Special Focus Session SF 1b

SPECIFYING INTENSITY FORECAST UNCERTAINTY

Session Leader: Russell L. Elsberry
Naval Postgraduate School
Monterey, CA 93943 USA
Email: Elsberry@nps.edu
Telephone: +1-831-656-2373

Session Contributors: *Jim Goerss (NRL – USA); Charles Sampson (NRL – USA); Hsiao-Chung Tsai (NPS – USA)*

SF 1b.1 Introduction

The theme of IWTC-VIII is "Quantifying and Communicating Forecast Uncertainty." Most tropical cyclone (TC) warning centres include some indication of TC track forecast uncertainty in terms of a cone about their track forecasts. The theme of this Special Focus Session 1b is that warning centres should also provide some indication of TC intensity forecast uncertainty about their intensity forecasts.

In the past, the U. S. National Hurricane Center (NHC) suggested to their customers that they should add (or subtract?) one Saffir-Simpson intensity category as a measure of the likely uncertainty in the NHC intensity forecast. In a NHC forecast advisory for Tropical Storm Simon in the eastern North Pacific on 2 October 2014, the intensity uncertainty was given as 15 kt. Given that there is uncertainty in the initial intensity even with *in situ* aircraft observations, what is the actual intensity uncertainty, and how does it vary with intensity categories? We typically assume uncertainty increases with increasing forecast interval – how does that apply to TC intensity forecasts? In the short-term (say 12 h – 36 h), and in the absence of rapid intensification (RI; 30 kt in 24 h), what is the intensity forecast error growth rate? While forecasters may not know the timing and magnitude of the RI (or an eyewall replacement cycle) event, should that keep warning centres from providing an estimate (say the 68% likelihood that the verifying intensity will lie within the intensity spread) of intensity uncertainty just as they do for track forecast uncertainty?

While the track uncertainty is clearly the most important factor in assessing the risk of damage from a TC, the range of tracks within the track uncertainty cone also implies a range of intensities, especially on the 3 – 5 day time scales. Some substantial intensity changes associated with track changes include: (i) Landfall versus non-landfall, or a return to open ocean after a landfall; (ii) Recurvature with a subsequent intensity decrease versus a continued westward track in the deep tropics, or northwestward into the subtropics, with an expectation of continued intensification; and (iii) Passage over a cold ocean eddy or a cold wake from a previous TC that may interrupt the intensification. If one of these substantial intensity changes may (may not) occur given the official track forecast, should the warning centre inform their customers that a larger (smaller) intensity forecast uncertainty applies?

The objective of this Special Focus Session SF 1b is to inform forecasters that guidance products are becoming available (specifically at the Joint Typhoon Warning Center [JTWC] for the western North Pacific) that will quantify intensity forecast uncertainty. The first of these guidance products is the Goerss and Sampson (2014) prediction of consensus TC intensity forecast error, which is based on the consensus (called S5YY) intensity model used by JTWC forecasters. The development and benefits of various consensus approaches for TC intensity forecasting are described in Rapporteur report 2.7. The advantage of the Goerss and Sampson consensus

approach to provide an estimate of the intensity uncertainty is that it is based on the same intensity guidance products that are included in the S5YY consensus intensity product, which is one of the primary tools for generating the JTWC official intensity forecast.

The second intensity uncertainty product is the Tsai and Elsberry (2014a, b) Weighted ANalog Intensity (WANI) prediction technique that is based on the 10 best historical analogs to the JTWC official track forecast and the current TC intensity. In the latest WANI version, an objective detection of substantial intensity bifurcation situations (e.g., landfall versus non-landfall, or recurvature versus non-recurvature) is provided with two clusters for the WANI intensity forecasts and intensity spreads. These intensity bifurcation situations are examples of where the forecaster can add value by selecting the correct cluster WANI intensity and also provide a smaller (and better) intensity uncertainty. Because this analog technique of Tsai and Elsberry (2014a, b) requires only a historical data base and can be run on a desktop computer in a few seconds, it could be applied locally in any TC warning centre.

The Goerss and Sampson (2014) technique is described in Section 2, and the Tsai and Elsberry (2014a, b) approach will be described in Section 3. Discussion and recommendations are given in Section 4.

SF 1b.2 Intensity uncertainty derived from predictions of consensus intensity forecast error

Goerss and Sampson (2014) developed a method for estimating the intensity "half-width" to be added to a consensus forecast of the intensity to include 67% of the observed intensity values. Only their application to the western North Pacific consensus intensity forecast called S5YY will be described here (Goerss and Sampson also considered the Atlantic and eastern North Pacific basins) because the Tsai and Elsberry (2014a, b) approach for intensity uncertainty has only been developed for the western North Pacific (although a preliminary version exists for the entire Southern Hemisphere).

Since it is important to understand the S5YY consensus intensity forecast upon which the Goerss and Sampson approach is based, a brief description is provided. The first step in creating the S5YY intensity consensus is to construct interpolated track forecasts using numerical weather prediction (NWP) model forecasts, because these NWP model forecasts from a given initial time are available too late to be used by forecasters during that particular forecast cycle. A post-processor called "the interpolator" (see Goerss et al. 2004; Sampson et al. 2008) generates both track and intensity forecasts that are available at approximately synoptic time + 1.5 hours, which is then in time to be used for the official forecast. The interpolated intensity forecasts from three NWP models were found to provide reasonable intensity skill (the first three aids in Table 1).

The other six aids utilized in S5YY are variations of the SHIPS (DeMaria et al. 2005) and LGEM (DeMaria et al. 2009) intensity techniques that are provided forecast tracks from three members of the operational track consensus used at JTWC. Ideally, these intensity consensus members generated from the SHIPS and LGEM should be provided with the thermodynamic and dynamic inputs from the NWP model corresponding to the interpolated forecast track. These thermodynamic and dynamic fields would then be consistent with the vortex structure as well as with the interpolated track, and thus should provide for more realistic SHIPS/LGEM computations (e.g., vertical wind shear computation) for that member. Goerss and Sampson (2014) could not obtain complete model fields for all of the member models, so they substituted other model fields. For the GFS interpolated model tracks (AVNI), SHIPS/LGEM is run with dynamic fields (u and v wind components) from GFS, but with the Navy Operational Global Atmospheric Prediction System (NOGAPS; Hogan and Pauley 2007) temperature, relative humidity and geopotential height fields as required for the input variables in SHIPS and LGEM. For WBAR (Weber 2001) and NOGAPS

interpolated tracks (WBAI and NGPI, respectively), both SHIPS and LGEM are provided NOGAPS dynamic and thermodynamic fields as inputs.

Table 1. Intensity forecast aids used by Goerss and Sampson (2014) for consensus (S5YY) intensity error prediction in western North Pacific. First column is the member name, and second is a description. The consensus members that utilize the SHIPS or the LGEM intensity techniques are denoted with "*".

Forecast Aid ID	Description
CHII	Interpolated Coupled Hurricane Intensity Prediction System (CHIPS; Emanuel et al. 2004)
COTI	Interpolated Coupled Ocean-Atmosphere Mesoscale Prediction System – Tropical Cyclone (COAMPS-TC; Doyle et al. 2012)
GFNI	Interpolated Navy GFDL model (Rennick, 1999)
LGEN*	LGEM, NOGAPS fields and NGPI track
DSHN*	SHIPS, NOGAPS fields and Next Generation Product Identification (NGPI) track
LGEA*	LGEM, GFS/NOGAPS fields and AVNI track
DSHA*	SHIPS, GFS/NOGAPS fields and AVNI track
LGEW*	LGEM, NOGAPS dynamic fields and WBAI track
DSHW*	SHIPS, NOGAPS dynamic fields and WBAI track

Note the variety of track forecasts that are provided as inputs to the S5YY consensus intensity forecast (Table 1): (i) CHIPS technique is provided the JTWC official track forecast; (ii) dynamical models COAMPS-TC and Navy GFDN of course predict the track as well as the intensity; and (iii) three tracks (NGPI, AVNI, and WBAI) used for the SHIPS and LGEM intensity techniques. Thus, track forecast uncertainty is an important input to the S5YY consensus intensity forecast. However, dynamic and thermodynamic inputs (albeit interpolated from slightly different track positions) for the SHIPS and LGEM techniques come from NOGAPS, and thus are not very independent intensity information. Since these six intensity estimates make up two-thirds of S5YY intensity consensus, it might be expected that the intensity spread among the nine consensus members will be under-determined (at least in the early portion of the forecast).

Goerss (2007) had previously used a multiple linear regression with various predictors (including the track spread among a number of NWP models) to predict the track error from a consensus of skillful deterministic model track forecasts. Goerss and Sampson (2014) use a similar approach to predict the magnitude of the intensity forecast error from the S5YY using predictors that are available in the real-time files of the Automated Tropical Cyclone Forecast System (ATCF; Sampson and Schrader 2000) prior to official forecast construction at the operational centres. Consensus model intensity spread is defined to be the average of the absolute differences between the member intensity forecasts and the consensus intensity forecast. The possible predictors Goerss and Sampson (2014) examined were consensus model spread; forecast TC intensity and intensity change; initial TC intensity and position; TC speed of motion; and the number of the nine consensus members in Table 1. Forecast TC intensity and intensity change are determined using the interpolated JTWC official forecasts (JTWI). All cases where an official intensity forecast was made and verified were used in the Goerss and Sampson study.

Consider first the relationships between the predictors and the S5YY TC intensity forecast errors. For S5YY in the western North Pacific basin, initial TC latitude was the leading predictor for the 24-h to 72-h forecasts with correlations ranging from -0.27 to -0.34, and was the second leading predictor for the 12-h and 96-h forecasts. Initial TC longitude (LONI) was the leading predictor for the 12-h forecast with a correlation of -0.20 and the second leading predictor for the 24-h to 48-h forecasts. Note that the consensus model intensity spread was not selected as a predictor in these short-range forecasts. However, a consensus model spread was the leading predictor for the 96-h and 120-h forecasts with correlations of 0.27 and 0.24, respectively, and the second leading predictor for the 72-h forecasts. As an example, the relationship between the 96-h consensus model spread and forecast absolute intensity error is shown in Figure 1. The reason the correlation is only 0.27 is there is a large range of intensity errors for consensus model intensity spreads in the range of 5 – 20 kt.

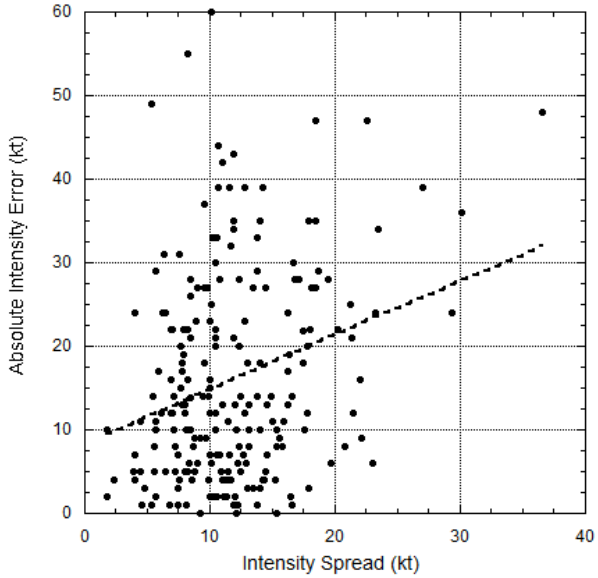

Figure 1. S5YY 96-h absolute intensity forecast error vs. consensus model intensity spread for the 2012 western North Pacific season (Goerss and Sampson 2014). The dashed line represents the linear regression fit to the data.

