The Use of the Conway–Maxwell–Poisson in the Seasonal Forecasting of Tropical Cyclones

TIMOTHY D. MITCHELL^a AND JOANNE CAMP^a

^a *Met Office, Exeter, United Kingdom*

(Manuscript received 10 September 2020, in final form 11 March 2021)

ABSTRACT: The Conway–Maxwell–Poisson distribution improves the precision with which seasonal counts of tropical cyclones may be modeled. Conventionally the Poisson is used, which assumes that the formation and transit of tropical cyclones is the result of a Poisson process, such that their frequency distribution has equal mean and variance (“equi-dispersion”). However, earlier studies of observed records have sometimes found overdispersion, where the variance exceeds the mean, indicating that tropical cyclones are clustered in particular years. The evidence presented here demonstrates that at least some of this overdispersion arises from observational inhomogeneities. Once this is removed, and particularly near the coasts, there is evidence for equi-dispersion or underdispersion. To more accurately model numbers of tropical cyclones, we investigate the use of the Conway–Maxwell–Poisson as an alternative to the Poisson that represents any dispersion characteristic. An example is given for East China where using it improves the skill of a prototype seasonal forecast of tropical cyclone landfall.

SIGNIFICANCE STATEMENT: It is challenging to make seasonal forecasts of tropical cyclones (also known as hurricanes or typhoons). There is good skill for predicting numbers of tropical cyclones within an ocean basin. However, skill is much more limited when forecasting where tropical cyclones will make landfall. Yet this is where their most severe impacts are felt and so where seasonal forecasts could be most useful. This paper contributes to the development of seasonal forecasts by identifying a statistical distribution, the Conway–Maxwell–Poisson, which can improve their accuracy. Its usefulness depends on the preexistence of a source of predictability.

KEYWORDS: Tropical cyclones; Statistical techniques; Seasonal forecasting

1. Introduction

Seasonal forecasts of tropical cyclones have been issued since the 1980s (Nicholls 1979; Gray 1984). Statistical methods are now complemented by dynamical methods using the higher-resolution models that are available (Camp et al. 2015). However, these models may better represent the wider environmental conditions that influence tropical cyclones than they do the tropical cyclones themselves (Camargo and Wing 2016). Therefore hybrid methods are often used, where a dynamical model predicts the large-scale flow and a statistical model relates that to the number of tropical cyclones (Choi et al. 2016; Zhang et al. 2017). These forecasts require statistical methods that precisely capture the observed year-to-year variability of tropical cyclone counts. The need for precision is particularly acute where the forecast is of tropical cyclones making landfall in a relatively small region, where the counts are low. A recent review remarked that “skillful forecasts of the probability of occurrence of landfalling storms for specific coastal regions would be a huge improvement from basin-wide quantities” (Camargo and Wing 2016, 221–222). In this paper we examine whether, in certain cases, we might improve the skill of such a forecast by using the Conway–Maxwell–Poisson (CMP) distribution (Conway and Maxwell 1962; Shmueli et al. 2005).

The formation of a tropical cyclone, and its arrival in a particular region, might be reasonably considered as Poisson processes. Therefore from early on the Poisson distribution was adopted to describe tropical cyclone frequencies (Thom 1960; Cry 1965; Mooley 1980), and it has become standard practice to use the Poisson, which is “a natural probability model for count data” (Wilks 2011, p. 242) and “the tool of choice” for tropical cyclone counts (Elsner 2003, p. 354). However, the Poisson may not always be the best choice, being a univariate distribution in which the variance always equals the mean (“equi-dispersion”), whereas real-world data might be either underdispersed or overdispersed. Underdispersion represents tropical cyclones occurring more regularly (as in a metronome) than would happen under a theoretical Poisson process; overdispersed tropical cyclones occur in clusters. Empirical departures from equi-dispersion have interested statisticians for many years (Lexis 1879; Student 1919). The mechanisms that lead to under and overdispersion may be categorized into experimental and observational (Xekalaki 2014). For tropical cyclone counts, these categories correspond to mechanisms inherent to the climate system, and mechanisms arising from inhomogeneous observational practices. This distinction is important, as overdispersion arising from inhomogeneities in the observations ought not to be reproduced in a predictive model. On the other hand, if the under or overdispersion is inherent to the phenomenon being predicted, then using the Poisson in a predictive model will impair its performance. The model performance might be improved by using a distribution that reproduces the under or overdispersion.

Previous studies have used the negative binomial distribution where overdispersion has been found (Thom 1960; Villarini et al. 2010). For underdispersion there is no established

Denotes content that is immediately available upon publication as open access.

Corresponding author: Timothy D. Mitchell, tim.mitchell@metoffice.gov.uk

replacement for the Poisson: the negative binomial, the gamma, and Consul's generalized Poisson all have to be ruled out (Sellers et al. 2012; Lord and Guikema 2012). However, there is a distribution that has been developed specifically for such data, the Conway–Maxwell–Poisson. This distribution was originally conceived as a model for queueing systems with state-dependent service times (Conway and Maxwell 1962), but was subsequently ignored and largely forgotten. It was revived to address the lack of a suitable distribution for modeling data that resemble the Poisson, but are differently dispersed (Shmueli et al. 2005). Its revival has depended on advances in computational power since the 1960s. Subsequently it has been developed and refined (Sellers and Shmueli 2010; Sellers and Raim 2016). Examples of its use may now be found in fields as widely dispersed as ecommerce, linguistics, vehicle crash analysis, and ecology (Sellers et al. 2012).

