Original

Barcikowska, M.; Feser, F.; Storch, H.v.:

Usability of best track data in climate statistics in the western North Pacific

In: Monthly Weather Review (2012) AMS

DOI: 10.1175/MWR-D-11-00175.1

Usability of Best Track Data in Climate Statistics in the Western North Pacific

MONIKA BARCIKOWSKA, FRAUKE FESER, AND HANS VON STORCH

Helmholtz-Zentrum Geesthacht Centre for Materials and Coastal Research, Geesthacht, and Cluster of Excellence, Integrated Climate System Analysis and Prediction, University of Hamburg, Hamburg, Germany

(Manuscript received 20 July 2011, in final form 26 February 2012)

ABSTRACT

Tropical cyclone (TC) activity for the last three decades shows strong discrepancies, deduced from different best track datasets (BTD) for the western North Pacific (WNP). This study analyzes the reliability of BTDs in deriving climate statistics for the WNP. Therefore, TC lifetime, operational parameters [current intensity (CI) number], and tracks are compared (for TCs identified concurrently) in BTD provided by the Joint Typhoon Warning Center (JTWC), the Japan Meteorological Agency (JMA), and the China Meteorological Administration (CMA).

The differences between the BTD are caused by varying algorithms used in weather services to estimate TC intensity. Available methods for minimizing these discrepancies are not sufficient. Only if intensity categories 2–5 are considered as a whole, do trends for annually accumulated TC days show a similar behavior. The reasons for remaining discrepancies point to extensive and not regular usage of supplementary sources in JTWC. These are added to improve the accuracy of TC intensity and center position estimates. Track and CI differences among BTDs coincide with a strong increase in the number of intense TC days in JTWC. These differences are very strong in the period of intensive improvement of spatiotemporal satellite coverage (1987–99).

Scatterometer-based data used as a reference show that for the tropical storm phase JMA provides more reliable TC intensities than JTWC. Comparisons with aircraft observations indicate that not only homogeneity, but also a harmonization and refinement of operational rules controlling intensity estimations, should be implemented in all agencies providing BTD.

1. Introduction

In recent years, tropical cyclone (TC) activity, which poses a risk for coastal populations, gained much attention in the environmental research community (Emanuel 2005; Landsea 2005; Webster et al. 2005; Wu et al. 2006). The long-term variability in TC activity became a subject of interest in atmospheric science pointing to changes in atmospheric rotational flow, vertical wind shear, or sea surface temperature (SST) over the last decades (Chan and Liu 2004; Trenberth 2005). Using tropical cyclone best track datasets (BTD), Webster et al. (2005) and Emanuel (2005) claimed there would be an increase in the occurrence of the most intense TCs in the western North Pacific (WNP). However, according to Wu et al. (2006), who used several BTDs provided by different institutes, neither the numbers of the most intense TCs nor the power dissipation index (PDI) defined by Emanuel (2005) show an increasing tendency.

Comparing three BTDs, Ren et al. (2011) confirmed increasing TC tendencies for the Joint Typhoon Warning Center (JTWC) BTD, but they found decreasing tendencies in the data of the Japan Meteorological Agency (JMA) and the China Meteorological Administration (CMA). Kamahori et al. (2006) found increasing numbers of TC days for categories 2 to 3 of the Saffir–Simpson Hurricane Scale (SSHS; Simpson 1974) and decreasing numbers in higher categories for JMA, while opposite trends were detected for the JTWC dataset. All these studies indicate a great dependency of the detected TC trends on the chosen BTD, pointing to data inhomogeneity and quality deficiencies in the WNP region.

Knaff and Sampson (2006) considered any detected intensity trend questionable before reanalyses employing datasets of TC intensity estimated with alternative

Corresponding author address: Monika Barcikowska, Helmholtz-Zentrum Geesthacht Centre for Materials and Coastal Research, Max-Planck-Str. 1, 21502 Geesthacht, Germany.
E-mail: monika.barcikowska@hzg.de

DOI: 10.1175/MWR-D-11-00175.1

techniques are incorporated. Others attempted to identify the reasons for the differences between BTD that affect TC activity trends (Kamahori et al. 2006; Nakazawa and Hoshino 2009; Song et al. 2010). Many studies highlighted the different operational procedures used by the individual meteorological agencies to estimate TC intensity as a main cause for differing TC activity results. Knapp and Kruk, (2010) attempted to minimize discrepancies among BTD by applying unified algorithms to operational data from all centers, resulting in more comparable BTDs.

In this paper we assess the reliability of BTD in deriving TC activity trends. In the first part of the study we evaluate the skill of current solutions for achieving homogeneity between the individual datasets. In the second part, the remaining discrepancies between BTDs are analyzed and evaluated using independent reference datasets. All data and methods used are described in section 2. Results and discussion of comparisons are presented in section 3. Section 4 summarizes and concludes the article.

2. Data and methods

Four different BTDs were analyzed in this study. They were provided by the following independent agencies: the CMA (http://www.typhoon.gov.cn), the Regional Specialized Meteorological Center (RSMC), Tokyo, Japan, of the JMA (http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/besttrack.html), and the JTWC (www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/wpindex.html). In addition, the International Best Track Archive for Climate Stewardship (IBTrACS; http://www.ncdc.noaa.gov/oa/ibtracs/index.php?name=ibtracs-data) was used. This product combines BTD from different operational centers to create a global best track dataset (Knapp et al. 2010). Although IBTrACS cannot serve as independent data, it provides useful information as it gives a merged BTD solution for which a data quality control was applied. BTDs for the WNP contain TC center, maximum sustained wind, and central pressure at 6-h intervals. JTWC and CMA intensity values start with tropical depression (TD) strength, and JMA starts with the tropical storm (TS) category.

From 1977 JMA began recording maximum sustained wind speeds using the Dvorak technique (DT; Dvorak 1972, 1973, 1975). Since 1987, when aircraft reconnaissance flights ended in the WNP, this method became the main tool for compiling BTDs. The technique estimates TC position and intensity using visible and infrared imageries from geostationary and polar-orbiting weather satellites. However, procedural rules to process the data for BTD within meteorological agencies were evolving differently. Dvorak parameters [T number and current intensity (CI) number], estimated operationally on basis of identified cloud patterns, are related to TC intensity through conversions, which were independently established for differing wind speed definitions in each operational center. While the JTWC uses 1-min mean sustained 10-m wind speed, as designed originally by the Dvorak technique, other agencies use 10-min-averaged values. JMA established a new conversion table in 1990 (Koba et al. 1991) that transfers operational parameters (CI) directly to TC intensity described as 10-min maximum sustained wind speed.

The CMA dataset specifies intensity in terms of "2-min mean maximum sustained wind speed (m s^{-1}) near the storm center." However, this procedure contradicts the description in Yu et al. (2007), which states that the CMA agency uses an empirically established linear relationship between 1- and 10-min-averaged values and multiplies wind values by a factor of 0.871. The assumed application of a 10-min-average definition in the CMA dataset is supported by findings of relatively small differences among JMA and CMA (Knapp and Kruk 2010). IBTrACS data use 10-min sustained wind speed.