The regression equations and scatter plots displaying the relationship between S5YY absolute intensity forecast error and predicted error for the 12-h through 48-h forecasts for the western North Pacific basin are shown in Figure 2. The correlations of the regression fit for the western North Pacific basin ranged from 0.25 to 0.39. Using this linear regression model, the percent variance of the S5YY TC absolute intensity forecast error that could be explained for the 2012 western North Pacific season ranged from 6%-15%.

The regression equations and scatter plots displaying the relationship between S5YY absolute intensity forecast error and predicted error for the 72-h through 120-h forecasts for the western North Pacific basin are shown in Figure 3. The correlations for the western North Pacific basin ranged from 0.24 to 0.44. The percent variance of S5YY absolute intensity forecast error that could be explained for the 2012 western North Pacific seasons ranged from 6%-19%.

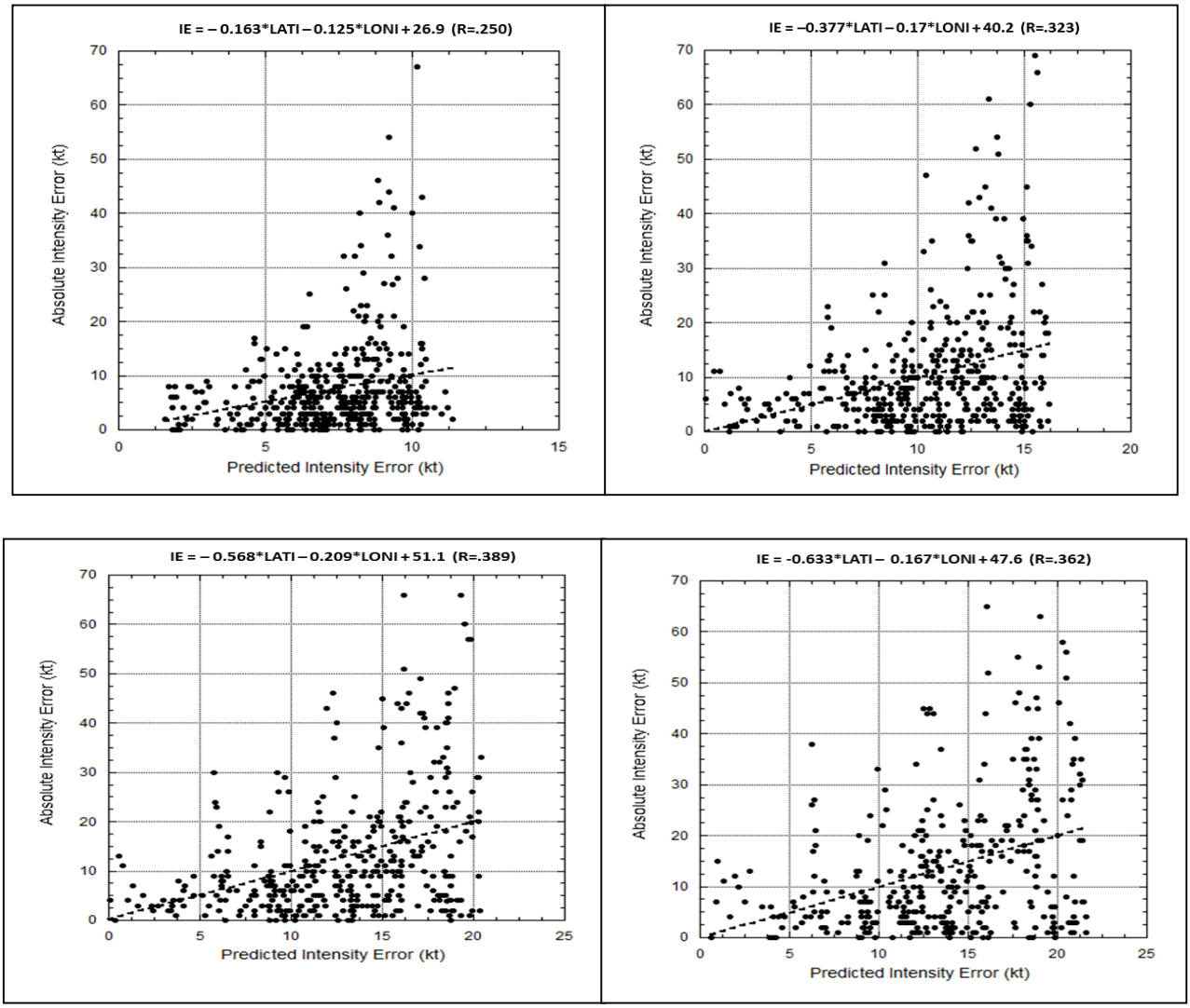

Figure 2. S5YY 12-h (top left), 24-h (top right), 36-h (bottom left), and 48-h (bottom right) absolute intensity forecast error vs. predicted error for the western North Pacific basin during 2012 (Goerss and Sampson 2014). The equation for the predicted error (IE) is shown at the top of each panel. The dashed lines represent the linear regression fit to the data.

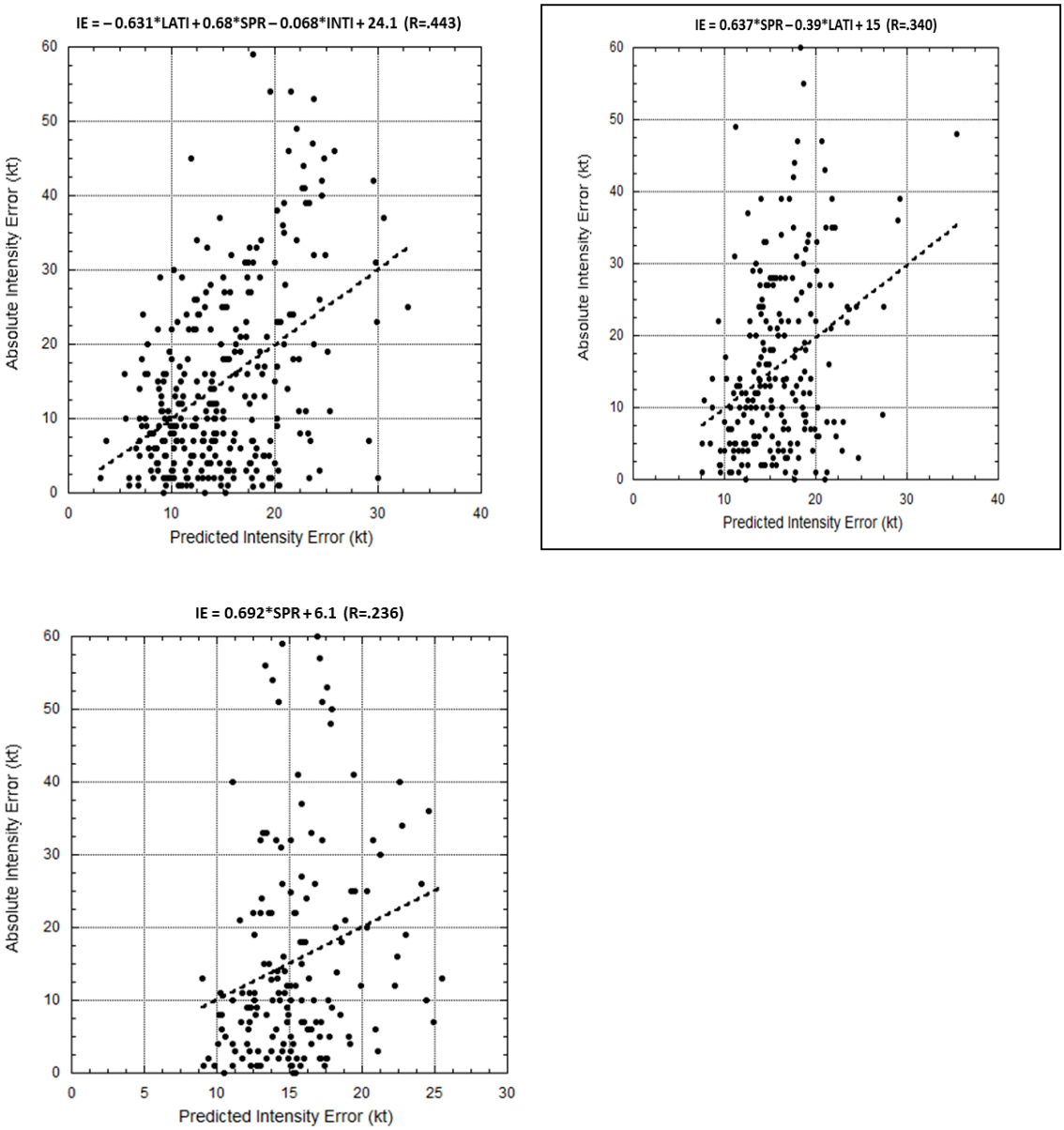

Figure 3. S5YY 72-h (top left), 96-h (top right), and 120-h (bottom) absolute intensity forecast error vs. predicted error (top left), for the western North Pacific basin during 2012 (Goerss and Sampson 2014). The equation for the predicted error (IE) is shown at the top of each panel. The dashed line represents the linear regression fit to the data.

Goerss and Sampson (2014) then construct intervals centred on the S5YY consensus intensity forecasts that will contain the verifying TC intensity roughly 67% of the time. These intensity intervals were "half-widths" that were computed by adding a constant for each forecast interval to the predicted S5YY consensus intensity forecast error that was calculated from the linear regression equations. The constants chosen so that the verifying TC intensities in the training set would be contained within the interval centred on the S5YY forecast intensity roughly 67% of the time were 0.5 kt at 12 h, 1.5 kt at 24 h, 1.4 kt at 36 h, 2.4 kt at 48 h, 2.9 kt at 72 h, 3.6 kt at 96 h, and 1.4 kt at 120 h. The results of applying these intensity confidence intervals to the S5YY forecasts for the 2012 western North Pacific season are summarized in Table 2.

Table 2. Verification summary for S5YY-derived intensity confidence intervals for the 2012 western North Pacific training set

Forecast Length (h)	12	24	36	48	72	96	120
Average ½ Range (kt)	7	12	14	16	17	19	17
Minimum ½ Range (kt)	3	3	3	3	5	11	10
Maximum ½ Range (kt)	11	18	22	24	35	39	29
Forecasts within Range (%)	67	67	67	67	67	67	67
Number of Forecasts	489	448	413	375	296	220	159

Note that the 2012 western North Pacific results are for a dependent (i.e., training) data set. Consequently, the verification intensities at all forecast intervals are within the range of 67% (Row 5), because the additive constraints were selected to satisfy that desired percentage. The most important numbers in Table 2 are the Average ½ ranges (Row 2) as a function of forecast interval. On average, ± 7 kt would need to be added to each S5YY consensus model intensity forecast at 12 h to achieve the goal of 67% of the verifying intensities to be within this intensity spread. By contrast, a Maximum ½ range (Row 4) of ± ll kt would be required to include 100% of the verifying 12-h intensities. Note also that the Average ½ range rapidly increases to 16 kt at 48 h and then approximately levels off at around 17 kt to 120 h. Similarly, the Maximum ½ range rapidly increases and actually would need to be 39 kt at 96 h in order that all verifying intensities would be included.

Goerss and Sampson (2014) were able to do an independent test of the regression equations that were computed for the Atlantic and eastern North Pacific basins for the 2008-2011 seasons. The independent test for the 2012 Atlantic season indicated the predicted intensity intervals were too large as they included the verifying intensities from 71% - 81% of the time, compared with the desired target of 67%. The primary reason for the over-prediction is the unusually small size of the Atlantic consensus model intensity forecast errors for 2012. For the 12-h to 48-h forecasts, the errors ranged from 5.5 to 11 kt compared with a range of 7 to 13 kt for the 2008-2011 dependent sample from which the regression equations were derived. For the 72-h to 120-h forecasts, the errors ranged from 13 kt to 14 kt compared with a range of 15 to 16 kt for the dependent sample.

