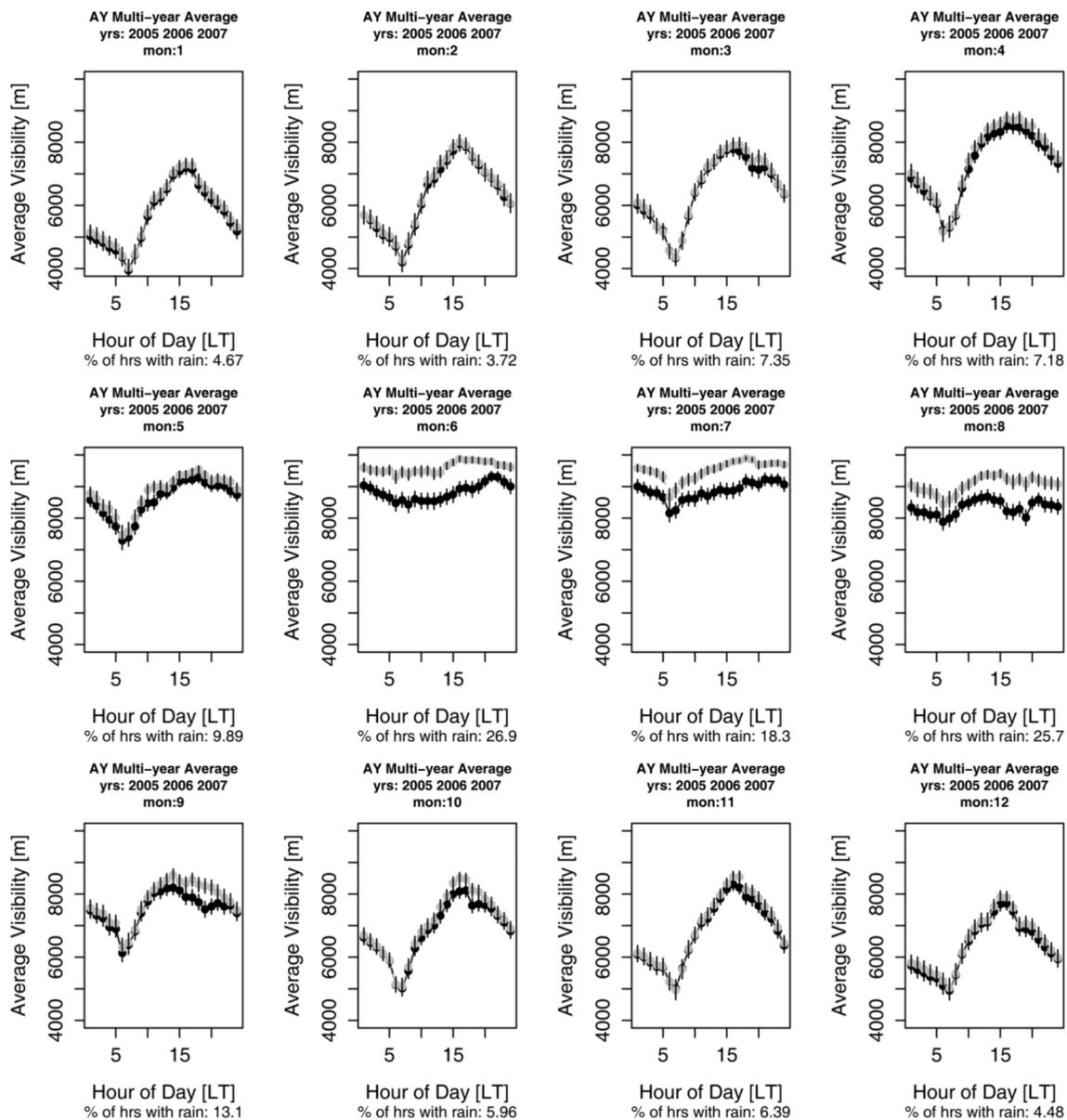


Fig. 2. Monthly averaged diurnal cycle of visibility (years 2005 ~ 2007) at AY (Gangshan). The x-axis is in local time (LT). The black points derive from the entire dataset, while the gray points come from a dataset when hours with rainfall are filtered out. The percentage of hours in the month when rainfall occurred is also indicated. The error bars indicate the standard error (which is the standard deviation divided by  $N^{1/2}$ , where  $N$  is the sample size).

nounced rainfall effect is observed throughout the year, instead of being restricted to the summer months as it is for AY and other west coast stations. The qualitative features observed in YU are also observed for the other two east coast stations - CS and ZN.

Due to the fact that the higher levels of visibility do not cause significant impact on aviation, we focused on the times with the most pronounced drop in visibility. Half of the year (May ~ October), in which high visibility was observed over most of Taiwan (Fig. 1), was excluded from subsequent statistical analyses. Furthermore, we focused on the hours 0400 ~ 0800 LT, when especially low visibility is observed over a majority of stations (e.g., Fig. 2).

### 3.2 Observed Patterns in PM

The spatial distribution of mean PM<sub>2.5</sub> concentrations in each month (average of years 2005 ~ 2007) is shown in Fig. 4. To some degree this figure shows semblance to the mean visibility (Fig. 1), in which low visibility often corresponds to high PM concentrations (note flipped color scale). MQ was a notable exception during January, March, and April, when low visibility was not associated with particularly high PM concentrations. MQ is situated at a higher elevation, at > 200 m asl on the eastern side of Mt. Dadu and separated from the EPA's PM station, which is only 5 km away but at a lower elevation of < 100 m asl and to the west of Mt. Dadu. We

