

# Track-dependency of tropical cyclone risk in South Korea

Chaehyeon C. Nam<sup>1</sup>, Doo-Sun R. Park<sup>1</sup>, Chang-Hoi Ho<sup>1</sup>, Deliang Chen<sup>2</sup>

<sup>1</sup>School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea

<sup>2</sup>Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden

5 *Correspondence to:* Doo-Sun R. Park (dsrpark@cpl.snu.ac.kr)

**Abstract.** In tropical cyclone (TC) risk assessment, many of previous studies have attempted to construct the statistical relationship of TC damage and three risk elements—exposure, vulnerability, and hazard. For hazard parameters, central minimum pressure, maximum wind speed, and size have been widely utilized while the track was not mainly considered as one of critical hazard parameters. This study shows that, in a decision tree analysis for TCs that made landfall in South Korea during 1979–2010, track is the primary factor to predict damage occurrence. Small track deviations  $\leq 250$  km could distinctively change the damage maps for South Korea. This significant track-dependency of TC risk exists because TC track is responsible for the realization of the other hazards developing from a potential to an active hazard, such as wind gusts or downpours at a given settlement. TC track determines the overall severity and spatial distribution of the active hazards by changing the combination of multiple factors: physical geography experience, duration of influence, and relative position of dangerous semicircle side of the TC. These results indicate that large uncertainty in future track projection may seriously mislead future TC risk modeling because trivial track projection error alone can produce severe errors in damage projection even when TC frequency and intensity value is precise and accurate.

## 1 Introduction

Tropical cyclones (TCs) are among the biggest concerns for disaster management since TC is the most costly natural disaster as a single natural hazard worldwide (<http://emdat.be>). Many researchers have tried to understand and predict TC activity and associated risk. In TC risk studies, the risk triangle concept, which describes risk as comprising three major elements (i.e., hazard, exposure, and vulnerability), is widely adopted (e.g., Mendelsohn et al. 2012; Peduzzi et al. 2012). To estimate the risk quantitatively, actual damage is used as a response variable in an empirical statistical model (Pielke et al. 2008; Park et al. 2015), although damage is more likely a materialization of risk in the strictest sense (Cardona et al. 2012). The three risk elements are often used as explanatory variables. Exposure and vulnerability are usually expressed by the number of residents and regional gross domestic product (GDP) in the area of interest, respectively (Pielke et al. 2008). Hazard is typically represented by a TC-based hazard parameter, such as central pressure, maximum wind speed, and TC size (e.g., Nordhaus 2010; Hsiang and Narita 2012; Czajkowski and Done 2014; Zhai and Jiang 2014).

Using TC-based hazard parameters, however, is insufficient for estimating TC damage for the following reason. TC-based hazard parameters represent the amount of the potential hazards rather than the active hazards because they just indicate only

one representative intensity value for a TC. Several different modes of hazards are recognized. Most hazards are ‘dormant’ or ‘potential,’ simply posing a level of threat to life, property, and/or environment, but once a hazard turns ‘active,’ it becomes an incident/emergency (MaCollum 2011). With this classification, TC-based hazard parameters, which indicate the intensity of a TC regardless of the possibility of actual impact occurrence, are labeled as ‘dormant’ or ‘potential’ mode if the TC is in a position of approaching a settlement. Meanwhile, ‘active’ hazards that are driven by a TC can be listed as rainfall, wind gusts, wind waves, and storm surges, all of which are localized and directly affect residents. Following their definitions, active hazards should be more closely correlated with damage than potential hazards.

In fact, the realization of potential hazards of a TC into active hazards seems to be largely dependent on the TC track. In other words, track has a key role in determining 1) whether the potential hazard of a TC will become ‘activated’ for a given settlement and 2) in how intense the active hazard would be. Record-breaking rainfall in Gangneung city, South Korea was recorded, because the track of Typhoon RUSA (2002) was optimal to strengthen the orographic effect on precipitation over the region (Park and Lee 2007). Also, the deadliest damage by typhoon Haiyan in the Philippines in 2013 was mainly because the TC penetrated Tacloban city, which is located in a low-lying area near the ocean, such that most of the damage arose from storm surge (Ching et al. 2015). In both cases, if the TCs went through a different area, avoiding the mountains and lowland, the result could have been much less devastating.

The present study focuses on the importance of track in the TC risk determination process, and purposes to show the role of track to mediate the TC hazards by systematically and directly comparing risk patterns of TC landfall cases in South Korea during 1979–2010. In a sense, the importance of track in TC risk assessment is well-acknowledged by many previous studies on TC risk, so that they used track information as the fundamental basis of their risk assessments. For example, Lin and Emanuel (2016) and Hall and Sobel (2013) could evaluate risk of very rare and extreme TCs in specific regions, whose return periods are much longer than the historical records, based on their models generating numerous synthetic tracks using Monte Carlo simulation. The present study, however, tries to directly compare the priority among the risk elements by various statistical analysis of the historical TC records and damage data from South Korea, and explicitly shows the significant track-dependency in TC risk. Moreover, we explain the role of track in the TC risk triangle framework and how TCs having similar intensity and size, but with slightly different track pattern could bring drastically different risk patterns in South Korea.

## **2 Materials and Method**

### **2.1 Data source and Processing**

The present study utilized several data sets: 1) TC information, 2) national survey data of TC damage, 3) national survey data of regional wealth, and 4) surface wind, precipitation observations.

First, TC information including track, intensity, and size were obtained from the Regional Specialized Meteorological Center (RSMC) best-track data. For intensity, we used the maximum wind speed and central pressure data. For TC size, we used the

longest radius of 30 knot winds or greater storm radius data. The 6-hour interval of RSMC was interpolated to a 1-hour interval to obtain precise hazard values at landfall (Park et al. 2011).

Second, damage data were collected from the National Disaster Information Center (NDIC) of the Korean government (<http://www.safekorea.go.kr>) dataset and used after following procedures. NDIC property loss data consists of monetary damages to industrial, public, and private facilities, standardized to the value of money in 2005 and taking inflation into account. The loss data were collected by local governmental offices so that most losses can be collected regardless of whether they were insured and uninsured. There may, however, be some light losses not reported to the local offices by the victims. The raw dataset was including loss data caused by all types of extreme weather such as TCs, heavy rainfall, heavy snowfall, or high waves. Some cases were not classified into specific damage sources, and some cases were stated as high wave damage, when in fact the damage was due to a TC. Therefore, we matched the loss data to each TC affecting Korea, comparing the period of damage in the NDIC dataset with the influence period in the National Typhoon Center (NTC) White Book (NTC 2011) and the period a TC remained within 3° of the Korean coastline. If any day of NDIC damage period overlapped with the RSMC or White Book influence period of a TC, the loss was attributed to the TC. Then, to confine the origin of the loss data to one TC, we excluded cases whose damage period exceed five days from landfall, as NDIC usually aggregates the damage amounts and periods for multiple successive extreme phenomena. “Undamaged” and “Damaged” cases were later categorized based on if there exist any economic loss records reported by NDIC for the given province for the given TC event.

