Preliminary Performance Evaluation of A Rainstorm Nowcasting System

Edwin S.T. Lai and P. W. Li

Hong Kong Observatory, 134A Nathan Road, Kowloon, Hong Kong

ABSTRACT

A rainstorm nowcasting system SWIRLS (Short-range Warning of Intense Rainstorms in Localized Systems) is currently under development at the Hong Kong Observatory. Phase I of SWIRLS went through operational trial in 1998 and was officially put into service in April 1999. Movement and intensity change of individual radar echo are analyzed by TREC (Tracking Radar Echoes by Correlation). Extrapolation based on the respective TREC vector provides the forecast echo position and intensity, but development of the latter has the option of being constrained by an idealized profile. Meanwhile, making use of the dense raingauge data over Hong Kong, SWIRLS adjusts the reflectivity-rainfall (Z-R) relationship in real time. Combining the echo forecast and the time-adjusted Z-R relationship, SWIRLS generates short-range quantitative precipitation forecasts (QPF) over Hong Kong every 6 minutes, i.e. time interval of radar scans.

QPF by SWIRLS during several heavy rain episodes in 1998 and 1999 is objectively evaluated using common performance measures such as NAP (No Alarm Probability), FAR (False Alarm Rate), Critical Success Index (CSI) and Heidke Skill Index (HSI). Results from this study help to: (a) assess the reliability and robustness of SWIRLS QPFs; (b) reveal strength and weaknesses in SWIRLS QPFs for forecasters' guidance and reference in operating the rainstorm warning system; (c) focus researchers' attention to rain characteristics peculiar to Hong Kong, hence pointing the way forward in SWIRLS development work.

1. Introduction

In Hong Kong, heavy rain in the summer is usually associated with the southwest monsoon, monsoon troughs and tropical cyclones. A three-tiered Rainstorm Warning System (RWS) is in operation based on hourly rainfall rate of 30 mm (amber warning), 50 mm (red warning) and 70 mm (black warning) as recorded by 10 raingauges or more over Hong Kong. While perfect matching of predicted and actual rainfall distribution over a small area like Hong Kong (in the order of 50 x 50 km) remains an illusive goal, operational forecasting systems such as SWIRLS should at least demonstrate a minimum level of skill and capability for forecasters operating the RWS.

A host of performance measures, namely NAP (No Alarm Probability), FAR (False Alarm Rate), Critical Success Index (CSI) and Heidke Skill Index (HSI) are applied to the critical rainfall criteria operationally used in the RWS. Owing to the rarity of black warning cases during 1998 and 1999, this evaluation exercise will focus on the amber (30 mm per hour) and red (50 mm per hour) criteria. But as we would also like to know how SWIRLS performs in leading up to the amber warning, an additional level at 20 mm per hour, say pre-amber threshold, will also be assessed in this study. Additional assessment is also made for criteria as applied to 5 gauges instead of the operational prerequisite of 10 gauges. The idea is again to see how effective SWIRLS is in the build-up stages of heavy rain events.

2.   Echo movement by TREC

Movement of individual radar echo between two consecutive CAPPI scans at 6-minute interval is derived using correlation method, i.e. TREC – Tracking Radar Echoes by Correlation.   The main advantage of using correlation method is that it requires only information from one single radar at two different observation times.   Although the derived motion field is essentially two-dimensional in the horizontal and does not include strong vertical motion normally expected in convective cloud systems, the approximation should be good enough given that the effect of horizontal advection should be more pronounced if only short-term motion is considered.   Once the motion vectors are computed, the entire echo field can be extrapolated to produce a short-term (say, one-hour) forecast field.

TREC method is applied to two successive CAPPI reflectivity fields at 6-minute interval.   The first field (at time T_1) is divided into a number of equally sized "boxes" or two-dimensional arrays of pixels.   Computation can be repeated for all possible arrays found at T_2 to determine which array results in the highest correlation, and the centre of this second array will be the end point of motion vector (i.e. TREC vector).   In practice, to reduce computation time, a search area centred on the first array is prescribed.   The radar reflectivity field has a resolution of 480 x 480 pixels.   The arrays are spaced 5 pixels apart, giving a TREC vector field on a 93 x 93 grid.   The equivalent spatial resolution is of the order of 1 to 5 km depending on the range of scan.