For the eastern North Pacific, the predicted intensity intervals were too large as they included the verifying intensity from 73%-85% of the time. For the 12-h to 48-h forecasts, the consensus model intensity forecast errors for 2012 ranged from 6 to 12 kt compared with a range of 7 to 16 kt for the dependent sample. For the 72-120 h forecasts, the intensity errors ranged from 12 to 13 kt compared with a range of 18 to 19 kt for the dependent sample.

Goerss and Sampson (2014) suggest that forecasters can use these average half-widths as intensity confidence intervals to determine in real-time how much (or little) confidence can be ascribed to individual S5YY consensus intensity forecasts. In the future, this new guidance may be employed in the Monte Carlo Wind Speed Probability Product (DeMaria et al. 2009) to reduce or expand the wind probability plume for given situations. One of the differences between the present version of the S5YY and the WANI intensity spread of Tsai and Elsberry (2014a, b) is that

the latter approach is situation-dependent (i.e., the WANI intensity spread varies with every forecast).

SF 1b.3 Situation-dependent intensity and intensity spread guidance

Elsberry and Tsai (2014) developed a new "situation-dependent" intensity prediction technique for the western North Pacific that was provided knowledge of the TC track (for the development sample, the JTWC best-track file was utilized). Ten best historical track analogs to the JTWC track were selected and the simple arithmetic average intensity out to 120 h corresponding to these 10 tracks was called the Situation-Dependent Intensity Prediction (SDIP). Since this SDIP may be considered as the intensity expected to occur based only on past events, Elsberry and Tsai (2014) proposed use of the SDIP as a new skill measure that takes the track forecast into account. Comparison of the various intensity guidance techniques relative to SDIP would thus illustrate the inherent skill due to prediction of the physical mechanisms leading to intensity change beyond knowledge of the track forecast. This use of SDIP as the skill measure may be useful in evaluating the source of the skill for the recently improved dynamical model intensity forecasts that are described in Rapporteur report 2.7. That is, how much of that improvement in intensity forecasts is due to the also improved track forecasts for those dynamical models?

The basic premise of the SDIP is that the track of the TC is a primary determinant of the intensity on timescales of five days or longer, and the spread among the 10 best historical analogs provides a measure of the range of environmental influences that may be occurring along that track. It is of course also important to consider at what stage in the life cycle the TC is at, i.e., is the TC still a tropical depression or has it already become a tropical storm or a typhoon? For westward tracks at low latitudes during the middle of the season, the large areal extent of warm ocean in the western North Pacific will favour intensification unless landfall is made on the Philippines. For recurvature situations, maximum intensity is expected to occur at or shortly after the time of recurvature because TCs tend to intensify over the warm tropical ocean and then weaken in the post-recurvature phase in response to both the increasing vertical wind shear associated with the midlatitude trough and the decreasing sea-surface temperatures to the north.

In the SDIP (Elsberry and Tsai 2014), the 10 best analogs were selected from JTWC best-tracks within ± 30 days of the current date and ranked according to the average track difference (d_T) between the target storm and the candidate analog storm plus the initial intensity difference (d_V). Although the 10 candidate analog storms were selected according to their ranking with a weighting factor of 0.8 for d_T and 0.2 for d_V, each of the 10 analog storms was then considered to be equally likely in the SDIP technique since a simple arithmetic average of the corresponding intensities was utilized.

Tsai and Elsberry (2014a) made three modifications to the SDIP to develop the Weighted Analog Intensity (WANI) technique for operational use. First, the 10 best historical track analogs are searched from the same historical data base of JTWC best-track files from 1945-2009 as for the SDIP. However, the JTWC official track forecasts during the 2010-2012 seasons rather than the JTWC best-tracks are matched to form a training dataset. This inclusion of the JTWC official track forecast information is analogous with the most skillful consensus intensity forecasts available at JTWC being based on global or regional model tracks that are combined to form the JTWC track forecasts (DeMaria et al. 2014).

The second modification applied to SDIP to derive the WANI technique was to calculate the average track difference (d_T) between the JTWC track forecast and the candidate analog storm with a linearly varying weighting factor of 1.0 at the initial time to 2.0 at 72 h and then a constant weighting factor of 2.0 in the 72 h to 120 h interval. As in the development of SDIP, a range of

weighting factors for d_T and the initial intensity difference d_v was tested for the sample of JWTC official forecast tracks in the development of WANI. Since a similar small spread in the Mean Absolute Errors (MAEs) and correlation coefficients was found as in Figure 2 of Elsberry and Tsai (2014), the same relative weighting factors of 0.8 for d_T and 0.2 for d_v were selected for ranking of the potential analogs. Thus, the final ranking of the candidate analogs in WANI is according to

$$\text{Rank}_{\text{analog}} = W_T(\text{Rank}_T) + W_V(\text{Rank}_V), \quad (1)$$

where $W_T + W_V = 1.0$. The $\text{Rank}_{\text{analog}}$ is then sorted is ascending order to select the 10 best analogs. Selection is limited to only 10 analogs because it may be difficult to find a larger number when the ranking is by similarity with the JTWC track forecast, initial intensity difference, and within ± 30 days. The second conditioning on the initial intensity (V_0) in the SDIP technique for subsamples of $V_0 \leq 35$ kt and $V_0 > 35$ kt was also applied in WANI, because it was found that the MAEs and correlation coefficients had similar differences in time for these V_0 subsamples as in Elsberry and Tsai (2014, Figure 3).

The third modification of SDIP for the WANI technique was to give greater weight to those analog tracks that most closely match the JTWC track forecast, i.e., match the path leading to the landfall position and timing. The weighted mean intensity (V_w) at each time *t* is:

$$V_w = \sum_{i=1}^{n} (w_i \cdot V_i) / \sum_{i=1}^{n} w_i \ , \quad (2)$$

where V_i is the intensity of the i-th track analog, and $w_i = (1/d_{T,i}) / \sum_{i=1}^{n} (1/d_{T,i})$.

Elsberry and Tsai (2014) also developed intensity spread guidance by utilizing the 10 intensities at each forecast interval to 120 h from the 10 best historical track analogs to the JTWC best-tracks. First, an intensity bias correction was developed to reduce the bias in the SDIP forecasts arising from the simple arithmetic average of the 10 best analogs. The second step was to calibrate the spread among the 10 intensity estimates from the track analogs such that at each 12 h the probability of detection was at least 68.26% since the intensities were assumed to be normally distributed about the average intensity.

Tsai and Elsberry (2014a) used a similar methodology as in Elsberry and Tsai (2014) to develop intensity spread guidance for the weighted analog intensity predictions in the WANI. As in Elsberry and Tsai (2014), the "raw" intensity spread at each 12 h among the 10 best historical analogs was not adequate to represent the full probability density function and thus a forecast spread calculation was required. However, it was first necessary to apply a bias correction to reduce any intensity bias in the WANI forecasts that arises from the weighted average of the 10 best analogs. From the approximately 1200 WANI forecasts during the 2010-2012 season, 70% were randomly selected for the training sample and 30% were retained for an independent verification. A linear regression was used at each 12 h to correct for the bias in the intensity

$$V_m = a' \mathbf{X} + b', \quad (3)$$

where V_m is the bias-corrected intensity, **X** represents various predictors, and a´ and b´ are the regression coefficients. In addition to the arithmetic average of the 10 intensities from the analogs, latitude, longitude, initial intensity, spread of the initial intensities, and spread among the 10 analog tracks are provided as potential predictors. This bias correction was successful as the WANI mean

intensity bias was reduced to less than 1 kt over most of the forecast intervals. By contrast, the intensity forecast biases of the raw WANI forecasts generally became more negative (under-forecast) with increasing forecast interval.

Although the ranking of the historical analogs includes the initial intensity with a weighting factor of 0.2, a range of intensities (hereafter raw intensity spread) exists even at the initial time. Elsberry and Tsai (2014, Figure 8) indicated that the probability of detection (PoD) that the actual intensity will lie within the raw intensity spread for the SDIP was 100% at the initial time, but then decreased rapidly with forecast interval. A similar PoD plot for the WANI beginning at 12 h indicated a rapid decrease with forecast interval for both the $V_0 \le 35$ kt and $V_0 > 35$ kt subsamples as well as the ALL sample (Tsai and Elsberry 2014a, Figure 5a). Relative to the desired PoD of 68.26%, the $V_0 > 35$ kt subsample was over-determined until 36 h and under-determined (raw spread too small) beyond that forecast interval. The $V_0 \le 35$ kt subsample was over-determined to about 48 h and was under-determined beyond that time, but was much less under-determined than the $V_0 > 35$ kt subsample in the 96 h to 120 h forecast interval.

The objective of the intensity spread calibration is to achieve as closely as possible the goal of a 68.26% PoD since the intensities are assumed to be normally distributed about the mean. That is, the P[|z| ≤ 1.0] where z is the normal distribution z – score $(x - \mu)/s$. Here x is the variable, μ is the mean, and s is the standard deviation.

In the WANI technique, the raw intensity spreads (σ) are also weighted:

$$\sigma = \sqrt{\sum_{i=1}^{n} (w_i (V_i - V_w)^2) / \sum_{i=1}^{n} w_i} \,, \tag{4}$$

where i=1-n analogs, and $w_i = (1/d_{T,i}) / \sum_{i=1}^{n} (1/d_{T,i})$. The intensity spread is calibrated at each time $t = t_i$

$$\sigma' = \left| a \cdot \sigma + b \right| , \tag{5}$$

where σ′ is the calibrated intensity spread, σ is the original spread, and a and b are the calibration factors to be determined. The calibration factor a is constrained as $0.5 < a < 1.5$ to avoid excessive reductions within the over-determined region or excessive amplifications in the under-determined region. The calibration factor b is constrained to $-0.5\,\overline{\sigma} \le b \le 0.5\,\overline{\sigma}$, where $\overline{\sigma}$ is the overall sample mean forecast spread at forecast interval t_i to ensure realistic values.

A cost function is calculated from the development sample

$$J_t = J_{1t} + J_{2t} + J_{3t} \,, \tag{6}$$

where t is the forecast interval, J_{1t} is the probability of PoD ≥ 68.26%, J_{2t} is the correlation coefficient r between σ'_i and E_i, which is the intensity forecast error defined as forecast intensity minus observed intensity, and J_{3t} is the penalty term, which is the probability of $L_{ratio} = \sigma'_i / E_i \ge 2.0$. Given the different units in the three cost functions, each of them is normalized by dividing by their standard deviation to define a modified cost function J′ as

$$J_t' = J'_{1t} + J'_{2t} + J'_{3t}. \quad (7)$$

This modified cost function is then minimized to obtain the calibration factors *a* and *b* at each 12 h forecast interval.

As expected for the training dataset, the WANI spread calibration procedure was successful in adjusting to about 68% the over-determined and under-determined raw intensity spreads for the full dataset (ALL) and the $V_0 \leq 35$ kt and $V_0 > 35$ kt subsamples. For the independent data set (30% of the 1200 cases), the calibration procedure was largely successful in reducing the PoD for the over-determined intensity spreads at forecast intervals less than 36 h as well as increasing the PoD for the under-determined intensity spreads at forecast intervals beyond 36 h (Tsai and Elsberry 2014a, Figure 5b).