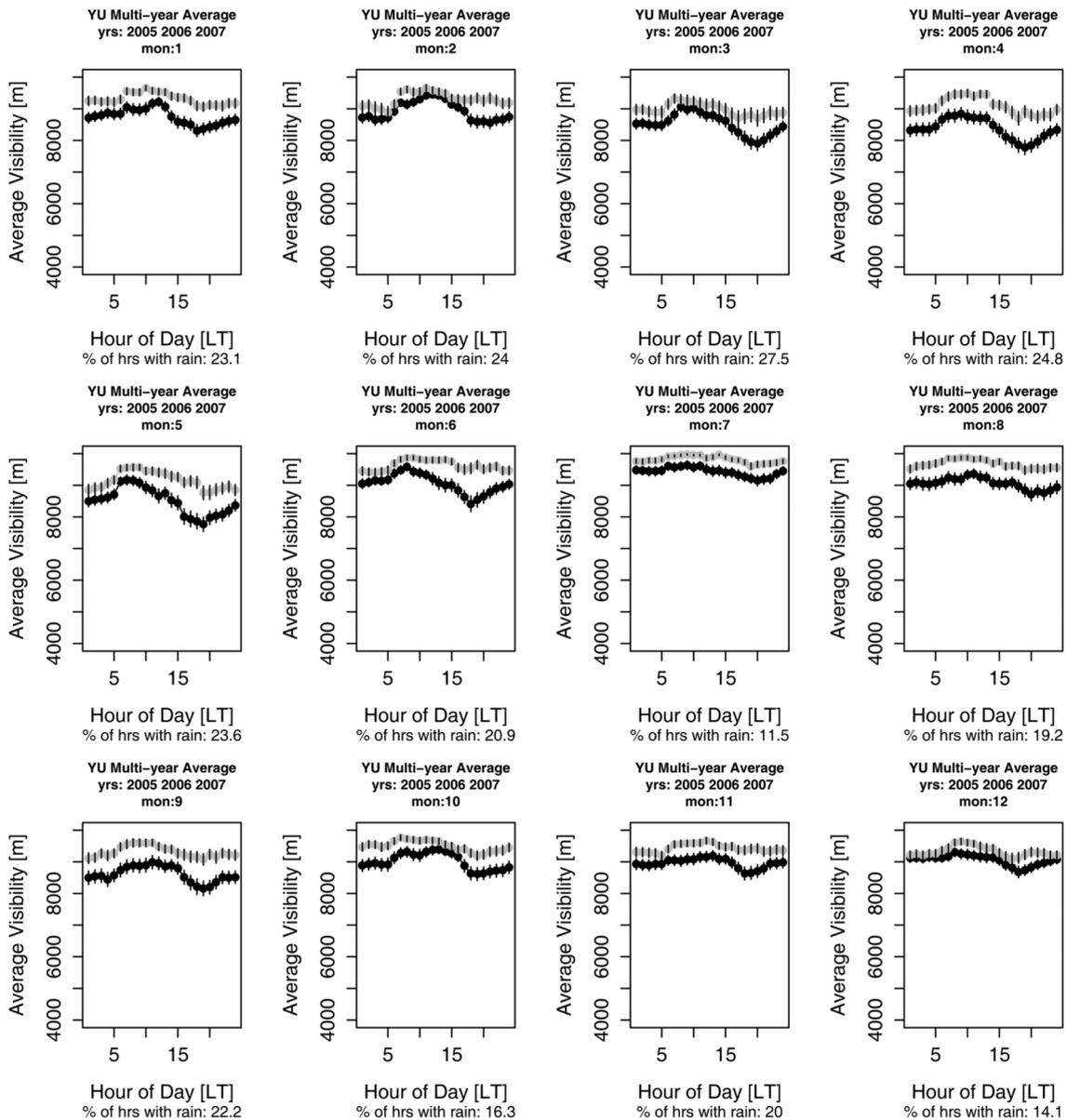


Fig. 3. Same as Fig. 2, but for YU (Hualien).

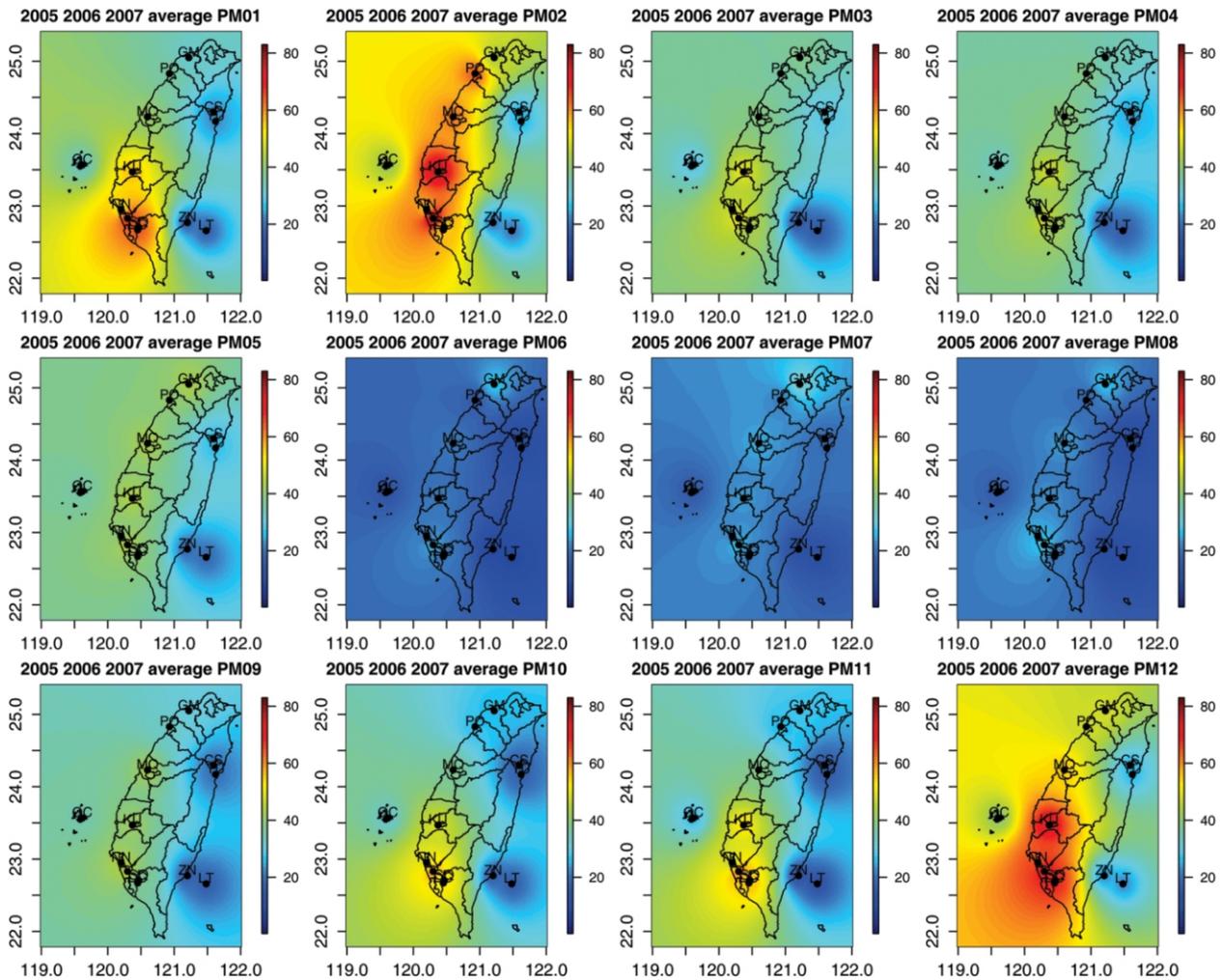


Fig. 4. Monthly average concentrations of  $PM_{2.5}$  (years: 2005 ~ 2007) based on the 13 observational stations; similar to Fig. 1.

believe the lack of correspondence between PM and visibility at MQ to be primarily due to this spatial separation.

Concentrations of  $PM_{2.5}$  are lower in the summer (June ~ August), presumably due in part to rainfall serving as a sink of PM through wet deposition. The highest average  $PM_{2.5}$  concentrations are found in the southwestern part of Taiwan, at stations such as AY, SQ, DC, KU, and NN. For coarse particles ( $PM_{10-2.5}$ ) the highest levels are found at the same stations (Table 1), and the spatial distributions at various months are very similar, with significantly lower concentrations on the east coast.

