



Combining hazard, exposure and vulnerability data to predict historical North Atlantic hurricane damage

Alexander F. Vessey^{1*}, Alexander J. Baker^{2,3*}, Vernie Marcellin-Honore², and James Michelin²

¹AXA XL, 20 Gracechurch Street, London, United Kingdom

²Department of Meteorology, University of Reading, Reading, United Kingdom

³National Centre for Atmospheric Science, Reading, United Kingdom

Correspondence to: Alexander F. Vessey (alecvessey@hotmail.co.uk) and Alexander J. Baker (alexander.baker@reading.ac.uk)

Abstract. Hurricanes are among the most destructive natural hazards globally. Accurate risk assessment requires integrated hazard, exposure, and vulnerability information, yet the widely used Saffir–Simpson scale, while an effective public-communication tool, is based on a single hazard quantity (wind speed) and is not well correlated with historical economic losses, limiting its predictive value. This study develops a statistical model to predict economic damage from landfalling North Atlantic hurricanes using optimally weighted, normalised-rank variables representing hazard, exposure, and vulnerability. The model significantly reduces root-mean-square error between predicted and observed losses from US\$35.6 billion (when using landfall wind speed) to US\$7.0 billion, and substantially outperforms single-parameter predictions, including landfall wind speed maxima and central pressure minima. To improve communication of financial risk, we introduce a loss-based ‘Hurricane Predictive Damage Scale’ to more directly link hurricane characteristics to economic impacts. Our results demonstrate that integrating exposure and vulnerability data with hazard observations yields markedly better estimates of historical hurricane economic impacts, and this approach is readily applicable to future forecast hurricanes, allowing assessment of how damage from an imminent landfall may rank among historical events. This framework is transferable to other cyclone-prone regions and highlights the critical need for open exposure and vulnerability data to advance climate risk quantification and inform policy.

Significance statement. Hurricanes are a destructive natural hazard. Saffir–Simpson category is used to convey impact but is not well correlated with losses. We combined hazard, exposure, and vulnerability data to predict losses from North Atlantic hurricanes. Our model significantly reduces errors between predicted and observed losses and is more skilful than hazard-only predictions. We introduced a loss-based scale to link hurricane characteristics to economic impacts to help quantify climate risk and inform policy.



1 Introduction

35 Intense tropical cyclones are the most damaging meteorological hazard worldwide (Gallagher Re, 2025; Willis Towers
Watson, 2024, AON, 2024). Between 1980 and 2025, global economic damage due to tropical cyclone (TC) landfalls totals
U.S.\$ 2.9 trillion, primarily reflecting building damage (National Oceanographic and Atmospheric Administration, 2024). In
addition to financial damage, tropical cyclones also pose a significant threat to life. The most important threats in the United
States being from storm surge and intense rainfall (Rappaport, 2014), and the average TC leads to an estimated mortality
40 burden of 7,000–11,000 mostly non-immediate excess deaths (Young and Hsiang, 2024).

Studies of hurricane damage are challenging because damage depends primarily on lower-frequency and impactful
landfalling TCs, so uncertainties are generally higher than those associated with basin-wide metrics of TC activity (Emanuel,
2011). The most financially damaging U.S. hurricanes include the Great Miami Hurricane (1926), Katrina (2005), and
Harvey (2017). There is also evidence that economic financial damage from hurricanes has increased over time (Grinsted *et*
45 *al.*, 2019; Klotzbach *et al.*, 2022b), highlighting the urgency of understanding impacts and losses, enhancing disaster
preparedness and supporting mitigation efforts. However, studies of hurricane damage are challenging because damage
depends primarily on low-frequency, high-impact landfalling TCs, so uncertainties are high compared with basin-wide
metrics of TC activity (Emanuel, 2011). Uncertainty remains over how cyclone-related damage will evolve in a warming
climate (Jewson, 2023), so understanding the key factors responsible for, and predicting, hurricane damage are critical steps
50 in mitigating and preparing for future impacts.

Recent work has demonstrated skilful multi-year predictions of North Atlantic hurricane activity and U.S. hurricane damage,
but individual high-damage events, particularly those occurring during periods of generally low activity, are not well
predicted (Lockwood *et al.*, 2023). Each TC's damage is the result of a potentially unique combination of meteorological
and socioeconomic factors; in other words, a combination of hazard, exposure and vulnerability variables (Intergovernmental
55 Panel on Climate Change, 2012). Hazard variables describe the TC's physical characteristics, including intensity, duration,
size, and the magnitude and extent of (sub-)perils, such as coastal flooding caused by storm surge and inland flooding caused
by TC-related rainfall. Exposure variables describe the assets (e.g., buildings) impacted by a TC, and include the location
and values of residential, industrial and commercial buildings within the TC's impact footprint. Vulnerability variables
describe the damage-susceptibility of these assets, and are influenced by factors such as construction type (e.g., materials),
60 building design (e.g., height), and building age. Quantifying TC impacts at landfall requires a comprehensive understanding
of these hazard, exposure and vulnerability components of risk (Ward *et al.*, 2020). To this end, complex and high-
resolution catastrophe models (e.g., Florida Cat Model) are developed, but these are computationally expensive and unsuited
to estimating losses for a forecast TC, as forecasts evolve on sub-daily timescales.



65 North Atlantic TCs are categorised using the Saffir–Simpson Hurricane Wind Scale, which indicates potential damage based
on 1-minute, near-surface wind speed (NOAA, 2021). This scale is an effective public-communication tool. However, as it is
based on a single hazard quantity, it does not predict damage sufficiently skilfully. At landfall, central sea-level pressure
minima are found to be more strongly correlated with normalised historical hurricane damage than wind speed maxima
(Klotzbach *et al.*, 2020; Klotzbach *et al.*, 2022a), likely due to central pressure being physically related to both TC intensity
and size (Chavas *et al.*, 2025). This suggests that the widely used Saffir–Simpson scale misrepresents potential hurricane
70 damage, and there are calls for modifying the scale itself (Wehner and Kossin, 2024) and to adopt multidisciplinary (i.e.,
hazard, exposure and vulnerability) approaches to understanding hurricane impacts (Camelo and Mayo, 2021).

To provide a hurricane categorisation that better reflects damage, Bloemendaal *et al.* (2021) devised a ‘Tropical Cyclone
Severity scale’, which categorised TCs by combining wind speed, storm surge and accumulated rainfall. TCs ranked by this
scale corresponded better with historical financial damage compared with the Saffir–Simpson scale. However, several TC
75 events were still ascribed to a low category despite their high damage. For example, Hurricane Sandy (2012) caused an
estimated U.S.\$ 70 bn in normalised economic losses, but its classification was only changed from category 1 in the Saffir–
Simpson scale to category 2 in the ‘Tropical Cyclone Severity scale’. Beyond using hazard information alone, Pilkington and
Mahmoud (2016) used an artificial neural network model to forecast the economic impact from hurricanes using hazard and
exposure data, including landfall location, population affected, wind speed, central pressure, precipitation and storm surge.
80 Baldwin *et al.* (2023) showed the importance of differences in vulnerability between conurbations and rural areas for
accurately modelling TC risk across the Philippines. Their model included wind hazard, exposure data, and a vulnerability
layer to link a given wind speed to a percentage of exposed assets destroyed. These recent studies add to a growing body of
evidence that combining these factors is necessary to capture risk (Ward *et al.*, 2020).

An open question is whether a scale that includes TC hazard as well as landfall-centred exposure and vulnerability
85 information, which all determine risk, can predict historical TC damage. Moreover, there remains a need for a TC scale that
accurately captures damage potential that (i) can help understand historical events and (ii) may be used to estimate and
communicate how damaging a forecast TC may be (in the context of historical events). While the Saffir–Simpson scale may
be easily interpretable and an important component of public early-warnings, several sectors, particularly financial (e.g., (re-)
insurance and banking) and energy, require a skilful estimation of likely damage to inform decision-making, ideally without
90 significant computational cost.

Focussing on historical North Atlantic hurricanes, this study examines numerous hurricane hazard, exposure and
vulnerability quantities to determine whether the inclusion of socio-economic data into a novel TC classification scheme
improves our ability to represent—and forecast—their damage. Such a scheme may be used prior to a forecast TC landfall to



communicate potential damage accurately and in a usable way to key stakeholders, such as governments, disaster-
95 management agencies, and financial sectors. Here, ‘usable’ refers to the need to communicate the likely rank, and the
corresponding financial damage, of a forthcoming TC will have in the context of damage from historical events, allowing
responding agencies and stakeholders to make appropriate preparations and (re-)insurance to mobilise sufficient capital,
based on historical experiences. With this aim, we address the following research questions:

- Which single hazard, exposure and vulnerability variable has the highest correlation to historical hurricane damage?
- 100 • Is a combination of hazard, exposure and vulnerability data more skilful in predicting historical damage?
- Does a novel ‘Hurricane Predictive Damage Scale’ better represent damage than the Saffir–Simpson Scale?

This study is structured as follows: datasets and methods are described in section 2, results presented in sections 3–5, with a
novel TC scale evaluated in section 6, and discussion and conclusions presented in section 7.

2 Data and methodology

105 In this section, we first describe the datasets we use, providing loss, hazard, exposure and vulnerability information for
historical hurricanes affecting the U.S (section 2.1). Fig. 1 provides a summary of datasets used in this study and the
temporal coverage spanned by each. We then describe the methods (section 2.2) employed to quantify these quantities within
the footprint of each historical hurricane in our analysis and the statistical approach to investigate and predict historical
damage (section 2.2).



Vulnerability

Exposure

Hazard

Loss

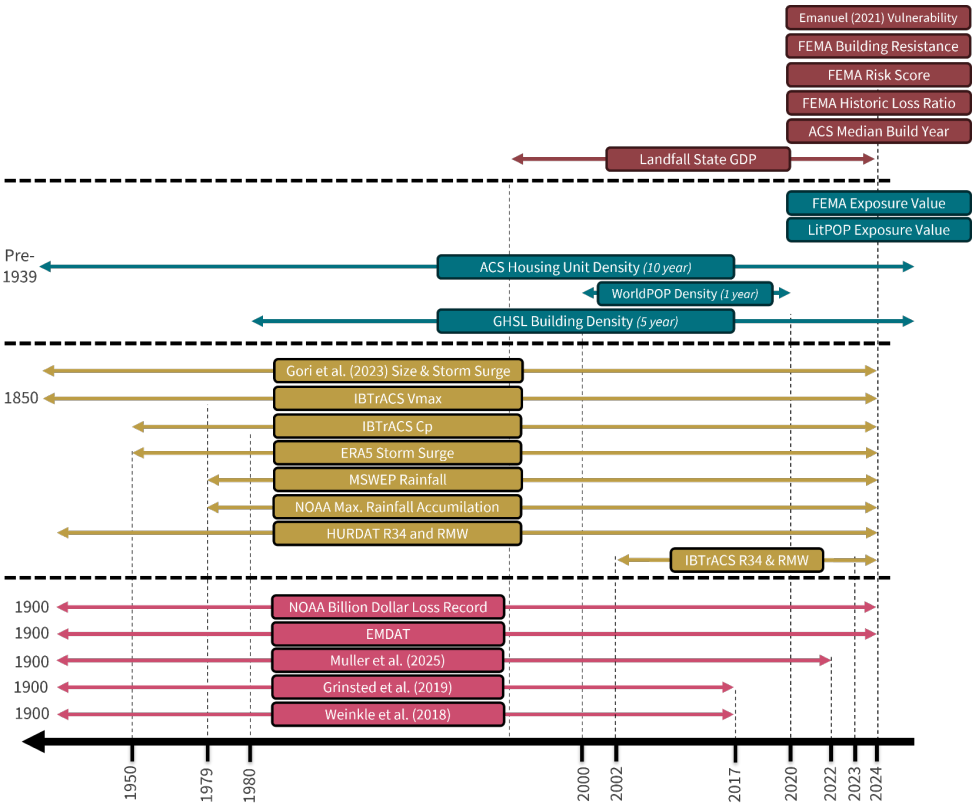


Figure 1. Overview of datasets used in this study. Hazard, exposure and vulnerability predictors are combined, with loss estimates being the predictand. Note that the x-axis timeline is not linear.

