COMP4802 Extended Final Year Project

Project Plan

Name: FONG Kwan Ching
UID: 3035564902
Project title: A Study on Forecasting Tropical Cyclone Properties
Supervisor: Dr Beta C. L. Yip

Table of Contents

Chapter 1: Introduction

The term tropical cyclone (TC) refers to storms forming over tropical seas, the stronger categories of which are called typhoons in Eastern Asia [1]. On average, 16 TCs appear within the area of responsibility of the Hong Kong Observatory (HKO) every year, among which six affect Hong Kong, necessitating the issuing of TC warning signals [2, pp.34-35]. The adverse weather such TCs bring about oftentimes disrupt Hong Kong citizens' daily activities such as commuting, upset economic activities, and in rare cases, result in casualties [3]. Thus, an accurate TC forecasting method and a reliable TC warning system must be in place, to truthfully inform the general public of an incoming TC's threat and potential impact. This final year project (FYP; "this project" henceforth) intends to explore the ways to evaluate the probability of a TC affecting Hong Kong and the corresponding extent of the impact and envisions the development of a forecast model for such purpose.

1.1 Prior Research in the Field

N.B.: Please see Chapter 2 for a comprehensive literature review.

Over the previous decades, numerous TC forecasting techniques were devised, and they significantly progressed over the years. On one hand, manual methods relying on the experience and knowledge of human forecasters have always been in use; on the other hand, objective methods involving statistics and atmospheric physics were readily being developed over the years [4]. The earliest statistical methods such as climatology and persistence (CLIPER) and the Veigas-Miller method [5] are now no longer operational, as they were replaced by numerical weather prediction (NWP; also referred to as "dynamical" methods) models that simulate the atmosphere [4] and provide detailed computerized analyses. NWP models, being the key advancement in weather forecasting, are now the state-of-the-art means [6] in TC track and intensity forecasts.

In recent years, interest arose in creating probabilistic models that, instead of giving deterministic values, provide a probability of an event, e.g. the possibility of a TC moving to a certain location (see [7]), or more complicatedly, the likelihood of gale-strength wind occurring at a location (see [8], [9]). Probabilistic forecasts are superior to deterministic ones because they account for the inherent uncertainty of TC behaviour [10]. Apart from track and intensity forecasts, strike risk forecasts have also been devised to evaluate and mitigate the impact of TCs, in particular on a season-wide basis. Examples of such forecasts include [11], [12] and [13]. Besides, the introduction of randomization via Monte Carlo methods opened new opportunities for probabilistic TC forecasts, an example of which being the TC wind field forecast by [8].

Combinations of statistical and dynamical methods are also an interest of modern research, with machine learning oftentimes coming into play as well. For example, [14] combined the statistical CLIPER method with dynamical atmospheric data to forecast TCs in the Southern hemisphere; while [15] summarizes the wide array of machine learning applications developed to augment preexisting forecasts. In short, there is a huge arsenal of tools at meteorologists' disposal nowadays, thanks to the research done heretofore.

It is observed that TC strike probability forecasts are not new, as the HKO developed one in the 1970s by statistical means [16]. However, it is noted that there is only one operational model that predicts the probability of a specific TC affecting Hong Kong, namely [17] which provides probabilities of a TC reaching certain distances from Hong Kong.

Strike Probability for **Hong Kong** regarding **14W DUJUAN**
14W DUJUAN 相對於 **Hong Kong** 之機率分佈表

Date/Time (UTC)	Current tau	Forecast tau	Lat	Long	Distance (km)	100 km			200 km			400 km			800 km			Strength kt / kph
						T	I	C	T	I	C	T	I	C	T	I	C	
Aug 31, 18:00	-12 hr	+0 hr	20.40	126.70	1313	<1	--	--	<1	--	--	<1	--	--	<1	--	--	100 / 185
Sep 01, 00:00	-6 hr	+6 hr	20.65	125.20	1155	<1	<1	<1	<1	<1	<1	<1	<1	<1	<1	<1	<1	103 / 190
Sep 01, 06:00	+0 hr	+12 hr	20.90	123.70	997	<1	<1	<1	<1	<1	<1	<1	<1	<1	<1	<1	<1	105 / 194
Sep 01, 12:00	+6 hr	+18 hr	21.20	122.10	827	<1	<1	<1	<1	<1	<1	<1	<1	<1	37	37	37	105 / 194
Sep 01, 18:00	+12 hr	+24 hr	21.50	120.50	659	<1	<1	<1	<1	<1	<1	1	1	1	86	87	87	105 / 194
Sep 02, 00:00	+18 hr	+30 hr	21.80	118.90	490	<1	<1	<1	1	1	1	22	22	22	98	99	99	95 / 176
Sep 02, 06:00	+24 hr	+36 hr	22.10	117.30	323	3	3	3	14	14	14	58	59	59	**100**	**100**	100	85 / 157
Sep 02, 12:00	+30 hr	+42 hr	22.35	115.80	168	9	9	9	31	32	32	78	80	80	**100**	**100**	100	80 / 148
Sep 02, 18:00	+36 hr	+48 hr	22.60	114.30	36	**10**	**11**	12	**33**	**35**	37	**80**	**82**	85	**100**	**100**	100	75 / 139
Sep 03, 00:00	+42 hr	+54 hr	22.75	113.05	125	7	8	14	25	30	44	69	76	95	99	**100**	100	69 / 127
Sep 03, 06:00	+48 hr	+60 hr	22.90	111.80	252	4	6	16	16	21	51	52	61	100	96	98	100	63 / 116
Sep 03, 12:00	+54 hr	+66 hr	23.05	110.55	380	2	3	17	10	13	56	36	44	100	88	92	100	56 / 104
Sep 03, 18:00	+60 hr	+72 hr	23.20	109.30	509	1	2	18	6	7	59	23	29	100	74	81	100	50 / 93
Sep 04, 00:00	+66 hr	+78 hr	23.05	108.30	608	1	1	19	4	5	61	16	20	100	62	68	100	43 / 79
Sep 04, 06:00	+72 hr	+84 hr	22.90	107.30	708	1	1	19	2	3	63	11	14	100	50	56	100	35 / 65
Sep 04, 12:00	+78 hr	+90 hr	22.75	106.30	809	<1	1	19	2	2	64	8	10	100	39	44	100	28 / 51
Sep 04, 18:00	+84 hr	+96 hr	22.60	105.30	911	<1	<1	**19**	1	1	**64**	5	7	**100**	30	34	**100**	20 / 37

Figure 1.1: Strike probability forecast sample from [17]

Figure 1.1 shows the sample output of [17]. The forecast was made on 31st August 2003 for Typhoon Dujuan, and the charts were valid on 1st September 2003. The forecasted track, distance from Hong Kong, and intensity (strength) of the typhoon were provided, along with four tables detailing the strike probabilities. Each table represented the probability of the TC entering a certain radius of Hong Kong. The columns, from left to right, stood for instantaneous, intervallic, and cumulative probabilities respectively. Intervallic probability meant the likelihood of the TC is in the radius during that 6-hour interval and cumulative probability evaluates the likelihood from start to that forecast hour. For example, the row labelled “Sep 02, 18:00” says that there is a 10% chance the TC would be instantaneously in a 100 km radius from Hong Kong, 11% chance in the same area within the six hours (18:00 to 00:00), and the chance of the TC being in this area between Aug 31 18:00 and Sep 03 00:00 is 12%. As seen in the tables, this forecast was very thorough indeed.

However, the basis of the forecast cannot be verified because the paper was lost. Thus, there exists room for additional research and development.

1.2 Problem Statement and Proposed Solution

This project explores and tries to solve a dual problem. Firstly, despite its value (see the next section for more details), there is only one operational model that forecasts a TC’s chances of affecting Hong Kong (also termed “strike probability”); secondly, the said model cannot be verified, making it difficult for future research and development to build thereupon.

The primary product of this project is a TC forecasting tool for the general public of Hong Kong which needs a correct and easy-to-interpret strike probability forecast. The main benefit of the tool, in general, is that it flexibly evaluates the strike probabilities, using four predictands corresponding to minimal, limited, substantial damage and direct TC strike. This shall be the principal objective of this project.

As an advantage over [17], this project intends to produce a verifiable tool that builds upon the latest research, by making its methods and results available. This way, the amateur enthusiast and the professional research communities will also benefit from this product in principle, because it is a new asset for TC forecasting and assessment that is verifiable with relative ease.

The tool will consist of the forecasting models at its core. The models will be statistical-dynamical and perform statistical analyses on best track datasets, NWP-generated atmospheric data, and HKO warnings issuance records. Three such models are envisaged, the first being a naïve model serving as a baseline for skill verification and the other two contributing to the output as a small-scaled ensemble. Different methodologies for these forecast models will be explored iteratively.

It is expected that with each model (ensemble member) built, the error decreases and the relative skill increases. As there exists no directly comparable literature, there will, unfortunately, be no objective basis to evaluate the project's technical success.

1.3 Justification of Project Feasibility and Distinctiveness

It is believed that the fabrication of a new TC strike probability model specifically for Hong Kong through this project is feasible and valuable. There are several reasons why:

To begin with, models that evaluate the probability of a TC affecting Hong Kong are important because they can give the general public an intuitive but correct means to assess the threat posed by TCs, and minimize losses caused by improper interpretations of the forecasts. This also helps to evaluate the associated adverse weather such as storm surges [18]. This can also augment preexisting deterministic forecasts by introducing a new point of view from which the unpredictable nature of TCs is considered [10].

The proposed model will not be redundant despite the existence of [17]. Firstly, it provides extra value compared to [17], because this model will be easier to validate and devise improvements. It is intended that the results, methods, and references of this project be published publicly, such that future interested parties will have a way to verify everything related to this project, unlike [17] which has unresolvable references to literature and provides no information about the algorithm used. Secondly, this project plans to use a different methodology (see Chapter 4. Note: there is no certain way to ascertain that the methodologies are indeed different.) developed based on the author's understanding of contemporary research, thus this project will produce substantially different results.

The proposed model also does not directly conflict with the HKO's TC track probability forecast (namely [7]), because this project plans to evaluate the likelihood of a TC affecting Hong Kong, instead of providing general track forecasts. The focuses of the two systems are drastically different and are valuable in different ways.

To summarize, this project is an exploratory project that envisions a statistical-dynamical TC strike probability tool to help assess TC impact.

The remainder of this project plan is organized as follows: Chapter 2 is the full literature review on the evolution of TC forecasting methods and research on strike and wind probabilities. Chapter 3 is a discussion about this project's objectives, scope, expected outcome, and intended deliverables. Chapter 4 discusses the methodologies, tentative technical plans, and anticipated problems of the project in-depth. Chapter 5 focuses on the project schedule and milestones. Chapter 6 summarizes the project plan and is followed by the Appendices.

Chapter 2: Literature Review

This chapter looks into the research on TC forecasting over the last few decades. Emphasis is placed in particular on probabilistic forecasts, an example of a purely statistical method called CLIPER, and novel techniques. All summaries and comments on each source are tabulated in Appendix 1.

Chin [5] summarized the subjective and objective techniques used at the HKO in the late 1970s when only 24-hour forecasts were produced and observation data were relatively scarce. The comparatively primitive methods were replaced by new technology, most notably NWP, in the subsequent years as summarized by Elsberry [6].