The main sources of  $PM_{10}$  have been identified as vehicle emissions, industry, secondary aerosols, crustal dust, biomass burning, marine spray, and combustion from a previous study in central Taiwan (Chio et al. 2004). Most of these sources arise from anthropogenic and industrial activity. Most of the Taiwanese population can be found along the western plain region, particularly in the three urban regions of Taipei, Taichung, and Kaohsiung in the northern, central, and southern parts of the western plain, respectively (Chen

et al. 1999). More than 60% of Taiwan's heavy industries are found in Kaohsiung (Yuan et al. 2006), near which the stations AY, SQ, and DC are all found. The high level of PM concentrations observed in southwestern Taiwan naturally follows from these strong emissions. This is in accordance with a prior observational study that showed Kaohsiung to have the highest PM concentrations, on average, among the three major urban regions (Chen et al. 1999).

## 4. RESULTS

### 4.1 Which Variables are Most Important?

Table 2 shows results from an attempt - based on the AIC - to determine which variables are the most important (in a statistical sense) in predicting visibility. "Important" is defined here as the highest incurred "cost" (most increase in AIC; see section 2.3.1) when a variable is removed from consideration in a regression model. In all of the west coast stations the three most important variables are  $M_i$ ,  $PM_{2.5}$ , and  $(100 - RH)$ . For the majority of stations the monthly factor  $M_i$

Table 2. The order of “importance” of each potential explanatory variable, i.e., its usefulness in predicting visibility, to be incorporated into the multiple regression. A variable was regarded as being more important in predicting visibility when a larger increase in Akaike Information Criterion (AIC; section 2.3.1) resulted after the variable is removed from the regression. The shaded area marks out the variables in which the regression coefficient for a variable cannot be rejected at the 0.01 significance level, based on an  $F$ -test. These results are derived for the subset of data used in the multiple linear regression (0400 ~ 0800 LT in November ~ April during 2005 ~ 2006). The “wind sector” denotes classification of wind direction data into the four cardinal directions (N, E, S, W).

	month	RH.unsat	PM <sub>2.5</sub>	rain	PM <sub>10-2.5</sub>	windspd	windsector
PO	1	2	3	4	5	6	7
GM	1	3	2	4	5	7	6
MQ	1	2	3	6	4	5	7
KU	1	3	2	6	4	5	7
NN	1	3	2	5	4	6	7
AY	2	3	1	4	5	6	7
SQ	3	2	1	4	5	6	7
DC	3	2	1	4	5	6	7
QC	1	3	2	4	7	6	5
YU	1	4	3	2	6	7	5
CS	1	4	3	2	6	7	5
ZN	1	4	2	3	6	7	5
LT	1	3	4	2	6	7	5

was the most significant, suggesting that systematic month-to-month variation in visibility exists. Rainfall is more important on the east coast, coming in second at YU, CS, and LT and third at ZN. The fine particles PM<sub>2.5</sub> appear to be more closely associated with visibility than the coarse particles PM<sub>10-2.5</sub>.

The wind-related variables (wind sector and windspeed) are generally less important. The “wind sector” refers to classification of wind direction data into the four cardinal directions (N, E, S, W). At 7 out of the 13 stations wind sector is the least important variable; at 5 of them it was wind speed. At 6 stations wind sector is not statistically significant at the 0.01 level according to the  $F$ -test (null hypothesis: regression coefficient = 0). The same is true for windspeed at 7 stations.

We removed variables revealed as statistically insignificant in Table 2 from the multiple regression to be discussed in the next Section. Furthermore, wind sector was removed from consideration altogether. Other than the fact that wind sector seemed lower in importance than most other factors, the reasons are twofold. In some places almost all of the data come from a single sector; in other cases the wind sector varied systematically from month to month, suggesting strong correlation and lack of independence from the month factor.

In addition, PM was not incorporated in the linear model for LT, because PM measurements were carried out far away from this off-shore island, at Taitung (Table 1). The large separation distance likely contributed in PM being ranked as only the fourth important variable, the lowest ranking among all of the stations.

## 4.2 Regression Results

We discuss results from the multiple regression (shown in Table 3), starting from diagnostics and then moving on to values of the fitted regression coefficients.

### 4.2.1 Model Evaluation: $R^2$ and Residual Standard Error

A rough assessment of the statistical model’s performance can be determined by examining  $R^2$  and residual standard error. Both diagnostics were calculated by comparing 2007 observations versus simulated values. Because the regression coefficients were fitted with 2005 ~ 2006 data, the 2007 observations provide an independent dataset against which to evaluate the regression. By using a separate year, the diagnostics also test the regression model’s capability to extrapolate to years not used in the fitting process.

The regression model accounted for roughly 30% to 60% of the variance, depending on location. The residual standard error is roughly 1000 ~ 2000 m, indicating that the regression-predicted visibility deviates from measurements by this amount, on average.

The lower  $R^2$  values on the east coast (YU, CS, ZN) suggest that visibility at those locations is controlled by processes that are captured only to a limited extent by the regression model. The associated lower residual standard error is a reflection of the fact that visibility on the east coast exhibits less amplitude to begin with (Fig. 3).

Table 3. Multiple regression results for different stations, based on the statistical model [Eq. (2)]:  $V_{isib} = M_i + B_{(100 - RH)} * [100 - RH] + B_{PM_{2.5}} * [PM_{2.5}] + B_{PM_{10-2.5}} * [PM_{10-2.5}] + B_{rain} * [Rain] + B_{wind} * [Windspeed]$ . Variables deemed not statistically significant are removed from consideration, following the results in Table 2. Furthermore, due to the absence of PM observations on Lyudao Island, PM was not considered as a predictor variable there.

Station	$R^2$	residerr [m]	$M_{01}$ [m]	$M_{02}$ [m]	$M_{03}$ [m]	$M_{04}$ [m]	$M_{11}$ [m]	$M_{12}$ [m]	Fitted Regression Equation
PO	0.61	1774	5854	5629	6800	6921	6901	6953	$V_{isib} = M_i + 161.1 * (100 - RH) - 58.39 * PM_{2.5} - 26.54 * PM_{10-2.5} - 1217 [if\ Rain] + 45.37 * Windspeed$
GM	0.56	1898	5955	6400	6846	7063	7890	7491	$V_{isib} = M_i + 139.6 * (100 - RH) - 62.58 * PM_{2.5} - 24.18 * PM_{10-2.5} - 1981 [if\ Rain] + 79.01 * Windspeed$
MQ	0.44	2325	4096	3299	4734	4520	5517	5348	$V_{isib} = M_i + 186.5 * (100 - RH) - 30.93 * PM_{2.5} - 33.74 * PM_{10-2.5} - 1135 [if\ Rain] + 77.38 * Windspeed$
KU	0.48	2060	5808	5107	5756	6766	6492	6933	$V_{isib} = M_i + 190.0 * (100 - RH) - 46.72 * PM_{2.5} - 12.28 * PM_{10-2.5} - 935 [if\ Rain] + 92.40 * Windspeed$
NN	0.50	1769	5790	5656	6135	6167	6570	6803	$V_{isib} = M_i + 176.2 * (100 - RH) - 47.48 * PM_{2.5} - 19.89 * PM_{10-2.5} - 897 [if\ Rain] + 45.01 * Windspeed$
AY	0.54	1716	7010	6901	6928	6844	7472	7609	$V_{isib} = M_i + 176.1 * (100 - RH) - 55.52 * PM_{2.5} - 13.37 * PM_{10-2.5} - 1100 [if\ Rain] - 64.12 * Windspeed$
SQ	0.35	1818	4777	4907	4945	4796	5007	5456	$V_{isib} = M_i + 167.1 * (100 - RH) - 41.91 * PM_{2.5} - 1293 [if\ Rain]$
DC	0.33	1844	4699	4774	4913	4726	4925	5429	$V_{isib} = M_i + 169.7 * (100 - RH) - 41.86 * PM_{2.5} - 1204 [if\ Rain]$
QC	0.47	1911	6021	5952	6806	7105	7203	7094	$V_{isib} = M_i + 114.4 * (100 - RH) - 51.65 * PM_{2.5} - 568 [if\ Rain]$
YU	0.31	1336	9387	9469	9640	9428	9693	9713	$V_{isib} = M_i + 59.77 * (100 - RH) - 55.09 * PM_{2.5} - 1902 [if\ Rain]$
CS	0.33	1295	9304	9396	9575	9288	9559	9676	$V_{isib} = M_i + 63.89 * (100 - RH) - 51.18 * PM_{2.5} - 1825 [if\ Rain]$
ZN	0.35	960.5	9430	9253	9341	9309	9487	9522	$V_{isib} = M_i + 38.10 * (100 - RH) - 53.15 * PM_{2.5} - 5.711 * PM_{10-2.5} - 1685 [if\ Rain]$
LT	0.32	1588	6164	6228	6562	6482	7080	6757	$V_{isib} = M_i + 126.2 * (100 - RH) - 1619 [if\ Rain]$