2.1 Data

2.1.1. Normalised historical hurricane losses

115 Historical hurricane economic loss estimates were collated from various government agencies and published studies: Blake
et al. (2011); Grinsted *et al.* (2019); Muller *et al.* (2025); Weinkle *et al.* (2018) (Fig. 1). Loss estimates from Blake *et al.*
(2011) are regularly updated by the National Oceanographic and Atmospheric Administration (NOAA), providing economic
losses for hurricanes which exceed U.S.\$ 1 bn (National Centers for Environmental Information, 2025). Loss estimates are
available from 1965, when Hurricane Betsy caused U.S.\$ 1.4 bn in economic loss, to the end of 2024, where Hurricane
120 Milton caused U.S.\$ 34.3 bn (National Centers for Environmental Information, 2025). Loss estimates from other studies are



limited to various time-periods (Table 1), and were collated from various sources including U.S. National Hurricane Center reports, Monthly Weather Review discussions, and National Centers for Environmental Information reports for individual storms. Damage estimates vary between sources and depend on whether (sub-)perils (e.g., coastal and inland flooding) are included alongside wind. Historical loss estimates are also provided by the Emergency Events Database (EM-DAT), in which estimates are collated from United Nations, governmental and non-governmental agencies, the (re-)insurance sector, research institutions, and the press (Delforge *et al.*, 2025).

We collated hurricane loss estimates from multiple sources to maximise sample size and to include differing loss estimates for storms included in multiple datasets (Table 1). Records from Weinkle *et al.* (2018) and Grinsted *et al.* (2019) end in 2017, and records from Muller *et al.* (2025) end in 2022. EM-DAT loss estimates from the 2024 hurricane season from Delforge *et al.* (2025) appear incomplete or are much lower than reported by NOAA (e.g., Helene and Milton (both 2024) are shown as causing less than U.S. 1bn in economic loss in EM-DAT). Weinkle *et al.* (2018) and EM-DAT provide estimates of the total economic damage across the whole hurricane track, whereas Grinsted *et al.* (2019) and Muller *et al.* (2025) provide a breakdown of economic damage per hurricane landfall, if a hurricane makes multiple landfalls across the U.S. along its track. Due to the lack of standardised damage recording in the U.S., economic loss estimates for every landfalling TC that occurred from 1979 to present (2024) are not available. Instead this study is limited to loss estimates included from these various sources (Table 1).

Hurricane economic loss estimates in this study represent average damage per hurricane landfall from various these sources (Delforge *et al.*, 2025; Grinsted *et al.*, 2019; Muller *et al.*, 2025; National Centers for Environmental Information, 2025; Weinkle *et al.*, 2018), hereafter termed ‘average loss’. Where damage estimates are available for multiple landfalls of a hurricane, the average economic damage estimate is taken from only Grinsted *et al.* (2019) and Muller *et al.* (2025), providing damage estimates per landfall. This approach increases the sample size for which financial loss estimates are available (Table 1). When loss estimates from these studies are merged and averaged, loss estimates are available for 134 landfalling (including multiple landfalls) and bypassing hurricanes between 1979 and 2024 (Table 1).



Table 1. Summary of hurricane loss datasets and the number of named storms (wind speed ≥ 39 mph) making landfall in (including multiple landfalls) or bypassing the U.S. from 1979 to 2024 for which an economic loss estimate is available.

Dataset reference	Dataset coverage	Landfalling TCs (1979–2024)
National Centers for Environmental Information (2025) ($\geq$ bn loss)	1965–present	58
Delforge <i>et al.</i> (2025)	1900–present	74
Muller <i>et al.</i> (2025)	1979–2022	33
Weinkle <i>et al.</i> (2018)	1900–2018	59
Grinsted <i>et al.</i> (2019)	1900–2018	102
<i>Average loss from above sources</i>	<i>1979–present</i>	<i>134</i>



Studies typically normalise historical economic loss by calibrating estimates to present-day levels of damage and apply
150 adjustment factors for changes in inflation, wealth and building density (Weinkle *et al.*, 2018). Simple normalisation applies
country-wide inflation and wealth adjustments and quantifies changes in building density as changes in residential housing
units. This does not account for regional variations in wealth and inflation changes, changes in building vulnerability over
time. Industrial or commercial building density changes may also not be accounted for, nor the impacts of climate change.
Nevertheless, normalisation is needed to help quantify present-day risk and we employ an updated approach based on
155 Weinkle *et al.* (2018) and Muller *et al.* (2025), by adjusting loss estimates using a country-wide inflation factor, a country-
wide real-wealth-per-housing-unit factor (both as of 2024), and county-wide (rather than country-wide) changes in housing
unit density factors. We perform this within a TC's radius (estimated using the radius of 34 kt wind speed, R34) (Eq 1).

$$D_{2024} = D_y \cdot \frac{I_{2024}}{I_y} \cdot \frac{W}{Hn_{1+\frac{(2024-y)}{y}}} \cdot Hn_{1+\frac{(2024-y)}{y}} \quad (\text{Eq. 1})$$

where D is damage per hurricane (and year), y , I is inflation, W is real national wealth per housing unit, and Hn is housing
160 unit density. I is determined using the annual implicit price deflator for gross domestic product provided by the U.S. Federal
Reserve Bank of St. Louis for the period 1979–2024 (U.S. Bureau of Economic Analysis, 2023). H density is determined
within R34 using U.S. housing unit data (U.S. Census Bureau, 2024). W is quantified using an estimate of current-cost net
stock of fixed assets and consumer durable goods (U.S. Bureau of Economic Analysis, 2025). Changes in housing unit
density per hurricane require storm size estimates, so using regional changes in housing unit density constrains the sample
165 size because TC size information in best-track data is not available for all TCs (further discussion below).

2.1.2. Hazard data

Hurricane track location, intensity and size information was obtained from the International Best Track Archive for Climate
Stewardship (IBTrACS) v04r01 (Gahtan *et al.*, 2024), provided by the National Hurricane Center (NHC), part of the
National Oceanographic and Atmospheric Administration (NOAA). For each historical hurricane, we obtained maximum
170 wind speed, $v_{\max}$, minimum central sea-level pressure, c_p , translation speed, radius of maximum wind (from the storm
centre), RMW, and the outermost (i.e., beyond RMW) radii of 34-, 50- and 64-knot wind speeds from the storm centre,
respectively, R34, R50 and R64. Each quantity was determined at the timestep before the storm centre crosses over land, so
atmospheric fields are not impacted by land-surface interactions. However, IBTrACS data are incomplete (i.e., not all hazard



variables are available at every timestep for every TC). Therefore, we supplemented IBTrACS with data from NOAA's
 175 HURDAT2 reanalysis (Landsea and Franklin, 2013)—specifically, the “U.S. Hurricane Impacts / Landfalls” table
 (Hurricane Research Division, 2025) of landfall information collected by NOAA reconnaissance aircraft. Where data are
 missing in IBTrACS, HURDAT2 data are substituted, if available. Where data are available in IBTrACS and HURDAT2 at
 a given timestep, HURDAT2 data were prioritised.

Storm surge and rainfall cause damage through coastal and inland flooding but are not included in IBTrACS and
 180 HURDAT2. In this study, historical hurricane storm surge (maximum storm surge height) data are taken from the storm
 surge residual product (Copernicus Climate Change Service, 2022), derived using the Global Tide Surge Model (version 3.0)
 forced by European Centre for Medium-Range Weather Forecasts' fifth-generation reanalysis (ERA5; Hersbach *et al.*,
 2020). This provides hourly reconstructed historical storm surge height from 1950–present. Storm surge residual is
 calculated as the difference between the total water level and simulated storm-tide elevation, including the influence of storm
 185 surge and tide. Storm surge may be larger in the hours before or after a hurricane makes landfall, depending on antecedent
 tidal height. To account for this, we defined the maximum storm surge residual as the maximum along the U.S. coastline
 within a 1,000 km radius of the TC's central coordinate and 24 hours before and after a hurricane makes landfall. An
 example maximum storm surge residual for Hurricane Katrina (2005) is shown in Fig. S1.

Historical rainfall footprints per hurricane were derived from the Multi-Source Weighted-Ensemble Precipitation (MSWEP)
 190 dataset (Beck *et al.*, 2019), which assimilates gauge observations and satellite data to reconstruct 3-hourly rainfall 1979–
 present. For each historical hurricane, we determined total accumulated rainfall along the hurricane track, maximum 3-
 hourly rain rate along the track, and the maximum total rain accumulation per grid-point along the track, within a 500 km
 radius of the hurricane centre for each timestep. (Location taken from IBTrACS.) An example rainfall accumulation
 footprint for Hurricane Katrina is shown in Fig. S2. To complement MSWEP, the maximum rainfall accumulation along the
 195 TCs track is also provided by NOAA (National Oceanographic and Atmospheric Administration, 2025) was also collated per
 historical storm. Only a single maximum rainfall value is provided by NOAA for most storms, preventing differentiation
 between multiple landfalls.

2.1.3. Exposure data

We took county-level building value information across the U.S. from the National Risk Index (Federal Emergency
 200 Management Agency, 2024), derived from Hazus 6.0 (Federal Emergency Management Agency, 2022), providing 2022-
 relative valuations per county based on the 2020 U.S. Census. Building values are time-invariant. Near present-day (2019)



building value is also quantified using the LitPOP dataset (Eberenz *et al.*, 2020), providing global aggregated building value estimates at a 1-km spatial resolution. Building value estimates vary between studies, hence we used two building value datasets.

205 We also used decadal, county-level housing unit density data (U.S. Census Bureau, 2024), available 1950–present as well as time-varying (every five years) building density estimates from the Global Human Settlement Layer (GHSL), where built-up surface area data are derived from Sentinel-2 satellite composite and Landsat satellite imagery, providing from 1975–present at 1-km spatial resolution (Pesaresi and Politis, 2023). Annual gridded population density between 2000 and 2020 at 1-km spatial resolution data were obtained from WorldPop (WorldPop, 2018). For hurricane landfalls outside the temporal
 210 coverage of GHSL and WorldPop data, we used data for the closest available year.