To evaluate the BTD additional observational datasets were tested for their ability to serve as a reference. Blended Sea Winds provided by the National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (NCDC; http://www.ncdc.noaa.gov/oa/rsad/air-sea/seawinds.html) contain ocean surface wind speed on a global 0.25° grid in 6-hourly time steps (Zhang et al. 2006a,b). The data are created by blending observations from multiple satellites with a simple spatiotemporally weighted interpolation. The quality of the blended product is related to the accuracy of the input data and sampling scheme of the observations. The number of long-term U.S. satellites providing wind observations increased from one in 1987 to five in 2000. In this study, years 2000 to 2008 were analyzed as they constitute a rather homogeneous temporal and spatial coverage. For this period wind observations are retrieved from the Quick Scatterometer (QuikSCAT), the Defense Meteorological Satellite Program (DMSP) Special Sensor Microwave Imager (SSM/I), the Advanced Microwave Scanning Radiometer (AMSR-E) of the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS), and the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI). Scatterometers measure instantaneous ocean surface wind vectors at 10-m height with a grid-typical resolution of 25 km and are widely used in operationally prepared analyses and forecasts

(Bourassa et al. 2010; Brennan et al. 2009; Hoffman and Leidner 2005). They are intended to provide accurate ocean surface winds in all weather conditions except for rain conditions that occur often during high winds. QuikSCAT data, evaluated against buoys, is adhered to an 8-min average. QuikSCAT winds were shown (Brennan et al. 2009) to have high skills in intensity estimation for tropical storms strength. However, enhanced backscattering by rain may introduce a positive bias during tropical depressions and rain attenuation causes large negative biases for very high winds. Microwave observations flagged as contaminated by precipitation were excluded from the analysis.

As reference data for the TCs of the strongest intensity, aircraft measurements were used. For the analyzed period 2000–08 The Observing System Research and Predictability Experiment (THORPEX) Pacific Asian Regional Campaign (T-PARC-2008) aircraft campaign took place in the WNP, which provided measurements of wind speed during TC events. Observations were obtained from Stepped-Frequency Microwave Radiometer (SFMR). Additionally we used the measurements from a field experiment in 2010: Impacts of Typhoons on the Ocean in the Pacific (ITOP-2010). (The databases for both campaigns are available online http://www.aoml.noaa.gov/hrd/data_sub/hurr.html.)

a. Quantifying TC trend differences derived from BTDs

The TC trends for the period 1977–2008 were derived from several BTDs and compared in the form of annual number of TC days categorized by the SSHS scale. The analysis is constrained to TC observations recorded concurrently in all independent BTDs. This excludes contributions of differing TC frequency among BTDs to trend discrepancies and enables the identification of the reasons for differences in estimated intensity.

Discrepancies among trends derived from 1-min (JTWC) and 10-min (JMA) sustained wind speed are discussed with regard to the impact of intensity definition on the derived climate statistics. The accuracy and effectiveness of two methods unifying wind definitions is assessed with respect to minimizing trend discrepancies.

The methods adjusting TC intensity definitions from 10–1-min-averaging period were applied to JMA and CMA. The first method is based on the statistical, linear relationship between 10- and 1-min-averaged intensity (Atkinson 1974). The data from JMA and CMA (for CMA a 10-min average is assumed as stated in the previous section) multiplied by a factor of 1.14 are hereafter referred to as JMA*1.14, and CMA*1.14.

Knapp and Kruk (2010), Song et al. (2010), and Wu et al. (2006) highlighted the problem of different algorithms used among the various BTD to convert CI parameters (derived from satellite imageries) to wind speed. An alternative method was proposed by Knapp and Kruk (2010) and Kruk et al. (2011), reversing intensity values back to operational parameters (CI) and then applying a single conversion table to all datasets resulting in more homogenous intensity values. Following these guidelines, the JMA dataset was reverted to CI numbers, using the conversion tables described in Koba et al. (1991). In a second step, we derive wind speed from CI numbers by applying the original Dvorak conversion table (Dvorak 1984) used in JTWC (hereafter DT conversion). It is possible that the Koba conversion table was applied only to intensity records starting in 1991 and previous years were not updated to the new procedures (Nakazawa and Hoshino 2009). However, the remapping method using the Koba conversion table was applied for the complete analysis. Consequently, years before 1987 should be analyzed with extreme caution and have only minor impact on the conclusions derived in this article.

The remaining reasons for BTD trend discrepancies are examined by comparing datasets with the same wind speed definition (JTWC and JMA/CMA adjusted to 1-min-averaging period). The statistical analysis additionally includes yearly mean differences for TC center locations, annual distributions of differences between BTDs for CI-numbers, and TC center locations. The difference in TC location is estimated by a measure of distance (ΔP) between two geographical points (x_1, y_1) and (x_2, y_2) on the earth's surface:

$$\Delta P = r_0 \cos^{-1}\{\sin(y_1)\sin(y_2) + \cos(y_1)\cos(y_2)\cos(x_1 - x_2)\}, \qquad (1)$$

where x and y are longitude and latitude and r_0 is the radius of the earth.

b. BTD-reference data comparison methods

Independent reference data were employed to evaluate the remaining discrepancies between BTDs. Because of a positive bias which occurs in QuikSCAT data for tropical depressions (Hoffman and Leidner 2005) and frequently changing procedures in operational centers to identify this phase, the analysis focuses on concurrent records in BTDs during tropical storm stage. As JTWC and JMA provide information about conversion tables in use, we use the JMA dataset remapped to 1-min-averaged wind speed using the DT table (as described in the previous section). Concurrent

TC observations in BTD were compared with the NOAA wind data for the period 2000–08, when QuikSCAT had a large impact. To derive maximum TC wind speeds from NOAA, the center positions given by JMA were used. The TC circulation in developed systems vanishes at a finite horizontal radius with an upper boundary of approximately 1000 km (Dean et al. 2009). For small, developing or already dissipated cyclonic systems, it was assumed that the maximum wind speed is within a 500-km radius around a given location. Maximum intensities between the two datasets were compared for all concurrent TC cases.

As microwave signal is vulnerable to heavy rain conditions, the NOAA data exclude such values of reduced accuracy. Therefore time steps with a number of missing values around a TC center potentially high enough to mask a region of maximum wind speeds were also excluded from the comparison.

For the comparison of the highest-intensity typhoons the SFMR observations were used. Observations were obtained during several flights targeting TC centers of Typhoons Sinlaku (2008), Jangmi (2008), and Megi (2010). SFMR measures wind speed values in 1-s intervals. To use these wind speeds compatible with BTD, the values were used in two forms: averaged over a 10-s and 1-min intervals. Similar to the previous method, the value of the maximum wind speed was derived by choosing the highest value within a certain radius from the TC center given by JMA.

3. Results and discussion

a. Are the current methods able to minimize discrepancies among TC activity trends in BTDs?

The damage potential posed by TCs depends on their frequency and duration. To evaluate the skill of the methods in reducing the differences between TC activity trends, we examined the annually integrated TC lifetime for original and modified BTDs. The analysis was conducted for the period from 1977 to 2008 and only for concurrent observations. Therefore the total TC-day number is the same for every dataset and differs only in the number of records falling into individual intensity categories. The values in the categories of the highest winds are the most significant for socioeconomic consideration. Therefore, we focused on categories 2–5 and TS separately.

Figure 1a presents annually accumulated records of TC days for categories 2–5. Original datasets are IBTrACS, JMA, and JTWC (reporting 1-min sustained wind speed). Datasets adjusted to a 1-min-averaging period are JMA*1.14 and CMA*1.14 (which result from applying a multiplication factor to JMA and CMA) and JMADT (where a remapping method—using the original Dvorak (1984) conversion table—was applied to JMA CI- numbers). JMA, IBTrACS, and CMA (not shown) show very similar TC-day numbers, with slightly higher numbers for CMA in the first years of the analysis. It was already demonstrated in previous studies (Knapp and Kruk 2010; Song et al. 2010; Ren et al. 2011) that JTWC and JMA wind speed values show the largest discrepancies among the original datasets. The application of methods to unify the wind averaging period significantly reduced these differences. The average of annual relative differences for the considered period exceeds 0.77 for JMA, and 0.57 for CMA in relation to JTWC. Multiplying JMA and CMA data by 1.14 (JMA*1.14 and CMA*1.14) results in much smaller values, 0.19 and 0.22, respectively. Consequently, JMA*1.14 showed a stronger increasing tendency, similar to JTWC. Recalculating JMA TC intensities from CI parameters with the original Dvorak conversion table (JMADT) also reduced the differences and increased the TC activity trend from 0.18 to 0.45, while JTWC shows the highest trend of 0.65.

