Short-term Climate Prediction of Mei-Yu Rainfall for Taiwan: Results of Ensemble Forecasting for the 1998 Season

Final Report to the Central Weather Bureau
(CWB 87-3M05)

Pao-Shin Chu
Climate Systems Enterprise
Honolulu, Hawaii
U. S. A.

June 9, 1998

1. Introduction

Taiwan is experiencing a population increase that places an increasing demand on water for drinking, recreation, food production, and other needs. If the current population trend continues, the island may be faced with a shortage of fresh water in the near future. Further compounding this problem is the large interannual variability of rainfall. The prolonged and devastating drought that occurred successively from summer 1995 to the first few months of 1996 brought renewed attention to the far reaching effects of water scarcity on the socio-economic welfare and environmental quality of the country.

Drought and heavy deluges have been recurrent and troublesome problems for Taiwan. If the nation is to come to grip with these incessant problems, it must also improve water management. One of the important components in water management is to develop better predictive methods so that climate forecasts can be used for strategic planning programs well before the situation has reached a crisis.

The Mei-Yu season in Taiwan occurs from May to June during which the Mei-Yu front becomes quasi-stationary. This front is a rain-bearing synoptic/mesoscale system which constantly moves from southern China toward Taiwan. Southwesterly moist flows often accompany the frontal system. Because of the blocking effect by the Central Mountain Range, which is basically oriented from north to south with an average elevation of 2 km, rainfall during the Mei-Yu season is higher on the west side (windward) of the island than on the east side (leeward). Climatologically, from Hsinchu to Kaohsiung on the western side of the island, which is densely populated, Mei-Yu rainfall constitutes approximately one third of the annual rainfall totals. This is also the region where the interannual variability of rainfall is large.

In reponse to an increasing demand for water, the uncertainty in Mei-Yu rainfall from year-to-year, and the importance of Mei-Yu rainfall in annual rainfall totals, we have successfully developed a Canonical Correlation Analysis (CCA) model for Mei-Yu rainfall in Taiwan (Chu and Wang, 1997; Chu, 1998). The Pacific sea surface temperature (SST) and 500 hPa height over East Asia/western Pacific in November/December are used as predictor variables and the May/June rainfall in the following year from eight major stations in Taiwan are employed as predictand variables. The forecast lead time is four months. Forecast skills are assessed for both an independent period (1986-95) and for the entire period (1956-95) via a cross-validation technique.

In our previous studies, the predictor variables are fixed in time (i.e., always November/December) and the forecasts are expressed deterministically (e.g., 750 mm rainfall in May/June 1997). Because prediction involves future variations and such variations cannot be completely described by the model (even with a perfect dynamic model) and by using the past records, forecast values would be better expressed in stochastic forms such as by a probability distribution (Lorenz, 1993). The ensemble forecasting procedure used in NWP forecasts (Tracton and Kalnay, 1993) draws a finite sample from the probability distribution describing the uncertainty of the initial state of the atmosphere. Using this idea, ensemble forecasting can also be implemented on statistical-based models by changing the initial state of predictor variables repeatedly and making a collection of forecasts. Among other things, ensemble forecasting yields information about the magnitude of the uncertainty in a forecast. Using ensemblle forecasting with a statistical model is a new approach. This is a computer-intensive and laborious technique, however, because in ensemble forecasting many forecasts are generated for a single time period. In this study, we will present a case of the 1998 Mei-Yu rainfall prediction using CCA with ensemble forecasting procedures. To the best of my knowledge, this approach is quite novel and has not been done before.

2. Data and data processing

The monthly mean SST data at 2 degrees resolution over a large portion of the Pacific (50°N-40°S, 120°E-90°W) from January 1955 to February 1998 were used. This dataset was then converted to a 10° latitude/longitude resolution, yielding 125 grid points over the Pacific. The monthly mean 500 hPa height records over Asia and the western Pacific were obtained from the Central Weather Bureau through the courtesy of Mr. Guay-Hong Chen. The domain covers from 20°N to 60°N and 60°E to 160°E, and the data period is identical to the SST. Long-term monthly rainfall totals for 16 stations were also provided by Mr. Chen. The stations include Keelung, Taipei, Hsinchu, Taichung, Tainan, Kaohsiung, Hualien, Taitung, Tanshui, Jihyuehtan, Alishan, Penghu, Hengchun, Ilan, Chengkung, and Tawu. The first eight stations were used in our previous studies (Chu and Wang, 1997; Chu, 1998). We extracted only May/June rainfall from these 16 stations for the period from 1956 to 1997.