The following is a categorized review of such new technologies and their relevance to this project:

2.1 Statistical Method: CLIPER

CLIPER was developed by Neumann in 1972 to forecast TC tracks. The method combined climatology (trend observed in historical data) and persistence (extrapolating the TC's velocity) to produce forecasts [19]. It is now used as a baseline against which forecast models are validated [20]. Its legacy is significant, as many similar models bear the term "CLIPER" in their names, such as [21] and [22]; and it serves as the methodological basis for other models, including its successor [23], a wind field forecaster [24] (see also [25]), an intensity model for the Southern hemisphere [26], and many others.

The original CLIPER model did not consider atmospheric factors and relied solely on historical tracks to produce its climatological segment [19]. However, with correct assumptions and factors (predictors), CLIPER-like models can be robust options. For example, [24] models TC wind circulations using a Rankine vortex instead of blindly regressing historical data, thus achieving reasonable success.

2.2 Dynamical Method: NWP

TC track probability products generated by NWP models, which simulate the whole atmosphere following its physical rules, are relatively common. For instance, the European Centre for Medium-Range Weather Forecasts (ECMWF), one of the leading institutions in NWP, generates such forecasts [27] using an ensemble [28], i.e. an NWP model is run multiple times and the outcomes are collectively considered to produce the output [10]. [10] and [29] both suggest that ensembles yield better results than individual models; whereas [30] and [31] discussed the strategies to select the best performing ensemble members to reach consensus.

While NWP models may not be directly available to this project, dynamical data can be extracted from NWP analyses and be employed in conjunction with other data sources. For example, the ECMWF [32], Global Forecasting System (GFS) [33], and the NCEP FNL Operational Model at the National Oceanic and Atmospheric Administration (NOAA) in the US [34] all provide datasets for public use. Also, the ensemble approach can be used to obtain more comprehensive forecasts.

2.3 Statistical Method: Machine Learning

Over the last decade, machine learning became commonplace in TC forecasting. Chen et al. give a complete picture of using machine learning to forecast and study various aspects of a TC, ranging from tracks and intensities to wind fields and storm surges [15]. The studies by Zhang et al. analyzed TC recurvatures (making sharp turns) and landfall patterns using data mining, achieving promising results [35-36]; while Tan et al. built an ensemble forecast using the machine learning technique of decision trees along with CLIPER predictors and meteorological data [37]. In addition, probabilistic forecasts of wind fields [38] and damage estimation [39] are both possible using machine learning methods.

It is noteworthy that the majority of the abovementioned studies are hybrid in nature: they combine statistical methodologies provided by machine learning with dynamical data obtained via NWP. This enables the underlying meteorological rules to be modelled and avoids blindly modelling the superficial outcomes as shown in historical data.

It should also be noted that many of these studies used relatively small datasets, which make them prone to overfitting and thereby losing flexibility and accuracy facing unseen data.

2.4 Specific Models for Strike Probabilities or Damage Estimates

This section goes over the research that particularly targets strike probabilities and estimation of TC damages.

The study by Devaraj et al. [39] treats TC damage as a classification problem, such that each building in a satellite image is labelled “damaged” or the opposite. There are numerous more complicated approaches, such as that by Wang et al. which evaluates TC damages in Hong Kong in terms of economic loss [11], [40]; computing the season-wide TC occurrence probabilities for a region using climatological and dynamical factors [13], [41]; using Monte Carlo methods to estimate the chance of TCs hitting specific locations over a year [12], and calculating TC occurrence probability for select cities over certain periods [16].

None of such research, with the notable exception being that referenced by [17], particularly focuses on finding the probability of a TC affecting (or more strictly, striking by making landfall) a specific place. Therefore, the field this project looks into is presumably new and unique.

2.5 Miscellaneous Methods and Techniques

To close this chapter, the remaining relevant methods found are summarized below.

Several techniques were devised to forecast winds brought by TCs. The HKO has two manual techniques called “kidney” and “beach ball” to estimate wind speeds and directions in Hong Kong [42], but it does not scale to other situations. DeMaria et al. at the NOAA devised a technique [8] to forecast TC wind fields, combining the CLIPER way of regressional analyses [24] and the Monte Carlo method to generate TC track and intensity information. It was further validated and updated by [43-45]. Randomization also has an important role in [46], where the frequencies of global TC wind speed occurrences and strike events are estimated using a dataset worth “10000 years of data”. Winds observable at specific weather stations can also be forecasted using statistical methods as described in [47], which notably mapped the winds and TC location information on a polar coordinate plane.

The paper by Kim et al. [48] describes a statistical-dynamical method to evaluate the climatology of seasonal TC activity. Hybrid models combining CLIPER with dynamical predictors are developed to forecast cyclones of the Southern hemisphere [49].

Finally, a paper by Heming [50] provides several useful measurements to evaluate TC forecasters. For example, by taking the improvement over simple CLIPER in forecast error as a percentage, the forecast skill of the new model can be computed.

The above shows that the inclusion of randomness via Monte Carlo methods ([8]) and data augmentation ([46]) opens new possibilities for TC forecasting, and statistical-dynamical models are enjoying increasing popularity.

Chapter 3: Project objectives

The objectives of this project are summarized in the following full vision statement:

For the general public of Hong Kong which needs an accurate and intuitive means to assess the probability of a TC affecting Hong Kong, and for the TC research community which needs a new verifiable perspective to assess TC threat, the product of this project ("the Model") is a statistical-dynamical ensemble forecast that produces separate probabilities of a TC resulting in minimal, limited, substantial levels of damage or a direct strike at Hong Kong.

Unlike its direct competitor [17], this Model will provide documentation that does not reference literature that can not be found and directly outlines the actual methods and algorithms instead, such that the results are transparent and verifiable. The Model is also unlike [7] because this Model evaluates strike probabilities in a Hong Kong-centric manner, rather than giving general track forecasts.

By "separate", it is meant that the probabilities of each category (e.g. minimal vs limited) are evaluated independently.

This project is also exploratory in nature. That is, this project will have several methodologies to investigate, and specifics may change over time in response to the latest results and discoveries. On the whole, the project plans to iteratively deepen the analyses and methodology.

3.1 Scope and Deliverables

This project has the following major features, each mapping to a stakeholder mentioned above:

1. An ensemble forecasting model is built to evaluate TC strike probabilities for Hong Kong via four predictands.
2. Verification of Model performance and methodology is possible, easing future development.

The former is prioritized because the latter is achievable simply by releasing all documentation of the model with the Model itself.

Further breaking down the two features, the project has the following deliverables, excluding the formalities dictated by the FYP guidelines:

- The Model itself, which minimally shall consist of
  - A baseline statistical model for evaluation purposes
  - Two statistical-dynamical models that serve as ensemble members
  - Other necessary modules that make the above work together, i.e. glue code
- Technical documentation outlining the Model's methodologies
- Report(s) on the outcome of experimentations during the Model's development

Note the latter two may be covered by the obligatory FYP reports. The details of the Model's high-level design and the forecasting modules' methodologies will be discussed in the next chapter.

The list of artefacts may be revised in the future, depending on the project's progress speed.

In the next chapter, viable methodologies that can be tried as ensemble members will be listed and ordered by priority, wherefrom two will be selected to build the ensemble members. There is no guarantee that the chosen methodologies will not be replaced in case of major technical difficulties. It cannot be assured either, that extra ensemble models will not be introduced in the event excess resources are present. The realization of all viable methodologies will thus be outside of the scope of this project.

The presentation of the Model's output will minimally be text-only. Production of forecast graphics and a more accessible user interface will only be considered if the progress of the project allows so.

3.2 Limitations and Exclusions

This section discusses the items that certainly lie beyond the scope of this project, regardless of future revisions.

Purely dynamical methodologies will not be considered. This is because running NWP forecasts require a large number of actual weather observations as input, which are unavailable to this project. Instead, this project will employ data analyzed by authoritative NWP models as its dynamical component.

The project on the whole has no objective quantitative metric of success, as a part of its nature. For instance, it is impossible to compare the Model's performance with that of [17], because one cannot run historical datasets through [17] to collect performance statistics.

Lastly, regardless of the project's outcome, the produced forecasts should be used for reference only, and they are by no means capable of standing in as an alternative to authoritative forecasts and advisories by the HKO. This project overlooks numerous technical intricacies and makes assumptions to simplify the problem, thus its results should not be wholly trusted.

Chapter 4: Project Methodology

This chapter discusses the project methodology in depth. The high-level design, definitions and considerations will be outlined. After that, several methods that this project intends to examine will be described, along with an analysis of the problems expected to arise.

4.1 Problem Definition and Solution Approach

The problem at hand, generating TC strike probabilities, can be rephrased in multiple ways, depending on the exact method used to tackle it. However, on the most empirical level, it can be defined as follows:

4.1.1 Explanation of the Predictands

The four predictands are probability values corresponding to the following four situations:

- The TC in question meets the minimal requirements for affecting Hong Kong, which is equivalent to TC signal number 1. This roughly means a TC is within an 800km radius of Hong Kong. [51]
- Strong winds at speeds of 41-62 kmph blow continuously in Hong Kong, roughly equivalent to TC signal number 3 [51]. This may be interpreted as "limited impact should be expected".
- Gale, storm, or hurricane winds of speeds 63 kmph or more. This roughly translates to TC warning signals numbers 8 to 10 [51]. This may be interpreted as "substantial impact should be expected".
- The TC is expected to pass through the circle with a 100 km radius and centred at Hong Kong, i.e. a direct strike.

These forecasts will be **valid for up to three days**. In other words, the probabilities evaluate the likelihood of the events occurring within the upcoming 72 hours and no further. This is to account for uncertainties in the long run when the forecasts become too inaccurate to use and to give a more realistic picture for TCs that are still far away from Hong Kong.

There are several special considerations regarding these four predictands.

It may seem natural to create two predictands for TC warning signals number 9 and 10 respectively, however, it is more reasonable to subsume them into that of signal number 8 (gale-force winds). This is because TC signal number 9 has no objective rule of issuance [51] and number 10 has too few occurrences. In addition, it is unlikely that Hong Kong will see widespread hurricane winds (signal number 10's level, wind speed > 117 kmph [51]) given its complicated terrain and high level of urbanization, as the average wind speed in Hong Kong decreased significantly over the past decades [52].

The fourth predictand, direct strike likelihood, is valuable because a TC's most intense winds and rainstorms are clustered around its centre, so the closer a TC's centre to Hong Kong is, the more damage can be expected. This compensates for the lack of equivalent measurements for signals number 9 and 10 when the user considers both the "substantial impact to expect" case and this "direct strike" case.

Note that the final predictand does not translate to wind strengths and damage levels, and the aforementioned three situations should be taken into consideration. For instance, a weak TC can pass through Hong Kong itself, causing little damage and disturbance, like the one on 18 June 2000 [53, pp. 42-49]; but a major typhoon can lead to a TC signal number 10 while passing by Hong Kong at a distance of just above 100 km, like Typhoon Vicente of 2012 [54, pp. 68-83]. In other words, the predictands are not independent.

4.1.2 Statistical-Dynamical Approach

The solution will use a statistical-dynamical approach. There are many ways to realize such a hybrid approach, such as selecting dynamic factors as inputs to a statistical model or using the output of NWP as inputs to statistical models [15]. This project plans to combine best track data, records of historical TC positions and intensities, with a select subset of dynamic NWP observation data in statistical models, thereby approximately fitting into the definitions of statistical-dynamical methods.