2.1.4. Vulnerability data

Vulnerability—the susceptibility of an asset to damage—is influenced by building age, often used as a proxy for building condition and resilience, and related to changes in building codes (regulated construction standards that include a minimal resistance to extreme weather). We use county-level average building age data (U.S. Census Bureau, 2024) and a county-
 215 level indicator of building resistance to extreme weather (Federal Emergency Management Agency, 2025). County-level Hurricane Risk Score (defined as average hurricane social vulnerability and community resilience as a percentage) and Hurricane Historical Loss Ratio (defined as the percentage of a location’s exposed value damaged by past hurricanes) data were also used (Federal Emergency Management Agency, 2024; Zuzak *et al.*, 2021), and we quantified the mean Hurricane Risk Score and Hurricane Historical Loss Ratio across all counties within the each hurricane’s landfall footprint. GDP data
 220 per affected U.S. state per year, which may indicate a state’s resilience resources, for the period 1998–2024 (U.S. Bureau of Economic Analysis, 2025) were used, and 1998 GDP values applied to all landfalls prior to 1998.

Hurricane wind vulnerability may also be expressed as a logistic–cubic wind–damage function, relating wind speed to the fractional value of assets lost. This study uses a simple quantification of this (Eq. 2 and Eq. 3), which was deduced by (Emanuel, 2011).

$$225 \quad f = \frac{v_n^3}{1+v_n^3}, \quad (\text{Eq. 2})$$

where f is the fraction of the property value lost and v_n is defined as:



$$v_n = \frac{[(v - v_{\text{thresh}}), 0]}{v_{\text{half}} - v_{\text{thresh}}}, \quad (\text{Eq. 3})$$

where v is TC's wind speed, v_{max} and v_{thresh} and v_{half} are the thresholds at which no damage and half asset damage occur, respectively.

230 Building characteristics (height, construction type, age etc) influence v_{half} , estimated to be between 96 and 139 kts (Vickery
et al., 2006). In this study, damage estimates were computed for historical hurricanes using this single vulnerability function
for all buildings, with $v_{\text{thresh}} = 40$ kts and $v_{\text{half}} = 140$ kts. Multiple vulnerability functions for various buildings with different
properties e.g., age, construction type, cannot be used due to the lack of data on individual building characteristics. This asset
damage potential is applied to the total LitPOP exposure value and GHSL building density within each hurricane's footprint
235 to determine the percentage of LitPOP exposure value and GHSL building density damaged. At each timestep, v_{max} is used
with Eq. 2 and Eq. 3 (Emanuel, 2011) and the extracted exposure value and building density.

2.2. Methods

2.2.1. Statistical estimation of hurricane size

240 Hurricane size (i.e., R34, R50 and R64) data are only available in IBTrACS from 2002 onwards. This significantly
constrains the sample size of TCs available for studies of impact, as size is needed to determine the region incurring damage,
and, practically, the exposure and vulnerability impact footprints derived in this study require TC size estimate. Therefore,
we developed a simple hurricane size estimation method to reconstruct hurricane size prior to 2002, extending the landfalling
TC record for this study. Such estimates may be less accurate than direct observations, but this statistical approach is proven
245 to be a skilful supplement to missing IBTrACS data (Fig. 2 and Fig. S3) and allows for storm size estimation prior to 2002,
and more than doubling our sample size from TCs between 2002 and 2024 to TCs between 1979-2024.

We developed a skilful random forest statistical model to estimate R34, R50 and R64 for each hurricane and at each timestep
based on v_{max} , RMW, c_p and latitude (Fig. 2), with RMW being the most influential predictor. When R34 estimates from this
statistical model are evaluated against R34 observations for the period 2002 onwards, using a leave-one-out approach (i.e.,
250 training the prediction model on observations except one, and evaluating skill on the left-out observation) across 4,220



observations, a Spearman's correlation coefficient of 0.88 and a mean absolute error (MAE) of 21.8 nm was found (Fig. 2). Prediction models were developed to estimate R50 and R64 for historical storms where observations are available, with evaluation shown in Fig. S3. From 1979 to 2002, however, RMW information within IBTrACS is also incomplete, so RMW estimates from HURDAT2 and reconstructions from Gori *et al.* (2023) were used. Gori *et al.* (2023) estimated RMW by combining the TC wind model of Chavas *et al.* (2015) and ERA5 data. Where RMW is missing from IBTrACS, RMW is replaced by values from HURDAT2 or Gori *et al.* (2023), with HURDAT2 values preferred. For missing data over successive timesteps, RMW values from the previous timestep were used, assuming that large changes in horizontal size within 3 hours are rare. Of all 3-hourly timesteps where a TC is over land (approximately 10,000 timesteps), approximately 2,500 timesteps have a missing RMW estimate from each of these three datasets, precluding a random-forest model estimate of R34, R50 or R64 and the TC was therefore omitted from our analysis. Where size was not observed or estimated, TC footprints could not be accurately determined (see section 2.2.2.), and hence RMW observations from previous timesteps were used. Although this introduces additional uncertainty, R34, R50 and R64 estimates from the random forest model using RMW estimates from previous timesteps are also a function of c_p , v_{max} and latitude, quantities that are much less frequently missing in IBTrACS.

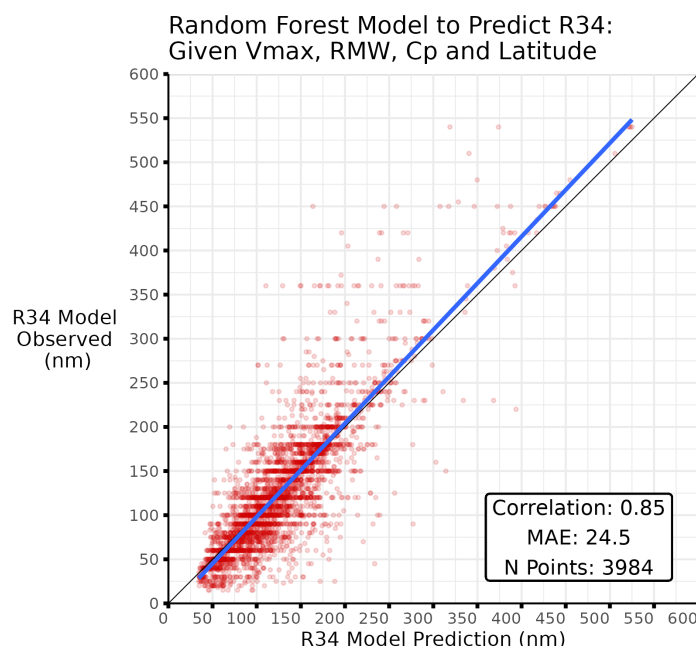


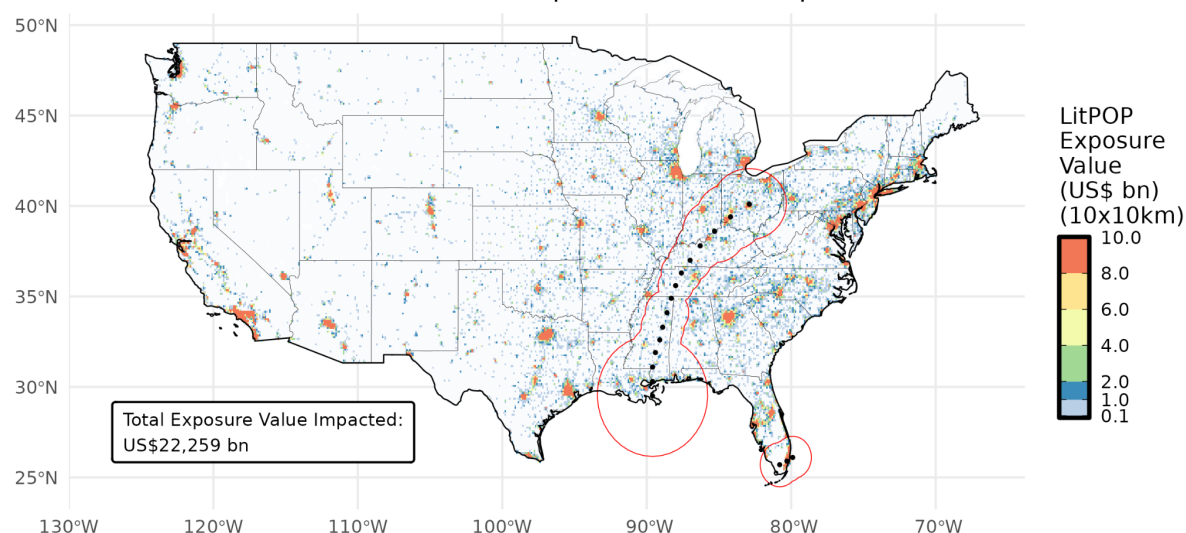


Figure 2. Comparison of estimated and observed hurricane R34 covering the period 1979–2023. Our random forest statistical model uses the hurricanes v_{max} , RMW , c_p and latitude as TC size predictors.

2.2.2. Extracting exposure and vulnerability impact footprints

270 For each hurricane, exposure and vulnerability data were extracted within the R34, R50 and R64 radii at each timestep, and accumulated to create two impact footprints: one along the full hurricane track (Fig. 3) and one from the immediate landfall coordinates (i.e., up to 12 hours after landfall). TC's generally weaken over land, and thus obtaining two impact footprints allows for an indication of impacts from the immediate landfall, where intensity and therefore impacts are likely to be greatest, and the full TC track.

Hurricane Katrina (2005) LitPOP Exposure Value Footprint



275

Figure 3. Example LitPOP exposure value impact footprint of Hurricane Katrina (2005). Hurricane locations from IBTrACS every three hours are shown as black dots, with the R34 radius around the hurricane centre indicated by the red line. Note multiple landfalls in Florida and the Gulf of Mexico coast are shown.



280 2.2.3. Predictive statistical approaches

In this study, we used a weighted combined-rank framework, linear regression framework and the random forest decision-tree framework to combine input predictors across hazard, exposure and vulnerability to predict historical hurricane damage. These methods were selected for their interpretability. Our target prediction variable is average damage for each TC, derived from multiple datasets (Grinsted *et al.*, 2019; Muller *et al.*, 2025; National Centers for Environmental Information, 2025; 285 Weinkle *et al.*, 2018), maximising the sample of hurricanes for which a loss estimate is available (Table 2). Overall, 106 hurricanes, for which all risk variables can be quantified, are available to train our predictive model.

To evaluate the predictive skill of the linear regression and random forest predictive models, a leave-one-out cross-validation was used, where each input case is treated once as the test case and the model trained on the remaining cases. The leave-one-out approach is better suited to evaluating the skill of single predictions than, for example, k-fold cross-validation, where 290 input data are split into training subsets. For the weighted combined-rank framework, which determines the combined damage rank from various hazard, exposure and vulnerability ranks, such validation is not appropriate, as the combined rank approach does not require training data. For this approach, the optimal combinations of weights between -10 and 10 are determined for each combination of input variables, to minimize a cost function, which in this case is the RMSE between predicted hurricane rank and observed hurricane damage rank. These approaches allow our model to be useful for predicting 295 damage due an impending hurricane, if input (hazard) predictors were obtained from forecast data, and exposure and vulnerability predictors were obtained from the forecast track and size data. To reiterate: a key aim of this study, to develop an approach to estimate expected damage for future forecast landfalling hurricanes.