In order to perform statistical ensemble forecasting, predictor variables at different time periods are needed. Accordingly, the January/February to November/December data from the two predictor variables (SST and 500 hPa height) and May/June rainfall are first summed to produce a bi-monthly total. The bi-monthly data are then normalized by taking the difference between the individual bi-monthly value and the long-term average of the bi-monthly value, and dividing this difference by the standard deviation of the bi-monthly series. This normalized method is applied to each grid point of the SST and height fields, and to each rainfall series.

To reduce the large dimensionality of the SST and height fields, an empirical orthogonal function analysis is first performed to each predictor field. The leading ten eigenmodes of the SST account for approximately 75% of the total variance. The leading six eigenmodes explain 60 to 80% of the total variance for the height field. These eigenmodes were retained for further analyses.

3. Statistical modeling in an ensemble forecasting framework

As described in the introduction, the ensemble forecasting is based on repeatedly running the NWP or GCM model, once for each member of the initial atmospheric state. In statistical modelings, ensemble forecasts can be produced by repeatedly running a statistical predictive model while constantly changing the initial state of the atmosphere/ocean state. In this sense, the initial state of the predictor variables, which represents atmosphere/ocean behavior, is set at different time periods. They vary successively from January/February of the preceding year to January/February of the same year as the predictand variable (May/June rainfall). This yields seven bi-monthly pairs (i.e., January/February, March/April, May/June, July/August, September/October, November/December, and January/February) for the predictor variables. Accordingly, the forecast lead times vary from two to 14 months ahead. Note that the predictor variables in March/April just preceding the target season were not used because by the time these data became available and were processed in Honolulu the Mei-Yu season had already begun.

In this experiment, time series of the leading modes of SST and a combination of SST/500 hPa height were adopted as the predictor variables. Because there are two sets of predictor variables and seven bi-monthly pairs in the predictor variables, we produced 13 ensemble members (ignoring the January/February

SST in the same year as the predictand). With 13 forecasts for each station, a probability distribution of the 1998 May/June rainfall forecasts made at different initial times can be constructed. For example, a CCA model is used to identify relationships between two multivariate datasets consisting of the leading eigenmodes of January/February SST and height for the years 1955-96, and the corresponding May/June rainfall indices for the period of 1956-97 (i.e., lagged by one year). The observed SST and height in January/February 1997 will then serve as inputs to the CCA model to forecast the May/June 1998 rainfall (i.e., 14 months ahead forecast).

CCA identifies new variables that maximize the relationship between two multivariate data fields. Thus it can be used to diagnose the coupled variability of the two fields (e.g., Wallace et al. 1992, Diaz et al., 1998). The model will not be described here because details are in Chu and He (1994) and Yu, Chu, and Schroeder (1997).

4. 1998 Mei-Yu rainfall forecasts

a. SST-based forecasts

Before showing the results of probabilistic ensemble forecasting, it would be interesting to demonstrate the deterministic forecasts because those were performed in previous studies. Figure 1 shows the 1998 rainfall forecasts for 16 individual stations and the average of all 16 stations using the Pacific SST of November/December 1997 as the only independent input to the CCA model. The lead time is four months. Forecasts call for more than normal rainfall for most stations, particularly for two southern stations, Tawu and Kaohsiung. Out of 16 stations, only Hualien and Keelung exhibit slightly less than normal rainfall. As a result, the average forecast for 16 stations is wetter than normal.