#### 4.2.2 Month Factor $M_i$

$M_i$  can be thought of as the “baseline visibility” when  $RH = 100\%$  and when no PM, rain, or wind is present [Eq. (2)].  $M_i$  also reflects influences from variables missing from the regression model that co-vary with the month. Examples include wind direction (section 4.1), synoptic conditions (weather system), and cloudiness.

From Table 3 we can see that the fitted values of  $M_i$  are lowest in the months of January and February, which are also the months with the lowest average observed visibility (Fig. 1). While the observed visibility can be thought of as the baseline visibility, plus the additional impact of explanatory variables [similar to Eq. (2)], it is not surprising to see the average visibility track the baseline.

#### 4.2.3 $B_{PM}$

The regression coefficients  $B_{var}$  in Table 3 represent the sensitivity of visibility to each explanatory variable  $var$ ; it is the “leverage” different physical variables have on visibility.

$B_{PM}$  is negative: i.e., visibility declines with increasing PM concentrations. The aerosols comprising particulate matter scatter and absorb radiation (Malm 1999), thereby reducing visibility. In other words,  $B_{PM}$  encapsulates the strength of aerosol light extinction due to absorption and scattering. Previous studies have attributed reductions in visibility to high concentrations of particles with relatively small diameter, ranging from 0.1 to 2 microns (Sloane et al. 1991). These particles are close in size to the wavelength of visible light, hence the higher scattering efficiency (Malm 1999), and comprise most of the fine particles comprising  $PM_{2.5}$ . Indeed,  $B_{PM_{2.5}}$  is more negative than  $B_{PM_{10-2.5}}$ .  $B_{PM_{10-2.5}}$  ranges from -6 to -34 m per  $\mu\text{g m}^{-3}$  of coarse particles but increases in magnitude to -31 ~ -63 in the case of  $PM_{2.5}$ . The sole exception was MQ, where  $B_{PM_{2.5}}$  and  $B_{PM_{10-2.5}}$  are comparable, at  $\sim -30$  m per  $\mu\text{g m}^{-3}$ . This is likely because of the spatial mismatch and effect from Mt. Dadu separating the visibility and PM observations at MQ (see section 3.2), with errors propagating into the derived  $B_{PM}$ .

The variability observed in  $B_{PM}$  at various sites may be due, to a large extent, to the varying chemical composition of particles. Among fine particles the light scattering coefficient is a strong function of chemical composition, with sulfate and nitrate species being particularly efficient scatterers (Sloane et al. 1991). The same result was confirmed in a previous study in Taiwan, based in Kaohsiung (Yuan et al. 2006).

On the other hand, the magnitudes of  $B_{PM_{2.5}}$  and  $B_{PM_{10-2.5}}$  derived in this study are significantly less - by an order of magnitude - than that from a previous study in Kaohsiung (Yuan et al. 2006), which was  $\sim -200$  m per  $\mu\text{g m}^{-3}$ . We suspect that the reason for this large difference is not just because of variability in chemical composition, but because the visibility observations in this study were reported with ma-

ximum values imposed at 9999 m, while the observations in Yuan et al. (2006) had no such ceiling. So any reduction of visibility above 9999 m (e.g., from 15 to 12 km) would be neglected.

#### 4.2.4 $B_{(100 - RH)}$

The values of  $B_{(100 - RH)}$  are positive, suggesting that as the atmosphere becomes drier (lower  $RH$ ), visibility increases. The lowering of visibility with higher  $RH$  is likely because fog formation (also see section 4.4) requires saturation conditions for water vapor to condense to the liquid phase, a fundamental principle of atmospheric thermodynamics (Bohren and Albrecht 1998). Furthermore, aerosol production is known to be accelerated at high humidity conditions (Tang et al. 1981). For instance, high humidity enables  $SO_2$  to more readily dissolve in droplets and oxidize to produce sulfate aerosols (Tsai and Cheng 1999). Furthermore, inorganic salts such as ammonium sulfate and nitrate are hygroscopic, undergoing conversion from solid particles to solution droplets at higher levels of  $RH$  and reaching a size that is especially effective in scattering light and reducing visibility (Malm 1999).

Previous localized studies in Taichung (Tsai and Cheng 1999) and Kaohsiung (Yuan et al. 2006) have both reported visibility reduction when  $RH$  increases. The reported coefficients of -190 m/% (Tsai and Cheng 1999) and -150 m/% (Yuan et al. 2006) are very much comparable to the values for west coast stations in this study. This suggests that unlike  $B_{PM}$ ,  $B_{(100 - RH)}$  was not significantly affected by the 9999 m ceiling. We believe this implies that under high visibility ( $> 9999$  m) conditions,  $RH$  affects visibility less than PM, which scatters/absorbs radiation regardless.

#### 4.2.5 $B_{rain}$

The values of  $B_{rain}$  indicate that the presence of rainfall decreases visibility on average by hundreds of meters, up to 2000 m (at GM and YU). The rainfall-induced fall in visibility appears to be large at all of the east coast stations and at LT, with magnitudes all greater than 1500 m. The impact of rainfall can also be discerned in Figs. 2 ~ 3, with the rainfall-induced visibility reduction being much more pronounced at YU ( $B_{rain} = -1902$ ) than at AY ( $B_{rain} = -1100$ ) during the months of November ~ April.