Third, Province-level aggregated wealth data were obtained from government statistical surveys (Korean Statistical Information Service, <http://kosis.kr/>). We aggregated the 17 current districts of South Korea into five provinces, because the administrative division had been changed during 1979–2010, and the size of 17 districts varies from city-size to province-size. The names of the provinces are Gyeong-gi (GG), Chung-cheong (CC), Jolla (JL), Gang-won (GW) and Gyeong-sang (GS) (See Fig. 1 for the distribution of the provinces). These five provinces have independent records of damage for every influential TC case and annual regional wealth. The temporal variation of wealth was considered through normalization of damage data to the reference year of 2005 with wealth per capita ( $\phi$ ). In general, the wealth of South Korea has consistently increased. However, there are significant differences in the growth rates of among provinces ( $\phi$ ), which affect the TC damage records of provinces. By normalization, the possible impact of regional differences in wealth trends was eliminated. The spatial disparity of wealth at a certain time, i.e., 2005, should be addressed when mapping the damage distribution.

Finally, from 60 weather stations throughout South Korea, near-surface wind speed and precipitation intensity were gathered. The influence duration of each station was also calculated by counting the number of days whose daily accumulated precipitation or daily maximum sustained wind speed exceeded the station’s critical thresholds, which we set as the 90<sup>th</sup> percentile of each station. Here, the range of duration was limited by the summation of three relevant periods; 1) the period of warning indicated by NTC White Book (NTC 2011), 2) the number of days that TC stayed within 3° of the Korean coastline based on the RSMC dataset, and 3) the period of damage in the NDIC dataset (See Supplementary Table 3 for the list of the names and years of all the target TCs).

## 2.2 Selection of TC

In the present study, the TCs that influenced South Korea were defined and distinguished. For national influential TC selection, it is common to use a certain criterion based on the distance between a TC center and a certain point on the coastline of interest. For example, influential TCs can be defined as TCs whose center was within 300 km of meteorological offices or coastlines of the given country (e.g., Park et al. 2011; Park et al. 2016). However, sizable TCs can influence farther than the defined criterion (e.g., 300 km), and vice versa. Relatively small TCs also can incur damages over areas larger than their size through storm surges and wind waves. Therefore, we found that using a fixed distance criterion for selecting TCs influential to a society was not the best approach.

The Korea Meteorological Agency (KMA) has considered a TC as affecting Korea when the center of the TC is located in the domain of 32°N–40°N and 120°E–138°E with a high probability of the occurrence of damages to the country. The term ‘probability of occurrence of damages’ depends on a weatherman’s subjective judgment at the time of forecasting. However, weathermen consider all the available observations at that time in making a decision, and this makes the influential TC list by the KMA more comprehensive. Hence, in this study, as in our previous study (Park et al. 2016), the influential TCs are selected by incorporating two complementary references: 1) the objective, strict definition - TCs entering the area within 3° of the coast of Korea and 2) the subjective, comprehensive decision of KMA weathermen - the official influential TC record in the Typhoon White book issued by the NTC (2011). We chose only the TCs that had maximum wind speeds greater than or equal to 17 m s<sup>-1</sup> (above or equal to Tropical Storm (TS) class) at the time of entering the 3° line. For the TCs that are from NTC White Book list and did not approach the 3° delineation, we checked whether their intensity was above TS or not given their maximum wind speeds at the time of their closest approach to the Korean coastline.

In total, 85 TCs were determined to be influential in South Korea from 1979 to 2010. The 85 influential TCs are then grouped according to their track patterns using the fuzzy c-means clustering method (FCM). The FCM is widely used for classifying widespread data with amorphous boundaries. Some previous studies have shown this method to be effective for grouping TC track patterns (e.g., Kim et al. 2011).

## 2.3 Statistical Analysis

### 2.3.1 Data mining methods

We clustered the track patterns, not for the whole tracks from genesis to disappearance, but for the part of the tracks in the domain of 28°N–40°N and 120°E–138°E (grey boxes in Fig. 1) so that we could divide tracks focusing on the paths near South Korea, whose national TC risk distribution was examined with respect to these clustered track patterns. The TCs were grouped into four types. The optimum cluster number were decided by five validity measures - partition coefficient, partition index, separation index, Xie and Beni index, and Dunn index (Kim et al. 2011), and four appeared to be the optimum number in our case. All of the five indexes still pointed to four as the optimum number of clusters even made slight changes to TC lists, such as different time frame (e.g., 1979–2015) or different clustering domain (e.g., 5° area from the Korean Peninsula

coastline). We checked the temporal trend and monthly distribution of the number of TCs, and they did not show any significant trend or discernable differences among the clusters.

We further introduced the decision tree analysis to decompose the relationships among risk elements. The decision tree method, a multi-variable technique, allowed us to explain, describe, classify, or predict a target as a result of the combined effects of multiple input variables beyond a one-cause and one-effect relationship. Compared to other multi-variable techniques, the decision tree method has its own advantage in that it is easy to use, robust with a variety of data, and most of all, intuitively interpretable. It helps decision analysts to structure the decision process in a graphical sequence.

Among several famous decision tree algorithms, this study applied See5/C5.0 as a classification method for TC risk materialization. The See5/C5.0 algorithm is an improved version of C4.5 (Quinlan 1993) in terms of accuracy, speed, and computer memory consumption. Also, C4.5 algorithm is advantageous because it can accommodate all of the class, binary, and continuous variable types that we needed (see Supplementary Table 1). See5/C5.0 calculates the information gain at each node based on the entropy concept in order to choose the most efficient attribute for splitting the training samples into two branches.

To prevent over-fitting, we introduced pruning and cross-validation. First, we required that branches have a sample size of at least five. The number five was determined through retrospective pruning process. Second, a ten-fold cross-validation was conducted, and we checked that the decision tree results (e.g., model accuracy, tree size, or attribute usage) were stable and consistent given the ten different cross-validation sets.

### 2.3.2 Significance tests

For all the statistical analysis of risk comparison among track groups, nonparametric methods were used (Sawilowsky 1990). Medians were used rather than means, and rank-based procedures were conducted for any significance test. This is because we cannot regard the TC damage as following the normal distribution; rather, damage shows an extreme distribution. Thirty percent of TCs have zero losses recorded, and 30% of all the accumulated damages are attributed to a single TC, typhoon Rusa in 2012. The Kruskal-Wallis test, or the one-way Analysis of variance (ANOVA) on ranks, was used to determine if there are statistically significant differences for a variable between track-groups. Spearman's rank correlation coefficient, which measures the linear relationships between the rankings of two variables, was used instead of the most commonly used Pearson product-moment correlation coefficient, which measures linear relationships between the raw values of two variables.

## 3 Results

In order to objectively evaluate the effect of each TC track on damage, a total of 85 TCs that influenced South Korea during 1979–2010 were grouped into four track patterns. The four TC track patterns can be characterized as 1) east-short, 2) east-long, 3) west-long, and 4) west-short types based on the position and length of the TC tracks around the Korean Peninsula

(Fig. 1). Although zonal distances between east-types (i.e., east-short and east-long) and west-types (i.e., west-short and west-long) are only about 250 km, the estimated hazards (both TC-based and active) and damages caused by the TCs are significantly different depending on the four TC track patterns at the 99% confidence level based on the Kruskal-Wallis test (Fig. 2). This is striking because the short distance around 250 km is somewhat trivial considering that average track errors in the northwest Pacific, as determined by many frequently used dynamic or statistical-dynamical techniques, are about 200 and 400 km for 24 and 48 h, respectively (Roy and Kovordanyi 2012).