As lead times increase, however, more stations show drier than normal forecasts. When lead times increase to 12 months ahead (i.e., using March/April of the preceding year), a drier than normal forecast exists for all 16 stations (Fig. 2). This example illustrates that forecasts are sensitive to the different initial time of the state of the boundary condition and are thus consistent with the NWP or GCM experiments. The different forecasts for 1998 May/June rainfall generated by using the Pacific SST in March/April and the SST in the subsequent months are of interest. This change in the predictor time periods which lead to quite different rainfall forecasts coincides with the onset of the very strong 1997-98 El Niño event which took place around May 1997. Anomalous SST was more than 1°C in a large area extending from the west coast of South America to the equatorial central Pacific. Prior to May 1997, SST anomalies in the tropical Pacific were only slightly above normal in limited areas.

b. SST- and height-based ensemble forecasts

Figure 3 shows the probabilistic forecasts of the 1998 Mei-Yu rainfall based on 13 ensemble members using both the antecedent SST and 500 hPa height as predictor variables. The forecasts are expressed in probabilities for three categories. If the station forecast is below the 30th percentile of the long-term observed May/June rainfall, it is classified as dry (or drought). If it is above the 70th percentile of the observed rainfall totals, it is classfied as wet. A forecast that falls between 30th and 70th percentiles is called normal.

For three northern stations (Keelung, Taipei, and Hsinchu), the chance for near-normal rainfall during this Mei-Yu season is relatively high, ranging from near 40% in Hsinchu to almost 80% in Taipei. Taichung has an almost equal chance (~40%) of being in the wet or near-normal rainfall category. Moving southward, the chance of being above the 70th percentile is rather high for Tainan (~70%) and Kaohsiung (>50%). Taitung, which is located in the southeastern corner of Taiwan, also has a good chance of being above the 70th percentile (>60%).

Because southwesterly flows bring moisture to the island during the Mei-Yu season, the above result suggests that heavy convective clouds embedded in southwesterly flows may prevail this season but their influence probably only extends to the southern and southeastern portions of the island. Generally speaking, the average forecast for 16 stations indicates a higher chance of being in the wet category (~60%) and a lower chance of being in the dry category (~15%). Probabilistic forecasts for some stations, such as Penghu and Alishan, are similar for all three categories; thus no single categorical forecast stands out than others.

5. Summary

Because of the dynamical chaos inherent in complex climate systems, future behavior of the atmosphere cannot be predicted with certainty. One way to provide a more realistic forecast is to let the determinstic dynamic equations operate on the probability distribution, which describes the uncertainty in observation and analysis of the initial state of the atmosphere and ocean. This idea of ensemble forecasting has recently become the main thrust in the NWP and GCM climate prediction arenas.

In this project, we carried ensemble forecasting in a different way from the above method by incorporating an advanced statistical model called Canonical Correlation Analysis (CCA). With initial values perturbed in time, we repeatedly ran the CCA model through a sizable series of experiments (13 ensemble members in this study). Given the different initial state of the atmosphere (500 hPa height) and ocean (SST), these experiments indicated the spread of the forecasted 1998 May/June rainfall for all 16 stations in Taiwan. Ensemble forecasting, when made with enough trials, allows the construction of probabilistic forecasts which are then expressed objectively in three categories (wet, normal, and dry). Because many runs are needed to describe the probability distribution (i.e., uncertainty) of a single forecast, however, this procedure is more laborious and requires more computer time.

Forecast results indicate that a majority of stations (15 out of 16 stations) are characterized by a low probability value, ranging from 8 to 38%, in the dry category. In other words, there is little likelihood of having a dry 1998 Mei-Yu season. On the other hand, forecast probabilities for eight stations range from 46 to 70% wetter than their 70th percentiles. More importantly, most of those stations are located in the southern (e.g., Tainan and Kaohsiung) and southeastern (e.g., Taitung, Hengchun, and Tawu) Taiwan. The two northernmost stations, Keelung and Taipei, show very low probabilities (15%) in the wet category. If these forecasts are correct, the implication is that the mesoscale convective systems embedded in southwesterly monsoonal flows will be active this season but their spatial extent will be limited to the southern and southeastern portions of Taiwan.

Our current work may be viewed as a first step in providing a scientifically based and state-of-the-art approach for predicting seasonal to interannual Mei-Yu rainfall for Taiwan. In the future, we intend to extend the CCA model in a similar ensemble forecasting framework to other seasons (e.g., typhoon, spring) and to provide forecast information to CWB personnel well before the season begins. We also plan to explore the predictability of seasonal to interannual rainfall.