The solution requires a baseline model as a control, against which the other models which will become ensemble members can compare. In practice, the baseline is typically CLIPER [20], but because of the resource constraints of this project, CLIPER-like methods may be useful as an ensemble member, therefore a different baseline will have to be found. To that end, a naïve machine learning model that only considers the best track data will be used and models that consider extra predictors and factors will be compared to it.

An ensemble is used because a more comprehensive forecast considering all viable (non-baseline) models can be obtained this way. The whole forecasting system will be more robust and be less prone to errors in any one of the member models. The weights of the member votes depend on a couple of factors, namely:

- The amount and range of input data used,
- How well does the method used represent the reality, and
- Relative skill of the model with respect to the baseline.

The input to each model, ensemble member or baseline, can vary and will be described in their respective sections. The output values are the four predictands described above.

4.1.3 Datasets for Best Track and Meteorological Data

The source of best track data will be chosen from among the dataset of the Joint Typhoon Warning Center (JTWC) [55], the *IBTrACS* dataset by the NOAA [56], and the best track dataset of the China Meteorological Administration (CMA) [57-59]. To decide which one should be the final choice, these criteria will be used:

- The spatial resolution of TC positions, i.e. how accurate the positions are,
- Record density, i.e. the duration between two consecutive records for the same TC,
- Time coverage, i.e. what are the years this dataset covers, and
- Ease of use, i.e. the amount of user support provided and the amount of irrelevant information that need to be cleaned up.

The necessary fields in these datasets include TC positions, intensities in terms of wind speed or pressure, and timestamps of the records.

The source of meteorological data extracted from NWP analyses, i.e. the data that will contribute to the dynamical part, will be chosen between the *NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999* [34] and the *ECMWF Reanalysis v5 (ERA5)* dataset [32]. More work is to be done to determine which one is better suited for this project. The metrics of consideration include:

- The comprehensiveness of the dataset, i.e. whether useful analyses such as surface wind velocities and geopotential height values are included,
- Grid resolution, i.e. the density of the observations,
- Time coverage,
- Ease of dataset access, i.e. dataset format and access restrictions by the provider, and
- Ease of obtaining the latest data, i.e. how to obtain the latest analyses from the same NWP model.

4.1.4 Input Data Considerations

The datasets have to be transformed into a dataset, containing in the right form strictly the data the system will need.

Each record in the transformed best track dataset should therefore contain:

- A timestamp consisting of the date (month and day) and time,
- A list of best track position and intensity data, recorded at the time above and within the prior two days,
- A list of meteorological data dated within the past two days, and
- The corresponding ground truth, i.e. the actual outcome associated with it.

Metadata that identify this sequence as a particular historical TC will not be needed, to avoid the models accidentally treating them as predictors.

The meteorological data of interest can also differ from model to model, especially if the experimentation reveals that some predictors turn out to be more useful than others. The following is a tentative list of interesting factors, referred to using their names in [34]:

- Geopotential height
- Surface winds
- Upper-level winds
- Convection
- Total precipitable water
- Sea surface temperature

These data will be available as grids, which can be imagined as the analyses plotted on a world map, with each data point every 110 km (1-degree latitude) or so. To feed only the relevant data into the models, further processing is necessary, such that the data's principal characteristics are summarized as an index (like the "400hPa subtropical high-intensity index" mentioned in [36]). That is, the meteorological data must be compressed to enable efficient processing. Also, only a portion of the grid will be needed because this project solely concerns Hong Kong and its neighbouring regions. More research has to be done to find an effective summarization of the meteorological data.

For all data with timestamps outside of the intersection between the best track data's time range and that of the NWP data, those will not be used when the other is needed. In other words, for each best track data record, if a corresponding NWP data point of roughly the same time cannot be found and NWP data is needed for the model, then this data record cannot be used.

Last but not least, the records have to be labelled with the relevant TC warning signal issuances. The source information will be obtained from the HKO's website. Following the assumption that warning signals are directly linked to threat and wind speed levels (see section 4.1.6), each TC best track record can be matched to the actual outcomes. For example, Typhoon Vicente of 2012 resulted in a TC signal number 10. By retrieving from the HKO the periods during which different warning signals were effective, each record of Vicente could be correspondingly labelled as "causing signal number 1 within the next 3 days", "causing signal number 3 within the next 3 days", etc. The ground truth for "direct strike" can be directly calculated using the coordinates of Hong Kong.

4.1.5 Metrics of Performance

A consistent metric to measure and compare the performance of the ensemble member models is needed, such that improvements can be made and the weighting between members can be decided.

At the moment, the F1-score is the planned metric to measure model performance. It takes both precision, how many true positives there are among positives, and recall, how many true positives are identified whenever they should be, into consideration [60]. For each of the four predictands, the F1-score will be computed and with that, the performance of the models in each category can be compared.

To tell positives (hit) apart from negatives, one also needs a threshold for the probabilities, above which it will be regarded as a positive (hit) and a negative otherwise. At this stage, it is surmised that using 0.5 is suitable.

It is also noted that the need to calibrate the outputs may arise. That is, the probability distribution found by the models is not well-aligned with the actual one [61], such that it tends to over or underestimate the probabilities. The need to calibrate the predictions will be evaluated on a case-by-case basis during the development of the models.

Lastly, a scoring rule may alternatively be used to assess the accuracy of the probabilistic predictions [62]. It is foreseeable that the implementation of a scoring rule to replace the F1-score will be troublesome and needs more thought, but if needed, the logarithmic score will be chosen due to its simplicity and adequate calibration performance, as illustrated in [63].

4.1.6 Assumptions

To simplify the problem such that it can be tackled using data and techniques available, several assumptions are made.

Firstly, it is assumed that Hong Kong is a point object and thus any wind speed estimates made will represent the average wind speed in Hong Kong. That means the exact rules of TC warning signal issuance are neglected. This is because it is impossible to model all weather stations in Hong Kong and analyze their data to obtain a truly illustrative model representing the actual wind intensity in Hong Kong.

Secondly, this project is largely built upon the TC warning system of the HKO. The first predictand (minimal impact) is equivalent to TC warning signal number 1, the second to signal number 3, and the third to signal numbers 8-10. This is because the warning system is assumed to be directly relevant to the actual impact a TC may bring upon Hong Kong and the signals genuinely represent the threat level. As a consequence, this project assumes that if the HKO hoists a certain signal representing some wind speed (e.g. number 8 for gale and storm), then it is true that such winds are observable in Hong Kong on average (see the first assumption).

Thirdly, it is assumed that the climatological patterns of TCs do not change over time, or at least it did not change in the timeframe the used data were collected. Otherwise, the statistical patterns the models discover will not truthfully represent the past patterns that manifested and shifted. However, climatological patterns do change over time, for example, [12] discovered that the average TC landfall position in China never stayed put.

4.2 Overall Model Design

The following describes the components of the TC strike probability model ("the Model") proposed and its supposed workflow.

4.2.1 Components and Structure

The following diagram illustrates the structure of the Model. Note that the internal designs and subcomponents of each ensemble member may vary.

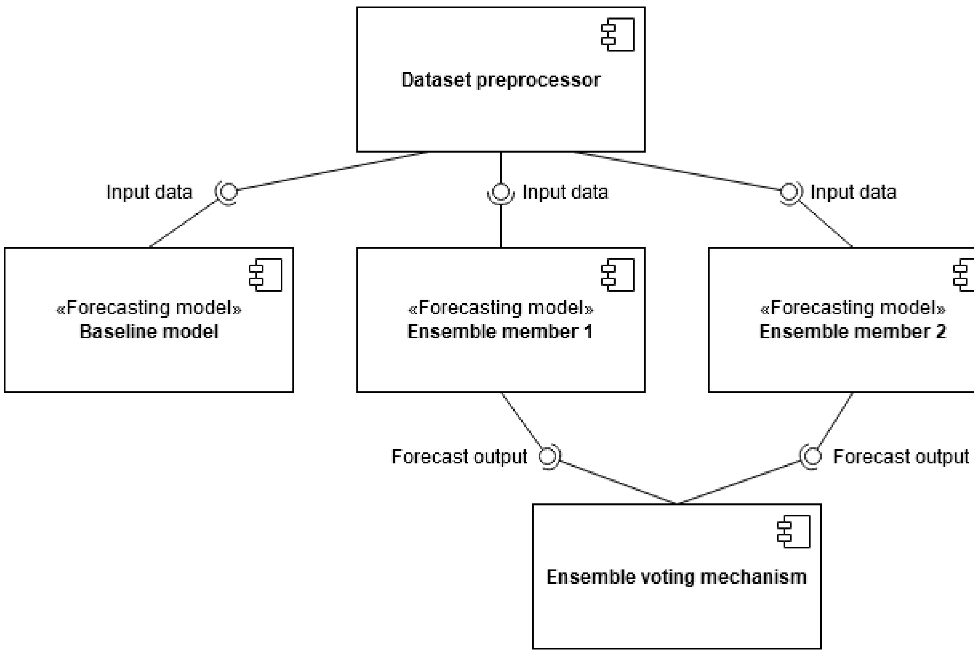

Figure 4.1: Forecasting system major components diagram

In Figure 4.1, the whole Model consists of five major components. There is a data preprocessor that prepares data from the datasets for the ensemble member models' use, three actual forecasting models, and the ensemble voting mechanism that gathers the output of the two ensemble members and performs the weighting. As mentioned before, the output of the baseline model will be neglected.

In practice, the system will have more components. There will be extra utilities that evaluate and compare the performance of each model, additional programs that fetch data from the sources for the preprocessor, and a command-line interface.

4.2.2 Ensemble Member Development Workflow

The following figure outlines the process of developing an ensemble member.