As expected, TC-based hazards display different results from active hazards. Although TC-based hazard parameters have been commonly used as the sole indicators in TC risk analysis (e.g., Nordhaus 2010; Hsiang and Narita 2012; Czajkowski and Done 2014; Zhai and Jiang 2014), it shows poor accordance with damage, especially when compared to active hazard (Fig. 2). TCs with longer tracks (i.e., east-long and west-long) show greater TC-based hazards, such as deeper central pressure, higher maximum wind speed, and larger size (Figs. 2A – C). The TC-based hazard ranking is as follows: east-long, west-long, west-short, and east-short track patterns. A TC with a stronger intensity should be more durable and able to maintain its system compared to weaker TCs under the same conditions of friction, shear, and/or energy source. In this sense, the longer tracks generally have higher wind speeds and deeper central pressures (e.g., Kim et al. 2011). In our analysis, more intense TCs appear to have a larger storm radius, but TC size is not necessarily correlated with intensity (Weatherford and Gray 1988; Knaff et al. 2014). High correlation (correlation coefficient  $r = 0.58$ ) between size and intensity may be one of the characteristics of TCs that have affected Korea. On the other hand, for active hazards, west-types generally have higher values than east-types (Figs. 2D – F). For accumulated precipitation and influence duration, and damage, the ranking is in order of west-short, west-long, east-long, and east-short track patterns. For near-surface wind, the ranking is in order of west-long, west-short, east-long, and east-short. The damage ranking is in order of west-short, west-long, east-long, and east-short track patterns, which is exactly the same as the ranking for accumulated precipitation and influence duration. In addition, all of the active hazard parameters considered here show much higher correlations with damages than TC-based potential hazard parameters, even if most of TC-based hazards display statistical significance at the 95% or 99% confidence (Table 1). The average of absolute correlation coefficient ( $|r|$ ) for all active hazards and for all track patterns is 0.62, while that of potential hazards is just 0.29. Among the active hazards, higher correlation coefficients for accumulated precipitation and affected duration compared to near-surface wind imply that rainfall can better account for damage than wind in the study area, as suggested by some previous studies (Lin et al. 2002; Park et al. 2016).

The discrepancy of the rankings between TC-based hazards and active hazards can be explained in terms of TC tracks. By examining the spatial distribution of local active hazards with respect to topography for the four track patterns (Fig. 3), we find evidence that there is a close relationship between track patterns and active hazards. We can see that track pattern determines the affected location; west-types have higher active hazards in western areas compared to the east-types (Fig. 3). Furthermore, as seen in Fig. 2, comparable near-surface wind speed in west-types and east-long tracks is found over the entire country, particularly along the coast (Figs. 3A – D) despite the fact that TC-based hazards of west-types are weaker than those of east-long (Figs. 2A – F). This is attributed to the fact that west-types cause the Korean Peninsula to fall into a

dangerous semicircle (right-hand side of TC center), in which the TC translation speed and rotational wind field are additive, whereas east-types place the country in a navigable semicircle (left-hand side semicircle). Furthermore, near-surface wind induced by east-types generally does not get to the two mountains of Sobaek and Taebaek (see Fig. 1C) owing to the long distance between the storm center and the mountains. Much heavier precipitation is found in west-types compared to east-types, particularly along the eastern side of the mountainous area (Figs. 3E – H). This is attributed to longer influence duration and orographic forcing (Figs. 3E – L). West-types may have longer influence duration compared to east-types since most TCs move from west to east in the mid-latitudes (see Fig. 1). For the same reason, accumulated precipitation throughout the influence duration can be also higher for the west-types than the east-types. Actually, the spatial patterns of influence duration and precipitation are almost coincident. One more possible cause of heavy precipitation in west-types is the orographic impact of southern and eastern mountains on precipitation (i.e., the Sobaek and the Taebaek mountains). When a TC is located in southwest South Korea, the east sides of the mountains become the upstream slope of the tangential wind of the TC such that precipitation induced by the TC can be even more than the inherent rainband of the TC (e.g., Park and Lee 2007; Lin et al. 2002). Thus, the orographic impact can be maximized by west-types but not by the east-types. The precipitation impact of the west-short group is even larger than the west-long group, and this point can be partly explained by the translational speed of a TC. The west-short type is the slowest among the four track groups (significant at the 95% confidence level, Kruskal-Wallis test), and thus it can stay the longest over the country, as shown in Fig. 3L. Moreover, it is possible that the west-short has the largest interaction time with the mountains. All of the above results suggest that track effectively bridges the gap between a TC in potential state and the active hazards.

In addition to the ranking, the map of economic loss (Fig. 4) also matches that of the active hazards well (compare Fig. 3 and Fig. 4). Southern parts, where TCs hit most frequently and both precipitation and wind speed were high, are the riskiest regions among all the groups. However, as Fig. 4 shows, for west-types, southwestern provinces (Jeonla, JL) records less economic losses than southeastern provinces (Gyeongsang, GS) despite the fact that more hazardous wind and rainfall were recorded in JL than GS (Fig. 3). This discordance is partly explained by an exposure disparity. GS possesses higher wealth compared to JL. After the damage is divided by regional wealth, the regional risk distribution becomes more analogous to the hazard distribution (as seen in parentheses in Fig. 4), although the damage in GS is still slightly higher than in JL for the west-short tracks (Fig. 4D). The unexplained parts of hazard–risk discordance may be related to regional vulnerability to TCs. More mountainous area in GS might be a possible reason that the vulnerability of GS is higher than that of JL. These results demonstrate that exposure/vulnerability factors are still important in TC risk analysis in addition to track and hazard. In this regard, we conducted a decision tree analysis considering all three risk elements plus track to determine which element is the primary factor for assessing TC risk.

Through the decision tree analysis, the following three questions could be answered: 1) what is the most effective factor for classifying damaged and undamaged TC cases, 2) how do different factors in combination determine damage occurrence, and 3) what are the critical values of the factors that can be used as quantitative guidelines related to TC damage occurrence. Here, the decision tree model was designed to objectively classify whether a TC will bring damage to a province or not; the

decision tree uses best-track data based on TC information (maximum wind speed, minimum central pressure, size, and track) as input variables (see Supplementary Table 1 for more information of input variables). Overall, we have 355 effective cases, comprising 160 damaged cases and 195 undamaged cases from five provinces (see Supplementary Table 2 for detailed information of damage cases). Once again, undamaged cases refer to the cases that the no economic losses have been recorded from the given settlement for the given period under the TC's influence by the official disaster information center of Korean government.