300



Table 2. Number of named storms (wind speed >39 mph) from 1979 to 2024 making landfall for which all hazard, exposure and vulnerability characteristics can be derived from various studies including damage estimates.

Loss dataset	# TCs with all loss variables
Weinkle <i>et al.</i> (2018)	57
Grinsted <i>et al.</i> (2019)	76
Muller <i>et al.</i> (2025)	33
EM-DAT (2025)	21
National Centers for Environmental Information (2025)	57
Average loss	106 (out of 134 — see Table 1)

305 We quantified model performance with Spearman correlation coefficient, ρ , (testing the ordering of prediction according to correlation between ranked values—i.e., whether one storm is more damaging than another), Pearson correlation coefficients, r , (testing non-ranked correlation), and model root-mean-square error (RMSE). Spearman’s ρ and RMSE are



both used, as ρ is a good indicator of the capability of accurately predicting damage rank, but not suitable for identifying cases where large differences occur in the damage prediction of a single hurricane.

310 Two approaches were considered for representing input variables: raw values and ranked values. As input variables span disparate ranges (for example, c_p spans 900–1000 hPa and LitPOP exposure spans U.S.\$ 10 million to 1 trillion), raw values were normalised to the same scale as ranked values 1– n (i.e., 1–106). In this normalisation, each variable is linearly rescaled so that the maximum value is assigned a score of 1 and its minimum a score of 106, with intermediate values mapped proportionally between these bounds. Linear and normalised input variable ranks were derived, with rank 1 corresponding to
 315 the costliest event and rank 106 to the least costly. To predict each hurricane's loss, the alike loss ranks and predicted loss ranks were identified and the loss associated with that observed loss rank is assigned to the hurricane with the alike predicted loss rank.

3 Historical relationship between hurricane $v_{\max}$ and damage

320 Economic damage across the U.S. from landfalling hurricanes has generally increased over time and exhibits large interannual variability (Fig. 4). The most destructive year for hurricane-related damage was 2005, with U.S.\$ 153 bn in un-normalised (and approximately U.S.\$ 350 in normalised) economic damage (Fig. 4). The most damaging hurricane was Hurricane Katrina (Fig. 5), although uncertainty is evident across available loss estimates (Grinsted *et al.*, 2019; Muller *et al.*, 2025; National Centers for Environmental Information, 2025; Weinkle *et al.*, 2018). If Katrina occurred today, economic
 325 damage may be in the range of U.S.\$ 190–290 bn. Hurricane Harvey (2017) is the second-most damaging hurricane, but uncertainty is more considerable, with estimates ranging from U.S.\$ 90 to U.S.\$ 190 bn. The Great Miami Hurricane (1926), although not considered in this study, caused an estimated U.S.\$ 105 bn in un-normalised economic damage (Weinkle *et al.*, 2018), and would likely surpass Katrina were this event to occur today.

330

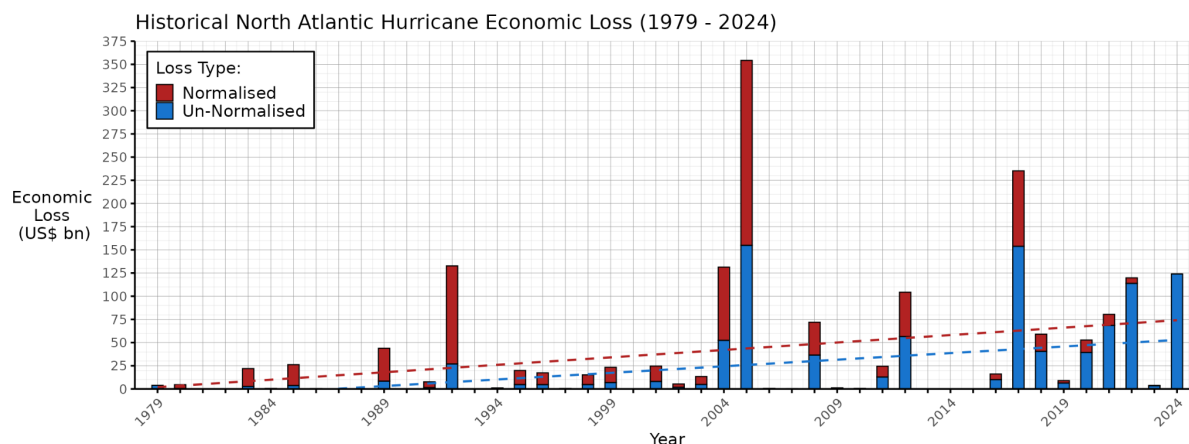


Figure 4. Average historical North Atlantic hurricane economic loss normalised to 2024 (red) and un-normalised by inflation, wealth and housing unit density (blue).

Hurricane Katrina, however, is one of many high-impact TCs whose Saffir–Simpson category is at odds with the magnitude of damage caused. Katrina was classified as category 3 at landfall using the Saffir–Simpson, despite causing unprecedented and record-breaking damage (Fig. 5). Other damaging cases which are mismatched with their Saffir–Simpson category at landfall, which include category 1 Hurricane Sandy (2012) and category 2 Hurricane Ike (2008), which each caused U.S.\$ 99.5 and 57 bn in economic damage respectively. The historical record reveals the limitation of an event’s Saffir–Simpson category in conveying its economic damage (Fig. 5), which has also been found in other studies (e.g., Bloemendaal et al., 2021).

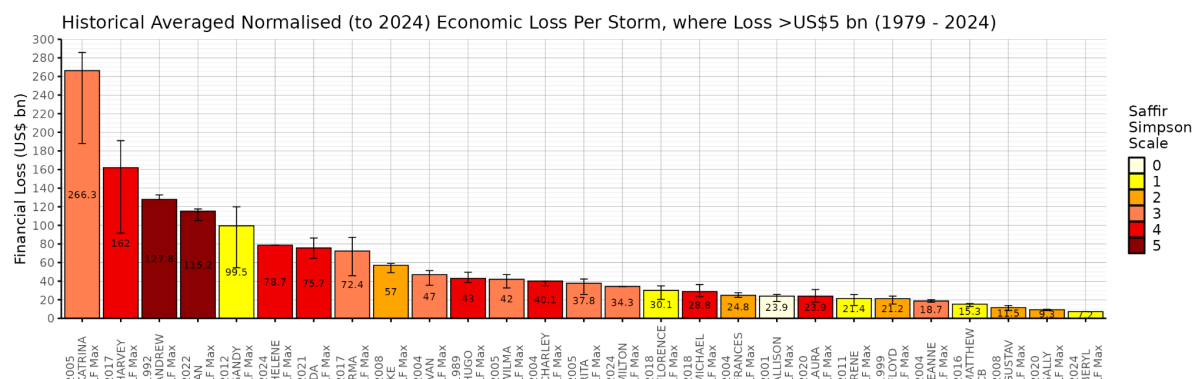


Figure 5. Average economic financial loss per hurricane, normalised to 2024, for loss events exceeding U.S.\$ 5 bn. Error bars indicate the range in loss estimates across datasets (Grinsted *et al.*, 2019; Muller *et al.*, 2025; National Centers for Environmental Information, 2025; Weinkle *et al.*, 2018) and the colour indicates landfall Saffir–Simpson category.

Although economic damage generally increases with the Saffir–Simpson landfall category (and therefore wind speed), the Saffir–Simpson scale has limited skill in predicting economic damage (Fig. 6a), with a correlation across 129 past hurricanes of $\rho = 0.54$ (RMSE of U.S.\$ 17.1 bn) when using linear fit. When using $v_{\max}$ to predict normalised hurricane economic damage, the correlation ($\rho = 0.51$) and (RMSE = U.S.\$ 17.6 bn) are slightly lower (Fig. 6b). This result is consistent with Klotzbach *et al.* (2020). Notably, landfall $v_{\max}$ is significantly less skilful at predicting economic damage for more extreme storms, with generally higher error at more extreme $v_{\max}$ values (Fig. 6b), which is particularly problematic as an inaccurate forecast for these more intense storms would produce larger errors. The relationship between $v_{\max}$ and damage is potentially nonlinear but spread in damage generally increases with $v_{\max}$. Performing a rank correlation between observed economic damage and damage predicted from $v_{\max}$ reveals significant spread (and heteroscedasticity) across the observed damage range (Fig. 7a), and this is found for all events as well as those for which damage exceeds U.S.\$ 1 bn (Fig. 7b).

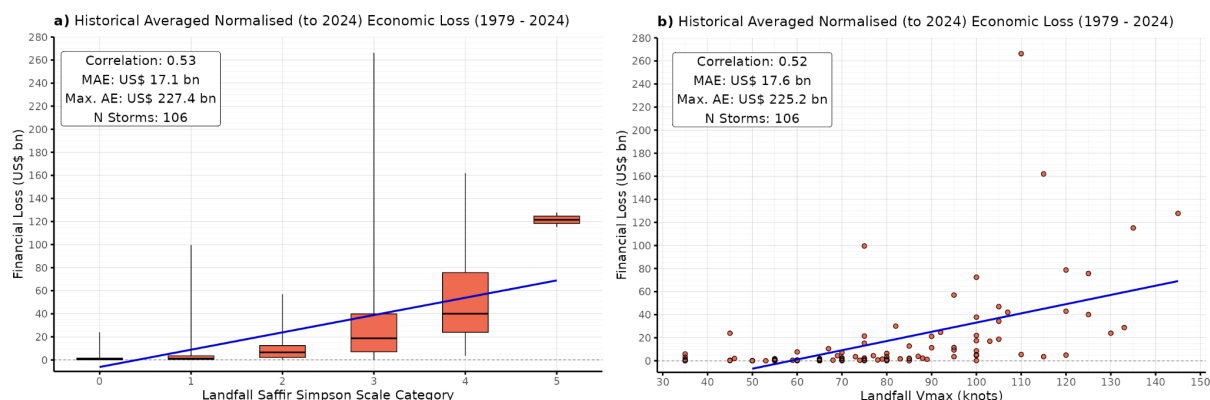


Figure 6. Average economic hurricane financial loss (normalised to 2024) versus a) Saffir–Simpson category and b) landfall v_{max} of U.S. landfalls over the period 1979–2023. Blue lines indicate linear fits and an indication of goodness of fit is given in each legend: mean absolute error (MAE), absolute error (AE), and sample size, N .