Tsai and Elsberry (2014a, Figure 6) also show the average magnitudes of the calibrated and non-calibrated intensity spreads and for the overall training set and the independent data set for the WANI forecasts are similar. Since these calibrated intensity spreads are quite successful in representing the uncertainty in the situation-dependent WANI technique, Tsai and Elsberry (2014a) proposed use of the WANI spread guidance for the PoD for the official intensity forecast. The objective of this intensity spread guidance is to provide useful uncertainty information for the forecasters, decision-makers, and informed members of the public.

Tsai and Elsberry (2014a) provide several intensity examples of WANI intensity and intensity spread forecasts that are excellent examples of intensity bifurcations. In their first example (Tsai and Elsberry 2014a, Figure 7) of Tropical Storm (TS) Fanapi (12W), the 10 best historical tracks have a large spread between westward-moving and recurving tracks (Figure 4a). Consequently, the intensities associated with the 10 best historical tracks also had a large spread from achieving supertyphoon intensity to continued tropical storm intensities (Figure 4b). Therefore, the WANI intensity spread guidance at 60 h ranged from 30 kt to 120 kt (Tsai and Elsberry 2014a, Figure 7c). The WANI intensity forecast (Figure 4b, red circles) and the JTWC intensity forecast (Figs. 4b, 4c, and 4d, black lines) "went down the middle," and thus considerably under-forecast the actual maximum intensity (Figs. 4c and 4d, boxes). Also shown in Figs. 4c and 4d (dark blue lines) is the intensity evolution for the highest-ranked analog (19W from 1980 season), which achieved a maximum intensity of 120 kt.

The objective of this section is to describe the objective procedure of detecting an intensity bifurcation situation such as in Figure 4b for TS Fanapi, and how two cluster intensity scenarios such as in Figure 4c and Figure 4d are calculated with separate intensity spreads that are considerably smaller than the original WANI spread. The intensity cluster 1 (C1, Figure 4c) is defined to be that cluster that has the larger maximum intensity, and intensity cluster 2 (C2, Figure 4d) is the alternate solution with a lower maximum intensity.

For this TS Fanapi case, the WANI cluster C1 would have been the more accurate intensity forecast and all of the verifying intensities (Figs. 4c and 4d, boxes) would have been within cluster C1 intensity spread guidance. This C1 intensity forecast would have also been more useful than the original WANI forecast or the JTWC intensity forecast, and certainly more useful than the cluster C2 intensity and intensity spread guidance. More examples of intensity bifurcation situations will be given after a description of how the WANI intensity bifurcation situations are detected.

As illustrated in Figure 4b, the 10 best historical analogs may have some large intensity spreads. The flow chart in Figure 5 summarizes the objective detection procedure for an intensity bifurcation situation utilizing the 10 analog intensities. First, the weighted-mean spread (WMS_i) is calculated for each individual case

$$WMS_i = \sum_{t=0h}^{120h} (w_t \cdot \sigma_{i,t}) / \sum_{t=0h}^{120h} w_t \ , \tag{8}$$

where i = 1-n samples, $\sigma_{i,t}$ is the intensity spread each 12 h, and w_t is the same weighting factor as for the weighted-mean intensities in Eq. (2). That is, increasing weight is given to the intensities from the initial time to 72 h, and from 72 h to 120 h the weighting factor is twice the value at t = 0 h. The weighted-mean spread (WMS) for the overall sample is

$$WMS = \sum_{i=1}^{n} WMS_i / n \ . \tag{9}$$

The code allows setting threshold values of the sample mean WMS, the sample mean plus 0.5 times the standard deviation, or the sample mean plus one standard deviation. In the following, the threshold value for detecting an intensity bifurcation is set as the sample mean in Eq. (9).

If the WMS_i over the 120 h forecast period for an individual case of 10 best-analog intensities exceeds the threshold WMS value, the second step in the flow chart (Figure 5) is to apply a hierarchical cluster analysis (Wilks 2011) to the 10 analog intensities to separate them into two clusters. If there are at least three analogs in each cluster, the WANI technique is separately applied to the two clusters to produce weighted-mean intensities (Eq. (2)) and weighted-mean intensity spreads (Eq. (4)) each 12 h to 120 h. Note that the same bias correction as in Eq. (3) and calibration as in Eqs. (5) – (7) are applied as for the original WANI forecast.

The third step in the flow chart (Figure 5) is a requirement that the cluster 1 WANI and cluster 2 WANI intensities must be substantially different. The motivation for this condition is simply that some meaningful intensity difference should exist between the cluster intensity forecasts that would justify having the forecaster search for alternate scenarios. Such a requirement is obviously met for the TS Fanapi case in Figs. 4c and 4d and the JTWC forecaster should be motivated to search for a justification for either a forecast of intensification to a typhoon as in cluster C1 versus a forecast of a continued TS stage. Since the JTWC track forecast for TS Fanapi was for a landfall on Taiwan, the forecaster may decide that the most appropriate warning would be the cluster C1 intensification forecast.

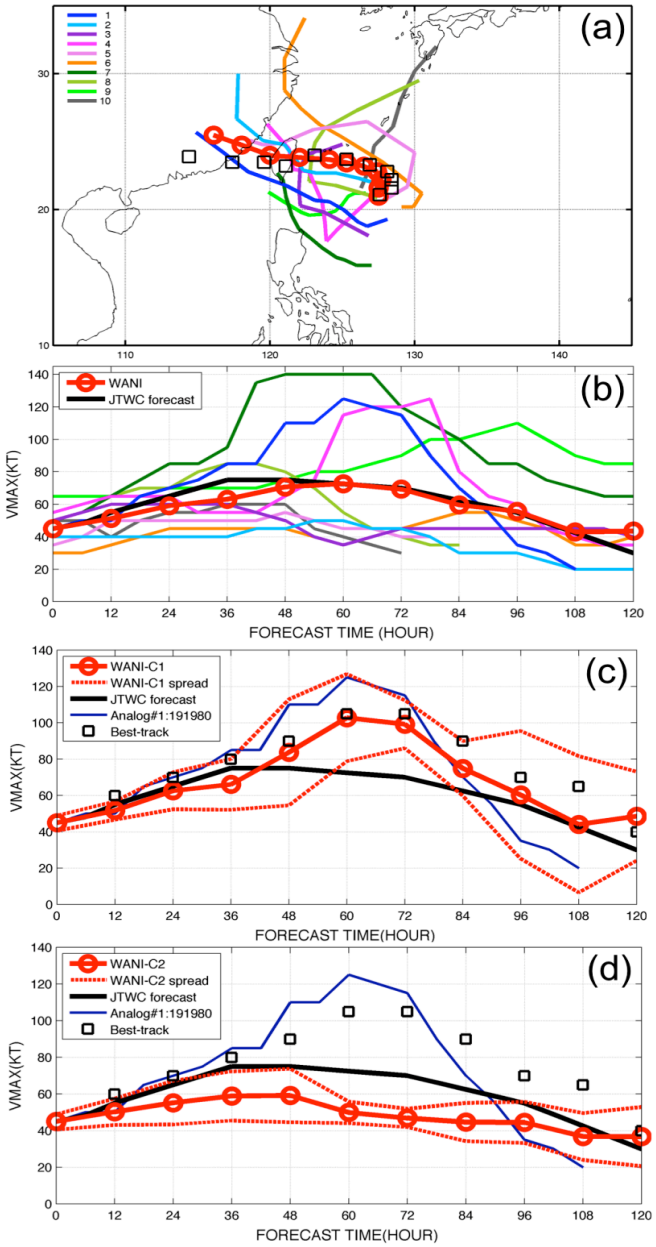

Figure 4. Example of WANI intensity bifurcation and cluster prediction. (a) JTWC official track forecast (Red circles each 12 h) for TS Fanapi (12W) from 1000 UTC 15 September 2010 and 10 best historical analog tracks (rankings indicated by colour bars) selected from JTWC best-tracks from 1945-2012 but excluding the target storms, plus the verifying track positions (boxes each 12 h). (b) Corresponding intensity (V_{max}, knots) evolution for the JTWC forecast (heavy black line) and the 10 best historical track analogs, plus the original WANI intensity forecast (red circles each 12 h) to 120 h. (c) Repeat of the JTWC intensity forecast but now with the cluster C1WANI forecast (red circles) and cluster C1 WANI intensity spread indicated by dashed red lines, plus the intensity evolution (heavy blue line) of the single-best analog (19W in 1990) among the 10 best analogs in panel (a), and also the verifying intensities each 12 h (boxes). (d) As in panel (c), except for the cluster C2 WANI forecast and intensity spread.

Testing for this substantial bifurcation condition is based on the intensity difference

$$VD = |WANI\_C1 - WANI\_C2|, \quad (10)$$

at each 12 h forecast interval, and the code allows for VD thresholds of 10 kt, 15 kt, and 20 kt. The two cluster WANI forecasts are considered to represent a substantial intensity bifurcation if the VD threshold is exceeded for at least 25% of the 12 h forecast intervals within the 120 h forecast period. By contrast, if this VD threshold is not satisfied for 25% of the 12 h forecast intervals, it is considered that no substantial intensity bifurcation exists and the original WANI based on all of the available analogs should be used. After some testing, this VD threshold was set at 15 kt.

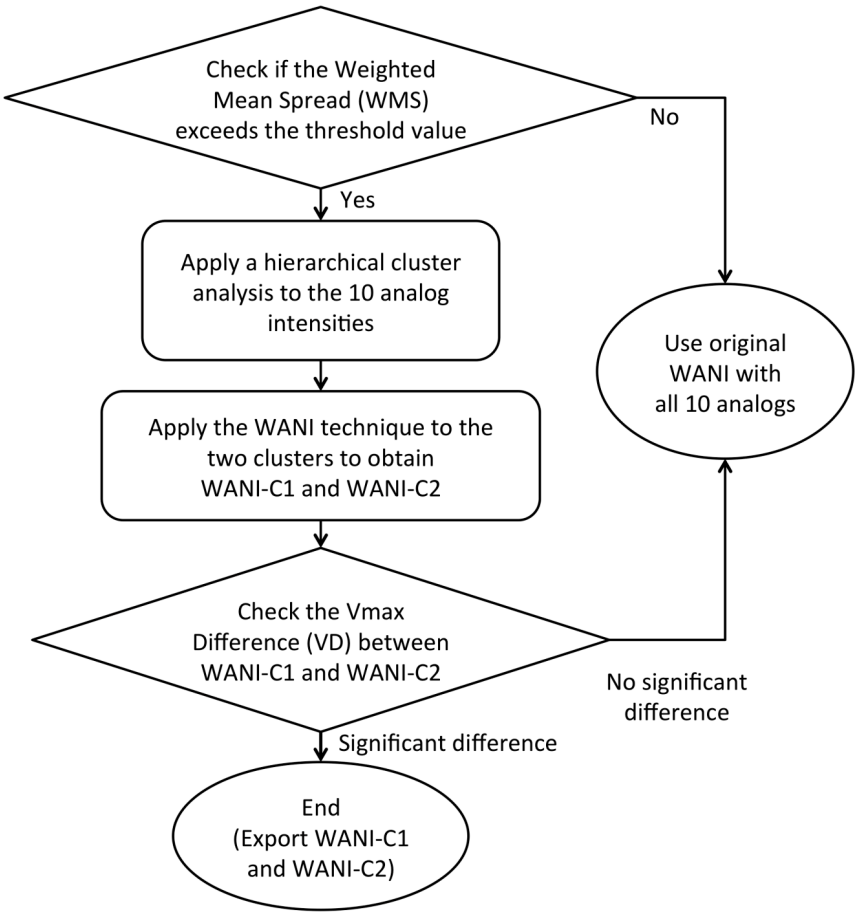

Figure 5. Flow chart of the decisions to objectively detect the existence of a WANI intensity bifurcation situation and thereby use a hierarchical cluster analysis of the 10 best analog intensities to determine two clusters for which the WANI intensities have differences that meet the threshold condition for a substantial intensity bifurcation (see text for definition).