According to the decision tree, track information acts as the primary determinant of TC risk. Information of track pattern is nominated as the first split attribute in the best-track based decision tree (see Fig. 5). This means all 355 cases should be classified by track group prior to all other decision nodes in order to reach the end nodes (damage occurrence). The detailed process is as follows. First, the model simply sends all west-type TCs to the end node of "damaged." On the other hand, the model requires east-type TCs to satisfy additional criteria to determine damage occurrence. At the national scale, it is true that all west-type TCs have caused damages, and all TCs that did not generate damage in any of the provinces in South Korea were from either east-short or east-long (see Supplementary Table 2). However, at the province scale, there were 32 undamaged cases among the 110 west-type cases. Next, the east-type TC cases are assessed according to province and TC intensity (maximum wind speed). For a TC in the east-long group to cause damage, it should be in JL, GS, or GW province. Additionally, for GW province, the maximum wind speed should be greater than  $41.1 \text{ m s}^{-1}$ . East-short cases, unlike east-long cases, are sent to the intensity criterion before the province criterion. East-short TCs with weak intensity (maximum wind speed less than  $25.7 \text{ m s}^{-1}$ ) are directly linked to the "undamaged" node. The east-short TCs with satisfactory intensity (maximum wind speed greater than  $25.7 \text{ m s}^{-1}$ ) are sent to the province criterion; the critically strong east-short TCs can incur damage in the southern provinces (See Fig. 5).

The relative importance of the variables in each decision tree is offered quantitatively in terms of the usage rate by the See5/C5.0 algorithm. When an attribute is the most-related variable to the target variable, the attribute should be used most frequently for classification by a decision tree model. In our decision tree (Fig. 5), the track group variable is used 100% of the cases; then, province and best track maximum wind follow with usage rates of 48% and 37%, respectively. Therefore, we can say that for risk determination, TC track, other than the local surface observations, is the most important attribute. TC intensity information is the third most important attribute. TC size is not utilized by any model as an effective classifier to distinguish between damaged and undamaged cases.

On the other hand, the use of the province variable by the TC best-track based decision tree is mainly related to the relative location of the province with respect to the TC center along the track. Southern provinces are generally closer to the TC center regardless of the four track types because TCs move from the south (low-latitude) to the north (high-latitude). However, in addition to the distance from the hazard center, the province variable can be interpreted as showing the exposure and/or the vulnerability of the provinces. If we put local hazard parameters into the decision tree model, we find that the province information acts as exposure/vulnerability (Supplementary Fig. 1). In the local-hazard based decision tree model, only the northwestern provinces are determined to be "damaged" with the same wind and precipitation conditions



(wind weaker than  $12.2 \text{ m s}^{-1}$  and precipitation greater than  $70.3 \text{ mm per day}$ ). We interpret that exposure to be the highest in the northwestern province considering that this region is the wealthiest and most populated area over South Korea.

#### 4 Summary and Disucssion

The present study firstly compares and contrasts the amount and spatial distribution of risk elements with respect to TC track types for TCs that influenced South Korea over the last three decades. The results show that TC damage is about two times more correlated with local active hazard compared to potential hazard. Then, track is suggested as the main reason why the localized active hazard substantially disagrees with potential hazard. The track is a sequence of TC locations, and location differences cause substantial changes to local active hazard distributions in association with 1) dangerous/navigable semicircle side differences, 2) interaction with inhomogeneous topography, and 3) duration of influence changes. Second, we analyzed the priority structure of the TC risk determination process, including track as an independent factor through decision tree analysis. When local active hazard information is missing, TC track acts to bridge the information gap between the TC system and local risk. TC intensity or size information plays peripheral roles for filling the information gap.

A flowchart of TC risk materialization is drawn as a concluding remark (See Fig. 6). This framework includes track as a separate risk element, and it implies that TC hazard can be activated only through track, in other words, only through the “conflicts at the interface between geophysical processes and human societies” (Alexander 2000). Then, the risk triangle is applied not to the dormant hazard (TC intensity and size) but to active hazard, which is a product of a combination of TC characteristics (i.e. dormant hazards) and local geography experiences through track. Note that not only local geography experience is dependent on track patterns, but TC characteristics also appeared to differ among track patterns (Figs. 2A – C). Therefore, we suggest that the integral TC hazard is highly dependent on track. One may argue that if track is taken into account within a risk framework, it overlaps with the concept of exposure. This rebuttal can arise from the fact that both parameters are linked to the possibility of actual influence by the TC. However, exposure refers to population and wealth at a certain fixed point. Track determines active hazards, i.e., how much local wind and rainfall at a fixed point will be caused by a TC with a given track. Track, however, cannot change exposure at a given point.

The small longitudinal track deviations, less than  $250 \text{ km}$ , could distinctively change the damage maps for South Korea. It implies that, even for the national scale (especially for countries having complex terrains like the Korean Peninsula), it is required to consider this track-dependency in risk realization for a reliable TC risk assessment rather than to formulate a statistical relationship between the whole national amount of damage from a TC and the intensity or size of the TC (e.g. Nordhaus 2010; Hsiang and Narita 2012). Thus, using the local active hazard parameters (rainfall, wind, and surge) would be the best to make the risk equations. When the historical observations of local hazards are not available for the case, one can adopt reconstructed active hazard values by simple statistical and dynamic models that, in turn, calculate the active hazard values from track, local geography and TC information [e.g., R-CLIPER model for rainfall (Cheung et al. 2008); SLOSH

model for storm surge (Lin and Emanuel 2016)]. Having this additional step between dormant hazard and active hazard, similar to a downscaling for numerical models, would much improve the accuracy of the risk assessment.

A cautious consideration of track variances is also suggested for the climate change scale future TC damage projection studies. Anthropogenic contributions to TC track changes have been reported (Colbert et al. 2013; Park et al. 2014; Kossin et al. 2016). Also, it has been shown that interdecadal changes of TC tracks in the western North Pacific are associated with the westward expansion of the subtropical northwestern Pacific high (Ho et al. 2004). Mendelsohn et al. (2012) estimated future TC damage including the track variance in future climate by combining the economic models with the statistical/deterministic hurricane model generating synthetic TCs under given future climate environments. Still, this method does not model the hazard activation, which differs sensitively depending on the landfall locations and from the TC approaching directions to the locations. Thus, we suggest that the uncertainty rising from track changes in future climate should be considered more in details for TC risk research, as it is shown that a slight change in TC track distribution can cause the total amount of damage to be much larger or smaller given the same number and intensity of TCs in a particular basin.

As a final implication, our findings highly support the needs of the impact-based forecast which has been recently promoted by WMO (2015). The impact-based forecast is a transition from weather forecast to warning services, on the basis that warnings should deliver information that is directly related to possible damages. It has been reported that avoidable casualties and economic losses have been caused by hurricanes and cyclones because the forecasts had an excessive focus on the intensity of the cyclones rather than the local active hazards caused by them (Vinet et al. 2012; Meyer et al. 2014). Communication on TC risk should focus more on actual impacts that residents in different areas are likely to experience. In accordance, our results point out that local hazards, penetrating into the residents, often fairly disagree with the maximum intensity of the TC, observed in the small area near the center. Also, the methodology (combining track clustering and decision tree model) and the regional threshold values for damage occurrence presented here, could be practically used for the impact-based forecasts for TCs over South Korea.