The factors controlling  $B_{rain}$  have been known at least since the early 1920s. Preston (1920), using a simple physical model, showed that the amount of visibility attenuation during rainfall is “simply proportional to the number of (drops falling onto a unit area of the earth’s surface) per second.” This work is updated in Achour (2002). These authors showed that the degree to which rainfall decreases visibility is magnified when more raindrops are found in

the atmosphere scattering sunlight. Given the same amount of precipitation accumulating over an hour, the visibility attenuation would be less if the raindrops are larger (less number of droplets in the atmosphere). Since observations of raindrop size distributions are not available, we can only speculate on this point and not attempt to verify the correspondence between  $B_{rain}$  and rain drop sizes.

#### 4.2.6 $B_{wind}$

Other than AY, the values of  $B_{wind}$  were positive at the 5 other stations where windspeed was deemed statistically significant. This suggests that visibility increases along with windspeed, a result also previously observed in southern (Yuan et al. 2006) and central (Tsai and Cheng 1999) Taiwan.

Higher wind induces mechanical production of turbulence, leading to dispersion as well as a higher boundary layer height (Stull 1988). Conversely, low wind conditions lead to buildup of moisture and PM in a shallower boundary layer, resulting in low visibility events. Indeed, one of the worst pollution episodes in southern China was associated with stagnant, low wind conditions (Fu et al. 2008).

$B_{wind}$  was negative at AY, making it an exception to the above. We suspect this is due to fog being advected to AY, which was the explanation given for the negative correspondence between visibility and wind at some sites in Finland (Hyvärinen et al. 2007).

### 4.3 Uncertainty Analysis

In this section we examine errors in visibility that arise from different sources, particularly if the regression model is used in a forecast setting. The total error  $\sigma_{tot}$  (i.e., the standard deviation of the total deviation between observed versus regression-simulated visibility) is partitioned into the sum of error in measuring ( $\sigma_{msm}$ ) and modeling ( $\sigma_{sim}$ ) visibility.  $\sigma_{sim}$ , in turn, incorporates inadequacies in the fitted statistical model ( $\sigma_{fit}$ ) as well as uncertainties in the values of different explanatory variables ( $\sigma_{var}$ ). Hence the error terms are as follows:

$$\sigma_{tot}^2 = \sigma_{msm}^2 + \sigma_{sim}^2 = \sigma_{msm}^2 + \sigma_{fit}^2 + \sum_{var=1}^{N_{var}} \sigma_{var}^2 \quad (4)$$

where  $N_{var}$  is the number of explanatory variables, and  $var$  refers to  $RH$ ,  $PM$ , ... etc. Equation (4) assumes that the different error terms are independent of one another (Taylor 1997).  $\sigma_{fit}$  arises from the statistical model's limitations. For instance, for the multiple linear regression adopted here [Eq. (2)] such restrictions would include the deviation from linearity and the missing impact of omitted variables.  $\sigma_{var}$  can be distinguished between the case of a) visibility "hind-

casts", in which measurements of explanatory variables are used; b) visibility "forecasts", in which predictions of the explanatory variables are used ( $\sigma_{var,fcst}$ ). The total forecast error when using the regression ( $\sigma_{sim,fcst}$ ) is, then:

$$\sigma_{sim,fcst}^2 = \sigma_{fit}^2 + \sum_{var=1}^{N_{var}} \sigma_{var,fcst}^2 \quad (5)$$

To solve for  $\sigma_{sim,fcst}$ , the value of  $\sigma_{fit}$  needs to be first established by rearranging Eq. (4):  $\sigma_{fit}^2 = \sigma_{tot}^2 - \sigma_{msm}^2 - \sum_{var=1}^{N_{var}} \sigma_{var}^2$ .

$\sigma_{tot}$  is set to the values of the residual standard error shown in Table 3.  $\sigma_{msm}$  is estimated to be approximately 10% of the measured visibility, as a rough estimate of the measurement error following personal communication with personnel in the Taiwanese Air Force familiar with the measurements. One non-negligible contribution to  $\sigma_{var}$  in the case of hindcasts is the error due to the fact that temperature and dewpoint are reported without decimal points. We estimated this round-off error cause uncertainties of roughly 2% in  $RH$ . This 2% is then multiplied with  $B_{(100-RH)}$  in Table 3 to calculate  $\sigma_{var}$ .

We derived errors in forecasted predictor variables from the published literature, and this error is multiplied by values of  $B_{var}$  as specified in Table 3 to calculate  $\sigma_{var,fcst}$ . We calculate errors of forecasts 24 hours into the future, since this is a typical timeframe with which flight decisions are made. The 24-hour forecast error of  $RH$  and windspeed are taken to be 17.5% and  $3 \text{ m s}^{-1}$ , respectively, following the results of an ensemble mesoscale modeling study over the Taiwan area by Chien et al. (2006). The forecast error of hourly  $PM_{2.5}$  is  $11 \text{ g m}^{-3}$ , as seen in a US study based on a state-of-the-art air quality modeling system (Yu et al. 2008). We are not aware of published values for forecast errors in coarse fraction PM ( $PM_{10-2.5}$ ). The forecast error of daily averaged  $PM_{10}$  in a European study (Honore et al. 2008), however, was identical to that reported for daily averaged  $PM_{2.5}$  (Yu et al. 2008). This suggests that the hourly forecast error for coarse fraction PM is likely similar to that for  $PM_{2.5}$ . In the absence of further information, we take the error of  $PM_{10-2.5}$  to be equal to that of  $PM_{2.5}$ , at  $11 \text{ g m}^{-3}$ .

Results of the uncertainty analysis are shown in Fig. 5.  $\sigma_{RH,fcst}$  is by far the largest source of error in regression-predicted visibility on the west coast of Taiwan, with values  $> 1500 \text{ m}$  and even up to  $3000 \text{ m}$ . The contribution from uncertainties in predicting  $PM_{2.5}$  is next in importance, followed by  $PM_{10-2.5}$  and windspeed, which are comparable in magnitude. The much lower sensitivity to  $PM_{10-2.5}$  (Table 3) clearly caused the lower contribution from these coarser PM.

Due to the lack of suitable estimates of uncertainties in predicting rainfall events, we currently do not have a quantitative estimate of  $\sigma_{rain,fcst}$ . However, it is clear that the values

of  $B_{rain}$  in Table 3 provide an upper bound for this uncertainty.

The smallest forecast errors are found at the three east coast stations of YU, CS, and ZN. This follows from the fact that visibility is generally high on the east coast (Fig. 1), with smaller variability and thus smaller residual error  $\sigma_{fit}$ . Fur-

thermore, visibility covaries to a lesser extent with  $RH$  on the east coast (Table 3), so even with the large error in forecasting  $RH$ , a smaller error in visibility would result.

#### 4.4 Logistic Regression: What Controls Probability of Extremely Low Visibility?

Results of the logistic regression are shown in Table 4. The east coast stations and off-shore islands (YU, CS, ZN, LT, QC) were not included in the analysis, since frequencies of occurrence of visibility  $\leq 1600$  m were too low (Table 1). Also, from analyses of AIC (section 2.3.2) we see that neither rain nor wind speed appears to be related significantly to occurrence of low visibility.