The overall sample size for testing for WANI intensity bifurcations was 1136. A total of 526 WANI forecasts have a WMS_i that exceeded the sample mean WMS in Eq. (9). With a VD threshold of 15 kt, 218 of those WANI forecasts were considered to have a substantial intensity bifurcation. Thus, about 19% of the overall sample of WANI forecasts had a substantial intensity bifurcation according to these threshold conditions, and these objectively-detected bifurcation cases should (and in the future will) be highlighted for the JTWC forecaster to consider two possible intensity and intensity spread scenarios.

The potential impact of the cluster WANI forecasts and intensity spread forecasts will be verified for two assumptions: (i) Forecaster was perfect (P) and always picked the correct cluster, and (ii) Forecaster always picked the wrong (W) cluster. Based on these assumptions, the mean absolute errors (MAEs) for the cluster WANI forecasts are compared in Figure 6a with the original WANI forecasts based on all 10 of the best analog intensities and the JTWC official intensity forecasts. As in the larger sample of WANI forecasts in Tsai and Elsberry (2014a, Figure 2a), the original WANI MAEs for this smaller sample of bifurcation situations are only slightly larger than the JTWC official intensity forecast errors through 84 h, but are 4 kt smaller than the JTWC errors at 120 h.

The most important conclusion from Figure 6a is that if the JTWC forecaster could perfectly pick the correct cluster in these 19% of the overall sample with an intensity bifurcation, the MAEs would be smaller than the JTWC intensity errors already by 24 h and the advantage of picking the correct cluster forecast would steadily increase to an error reduction of 8 kt at 120 h. This large improvement may be attributed to the requirement that these intensity bifurcation situations have substantial (15 kt) maximum velocity differences between the clusters. Consequently, the JTWC intensity forecasts, which are typically similar to an original WANI ("down the middle") forecasts, may be also substantially different from the correct cluster WANI forecast. By contrast, the penalty for always picking the wrong cluster is very large (Figure 6a). Already by 24 h, the wrong cluster errors would be 8 kt larger than for the correct cluster errors or the JTWC intensity forecast errors. The wrong cluster forecast errors are already about 38 kt by 72 h, which vividly illustrates the benefit of picking the correct cluster in these 19% of the overall sample of WANI forecasts with a substantial intensity bifurcation.

One contribution to the MAEs among the original WANI, the two cluster WANI, and the JTWC intensity forecasts is the bias error (Figure 6b). In this intensity bifurcation sample, the WANI bias is surprisingly large (-8 kt) in the 84 h – 108 h forecast interval, which means the WANI bias correction for the overall sample does not apply as well for the bifurcation cases and an under-forecast of intensity is the result. The JTWC forecasts have the smallest (-4 kt) bias and the perfect cluster selection WANI forecasts have the next smallest at -6 kt in this forecast interval. The penalty from selecting the wrong cluster WANI forecast is an intensity bias of -16 kt in the 96 h – 108 h forecast intervals. It is likely that this bias is a substantial contribution to the large MAEs in Figure 6a for the sample of wrong cluster WANI forecasts.

Another demonstration of the favourable impact of a perfect selection of the correct cluster in WANI intensity bifurcation situations is the correlation coefficients with the verifying intensities (Figure 6c). In the larger independent sample of WANI forecasts examined by Tsai and Elsberry (2014a, Figure 2b), the correlation coefficients for the WANI forecasts were essentially the same as for the JTWC intensity forecasts through 84 h and then the WANI correlation coefficients remained near 0.7 through 120 h. Somewhat similar coefficients were found here for the original WANI (based on all 10 analogs) and the JTWC forecasts for this smaller sample of intensity bifurcation situations (Figure 6c). If the forecaster could perfectly select the correct cluster WANI forecasts, the correlation coefficient drops off more slowly and levels off at about 0.8 at 72 h and longer forecast intervals. Whereas this always correct cluster WANI intensity forecasts would explain 64% of the variance in the verifying intensities at 120 h, the JTWC forecasts for this bifurcation sample would only explain 34% of the variance. By contrast, correlation coefficients decrease very rapidly if the wrong cluster WANI forecasts are always selected (Figure 6c). Indeed, the correlation coefficient at 24 h is only 0.7 (explained variance of 49%) and is less than 0.5 by 36 h, which is consistent with the rapid increase in intensity MAEs (Figure 6a) if the wrong cluster WANI forecast is always selected.

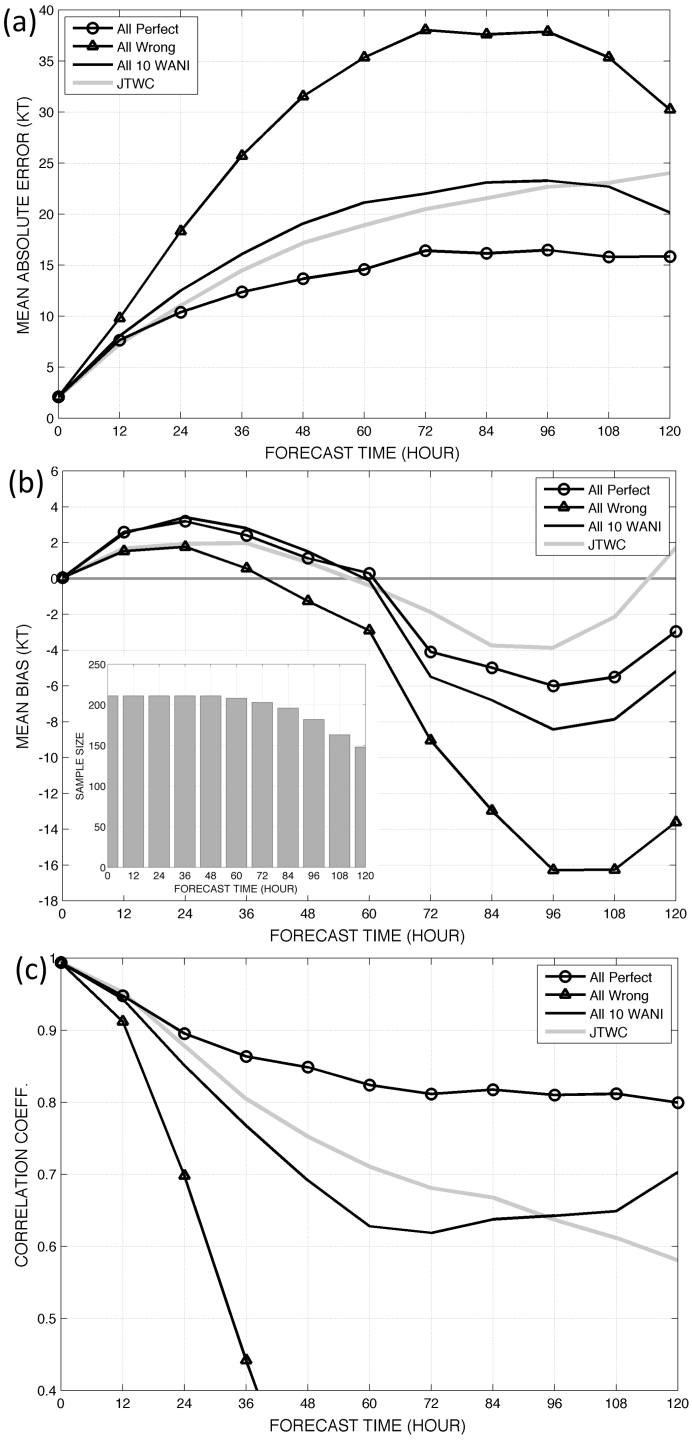

Figure 6. Homogeneous comparison for a sample of substantial WANI intensity bifurcation situations in which either a Perfect choice of the correct cluster WANI forecast or always the Wrong choice of the cluster WANI forecast is made: (a) Intensity mean absolute errors (kt), and (b) Intensity bias (forecast-observed) errors (kt), and (c) Correlation coefficients with verifying intensities as a function of forecast times.

The overall sample mean intensity spreads for the original WANI forecasts based on all 10 of the best-analogs (Figure 7) are consistently 4-5 kt larger for these intensity bifurcation situations than for the larger (ALL) sample of WANI forecasts in Tsai and Elsberry (2014a, Figure 6). The explanation is simply that these bifurcation situations by definition have larger intensity

spreads, and the objective in this study has been to develop two cluster WANI forecasts that represent two separate intensity forecast scenarios. The overall subsample mean intensity spreads about the perfect selection of the cluster WANI forecasts and about the always wrong selections are relatively small compared to the overall sample means of the original WANI forecasts, and are surprisingly similar in magnitude. At the 72 h forecast interval, each of the cluster WANI forecasts has a subsample mean intensity spread of 15 kt, which is one half the overall sample mean intensity spread for the original WANI forecasts. That is, the very large WANI intensity spreads for these intensity bifurcation situations are being split between the two cluster WANI intensity spreads.

These smaller mean intensity spreads for the two clusters (Figure 7) would indicate much higher confidence limits (less uncertainty) in the cluster WANI forecast than for the original WANI forecast, which again were selected for their large intensity spreads in these intensity bifurcation situations. However, Tsai and Elsberry (2014a, Figure 4) demonstrate the intensity spreads for the perfect-selection cluster WANI forecast are somewhat too small (under-determined relative to the desired 68%). Some additional intensity bias correction and calibration of these cluster WANI forecasts is therefore indicated to address this under-determination issue. As indicated above, the always-wrong selection of cluster WANI forecasts have very large intensity MAEs that also contribute to the under-determined intensity spread.

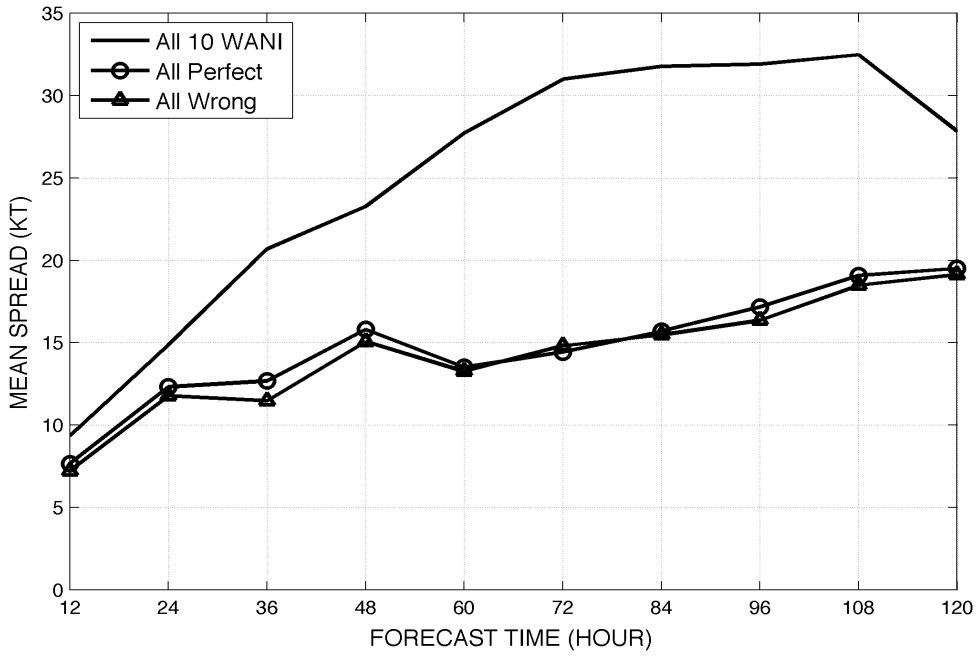

Figure 7. Bifurcation sample mean WANI intensity spreads (kt) for the original WANI forecasts based on all 10 best analogs and subsample mean WANI intensity spreads for the All Perfect selections and All Wrong selections of the cluster WANI forecasts as a function of forecast interval.