The first stage of our analysis will re-evaluate the dispersion of the observed records to assess whether underdispersion might be present. We will examine the North Atlantic because, although we will subsequently make a seasonal forecast for the western North Pacific, there are important divergences between the observational records in the western North Pacific (Song et al. 2010); the reanalysis project that the basin needs, while often proposed (e.g., Knapp et al. 2013), is yet to happen. For the North Atlantic, in contrast, the record of tropical cyclones is undergoing careful revision. The well-used HURDAT (Hurricane Database) was built in the early 1980s (Jarvinen et al. 1984) and subsequently reissued in a revised format as HURDAT2 (Landsea and Franklin 2013). The Atlantic Hurricane Database Reanalysis Project has thus far revisited the underlying records for 1851–1965 to identify and correct inhomogeneities (Landsea et al. 2004; Delgado et al. 2018). Other significant work has evaluated the likely effects of the historical changes in the technology and practices used to detect and measure tropical cyclones (Landsea 2007; Vecchi and Knutson 2008). These assessments of the quality of the observational record give us the opportunity to evaluate the possible role of changes in observational practice in influencing our evaluation of dispersion.

Previous studies for the North Atlantic have tended to find overdispersion rather than underdispersion (Villarini et al. 2010). However, a common cause of overdispersion in any dataset is inhomogeneity (Xekalaki 2014). It may be significant, therefore, that the overdispersion has been identified most strongly in basinwide measures (Thom 1960) and in systems that do not reach hurricane strength (Mumby et al. 2011), where the quality of observations is likely to be less consistent. Since the HURDAT2 reanalysis has identified and corrected some inhomogeneities in the existing records, we will compare the dispersion statistics of the uncorrected and corrected observations to establish whether inhomogeneities have contributed to the overdispersion that earlier studies found.

Another potential source of inhomogeneity is missing events. Historical changes in observational practices have probably had a material effect on the proportion of tropical cyclones that have been observed and recorded (Landsea 2007; Vecchi and Knutson 2008). While the reanalysis project has improved the records of observed events, “such efforts will not be able to recover observations of open-ocean tropical cyclones that were

just never taken” (Landsea 2007, p. 200). This is a time-varying destruction of theoretically observable events that will affect our estimates of dispersion (Xekalaki 2014). We cannot remove this observational inhomogeneity. We will, however, evaluate its likely effect by adding back into the counts the events that are thought to be missing, so as to calculate the effect on dispersion of varying numbers of missing tropical cyclones over the course of the record.

If estimates of dispersion are affected by observational methods, those estimates may be different near the coast compared to over the open ocean. Therefore, following Mumby et al. (2011), we will examine the spatial distribution of the dispersion statistic in the North Atlantic and the western North Pacific to identify regions where there may be underdispersion.

The second stage of our analysis will examine whether using an alternative to the Poisson might improve the skill of a seasonal forecast. The most significant impacts of tropical cyclones are where they make landfall, but building seasonal forecasts of landfall is even more demanding than for basinwide activity (Vecchi and Villarini 2014). It is therefore particularly pertinent to determine if tropical cyclone counts at the coast are underdispersed. In such cases the statistical model should fit that underdispersion.

The immediate need for such a model has arisen via the opportunity to build a climate service around the issuing of seasonal forecasts of tropical cyclone landfall for East China (Hewitt et al. 2020). The forecast relies on the influence on storm tracks of the Western Pacific subtropical high (WPSH; Wang et al. 2013). The interannual variability in the westward extent of the WPSH in summer is skilfully predicted by GloSea5, a seasonal forecast model (Camp et al. 2019). A prototype forecast has already been developed using a hybrid method in which the model forecast of the WPSH was related to observed tropical cyclone activity through linear regression (Camp et al. 2020).

We seek to improve the skill of this prototype. To do this, we investigate the use of the Poisson distribution, since it might better represent the observed distribution. However, if the observed series is not equi-dispersed, an alternative such as the Conway–Maxwell–Poisson may outperform the Poisson. We therefore build predictive models using each distribution and compare their skill.

In section 2 we describe our data and methods. In section 3 we present the results. In section 4 we draw some conclusions concerning the best practices to adopt when making seasonal forecasts of tropical cyclone counts.

2. Data and methods

a. Dispersion

The annual counts for North Atlantic named tropical storms, hurricanes, and hurricanes making landfall in the United States are from HURDAT (Jarvinen et al. 1984) and HURDAT2 (Landsea and Franklin 2013). HURDAT2 now includes corrections that have been made incrementally over recent years as storm tracks have been reanalyzed, and differs from the original HURDAT up to 1965 (NOAA 2020). We also consider a partly revised version as it stood in 2008 (Vecchi and Knutson 2008).

Following Fisher (1950), the index of dispersion is the ratio of the sample variance (using $n - 1$) to the sample mean:

$$D = s^2/\bar{x}. \quad (1)$$

Where appropriate, we test an observed series for the significance of any departure from equi-dispersion ($D = 1$) using a bootstrap method (Wilks 2011, section 5.3.5). In this test the observed series is resampled with replacement 10 000 times; the samples are ordered by D ; the $(1 - \alpha)\%$ confidence interval for D is given by the central $\alpha \times 10\,000$ samples. If this confidence interval does not span equi-dispersion, the observed series is significantly under or overdispersed.

The Conway–Maxwell–Poisson generalizes the Poisson distribution so as to model departures from equi-dispersion. Following Shmueli et al. (2005), its probability function is

$$P(X = x) = \frac{\lambda^x}{(x!)^\nu} \frac{1}{Z(\lambda, \nu)}, \quad x = 0, 1, 2, \dots, \quad (2)$$

$$Z(\lambda, \nu) = \sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^\nu}, \quad \text{for } \lambda > 0 \quad \text{and} \quad \nu \geq 0. \quad (3)$$

The dispersion parameter is ν : the Poisson is a special case where $\nu = 1$; where $\nu > 1$ there is underdispersion, and where $\nu < 1$ there is overdispersion.

We fit a Conway–Maxwell–Poisson distribution to the observations for selected periods. Following Shmueli et al. (2005, section 3.2) we adopt an iterative method in which we optimize $Z(\lambda, \nu)$ [Eq. (3)] so that the mean and variance of the fitted distribution [Eq. (2)] converge on and precisely match $\bar{x}$ and s^2 [Eq. (1)]. We obtain the moments of a fitted distribution using the method of Sellers and Shmueli (2010), implemented in the COMpoissonReg package for R (Sellers et al. 2018), to describe its probability density function [Eq. (2)]. We optimize the fit using a Nelder–Mead method from the optimx package (Nash et al. 2020) that we initialize with $Z(\bar{x}, 1)$.