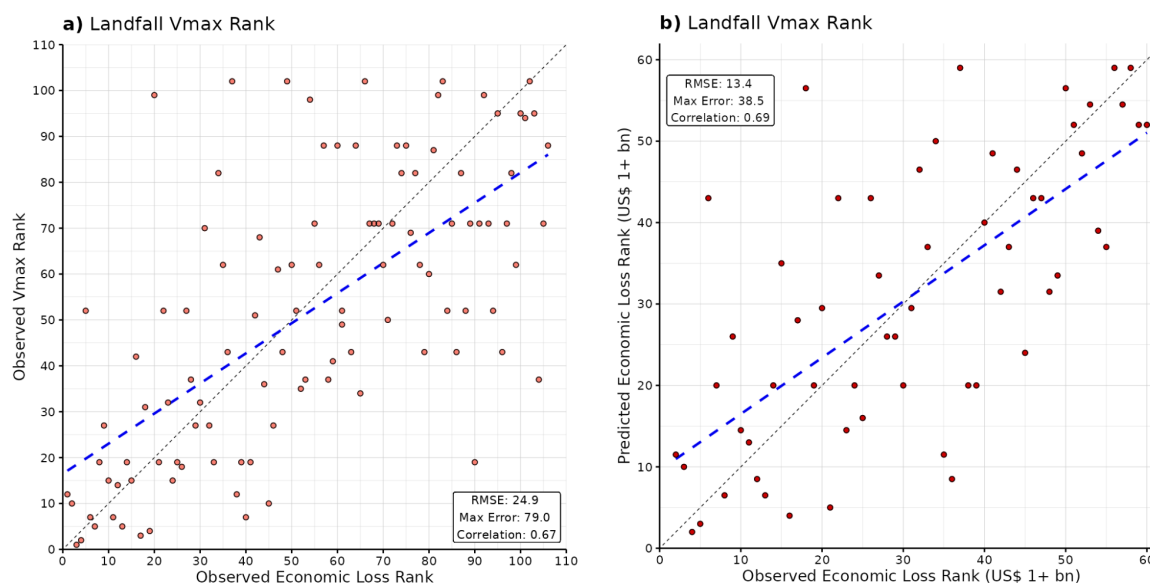


Figure 7. Leave-one-out evaluation using a generalised linear prediction model to predict historical hurricane loss rank using landfall v_{max} for a) all hurricanes for which all landfall variables could be derived ($n = 106$) and b) hurricanes that caused $\geq \$1$ bn normalised economic damage ($n = 60$). Rank 1 denotes the most intense hurricane at landfall and the highest normalised economic damage.



4 Relationships between historical hurricane damage and risk-related variables

We next analysed linear correlations between single risk-related predictors and historical hurricane damage over the period 1979–2024, using both raw values (Fig. 8a) and ranks (Fig. 8b). This analysis demonstrates that high prediction skill may be obtained across numerous hazard, exposure and vulnerability TC quantities, with landfall c_p rank yielding the highest correlation ($r = -0.78$), followed by GHSL building density damage percentage ($r = 0.77$) and LitPOP exposure value damage percentage ($r = 0.74$) (Fig. 8a). Overall, using the value rank per storm yields higher correlations than using raw values (Fig. 8c) because using ranks normalises to a linear scale, suggesting that landfall attribute ranks may provide more skilful damage predictions. This complements the analysis of Klotzbach *et al.* (2022a), who showed that landfall c_p better correlates with damage rank than v_{max} or accumulated cyclone energy (ACE). Here, we further show that landfall c_p outperforms other hazard variables and that landfall exposure and vulnerability variables yield higher correlations with damage rank than landfall v_{max} .

In this study, hazard variables at landfall (v_{max} , c_p , storm surge, R34, translation speed, and RMW) remain unchanged, but variables quantified within a hurricane footprint depend on chosen hurricane size metric (i.e., R34, R50 or R64) and those quantified for along hurricane track depend on timeframe (i.e., full track or 12 hours post landfall). For landfall variables with significant correlation to damage (i.e., ≥ 0.3), a 12-hour post-landfall track yields similar (or somewhat higher) correlation compared to using the full track where winds exceed 34 kts (Fig. S4a). This indicates that including the full track, where winds weaken over land, reduces correlation. But focussed studies of inland impacts may require full-track analysis and therefore may yield lower skill than landfall studies.

Additionally, correlations are generally higher when considering normalised rather than unnormalised economic damage (Fig. S4b). The correlation between damage rank and variable rank is lower when using R64 to define impact radius compared with R34 and R50 (Fig. S4c). R64 is typically less than 100 nm, which may be too small to capture hurricane impacts, especially for lower-resolution data (e.g., county-level housing unit data). Correlations using R34 and R50 are similar. Notably, the percentage of damaged GHSL building density across the full track has a strong correlation with damage rank ($\rho = 0.78$), equalling that of landfall c_p (Fig. 8). The damage density correlation using R50 ($\rho = 0.79$) slightly outperforms that of landfall c_p .

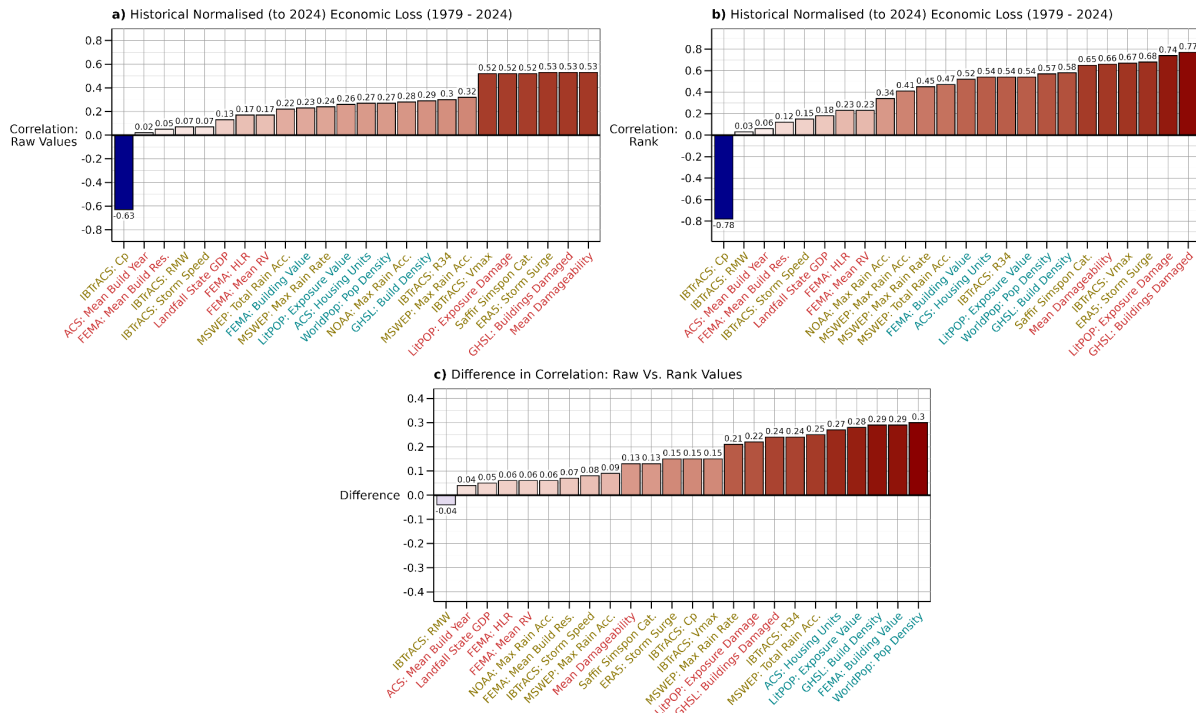


Figure 8. Pearson's correlation coefficients between historical hurricane landfall variables quantified at landfall and within 12 hours of landfall and their NOAA financial loss for the period 1979–2024. Shown are a) raw values versus loss, b) variable rank values versus loss, and c) the coefficient difference between raw versus ranked values (i.e., ranked minus raw values). Note that the x-axes differ between the panels, based on the correlation order. Colours indicate if the variable is a hazard (yellow), exposure (blue) or vulnerability (red) variable (refer to Fig. 1).

5 Statistical prediction of historical hurricane damage

Using landfall $v_{\max}$ rank to predict storm damage, by assigning the loss from the alike loss rank, yields $\rho = 0.67$ and RMSE = U.S.\$ 36.2 bn (Fig. 9), so $v_{\max}$ has limited skill in historical damage. Among the most underestimated damage events are hurricanes Katrina (2005) and Harvey (2017), which were category-3 and category-4, respectively, at landfall (Fig. 4), but are the two most damaging hurricanes since 1979 (Fig. 10a). Furthermore, using $v_{\max}$ alone leads to damage predictions exceeding U.S.\$ 30 bn for several less-damaging ($\leq$ U.S.\$ 10 bn) events (Fig. 10a), including Michael (2018), Laura (2020),



Dennis (2005), Andrew (1992; second landfall) and Idalia (2023). In our analysis (Fig. 8), and in recent work (Klotzbach *et al.*, 2020; Klotzbach *et al.*, 2022a), landfall c_p is found to be the most skilful single hazard quantity to predict historical damage, and using landfall c_p rank to predict damage rank improves this correlation (Spearman correlation coefficient 0.78 and RMSE U.S.\$ 26.3 bn) (Fig. 9). However, several events, across a range of actual damages, remain poorly predicted: Hurricane Andrew (1992), hurricanes Michael (2018), Rita (2005), Hugo (1989), Dennis (2005), Allen (1980) and Idalia (2023) are all overestimated (Fig. 10b). This similarity between $v_{\max}$ and c_p in predictive power is due to the physical pressure–wind balance intrinsic to hurricanes (Chavas *et al.*, 2017), indicating that intensity metrics alone cannot accurately model damage.

We now combine landfall hazard, exposure and vulnerability quantities to derive more skilful models to predict historical hurricane damage, using three statistical approaches: multiple linear regression, random forest, and weighted combined rank (Fig. 9). Using a random forest model based on the optimal combination of hazard, exposure and vulnerability variables improves the damage prediction model skill ($\rho = 0.87$; RMSE = U.S.\$ 18.7 bn), with a further improvement being obtained from using a linear regression model ($\rho = 0.87$; RMSE = U.S.\$ 17.1 bn) (Fig. 9). Even further improvement in skill is found when using a weighted combined-sum approach ($\rho = 0.89$; RMSE = U.S.\$ 7.0 bn) (Fig. 9). Overall, effectively combining hazard, exposure and vulnerability predictors is found to reduce the RMSE of TC damage predictions by as much as ~67% compared with using landfall $v_{\max}$ rank alone, and by ~55% when using landfall c_p rank alone (Fig. 9).

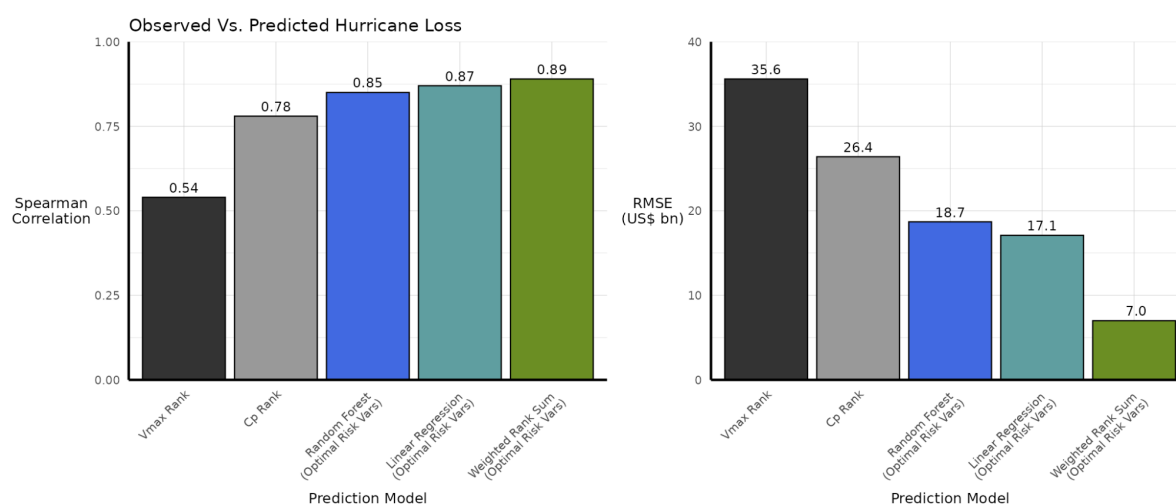


Figure 9. The a) Spearman correlation coefficients and b) root mean squared error (RMSE) of each tested historical hurricane damage prediction model.