Before the whole vector field is analyzed objectively, each motion vector has to pass a preliminary consistency check.   If a TREC vector deviates more than 25 degrees from the local mean of its neighbourhood (which is calculated from the set of 25 vectors located around it), it will be replaced by this mean vector.   In this way, the "noisiness" as a result of sole consideration of maximum correlation in deriving the motion vector can be reduced at an early stage.   To produce a continuous gridded vector field, a three-pass objective analysis using modified Cressman weighting function is then applied (Cressman 1959).   Results have shown that TREC is able to reproduce realistic wind fields in major weather systems.   In fact, during the close passage of several tropical cyclones in 1999, TREC has demonstrated its usefulness in radar eye fix and timing of gale onset.

The forecast radar reflectivity pattern is obtained by extrapolation based on the smoothed TREC vectors.   Temporal integration makes use of the two-time-level semi-Lagrangian scheme (Staniforth and Cote, 1991).   As the vector field is computed from consecutive radar CAPPI images six minutes apart, the integration scheme can output forecasts in 6-minute interval, and each successive forecast pattern is then used for the next step of integration.

3.   Echo intensity change

The next problem relates to growth or decay in echo size and intensity.   TREC winds provide a basis to estimate the extent of growth and decay associated with each echo pixel.   As the TREC vector traces the trajectory of individual echo, the difference in echo intensity along the flow between two successive radar scans can also be deduced.   This information can be suitably included in the forecast if so desired.

Keeping things relatively simple to start with, there are basically three approaches in predicting short-term echo intensity changes: (a) assuming no change as from the latest scan; (b) assuming linear extrapolation from the latest scans; (c) imposing an idealized intensity profile upon the intensity trends as observed from the latest scans.

Option (a) may seem simplistic at times, but given the complexity of the problem, it has the advantage of providing forecasts with good continuity, or even reasonable estimates in some cases should rain generation and degeneration balance out on the broader scale. Option (b) in general gives reasonable extrapolation within the first hour. But given the short life cycle of convective cells, linear extrapolation will soon become unrealistic and rainfall forecasts based on such an assumption will soon become disproportionately large. Option (c) provides a possible, though crude, remedy by imposing an idealized empirical intensity profile in the Gaussian form. SWIRLS operationally generates both 1-hour and 2-hour rainfall forecasts assuming no intensity change, i.e. option (a), as well as based on forecast intensity change deduced from option (c).

4. Radar and raingauge rainfall re-analysis

Conventionally, the Marshall-Palmer relationship is used in converting echo intensity (i.e. reflectivity, Z) to rainfall, R. But it is well known that the relationship is much generalized and does not apply well in many cases. To obtain a better estimate, SWIRLS makes use of raingauge data over Hong Kong to calibrate radar reflectivity in real time. In other words, the Z-R relationship will be adjusted in time as the rain event unfolds.

To optimize the parameters a and b in the conventional $Z = aR^b$ relationship, radar reflectivity is correlated every five minutes with the rainfall recorded by the raingauges underneath. The adjusted Z-R relationship is used subsequently in the forecast module, based on TREC winds, to convert the forecast radar reflectivity into rainfall over the raingauge positions. The procedures adopted largely follow Zawadzki *et al.* (1987). An example of SWIRLS forecast during the passage of Typhoon York on 16 September 1999 is shown in Figure 1.

When there are not enough reporting radar-raingauge pairs, such as when a rainstorm is still outside Hong Kong, either the standard Marshall-Palmer relationship ($a = 200$, $b = 1.6$) or the Z-R relationship from the previous rain episode will be used as first guess. As more radar-raingauge pairs become available, the Z-R relationship will be updated accordingly.