To analyze the effectiveness of the methods for different TC intensity categories we analyzed the trends for categories 2–3 and 4–5 separately (Figs. 1b,c). The JMA

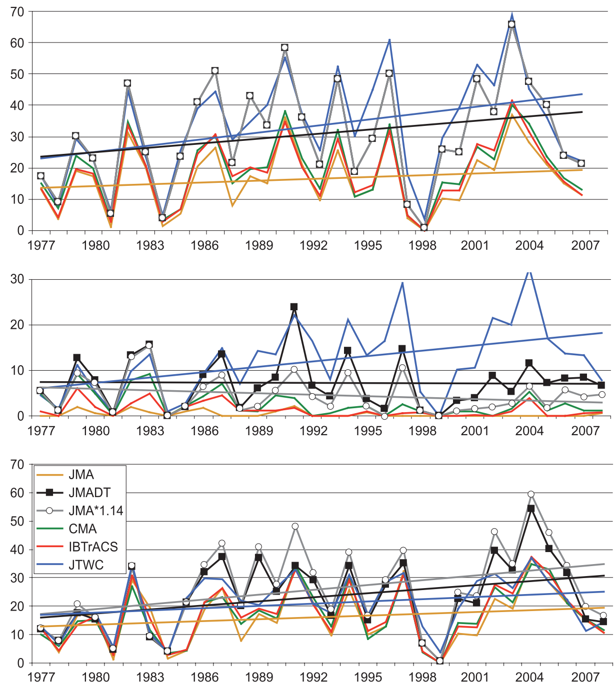

FIG. 1. TC-day numbers for SSHS intensity categories: (top) 2–5, (middle) 4–5, and (bottom) 2–3 for original BTDs: JMA, CMA, JTWC, IBTrACS and modified BTDs: JMA*1.14 (JMA multiplied by a factor), JMADT (JMA using the Dvorak conversion table).

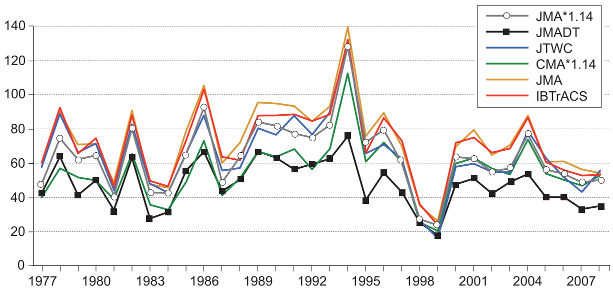

FIG. 2. TC-day numbers in the TS intensity category for original BTDs: JMA, JTWC, IBTrACS, and modified BTDs: JMA*1.14, CMA*1.14 (JMA and CMA multiplied by a factor), and JMADT (JMA using the Dvorak conversion table).

TC-days trend for categories 2–3 is high (0.22) and increases to 0.56 when using the multiplicative factor (JMA*1.14). This is twice as high as the JTWC trend. The method has less effect in categories 4–5. Trends in modified datasets (JMA*1.14 and CMA*1.14) still retain the decreasing character of 10-min wind speed BTD (JMA, CMA, and IBTrACS). In contrast, 1-min wind speed BTD (JTWC) shows upward trends. The results for CMA*1.14 are almost identical to JMA*1.14, which suggests that 10-min-averaged wind speed values were used in CMA (see chapter 2). The results indicate that usage of the multiplicative factor increases intensity values sufficiently to upgrade them to categories 2–3. However, it is still too small for upgrading values to categories 4–5, and therefore leads to accumulated TC records in the lower range.

Applying the original Dvorak conversion table to JMA leads to higher numbers of TC days in both categories 2–3 and 4–5. It increases the trend for categories 2–3 to 0.46, which is already lower than JMA*1.14 (0.56), but still significantly higher than in JTWC. This method upgrades intensity values to categories 4–5 and partially reduces the differences in TC-day numbers in comparison to JTWC. While JTWC features an increasing trend of 0.39 (Fig. 1c), JMA*1.14 presents the strongest decrease and JMADT shows no trend.

The interagency differences also change in time for the lower wind categories. Figure 2 presents annually accumulated TC-day records for BTDs for the TS category, where the highest differences occur for the middle period of the analysis (1987–98). JMADT shows systematically lower numbers of TC-day records than JMA, as applying the Dvorak conversion degrades over 30% of all records from the TS to the TD category. In contrast, application of the multiplication factor upgrades values to higher categories. Therefore, both methods result in smaller TC-day numbers for the TS category. Nevertheless the TC activity tendencies of the analyzed records are in good agreement showing a slight increase until the mid-1990s and a decrease for the last decade.

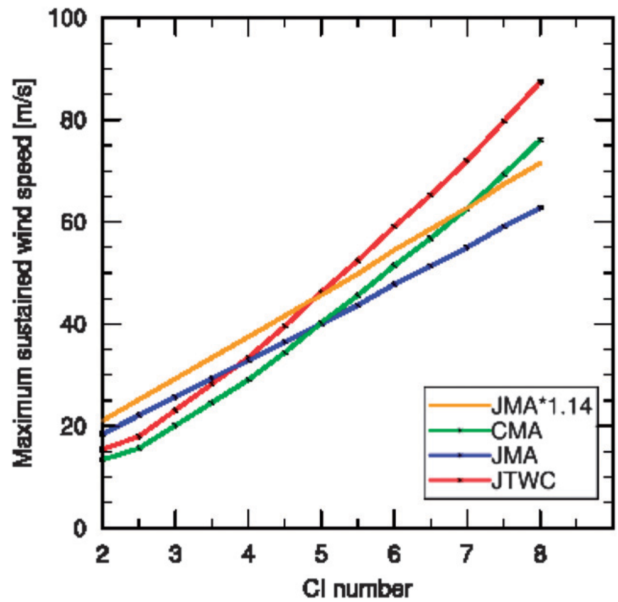

FIG. 3. Relationship of CI parameter to maximum sustained 10-m wind speed (m s^{-1}) using conversion tables used in JTWC (Dvorak), JMA (Koba), JMA*1.14 (Koba multiplied by a factor of 1.14), and CMA (Dvorak multiplied by a factor of 0.87).

b. Impact of unification of conversion tables in BTD on climate statistics

Knapp and Kruk (2010) found discrepancies among BTD intensity records to be highly linear, and demonstrated that they can be minimized using a remapping method. Our analysis shows that both methods have the skill to reduce discrepancies among TC activity trends for categories 2–5 and lead to increasing TC trends. Figure 3 presents the functions for converting CI parameters to TC intensity that are used in operational centers in the WNP region. It is visible that for wind speed of category 1 and higher, both conversions—the Dvorak table used in JTWC (DT) and the linear factor (JMA*1.14)—provide higher wind speed values for the same CI parameter than the Koba conversion. Therefore, the application of such methods reduces the differences in comparison to JTWC due to increasing wind values and due to shifting more low-category TC records toward categories 2–5.