We examine the influence of missing events by inflating the observed series in two different ways. First, we inflate the then-current version of HURDAT by the annual number of missing named storms estimated by Vecchi and Knutson (2008). Second, we inflate the present HURDAT2 using the annual average number of missing tropical cyclones (x_m) estimated by Landsea (2007). We use x_m to estimate the population mean $\hat{\mu}$ and variance $\hat{\sigma}^2$:

$$\hat{\mu} = \bar{x} + x_m, \quad \hat{\sigma}^2 = s^2(\hat{\mu}/\bar{x}). \quad (4)$$

We fit a Conway–Maxwell–Poisson distribution to $\hat{\mu}$ and $\hat{\sigma}^2$ for each period of relative stability in observational practice (defined in section 3a). The distribution for combined periods is a mixture distribution (Wilks 2011, section 4.4.6) that we obtain by weighting the probability density function from each constituent period.

We examine the spatial distribution of the dispersion statistic using the best tracks from HURDAT2 (Landsea and Franklin 2013) for the North Atlantic and from the U.S. Navy's Joint Typhoon Warning Center (JTWC) for the western North Pacific (Chu et al. 2002). We limit the period to that dominated by satellite observations: 1966–2001 and 1972–2006, respectively. This restriction is in order to reduce, as far as possible, the influence of any changes in observational practice. Following a sensitivity analysis that indicated that the pattern of the results was scale independent, but small boxes made the results noisier

(cf. Mumby et al. 2011), we chose to maximize the robustness of the results by calculating counts for large boxes of 20° latitude and longitude. All tracks are included for the western North Pacific, as the JTWC record does not distinguish between tropical and nontropical systems prior to 2004. Only the tropical cyclone portions of tracks are included in the North Atlantic; HURDAT2 observations assigned a nontropical type are excluded. Any qualifying track is counted as passing through a 20° box if a 6-hourly observation with a 1-min maximum sustained wind speed of 34 kt (39 mph; 17.5 m s^{-1}) or higher falls within that box.

b. Forecasts

We build hybrid models from which we make seasonal forecasts of tropical cyclone landfall risk for East China. The predictand t is the observed count of tropical cyclones that transect a latitude–longitude box during summer (JJA). We count all best tracks from JTWC that intersect the box, excluding any that do not satisfy the wind speed criterion (see above) in one or more 6-hourly observations during their lifetime; so both tropical storms and typhoons are included. The box is truncated, relative to the earlier prototype (Camp et al. 2020, Fig. 1), to 27° – 40°N , 119° – 124°E , so as to capture the area of East China most influenced by the WPSH. Since the box includes ocean, some tropical cyclones will be counted whose centers do not pass over land and therefore are not counted in landfall statistics. This choice permits the inclusion of tropical cyclones that had an impact on land but whose centers remained offshore.

The dynamical element of the hybrid forecast is GloSea5, the Met Office Global Seasonal forecast system (MacLachlan et al. 2015), which we use to forecast the westward extent of the WPSH. For each year of operation (2017–20) there is a 42-member forecast ensemble for that year with 2 members initialized on each of 10–30 April. For each year there is also a 28-member hindcast ensemble with 7 members initialized on each of 9, 17, and 25 April, and 1 May. The 9 April members replace the 9 May members used in the earlier prototype (Camp et al. 2020) so as to correspond more closely to the initialization dates used when issuing a real-time forecast. The hindcast spanned 1993–2015 in 2017 and 2018, and 1993–2016 in 2019 and 2020. Following Wang et al. (2013), the WPSH index w is defined as the standardized anomaly of the summer (JJA) ensemble mean 850-hPa geopotential height in a box (15° – 25°N , 115° – 150°E) located where the interannual variability of the WPSH is greatest. We examine the model's performance by comparing its predicted WPSH index with the index from the ERA5 reanalysis (Hersbach et al. 2020).

We relate the predicted WPSH index to tropical cyclone counts by building three alternative types of statistical model:

$$t = k_1 + k_2 w, \quad (5)$$

$$\mu = \exp[k_3 + k_4 w], \quad (6)$$

$$\lambda = \exp[k_5 + k_6 w], \quad \nu = k_7. \quad (7)$$

The first [Eq. (5)] uses simple linear regression, as in the earlier prototype (Camp et al. 2020). The second [Eq. (6)] uses regression to express the Poisson mean μ as a nonlinear function of w (Wilks 2011, section 7.3.3). The third [Eq. (7)] fits a Conway–Maxwell–Poisson regression model by using

TABLE 1. The dispersion statistic (D) from HURDAT named storms, hurricanes, and hurricanes making landfall over the continental United States, for periods of relatively consistent observational practice. The dispersion is calculated for the original record (“HURDAT”), following all corrections in HURDAT2 made to present (“HURDAT2 2019”), following only those corrections made to 2008 (“HURDAT2 2008”), and following inflation of the series as it stood in 2008 (“INFL. 2008”) using the estimates of Vecchi and Knutson (2008). Significant differences from equi-dispersion (1.0) are indicated by a two-tail test at the 90% (*), 95% (**), and 99% (***) levels.