We found that the optimal model to predict historical hurricane damage uses the combined normalised rank of landfall c_p , maximum rainfall accumulation rank, and the percentage of total GHSL building density damage within R64 of the hurricane centre and within 12 hours of the hurricane's landfall (Fig. 10c). This model includes two hazard quantities
 430 representing hurricane wind and inland flooding intensity (c_p and rainfall), with hurricane radius implicitly included in the percentage of total GHSL building density damage (with storms with larger R64 having higher impacted building density values). This also includes exposure values of impacted building density and applies a vulnerability function, by determining the percentage of damaged buildings proportional to $v_{\max}$ (at each timestep). Overall, this model demonstrates the extent to which damage prediction skill can be improved when accounting for hazard, exposure and vulnerability aspects of hurricane
 435 risk.

Across the 106 historical hurricanes in our sample, this prediction model has the lowest RMSE (U.S.\$ 7.0 bn), when these hazard, exposure and vulnerability hurricane landfall attributes are ranked and summed, while also applying the optimal weights (Fig. 9c). Moreover, considering highly damaging historical hurricanes ($\geq$ U.S.\$ 20 bn), $\rho = 0.9$, which is greatly improved compared with using only landfall $v_{\max}$ ($\rho = 0.45$) or landfall c_p ($\rho = 0.66$) (Fig. 9). This model results in the five
 440 most-damaging hurricanes (Katrina, Harvey, Andrew, Ian, and Sandy) being ordered according to observed losses, and thus their damage is very well captured.

Some discrepancies remain across all historical storms between predicted and observed damage ranks, but for higher-loss events ($\geq$ U.S.\$ 20–80 bn), damage rank is very skilfully predicted. The largest errors include hurricanes Michael (2018), Opal (1995) and Allen (1980), all of which have a high landfall c_p rank (i.e., high intensity) but did not cause (relatively)
 445 high damage. Overall, this model is markedly more skilful than using the single variables of landfall $v_{\max}$ and c_p to predict a storm's damage (Fig. 9 and Fig. 10).

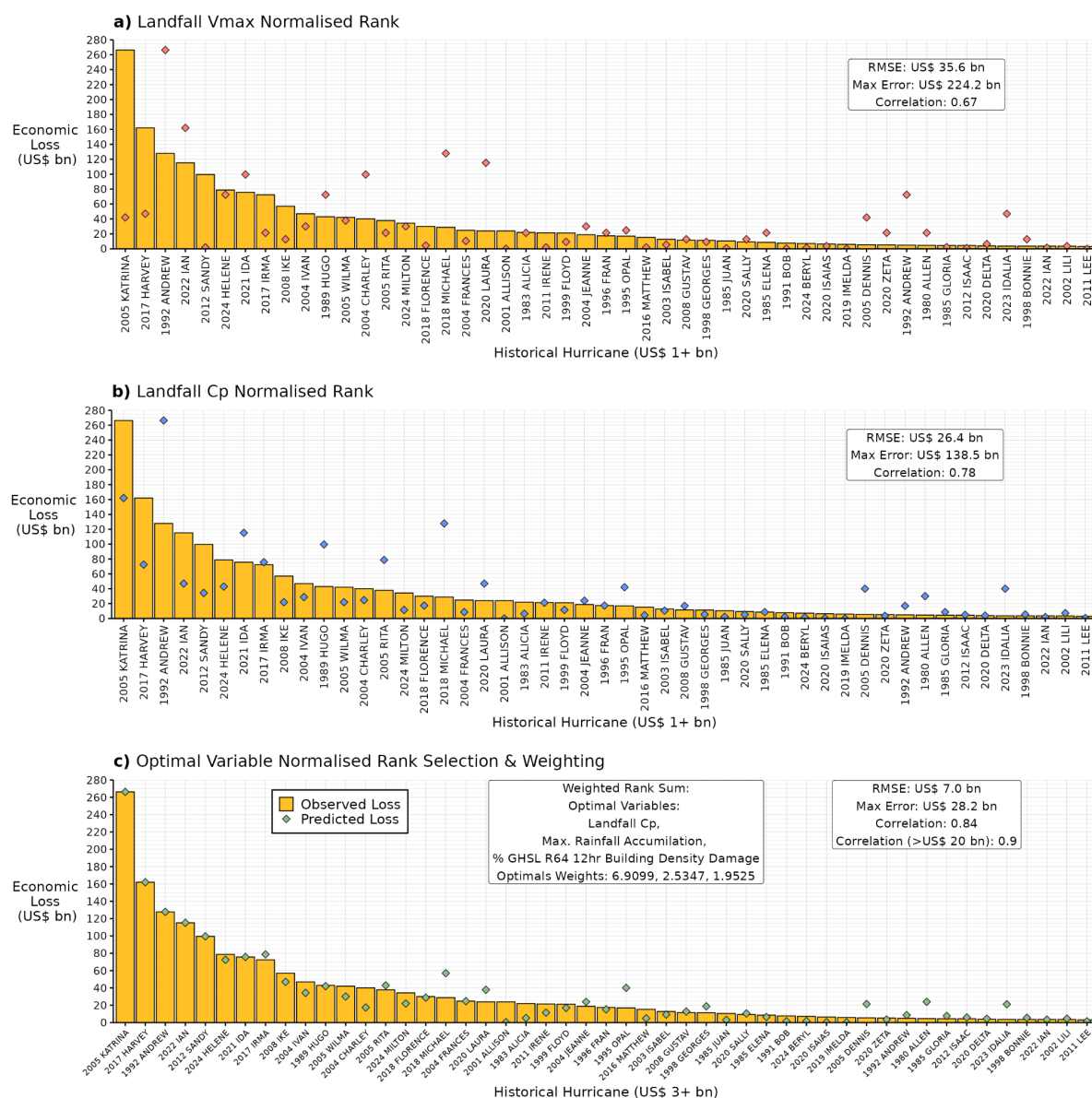


Figure 10. Comparisons of landfall variable rank with loss rank to predict past hurricane damage given past economic damage observations, using a) landfall v_{max} rank, b) landfall c_p rank, and c) optimally selected and weighted (refer to Fig. 8) hazard, exposure and vulnerability linear rank variables at landfall. The yellow bars indicate observed historical hurricane economic damage, and the coloured diamonds indicate predicted historical hurricane economic damage. Note that the optimal risk variables and their optimal weight are given in panel c).



455 6 A damage-based hurricane scale

The Saffir–Simpson scale is an effective communication tool and a key component of early-warning dissemination (Camelo and Mayo, 2021; Oliver-Smith, 2020; Wehner and Kossin, 2024), but, being based on $v_{\max}$ alone, does not correspond adequately with historical damage ranking (Fig. 4, Fig. 5). Based on our analyses, we devised a damage-based hurricane categorisation, termed the ‘Hurricane Predictive Damage Scale’ (HPDS). We intend this novel scale to complement the

460 Saffir–Simpson scale, as well as published work (e.g., Bloemendaal *et al.*, 2021; Pilkington and Mahmoud, 2016), and provide a scale based on financial loss that may be used by stakeholders in financial industries (e.g., (re-)insurance, banking and energy) to predict damage. A key application of this scale is in the (re-)insurance sector, for which there is significant value in knowing where a forecast TC may rank compared with historical events, and an indirect public benefit that results from greater recovery benefit from a more efficient (re-)insurance sector.

465 Table 3 shows the historical frequency proportions of landfalling hurricanes in each Saffir–Simpson category. Since 1900, 39% of landfalling hurricanes have been category 1, with only 12% and 3% being category 4 and 5, respectively (Table 3). We applied the same observed relative proportions in each Saffir–Simpson category to the averaged historical loss data since 1979 (Delforge *et al.*, 2025; Grinsted *et al.*, 2019; Muller *et al.*, 2025; National Oceanographic and Atmospheric Administration, 2024; Weinkle *et al.*, 2018) to derive historical loss categories (Table 3), which form our TCDPS. A cyclone

470 that caused loss less than U.S.\$ 1.3 bn would be a ‘loss category 1’ event, and a cyclone causing loss greater than U.S.\$ 119.5 bn a ‘loss category 5’ event.

475



Table 3. Percentage of landfalling historical storms within each Saffir–Simpson category, since 1900 from Hurricane
 480 Research Division (2025). Loss category thresholds define our proposed TCDPS to complement the Saffir–Simpson scale.

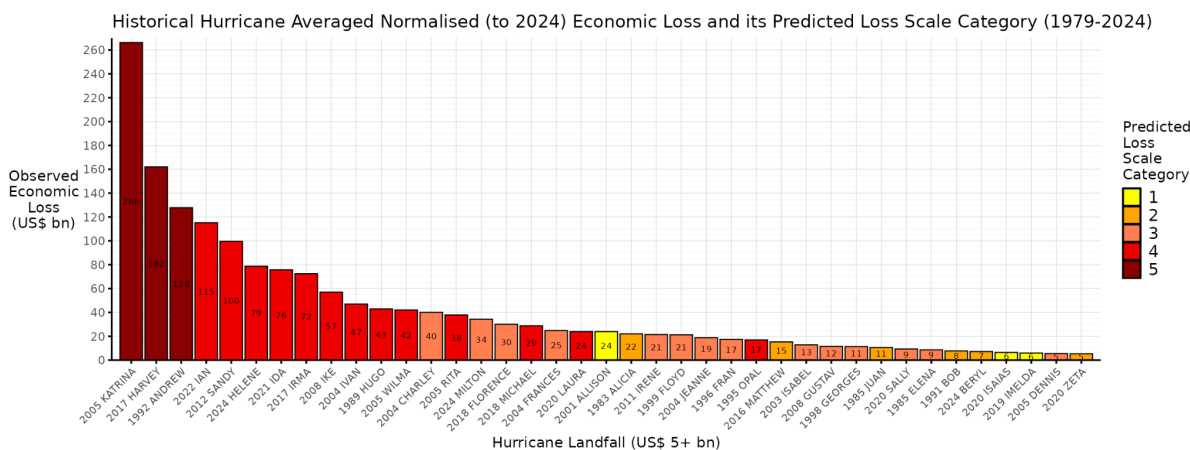
Saffir–Simpson category	Frequency	% of all TCs	$v_{\max}$ threshold (kt)	TCDPS loss category threshold (U.S.\$ bn)
1	84	38.9	64-83	0-1.3
2	54	25.0	83-96	1.3-5.3
3	46	21.3	96-113	5.3-29.5
4	26	12.0	113-136	29.5-119.5
5	6	2.8	136+	119.5+

Using our skilful hurricane loss prediction model, based on hazard, exposure and vulnerability attributes, we determined the TCDPS loss category derived from our model’s predicted loss for each historical hurricane (Fig. 10c and Fig. 11). The ranks of the most damaging hurricanes (Katrina, Harvey, and Andrew) are correctly predicted as being ‘loss category 5’. The next most historically damaging storms (from hurricanes Ian to Wilma) are represented as ‘loss category 4’ events, each causing
 485 between U.S.\$ 115-42 bn (Fig. 11). Hurricane Sandy (2012), which has been notably under-categorised by the Saffir–Simpson scale (category 1) and the ‘Tropical Cyclone Severity Scale’ of (Bloemendaal *et al.*, 2021) (category 2), is now



predicted to be a ‘loss category 4’ storm. Combining hazard, exposure and vulnerability quantities has provided a more skilful damage-based categorisation for this case, and many other cases, than is possible based only on hazard information.