a. Example of low latitude formation and rapid intensification case

Another intensity bifurcation situation in Tsai and Elsberry (2014a, Figure 9) was the WANI forecast for TS Bopha (26W) at 0000 UTC 28 November 2012. Because the JTWC official track forecast begins at such a low latitude (~ 5°N) and remains below 10°N (as does the actual track), nearly all of the 10 best historical track analogs are to the north (Figure 8a). Consequently, the corresponding intensity evolutions of the analogs (Figure 8b) are of two types: (i) Smaller initial intensities and non-intensifying; and (ii) Larger initial intensities and intensifying. Due to the WANI ranking procedure that considers the initial intensity, the original WANI forecast based on all 10

analogs is mainly governed by the intensifying analogs. As is common for these intensity bifurcation situations, the original WANI forecast tends to be down the middle between the two intensity evolutions, but is on the low side of these 10 analog intensities due to an almost equal number of non-intensifying analogs (Figure 8b, red circles). This original WANI forecast is similar to the corresponding JTWC intensity forecast (Figure 8b, black line). However, this intensity bifurcation situation with two distinct intensity evolutions leads to an increasing WANI intensity spread with forecast interval that at 108 h ranges from 45 kt to 150 kt (Tsai and Elsberry 2014a, Figure 9). Such a large intensity spread is not a useful representation of the uncertainty of the JTWC intensity forecast that is clearly based on an intensification scenario.

The two cluster WANI intensity and intensity spread forecasts for this TS Bopha case are shown in Figs. 8c and 8d. The cluster C1, which by definition has the larger intensities, contains only the more rapidly intensifying analogs and this results in a forecast of steady intensification to 135 kt at 120 h (Figure 8c). This C1 intensification forecast follows very clearly the intensity evolution of the single-best analog (23W from the 1959 season) that intensified to 120 kt by 108 h (Figure 8c, dark blue line). The cluster C1 WANI intensity spread, which of course is centred on the C1 WANI forecast, increases with forecast interval but only ranges from 95 kt to 160 kt at 108 h (Figure 8c, blue dashed lines) compared to the 45 kt to 150 kt range for the original WANI intensity spread (Tsai and Elsberry 2014a, Figure 9). Although this C1 WANI intensity spread may seem to be too large, the skill in predicting large intensity changes is not very high, so this intensity spread may be an appropriate estimate of the uncertainty for the JTWC intensity forecast. In fact, TS Bopha did rapidly intensify from 55 kt at 48 h to 115 kt at 72 h (Figure 8c, boxes). While the 115 kt intensity slightly exceeded the upper intensity spread value for the original WANI forecast (Tsai and Elsberry 2014a, Figure 9), it is within the intensity spread for the C1 cluster WANI (Figure 8c).

The cluster C2 WANI forecast (Figure 8d) also includes some intensification at longer forecast intervals, which may seem surprising as this cluster includes several analogs that did not intensify (Figure 8b). However, these were also the analogs that had some low initial intensities. When these low initial intensity evolutions have been adjusted to have the actual TS Bopha initial intensity of 50 kt, the cluster C2 WANI intensity spread is rather small. If this was the “perfect” choice between the clusters instead of the wrong choice as described in section 3, this would likely be a case in which the cluster WANI intensity spread would be too small.

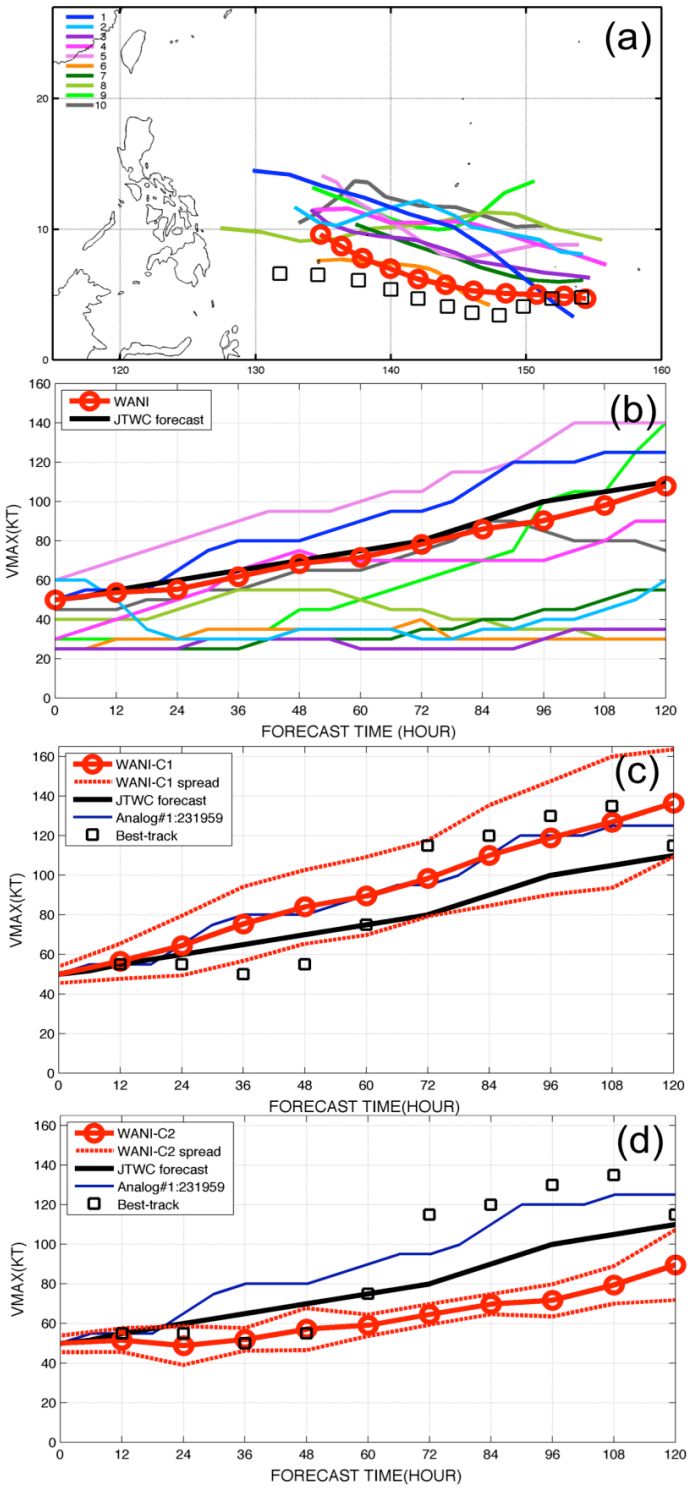

Figure 8. Original and cluster WANI forecasts and intensity spreads for intensity bifurcation situation as in Figure 4, except for TS Bopha (26W) at 0000 UTC 28 November 2012, and the single-best analog is 23W from the 1959 season

b. Example of classic formation and intensification case

The pre-formation WANI forecast from 0000 UTC 3 July 2014 of a tropical disturbance (according to the Regional Specialized Meteorological Center, Tokyo-Typhoon Center) that would later become Supertyphoon (STY) Neoguri (08W according to the JTWC counting) is summarized in Figure 9. As the JTWC forecast track is for a classic path toward a recurvature near 25°-30°N sometime after the end of the five-day forecast, the 10 best track analogs are closely matched to that JTWC forecast track, except a few analog tracks turn to the west near the end of the forecast period (Figure 9a). In contrast to the late-season 26W case in Figure 8b, all of the 10-best analogs in this example begin at the observed initial intensity of 25 kt (according to JTWC). However, a wider range (30 kt to 145 kt) of analog intensities exists at 120 h (Figure 9b), so this is a substantial intensity bifurcation situation. Whereas the original WANI forecast is again in the middle of this range of analog intensities, the JTWC intensity forecast is for consistently larger intensities.

The cluster C1 WANI forecast (Figure 9c) agrees well with the JTWC intensity forecast except the WANI forecast is 20 kt higher at 96 h – perhaps due to the influence of the single-best historical analog (08W in the 1976 season) that had rapidly intensified from 70 kt to 110 kt leading up to the 96 h forecast interval. Due to the rapid intensification of STY Neoguri, the preliminary (working best-track) intensities exceed the cluster C1 WANI intensity spread (Figure 9c). Nevertheless, the C1 WANI forecast would have been a better choice than the original WANI forecast based on all 10 analogs.

The cluster C2 WANI forecast includes two non-intensifying analogs and several moderate intensifying analogs (Figure 9d). Thus, the C2 forecast does not intensify 08W to even the typhoon stage. Furthermore, the C2 WANI intensity spread is much larger for this example than in the TS Bopha example in Figure 8d, which is a reflection of the wider range of intensities for the analogs in the C2 cluster, which all started with the same initial intensity and thus did not need adjustment as in the TS Bopha example.

Because the JTWC forecasters clearly had evidence of a likely intensification (see JTWC forecast in Figs. 9b, 9c, and 9d), the choice of the C1 WANI forecast rather than the C2 WANI forecast would have been obvious in this example. The corresponding C1 WANI intensity spread would obviously have been a better representation of the intensity uncertainty in this intensity bifurcation example than the intensity spread of the original WANI forecast that included all 10 best analogs.

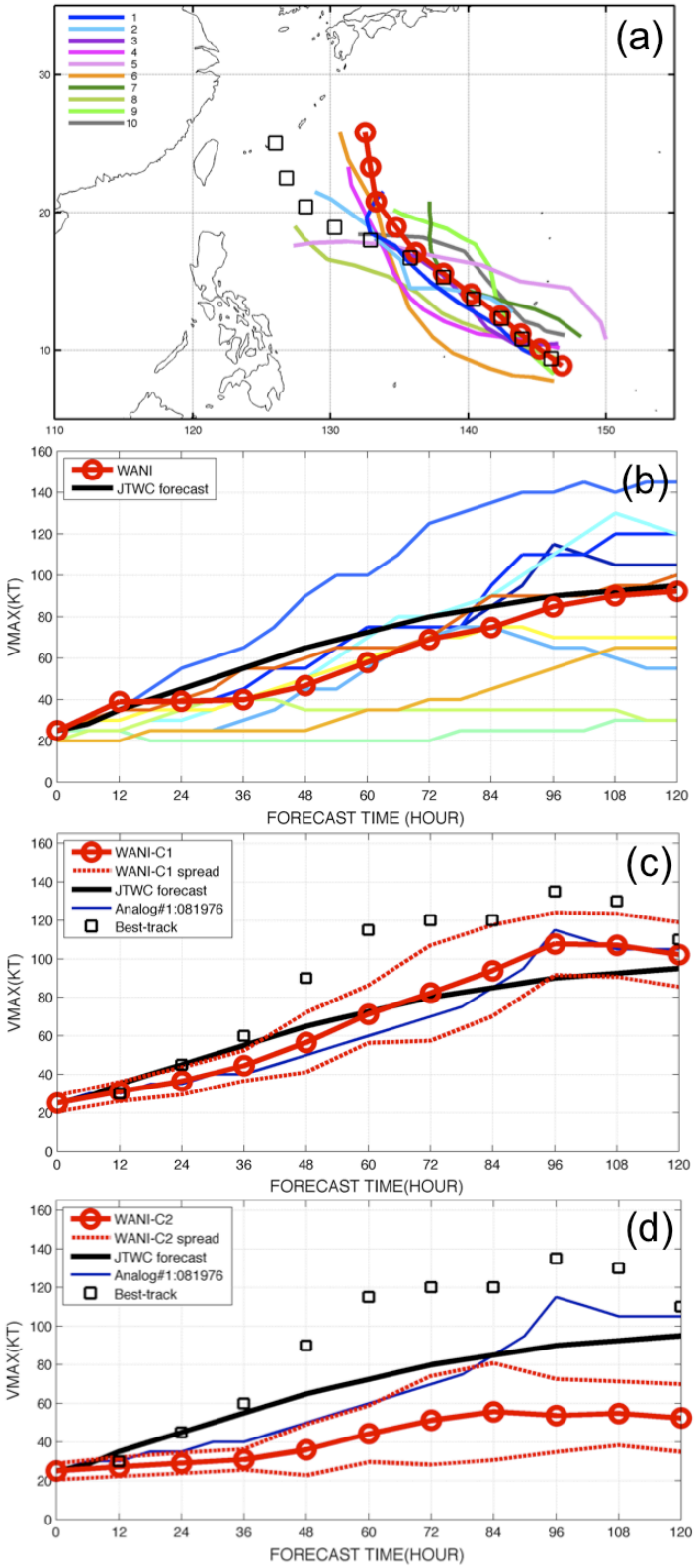

Figure 9. Original and cluster WANI forecasts and intensity spreads for intensity bifurcation situation as in Figure 4, except for pre-Neoguri (08W) at 0000 UTC 3 July 2014, and the single-best analog is 08W from the 1976 season