	Period	Named storms	Hurricane	U.S. hurricanes
HURDAT				
Ship	1900–43	**2.15	1.39	1.07
Aircraft	1944–65	**0.64	*0.72	0.80
Satellite	1966–2001	1.15	1.04	1.16
New technology	2002–18	1.48	1.53	2.02
Combined	1900–2018	***2.08	**1.41	1.18
HURDAT2 2019				
Ship	1900–43	***1.90	1.19	1.10
Aircraft	1944–65	**0.65	*0.68	0.79
Combined	1900–2018	***1.79	1.26	1.18
HURDAT2 2008				
Ship	1900–43	**2.07	—	—
Aircraft	1944–65	**0.64	—	—
Combined	1900–2018	***2.00	—	—
INFL. 2008				
Ship	1900–43	**1.84	—	—
Aircraft	1944–65	**0.64	—	—
Combined	1900–2018	***1.75	—	—

the method of Sellers and Shmueli (2010), implemented in the COMpoissonReg package for R (Sellers et al. 2018), to express $Z(\lambda, \nu)$ [Eq. (3)] as a nonlinear function of w . We also build an equivalent model from climatology to provide a baseline against which these statistical models may be compared. The model from climatology is the observed probability distribution of tropical cyclone counts in the East China box from the period used to build the statistical models.

We compare these statistical approaches by building an instance of each type of model for each year to be predicted. For 1993–2016 these use the 2020 hindcast from GloSea5, omitting the year to be predicted, so as to implement a leave-one-out approach (Wilks 2011, section 7.4.4). For 2017–20 these use the full hindcast from each respective year. For 2017–18 these are retrospective forecasts made after the event; for 2019–20 these were real-time forecasts issued on 1 May, using the Gaussian in 2019 (Camp et al. 2020) and the Conway–Maxwell–Poisson in 2020. So each prediction has been made from a model built without knowledge of the outcome of the year being predicted.

We quantify the skill of these models using the ranked probability skill score (RPSS; Wilks [(2011), their Eq. (8.52)]. This measure of skill is on a scale from negative infinity to 1.0. A prediction that copied the climatological probability of each tropical cyclone count would yield a skill of 0.0; a perfect prediction that assigned 100% probability to the number observed would yield a skill of 1.0.

3. Results

a. Dispersion

We reevaluate the dispersion statistic D in the observed records. First we identify the effect of known inhomogeneities

by comparing the dispersion statistics for the original and corrected versions of HURDAT (Table 1). Following Landsea (2007), we have divided the twentieth–twenty-first century record into periods by the year of introduction of new technology to observational practice: ships (1900–43), aircraft (1944–65), satellites (1966–2001), and the new technology (2002–) of advanced microwave sounding, QuikSCAT and cyclone phase space analysis. These were periods in which the technology used to detect and measure tropical cyclones remained relatively consistent. The revisions made to HURDAT were greatest for the era of ship observations, for tropical storms, and for tropical cyclones that did not make landfall (Landsea et al. 2012). Table 1 (“HURDAT,” “HURDAT2 2019”) shows that these corrections reduced overdispersion in the basinwide measures within the ship era, and most substantially for weaker systems. Overdispersion may also be caused by combining periods with differing observational practices. The HURDAT corrections have reduced the effect of the differences in practice between different periods. Table 1 shows that these corrections reduced overdispersion in basinwide events in the “combined” period. These comparisons demonstrate that at least part of the overdispersion found in studies based on the original HURDAT was due to observational inhomogeneities, assuming that the HURDAT2 revisions are correct. This is a valuable insight since departures from equi-dispersion might otherwise be attributed to multi-decadal variability.

We pursue this further by considering the potential effect on the dispersion statistic of missing tropical cyclones. We employ two alternative estimates of missing storms from Vecchi and Knutson (2008) and Landsea (2007). First, Table 1 (“2008”) compares the partly corrected record of HURDAT from 2008 with an inflated version of that record obtained by adding the

estimate of missing named storms from Vecchi and Knutson (2008). The uninflated record from 2008 has dispersion statistics part-way between those for the original HURDAT and the current record. The inflation proposed by Vecchi and Knutson (2008) averages 1.08 in 1900–43 and 0.57 in 1944–65. The inflation has negligible effect during 1944–65, but substantially reduces the overdispersion during 1900–43, and consequently during 1900–2018.

Second, Landsea (2007) estimates that compared to the present day, the observational record from the satellite era is missing on average 1.0 tropical cyclone per year, and the pre-satellite twentieth century is missing 3.2 tropical cyclones per year. In Table 2 we inflate the observed statistics from the current record to add these missing tropical cyclones [Eq. (4)]. In Table 3 we show the effect of adding the missing tropical cyclones on the dispersion statistic over the combined periods. The first column gives the dispersion statistic in the observed record (cf. Table 1 “2019”). The second column gives simulated values obtained by precisely representing the observed distribution of each constituent period, then combining these in a mixture distribution. These D are slightly higher than the observed D from a single calculation (first column) because the variance of a mixture distribution includes a contribution arising from the differences of the mean between the constituent distributions (Wilks 2011, Eq. 4.68). The third and fourth columns also give simulated values, using inflated distributions for their constituent periods (from Table 2), and holding constant in turn the observed D and s^2 . A comparison of these columns with the second column shows that the addition of the missing tropical cyclones reduces overdispersion, most markedly over the full period, but also for the relatively recent period since the introduction of satellites. This demonstrates that at least part of the overdispersion still present in HURDAT2 is due to observational inhomogeneities, assuming that the estimates of missing tropical cyclones are very broadly correct.

We examine the spatial distribution of the dispersion statistic in the North Atlantic basin (Fig. 1). The observed record is particularly vulnerable to inhomogeneity where observational practices have changed. Therefore we have restricted the data to the satellite era in order to reduce the effects of inhomogeneity as far as possible. Figure 1 shows that overdispersion is greatest in the open Atlantic where observations are likely to be least complete. Since the best estimate is that one tropical cyclone per year is missing (Landsea 2007), and since this is most likely to be far from land, part of the overdispersion in the open ocean in Fig. 1 is likely to be due to observational inhomogeneity. Conversely, in the southern United States and Gulf of Mexico where coverage is likely to be most complete, there is equi-dispersion. This is consistent with Table 1, where landfalling hurricanes are equi-dispersed.