However, the most damaging hurricane that is misrepresented by the TCDPS scheme is Hurricane Charley (2004), whose
490 observed damage (U.S.\$ 40.1 bn) is underpredicted by our model (Fig. 9c). Charley should be ‘loss category 4’ (Fig. 4 and Table 3), but its predicted damage (~U.S.\$ 20 bn) falls in ‘loss category 3’. Additionally, Hurricane Allison (2001) is ascribed ‘loss category 1’, which stands out as a poor model prediction (Fig. 11).



495 *Figure 11. Predicted ‘loss category’ of each historical landfalling hurricane, using the optimal prediction model (Fig. 10c) and the TCDPS (Table 3).*

Finally, we compared the success rates of both the Saffir–Simpson categories and ‘loss categories’ based on our predicted losses (Fig. 10c) with observed ‘loss categories’ (i.e., TCDPS categories derived from observed damage). We found 74 storms (70%) of landfalling hurricanes (with an available economic loss estimate from 1979) are correctly classified by the
500 TCDPS (Table 4). However, 28 storms (26%) are misrepresented by ± 1 ‘loss category’ and 4 (4%) are misrepresented by ± 2 ‘loss categories’ (Table 4). Overall, these results markedly outperform using landfall $v_{\max}$ and therefore the Saffir–Simpson scale to represent damage, for which the ‘loss categories’ of just over half (55%) of events are correctly represented, and 31%, 12% and 2% are misrepresented by ± 1 , ± 2 and ± 3 ‘loss categories’, respectively (Table 4).



505 *Table 4. Quantified differences between Saffir–Simpson category or predicted ‘loss category’ using our model (Fig. 10c) and the observed ‘loss category’ from the TCDPS (i.e., derived from observed damage).*

Difference from observed ‘loss category’	Saffir–Simpson category	Predicted ‘loss category’
0 (i.e., correct prediction)	58 (55%)	74 (70%)
±1	33 (31%)	28 (26%)
±2	15 (12%)	4 (4%)
±3	5 (2%)	0 (0%)

7 **Summary and discussion**

510 Hurricanes have historically been among the most destructive natural hazards globally. It is critical to understand hurricane risk, which is determined by various hazard, exposure, and vulnerability attributes, for effective mitigation of their impacts. This study explores statistical relationships between historical North Atlantic hurricane economic damage and hazard, exposure and vulnerability data, with two aims. First, to determine whether the inclusion of socioeconomic information into a predictive model for economic damage from U.S. landfalls yields significant additional skill than only using hazard
515 information. Second, to devise a damage-based hurricane classification scheme to improve both the representation of



historical damage events and the rapid assessment of damage potential for forecast events, complementing existing classifications, such as the Saffir–Simpson scale, which is important for public communication.

For historical hurricanes, we derived storm-centred hazard, exposure and vulnerability quantities, and limited our analysis to cases for which observed loss estimates are available and storm radius is either observed or could be skilfully estimated statistically from other observed size information. Variables found to significantly influence a given storm’s loss include storm intensity (wind, central pressure), inland flooding and coastal flooding from surge, storm size, the coverage and density of buildings, population and financial exposure impacted, and the vulnerability associated with the affected region (storm footprint). A statistical model, based on a random forest approach, was developed to predict hurricane damage, which outperforms other hazard-only methods to represent storm damage potential. Finally, a new ‘Hurricane Predictive Damage Scale’ was derived to categorise tropical storms based on their economic consequences (i.e., loss and damage) and is intended for use by stakeholders concerned with hurricane risk, including governmental agencies coordinating hurricane response and the (re-)insurance sector.

7.1 Key results and limitations

7.1.1 *Integrating hazard, exposure, and vulnerability data predicts historical hurricane damage more skilfully than hazard data alone*

Although economic damage generally increases landfall wind speed (and with Saffir–Simpson landfall category), hurricane wind intensity has limited skill ($\rho = 0.67$) in predicting economic damage, with a large RMSE across all storms, a finding which substantiates previous research (Klotzbach *et al.*, 2020; Klotzbach *et al.*, 2022a). Landfall c_p is a more skilful single predictor yet using both intensity quantities to predict damage results in many erroneous predictions, especially for the more damaging events. A range of hazard, exposure and vulnerability hurricane quantities at landfall have strong correlations with historical damage ($r > 0.5$), providing evidence that these should be combined to represent and predict hurricane damage. By using a weighted-rank-sum approach to optimally combine this hazard, exposure and vulnerability information yields markedly more skilful predictions of damage, with very high correlation between loss observations and predictions ($\rho = 0.89$ across all storms). This increase in skill is particularly evident in the improved predicted damage rank of the most impactful historical cases, resulting in a very skilful financial damage prediction for each past storm, with a significantly reduced RMSE of U.S.\$ 7.0 bn. This is a notable improvement from using landfall $v_{\max}$ and c_p to predict financial damage, which yielded RMSE of U.S.\$ 35.6 bn and U.S.\$ 26.4 bn respectively.



Our optimal prediction model was derived by determining the weighted sum of normalised ranks of landfall c_p , maximum rainfall accumulation, and the percentage of total GHSL building density damage within R64 and within 12 hours of landfall.

545 This model includes two hazard attributes representing hurricane intensity (c_p and rainfall), with hurricane radius included implicitly in determining percentage of total GHSL building density damage; includes exposure values of building density impacted by each storm; and applies a vulnerability function by determining the percentage of building damage proportional to v_{max} . Important limitations are related to the observational uncertainties in both input predictors and our target loss predictand. A key limitation is the lack of complete cyclone size information prior to 2002, which necessitated our statistical

550 estimation approach, and this could impact predictive skill for cases during this earlier period in our study's sample. An additional data gap exists in the availability of vulnerability information, and in its temporal granularity. Additional uncertainty around historical losses required our consideration of multiple datasets and, where required, calculating an average loss estimate.

7.1.2 *A damage-based hurricane scale can effectively communicate economic impacts*

555 We derived a 'Hurricane Predictive Damage Scale', in which hurricanes are ascribed a 'loss category', to categorise historical hurricanes by their economic damage and to help communicate the damage potential of forecast storms. This scale can complement the Saffir–Simpson scale, as well as published work (e.g., Bloemendaal *et al.*, 2021; Pilkington and Mahmoud, 2016), to provide an indication of a historical and forecasted storm's financial loss, that may be more effective for decision making, particularly within the (re-)insurance, banking and energy sectors. Evaluating this scale revealed that our

560 model-predicted 'loss categories' matched the observed categories determined from loss data in 70% of cases, which compares favourably with the Saffir–Simpson scale when categories are ranked.

The damage order of historical storms is much better represented and communicated using this 'Hurricane Predictive Damage Scale'. Notably, among many other cases, Hurricane Sandy (2012), a lower-intensity, larger-size storm, is appropriately reclassified in our scheme as a 'loss category 4' case, which more appropriately describes its historical damage

565 ranking. We anticipate that the TCDPS may be of significant utility within the disaster-response community, including government agencies, as well as financial sectors, such as (re-)insurance, banking, and energy, for which understanding historical and future risk are important, as is the season-to-season requirement of preparing for the potential consequences of an impending storm. This study provides a skilful method for determining the damage potential of forecast landfalls, notwithstanding the current limitations of numerical weather-prediction models in capturing, for example, the intensification

570 rates of TCs (Trabing and Bell, 2020).



7.2. Outlook

There are several clear data needs, which our study highlights. First, vulnerability data of higher granularity are needed, both spatial (e.g., building year and construction material per unit) and temporal (e.g., data of at least yearly frequency). There is potential for the development of novel, higher-resolution datasets using emerging technologies, such as machine learning and satellite imagery, to better describe building attributes, and vulnerability more generally. Second, there is a need for TC size information for historical events. This is a key physical quantity that determines a storm's directly impacted area, and reconstructions of size would be considerably beneficial. TC outer size is not projected to change over this century (Schenkel *et al.*, 2023), although supporting evidence from studies of high-resolution climate models, which adequately capture intensification and horizontal structure, would help substantiate this projection and evaluate its past and future risk implications.

Furthermore, there is potential to refine our approach for application to a variety of sectors and stakeholders, potentially developing sector-specific schemes to categorise historical events based on values that are meaningful. We have developed an initial scale based on financial damage, which would be most useful to financial industries (e.g., (re-)insurance and banking). For other sectors, a scale based on other impact metrics may be more suitable; for example, a scale based on population impact may be more suitable for governments and relief charities. There is a broad public need for open-source tools to provide these varied initial assessments of damage potential, and different tools may need to be developed to suit specific sectors. Such tools need to be readily understandable and transparent (input data, methods etc) and physically interpretable, reflecting the various ways that TCs cause damage.

Finally, our predictive model may be a helpful tool for investigating how TC risk may change in a warming climate. Applying our method to simulated future tropical cyclones, which may be more intense and induce heavier precipitation (Knutson *et al.*, 2020), may help quantify the extent to which U.S. hurricane-related losses may change with changes in hazard, exposure and vulnerability, and our methodology may be expanded to other TC-prone regions, where necessary data are available. This study provides an approach to more skilful quantification of hurricane losses and communication of risk, aiding more effective future TC risk management.

8 Acknowledgements

AJB is supported by the National Centre for Atmospheric Science and by the nextGEMS (European Union Horizon2020 grant agreement 101003470) and Huracán (NERC–NSF grant agreement NE/W009587/1) projects. Data analysis from this



600 study was facilitated by JASMIN (jasmin.ac.uk), which is operated by the Science and Technology Facilities Council on behalf of the Natural Environment Research Council.

9 Author contributions

605

AFV and AJB conceived the study. AFV performed data analysis and figure preparation, supported by AJB, VMH and JM. All authors interpreted results. AFV and AJB wrote the manuscript, led by AFV, and all authors approved the final draft.

610 10 Competing interests

The authors declare no competing interests.