### **Acknowledgement**

This work was supported by the National Research Foundation of Korea (NRF-2013R1A1A2064211) and the Korea Ministry of Environment (MOE) as “Climate Change Correspondence Program.”

### **References**

Alexander, D.: *Confronting catastrophe: New perspectives on natural disasters*. New York: Oxford University Press, 283 pp, 2000.

- Cardona, O. D, and coauthors: Determinants of risk: Exposure and vulnerability in managing the risks of extreme events and disasters to advance climate change adaptation. A special report of working groups I and II of the IPCC, Cambridge University Press, 65-108, 2012.
- Cheung, K. K. W., Huang, L.-R., and Lee, C.-S.: Characteristics of rainfall during tropical cyclone periods in Taiwan, *Nat. Hazards Earth Syst. Sci.*, 8, 1463-1474, doi:10.5194/nhess-8-1463-2008, 2008.
- Ching, P. K., de los Reyes, V. C., Sucaldito, M. N., and Tayag, E.: An assessment of disaster-related mortality post-Haiyan in Tacloban city. *Western Pacific Surveillance and Response Journal*, 6, 34–38, doi:10.5365/wpsar.2015.6.2.HYN\_005, 2015.
- Colbert, A. J., Soden, B. J., Vecchi, G. A., and Kirtman, B. P.: The impact of anthropogenic climate change on North Atlantic tropical cyclone tracks. *J. Climate*, 26, 4088–4095, doi: 10.1175/JCLI-D-12-00342.1, 2013.
- Czajkowski, J., and Done, J.J.: As the wind blows? Understanding hurricane damages at the local level through a case study analysis. *Wea. Climate Soc.*, 6, 202–217, doi: 10.1175/WCAS-D-13-00024.1, 2014.
- Hall, T.M., and Soble, A. H.: On the impact angle of Hurricane Sandy's New Jersey landfall. *Geophys. Res. Lett.*, 40, 2312-2315, doi:10.1002/grl.50395, 2013.
- Ho, C.-H, J.-J. Baik, J.-H. Kim, D.-Y. Gong, and C.-H Sui, 2004: Interdecadal changes in summertime typhoon tracks. *J. Climate*, 17, 1767–1776, doi:10.1175/1520-0442(2004)017<1767:ICISTT>2.0.CO;2
- Hsinag, S. M., and Narita, D.: Adaptation to cyclone risk: Evidence from the global cross-section. *Climate Change Economics*, 3, 1250011, doi:10.1142/S201000781250011X, 2012.
- Kim, H.-S., Kim, J.-H., Ho, C.-H., and Chu, P.-S.: Pattern classification of typhoon tracks using the fuzzy c-means clustering method. *J. Climate*, 24, 488–508, doi:10.1175/2010JCLI3751.1, 2011.
- Knaff, J. A., Longmore, S.P., and Molenaar, D. A.: An objective satellite-based tropical cyclone size climatology, *J. Climate*, 27, 455–476, doi:10.1175/JCLI-D-13-00096.1, 2014.
- Kossin, J, Emanuel, K. A., and Camargo, S. J.: Past and projected changes in western North Pacific tropical cyclone exposure, *J. Climate*, 29, 5725-5739, doi:10.1175/JCLID-16-0076.1, 2016.
- Lin, N., and Emanuel, K.A.: Grey swan tropical cyclones. *Nat. Climate Change*, 6, 106-111, doi:10.1038/nclimate2777, 2016.
- Lin, Y.-L., Ensley, D. B., Chiao, S., and Huang, C.-Y.: Orographic influences on rainfall and track deflection associated with the passage of a tropical cyclone. *Mon. Wea. Rev.*, 130, 2929–2950, doi:10.1175/1520-0493(2002)130<2929:OIORAT>2.0.CO;2, 2002.

- MacCollum, D.: Construction safety engineering principles: designing and managing safer job sites. McGraw-Hill Education, 416 pp, 2007.
- Mendelsohn, R, Emanuel, K. A., Chonabayashi, S., and Bakkensen, L.: The impact of climate change on global tropical cyclone damage. *Nat. Climate Change*, 2, 205–209, doi: 10.1038/nclimate1357, 2012.
- 5 Meyer, R. J., Baker, J., Broad, K., Czajkowski, J., and Orlove, B.: The dynamics of hurricane risk perception, real-time evidence from the 2012 Atlantic hurricane season, *Bull. Amer. Meteor. Soc.*, 95, 1389–1404, doi: 0.1175/BAMS-D-12-00218.1, 2014.
- National Typhoon Center: Typhoon White Book (in Korean). The Korean Meteor. Administration, 358 pp, 2011.
- Nordhaus, W. D.: The economics of hurricanes and implications of global warming. *Climate Change Economics*, 1, 1–20,  
10 doi:10.1142/S2010007810000054, 2010.
- Park, D.-S. R., Ho, C.-H., Kim, J.-H. Kim, and Kim, H.-S.: Strong landfall typhoons in Korea and Japan in a recent decade, *J. Geophys. Res.-Atmos.*, 116, Artn D07105., doi: 10.1029/2010JD014801, 2011.
- Park, D.-S. R., Ho, C.-H., and Kim, J.-H.: Growing threat of intense tropical cyclones to East Asia over the period 1977–2010. *Environ. Res. Lett.*, 9, Artn 014008, 2014.
- 15 Park, D.-S. R., Ho, C.-H., Nam, C. C., and Kim, H.-S.: Evidence of reduced vulnerability to tropical cyclones in the Republic of Korea. *Environ. Res. Lett.*, 10, Artn 054003, 2015.
- Park, D.-S. R., Ho, C.-H., Kim, J., Kang, K.-R. and Nam, C. C.: Highlighting socioeconomic damages caused by weakened tropical cyclones in the Republic of Korea, *Natural Hazards*, 82, 1301-1315, 2016.
- Park, S. K., and Lee, E.: Synoptic features of orographically enhanced heavy rainfall on the east coast of Korea associated  
20 with Typhoon Rusa (2002), *Geophys. Res. Lett.*, 34, Artn L02803, doi: 10.1029/2006GL028592, 2007.
- Peduzzi, P., Chatenoux, B., Dao, H., De Bono, A., Herold, C., Kossin, J., Mouton, F., and Nordbeck, O.: Global trends in tropical cyclone risk. *Nat. Climate Change*, 2, 289–294, doi: 10.1038/nclimate1410, 2012.
- Pielke, R. A. Jr., Gratz, J., and Landsea, C. W.: Normalized hurricane damages in the United States: 1900–2005. *Natural Hazards Review*, 9, 29–42, doi: 10.1061/(ASCE)1527-6988(2008)9:1(29), 2008.
- 25 Quinlan, J. R.: C4.5: programs for machine learning. Morgan Kaufmann Publisher Inc., 270 pp, 1993.
- Roy, C., and Kovordanyi, R.: Tropical cyclone track forecasting techniques—A review, *Atmos. Res.* 104-105, 40–69, doi: 10.1016/j.atmosres.2011.09.012, 2012.
- Sawilowsky, S. S.: Nonparametric-tests of interaction in experimental-design. *Rev. Educ. Res.*, 60, 91–126, 1990.