c. Example of uncertain or delayed intensifications

One source of intensity bifurcations arises from the uncertain timing of when an early stage TC will begin to intensify. An example of the cluster WANI forecasts and intensity spreads for STY Rammasun (09W, according to JTWC) at 0000 UTC 11 July 2014 is shown in Figure 10. At this time, the JTWC track forecast was for STY Rammasun to be making landfall on Luzon Island at t = 120 h (Figure 10a). A relatively tight cluster of the 10 best historical analogs was available, except that seven of the 10 analogs began to the west and southwest of STY Rammasun and had tracks toward the northwest rather than the west-northwest as in the JTWC forecast. The actual track for STY Rammasun was toward the west and it passed over southern Luzon (Figure 10a, boxes).

Because these 10 best analogs predominantly had immediate intensifications, the original WANI forecast (Figure 10b, red circles) also had an immediate intensification with an intensity forecast of 115 kt at 120 h. This forecast was strongly influenced by the single-best analog (07W from the 1986 season, upper-most intensity plot in Figure 10b) that had an intensification to 100 kt by t = 48 h and had an intensity of 140 kt at t = 84 h. Recall that these analog intensities are weighted as in Eq. (2), with the single-best analog intensity being given the highest weighting as it will have had the closest match to the JTWC track forecast.

Notice the JTWC intensity forecast (Figure 10b, black line) is almost identical to the original WANI forecast, which indicates that the JTWC forecasters had concluded that in this situation that intensification would immediately begin and would be sustained throughout the 120 h forecast period. Since landfall on Luzon was being forecast at 120 h, such an intensification would prompt warnings of the threat to Luzon within five days even though Rammasun was only at minimal TS stage at this time. Notice also that a climatological-type intensification rate was being forecast, rather than a more rapid intensification that might have been forecast had the JTWC had available the WANI forecast and there was supporting evidence for a more rapid intensification.

The cluster C1 WANI forecast (Figure 10c, red circles) would provide that alternate, more rapid intensification forecast, which would be 120 kt at 84 h rather than the JTWC forecast of 90 kt. However, STY Rammasun did not immediately begin to intensify (Figure 10b, boxes). Rather, the intensification was delayed for 48 h, but then a rapid intensification began at 60 h with an increase from 40 kt to 115 kt at t = 108 h. While the cluster C1 WANI forecast at 108 h (i.e., just before predicted landfall) would verify within 5 kt, this WANI forecast has intensities between 24 h and 96 h that are badly over-forecast (by 55 kt at 72 h). With this over-forecast and the moderate-size intensity spreads at each forecast interval, all of the verifying intensities between 12 h and 96 h lie outside of the lower bound of the intensity spread (Figure 10c, red dashed lines).

By contrast, the cluster C2 WANI forecast (Figure 10d, red circles) is excellent between 12 h and 60 h with errors of less than 5 kt. However, the cluster C2 forecast does not predict the rapid intensification period, and thus the 108 h forecast intensity of 55 kt is an under-forecast by 60 kt. Although the relatively large intensity spreads for the cluster C2 at 72 h – 84 h are broad enough to include the first 24 h of the rapid intensification, the 96 h – 108 h intensities during the second 24 h of the rapid intensification are far above the upper bounds of the cluster C2 intensity spreads.

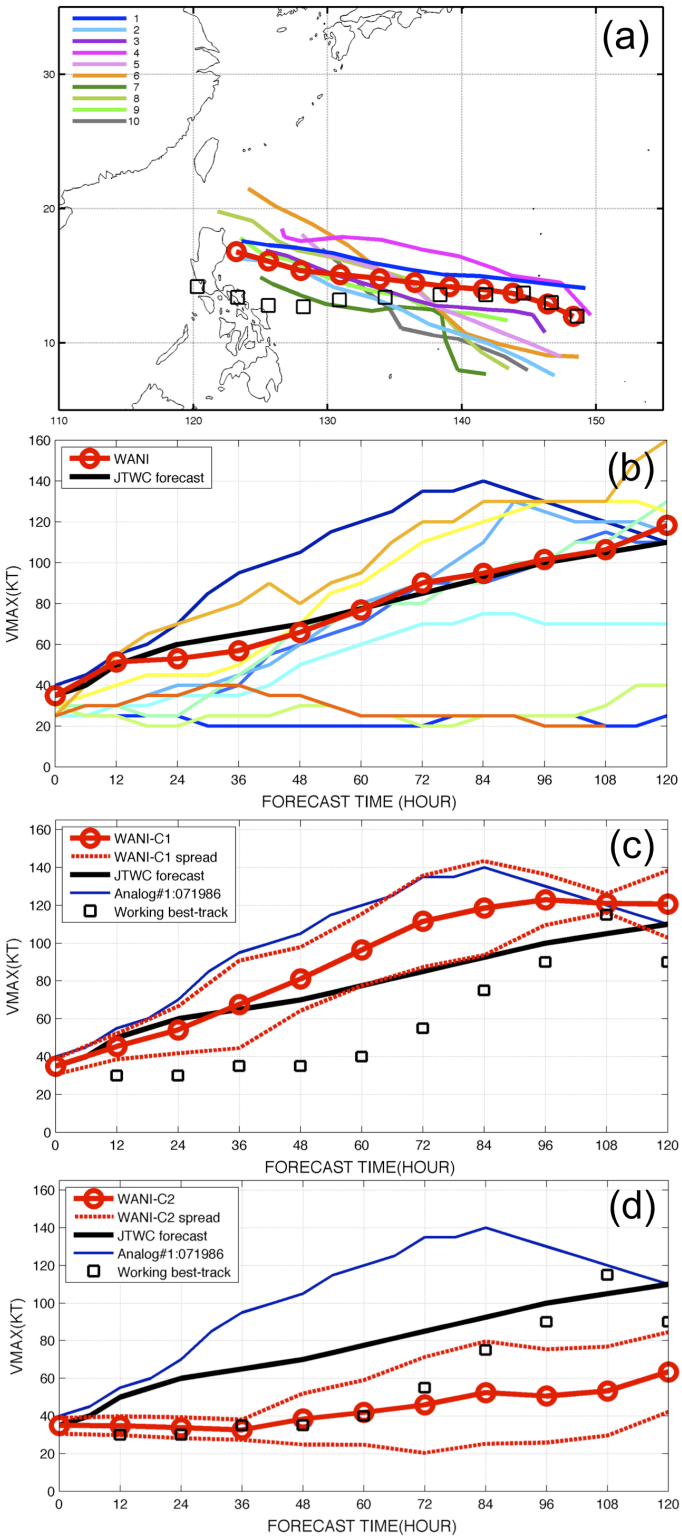

Figure 10. Original and cluster WANI forecasts and intensity spreads for intensity bifurcation situations as in Figure 4, except for STY Rammasun (09W) at 0000 UTC 11 July 2014, and the single-best analog is 07W from the 1986 season

Thus, the intensity forecast from 0000 UTC 11 July 2014 is a difficult challenge. The delayed intensification for the first 48 h was not expected as the JTWC had forecast an intensification by 35 kt during that period. Since the cluster C1 WANI forecast had an intensification of 45 kt, this would be categorized as a wrong forecast. By contrast, the cluster C2 WANI forecast would be categorized as perfect for this early period, but as a wrong forecast after 84 h because it did not account for a rapid intensification of 80 kt from 60 h to 108 h. The cluster C1 forecast also did not include this rapid intensification and only coincidentally had an intensity error of only 5 kt at 108 h just before the predicted landfall time.

d. Example of cluster WANI forecasts for JTWC invests

The JTWC monitors tropical disturbances in the western North Pacific that might become TCs and these "invests" are given an identification number WP*nn* where *nn* ranges from 90-99. A tentative track to five days is produced based on a consensus of global and regional number models (CONW). A preliminary intensity forecast is also produced based on the intensity changes predicted by those models after adjustment to an initial intensity assigned by JTWC (usually 20 kt).

Given that CONW track for an invest, an original WANI forecast has been generated based on the 10 best historical track analogs to provide JTWC forecasters an alternate intensity forecast to the CONW intensity forecast, which typically does not have a significant intensification because it is primarily based on relatively coarse resolution global models. Substantial intensity bifurcation situations frequently appear in these WANI forecasts for the JTWC invests simply due to different forecast intervals at which the 10 best analogs began to intensify.

An example of such an intensity bifurcation for invest WP93 at 1200 UTC 16 July 2014 is given in Figure 11. The CONW track forecast for WP93 results in 10 best analog tracks (Figure 11a) just as for the JTWC forecast tracks in Figure 4 and Figs. 8-10. Setting the initial intensity at 20 kt for WP93 results in the corresponding 10 intensity evolutions (Figure 11b) that also begin at 20 kt because the JTWC historical best track file includes tropical disturbances. Even though all of the analogs begin at 20 kt, a wide range of intensities exists at 120 h. While the original WANI forecast based on the 10 analogs (Figure 11b, red circles) is down the middle of these intensities, the CONW intensity forecast (Figure11b, black line) is for only a slight intensification.

Because the 10 analog intensities in Figure 11b satisfy the criteria for a substantial intensity bifurcation, the two cluster WANI forecasts in Figs. 11c and 11d are produced. The cluster C1 WANI forecast (Figure 11c) is based on those analogs that began to intensify earlier and thus attained the largest maximum intensity. Notice that the single-best analog (02W from the 1992 season) has strongly influenced the selection of analogs for cluster C1, and that the intensity spread for cluster C1 is quite small. The cluster C2 WANI forecast (Figure 11d) is based on those analogs that began to intensify later, and thus did not intensify as much by 120 h. Cluster C2 has a large intensity spread because it includes analogs that persisted as tropical disturbances through most of the five-day forecast interval. Note that the JTWC historical best-track file only includes tropical disturbances that at some time in their life cycle achieved Tropical Depression (TD, 25 kt) status. Consequently, any analog selected from that file will at some time intensify.