5. QPF Verification

Using notations of Y = Yes (rainfall criterion reached) and N = No (rainfall criterion not reached), we consider the four possible outcomes of YY, NY, YN and NN where LHS denotes forecast conditions and RHS denotes actual conditions (e.g. YN will translate as forecast criterion reached but actual criterion not reached). Then the four performance measures used are defined as:

NAP (No Alarm Probability) = NY / (YY+NY)
FAR (False Alarm Rate) = YN / (YY+YN)
CSI (Critical Success Index) = YY / (YY+NY+YN)

HSI (Heidke Skill Index) = [(YY+NN) – F] / [(YY+NY+YN+NN) – F]

In HSI, F can be any low-skill reference techniques such as persistence or climatology. In our study, F is a random forecast given by:

F = [(YY+NY)*(YY+YN)+(YN+NN)*(NY+NN)] / (YY+NY+YN+NN).

In both NAP and FAR, it is both desirable to have small values. In CSI, higher values mean better performance. In HSI, we want to stay on the positive side to beat the low-skill alternative.

And so it transpires, heavy precipitation in Hong Kong in the rain season of 1999 is very much dominated by visiting tropical cyclones (TC). Verification is carried out for Maggie, Sam, York and Cam, i.e. TCs in 1999, each over a 4-hour period. A total of 143 forecasts have been made (the maximum of 160 forecasts not reached due to some data losses in the cases of Sam and York). Results from NAP, FAR, CSI and HSI are tabulated in Tables 1, 2, 3, 4 respectively. Despite the shortage of good cases in 1999, we are also interested in obtaining a preliminary evaluation on how SWIRLS performs in the more volatile monsoon trough situations as compared to tropical cyclone rain.   As such, results of a prolonged heavy rain event on 9 June 1998, from a total of 80 forecasts over an 8-hour period, are also shown alongside.

6.   Discussion

From Tables 1 – 4, some interesting observations are made.

The amber level of 30 mm per hour stands out as the threshold SWIRLS is most comfortable with. While judgment on the red level is reserved owing to comparatively fewer cases, it is obvious that SWIRLS holds no obvious advantage in the pre-amber stages of lesser rain. If tuning of the system towards an optimal level is required, then the setting is more or less appropriate as it is since most of the forecast-based decision-making relating to the RWS revolves around the amber and pre-red stages.

In the build-up towards amber, an alternative way to gain some forecast lead time is to keep count of the number of gauges reaching amber criterion. Focussing on the 30 mm per hour column in the tables, it can be seen that reducing the number of gauges to, say, five in general gives better results.

Another major dilemma faced by the forecasters is whether to put his faith on QPF with intensity change or without intensity change. Indications from the verification suggest that it is essential to use the intensity change option. Doing so will significantly reduce NAP at the expense of only a slight increase in FAR. But ideally, the absolute values in FAR need to be reduced further in the next phases of SWIRLS. In terms of CSI and HSI, amber warning at the operational 10 gauges prerequisite based on forecast intensity change peaks at the values of 0.33 and 0.29 respectively. If a build-up stage of 5 gauges is inserted, the corresponding CSI and HSI values are as high as 0.54 and 0.48 respectively.

Preliminary comparison between tropical cyclone rain and monsoon trough rain as represented by the 9 June 1998 case seems to suggest no significant differences in skill level

at the amber level. However, i is interesting to note that while adding intensity change in the QPF generally gives better results in the red warning for the TC cases, improvement in the case of 9 June 1998 is only marginal. One hypothesis is that intense rain exceeding 50 mm per hour in the case of tropical cyclone rainbands is very much the result of rain concentration by advective processes, mechanisms which currently are well handled by SWIRLS. In contrast, intense rain in monsoon trough situations is due more to enhanced convection in localized areas triggered by smaller scale factors; and further development work in SWIRLS on these aspects would be required. Tracking of echo groups, pattern recognition, terrain correction, and real-time intensity profile fitting are just some of the approaches currently under investigation.