Vinet, F., Lumbroso, D., Defossez, S., and Boissier L.: A comparative analysis of the loss of life during two recent floods in France: the sea surge caused by the storm Xynthia and the flash flood in Var. *Natural Hazards*, 61, 1179–1201, doi: 10.1007/s11069-011-9975-5, 2012.

World Meteorological Organization: WMO Guidelines on Multi-hazard Impact-based Forecast and Warning Services. 23 pp, 2015.

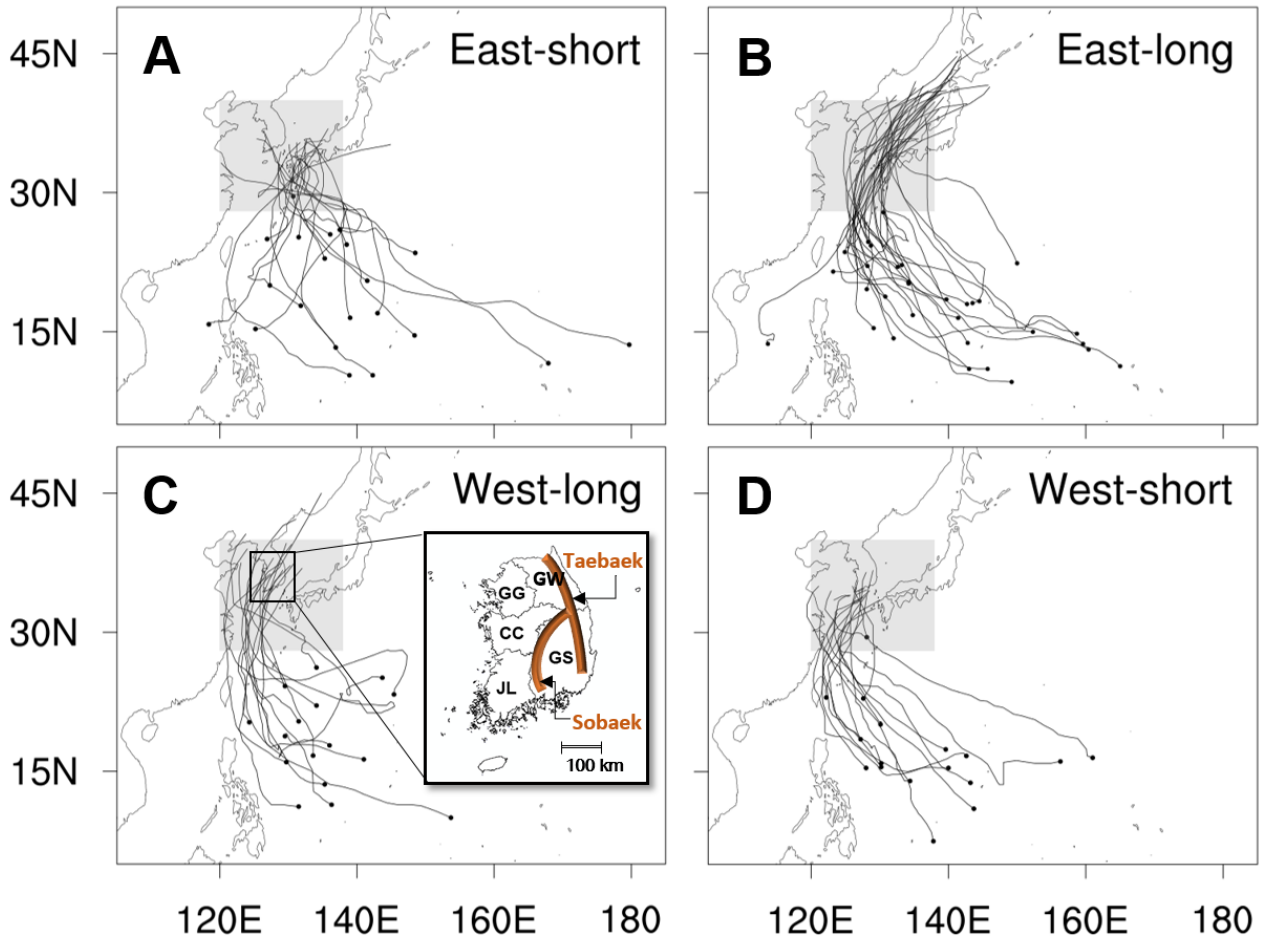
Weatherford, C. L., and Gray, W. M.: Typhoon structure as revealed by aircraft reconnaissance. Part II: Structural variability. *Mon. Wea. Rev.*, 116, 1044–1056, doi:10.1175/1520-0493(1988)116<1044:TSARBA>2.0.CO;2, 1988.

Zhai, A. R., and Jiang, J. H.: Dependence of US hurricane economic loss on maximum wind speed and storm size. *Environ. Res. Lett.*, 9, Artn 064019, 2014.

**Table 1: Each Spearman’s correlation coefficient of property losses with active and TC-based hazards.** Active hazards are maximum daily wind speed, maximum daily precipitation, and the sum of influenced periods for all 60 weather stations. TC-based, potential hazards are maximum wind speed, central pressure, and storm radius (30 knot) based on the RSMC best-track data for each track group. The significances of correlations are shown with asterisks.

	Four track groups				All
	East-short	East-long	West-long	West-short	
<b>Active hazard parameters (from weather station)</b>					
Daily max wind speed	0.45**	0.58**	0.66**	0.59**	0.62**
Daily precipitation	0.37**	0.66**	0.74**	0.80**	0.71**
Influence duration	0.48**	0.76**	0.59**	0.78**	0.76**
<b>Potential hazard parameters (from best-track data)</b>					
Maximum wind speed	0.39**	0.17*	0.27*	0.39**	0.29**
Central pressure	-0.40**	-0.16*	-0.35**	-0.41**	-0.27**
Storm radius	0.39**	0.08	0.16	0.30*	0.24**

\* Significant at the 95%, \*\* significant at the 99% confidence levels.

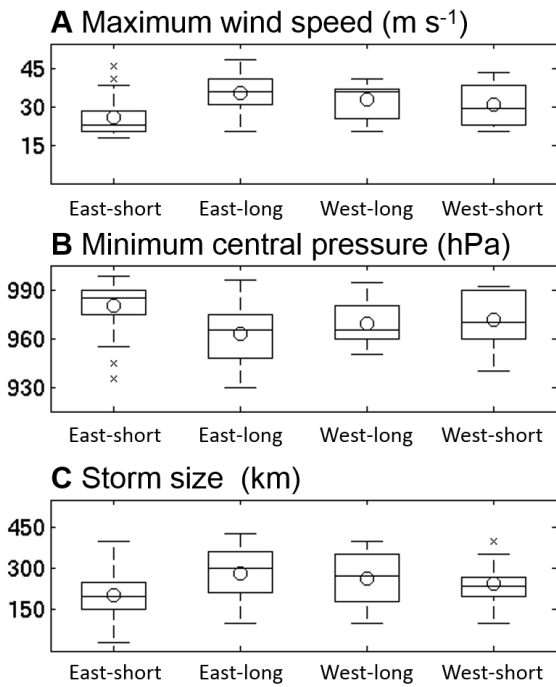


**Figure 1: Four groups of tropical cyclone tracks that made landfall over South Korea during 1979–2010.** Box shaded in grey, covering 28N – 40N and 120E – 138E, indicates the clustering domain for the fuzzy c-means clustering method. A map of the five aggregated provinces of South Korea is displayed in (C): Gyeong-gi (GG), Chung-cheong (CC), Jolla (JL), Gang-won (GW), and Gyeong-sang (GS). The Taebaek and Sobaek mountains are drawn with orange lines.