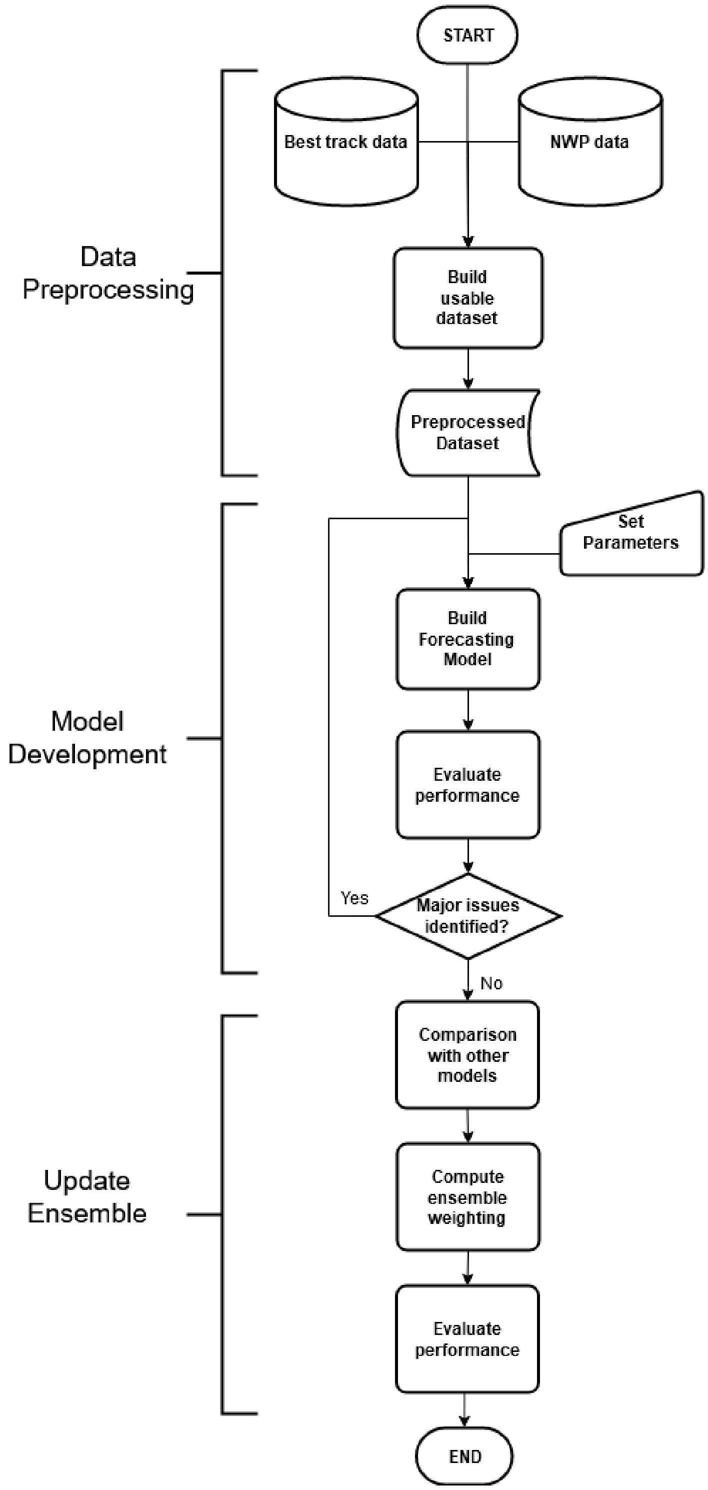

Figure 4.2: Ensemble-member development process flowchart

Figure 4.2 shows that the process has three steps. Firstly, the datasets are processed to fit the requirements described in section 4.1.4. The resultant dataset will then be used to build a forecasting model; the methodology may require some manually configured parameters. After that, the model is evaluated using the principles outlined in section 4.1.5. If the results are satisfactory and no major issues are found, then the model will be fit for admission into the ensemble. The weighting will be decided after considering the new model's specifics. Finally, the new ensemble's performance will be evaluated.

4.2.3 Ensemble Forecast Workflow

The workflow of the Model in deployment is different, as illustrated below:

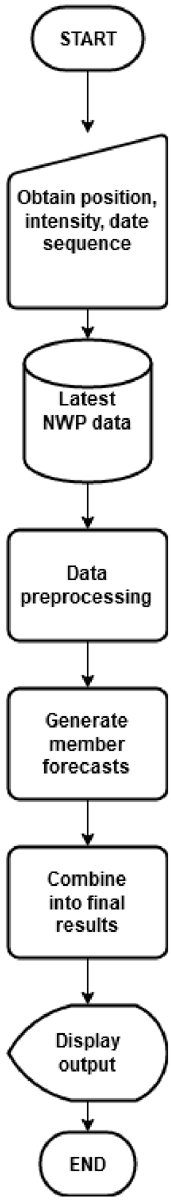

Figure 4.3: Ensemble forecast procedure flowchart

As shown in Figure 4.3, the procedure now consists of six steps. First of all, as the TC to forecast will not be in the best track dataset, the corresponding sequence of variables such as position, intensity, and date will be entered manually. Then, for dynamical data, the latest NWP analysis data will be obtained online from the same institution that provided the dataset. This is slightly risky because the quality and availability of the data can be different from before. The third step is to process these new data into formats the models can accept. After that, the data is given to the two ensemble members, who generate their forecasts that will next be weighted in the ensemble voting process. The final results will be output to the user via the command-line interface.

4.3 Methods for the Forecasting Models

4.3.1 Baseline Statistical Model: Multilayer Perceptron

A multilayer perceptron is a simple neural network consisting of input, output layers and at least one intermediate hidden layer. It is the basis of all artificial neural networks and is reasonably versatile, thanks to the nonlinear activation functions connecting neurons of different layers [64].

For machine learning methodologies like this and decision trees (section 4.3.2), the problem described in section 4.1 will be reframed as a classification problem. For each given input data record, the model should assign labels (classes) to it, but the classes are not necessarily mutually exclusive. In other words, the input will be classified as “minimal damage”, “limited damage”, etc. The model should also output confidence scores describing the likelihood of the input truly belonging to each class, which can be used as the output probabilities.

For this project, this method will be used for the baseline model. It is assumed that a simple multilayer perceptron is not complex enough to learn the rules of the dataset. To ensure it is indeed handicapped, it will only employ the best track data as input. The model is thus not expected to effectively identify the trends and rules associating best track data with strikes at Hong Kong.

The exact details as to how to implement one such contraption will be devised later. The experience gained working on this model and its data will be transferrable to the development of the two ensemble members.

4.3.2 Statistical-Dynamical Decision Tree

Decision trees sort data points into categories using explicit rules. The tree is a binary tree and at each stage (non-leaf node), the data is separated into two groups using one clear-cut rule which tests an attribute of the data. The output variables will be found at the leaf nodes [65].

Decision trees have been cited as useful for TC forecasting and rule discoveries according to [15] and [37] gives a working example of such methodologies. While decision trees are also prone to overfitting by nature, improvements can be made to alleviate overfitting. For example, the branches of the tree could be cut by pruning, so the tree does not grow to unnecessary complexities [66].

For this project, a decision tree may be built as one of the two ensemble members. It will use both best track and dynamical data as input, employ the said techniques against overfitting, and if possible, have its decision rules retrieved for analysis. The classification rules might directly reflect the actual meteorological patterns and this phenomenon can be used to check the effectiveness of the tree. Finally, provided that decision trees prove to be applicable, an ensemble of randomized decision trees, a “random forest”, can be built to fully employ the possibilities decision trees offer while also enjoying the benefit of overfitting avoidance [66]. Extra work will have to be done to make the trees produce a confidence score instead of a fixed classification.

4.3.3 Monte Carlo Method

The Monte Carlo method, or Monte Carlo simulations, is a probabilistic method to forecast outcomes based on randomization and an estimated probability distribution. It is similar to repeatedly doing the same thing to evaluate the probabilities behind, like throwing an unfair die with unknown probabilities numerous times. Given a probability distribution, the Monte Carlo method simulates its outcomes over and over with randomized initial states, then the results can be aggregated as a forecast [67].

As shown by [8] and [12], Monte Carlo methods can be used for TC forecasting. The paper by DeMaria et al. [8] is in particular interesting because it also outlines a method to forecast TC wind field radii, which measure the extent of the TC’s strong winds, gales, and hurricanes. With that, the probability of strong winds or gales blowing in Hong Kong (corresponding to predictands 2 and 3) can be computed. However, because wind speeds do not translate to the first and last predictands, this method cannot be directly used.

Nonetheless, provided a means to model the probability distribution of the input predictors in the first place, this is a very viable choice for an ensemble member. The output of each simulation is different, thus the density of the numerous simulation outputs can be directly used as the probability. The modelling of the probabilities is expected to be the main difficulty in implementing a Monte Carlo method for this project.

4.3.4 CLIPER-like Method

A third candidate for ensemble member model choice is a more comprehensive CLIPER. A combination of an appropriate dataset with the CLIPER methodology can yield decent models, as illustrated by [21], [22], [24] and [26]. While CLIPER is no longer operational [20] and extra care must be taken when handling it, CLIPER can at least provide a way to identify possible predictors, as in [8], [37] and [49].

The primary challenge of implementing CLIPER lies in the difficulty of computing the mathematics behind the climatology part of the model. In [19], the eight input variables (e.g. TC movement velocity, intensity, etc.) were arranged in more than a hundred combinations, under the assumption that these variables control the actual movements of the TC via a polynomial with degree 3. Each combination was then examined until a few candidates were identified as the most important predictors. After that, linear regression was carried out to match the output (TC future movement) to the predictors. Today, techniques such as principal component analysis can be used to overcome the hard work of evaluating the many potential predictors, but the selection and permutation of the input variables are possibly still manual tasks.

For this project, the predictors that constitute the climatology can be found with some patience, but the choice of a persistence measure will be problematic. The persistence (ongoing trend of the TC) and the climatology components are considered together to produce the output [19] but choosing the right vector of variables to serve as the persistence part will likely be difficult. For the case of [19], just the TC's velocity would suffice, because the predictand is the TC's future movements; but for this project, the definitions of the predictand implies there is likely no combination of input variables whose trend will assume important roles.

4.4 Anticipated Problems

It is expected that technical difficulties will recurrently arise throughout the project. To name a few, there will be:

Difficulty in processing NWP data. The gridded data are provided in GRIB2 format and has a large volume. The right computing resource and software to open the dataset will be necessary. The extraction of only the needed variables and the further compressing of the data values will also be challenging. At the moment, it is expected that further exploration will help resolve the issues.

Difficulty in implementing evaluation metrics, such as the logarithmic score. While it is possible to just follow the definitions, the intricacies of the data (e.g. the data types and formats) will most likely hinder rapid progress. Alternative metrics may have to be sought or the F1-score will remain the sole metric used, which makes the outcomes slightly less convincing.

Difficulty in implementing the operational forecast, i.e. the one described in section 4.3.2 that does not rely on datasets. The acquisition of NWP data can be bothersome because it depends on the NWP models' forecast hours and the data format they provide. This, however, can be programmatically handled.

Chapter 5: Project Schedule and Milestones

This chapter deliberates the management of the project. Timetables, risk, and other non-technical issues are conferred below.

5.1 Project Timetable and Artefacts

The following table summarizes the project schedule and milestones.

Date	FYP formalities	Tasks to complete	Artefact
3 Oct 2021	FYP Project Plan		
	FYP Webpage		
October 2021		High-level design	
		Dataset acquisition	
		Data engineering	
November 2021		Building baseline model	Baseline model
		Evaluation of model	
December 2021		Further data engineering	

		Building ensemble member 1	
January 2022	Interim report	Building ensemble member 1	Ensemble-member 1
		Project plan review	
		Evaluation of model	
February 2022		Further data engineering	
		Building ensemble member 2	
March 2022		Building ensemble member 2	Ensemble-member 2
		Evaluation of model	
		Building final ensemble forecast	
April 2022	Final report	Building final ensemble forecast	Final ensemble
		Evaluation of performance	
	Final presentation	Final review	Documentation
		Documentation	

Table 5.1: Project schedule and artefacts table

According to the table, the completion of each of the three models and the ensemble will serve as milestones to track progress. To complete each milestone, the model will be programmed and evaluated, after any necessary data engineering (e.g. to label ground truths, perform needed transformations, etc.) and preparation tasks are completed. The FYP reports and presentation will intersperse the development process. It is expected that the development of the first two models will be more time-consuming than the last one because the process to familiarize oneself with the data and the tool will take considerable time and effort.

5.2 Risk Management

The first risk factor is that the author is expected to be encumbered by coursework of other classes in particular in the first semester, which is crucial as this will be when the groundwork of the project is laid down. The tentative contingency plan is to make up for the lost time during December, but there is still doubt whether this will be sufficient.

The second risk factor is that unforeseen technical difficulties will further hinder progress as they inevitably arise, but it is impossible to judge the extent to which this will happen. If the challenge is so great that it will not be surmountable within a reasonable time, i.e. it will delay the delivery of an artefact significantly, then a change in methodology for that particular part of the project will be necessary.

A third risk factor is the sudden advent of a forecasting system whose scope fully includes that of this project. In that case, the *raison d'être* of this project may be nullified, if the new system proves to be superior. However, it is surmised that the approach of this project is sufficiently unique, and the likelihood of such cataclysmic events is negligibly low.