We can see that values of  $exp(m_i)$  vary greatly, ranging from below 1.0 to even over 3.0. This means that the “baseline probability” of low visibility varies significantly from month to month. However, at almost every site the largest values occur during the months of January and February, which are also the months when the lowest average visibilities are observed around Taiwan (Fig. 1), simply indicating that fog events contribute significantly to the low visibilities in January and February.

Values of  $exp[b_{(100 - RH)}]$  at the different stations are less than 1.0 and vary within a relatively tight range of 0.7 ~ 0.8. Thus as the atmosphere becomes drier, i.e., increased  $(100 - RH)$ , the odds of a low visibility event decreases. This can be seen in Fig. 6 for AY. The estimated dependence of  $p$  with respect to  $(100 - RH)$  is also shown as the solid line. High  $RH$  is certainly a requisite condition for fog formation; for instance, Croft et al. (1997) found that the best statistical predictor for dense fog in southern US was the surface-level  $RH$ . Likewise, Tsai and Cheng (1999) pointed out that the high  $RH$  in central Taiwan caused the presence of early

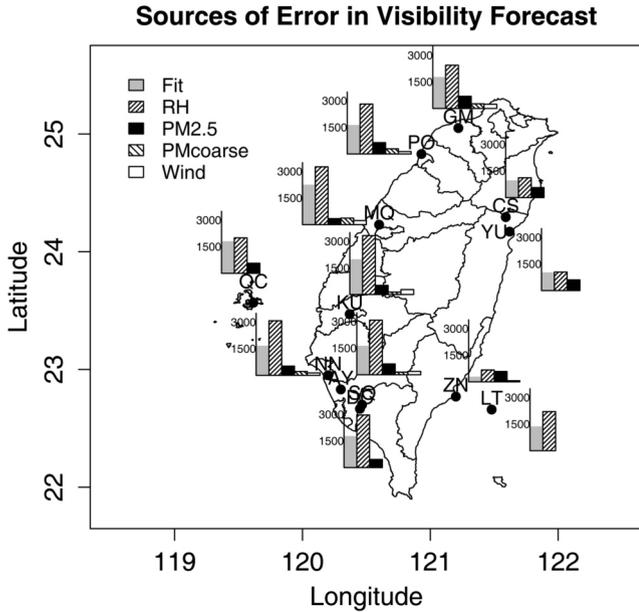


Fig. 5. Sources of uncertainties in forecasting visibility 24 hours into the future, based on the multiple linear regression at various stations around the Taiwan region. The various bars refer to uncertainties arising from inadequacies in the fitted statistical model (“Fit”) and errors in the forecasted values of explanatory variables (“ $RH$ ”, “ $PM_{2.5}$ ”, “ $PM_{coarse}$ ”, and “Wind”). Variables deemed statistically unrelated to visibility at a station are not included within the uncertainty analysis. Values are in units of meters and refer to the standard deviation (rather than variance)  $\sigma_{var,forecast}$  for explanatory variable  $var$ .

Table 4. Results from the logistic regression [Eq. (3)]:  $logit(p) = log_e[p/(1 - p)] = m_0 + b_{(100 - RH)} * [100 - RH] + b_{PM_{2.5}} * [PM_{2.5}] + b_{PM_{10-2.5}} * [PM_{10-2.5}]$ , where  $p$  is the probability of occurrence of extremely low visibility events ( $\leq 1600$  m). The regression coefficients are shown as  $exp(m_i)$  or  $exp(b_{var})$ , which is the change in odds [ $\equiv p/(1 - p)$ ] for a unit increase in  $RH$ ,  $PM_{2.5}$ , or  $PM_{10-2.5}$  (see section 2.3.2). Thus values  $> 1.0$  indicated increased  $p$  when the variable under consideration is increased, and vice versa. The east coast stations and off-shore islands (YU, CS, ZN, LT, QC) were not included in the analysis, since frequencies of occurrence of visibility  $\leq 1600$  m were too low (Table 1).

Station	$exp(m_{01})$	$exp(m_{02})$	$exp(m_{03})$	$exp(m_{04})$	$exp(m_{11})$	$exp(m_{12})$	$exp(b_{100 - RH})$	$exp(b_{PM_{2.5}})$	$exp(b_{PM_{10-2.5}})$
PO	1.900	2.291	0.957	0.292	0.595	0.360	0.773	-	-
GM	1.333	0.463	0.493	0.077	0.082	0.018	0.711	-	1.070
MQ	2.196	3.136	1.125	0.679	0.500	0.508	0.832	-	-
KU	0.137	0.175	0.039	0.025	0.046	0.028	0.692	1.035	1.015
NN	0.333	0.207	0.053	0.162	0.135	0.110	0.693	-	1.063
AY	0.103	0.057	0.028	0.032	0.051	0.105	0.716	1.027	1.034
SQ	1.564	1.464	1.250	0.220	0.768	1.199	0.764	-	-
DC	2.019	1.760	1.060	0.253	1.068	1.173	0.752	-	-

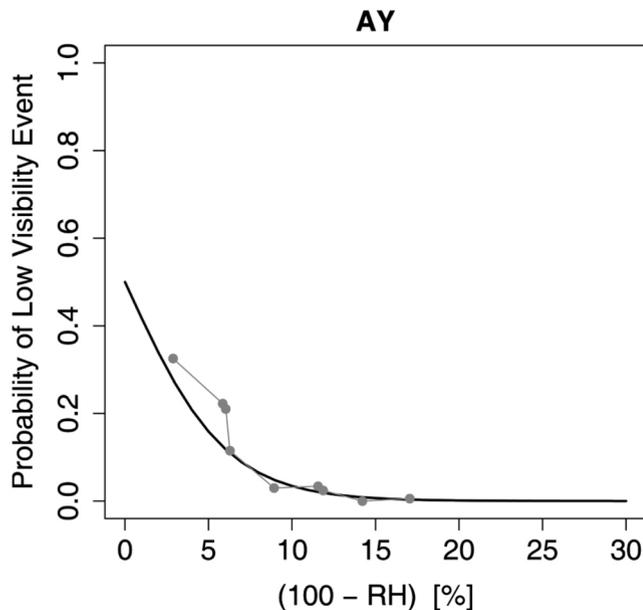


Fig. 6. Variation of  $p$ , the probability of occurrence of low visibility events ( $< 1600$  m) with the departure from saturation ( $100 - RH$ ) at AY. The curve is the fitted dependence from the logistic regression. The gray points indicate the observed probabilities.

morning fog in the late winter and early spring.

$PM_{2.5}$  and  $PM_{10-2.5}$  are statistically significant at some locations and increase  $p$ . PM is known to affect the micro-physical properties of fog. A larger number of particles lead to smaller, but higher number of droplets (Twomey 1977). Because the opacity of fog increases with droplet concentration, the visibility would also decrease with PM (Croft et al. 1997), thus increasing  $p$ .

## 5. CONCLUSIONS

Through statistical analyses, this study has revealed the following:

- Visibility exhibits systematic variation from month to month, with lowest values during the winter months;
- On the West Coast of Taiwan,  $PM_{2.5}$  and  $RH$  are related most to visibility, and to a lesser extent,  $PM_{10-2.5}$  and windspeed;
- On the East Coast, rainfall plays a more significant role in reducing visibility;
- Wind direction was found to be a poor predictor of visibility;
- $PM_{2.5}$  causes a larger drop in visibility than  $PM_{10-2.5}$ .