5

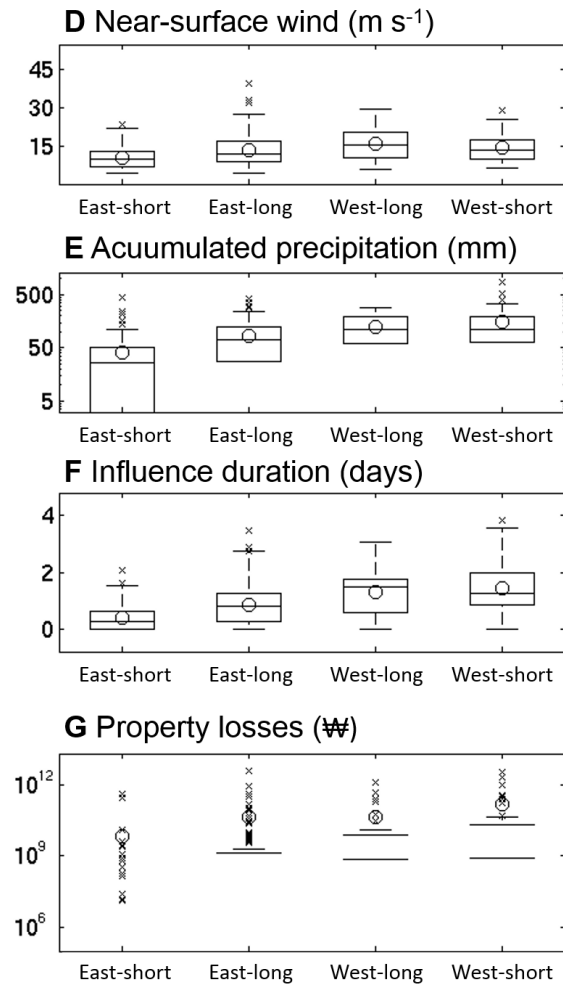
## RSMC best-track

### Parameters for potential hazard



## Regional observation

### Parameters for active hazard & damage



**Figure 2: Boxplots of the hazards and damages of track-pattern groups.** (A) Maximum wind speed, (B) central minimum pressure, and (C) storm size from RSMC best-track data at the point that the TCs entered the area within  $3^\circ$  of the coast of the Korean Peninsula, or for the TCs that did not enter the area within  $3^\circ$ , when they were closest to South Korea. (D) Daily maximum wind speed (10 min average), (E) daily accumulated precipitation, (F) influence duration from 60 weather stations, and (G) property losses ( $1\$ \approx 10^3\text{₩}$ ) over South Korea. The storm size is the longest radius of 30 knot winds or greater. Station maximum wind speed and precipitation are one maximum daily value in the whole influence duration. The upper (bottom) side of each box is the third (first) quantile. The bar inside the box represents the median, and the circle represents the mean. Outliers, which are located outside of the maximum whisker length, are drawn as 'x.' The maximum whisker length is 1.5 times the value of the third quantile minus the first quantile. The plotted whiskers extend to the adjacent value, which is the most extreme data point that is not an outlier.

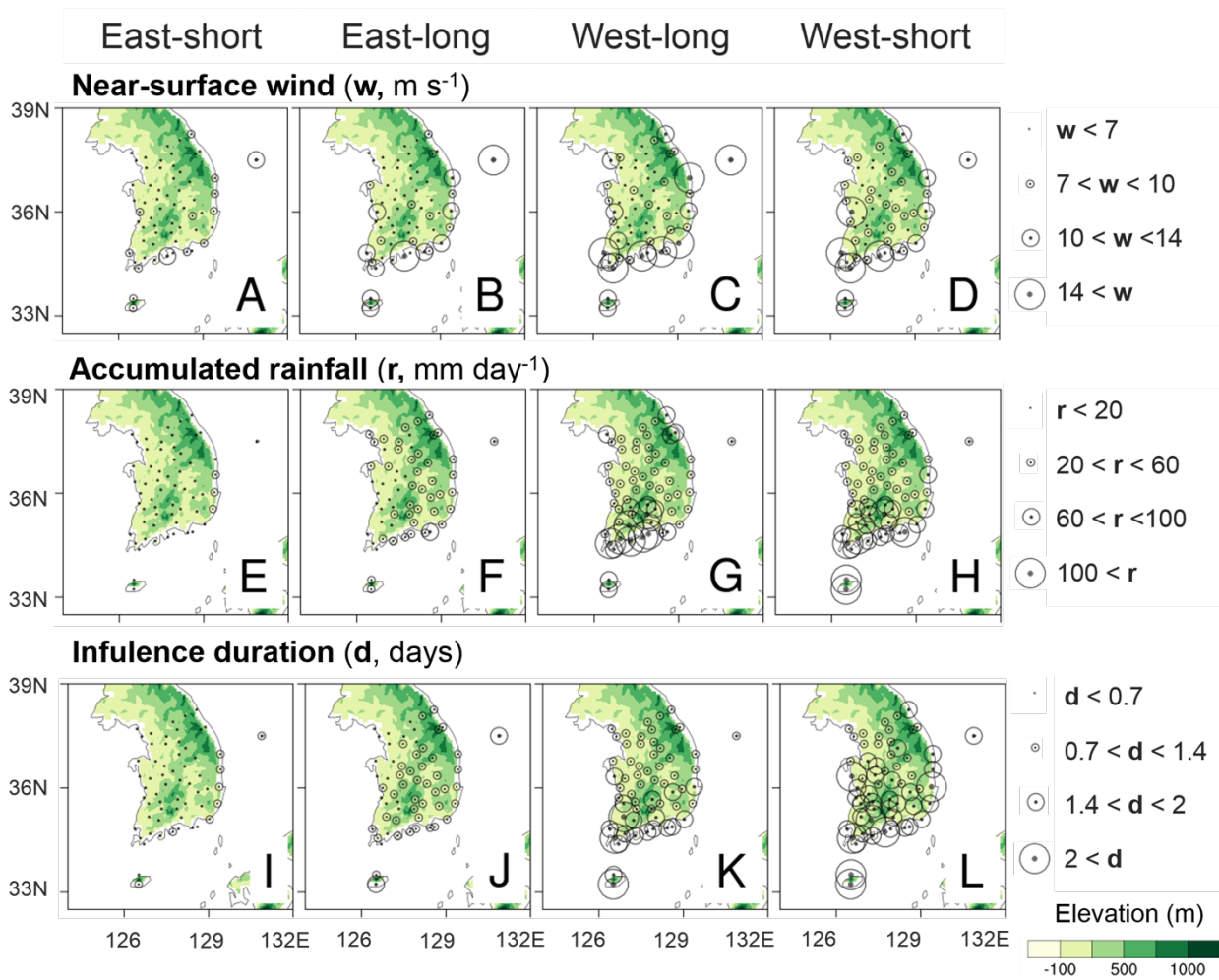
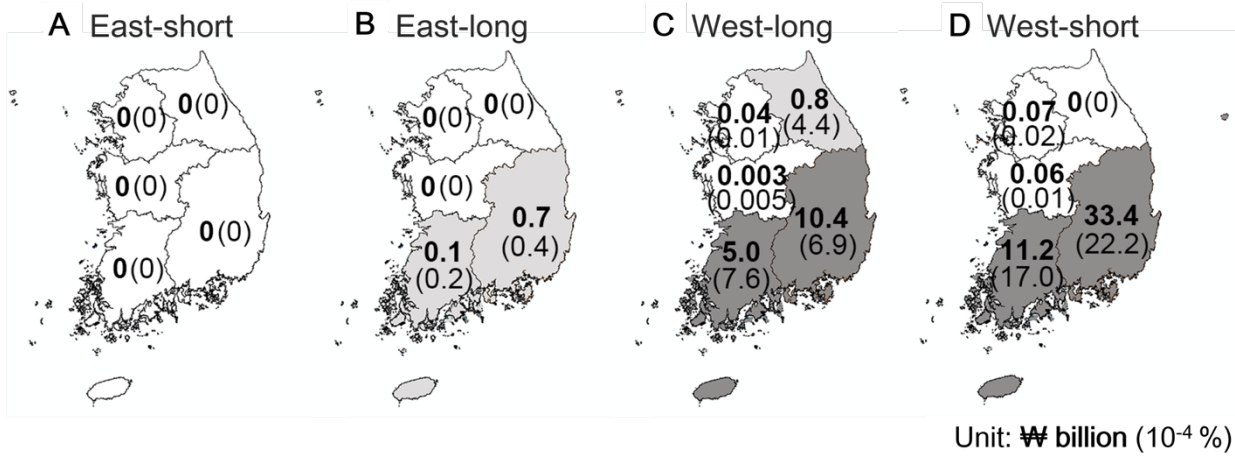