Acknowledgments

My deep thanks are extended to Dr. Mong-Ming Lu for her interest, encouragement, kind support, and for inviting me to give a series of lectures at the CWB in December 1997. She has constantly provided a pivotal role in bridging the link between the CWB and me. Mr. Guay-Hong Chen of the Forecast Center of CWB is also thanked for timely providing the data for the analysis in this project. I am very fortunate to have the opportunity to work with them and other research staff at the CWB.

References

Chu, P.-S., 1998: Short-term climate prediction of Mei-Yu rainfall for Taiwan using canonical correlation analysis. Int. J. Climatol., 18, 215-224.

Chu, P.-S., and Y. He, 1994: Long-range prediction of Hawaiian winter rainfall using canonical correlation analysis. Int. J. Climatol., 14, 659-669.

Chu, P.-S., and J. Wang, 1997: Towards the climate prediction of Mei-Yu rainfall for Taiwan. Proceedings of the Conference on Weather Analysis and Forecasting, March 3-5, 1997, Taipei, Taiwan, Central Weather Bureau, 29-31.

Diaz, A.F., C.D. Studzinski, and C. R. Mechoso, 1998:

Relationships between precipitation anomalies in Uruguay and southern Brazil and sea surface temperature in the Pacific and Atlantic Oceans. J. Clim., 11, 251-271.

Lorenz, E.N., 1993: The essence of chaos. University of Washington Press, Seattle, 227 pp.

Tracton, M.S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. Weather and Forecasting, 8, 379-398.

Wallace, J.M., C. Smith, and C.S. Bretherton, 1992: Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies. J. Clim., 5, 561-576.

Yu, Z.-P., P.-S. Chu, and T.A. Schroeder, 1997: Predictive skills of seasonal to annual rainfall variations in the U.S. affiliated Pacific islands: Canonical correlation analysis and multivariate principal component regression approaches. J. Clim., 10, 2586-2599.