We also examine the spatial distribution of the dispersion statistic in the western North Pacific basin (Fig. 2). The period chosen begins with the introduction of satellite information in 1972 and ends prior to a significant change in the wind-pressure relationship used in the JTWC algorithm in 2007 (Knaff and Zehr 2007; Knapp et al. 2013). Unlike in the Atlantic, these observations have not been subject to reanalysis and the ending of reconnaissance flights in 1987 may also have

TABLE 2. The mean and variance for basinwide named storms from HURDAT2: observed ($\bar{x}$, s^2) and inflated ($\hat{\mu}$, $\hat{\sigma}^2$).

	Period	$\bar{x}$	s^2	$\hat{\mu}$	$\hat{\sigma}^2$
Ship	1900–43	8.34	15.86	11.54	21.94
Aircraft	1944–65	11.64	7.58	14.84	9.66
Satellite	1966–2001	10.36	11.89	11.36	13.04
New technology	2002–18	15.18	22.40	—	—

introduced inhomogeneities to the data. Unlike for the Atlantic, the extratropical portions of the best tracks are included; although the results for the northernmost boxes are shown for completeness, these should be treated with caution. The results for the tropics and subtropics show a similar pattern to the Atlantic: significant overdispersion far from land, but not close to land. However, the departures from equi-dispersion are more marked: stronger overdispersion over the ocean, and significant underdispersion where the tropical cyclones make landfall over China. On the basis of the evidence from the Atlantic, where it has been possible to explore the effects of changes in observational technology and practice, and since the same factors are present in the North Pacific, it is reasonable to suggest that much of the overdispersion in the open Pacific may be due to observational inhomogeneities. Conversely the strong underdispersion close to the Asian coasts, where tropical cyclones are best observed, is not easily attributable to inhomogeneities and may be genuine.

The evidence from HURDAT is that a substantial part of the observed overdispersion in tropical cyclone counts is not real, but rather an artifact of changing observational practices. It is conceivable that the real-world numbers may in places be equi-dispersed or underdispersed. If there are dynamical mechanisms that encourage the formation of multiple tropical cyclones in certain years, and prevent them in others, there may be overdispersion. Conversely, if the mechanisms reliably generate a certain number of tropical cyclones each year, there may be underdispersion. Since large-scale patterns of interannual variability have a greater influence on the frequencies of more intense storms, overdispersion may be more likely for more intense storms.

b. Forecasts

Where the observed tropical cyclone counts are underdispersed, seasonal forecasts might be improved by using a distribution such as the Conway–Maxwell–Poisson that captures that underdispersion. We illustrate this by reconsidering the hybrid model adopted in a prototype seasonal forecast of tropical cyclone landfall risk in East China developed by Camp et al. (2020).

The skill of the hybrid model depends first on the skill of the dynamical model in predicting the WPSH index. We compare the WPSH index predicted by GloSea5 on 1 May with the outcome from ERA5 reanalysis (Fig. 3). Similarly to Camp et al. (2019), we find that the GloSea5 hindcast on 1st May predicts well ($r = 0.75$) the interannual variations in the westward extent of the WPSH in summer (JJA).

We then build three alternative statistical models to relate the predicted WPSH index to the observed tropical cyclone

TABLE 3. The dispersion statistic D for basinwide named storms from HURDAT2: 1) observed; simulated by a mixture distribution obtained by combining distributions fitted to 2) $\bar{x}$, s^2 , 3) $\hat{\mu}$, $\hat{\sigma}^2$, and 4) $\hat{\mu}$, s^2 from the constituent periods.

Period	Observed $\bar{x}$, s^2	Simulated $\bar{x}$, s^2	Simulated $\hat{\mu}$, $\hat{\sigma}^2$	Simulated $\hat{\mu}$, s^2
1900–2018	1.79	1.82	1.57	1.33
1944–2018	1.37	1.40	1.32	1.23
1966–2018	1.68	1.71	1.53	1.47

count in the East China box using the 2020 hindcast runs from GloSea5. Figure 4 provides a visual comparison between the three types of statistical model and allows each to be compared to the data used in model building: the WPSH index predicted from the 2020 hindcast from GloSea5 and the observed tropical cyclone counts, both covering 1993–2016. We also plot values for 2017–20 using the WPSH index predicted from the GloSea5 forecasts. A version of this plot, without the 2020 observed value, was produced as part of the real-time forecast issued on 1 May 2020. We present these predictive models by displaying a central prediction and a prediction interval across the range of WPSH values observed. In the linear Gaussian case these are the mean and a 75% prediction interval from a rule-of-thumb approach (Wilks 2011, section 7.2.7). The Poisson and Conway–Maxwell–Poisson models predict integers; so for these we present the median and a prediction interval of at least 75% that is defined by the counts predicted for 12.5% and 87.5% of the cumulative probability.

Figure 5 shows the prediction made using each type of statistical model for 1993–2020 and compares it with the observed outcome. An instance of each statistical model is calculated separately for each predicted year. For 1993–2016, the models are built from the 2020 hindcast, less the year being predicted.

For 2017–20, the models are built from each respective year’s entire hindcast. For example, on 1 May 2020 we predicted a WPSH index of 1.33 for summer 2020. Having built the Conway–Maxwell–Poisson model shown in Fig. 4, we forecast an 87.2% (12.7%) chance of 0 (1) tropical cyclones in the East China box. The observed outcome was 1: Typhoon Hagupit.

The prediction from climatology (Fig. 5: top left) varies a little from year to year, as the year being predicted is omitted from the climatology. The linear regression model, as was used for a real-time forecast in 2019 (Camp et al. 2020), is skilful compared to the prediction from climatology, as the skill score from the hindcast period indicates (Table 4). However, this model employs a linear fit to a nonlinear relationship (Fig. 4: green) and so predicts negative landfall counts (Fig. 5: top right), which is a physical impossibility. The Poisson regression improves on the Gaussian by only predicting positive integers. However, since the observed tropical cyclone counts are underdispersed ($D = 0.83$), the Poisson regression gives a sub-optimal fit, such that the prediction intervals from the model are too wide (Fig. 4: blue). Consequently the predictions made by the model are less sharp (Fig. 5: bottom left), and the forecast skill barely improves on the Gaussian (Table 4). The Conway–Maxwell–Poisson allows dispersion to be modeled.