When categories 2–3 and 4–5 are regarded separately, the application of the linear relationship has obvious drawbacks, as it introduces high uncertainties to TC trends. The multiplication factor enhances wind speed values linearly, for the whole dataset distribution. However, the nonlinear sensitivity of wind speed to the

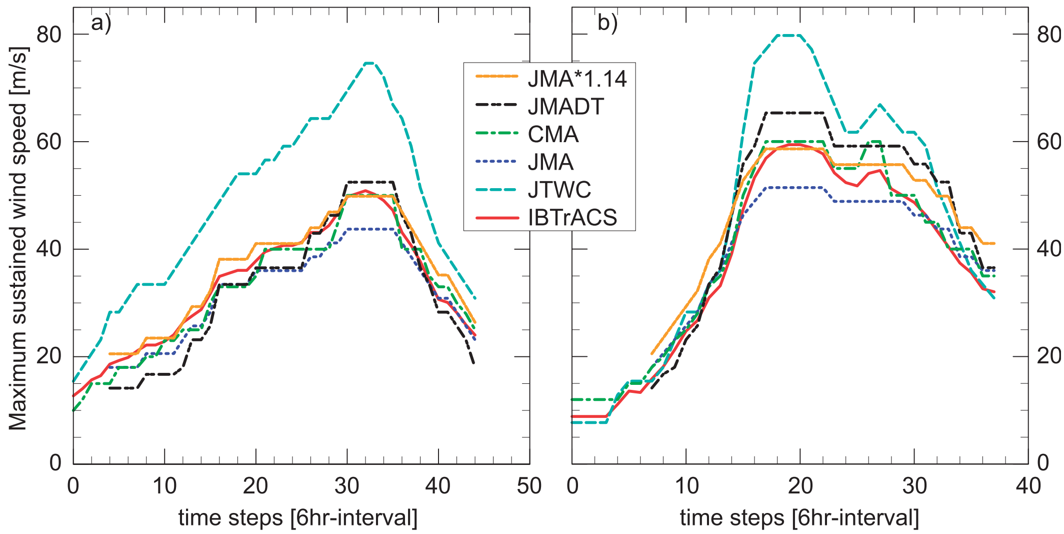

FIG. 4. Wind speed time series for two TC events: (a) Isa (1997), (b) Dianmu (2004) for original BTDs: JMA, CMA, JTWC, IBTrACS, and modified BTDs: JMA*1.14 (JMA multiplied by a factor) and JMADT (JMA using the Dvorak conversion table).

averaging period, which makes Atkinson's (1974) linear relation less accurate, creates the risk of overestimating values in the lower-intensity categories (2–3) and underestimating the highest ones. Kamahori et al. (2006) confirmed our findings, showing high discrepancies in trend tendencies between JTWC and linearly modified JMA, but this comparison included all identified TCs in both datasets and not only the concurrent ones. They also found a strong increase in JMA TC days for categories 2–3 and a decrease for categories 4–5, while JTWC showed opposite tendencies.

Applying the original Dvorak conversion considers the nonlinear effects of the averaging time interval. The remapping method using the DT conversion reduces the trend discrepancies between JMA and JTWC more efficiently for categories 2–3 and has higher skill in the extreme wind range by upgrading more records to categories 4 and 5. However, the trends in category 4–5 still differ.

Song et al. (2010) suggested that the main reasons for differences in BTD intensity over the WNP are different conversion algorithms. Following this hypothesis, applying the same algorithm to all deduced BTD operational parameters should reduce the difference in wind speed to zero, assuming that the same CI parameters were provided by the meteorological agencies. We found that this remapping method leads to enhanced agreement in TC days for the highest wind speeds, but relatively high differences are still present. This indicates that there are additional contributing factors, which, in the earlier TC intensity estimation stage, cause discrepancies in operational parameters (T and CI).

The differences among BTD show temporal variation. High agreement in TC-day records is visible in the first years of the analysis (1977–87). As a possible explanation, Knapp and Kruk (2010) suggested that the same Dvorak procedures (e.g., the same conversion algorithm) were applied for this period. In the second period (1988–97) numbers and trends among original BTDs differ a great degree. However, unifying wind speed definitions (application of the Dvorak table to BTD) did not efficiently resolve differences in the highest categories. Discrepancies among BTDs in this period are increased, very similar to the strong increase of TC-day records in JTWC. In contrast, TC activity for the last decade shows good agreement, and an increasing trend for the categories 4–5 for JTWC, JMA*1.14, and JMADT. We conclude that unifying the conversion algorithms, and thus wind speed definitions, is necessary for an accurate assessment of BTDs. However, the trend statistics derived from the given datasets remain inconsistent. This requires an explanation of the remaining differences, as offered in the following.

c. *Can the reasons for discrepancies between BTD and the discrepancies themselves be evaluated?*

Here we focus on the trends derived from JTWC and JMADT in search of additional reasons for the remaining differences. The resulting discrepancies indicate that there are differences among CI numbers provided by the BTD agencies. To visualize the problem, which cannot be resolved by applying the same DT algorithm, two intense typhoons, Isa (1997) and Dianmu (2004), are presented in Figs. 4a,b. The figure shows a time series of

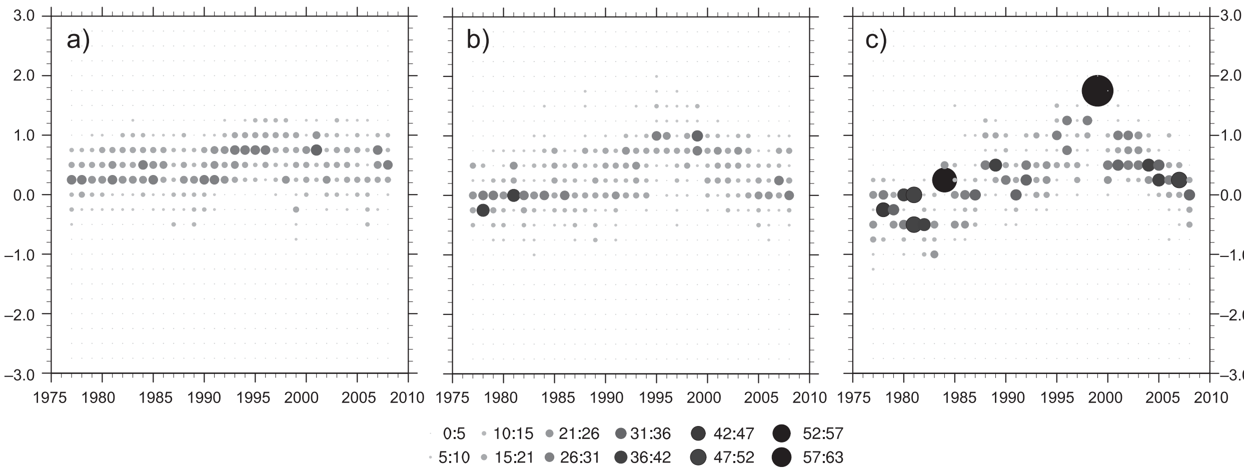

FIG. 5. Distribution of CI-number differences for JTWC − JMADT assigned to intensity categories: (a) TD, TS, 1; (b) 2–3; and (c) 4–5. The circles indicate the percentage of the occurrence number, counted for each year separately.

maximum wind speed given by different BTD. The differences between original 10-min JMA data and 1-min JTWC reach 30 m s^{-1} during peak winds. Adjusting JMA to 1-min wind speed using a multiplication factor reduces the difference to 25 and 20 m s^{-1} for Isa and Dianmu, respectively. After applying the same Dvorak conversion table a difference of 20 and 15 m s^{-1} still remains, which corresponds to a difference in CI parameters of 1.75 and 1 (Fig. 3). For TC Isa, a high discrepancy is noticeable during the whole TC lifetime. For Dianmu, the main differences occur during the highest-intensity phase, when the TC in JMADT reaches the fourth category. It is also worthy to note how the multiplication factor shapes the values during the TC lifetime. JMA*1.14 shows higher intensity than JMA/JMADT intensities in the categories TD, TS, and 1, but lower intensity than JMADT in the peak categories.