5.3 Resource Requirements

The storage and processing of the datasets will likely require a compute resource at the Department of Computer Science. This can be satisfied either by applying for a virtual machine dedicated for FYP purposes, or by employing the department's GPU farm. The latter is in particular useful for machine learning models that may be used.

The datasets themselves can be acquired from third parties on the Internet. Authoritative datasets for best track data are readily available at various authorities, while chapter 2 describes three sources for atmospheric data.

Apart from these, there are no other special resource requirements identified at the moment.

Chapter 6: Conclusion

This is an exploratory project that aims to build a statistical-dynamical ensemble forecast for TC strike probabilities and explore the methods meanwhile. Using best track data, NWP analyses archives, and HKO historical TC warning signal issuance records, two ensemble members will be developed and evaluated against a dummy statistical model, before their results are combined to produce final output as a weighted average.

It is envisaged that the resulting forecasts will be useful as an intuitive and reasonably accurate tool for the public to assess the risks of TCs affecting Hong Kong, while also being verifiable such that future research and developments thereupon is plausible.

Appendices

Appendix 1: Full Literature Review Tables

- General literature on TC forecasting

Citation of the Source	Summary	Comment
[5] P. C. Chin, “The Royal Observatory tropical cyclone prediction programme, 1977-1978,” Royal Observatory, Hong Kong, Occasional Paper No. 37, Jul. 1977.	This paper summarizes the TC forecast techniques used at the Hong Kong Observatory (HKO) in 1977. These include climatology and persistence (CLIPER) with 2.5-degree grids, which was considered less accurate if TCs recurve; the Veigas-Miller statistical method with 5-degree grids and sea-level air pressure data, which was decent but error-sensitive and oftentimes ineffective against recurving TCs; the manual subjective Tse’s Method which sensitively relied on upper-air observation data; the control point method which also needed upper-air data; and finally the experimental regression	The methods of the 1970s were unable to consider atmospheric conditions and statistics as comprehensively as nowadays when compared to the multitude of options available today. However, the statistical and experiential techniques are certainly impressive achievements of their time.

	method which employed data between 1884 and 1970. As of 1977, the HKO did not issue forecasts beyond 24 hours.	
[6] R. L. Elsberry, "Advances in research and forecasting of tropical cyclones from 1963–2013," *Journal of the Korean Meteorological Society (한국기상학회지)*, vol. 50, no. 1, pp. 3-16, 2014. DOI: 10.1007/s13143-014-0001-1.	This article summarizes the progress in TC forecasts over the past 50 years. The key advance was in the field of numerical models and some progress was made to forecast TC impacts, but there was little improvement in TC intensity forecasting.	This article points out that NWP is the best TC forecasting means to date.

Table 7.1.1: General literature on prior TC forecasting techniques and progress

- Statistical method: CLIPER

Citation of the Source	Summary	Comment
[19] C. J. Neumann, "An Alternative to the HURRAN (Hurricane Analog) Tropical Cyclone Forecast System," National Hurricane Center, Miami, FL, USA, NOAA Tech. Memo. NWS SR-62, Jan. 1972.	This technical memorandum outlines a statistical method to forecast TC tracks, in light of the deficiencies of earlier systems like HURRAN, using data from 1931-1970. The regressional analyses found two equations describing TC future displacements based on a least-squares fit to 164 predictors. The climatology found, i.e. historical patterns, was combined with	CLIPER models are fully statistical and are now no longer operational [20, pp.17-18], having been superseded by dynamic NWP methods. However, it is still used as a baseline of comparison against other techniques [20, pp.18] and as a basis to develop new statistical methods, as many models bear the name "CLIPER", e.g. in aerosol forecasts [21] and rainfall forecasts [22]. The old CLIPER model

	the TC’s immediate velocity, which served as the persistence component, to form a forecasting model. This new method was called CLIPER (Climatology and Persistence).	and similar methods can be used for this project, as long as the probability of a TC affecting Hong Kong can be calculated from the TC’s forecasted track, intensity, and other attributes.
[23] J. A. Knaff, M. DeMaria, C. R. Sampson and J. M. Gross, “Statistical, 5-Day Tropical Cyclone Intensity Forecasts Derived from Climatology and Persistence,” *Weather and Forecasting*, vol. 18, no. 1, pp. 80-92, Feb. 1, 2003. DOI: 10.1175/1520-0434(2003)018<0080:SDTCIF>2.0.CO;2.	This paper builds upon earlier statistical methods and CLIPER models to construct a new CLIPER model for five-day forecasts. For the western North Pacific, the new model which employed Joint Typhoon Warning Center (JTWC) best track data showed major improvement.	The model introduced in this paper is a successor to that in [19]. It is shown that considering CLIPER methods for Hong Kong’s case and forecasting up to five days ahead in case NWP is not available is possible.
[24] J. A. Knaff, C. R. Sampson, M. DeMaria, T. P. Marchok, J. M. Gross and C. J. McAdie, “Statistical Tropical Cyclone Wind Radii Prediction Using Climatology and Persistence,” *Weather and Forecasting*, vol. 22, no. 4, pp. 781-791, Aug. 1, 2007. DOI: 10.1175/WAF1026.1.	This study used CLIPER methods with best track data and wind distribution radii values mentioned in TC advisories to fit a Rankine vortex, which mathematically models wind as a vortex in some viscous fluid. Then the observed wind radii were included as the persistence component.	This paper shows that with a proper dataset, it is possible to also forecast wind distributions using CLIPER-like statistical means. As wind speeds are a factor in assessing a TC’s impact on Hong Kong, the method described in the paper is potentially viable. However, this approach has rather significant error margins of up to 37% in 72h.
[25] J. A. Knaff, C. R. Sampson and K. D. Musgrave, “Statistical Tropical Cyclone Wind	This paper is a follow up to [24], to address the limitations of the model in the western	The importance of choosing the right mathematical model is illustrated in the paper,

Radii Prediction Using Climatology and Persistence: Updates for the Western North Pacific," *Weather and Forecasting*, vol. 33, no. 4, pp. 1093-1098, Aug. 1, 2018. DOI: 10.1175/WAF-D-18-0027.1.	North Pacific. The vortex model previously used made the radii found too small and symmetric, such that a new vortex climatology needed to be devised.	even though the exact changes made were subtle. See also [24].
[26] J. A. Knaff and C. R. Sampson, "Southern hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning Center, Part I: control forecasts based on climatology and persistence," *Australian Meteorological and Oceanographic Journal*, vol. 58, no.1, pp. 1-7, Mar. 2009. DOI: 10.22499/2.5801.001.	This paper documents a CLIPER method to forecast TC intensities in the Southern hemisphere. The method presented employed data between 1980 and 2002 in its linear regression model.	This paper, once again, states that CLIPER is primarily a benchmark to verify other forecasts, despite being able to provide value. It nonetheless demonstrates how CLIPER is used to forecast intensities.

Table 7.1.2: Review of CLIPER-like TC forecasts

- Dynamical method: NWP

Citation of the Source	Summary	Comment
[10] H. A. Titley, R. L. Bowyer and H. L. Cloke, "A global evaluation of multi-model ensemble tropical cyclone track probability forecasts," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 726, pp. 531-545, Jan. 2020. DOI: 10.1002/qj.3712.	This study evaluated the performance of three dynamic ensemble forecast systems in different ocean basins. The TC track probability forecasts by each system were evaluated and compared,	This paper shows that using ensemble methods provides the best performance in TC forecasting, even if the performance of each ensemble member can vary from

	and it was concluded that different basins had different best-performing forecasts, but a combined ensemble of all three performed even better.	case to case. Also, the paper pointed out that probabilistic forecasts are valuable.
[28] "TC Show Guide." ECMWF. http://www.ecmwf.int/sites/default/files/TC_ShowGuide.pdf (retrieved Sep. 26, 2021).	The ECMWF regularly publishes tropical cyclone strike probability products [27]. By considering the probability at some particular places, the likelihood of a TC affecting a location may be evaluated. This document is a technical summary of the strike probability product, which combines the steering winds and the TC's prior trajectory in a weighted average to produce output.	ECMWF is one of the institutions that produce state-of-the-art weather forecast products. The documentation suggests several indicators for NWP, such as vorticity, warm core, and the aforementioned steering winds.
[29] S. J. Majumdar and P. M. Finocchio, "On the Ability of Global Ensemble Prediction Systems to Predict Tropical Cyclone Track Probabilities," *Weather and Forecasting*, vol. 25, no. 2, pp. 659-680, Apr. 1, 2010. DOI: 10.1175/2009WAF2222327.1.	This article compares the performance of two NWP ensembles in predicting TC track probabilities and found that both had close performances. It also concludes that ensembles enhance probabilistic forecasts.	This article evaluates the contemporary NWP systems' performance and suggests that combined methods and ensembles will likely give superior results.
[30] L. Qi, H. Yu and P. Chen, "Selective ensemble-mean technique for tropical cyclone track forecast by using	This paper attempts to build a TC track forecaster by finding a consensus track	This paper describes one way to establish and enhance ensemble

ensemble prediction systems," *Quarterly Journal for the Royal Meteorological Society*, vol. 140, no. 680, pp. 805-813, Apr. 2014. DOI: 10.1002/qj.2196.	among different NWP models, but the ensemble members' votes were selectively weighted. The outcomes showed that this approach can outperform simple deterministic forecasts.	forecasts based on NWP models.
[31] X. Zhang and H. Yu, "A Probabilistic Tropical Cyclone Track Forecast Scheme Based on the Selective Consensus of Ensemble Prediction Systems," *Weather and Forecasting*, vol. 32, no. 6, pp. 2143-2157, Dec. 2017. DOI: 10.1175/WAF-D-17-0071.1.	This study constructed an ensemble out of two ensemble prediction systems based on NWP and then generated a track probability forecast by selectively favouring individual ensemble members with potential. The resulting model was effective as it showed certain improvement over just the grand ensemble and an individual ensemble prediction system.	Similar to [30], this paper also uses a selective consensus approach. This one establishes a probabilistic method instead. However, its advantage over control methods diminished when forecast hours grew beyond 48h, therefore its improvement is likely limited.
[32] *ECMWF Reanalysis v5 (ERA5)*, ECMWF, n.d. [Online] Available: https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5.	This dataset consists of atmospheric reanalysis data from 1979 to the present, at the resolution of hourly 30-km grids. It also provides a preliminary extension that covers 1950-1979.	This dataset contains high-resolution data, but the data access procedures are more convoluted.
[33] *NCEP GFS 0.25 Degree Global Forecast Grids Historical*	This dataset is an archive of NWP analysis	This dataset is comprehensive and

Archive, National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2015, DOI: 10.5065/D65D8PWK.	data used by the Global Forecasting System (GFS) of the NOAA. Various meteorological data from 2015 to 2020 are archived.	comes from a reputable source, however, it only contains six years' worth of data.
[34] *NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999*, National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2000, DOI: 10.5065/D6M043C6.	This dataset is similar to [33], but the NWP model associated is different. The NCEP FNL model is based on GFS and collects more observation data than GFS before the model is prepared for forecasts. This dataset consists of archived data from 1999 to 2021 and is updated daily.	This dataset covers a substantially longer period than [33], but the resolution of the data (1-degree by 1-degree grids) is not as high as [33] (0.25-degree grids).
[50] J. T. Heming, "Tropical cyclone tracking and verification techniques for Met Office numerical weather prediction models," *Meteorological Applications*, vol. 24, no. 1, pp. 1-8, Jan. 2017. DOI: 10.1002/met.1599.	This article summarizes the history and development of the usage of NWP models at UK authorities, the usable metrics of error, and the current methods to produce TC forecasts.	Useful measurements to evaluate models are presented in this article.