FIG. 1. The dispersion statistic for the number of tropical cyclone tracks passing through a 20° latitude-longitude box in the North Atlantic per year (1966–2001). The circled number shows the total number of tracks over the period. A black (white) circle indicates a significant difference from equi-dispersion from a 95% (90%) two-tailed test.

FIG. 2. Dispersion statistic for the number of tropical cyclone tracks passing through a 20° latitude-longitude box in the northwest Pacific per year (1972–2006). The circled number shows the total number of tracks over the period. A black (white) circle indicates a significant difference from equi-dispersion from a 95% (90%) two-tailed test.

As Sellers and Shmueli (2010, Fig. 1) found, it provides a better fit to underdispersed data by narrowing the prediction intervals (Fig. 4: red) and sharpening the predictions (Fig. 5: bottom right). Consequently the Conway–Maxwell–Poisson markedly improves the skill of the model (0.47) over the period used to build it. This example demonstrates, therefore, that where the observed tropical cyclone counts are underdispersed, the skill of a forecast may be improved by using the Conway–Maxwell–Poisson.

Table 4 displays the statistical models from Fig. 4 as mathematical expressions and gives their skill scores, calculated from the hindcast period used to build them. A skill score is also given for each model type for the full period, calculated from the predictions shown in Fig. 5. There is a reduction in skill across all types of statistical model when the full period is evaluated, most strongly for the Conway–Maxwell–Poisson. The predictions for 2018 make an important contribution to this. In 2018 the seasonal forecast model correctly predicted a

FIG. 3. WPSH index in JJA from ERA5 reanalysis (black, 1993–2020), predicted from GloSea5 on 1 May using the 2020 hindcast (1993–2016, gray solid) and the 2017–20 forecasts (gray dash). The correlation coefficient (1993–2016, significance level in parentheses) compares the hindcast with ERA5.

FIG. 4. The observed tropical cyclone counts in summer (JJA) in the East China box related to the WPSH index predicted by the GloSea5 hindcast (1993–2016, open circles) and forecast (2017–20, solid circles). The dashed lines indicate the means (1993–2016). Three statistical models using Gaussian, Poisson, and Conway–Maxwell–Poisson regression are fitted to the hindcast data.

negative WPSH index, and consequently the statistical models predicted an enhanced probability of an active year. However, the predicted WPSH index magnitude was not sufficiently large, and 2018 was more active than any year in the hindcast

period used to train the models, and so none of the models predicted so active a year.

One way to view this reduction in skill is as the consequence of enlarging the sample of the distribution of tropical cyclones

FIG. 5. The probabilistic forecast and observed tropical cyclone counts in summer (JJA) in the East China box as predicted by climatology and the Gaussian, Poisson, and Conway–Maxwell–Poisson models.

TABLE 4. Statistical models fitted to the predicted WPSH index from the GloSea5 hindcast (2020 runs) and the observed tropical cyclone counts, the correlation coefficient (r), and the skill score (RPSS) for each model from the hindcast only (“hx,” 1993–2016). The skill score for the full period (“hx + fx,” 1993–2020) includes the predictions made for 2017–20 from equivalent models built using the hindcasts from those years.

Model	Model fit	r	RPSS (hx)	RPSS (hx + fx)
Gaussian	$t = 1.33 - 0.78w$	-0.75	0.38	0.31
Poisson	$\mu = \exp[-0.02 - 0.95w]$	—	0.39	0.34
CMP	$\lambda = \exp[1.56 - 2.65w]$ $\nu = \exp[1.40]$	—	0.47	0.37

that we are attempting to model. Our Conway–Maxwell–Poisson model in Table 4 was built using 1993–2016, when the tropical cyclones were underdispersed ($D = 0.83$). Augmenting this sample with 2017–20 makes it very slightly overdispersed ($D = 1.06$). We are therefore using a Conway–Maxwell–Poisson model tuned to underdispersion to predict for a period that is close to equi-dispersion. While this highlights how any model is dependent on the period chosen to build it, it also highlights how the Conway–Maxwell–Poisson can represent that period more precisely than the Poisson.

4. Conclusions

The Poisson has become well established as the statistical distribution of choice for modeling tropical cyclone counts. Previous studies of tropical cyclone counts have found either equi-dispersion ($D = 1$) or overdispersion ($D > 1$), and have generally elected to model these using the Poisson and the negative binomial distribution, respectively.

However, our review of HURDAT2 in the light of recent reanalysis has demonstrated that at least part of that overdispersion was due to observational inhomogeneities. Moreover, at least some of the residual overdispersion is probably due to inhomogeneities. HURDAT2 has been carefully revised up to 1965, and records from other basins are likely to present similar issues. Conversely, there is some evidence for underdispersion ($D < 1$), particularly in coastal areas. This is where the impacts of tropical cyclones are most severe, and hence where there is greatest interest in skilful seasonal forecasts. Therefore a method is needed that precisely models small numbers of tropical cyclones that may be either under or overdispersed. Poisson regression is likely to be deficient, as it assumes equi-dispersion.