Kruk et al. (2011) considered CI parameters for most TCs in the WNP to be almost identical between the BTD agencies, with the 95th percentile varying between 5.75 and 6.25 among BTDs. However, we would like to emphasize that, for the highest-intensity categories, noticeable differences are apparent, as shown by two example TCs. Figure 5 presents the CI parameter annual differences distribution for JTWC and JMADT, corresponding to the remaining intensity differences, separated into three categories: TD-1, 2–3, and 4–5. The distribution of TC lifetime discrepancies (Fig. 1b,c) reflects the differences of CI numbers. The most pronounced differences are visible for the highest parameters, especially in the second period, 1988–97. In this period the CI differences were increasing in time and reached the extreme high percentage of CI differences of 2 in 1997. Lower categories, although with smaller CI differences, retained similar features, as did the enhanced TC activity level in the second period, especially in the years 1995–97. In the early 2000s CI discrepancies are still higher, especially for categories 4–5. Two periods of the strongest CI discrepancies were also identified by Nakazawa and Hoshino (2009), who analyzed operational parameters from 1987–2006. They found a significantly higher numbers in JTWC for 1992–97 and 2000–05 in comparison to JMA.

The reasons for changes in time of CI discrepancies can be related to separately evolving practices and usage of different information sources by operational centers. JMA reports geostationary satellites to be the principal source of TC localization and intensity estimation. In contrast, JTWC emphasize supplementing these data with other: remotely sensed and in situ observations, that are useful for TC center identification, defining TC structure, and providing more direct intensity estimation. The distribution of differences in TC position among BTD shown in Fig. 6 might indicate that different satellite-based sources were used for intensity estimation. Figure 6b shows annual means of TC center differences provided by JTWC and JMA. The mean annual differences in TC center position decrease with increasing intensity. The highest discrepancies occur for weak TCs (CI range of 1–4.75), where often intensity and centers are difficult to estimate by low-resolved observations. In contrast, there is better agreement in locating the strongest TC centers.

The most striking values are visible for the period 1988–98, when the aircraft reconnaissance era in the WNP was replaced by intensively developing satellite measurements. In that time widely distributed differences in TC locations were up to 150 km with mean

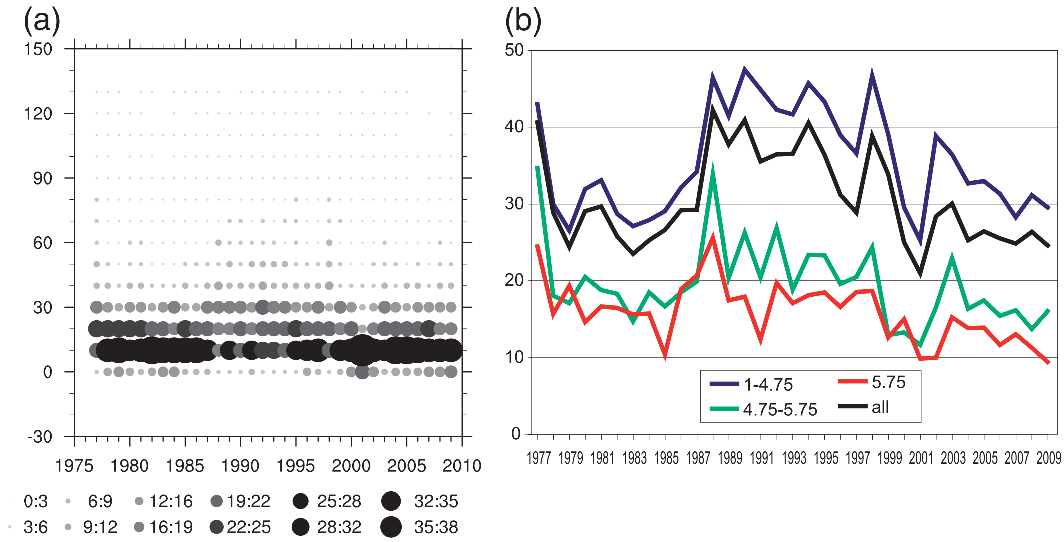

FIG. 6. (a) Distribution of TC position differences (unit: km) between JTWC and JMA. The circles indicate the percentage of the occurrence number, counted for each year separately. (b) Annual mean differences in TC position between JTWC and JMA for categories given by CI number.

annual differences varying between 30–50 km. After 1998 these differences are significantly smaller and do not exceed 30 km.

The relationships between BTD trends in these distinct three periods correspond well with those of annual CI differences and TC-days trends (Figs. 1 and 5). The larger TC location differences for the mid-period correlates well with strong CI discrepancies and opposite TC-days trends. In the last decade both TC location and CI differences show downward tendencies. The TC activity trends in that time are similar for JTWC and JMADT, even for the strongest categories.

d. Additional contributors for BTD inconsistencies

The analysis shows that differences in CI numbers and TC locations share a strong relationship. They are most distinguishable in the years 1987–98, when the aircraft reconnaissance terminated and development of the intense satellite measurements began. Such coincidence suggests the usage of different information sources by JTWC and JMA may be a reason for the given TC trend differences. JMA reports usage of geostationary imageries only as a source for intensity estimation. In contrast, JTWC's operational center uses all available satellite data to ascertain the location and underlying storm structure and therefore improves the information used for imagery processing with the Dvorak technique. Such practices in JTWC might increase intensity values and contribute strongly to increasing tendencies of intense TC days.

Increasing coverage of microwave observations (SSM/I) from 1987 onward, which reached the maximum in 1997, together with high-resolution scatterometer [i.e., the *European Remote Sensing Satellite-2* (*ERS-2*)] measuring in 1995–97, helped in TC center positioning and analysis of the lower-intensity systems. Enhanced radar usability and additional information of higher-resolution TRMM in 1997 improved the accuracy of Dvorak-based estimations in JTWC. Introducing more and better spatially resolved data certainly could affect the dataset homogeneity and statistical information concerning derived trends. Extensive and irregular use of additional supplementary sources by one operational center and not the other, might lead to large CI discrepancies and opposite trends of intense TCs activity in comparison to other BTD. The strong, increasing tendency in intense TC days found in JTWC, especially for the period 1987–99, might be severely biased by inhomogeneities introduced by changing procedures and different information sources applied in the operational centers.

We suggest that apart from differing methods for converting CI numbers to intensities, CI discrepancies are the main contributor to differences between TC activity trends. Our analysis indicates that discrepancies among operational parameters occur due to different data used as input for the Dvorak method applied in JTWC. However, to check the credibility of these parameters, they need to be compared with reference data.

e. Can the CI discrepancies between BTDs be evaluated? A NOAA-BTD–aircraft-BTD comparison

To evaluate CI discrepancies, records for the years 2000–08 in NOAA, JTWC, and JMADT were analyzed for the TS category. The main input of NOAA, QuikSCAT, is stated as having highly reliable values for moderate and high TS values, while slightly overestimating the wind of tropical depression strength. However, it provides data adhered to an 8-min average. For this reason, NOAA can underestimate values up to 2 m s^{-1} when comparing with 1-min wind speed values within the TS category.

Figure 7 presents computed annual mean differences for concurrent records between JTWC and JMADT. In this comparison JTWC reveals systematically higher values compared to JMADT. For less than 15% of all cases the absolute difference is smaller than 2 m s^{-1}, which according to Kruk et al. (2011), is within the range of the remapping method's accuracy. However, for the majority of cases (60%) JTWC is higher than JMADT by 2–8 m s^{-1}. For our comparison the data was divided into two groups according to these relationships. For the first one, representing almost 60% of cases, JMADT remains like JTWC within the TS category. For the second group, representing over 40% of the cases, JMADT is low enough to fall into the TD category. To assess which agency gives more reliable parameters, these two groups are compared with NOAA. They are analyzed separately, with a greater focus on the first one (TS) caused by high reference data reliability.