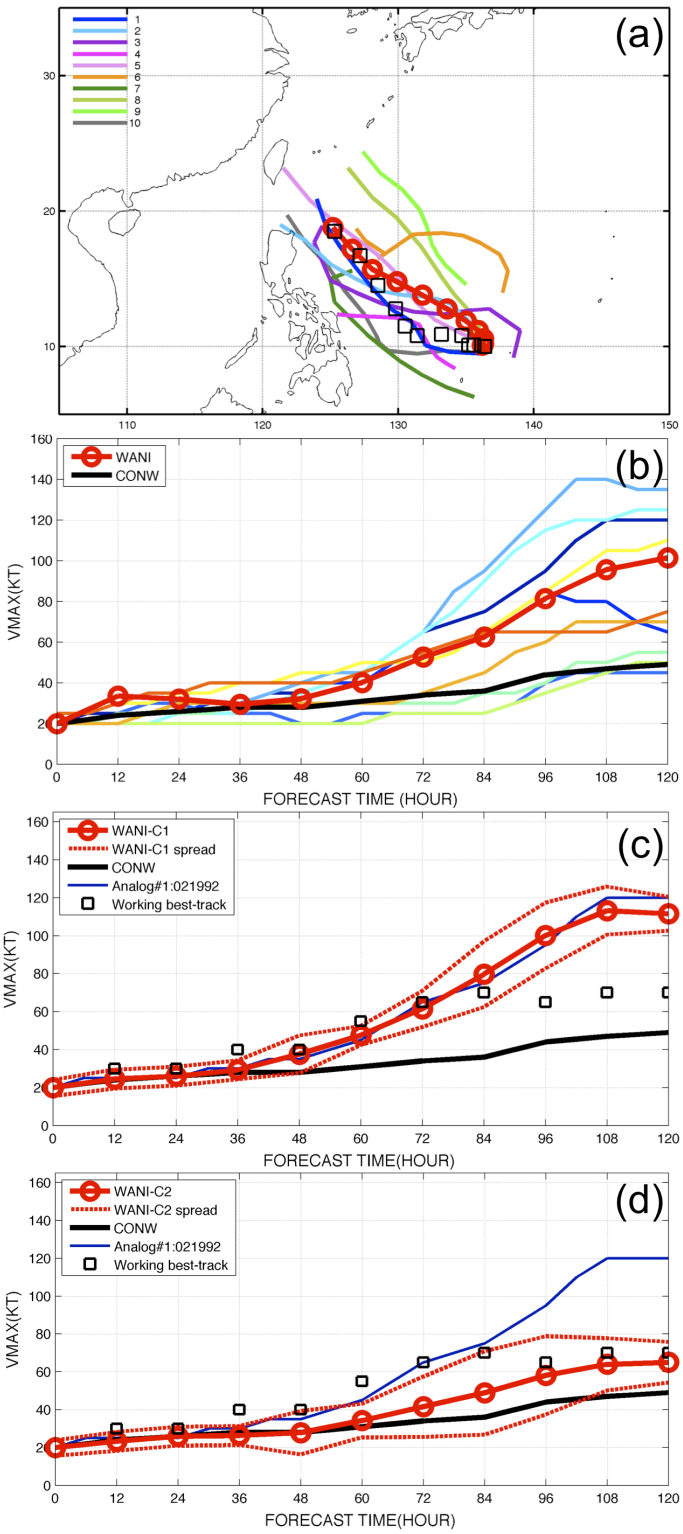

Figure 11. Original and cluster WANI forecasts and intensity spreads for intensity bifurcation situations as in Figure 4, except for a JTWC invest WP93 at 1200 UTC 16 July 2014, and the single-best analog is 02W from the 1992 season

It is proposed that such WANI forecast based on the CONW track forecasts for the JTWC invests may be useful as potential intensity evolutions should that invest develop. It is not being proposed that those WANI forecasts are predicting the actual time of intensification to a TD or TS. Rather, JTWC has an on-going program to utilize global model forecasts and other information to predict the probability of a formation within 48 h. Given a successful formation forecast at forecast time t, the original WANI forecast and/or cluster WANI forecast would provide an early outlook(s) intensity evolution for that invest.

SF 1b.4 Summary and discussion

Goerss and Sampson (2014) have developed a regression technique to predict the S5YY consensus model intensity forecast error. The predictors for this consensus intensity forecast error had to be quantities that are available prior to the official forecast deadline. Leading predictors were found to forecast the TC intensity and intensity change, initial intensity and latitude of the TC, and consensus model spread, which is defined to be the average of the absolute intensity differences between the member forecasts and the consensus forecast. Based on the regression model, intensity intervals were determined centred on the S5YY forecasts that contained the verifying TC intensity about 67% of the time. Based on the size of these intervals, a forecaster can determine the confidence that can be placed on the S5YY forecasts. Independent data testing yielded results only slightly degraded from those of dependent data testing, highlighting the capability of these methods in practical forecasting applications.

Tsai and Elsberry (2014a) developed and tested a new situation-dependent intensity prediction technique called WANI for operational intensity forecasts of western North Pacific TCs. The WANI was demonstrated to be almost as skillful as the JTWC intensity forecasts through 84 h and more skill at 96 h – 120 h. Tsai and Elsberry (2014a) also developed intensity spread guidance for the weighted analog intensity predictions in the WANI, and provided examples of large intensity spreads that clearly indicated bimodal distributions that Tsai and Elsberry referred to as intensity bifurcation situations.

Tsai and Elsberry (2014b) have developed an objective procedure for detecting these intensity bifurcation situations in a sample of 1136 WANI forecasts. Approximately 19% of the sample had conditions that met the thresholds for a substantial intensity bifurcation. A hierarchical clustering technique was applied to determine two clusters among the 10 best intensity analogs, and the WANI forecasts for these two clusters are provided as alternate intensity evolutions. If the forecaster could always pick the correct cluster WANI forecast, a large improvement in intensity forecasts would result relative to the original WANI forecast, which tends to be in the middle of all 10 analog intensity evolutions for these intensity bifurcation situations. Separate WANI intensity spreads that are calculated for the two WANI forecasts then tend to divide the large intensity spreads of the original WANI forecast based on all 10 analogs into two parts. However, the cluster WANI intensity spreads tend to be under-determined even with a perfect selection of the correct cluster WANI forecast.

The operational usefulness of these cluster forecasts in substantial bifurcation situations does not assume that the forecaster will rely only on the information provided by the WANI technique. Rather, these cluster WANI forecasts are intended to alert the JTWC forecaster to examine alternate intensity solutions based on all the other guidance available. In most cases, the JTWC forecasters will already be considering formation versus non-formation, the possibility of a rapid intensification, or the consequences of a landfall or non-landfall that might lead to an intensity bifurcation situation. A correct choice of the intensity forecast scenario in these bifurcation situations with the assistance of the cluster WANI technique will be the basis for improved warnings of the threat from western North Pacific TCs.

In the Special Focus session SF1b. at ITWC-VIII, a case study of intensity forecasts for the western North Pacific (WP11 according to JTWC) storm will be presented. This case was selected because it was a very difficult intensity forecast scenario with a very rapid intensification that was followed by a rapid decay over the open ocean, but then WP11 was maintained at typhoon intensity until landfall. This case study will illustrate that rapid intensification and rapid decay events will likely fall outside the WANI intensity spread guidance or the Goerss and Sampson half-width confidence intervals. The implication may be that these extreme intensity changes are not forecastable because their timing and magnitude depend on convective-scale and mesoscale processes that are not observable. However, just the documentation of when such events exceeded the expected forecast uncertainty may help the forecaster recognize future events in real-time, and also serve as case studies for researchers searching for better understanding of large intensity changes.

Recommendations

That successful guidance products to specify TC intensity uncertainty be shared, and that TC warning centres provide TC intensity forecast uncertainty as part of their warnings by 2018.

Acknowledgments

Dr H.-C. Tsai is a National Research Council post-doc at the Naval Postgraduate School. He and Professor R. L. Elsberry are supported by the Office of Naval Research Marine Meteorology section. Mrs Penny Jones provided excellent assistance in the manuscript preparation.

References

DeMaria, M., M. Mainelli, L.K. Shay, J.A. Knaff, J. Kaplan, 2005: Further improvement to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. and Forecasting,* **20**, 531–543.

DeMaria, M., J .A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria, 2009: A new method for estimating tropical cyclone wind speed probabilities. *Wea. Forecasting,* **24**, 1573-1591.

DeMaria, M.,C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc*., **95**, 387-398.

Doyle, J.D., Y. Jin, R. Hodur, S. Chen. H. Jin, J. Moskaitis, A. Reinecke, P. Black, J. Cummings, E. Hendricks, T. Holt, C. Liou, M. Peng, C. Reynolds, K. Sashegyi, J. Schmidt, S. Wang, 2012: Real time tropical cyclone prediction using COAMPS-TC. *Advances in Geosciences,* **28**, Atmospheric Science and Ocean Sciences (2012), Eds. Chun-Chieh Wu and Jianping Gan, 15-28.

Elsberry, R. L., and H.-C. Tsai, 2014: Situation-dependent intensity skill metric and intensity spread guidance for western North Pacific tropical cyclones. *Asia-Pacific J. Atmos. Sci*., **50**, 297-306.

Emanuel, K., C. DesAutels, C. Holloway, and R. Korty, 2004: Environmental control of tropical cyclone intensity**.** *J. Atmos. Sci***.,** **61**, 843–858.

Goerss, J. S., 2000: Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Wea. Rev.*, **128**, 1187-1193.

Goerss, J. S., 2007: Prediction of consensus tropical cyclone track forecast error. *Mon. Wea. Rev.*, **135**, 1985-1993.

Goerss, J. S., and C. R. Sampson, 2014: Predictions of consensus tropical cyclone intensity forecast error. *Wea. Forecasting*, **29**, 750-762.

Goerss, J. S., C. R. Sampson, and J. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting*, **19**, 633-638.

Hogan, T. F., and R. L. Pauley, 2007: The impact of convective momentum transport on tropical cyclone track forecasts using the Emanuel cumulus parameterization. *Mon. Wea. Rev.*, **135**, 1195–1207.

Knaff, J., C. R. Sampson, and M. DeMaria, 2005: An operational statistical typhoon intensity prediction scheme for the western North Pacific. *Wea. Forecasting,* **20,** 688-688.

Rennick, M. A., 1999: Performance of the Navy's tropical cyclone prediction model in the western North Pacific basin during 1996. *Wea. Forecasting*, **14**, 3-14.

Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (Version 3.2). *Bull. Amer. Meteor. Soc.,* **81**, 1231-1240.

Sampson, C. R., J. L. Franklin, J. A. Knaff, and M. DeMaria, 2008: Experiments with a simple tropical cyclone intensity consensus. *Wea. Forecasting*, **23**, 304-312.

Tsai, H.-C., and R. L. Elsberry, 2014a: Applications of situation-dependent intensity and intensity spread predictions based on a weighted analog technique. *Asia-Pacific J. Atmos. Sci.*, DOI: 10.1007/s13143-014-0040-7.

Tsai, H.-C., and R. L. Elsberry, 2014b: Improved tropical cyclone intensity and intensity spread prediction in bifurcation situations. *Asia-Pacific J. Atmos. Sci*., (submitted).

Weber, H. C., 2001: Hurricane track prediction with a new barotropic model. *Mon. Wea. Rev.*, **129**, 1834–1858.

Weber, H. C., 2005: Probabilistic prediction of tropical cyclones. Part II: Intensity. *Mon. Wea. Rev***.**, **133**, 1853–1864.

Wilks, D. S., 2011: Statistical Methods in the Atmospheric Sciences. 3rd Edition, Academic Press, Amsterdam, 676 pp.