We have illustrated the benefit of using the Conway–Maxwell–Poisson in the case of a prototype seasonal forecast of tropical cyclone landfall for East China. The opportunity to make a forecast arises from the influence of the WPSH on tropical cyclone tracks (Wang et al. 2013) and from the ability of a dynamical model to skilfully predict, some months ahead, the interannual variations in its westward extent (Camp et al. 2019). To maximize the forecast skill requires a statistical model that captures the relationship between the predicted WPSH and the observed numbers of tropical cyclones, which in this case are underdispersed. Least squares regression provides a linear fit that predicts negative tropical cyclone counts, which is a physical impossibility. The Poisson improves on the Gaussian by

predicting only nonnegative integers, but fails to represent the underdispersion and so gives predictions that are too widely spread. The Conway–Maxwell–Poisson represents the underdispersion, sharpens the predictions, and improves the predictive skill from 0.38 to 0.47 over the period used for model building. We have used it for our 2020 seasonal forecast of tropical cyclone landfall risk in East China as part of an evolutionary development to the 2019 prototype (Camp et al. 2020). This science-led change is accompanied by other, user-driven changes to the forecast product, and forms part of our evolution of a prototype toward a useful climate service (Hewitt et al. 2020).

This example demonstrates the benefit of using the Conway–Maxwell–Poisson to model small numbers of tropical cyclones that are underdispersed. We recommend therefore that those developing seasonal forecasts in such circumstances consider using it rather than the Poisson. Indeed, since the Poisson is a special case of the Conway–Maxwell–Poisson, and the latter can represent either under or overdispersion, the Conway–Maxwell–Poisson could be the default choice.

Acknowledgments. Supported by the U.K.–China Research and Innovation Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP) China as part of the Newton Fund.

Data availability statement. All datasets and statistical code libraries used are documented in the article and are publicly available.

REFERENCES

- Camargo, S. J., and A. A. Wing, 2016: Tropical cyclones in climate models. *Wiley Interdiscip. Rev.: Climate Change*, **7**, 211–237, <https://doi.org/10.1002/wcc.373>.
- Camp, J., M. Roberts, C. MacLachlan, E. Wallace, L. Hermanson, A. Brookshaw, A. Arribas, and A. A. Scaife, 2015: Seasonal forecasting of tropical storms using the Met Office GloSea5 seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 2206–2219, <https://doi.org/10.1002/qj.2516>.
- , and Coauthors, 2019: The western Pacific subtropical high and tropical cyclone landfall: Seasonal forecasts using the Met Office GloSea5 system. *Quart. J. Roy. Meteor. Soc.*, **145**, 105–116, <https://doi.org/10.1002/qj.3407>.
- , P. E. Bett, N. Golding, C. Hewitt, T. D. Mitchell, and A. A. Scaife, 2020: Verification of the 2019 GloSea5 seasonal tropical cyclone landfall forecast for East China. *J. Meteor. Res.*, **34**, 917–925, <https://doi.org/10.1007/s13351-020-0043-5>.