Figures 8a,b presents mean differences for NOAA minus JMADT and for NOAA minus JTWC, computed for the whole analyzed period, for both groups. For the group that contains data of both analyzed BTD within the TS category, NOAA remains closer to JMADT with 26% of the records remaining within absolute difference of 2 m s^{-1} and 50% within 4 m s^{-1}. However, NOAA presents slightly higher values then JMADT with a median for the differences in the range $<0, 2>$ m s^{-1}. In comparison with JTWC, NOAA has lower values for more then 60% of the records, with the median within the range of $<-4, -2>$ m s^{-1}.

For the second group, where JMADT indicates the TD phase, only 15% of the NOAA values remain within absolute difference of 2 m s^{-1} of JMADT. Here NOAA presents stronger tendencies toward higher values with a median of the difference in the range of $<4, 6>$ m s^{-1}. However, this might be caused by a positive bias introduced by scatterometer data during rainy conditions for tropical depressions. Despite this fact, JTWC still remains higher than NOAA in almost 50% of the cases.

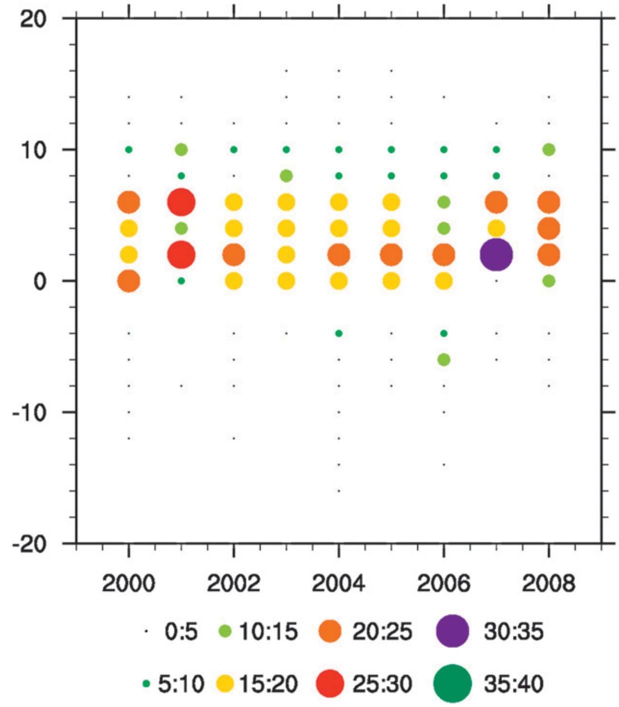

FIG. 7. Distribution of TC intensity differences (unit: m s^{-1}) between JTWC − JMADT for JTWC in the TS category. The circles indicate the percentage of the occurrence number, counted for each year separately.

Figure 9a presents a TC from 2008 where JTWC wind values were higher during the whole event, except for the TD and early TS phase when NOAA showed the highest values. For this TC the NOAA values remained noticeably closer to JMADT.

The highest discrepancies still remain in the highest wind categories, therefore an evaluation of adjustment methods for categories 4–5 is crucial for determining trends in TC activity. For two intense TCs, Jangmi in 2008 and Megi in 2010, aircraft-measured maximum sustained wind speed are available. For the TC Jangmi maximum wind speed estimates of JTWC (72 m s^{-1}) match the observed ones given by SMFR better than JMADT. For this case JMADT presents the highest values (79 m s^{-1}), while SMFR 60-s observations show 68 m s^{-1}. Figure 9c also shows the Supertyphoon Megi in 2010 for which maximum wind speed was measured during an aircraft campaign as well. For this event, SFMR measurements, even after averaging by a 60-s interval, show the highest values (90 m s^{-1}): 87 m s^{-1} for JMADT and 82 m s^{-1} for JTWC.

f. Accuracy of intensity estimations given by BTDs

To evaluate CI discrepancies, BTD records were compared with satellite-based NOAA data and aircraft

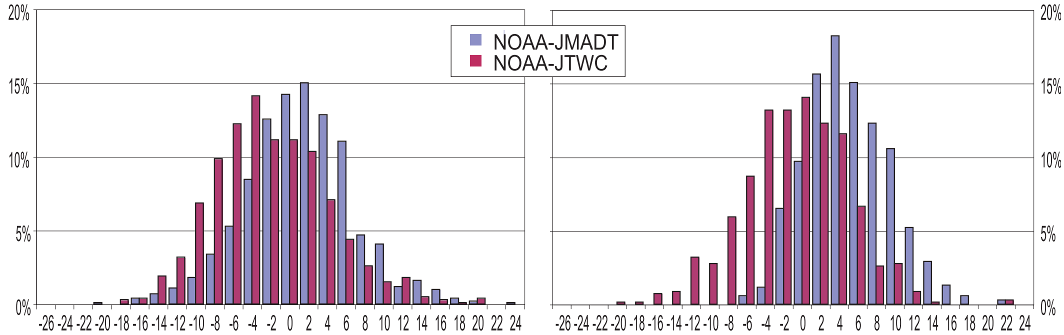

FIG. 8. Differences distribution of wind speed for NOAA − JMADT and NOAA − JTWC for (a) JTWC and JMADT in the TS category and (b) JTWC in TS and JMADT degraded to the TD category.

observations. NOAA serves as reference data for the lower-intensity categories, while aircraft observations are used for the highest wind speed evaluation.

Wind values derived from NOAA for the TS phase provide data with reliable accuracy and remain closer to JMADT than JTWC. Nevertheless, a wide spread of differences exists among the data. JTWC shows much higher values than NOAA and JMADT, even in the group where JMADT falls into the TD category and a possible positive bias in NOAA has been taken into account. This indicates possible intensity overestimations in JTWC due to an erroneous contribution of CI parameters. Such overestimations may also be caused by supplementary data usage of JTWC (e.g., QuikSCAT), which gives values averaged over a 25-km area and an 8-min interval. These values would be treated as the minimum threshold for estimation by a forecasted maximum wind speed. In the results, JTWC may increase the final wind estimates to compensate for possible underestimations due to wind retrieval limitations. Figure 9a shows time series of TC intensity for Typhoon Dolphin in 2008 and serves as an example for pronouncedly higher wind speed values of JTWC in comparison to reference data (NOAA) and alternative BTD. However, the indirect way of choosing the maximum wind speed for NOAA winds (which provide reliable information only for lower TC intensity categories), as well as the limited accuracy of the remapping method still contribute to the uncertainty in our estimation of BTD reliability.