While we hope to compile more verification statistics as SWIRLS gains further field experience, the preliminary results obtained so far are already found to be quite useful for operational guidance and research reference. Notwithstanding the best of theories and numerical techniques, there will always be a place for the engineering approach in the design of a QPF nowcasting system tailor-made for the specific locale, both in terms of rain characteristics and operational warning requirement. For example, CSI scores for the VSRF system in JMA ranges between 0.2 and 0.4 for rainfall rate of 10 mm per hour (JMA 1998). If verification for this criterion is run on SWIRLS, the CSI score will presumably be quite unsatisfactory. But in Hong Kong, forecasters are not so concerned about rain of 10 mm per hour anyway! And if systematic verification can guide us towards the optimal level and areas of interest, so much the better.

References

Cressman, G.P., 1959: An operational objective analysis system. *Mon. Wea. Rev.,* **87**, 367-374.

Japan Meteorological Agency (JMA), Forecast Department, Forecast Division, 1998: Analysis and Very-Short-Range Forecasting of Precipitation.

Staniforth, A. and J. Cote, 1991: Semi-Lagrangian integration schemes for atmospheric models – a review. *Mon. Wea. Rev.*, **119**, 2206-2223.

Zawadzki, I., C. Desrochers & E. Torlaschi, 1986: A Radar-raingauge Comparison. Preprints, *23rd Conf. on Radar Meteorology*, Snowmass, Colorado, Amer. Meteor. Soc., 121-124.

Table 1: NAP (No Alarm Probability)

(a)

10 gauges reaching criteria		20 mm per hour		30 mm per hour		50 mm per hour	
Intensity Change	No	TCs in 1999	0.58	TCs in 1999	0.83	TCs in 1999	1.00
		9 June 1998	0.78	9 June 1998	1.00	9 June 1998	1.00
	Yes	TCs in 1999	0.16	TCs in 1999	0.10	TCs in 1999	0.45
		9 June 1998	0.00	9 June 1998	0.07	9 June 1998	0.83

(b)

5 gauges reaching criteria		20 mm per hour		30 mm per hour		50 mm per hour	
Intensity Change	No	TCs in 1999	0.50	TCs in 1999	0.60	TCs in 1999	1.00
		9 June 1998	0.75	9 June 1998	1.00	9 June 1998	1.00
	Yes	TCs in 1999	0.20	TCs in 1999	0.04	TCs in 1999	0.38
		9 June 1998	0.08	9 June 1998	0.00	9 June 1998	0.85

Table 2: FAR (False Alarm Rate)

(a)

10 gauges reaching criteria		20 mm per hour		30 mm per hour		50 mm per hour	
Intensity Change	No	TCs in 1999	0.70	TCs in 1999	0.75	TCs in 1999	0.00
		9 June 1998	0.90	9 June 1998	1.00	9 June 1998	0.00
	Yes	TCs in 1999	0.69	TCs in 1999	0.66	TCs in 1999	0.73
		9 June 1998	0.88	9 June 1998	0.38	9 June 1998	0.80

(b)

5 gauges reaching criteria		20 mm per hour		30 mm per hour		50 mm per hour	
Intensity Change	No	TCs in 1999	0.77	TCs in 1999	0.44	TCs in 1999	1.00
		9 June 1998	0.92	9 June 1998	1.00	9 June 1998	0.00
	Yes	TCs in 1999	0.75	TCs in 1999	0.45	TCs in 1999	0.84
		9 June 1998	0.85	9 June 1998	0.73	9 June 1998	0.69

Table 3: CSI (Critical Success Index)

(a)

10 gauges reaching criteria		20 mm per hour		30 mm per hour		50 mm per hour	
Intensity Change	No	TCs in 1999	0.21	TCs in 1999	0.11	TCs in 1999	0.00
		9 June 1998	0.07	9 June 1998	0.00	9 June 1998	0.00
	Yes	TCs in 1999	0.29	TCs in 1999	0.33	TCs in 1999	0.22
		9 June 1998	0.12	9 June 1998	0.60	9 June 1998	0.10

(b)