- Choi, W., C. H. Ho, C. S. Jin, J. Kim, S. Feng, D. S. R. Park, and J. K. E. Schemm, 2016: Seasonal forecasting of intense tropical cyclones over the North Atlantic and the western North Pacific basins. *Climate Dyn.*, **47**, 3063–3075, <https://doi.org/10.1007/s00382-016-3013-y>.
- Chu, J.-H., C. R. Sampson, A. S. Levine, and E. Fukada, 2002: The Joint Typhoon Warning Center tropical cyclone best-tracks, 1945–2000. Naval Research Laboratory Tech. Rep. NRL/MR/7540-02-16, accessed 27 April 2020, <https://www.metoc.navy.mil/jtwc/products/best-tracks/tc-bt-report.html>.
- Conway, R. W., and W. L. Maxwell, 1962: A queuing model with state dependent service rates. *J. Ind. Eng.*, **12**, 132–136.
- Cry, G. W., 1965: Tropical cyclones of the North Atlantic Ocean: Tracks and frequencies of hurricanes and tropical storms, 1871–1963. U.S. Weather Bureau Tech. Paper 55, 152 pp., http://www.nws.noaa.gov/oh/hdsc/Technical_papers/TP55.pdf.
- Delgado, S., C. W. Landsea, and H. Willoughby, 2018: Reanalysis of the 1954–63 Atlantic hurricane seasons. *J. Climate*, **31**, 4177–4192, <https://doi.org/10.1175/JCLI-D-15-0537.1>.
- Elsner, J. B., 2003: Tracking hurricanes. *Bull. Amer. Meteor. Soc.*, **84**, 353–356, <https://doi.org/10.1175/BAMS-84-3-353>.
- Fisher, R. A., 1950: The significance of deviations from expectation in a Poisson series. *Biometrics*, **6**, 17–24, <https://doi.org/10.2307/3001420>.
- Gray, W. M., 1984: Atlantic seasonal hurricane frequency. Part II: Forecasting its variability. *Mon. Wea. Rev.*, **112**, 1669–1683, [https://doi.org/10.1175/1520-0493\(1984\)112<1669:ASHFPI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1984)112<1669:ASHFPI>2.0.CO;2).
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Hewitt, C., N. Golding, P. Zhang, T. Dunbar, P. E. Bett, J. Camp, T. D. Mitchell, and E. Pope, 2020: The process and benefits of developing prototype climate services—Examples in China. *J. Meteor. Res.*, **34**, 893–903, <https://doi.org/10.1007/s13351-020-0042-6>.
- Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886–1983: Contents, limitations and uses. NOAA Tech. Memo. NWS NHC 22, 24 pp., <https://repository.library.noaa.gov/view/noaa/7069>.
- Knaff, J. A., and R. M. Zehr, 2007: Reexamination of tropical cyclone wind–pressure relationships. *Wea. Forecasting*, **22**, 71–88, <https://doi.org/10.1175/WAF965.1>.
- Knapp, K. R., J. A. Knaff, C. R. Sampson, G. M. Riggio, and A. D. Schnapp, 2013: A pressure-based analysis of the historical western North Pacific tropical cyclone intensity record. *Mon. Wea. Rev.*, **141**, 2611–2631, <https://doi.org/10.1175/MWR-D-12-00323.1>.
- Landsea, C. W., 2007: Counting Atlantic tropical cyclones back to 1900. *Eos, Trans. Amer. Geophys. Union*, **88**, 197–202, <https://doi.org/10.1029/2007EO180001>.
- , and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, <https://doi.org/10.1175/MWR-D-12-00254.1>.
- , and Coauthors, 2004: The Atlantic hurricane database reanalysis project: Documentation for the 1851–1910 alterations and additions to the HURDAT database. *Hurricanes and Typhoons: Past, Present and Future*, R. J. Murnane and K.-B. Liu, Eds., Columbia University Press, 177–221.
- , S. Feuer, A. Hagen, D. A. Glenn, J. Sims, R. Perez, M. Chenoweth, and N. Anderson, 2012: A reanalysis of the 1921–30 Atlantic hurricane database. *J. Climate*, **25**, 865–885, <https://doi.org/10.1175/JCLI-D-11-00026.1>.
- Lexis, W., 1879: Über die Theorie der Stabilität statistischer Reihen. *Jahrbücher f. National-ökonomie u. Stat.*, **32**, 60–98.
- Lord, D., and S. D. Guikema, 2012: The Conway-Maxwell-Poisson model for analyzing crash data. *Appl. Stoch. Models Data Anal.*, **28**, 122–127, <https://doi.org/10.1002/asmb.937>.
- MacLachlan, C., and Coauthors, 2015: Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 1072–1084, <https://doi.org/10.1002/qj.2396>.
- Mooley, D. A., 1980: Severe cyclonic storms in the Bay of Bengal, 1877–1977. *Mon. Wea. Rev.*, **108**, 1647–1655, [https://doi.org/10.1175/1520-0493\(1980\)108<1647:SCSITB>2.0.CO;2](https://doi.org/10.1175/1520-0493(1980)108<1647:SCSITB>2.0.CO;2).
- Mumby, P. J., R. Vitolo, and D. B. Stephenson, 2011: Temporal clustering of tropical cyclones and its ecosystem impacts. *Proc. Natl. Acad. Sci. USA*, **108**, 17 626–17 630, <https://doi.org/10.1073/pnas.1100436108>.
- Nash, J. C., R. Varadhan, and G. Grothendieck, 2020: Package ‘optimx’ version 2020-4.2. Accessed 28 August 2020, <https://CRAN.R-project.org/package=optimx>.
- Nicholls, N., 1979: A possible method for predicting seasonal tropical cyclone activity in the Australian region. *Mon. Wea. Rev.*, **107**, 1221–1224, [https://doi.org/10.1175/1520-0493\(1979\)107<1221:APMFPS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1979)107<1221:APMFPS>2.0.CO;2).
- NOAA, 2020: North Atlantic hurricane basin (1851–2018) comparison of original and revised HURDAT: Revised in June 2019 to include the 2018 season. National Oceanic and Atmospheric Administration, accessed 31 March 2020, https://www.aoml.noaa.gov/hrd/hurdat/comparison_table.html.
- Sellers, K. F., and G. Shmueli, 2010: A flexible regression model for count data. *Ann. Appl. Stat.*, **4**, 943–961, <https://doi.org/10.1214/09-AOAS306>.
- , and A. Raim, 2016: A flexible zero-inflated model to address data dispersion. *Comput. Stat. Data Anal.*, **99**, 68–80, <https://doi.org/10.1016/j.csda.2016.01.007>.
- , S. Borle, and G. Shmueli, 2012: The COM-Poisson model for count data: A survey of methods and applications. *Appl. Stoch. Models Data Anal.*, **28**, 104–116, <https://doi.org/10.1002/asmb.918>.
- , T. Lotze, and A. Raim, 2018: COM-PoissonReg: Conway-Maxwell Poisson (COM-Poisson) Regression version 0.6.1. Accessed 29 November 2019, <https://CRAN.R-project.org/package=COM-PoissonReg>.
- Shmueli, G., T. P. Minka, J. B. Kadane, and P. Boatwright, 2005: A useful distribution for fitting discrete data: Revival of the COM-Poisson. *Appl. Stat.*, **54**, 127–142, <https://doi.org/10.1111/j.1467-9876.2005.00474.x>.
- 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, <https://doi.org/10.1029/2009JD013058>.
- Student, 1919: An explanation of deviations from Poisson’s law in practice. *Biometrika*, **12**, 211–215, <https://doi.org/10.2307/2331767>.
- Thom, H. C. S., 1960: The distribution of annual tropical cyclone frequency. *J. Geophys. Res.*, **65**, 213–222, <https://doi.org/10.1029/JZ065i001p00213>.
- Vecchi, G. A., and T. R. Knutson, 2008: On estimates of historical North Atlantic tropical cyclone activity. *J. Climate*, **21**, 3580–3600, <https://doi.org/10.1175/2008JCLI2178.1>.
- , and G. Villarini, 2014: Next season’s hurricanes. *Science*, **343**, 618–619, <https://doi.org/10.1126/science.1247759>.
- Villarini, G., G. A. Vecchi, and J. A. Smith, 2010: Modeling the dependence of tropical storm counts in the North Atlantic

- basin on climate indices. *Mon. Wea. Rev.*, **138**, 2681–2705, <https://doi.org/10.1175/2010MWR3315.1>.
- Wang, B., B. Xiang, and J.-Y. Lee, 2013: Subtropical High predictability establishes a promising way for monsoon and tropical storm predictions. *Proc. Natl. Acad. Sci. USA*, **110**, 2718–2722, <https://doi.org/10.1073/pnas.1214626110>.
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. . International Geophysics Series, Vol. 100, Academic Press, 704 pp.
- Xekalaki, E., 2014: On the distribution theory of over-dispersion. *J. Stat. Distrib. Appl.*, **1**, 19, <https://doi.org/10.1186/s40488-014-0019-z>.
- Zhang, W., G. A. Vecchi, G. Villarini, H. Murakami, R. Gudgel, and X. Yang, 2017: Statistical-dynamical seasonal forecast of western North Pacific and East Asia landfalling tropical cyclones using the GFDL FLOR coupled climate model. *J. Climate*, **30**, 2209–2232, <https://doi.org/10.1175/JCLI-D-16-0487.1>.