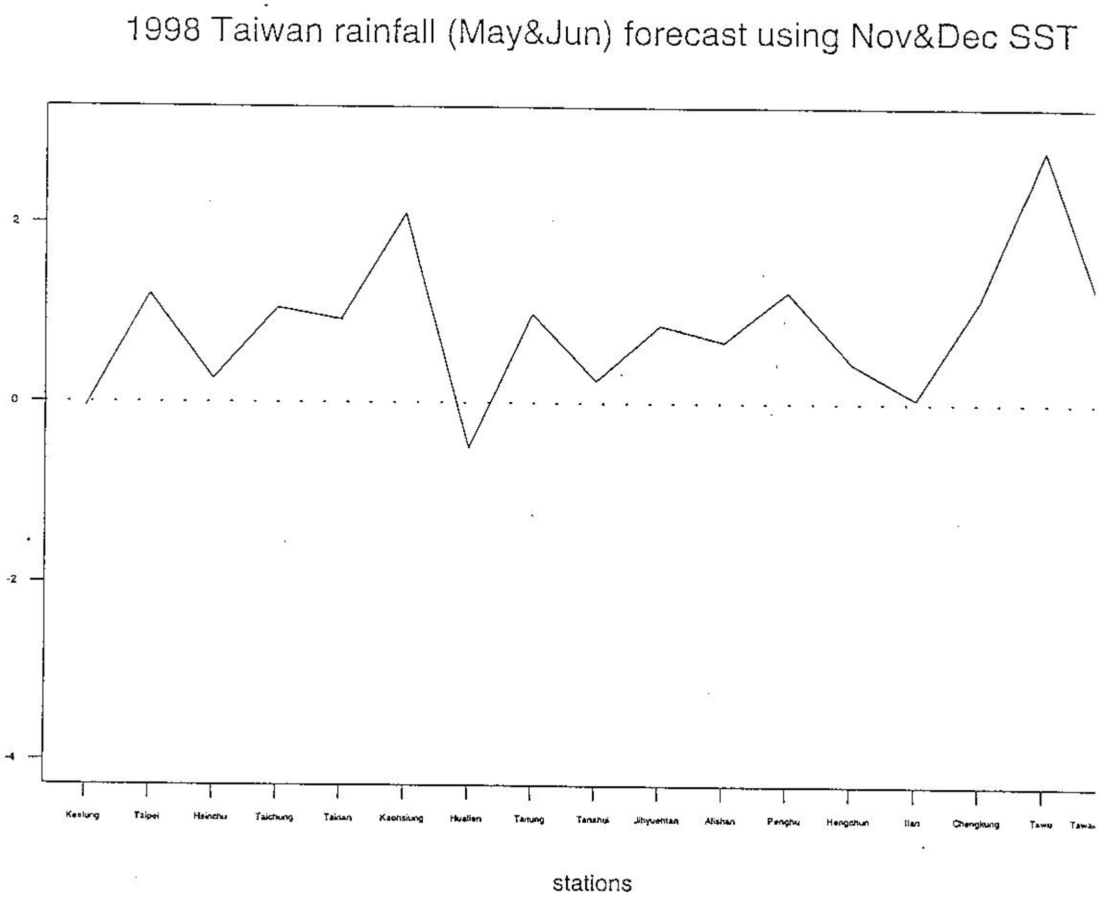
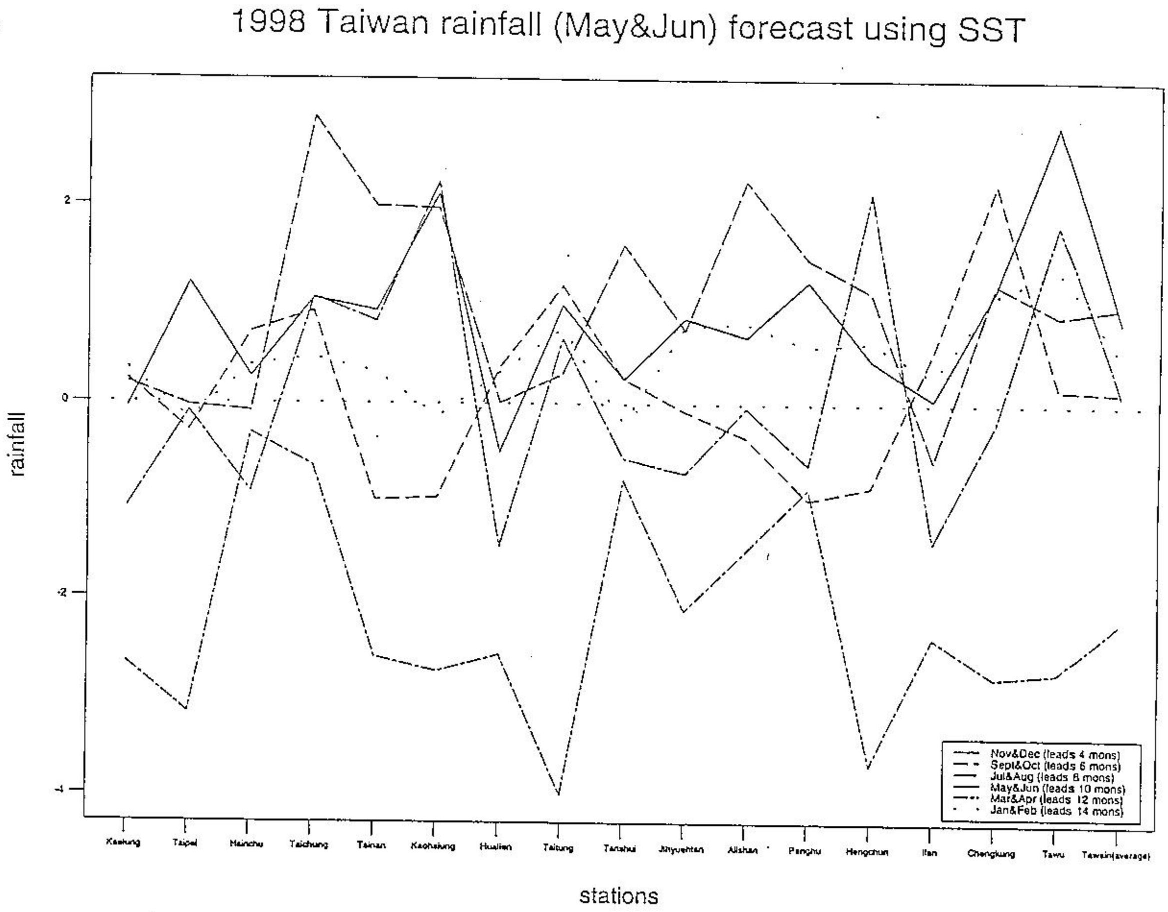


Fig. 1. The 1998 Mei-Yu rainfall forecast by CCA for 16 stations and the average of 16 stations in Taiwan. The vertical scale is the normalized rainfall index. Lead time is four months ahead and the Pacific SST is the only predictor variable.