Table 7.1.3: Review of NWP-based TC forecasts and open datasets

- Statistical method: Machine learning

Citation of the Source	Summary	Comment
[15] R. Chen, W. Zhang and X. Wang, "Machine	This article is a review of the recent	This article points to other research on using

Learning in Tropical Cyclone Forecast Modeling: A Review," *Atmosphere*, vol. 11, no. 7, pp. 676, Jun. 27, 2020. DOI: 10.3390/atmos11070676.	advancements in employing machine learning in TC forecasts. In particular, dynamic-statistical models that feed dynamic atmospheric factors into statistical models, statistically assist the correction of NWP models or simple joint models are deemed useful. These techniques are currently used for path forecasts, predictor searching, intensity forecasts and impact forecasts.	machine learning to forecast, simulate and estimate TC wind fields, future intensities, and track.
[35] W. Zhang, Y. Leung and J. C. L. Chan, "The Analysis of Tropical Cyclone Tracks in the Western North Pacific through Data Mining. Part I: Tropical Cyclone Recurvature," *Journal of Applied Meteorology and Climatology*, vol. 52, no. 6, pp. 1394-1416, Jun. 1, 2013. DOI: 10.1175/JAMC-D-12-045.1.	This paper is the prequel to [36]. In this paper, the conditions leading to a TC recurving (making a sharp turn around the subtropical high) are studied, using the same dataset as [36]. It is only in this paper the data sources are better explained.	According to this article, the meteorological data can be obtained from numerical weather prediction (NWP) institutions such as the GFS. While only the data from 1999 onwards are available, it can be a very useful resource. Please see the section about NWP for more information.
[36] W. Zhang, Y. Leung and J. C. L. Chan, "The Analysis of Tropical Cyclone Tracks in the Western North Pacific through Data Mining. Part II: Tropical Cyclone Landfall," *Journal of Applied Meteorology and Climatology*, vol. 52, no. 6, pp. 1417-1432, Jun. 1, 2013. DOI:	This paper uses data mining techniques to study the meteorological data (such as 400hPa subtropical high-intensity index) and TC best track data of 2000-2009 to discover the rules deciding whether a TC makes landfall in China or not, via a decision tree.	This paper is rather limited in that it only employs a decade's data, so the results may suffer from overfitting. However, its satisfying result (83% accuracy) illustrates that it is possible to use data mining and machine learning methods to tackle the problem of

10.1175/JAMC-D-12-046.1.		predicting a TC's impact on Hong Kong, provided that meteorological data is involved.
[37] J. Tan, S. Chen and J. Wang, "Western North Pacific tropical cyclone track forecasts by a machine learning model," *Stochastic Environmental Research and Risk Assessment*, vol. 35, pp. 1113-1126, Nov. 13, 2020. DOI: 10.1007/s00477-020-01930-w.	This paper describes an ensemble forecast for TC tracks using decision trees with 30 years' worth of data. The predictors were extracted from the preexisting Climatology and Persistence (CLIPER) techniques, against which the forecast's performance is compared.	This paper is similar to [36] in that they both used machine learning methods with meteorological data to build decent models. This further justifies the plausibility of machine learning being a potential approach.
[38] T. Loridan, R. P. Crompton and E. Dubossarsky, "A Machine Learning Approach to Modeling Tropical Cyclone Wind Field Uncertainty," *Monthly Weather Review*, vol. 145, no. 8, pp. 3203-3221, Aug. 1, 2017. DOI: 10.1175/MWR-D-16-0429.1.	This paper explores potential machine learning algorithms to represent TC wind distributions. It analyzed data generated using a numerical model (WRF), which was in turn initialized using historical data of 30 TCs, and a quantile regression forest model was created to model the key predictors. Then the probabilities of getting certain wind speeds were computed.	This study shows that it is feasible to combine machine learning with meteorological data prepared by NWP. However, it is believed that a small sampling of 30 TCs is insufficient to model all TCs.
[39] J. Devaraj, S. Ganesan, R. M. Elavarasan and U. Subramaniam, "A Novel Deep Learning-Based Model for Tropical Intensity Estimation and Post-Disaster Management of	This paper discusses using deep learning models to estimate TC intensities based on infrared satellite imagery and wind speed data and to predict the extent of the damage.	In this paper, the damage of the TC was represented by marking satellite images of Houston as "damaged" or "not damaged", which may not apply to densely developed

Hurricanes," *Applied Sciences*, vol. 11, no. 9, pp. 4129, Jan. 1, 2021. DOI: 10.3390/app11094129.		cities like Hong Kong, where TC damage cannot be visually identified from a satellite image.

Table 7.1.4: Review of machine learning applications for TC forecasting

- Literature on strike probabilities or damage estimates

Citation of the Source	Summary	Comments
[11] C. Wang and H. Zhang, "Probability-based estimate of tropical cyclone damage: An explicit approach and application to Hong Kong, China," *Engineering Structures*, vol. 167, pp. 471–480, Jul. 15, 2018. DOI: 10.1016/j.engstruct.2018.04.064.	This paper introduces probabilistic models to estimate TC damage and TC occurrences using historical data. The data employed in the study is available at [40].	The study used economic loss in HKD as a metric for TC damage. The paper focuses only on the amount of damage brought by a TC, instead of the likelihood of a TC being able to cause damage. The severity of TCs was quantified primarily by maximum wind speed, and the shortcomings of only relying on wind speeds have been acknowledged.
[12] X. Xie, B. Xie, J. Cheng, Q. Chu and T. Dooling, "A simple Monte Carlo method for estimating the chance of a cyclone impact," *Natural Hazards*, vol. 107, pp. 2573-2582, Jan. 13, 2021. DOI: 10.1007/s11069-021-04505-2.	This paper uses Monte Carlo techniques with historical data to evaluate the probability of a TC hitting a specific region in a year.	This study did not evaluate the probabilities of a specific TC hitting a specific location. However, the methodology used is novel and could open possibilities for data

		augmentation.
[13] S. S. Chand, K. J. E. Walsh and J. C. L. Chan, "A Bayesian Regression Approach to Seasonal Prediction of Tropical Cyclones Affecting the Fiji Region," *Journal of Climate*, vol. 23, no. 13, pp. 3425-3445, Jul. 1, 2010. DOI: 10.1175/2010JCL3521.1.	This paper uses Bayesian regression models to predict the annual TC activity around Fiji, Samoa, and Tonga, taking large-scale environmental factors such as the El Niño-Southern Oscillation (ENSO) index into consideration.	This study focused on counting the number of TCs that may affect the Fiji region, instead of the likelihood of individual TCs to affect Fiji after their geneses. This study also highlights how ENSO reduces forecast accuracies.
[16] P. C. Chin, "Tropical Cyclone Strike Probability Values for Ten Target Locations in Southeast Asia," Royal Observatory, Hong Kong, Occasional Paper No. 39, Nov. 1977.	This paper by the Hong Kong Observatory uses best track statistics to estimate the probability of a TC being detected inside each 5-degree by 5-degree squares. For the regions around the concerned location, probabilities of TCs detected inside 1-, 2-, and 3-degrees radii were computed. The seasonal factor (i.e. date) was quantified by dividing one year into 12 periods, each having the same number of historical TC occurrences.	This is one of the earliest statistical and probabilistic studies on TC strike probability available and despite the relatively crude methodology, this study is a laudable effort in this field. It is also worth noting that this paper does not include any evaluation of the computed probabilities.
[17] "Tropical Cyclone Strike Probability Help Document." Hong	Based on research by T. R. Metcalf,	The original paper by Metcalf is

Kong Weather Watch. http://www.hkww.org/weather/strikeprob/strikeprobhelp.html (retrieved Sep. 26, 2021).	an algorithm was implemented to calculate the probability of a given TC entering a certain radius from Hong Kong. The forecasts are made up to 96 hours in advance and expected positions and intensity values of the TC are also calculated.	missing and is hence unusable for this project, as there is no way to implement, evaluate or modify this approach. It should also be noted that this software considers position and intensity only.
[41] S. S. Chand and K. J. E. Walsh, "Modeling Seasonal Tropical Cyclone Activity in the Fiji Region as a Binary Classification Problem," *Journal of Climate*, vol. 25, no. 14, pp. 5057-5071, Jul. 15, 2012. DOI: 10.1175/JCLI-D-11-00507.1.	This paper models TC activity around Fiji, Samoa, and Tonga as a binary classification problem (high vs low TC activity), which was solved in the study using a probit regression model.	Like [3], this study was intended for season-wide forecasts of TC activity. Nonetheless, several useful TC activity indicators are revealed in the study, namely accumulated cyclone energy, low-level relative vorticity, upper-level divergence, and midtropspheric relative humidity.

Table 7.1.5: Various TC strike probability and activity level forecasts

- Miscellaneous methods and techniques

Citation of the Source	Summary	Comments
[8] M. DeMaria, J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson and R. T. DeMaria, "A New Method for Estimating Tropical Cyclone Wind Speed Probabilities,"	In this study by the NOAA, a Monte Carlo method-based model was built to forecast the probabilities of different levels of winds	This paper provides a new approach to the problem, namely forecasts may build upon each other, and older

Weather and Forecasting, vol. 24, no. 6, pp. 1573-1591, Dec. 1, 2009. DOI: 10.1775/2009WAF2222286.1.	distributed around TCs. Plausible tuples of future tracks and intensities were first generated, then CLIPER methods (see [24]) were used to find the wind structures for each of them before probabilities were computed. The resultant model went into operation in 2006.	techniques can be combined into the analysis. However, it is tricky to generate sufficiently accurate track and intensity forecasts in the first place.
[42] B. Y. Lee, "Kidney and Beach Ball (Wind Speed and Direction Forecast)." Hong Kong Observatory. https://www.weather.gov.hk/en/education/tropical-cyclone/forecasting/00164-kidney-and-beach-ball-wind-speed-and-direction-forecast.html (retrieved Sep. 26, 2021).	This article outlines a manual method based on experience to estimate the possibility of strong winds appearing in Hong Kong in the event of a TC passing through certain sea areas.	The manual method is certainly unusable for this project and the article does not mention how to scale the method up to estimate other wind strengths. Nonetheless, it illustrates that historical data may be used to derive a probabilistic means to predict the chances of Hong Kong being affected by TCs.
[43] M. E. Splitt, J. A. Shafer, S. M. Lazarus and W. P. Roeder, "Evaluation of the National Hurricane Center's Tropical Cyclone Wind Speed Probability Forecast Product," *Weather and Forecasting*, vol. 25, no. 2, pp. 511-525, Apr. 1, 2010. DOI: 10.1175/2009WAF2222279.1.	This paper evaluates the method described in [8] and the researchers discovered that the technique tended to overforecast probability values in general, except the 50-knots category where long-term forecasts more often underforecasted.	This paper reveals the limitations of the method in [8]. See also [8], [24] and [44].