The occurrence of low visibility ( $< 1600$  m) events such as fog is shown to:

- be rare on the East Coast and off-shore islands;
- increase with  $RH$  at all stations;
- increase with PM at some locations;
- be unrelated to other variables such as windspeed, wind

direction, and rainfall.

A quantitative analysis of uncertainties shows that large errors still exist if the regression model is used to forecast visibility 24 hours into the future. The largest error source is the forecast error in  $RH$  ( $\sigma_{RH,fcst}$ ), which can reach values as large as 3000 m. Next is the error in the statistical model ( $\sigma_{fit}$ ), which is approximately 1500 m. Given such large errors, the resulting predictions in visibility would have uncertainties (assuming each term being uncorrelated) of approximately  $\sqrt{1500^2 + 3000^2} = 3350$  m.

These large uncertainties directly point to future steps to improve visibility forecasts.  $\sigma_{fit}$  can be easily reduced by adopting a more sophisticated statistical model. Separate coefficients for each month can be fitted. Also, instead of relying entirely on 24-hour forecasts of explanatory variables, the statistical model can make use of observed values of these variables within the 24-hour window in a time-lagged correlation regression. Additional predictor variables can be incorporated that would likely have relevance for visibility, such as synoptic patterns (e.g., positions of pressure systems), turbulence strength, or cloud cover.

Reducing  $\sigma_{RH,fcst}$  is much more difficult. The significant error in  $RH$  of 17.5% reported by Chien et al. (2006) derives from an advanced, ensemble simulation of mesoscale models over Taiwan. Additional improvements may require significant investment in assimilating into mesoscale models different humidity observations - e.g., from radiosondes and satellites such as AIRS (Aumann et al. 2003). It is clear that without accurate predictions of  $RH$ , precise visibility forecasts are not possible. Determining the amount of liquid water suspended in the atmosphere attenuating light - the reason behind severe visibility reductions under foggy conditions - requires knowledge of  $RH$ .

The uncertainty due to  $PM_{2.5}$  ( $\sigma_{PM_{2.5},fcst}$ ) is on the order of several hundred meters, much less than that due to  $RH$  but larger than the errors due to the other explanatory variables. The 24-hour forecast error of  $11 \text{ g m}^{-3}$  for  $PM_{2.5}$  (Yu et al. 2008), based on state-of-the-art air quality modeling system in the US, is likely larger for Taiwan, where the population density ( $642 \text{ people km}^{-2}$ ) is vastly higher than in the US ( $31 \text{ people km}^{-2}$ ) (<http://www.wikipedia.org>; accessed August 2008) and where strong emissions associated with population centers are expected to translate into higher magnitudes and gradients of PM concentrations (Fig. 4), increasing the difficulty of PM prediction. Even in the US, where significant effort has been invested into compiling accurate inventories of PM sources, considerable uncertainties still exist (Simon et al. 2008).

PM forecasts can be improved by improving the PM emission inventories by combining the PM concentration observations with a model of atmospheric transport. Known as "inverse analysis", this method has been widely applied to other atmospheric species such as  $CO_2$  (Lin et al. 2004),  $CO$

(Miller et al. 2008), or CFCs (Hurst et al. 2006).

Ultimately, visibility should be a forecast variable directly predicted by numerical prediction models, bypassing the need to use statistical models. Towards this end the models need to incorporate more mechanistic descriptions of processes such as microphysics and radiative transfer, i.e., the transformation of hydrometeors plus the scattering/absorption of radiation by aerosols and water droplets. However, this is clearly beyond the capability of most models, with the large uncertainty in predicting *RH* being symptomatic of such deficiencies. Moreover, most numerical models inadequately resolve details of relevance for visibility, which are at the sub-gridscale (Hansen 2007). In the absence of direct numerical prediction capabilities for visibility, the statistical approach exemplified by the one described in this paper will continue to be of value, and the considerable uncertainties identified above need to be overcome.

**Acknowledgements** We thank the Taiwan Environmental Protection Administration in carrying out the PM measurements and graciously making available the dataset. We appreciate the dedicated and sleep-deprived personnel of the Weather Wing of the Taiwan Air Force in making possible the 24-hour, year-round measurements of visibility and meteorological variables.