5 gauges reaching criteria		20 mm per hour		30 mm per hour		50 mm per hour	
Intensity Change	No	TCs in 1999	0.19	TCs in 1999	0.31	TCs in 1999	0.00
		9 June 1998	0.06	9 June 1998	0.00	9 June 1998	0.00
	Yes	TCs in 1999	0.24	TCs in 1999	0.54	TCs in 1999	0.15
		9 June 1998	0.14	9 June 1998	0.27	9 June 1998	0.11

Table 4: HSI (Heidke Skill Index)

(a)

10 gauges reaching criteria		20 mm per hour		30 mm per hour		50 mm per hour	
Intensity Change	No	TCs in 1999	0.12	TCs in 1999	0.05	TCs in 1999	0.00
		9 June 1998	-0.02	9 June 1998	-0.03	9 June 1998	0.00
	Yes	TCs in 1999	0.20	TCs in 1999	0.29	TCs in 1999	0.29
		9 June 1998	0.02	9 June 1998	0.28	9 June 1998	0.12

(b)

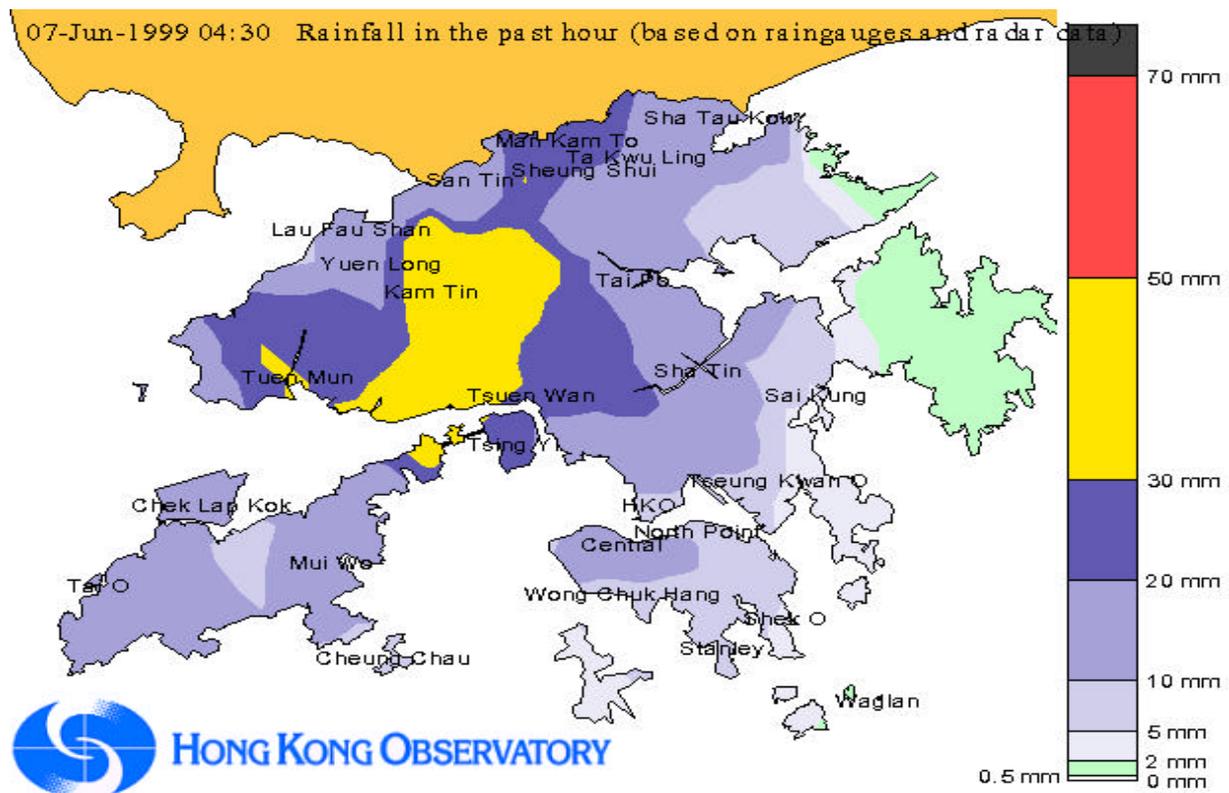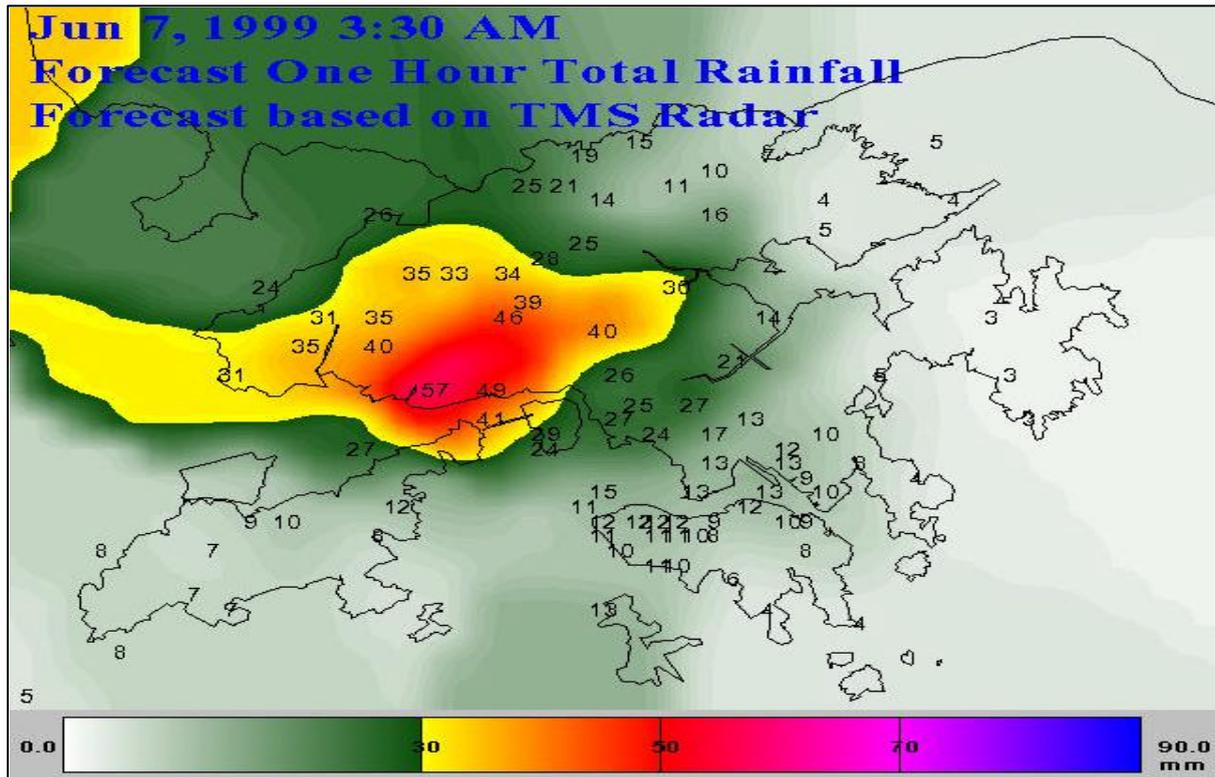| 5 gauges reaching criteria		20 mm per hour		30 mm per hour		50 mm per hour	
Intensity Change	No	TCs in 1999	0.04	TCs in 1999	0.27	TCs in 1999	-0.05
		9 June 1998	-0.14	9 June 1998	-0.07	9 June 1998	0.00
	Yes	TCs in 1999	0.09	TCs in 1999	0.48	TCs in 1999	0.13
		9 June 1998	-0.01	9 June 1998	0.08	9 June 1998	-0.09

Figure 1   1-hour rainfall forecast by SWIRLS for Typhoon Maggie issued at 3:30 HKT on 7 June 1999 (upper panel) and the corresponding one hour rainfall accumulation reported by the raingauges from 3:30 HKT to 4:30 HKT on 7 June 1999 (lower panel).

8