[44] M. DeMaria, J. A. Knaff, M. J. Brennan, D. Brown, R. D. Knabb, R. T. DeMaria, A. Schumacher, C. A. Lauer, D. P. Roberts, C. R. Sampson, P. Santos, D. Sharp and K. A. Winters, "Improvements to the Operational Tropical Cyclone Wind Speed Probability Model," *Weather and Forecasting*, vol. 28, no. 3, pp. 586-602, Jun. 1, 2013. DOI: 10.1175/WAF-D-12-00116.1.	This article is a follow-up to [8] and [43] and documents the verification of and improvements made to the Monte Carlo model introduced in [8]. The model was updated to also consider errors, to forecast more frequently and to correct overestimates.	This paper details the limitations of the model described in [8]; also see [8], [24] and [43].
[45] M. E. Splitt, S. M. Lazarus, S. Collins, D. N. Botambekov and W. P. Roeder, "Probability Distributions and Threshold Selection for Monte Carlo–Type Tropical Cyclone Wind Speed Forecasts," *Weather and Forecasting*, vol. 29, no. 5, pp. 1155-1168, Oct. 1, 2014. DOI: 10.1175/WAF-D-13-00100.1.	This paper is a follow-up to [43] and [44] Here, an alternative mathematical method to evaluate the Monte-Carlo TC wind speed forecast product is proposed.	This paper primarily describes evaluation and interpretation methods for models like [8] and [44], but the techniques are complicated to implement.
[47] Q. Li, P. Xu, X. Wang, H. Lan, C. Cao, G. Li, L. Zhang and L. Sun, "An Operational Statistical Scheme for Tropical Cyclone Induced Wind Gust Forecasts," *Weather and Forecasting*, vol. 31, no. 6, pp. 1817-1832, Dec. 1, 2016. DOI: 10.1175/WAF-D-16-0015-1.	This paper discusses using statistical methods to forecast the probability of gust gale occurrence at certain weather stations in Hong Kong and Shenzhen. The probabilities took into consideration the distance and azimuth of the inbound TCs from the stations and the model primarily used observations between 1968 and 2014.	The paper provides an appealing angle to look at TCs and their impact (in terms of gusts) by plotting the TCs on polar coordinate planes. However, in practice, the winds observed at stations inside urban centres decreased over the years, and wind speeds at merely 3 stations are questionable indicators of a TC's

		impact.
[46] N. Bloemendaal, H. de Moel, S. Muis, I. D. Haigh and J. C. J. H. Aerts, "Estimation of global tropical cyclone wind speed probabilities using the STORM dataset," *Scientific Data*, vol. 7, no.1, Nov. 10, 2020. DOI: 10.1038/x41597-020-00720-x.	This paper uses a synthetic dataset equivalent to 10000 years' worth of TC data to estimate the frequencies of TC occurrences, specific wind speed occurrences and TC strike events.	While the dataset of choice is intriguing in its synthetic nature, the paper does not estimate strike probabilities for individual TCs.
[48] H. Kim, C. Ho, J. Kim and P. Chu, "Track-Pattern-Based Model for Seasonal Prediction of Tropical Cyclone Activity in the Western North Pacific," *Journal of Climate*, vol. 25, no. 13, pp. 4660-4678, Jul. 1, 2012. DOI: 10.1175/JCLI-D-11-00236.1.	This paper describes a statistical-dynamical method to cluster historical TCs into seven patterns observable over each season and to plot a map showing the probability density of seasonal TC tracks over the ocean basin.	This paper is another example of combined statistical and dynamic methods being used in modern TC forecasting. However, this paper focuses more on season-wide climatology than forecasting individual TCs.
[49] J. A. Knaff and C. R. Sampson, "Southern hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning Center, Part II: statistical-dynamical forecasts," *Australian Meteorological and Oceanographic Journal*, vol. 58, no.1, pp. 9-18, Mar. 2009. DOI: 10.22499/2.5801.002.	This paper is the sequel to [14]. Here a model that combines CLIPER and actual observations such as vertical wind shear and other intensity forecasts is introduced.	This paper is similar to [8] as various methods and observations are combined and potential predictors are identified, before regressional analyses to fit those predictors to the dataset is done.

Table 7.1.6: Miscellaneous TC forecasting methods and techniques

Appendix 2: List of Abbreviations

Abbreviation	Full form
CLIPER	Climatology and persistence
CMA	China Meteorological Administration
ECMWF	European Centre for Medium-Range Weather Forecasts
FYP	Final Year Project
GFS	Global Forecasting System
HKO	Hong Kong Observatory
JTWC	Joint Typhoon Warning Center
NHC	National Hurricane Center
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical weather prediction
TC	Tropical cyclone

Appendix 3: List of Tables and Figures

Appendix 3.1: List of Tables

Appendix 3.2: List of Figures

Appendix 4: References

[1] "Hurricanes Frequently Asked Questions." NOAA's Atlantic Oceanographic and Meteorological Laboratory. https://www.aoml.noaa.gov/hrd-faq/#what-is-a-hurricane (accessed Sep. 30, 2021).

[2] "Tropical Cyclones in 2020," HKO, Hong Kong, Jul. 2021. Accessed Sep. 30, 2021. [Online]. Available: https://www.hko.gov.hk/en/publica/tc/files/TC2020.pdf.

[3] "Social and Economic Impact of Tropical Cyclones." HKO. https://www.hko.gov.hk/en/informtc/economice.htm (accessed Sep. 30, 2021).

[4] "Tracking and Forecasting Tropical Cyclones." HKO. https://www.hko.gov.hk/en/informtc/tracking.htm (accessed Sep. 30, 2021).

[5] P. C. Chin, "The Royal Observatory tropical cyclone prediction programme, 1977-1978," Royal Observatory, Hong Kong, Occasional Paper No. 37, Jul. 1977.

[6] R. L. Elsberry, "Advances in research and forecasting of tropical cyclones from 1963–2013," *Journal of the Korean Meteorological Society (한국기상학회지)*, vol. 50, no. 1, pp. 3-16, 2014. DOI: 10.1007/s13143-014-0001-1.

[7] "Tropical Cyclone Track Probability Forecast (Beta)." HKO. https://www.hko.gov.hk/en/probfcst/tc_spm.htm (accessed Sep. 30, 2021)

[8] M. DeMaria, J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson and R. T. DeMaria, "A New Method for Estimating Tropical Cyclone Wind Speed Probabilities," *Weather and Forecasting*, vol. 24, no. 6, pp. 1573-1591, Dec. 1, 2009. DOI: 10.1775/2009WAF2222286.1.

[9] "NHC Tropical Cyclone Graphical Product Descriptions." NHC and Central Pacific Hurricane Center. https://www.nhc.noaa.gov/aboutnhcgraphics.shtml#WINDPROB (accessed Sep. 30, 2021).

[10] H. A. Titley, R. L. Bowyer and H. L. Cloke, "A global evaluation of multi-model ensemble tropical cyclone track probability forecasts," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 726, pp. 531-545, Jan. 2020. DOI: 10.1002/qj.3712.

[11] C. Wang and H. Zhang, "Probability-based estimate of tropical cyclone damage: An explicit approach and application to Hong Kong, China," *Engineering Structures*, vol. 167, pp. 471–480, Jul. 15, 2018. DOI: 10.1016/j.engstruct.2018.04.064.

[12] X. Xie, B. Xie, J. Cheng, Q. Chu and T. Dooling, "A simple Monte Carlo method for estimating the chance of a cyclone impact," *Natural Hazards*, vol. 107, pp. 2573-2582, Jan. 13, 2021. DOI: 10.1007/s11069-021-04505-2.

[13] S. S. Chand, K. J. E. Walsh and J. C. L. Chan, "A Bayesian Regression Approach to Seasonal Prediction of Tropical Cyclones Affecting the Fiji Region," *Journal of*

Climate, vol. 23, no. 13, pp. 3425-3445, Jul. 1, 2010. DOI: 10.1175/2010JCL3521.1.

[14] J. A. Knaff and C. R. Sampson, “Southern hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning Center, Part II: statistical-dynamical forecasts,” *Australian Meteorological and Oceanographic Journal*, vol. 58, no.1, pp. 9-18, Mar. 2009. DOI: 10.22499/2.5801.002.

[15] R. Chen, W. Zhang and X. Wang, “Machine Learning in Tropical Cyclone Forecast Modeling: A Review,” *Atmosphere*, vol. 11, no. 7, pp. 676, Jun. 27, 2020. DOI: 10.3390/atmos11070676.

[16] P. C. Chin, “Tropical Cyclone Strike Probability Values for Ten Target Locations in Southeast Asia,” Royal Observatory, Hong Kong, Occasional Paper No. 39, Nov. 1977.

[17] “Tropical Cyclone Strike Probability Help Document.” Hong Kong Weather Watch. http://www.hkww.org/weather/strikeprob/strikeprobhelp.html (retrieved Sep. 26, 2021).

[18] J. Jarrell and S. Brand, “Tropical Cyclone Strike and Wind Probability Applications,” *Bulletin of the American Meteorological Society*, vol. 64, no. 9, pp. 1050-1056, Sep. 1983. [Online] Available: https://www.jstor.org/stable/26223426.

[19] C. J. Neumann, “An Alternative to the HURRAN (Hurricane Analog) Tropical Cyclone Forecast System,” National Hurricane Center, Miami, FL, USA, NOAA Tech. Memo. NWS SR-62, Jan. 1972. [Online] Available: https://repository.library.noaa.gov/view/noaa/3605.

[20] D. A. Zelinsky and R. J. Pasch, “Tropical Cyclone Track Prediction.” Severe Weather. http://severeweather.wmo.int/TCFW/RAIV_Workshop2021/20_TC-TrackForecasting_DaveZelinsky_Richardpasch.pdf (retrieved Sep. 27, 2021).

[21] J. Reid, J. Zhang, B. Sampson, J. Hansen and W. Sessions. “Climatology-Persistence Models (CLIPERs).” ICAP-Ensemble Meeting. http://icap.atmos.und.edu/EnsembleForecastsDataAssimilation/MeetingPDFs/May12/06_Reid-Cliper.pdf (retrieved Sep. 27, 2021).

[22] F. Marks, G. Kappler and M. DeMaria. “Development of a Tropical Cyclone Rainfall Climatology and Persistence (R-CLIPER) Model,” National Oceanic and Atmospheric Administration (NOAA) Hurricane Research Division, Miami, FL, Dec. 31, 2003. Accessed: Sep. 27, 2021. [Online] Available: https://origin.www.nhc.noaa.gov/jht/final_rep/R-CLIPER_final.pdf.

[23] J. A. Knaff, M. DeMaria, C. R. Sampson and J. M. Gross, “Statistical, 5-Day Tropical Cyclone Intensity Forecasts Derived from Climatology and Persistence,” *Weather and Forecasting*, vol. 18, no. 1, pp. 80-92, Feb. 1, 2003. DOI: 10.1175/1520-0434(2003)018<0080:SDTCIF>2.0.CO;2.

[24] J. A. Knaff, C. R. Sampson, M. DeMaria, T. P. Marchok, J. M. Gross and C. J. McAdie,

"Statistical Tropical Cyclone Wind Radii Prediction Using Climatology and Persistence," Weather and Forecasting, vol. 22, no. 4, pp. 781-791, Aug. 1, 2007. DOI: 10.1175/WAF1026.1.

[25] J. A. Knaff, C. R. Sampson and K. D. Musgrave, "Statistical Tropical Cyclone Wind Radii Prediction Using Climatology and Persistence: Updates for the Western North Pacific," *Weather and Forecasting*, vol. 33, no. 4, pp. 1093-1098, Aug. 1, 2018. DOI: 10.1175/WAF-D-18-0027.1.