## REFERENCES

- Achour, M., 2002: Simulating atmospheric free-space optical propagation: Part I, rainfall attenuation. *Proc. SPIE*, **4635**, 192-201, doi: 10.1117/12.464100. [[Link](#)]
- Akaike, H., 1974: A new look at statistical model identification. *IEEE Trans. Autom. Control*, **AU-19**, 716-723, doi: 10.1109/TAC.1974.1100705. [[Link](#)]
- Arpino, R., 1994: Automated weather observation systems find increasing acceptance at major airports. *ICAO J.*, **49**, 15-16.
- Aumann, H. H., M. T. Chahine, C. Gautier, M. D. Goldberg, E. Kalnay, L. M. McMillin, H. Revercomb, P. W. Rosenkrag, W. L. Smith, D. H. Staelin, L. L. Straw, and J. Susskind, 2003: AIRS/AMSU/HSB on the Aqua Mission: Design, science objectives, data products, and processing systems. *IEEE Trans. Geosci. Remote Sensing*, **41**, 253-264, doi: 10.1109/TGRS.2002.808356. [[Link](#)]
- Bohren, C. F. and B. A. Albrecht, 1998: Atmospheric Thermodynamics. Oxford University Press, 416 pp.
- Chang, S.-C. and C.-T. Lee, 2007a: Evaluation of the trend of air quality in Taipei, Taiwan from 1994 to 2003. *Environ. Monit. Assess.*, **127**, 87-96, doi: 10.1007/s10661-006-9262-1. [[Link](#)]
- Chang, S.-C. and C.-T. Lee, 2007b: Assessment of PM<sub>10</sub> enhancement by yellow sand on the air quality of Taipei, Taiwan in 2001. *Environ. Monit. Assess.*, **132**, 297-309, doi: 10.1007/s10661-006-9534-9. [[Link](#)]
- Chen, M.-L., I.-F. Mao, and I.-K. Lin, 1999: The PM<sub>2.5</sub> and PM<sub>10</sub> particles in urban areas of Taiwan. *Sci. Total Environ.*, **226**, 227-235, doi: 10.1016/S0048-9697(98)00407-0. [[Link](#)]
- Chien, F.-C., Y.-C. Liu, and B. J.-D. Jou, 2006: MM5 ensemble mean forecasts in the Taiwan area for the 2003 Mei-Yu season. *Weather Forecast.*, **21**, 1006-1023, doi: 10.1175/WAF960.1. [[Link](#)]
- Chio, C.-P., M.-T. Cheng, and C.-F. Wang, 2004: Source apportionment to PM<sub>10</sub> in different air quality conditions for Taichung urban and coastal areas, Taiwan. *Atmos. Environ.*, **38**, 6893-6905, doi: 10.1016/j.atmosenv.2004.08.041. [[Link](#)]
- Croft, P. J., R. L. Pfost, J. M. Medlin, and G. A. Johnson, 1997: Fog forecasting for the Southern Region: A conceptual model approach. *Weather Forecast.*, **12**, 545-556, doi: 10.1175/1520-0434(1997)012<0545:FFFTSR>2.0.CO;2. [[Link](#)]
- Dalgaard, P., 2002: Introductory Statistics with R. Springer, 267 pp.
- Fu, Q., G. Zhuang, J. Wang, C. Xu, K. Huang, J. Li, B. Hou, T. Lu, and D. G. Streets, 2008: Mechanism of formation of the heaviest pollution episode ever recorded in the Yangtze River Delta, China. *Atmos. Environ.*, **42**, 2023-2036, doi: 10.1016/j.atmosenv.2007.12.002. [[Link](#)]
- George, J. J., 1960: Weather Forecasting for Aeronautics. Academic Press, 637 pp.
- Hansen, B., 2007: A fuzzy logic-based analog forecasting system for ceiling and visibility. *Weather Forecast.*, **22**, 1319-1330, doi: 10.1175/2007WAF2006017.1. [[Link](#)]
- Honoré, C., L. Rouil, R. Vautard, M. Beekmann, B. Bessagnet, A. Dufour, C. Elichegaray, J.-M. Flaud, L. Malherbe, F. Meleux, L. Menut, D. Martin, A. Peuch, V.-H. Peuch, and N. Poisson, 2008: Predictability of European air quality: Assessment of 3 years of operational forecasts and analyses by the PREV'AIR system. *J. Geophys. Res.*, **113**, D04301, doi: 10.1029/2007JD008761. [[Link](#)]
- Hornig, C.-L. and M.-T. Cheng, 2008: Distribution of PM<sub>2.5</sub>, acidic and basic gases near highway in central Taiwan. *Atmos. Res.*, **88**, 1-12, doi: 10.1016/j.atmosres.2007.09.002. [[Link](#)]
- Hurst, D., J. C. Lin, P. A. Romashkin, B. C. Daube, C. Gerbig, D. M. Matross, S. C. Wofsy, B. D. Hall, and J. W. Elkins, 2006: Continuing global significance of emissions of Montreal Protocol-restricted halocarbons in the United States and Canada. *J. Geophys. Res.*, **111**, D15302, doi: 10.1029/2005JD006785. [[Link](#)]
- Hyvärinen, O., J. Julkunen, and V. Nietosvaara, 2007: Climatological tools for low visibility forecasting. *Pure Appl. Geophys.*, **164**, 1383-1396, doi: 10.1007/s00024-007-0224-5. [[Link](#)]
- ICAO (International Civil Aviation Organization), 2007: Annex 3: Meteorological Service for International Air Navigation, 16<sup>th</sup> Edition. Montreal, Quebec, Canada, 202 pp.
- Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, C. A. Grainger, B. B. Stephens, P. S. Bakwin, and D. Y. Hollinger, 2004: Measuring fluxes of trace gases at regional scales by Lagrangian observations: Application to the CO<sub>2</sub>

- Budget and Rectification Airborne (COBRA) study. *J. Geophys. Res.*, **109**, D15304, doi: 10.1029/2004JD004754. [[Link](#)]
- Malm, W. C., 1999: Introduction to Visibility. Report from Air Resources Division, National Park Service, Fort Collins, Colorado, USA, 70 pp.
- Martin, M. D. and P. W. Suckling, 1987: Winter fog and air transportation in Sacramento, California. *Climatol. Bull.*, **21**, 16-22.
- Miller, S., D. M. Matross, A. E. Andrews, D. B. Millet, M. Longo, E. W. Gottlieb, A. I. Hirsch, C. Gerbig, J. C. Lin, B. C. Daube, R. C. Hudman, P. L. S. Dias, V. Y. Chow, and S. C. Wofsy, 2008: Sources of carbon monoxide and formaldehyde in North America determined from high-resolution atmospheric data. *Atmos. Chem. Phys.*, **8**, 7673-7696.
- Müller, M. D., C. Schmutz, and E. Parlow, 2007: A one-dimensional ensemble forecast and assimilation system for fog prediction. *Pure Appl. Geophys.*, **164**, 1241-1264, doi: 10.1007/s00024-007-0217-4. [[Link](#)]
- Preston, F. W., 1920: Visibility of the landscape during rain. *Nature*, **106**, 343-344, doi: 10.1038/106343d0. [[Link](#)]
- R Development Core Team, 2005: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>. [[Link](#)]
- Simon, H., D. T. Allen, and A. E. Wittig, 2008: Fine particulate matter emissions inventories: Comparisons of emissions estimates with observations from recent field programs. *J. Air Waste Manage. Assoc.*, **58**, 320-343.
- Sloane, C. S., J. G. Watson, J. C. Chow, L. C. Pritchett, and L. W. Richards, 1991: Size-segregated fine particle measurements by chemical species and their impact on visibility impairment in Denver. *Atmos. Environ.*, **25A**, 1013-1024.
- Stull, R. B., 1988: An Introduction to Boundary Layer Meteorology. Kluwer, 680 pp.
- Tang, I. N., W. T. Wong, and H. R. Munkelwitz, 1981: The relative importance of atmospheric sulfate and nitrates in visibility reduction. *Atmos. Environ.*, **15**, 2463-2471, doi: 10.1016/0004-6981(81)90062-7. [[Link](#)]
- Taylor, J. R., 1997: An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements (Second Edition). University Science Books, 327 pp.
- Teixeira, J. and P. M. A. Miranda, 2001: Fog prediction at Lisbon Airport using a one-dimensional boundary layer model. *Meteorol. Appl.*, **8**, 497-505, doi: 10.1017/S135048270100411X. [[Link](#)]
- Tsai, Y. I. and M. T. Cheng, 1999: Visibility and aerosol chemical compositions near the coastal area in Central Taiwan. *Sci. Total Environ.*, **231**, 37-51, doi: 10.1016/S0048-9697(99)00093-5. [[Link](#)]
- Twomey, S., 1977: The influence of pollution on the shortwave albedo of clouds. *J. Atmos. Sci.*, **34**, 1149-1152.
- Venables, W. N. and B. D. Ripley, 2002: Modern Applied Statistics with S. Springer, 495 pp.
- Yu, S., R. Mathur, K. Schere, D. Kang, J. Pleim, J. Young, D. Tong, G. Pouliot, S. A. McKeen, and S. T. Rao, 2008: Evaluation of real-time PM<sub>2.5</sub> forecasts and process analysis for PM<sub>2.5</sub> formation over the eastern United States using the Eta-CMAQ forecast model during the 2004 ICARTT study. *J. Geophys. Res.*, **113**, D06204, doi: 10.1029/2007JD009226. [[Link](#)]
- Yuan, C.-S., C.-G. Lee, S.-H. Liu, J.-C. Chang, C. Yuan, and H.-Y. Yang, 2006: Correlation of atmospheric visibility with chemical composition of Kaohsiung aerosols. *Atmos. Res.*, **82**, 663-679, doi: 10.1016/j.atmosres.2006.02.027. [[Link](#)]