Figure 3: Three active hazard parameters – wind, precipitation, and duration, of tropical cyclones for each track-group observed at 60 weather stations.



## Economic losses (per regional wealth)



**Figure 4: Medians of regional economic losses from a tropical cyclone (regional economic losses divided by regional wealth).** The dark shading indicates provinces having median losses larger than ₩(KRW) one billion, and the light shading indicates provinces having median losses larger than ₩ 0.1 billion and smaller than ₩ one billion. More than the half of the east-short TCs are undamaging TCs, so the property loss medians of all provinces are zero.

# TC best-track based decision tree

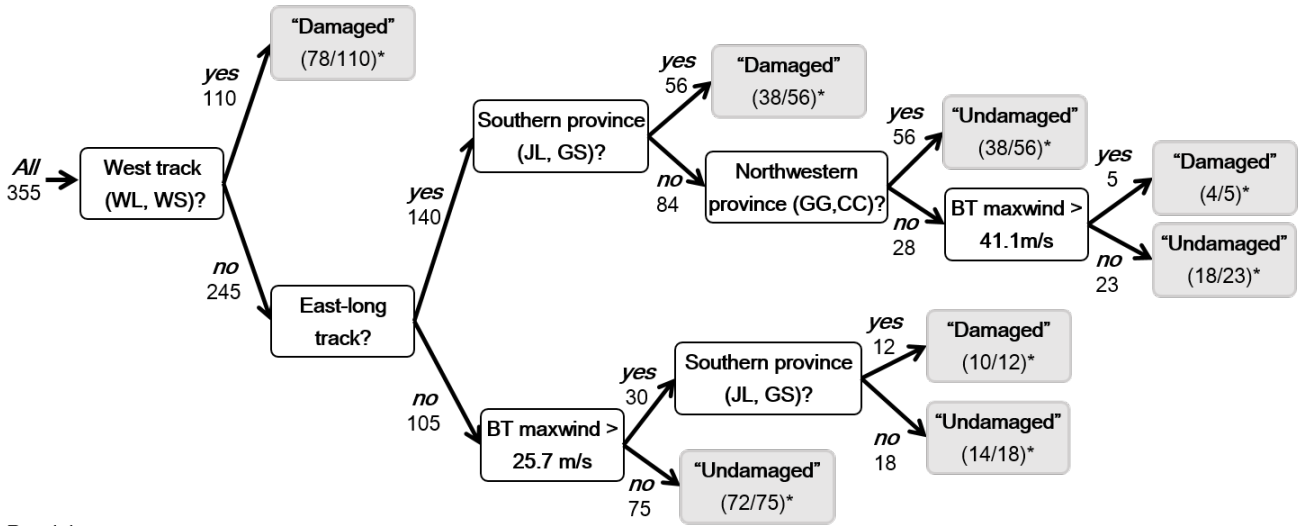


Figure 5: Decision tree model for damage occurrence using the four TC best-track attributes (maximum wind speed, central pressure, storm size, and track-group) and province information as input variables. The number of cases corresponding to each criterion is presented along each arrow. The shaded boxes indicate the final diagnosis boxes, in which the precision of the diagnosis is written in parentheses (the number of correctly identified cases / the number of cases diagnosed following the specific sequence of criteria).

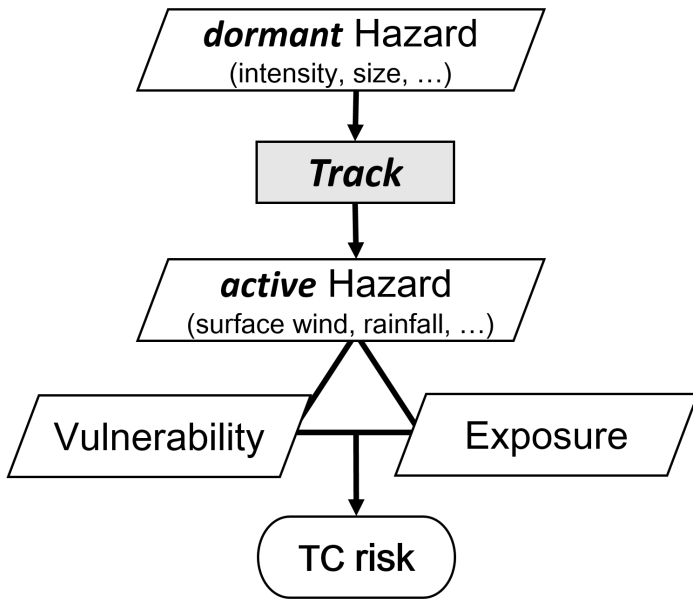


Figure 6: Flowchart for local risk materialization process with TC risk elements and their relationship.