1998 Taiwan rainfall (May&Jun) forecast using SST

Fig. 2. Same as Fig. 1 but the lead time varies from four months to 14 months ahead.

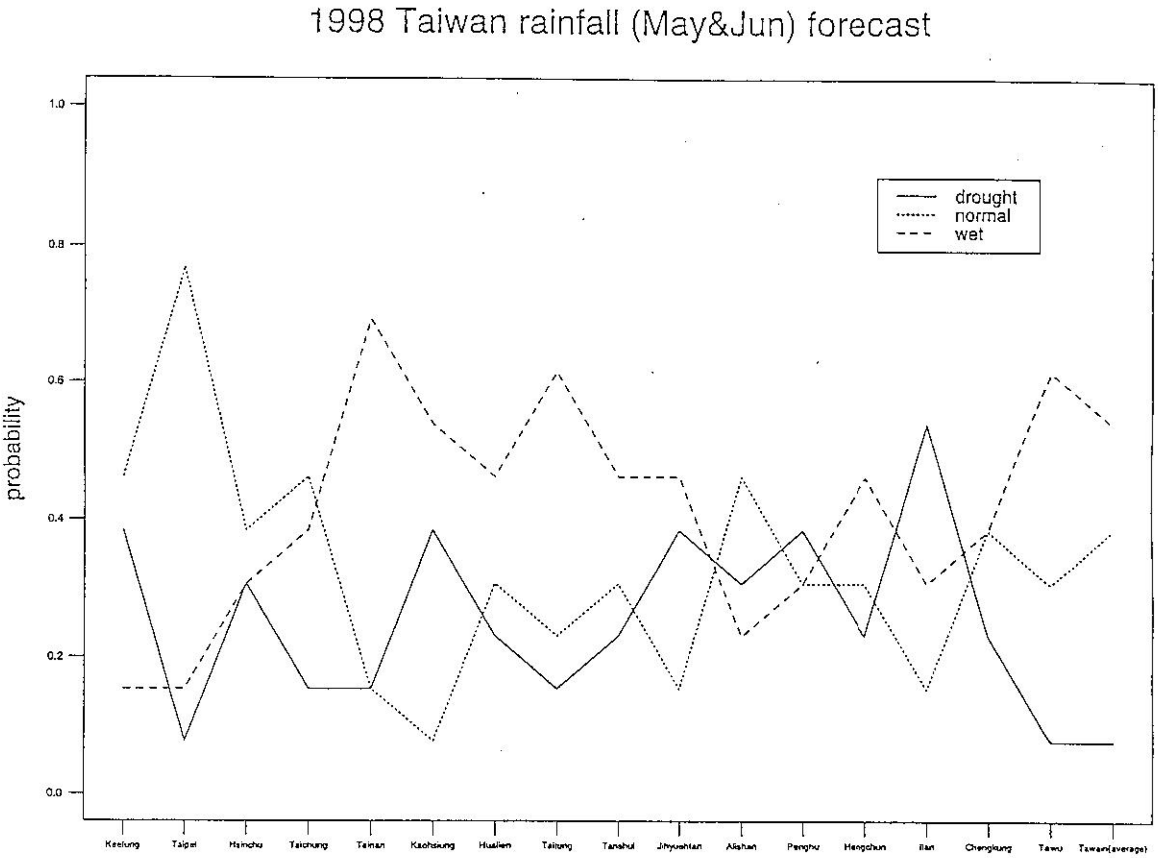


Fig. 3. Ensemble forecasting of the 1998 Mei-Yu rainfall by CCA for 16 stations and the average of 16 stations in Taiwan. Forecasts are expressed in terms of probabilities in three categories (dry, normal, and wet) from 13 ensemble members. For the dry category, the forecast is below the 30th percentile of the observations. For the wet category, the forecast is above the 70th percentile of the observations. For the normal category, the forecast is between the 30th and 70th percentiles of the observations. The Pacific SST and SST/500 mb height are the predictor variables.