We now focus on CI parameters in the higher part of the SSHS intensity scale, where the strongest discrepancies still remain. An evaluation of BTD for categories 4–5 is crucial for determining trends in TC activity. As aircraft sensors are unable to provide direct measurements of 10-m, 1-min sustained wind speed, they serve only as input to prepare surface wind analyses. Here the initialization conditions and assimilation techniques are crucial to construct reliable analyses. Figure 9b presents BTD, aircraft observations provided by SFMR taken

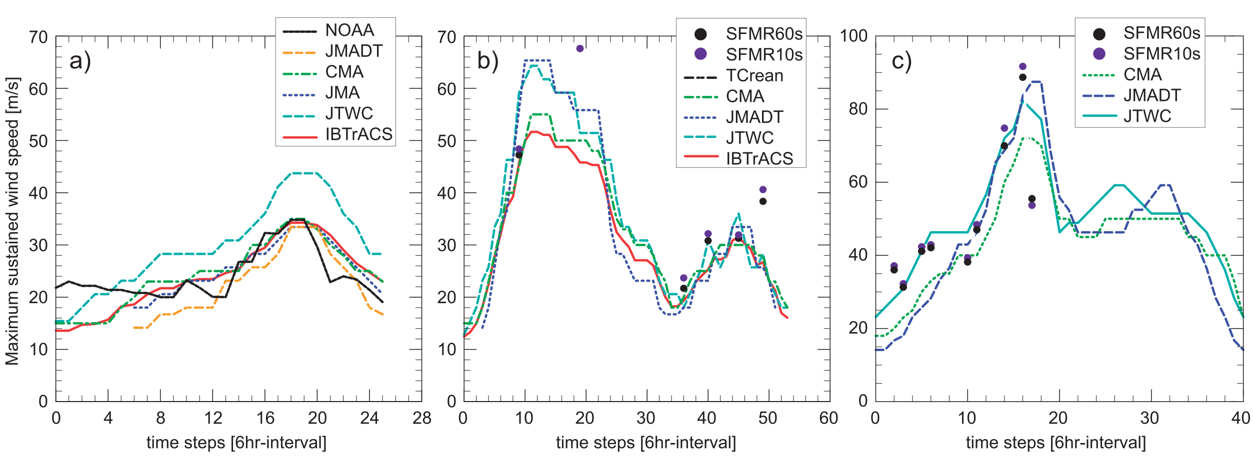

FIG. 9. Wind speed time series for the TC event: (a) Dolphin (2008) for different best track datasets and NOAA; (b) Sinlaku (2008), for BTDs, TC reanalysis including aircraft reconnaissance (TCrean), and aircraft observations; and (c) Megi (2010) for BTDs and aircraft observations (SFMR10s and SFMR60s). SFMR10s and SFMR60s are intensities averaged over a 10- and 60-s time interval.

during the TCS-08 2008 campaign, and an analysis reconstructed with those observations (Zhang et al. 2007) for Typhoon Sinlaku in 2008. The initialization scheme assimilates TC central minimum pressure given by JTWC, but the maximum wind speed for higher categories does not reach JTWC values. As the provided TC reconstruction may be also biased due to the 10-km horizontal resolution, this can complicate the evaluation of BTD. On the other hand, the JTWC report (Joint Typhoon Warning Center 2009) states, that the aircraft measurements themselves for this TC had decisive impact on intensity estimation. Aircraft reconnaissance in this case helped to identify the second intensification phase. While for the first intensification phase the Dvorak technique estimated intensity with good accuracy, it underestimated the TC intensity during the second phase. The reconstructed reanalysis for the second period matches the observed values.

For Typhoon Jangmi the flights during the T-PARC aircraft campaign occurred at the time of TC maximum intensity for which a mean 60-s value of 68 m s^{-1} was measured while JMADT estimated the highest values (79 m s^{-1}). For Megi, SFMR 60-s measurements show the highest values (90 m s^{-1}) of maximum wind speed. Additionally, the SFMR recorded the weakening of TC Megi faster than estimated by the Dvorak method. Landfalling TC situations, for which the reliability of Dvorak relationships is limited, require in situ observations. Nakazawa and Hoshino (2009) also noticed differences in operational (CI and *T*) numbers among various BTD, both for intensification and weakening phases. Differing weakening ratios, after reaching TC maximum intensity in BTDs, indicate that there may be differences between definitions for allowable intensity change (in the form of CI and *T* parameters). Such constraints (Dvorak 1984) were gradually relaxed by JTWC during the 1990s (Velden et al. 2006), allowing for a faster weakening of intense TCs. These procedural changes possibly contributed to the existing discrepancies among BTD.

Additionally, it is noticeable that in the developing stage of a typhoon, BTD in JTWC is strongly influenced by aircraft measurements (Fig. 9c). These were possibly used to supplementarily identify the early intensification phase.

4. Summary and conclusions

This paper assesses the reliability of BTD in climate statistics for the WNP region. We confirmed that the different methodologies to derive TC intensities used by the meteorological agencies producing BTD influence TC activity trends. Therefore, the skill of methods to minimize discrepancies between the individual datasets was evaluated. Both the commonly used linear factor multiplication method (used to homogenize BTD with different wind speed intervals) as well as the method of Knapp and Kruk (2010) show high skill to reduce trend discrepancies, but only when categories 2–5 are considered together. Then all BTD show increasing numbers of annually accumulated TC days for the period 1977–2008. However, when analyzing categories 2–3 and 4–5 separately, the methods' skills differ. We found that using a multiplication factor lead to overestimated trends of TC days for lower categories (2–3), while still underestimating the highest ones (4–5).

An alternative method, which reconstructs TC intensity by remapping CI parameters with a DT conversion (Knapp and Kruk 2010), reduces most discrepancies for categories 2–3. For the highest categories, the technique minimizes discrepancies only partly; TC activity trends in JMADT show no trend while strongly increasing trends are visible for JTWC.

The application of the same converting procedures to retrieve TC intensities should theoretically reduce the difference between the individual BTD to zero. However, remaining differences indicate that there are additional contributing factors leading to discrepancies in operational CI numbers.

The distribution of the CI discrepancies in time corresponds to the differences in TC center positions. The largest discrepancies occur in the 1980s when higher-resolution satellite observations were developing. Toward the latter half of the decade, the reduction and phasing out of aircraft data sources may have also had an influence.

This indicates that extensive and irregular use of additional supplementary sources by JTWC might cause huge CI discrepancies and opposite trends of intense TCs activity with other BTDs. The strong increasing tendency in intense TC days found in JTWC, especially for the period 1987–99, might be severely biased by inhomogeneities introduced by changing procedures and information sources. Using only the geostationary satellite imageries for intensity estimations by JMA limits its accuracy. On the other hand, this maintains homogeneity within the dataset, which makes this source more reliable for deriving climate statistics.

The CI numbers and wind intensity of JTWC and JMADT were compared to NOAA sea surface wind speeds and aircraft measurements to evaluate which BTDs provide more accurate estimations. JTWC shows a systematic overestimation of both NOAA and JMADT for the TS category, whereas NOAA data are considered to be very accurate. For the TS category JMADT wind speed values remain closer to NOAA, although visible differences still exist. Higher CI parameter estimates as well as subjective interpretation of additional sources in JTWC (e.g., microwave wind retrievals) likely contribute to such results. We conclude that JMA provides more reliable CI parameters than JTWC for the TS wind speed range.

Sparse in situ data limit the evaluation effort of BTD accuracy for the highest wind regimes. Aircraft campaign measurements in 2008 and 2010 show some agreement with maximum intensity estimations in BTDs. For a more complete evaluation, aircraft data for the earlier period would be needed, when the accuracy uncertainties were the highest.

Additionally, the analysis of some strong TC events like Sinlaku (2008) and Megi (2010) suggests that there are some deficiencies within the Dvorak technique procedures. Slow weakening ratios for BTDs in comparison to in situ observations indicate that not only the homogeneity has to be assured, but also that temporal nonchanging methods to estimate TC intensities should be applied in all operational centers.

We emphasize the importance of documenting those operational procedures that are applied for the Dvorak technique by the meteorological agencies; otherwise, the interpretation of the results can lead to misleading conclusions. This may happen when considering ambiguously specified wind speed definitions in CMA or intensity in JMA before applying the conversion table in Koba et al. (1991). We suggest paying special attention with regard to the highest wind regimes as the largest differences between BTDs were found here. The differences in TC activity trends may require academic agreement on a set of procedures and a reanalysis of existing storm data.