[26] J. A. Knaff and C. R. Sampson, "Southern hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning Center, Part I: control forecasts based on climatology and persistence," *Australian Meteorological and Oceanographic Journal*, vol. 58, no.1, pp. 1-7, Mar. 2009. DOI: 10.22499/2.5801.001.

[27] "Tropical cyclone strike probability." ECMWF. https://www.ecmwf.int/en/forecasts/charts/tcyclone/tc_strike_probability (retrieved Sep. 26, 2021).

[28] "TC Show Guide." ECMWF. http://www.ecmwf.int/sites/default/files/TC_ShowGuide.pdf (retrieved Sep. 26, 2021).

[29] S. J. Majumdar and P. M. Finocchio, "On the Ability of Global Ensemble Prediction Systems to Predict Tropical Cyclone Track Probabilities," *Weather and Forecasting*, vol. 25, no. 2, pp. 659-680, Apr. 1, 2010. DOI: 10.1175/2009WAF2222327.1.

[30] L. Qi, H. Yu and P. Chen, "Selective ensemble-mean technique for tropical cyclone track forecast by using ensemble prediction systems," *Quarterly Journal for the Royal Meteorological Society*, vol. 140, no. 680, pp. 805-813, Apr. 2014. DOI: 10.1002/qj.2196.

[31] X. Zhang and H. Yu, "A Probabilistic Tropical Cyclone Track Forecast Scheme Based on the Selective Consensus of Ensemble Prediction Systems," *Weather and Forecasting*, vol. 32, no. 6, pp. 2143-2157, Dec. 2017. DOI: 10.1175/WAF-D-17-0071.1.

[32] *ECMWF Reanalysis v5 (ERA5)*, ECMWF, n.d. [Online] Available: https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5.

[33] *NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive*, National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2015, DOI: 10.5065/D65D8PWK.

[34] *NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999*, National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2000, DOI: 10.5065/D6M043C6.

[35] W. Zhang, Y. Leung and J. C. L. Chan, "The Analysis of Tropical Cyclone Tracks in

the Western North Pacific through Data Mining. Part I: Tropical Cyclone Recurvature," *Journal of Applied Meteorology and Climatology*, vol. 52, no. 6, pp. 1394-1416, Jun. 1, 2013. DOI: 10.1175/JAMC-D-12-045.1.

[36] W. Zhang, Y. Leung and J. C. L. Chan, "The Analysis of Tropical Cyclone Tracks in the Western North Pacific through Data Mining. Part II: Tropical Cyclone Landfall," *Journal of Applied Meteorology and Climatology*, vol. 52, no. 6, pp. 1417-1432, Jun. 1, 2013. DOI: 10.1175/JAMC-D-12-046.1. \

[37] J. Tan, S. Chen and J. Wang, "Western North Pacific tropical cyclone track forecasts by a machine learning model," *Stochastic Environmental Research and Risk Assessment*, vol. 35, pp. 1113-1126, Nov. 13, 2020. DOI: 10.1007/s00477-020-01930-w.

[38] T. Loridan, R. P. Crompton and E. Dubossarsky, "A Machine Learning Approach to Modeling Tropical Cyclone Wind Field Uncertainty," *Monthly Weather Review*, vol. 145, no. 8, pp. 3203-3221, Aug. 1, 2017. DOI: 10.1175/MWR-D-16-0429.1.

[39] J. Devaraj, S. Ganesan, R. M. Elavarasan and U. Subramaniam, "A Novel Deep Learning-Based Model for Tropical Intensity Estimation and Post-Disaster Management of Hurricanes," *Applied Sciences*, vol. 11, no. 9, pp. 4129, Jan. 1, 2021. DOI: 10.3390/app11094129.

[40] "Study Data from University of Sydney Update Knowledge of Climate Change (Probability-based estimate of tropical cyclone damage: An explicit approach and application to Hong Kong, China)," *Global Warming Focus*, pp. 238, Jul. 2, 2018.

[41] S. S. Chand and K. J. E. Walsh, "Modeling Seasonal Tropical Cyclone Activity in the Fiji Region as a Binary Classification Problem," *Journal of Climate*, vol. 25, no. 14, pp. 5057-5071, Jul. 15, 2012. DOI: 10.1175/JCLI-D-11-00507.1.

[42] B. Y. Lee, "Kidney and Beach Ball (Wind Speed and Direction Forecast)." Hong Kong Observatory. https://www.weather.gov.hk/en/education/tropical-cyclone/forecasting/00164-kidney-and-beach-ball-wind-speed-and-direction-forecast.html (retrieved Sep. 26, 2021).

[43] M. E. Splitt, J. A. Shafer, S. M. Lazarus and W. P. Roeder, "Evaluation of the National Hurricane Center's Tropical Cyclone Wind Speed Probability Forecast Product," *Weather and Forecasting*, vol. 25, no. 2, pp. 511-525, Apr. 1, 2010. DOI: 10.1175/2009WAF2222279.1.

[44] M. DeMaria, J. A. Knaff, M. J. Brennan, D. Brown, R. D. Knabb, R. T. DeMaria, A. Schumacher, C. A. Lauer, D. P. Roberts, C. R. Sampson, P. Santos, D. Sharp and K. A. Winters, "Improvements to the Operational Tropical Cyclone Wind Speed Probability Model," *Weather and Forecasting*, vol. 28, no. 3, pp. 586-602, Jun. 1, 2013. DOI: 10.1175/WAF-D-12-00116.1.

[45] M. E. Splitt, S. M. Lazarus, S. Collins, D. N. Botambekov and W. P. Roeder,

"Probability Distributions and Threshold Selection for Monte Carlo–Type Tropical Cyclone Wind Speed Forecasts," *Weather and Forecasting*, vol. 29, no. 5, pp. 1155-1168, Oct. 1, 2014. DOI: 10.1175/WAF-D-13-00100.1.

[46] N. Bloemendaal, H. de Moel, S. Muis, I. D. Haigh and J. C. J. H. Aerts, "Estimation of global tropical cyclone wind speed probabilities using the STORM dataset," *Scientific Data*, vol. 7, no.1, Nov. 10, 2020. DOI: 10.1038/x41597-020-00720-x.

[47] Q. Li, P. Xu, X. Wang, H. Lan, C. Cao, G. Li, L. Zhang and L. Sun, "An Operational Statistical Scheme for Tropical Cyclone Induced Wind Gust Forecasts," *Weather and Forecasting*, vol. 31, no. 6, pp. 1817-1832, Dec. 1, 2016. DOI: 10.1175/WAF-D-16-0015-1.

[48] H. Kim, C. Ho, J. Kim and P. Chu, "Track-Pattern-Based Model for Seasonal Prediction of Tropical Cyclone Activity in the Western North Pacific," *Journal of Climate*, vol. 25, no. 13, pp. 4660-4678, Jul. 1, 2012. DOI: 10.1175/JCLI-D-11-00236.1.

[49] J. A. Knaff and C. R. Sampson, "Southern hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning Center, Part II: statistical-dynamical forecasts," *Australian Meteorological and Oceanographic Journal*, vol. 58, no.1, pp. 9-18, Mar. 2009. DOI: 10.22499/2.5801.002.

[50] J. T. Heming, "Tropical cyclone tracking and verification techniques for Met Office numerical weather prediction models," Meteorological Applications, vol. 24, no. 1, pp. 1-8, Jan. 2017. DOI: 10.1002/met.1599.

[51] *Hong Kong's Tropical Cyclone Warning Signals.* HKO, Hong Kong, n.d. Accessed: Oct. 3, 2021. [Online]. Available: https://www.hko.gov.hk/en/publica/gen_pub/files/tcws.pdf.

[52] L. Peng, Y. Li, F. Peng, M. Wong, P. Chan and T. Lee, (n.d.) Wind loss in Kowloon, Hong Kong from 1964 to 2010. Presented at Urban Meteorology and Climate Conference, Guy Carpenter Asia-Pacific Climate Impact Centre, City University of Hong Kong. [Online]. Available: http://www.cityu.edu.hk/gcacic/umcc/Abstracts/Abstract_Kowloon_PENG_LEI.pdf.

[53] "Tropical Cyclones in 2000," HKO, Hong Kong, Apr. 2001. Accessed Oct. 3, 2021. [Online]. Available: https://www.hko.gov.hk/en/publica/tc/files/tc2000.pdf.

[54] "Tropical Cyclones in 2012," HKO, Hong Kong, Jun. 2014 Accessed Oct. 3, 2021. [Online]. Available: https://www.hko.gov.hk/en/publica/tc/files/TC2012.pdf.

[55] *Western North Pacific Ocean Best Track Data*, JTWC, n.d. [Online]. Available: https://www.metoc.navy.mil/jtwc/jtwc.html?western-pacific.

[56] *International Best Track Archive for Climate Stewardship (IBTrACS)*, World Data Center for Meteorology, Asheville, n.d. doi:10.25921/82ty-9e16.

[57] *CMA Tropical Cyclone Best Track Dataset*, China Meteorological Administration, 2021. [Online]. Available: https://tcdata.typhoon.org.cn/en/zjljsjj_sm.html.

[58] M. Ying, W. Zhang, H. Yu, X. Lu, J. Feng, Y. Fan, Y. Zhu and D. Chen, "An overview of the China Meteorological Administration tropical cyclone database," *J. Atmos. Oceanic Technol.*, vol. 31, pp. 287-301, 2014. DOI: 10.1175/JTECH-D-12-00119.1.

[59] X. Q. Lu, H. Yu, M. Ying, B. K. Zhao, S. Zhang, L. M. Lin, L. N. Bai and R. J. Wan, "Western North Pacific tropical cyclone database created by the China Meteorological Administration," *Adv. Atmos. Sci.*, vol. 38, no. 4, pp. 690–699, 2021. DOI: 10.1007/s00376-020-0211-7.

[60] P. S. Koo, "Accuracy, Precision, Recall or F1?" Towards Data Science. https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9 (retrieved Oct. 3, 2021).

[61] J. Y. Wu, "Why Calibrators? Part 1 of the Series on Probability Calibration." Towards Data Science. https://towardsdatascience.com/why-calibrators-part-1-of-the-series-on-probability-calibration-9110831c6bde (retrieved Oct. 3, 2021).

[62] A. P. Dawid and M. Musio, "Theory and Applications of Proper Scoring Rules," *METRON*, vol. 72, no. 2, pp. 169-183, Jan. 2014. DOI:10.1007/s40300-014-0039-y.

[63] "1.16 Probability calibration." Scikit-learn 1.0 documentation. https://scikit-learn.org/stable/modules/calibration.html (retrieved Oct. 3, 2021).

[64] "Multilayer Perceptron." DeepAI. https://deepai.org/machine-learning-glossary-and-terms/multilayer-perceptron (retrieved Oct. 3, 2021).

[65] "Decision Tree." DeepAI. https://deepai.org/machine-learning-glossary-and-terms/decision-tree (retrieved Oct. 3, 2021).

[66] S. Kumar, "3 Techniques to Avoid Overfitting of Decision Trees." Towards Data Science. https://towardsdatascience.com/3-techniques-to-avoid-overfitting-of-decision-trees-1e7d3d985a09 (retrieved Oct. 3, 2021).

[67] "Monte Carlo Simulation." IBM Cloud Learn Hub. https://www.ibm.com/cloud/learn/monte-carlo-simulation (retrieved Oct. 3, 2021).