Acknowledgments. Work is supported (in parts) through the Cluster of Excellence ''CliSAP,'' University of Hamburg, funded through the German Science Foundation (DFG-EXC177), and through the Helmholtz-Zentrum Geesthacht, Centre for Materials and Coastal Research. The authors thank S. Bakan, W. Koch, and J. Horstmann for valuable discussions and B. Gardeike who helped with the preparation of the figures for this paper. The authors thank colleagues who provided best track datasets from the Japan Meteorological Agency (http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/besttrack.html), the China Meteorological Agency (http://www.typhoon.gov.cn), the Joint Typhoon Warning Center (http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/wpindex.html), and the International Best Track Archive for Climate Stewardship (http://www.ncdc.noaa.gov/oa/ibtracs/index.php?name=ibtracs-data). We appreciate being able to use the Blended Sea Winds provided by NOAA (http://www.ncdc.noaa.gov/oa/rsad/seawinds.html), the Typhoon Reanalysis for the western Pacific (http://apdrc.soest.hawaii.edu/datadoc/typhoon.php), and T-PARC-2008, ITOP-2010 field campaign measurements (http://www.aoml.noaa.gov/hrd/data_sub/hurr.html). This work is a contribution to the ''Helmholtz Climate Initiative REKLIM'' (Regional Climate Change), a joint research project of the Helmholtz Association of German research centres (HGF). We also thank the two anonymous reviewers for their constructive comments.

REFERENCES

Atkinson, G. D., 1974: Investigation of gust factors in tropical cyclones. FLEWEACEN Tech. Note JTWC 74-1, Fleet Weather Center, Guam, 9 pp.

Bourassa, M. A., and Coauthors, 2010: Remotely sensed winds and wind stresses for marine forecasting and ocean modeling. *Proc. ''OceanObs'09: Sustained Ocean Observations and Information for Society'' Conf.,* Vol. 2, Venice, Italy, ESA Publ. WPP-306, doi:10.5270/OceanObs09.cwp.08. [Available online at http://www.oceanobs09.net/proceedings/cwp/cwp08/index.php.]

Brennan, M. J., C. C. Hennon, and R. D. Knabb, 2009: The operational use of QuikSCAT ocean surface vector winds at the National Hurricane Center. *Wea. Forecasting,* **24,** 621–645.

Chan, J. C. L., and K. S. Liu, 2004: Global warming and western North Pacific typhoon activity from an observational perspective. *J. Climate,* **17,** 4590–4602.

Dean, L., K. A. Emanuel, and D. R. Chavas, 2009: On the size distribution of Atlantic tropical cyclones. *Geophys. Res. Lett.,* **36,** L14803, doi:10.1029/2009GL039051.

Dvorak, V. F., 1972: A technique for the analysis and forecasting for tropical cyclone intensities from satellite pictures. NOAA Tech. Memo. NESS 36, NOAA, 15 pp.

——, 1973: A technique for the analysis and forecasting for tropical cyclone intensities from satellite pictures. NOAA Tech. Memo. NESS 45, NOAA, 19 pp.

——, 1975: Tropical cyclone intensity analysis and forecasting from satellite imagery. *Mon. Wea. Rev.,* **103,** 420–430.

——, 1984: Tropical cyclone intensity analysis using satellite data. NOAA Tech. Rep. NESDIS 11, NOAA, 45 pp.

Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature,* **436,** 686–688.

Hoffman, R. N., and S. M. Leidner, 2005: An introduction to the near-real-time QuikSCAT data. *Wea. Forecasting,* **20,** 476–493.

Joint Typhoon Warning Center, 2009: 2008 annual tropical cyclone report. Joint Typhoon Warning Center, Pearl Harbor, HI, 116 pp. [Available online at http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/atcr/2008atcr.pdf.]

Kamahori, H., N. Yamazaki, N. Mannoji, and K. Takahashi, 2006: Variability in intense tropical cyclone days in the western North Pacific. *SOLA,* **2,** 104–107.

Knaff, J. A., and C. R. Sampson, 2006: Reanalysis of West Pacific tropical cyclone intensity 1966–1987. Preprints, *27th Conf. on Hurricanes and Tropical Meteorology,* Monterey, CA, Amer. Meteor. Soc., 5B.5. [Available online at http://ams.confex.com/ams/27Hurricanes/techprogram/paper_108298.htm.]

Knapp, K. R., and M. C. Kruk, 2010: Quantifying interagency differences in tropical cyclone best-track wind speed estimations. *Mon. Wea. Rev.,* **138,** 1459–1473.

——, ——, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010: The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying tropical cyclone best-track data. *Bull. Amer. Meteor. Soc.,* **91,** 363–376.

Kruk, M. C., P. A. Hennon, and K. R. Knapp, 2011: On the use of the Dvorak current intensity as a climate data record in the western North Pacific. Preprints, *23rd Conf. on Climate Variability and Change,* Seattle, WA, Amer. Meteor. Soc., 142.

Koba, H., T. Hagiwara, S. Osano, and S. Akashi, 1991: Relationships between CI number and minimum sea level pressure/maximum wind speed of tropical cyclones. *Geophys. Mag.,* **44,** 15–25.

Landsea, C. W., 2005: Hurricanes and global warming. *Nature,* **438,** E11–E13.

Nakazawa, T., and S. Hoshino, 2009: Intercomparison of Dvorak parameters in the tropical cyclone datasets over the western North Pacific. *Sci. Online Lett. Atmos.,* **5,** 33–36.

Ren, F., J. Liang, G. Wu, W. Dong, and X. Yang, 2011: Reliability analysis of climate change of tropical cyclone activity over the western North Pacific. *J. Climate,* **24,** 5887–5898.

Simpson, R. H., 1974: The hurricane disaster potential scale. *Weatherwise,* **27,** 109–186.

Song, J. J., Y. Wang, and L. Wu, 2010: Trend discrepancies among three best track data sets of western North Pacific tropical cyclones. *J. Geophys. Res.,* **115,** D12128, doi:10.1029/2009JD013058.

Trenberth, K., 2005: Uncertainty in hurricanes and global warming. *Science,* **308,** 1753–1754, doi:10.1126/science.1112551.

Velden, C., and Coauthors, 2006: The Dvorak tropical cyclone intensity estimation technique: A satellite-based method that has endured for over 30 years. *Bull. Amer. Meteor. Soc.,* **87,** 1195–1210.

Webster, P. J., G. J. Holland, J. A. Curry, and H. R. Chang, 2005: Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science,* **309,** 1844–1846, doi:10.1126/science.1116448.

Wu, M. C., K. H. Yeung, and W. L. Chang, 2006: Trends in western North Pacific tropical cyclone intensity. *Eos, Trans. Amer. Geophys. Union,* **87** (48), 537–538.

Yu, H., C. Hu, and L. Jiang, 2007: Comparison of three tropical cyclone intensity datasets. *Acta Meteor. Sin.,* **21,** 121–128.

Zhang, H.-M., J. J. Bates, and R. W. Reynolds, 2006a: Assessment of composite global sampling: Sea surface wind speed. *Geophys. Res. Lett.,* **33,** L17714, doi:10.1029/2006GL027086.

——, R. W. Reynolds, and J. J. Bates, 2006b: Blended and gridded high resolution global sea surface wind speed and climatology from multiple satellites: 1987–present. Preprints, *14th Conf. on Satellite Meteorology and Oceanography,* Atlanta, GA, Amer. Meteor. Soc., P2.23. [Available online at http://ams.confex.com/ams/Annual2006/techprogram/paper_100004.htm.]

Zhang, X., T. Li, F. Weng, C.-C. Wu, and L. Xu, 2007: Reanalysis of western Pacific typhoons in 2004 with multi-satellite observations. *Meteor. Atmos. Phys.,* **97,** 3–18.