615 11 Code / data availability

IBTrACS data available from ncei.noaa.gov/products/international-best-track-archive. All other datasets are available from the cited sources. Data analysis and visualisation code is available from github.com/ncas-metoffice-hrcm/hurricane_damage.

620 References

- AON. 2025. 2025 Climate and Catastrophe Insight. Available: <https://www.aon.com/en/insights/reports/climate-and-catastrophe-report> [Accessed 16 Sept 2025].
- Baldwin, J. W., Lee, C.-Y., Walsh, B. J., Camargo, S. J. and Sobel, A. H., 2023. Vulnerability in a Tropical Cyclone Risk Model: Philippines Case Study. *Weather, Climate, and Society* **15**, 503-523.
- 625 Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., McVicar, T. R. and Adler, R. F., 2019. MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society* **100**, 473-500.
- Blake, E. S., Landsea, C. and Gibney, E. J. 2011. The deadliest, costliest, and most intense United States tropical cyclones from 1851 to 2010 (and other frequently requested hurricane facts). NOAA technical memorandum NWS NHC-6.
- 630 Bloemendaal, N., de Moel, H., Mol, J. M., Bosma, P. R. M., Polen, A. N. and Collins, J. M., 2021. Adequately reflecting the severity of tropical cyclones using the new Tropical Cyclone Severity Scale. *Environmental Research Letters* **16**, 014048.



- Camelo, J. and Mayo, T., 2021. The lasting impacts of the Saffir-Simpson Hurricane Wind Scale on storm surge risk communication: The need for multidisciplinary research in addressing a multidisciplinary challenge. *Weather and Climate Extremes* **33**, 100335.
- 635 Chavas, D. R., Knaff, J. A. and Klotzbach, P., 2025. A Simple Model for Predicting Tropical Cyclone Minimum Central Pressure from Intensity and Size. *Weather and Forecasting* **40**, 333-346.
- Chavas, D. R., Lin, N. and Emanuel, K., 2015. A Model for the Complete Radial Structure of the Tropical Cyclone Wind Field. Part I: Comparison with Observed Structure. *Journal of the Atmospheric Sciences* **72**, 3647-3662.
- Chavas, D. R., Reed, K. A. and Knaff, J. A., 2017. Physical understanding of the tropical cyclone wind-pressure relationship.
 640 *Nature Communications* **8**, 1360.
- Delforge, D., Wathelet, V., Below, R., Sofia, C. L., Tonnelier, M., van Loenhout, J. A. F. and Speybroeck, N., 2025. EM-DAT: the Emergency Events Database. *International Journal of Disaster Risk Reduction* **124**, 105509.
- Eberenz, S., Stocker, D., Rösli, T. and Bresch, D. N., 2020. Asset exposure data for global physical risk assessment. *Earth System Science Data* **12**, 817-833.
- 645 Emanuel, K., 2011. Global Warming Effects on U.S. Hurricane Damage. *Weather, Climate, and Society* **3**, 261-268.
- Federal Emergency Management Agency. 2022. Hazus 5.1; Hazus Hurricane Model Technical Manual. Available: https://www.fema.gov/sites/default/files/documents/fema_hazus-hurricane-model-technical-manual-5-1.pdf [Accessed 18 May 2025].
- Federal Emergency Management Agency. 2024. National Risk Index: Future Risk Technical Documentation. Available:
 650 https://eelp.law.harvard.edu/wp-content/uploads/2025/03/NRI_Future_Risk_Technical_Document.pdf [Accessed 18 May 2025].
- Federal Emergency Management Agency. 2025. Building Code Adoption Tracking. Available: <https://www.fema.gov/emergency-managers/risk-management/building-science/bcat> [Accessed 18 May 2025].
- Gahtan, J., Knapp, K., Schreck III, C., Diamond, H., Kossin, J. and Kruk, M. C. 2024. International Best Track Archive for
 655 Climate Stewardship (IBTrACS), version 4r01. National Oceanographic and Atmospheric Administration National Centers for Environmental Information.
- Gallagher Re. 2025. Gallagher Re Natural Catastrophe and Climate Report 2024. Available: <https://www.ajg.com/gallagherre/-/media/files/gallagher/gallagherre/news-and-insights/2025/natural-catastrophe-and-climate-report-2025.pdf> [Accessed 18 May 2025].
- 660 Gori, A., Lin, N., Schenkel, B. and Chavas, D., 2023. North Atlantic Tropical Cyclone Size and Storm Surge Reconstructions From 1950-Present. *Journal of Geophysical Research: Atmospheres* **128**, e2022JD037312.
- Grinsted, A., Ditlevsen, P. and Christensen, J. H., 2019. Normalized US hurricane damage estimates using area of total destruction, 1900–2018. *Proceedings of the National Academy of Sciences*, 201912277.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R.,
 665 Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita,



- M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.-N., 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999-2049.
- 670 Hurricane Research Division. 2025. Detailed List of Continental United States Hurricane Impacts/Landfalls. Available: http://www.aoml.noaa.gov/hrd/hurdat/USHurrs_detailed.html [Accessed 18 May 2025].
- Intergovernmental Panel on Climate Change 2012. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. In: Field, C. B., V. Barros, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, M. D. Mastrandrea, K. J. Mach, G.-K. Plattner, S. K.
- 675 Allen, M. Tignor, and P. M. Midgley. (ed.). Cambridge University Press, Cambridge, UK, and New York, NY, USA.
- Jewson, S., 2023. The Impact of Projected Changes in Hurricane Frequencies on U.S. Hurricane Wind and Surge Damage. *Journal of Applied Meteorology and Climatology* **62**, 1827-1843.
- Klotzbach, P. J., Bell, M. M., Bowen, S. G., Gibney, E. J., Knapp, K. R. and Schreck, C. J., III, 2020. Surface Pressure a More Skillful Predictor of Normalized Hurricane Damage than Maximum Sustained Wind. *Bulletin of the American*
- 680 *Meteorological Society* **101**, E830-E846.
- Klotzbach, P. J., Chavas, D. R., Bell, M. M., Bowen, S. G., Gibney, E. J. and Schreck Iii, C. J., 2022a. Characterizing Continental US Hurricane Risk: Which Intensity Metric Is Best? *Journal of Geophysical Research: Atmospheres* **127**, e2022JD037030.
- Klotzbach, P. J., Wood, K. M., Schreck Iii, C. J., Bowen, S. G., Patricola, C. M. and Bell, M. M., 2022b. Trends in Global
- 685 Tropical Cyclone Activity: 1990–2021. *Geophysical Research Letters* **49**, e2021GL095774.
- Knutson, T., Camargo, S. J., Chan, J. C. L., Emanuel, K., Ho, C.-H., Kossin, J., Mohapatra, M., Satoh, M., Sugi, M., Walsh, K. and Wu, L., 2020. Tropical Cyclones and Climate Change Assessment: Part II: Projected Response to Anthropogenic Warming. *Bulletin of the American Meteorological Society* **101**, E303-E322.
- Landsea, C. W. and Franklin, J. L., 2013. Atlantic Hurricane Database Uncertainty and Presentation of a New Database
- 690 Format. *Monthly Weather Review* **141**, 3576-3592.
- Lockwood, J. F., Dunstone, N., Hermanson, L., Saville, G. R., Scaife, A. A., Smith, D. and Thornton, H. E., 2023. A Decadal Climate Service for Insurance: Skillful Multiyear Predictions of North Atlantic Hurricane Activity and U.S. Hurricane Damage. *Journal of Applied Meteorology and Climatology* **62**, 1151-1163.
- Muller, J., Mooney, K., Bowen, S. G., Klotzbach, P. J., Martin, T., Philp, T. J., Dhruvkumar, B., Dixon, R. S. and
- 695 Girimurugan, S. B., 2025. Normalized Hurricane Damage in the United States: 1900–2022. *Bulletin of the American Meteorological Society* **106**, E51-E67.
- National Centers for Environmental Information. 2021. The Saffir-Simpson Hurricane Wind Scale. Available: <https://www.nhc.noaa.gov/pdf/sshws.pdf>.



- National Centers for Environmental Information. 2025. U.S. Billion-Dollar Weather and Climate Disasters. Available:
 700 <https://www.ncei.noaa.gov/access/billions/>.
- National Oceanographic and Atmospheric Administration. 2024. Hurricane Costs. Available:
<https://coast.noaa.gov/states/fast-facts/hurricane-costs.html> [Accessed 18 May 2025].
- National Oceanographic and Atmospheric Administration. 2025. Tropical Cyclone Rainfall. Available:
<https://www.wpc.ncep.noaa.gov/tropical/rain/tcrainfall.html> [Accessed 18 May 2025].
- 705 Oliver-Smith, A., 2020. Hurricanes, Climate Change, and the Social Construction of Risk. *International Journal of Mass Emergencies & Disasters* **38**, 1-12.
- Pesaresi, M. and Politis, P. 2023. GHS-BUILT-S R2023A – GHS built-up surface grid, derived from Sentinel2 composite and Landsat, multitemporal (1975–2030). European Commission, Joint Research Centre.
- Pilkington, S. F. and Mahmoud, H. N., 2016. Using artificial neural networks to forecast economic impact of multi-hazard
 710 hurricane-based events. *Sustainable and Resilient Infrastructure* **1**, 63-83.
- Rappaport, E. N., 2014. Fatalities in the United States from Atlantic Tropical Cyclones: New Data and Interpretation. *Bulletin of the American Meteorological Society* **95**, 341-346.
- Schenkel, B. A., Chavas, D., Lin, N., Knutson, T., Vecchi, G. and Brammer, A., 2023. North Atlantic Tropical Cyclone Outer Size and Structure Remain Unchanged by the Late Twenty-First Century. *Journal of Climate* **36**, 359-382.
- 715 Trabling, B. C. and Bell, M. M., 2020. Understanding Error Distributions of Hurricane Intensity Forecasts during Rapid Intensity Changes. *Weather and Forecasting* **35**, 2219-2234.
- U.S. Bureau of Economic Analysis. 2023. Current-cost net stock of fixed assets and consumer durable goods [K1WTOTL1ES000]. Available: <https://fred.stlouisfed.org/series/K1WTOTL1ES000> [Accessed 1 February 2025].
- U.S. Bureau of Economic Analysis. 2025. National Data Fixed Assets Accounts Tables. Available:
 720 https://apps.bea.gov/iTable/?reqid=10&step=3&isuri=1&table_list=16#eyJhcHBpZCI6MTAsInN0ZXBzIjpbMSwyLDMsM10sImRhdGEiOltbInRhYmxlX2xpc3QiLCIxNiJdLFsiQ2F0ZWdvcmllcyIsIlB1YmxpY0ZBQSJdLFsiU2NhbnGUiLCItOSJdLFsiRmlyc3RfWWVhcnIsIjE5NTAiXSxbIkxhc3RfWWVhcnIsIjIwMjMiXSxbIlNlcmlncyIsIkEiXV19 [Accessed 18 May 2025].
- U.S. Census Bureau. 2024. American Community Survey Housing Units by Year Built Variables — Tract. Available:
 725 https://hub.scag.ca.gov/datasets/41c90eef3a12451faff10c2a6b26fc46_2/about [Accessed 18 May 2025].
- Vickery, P., J., Skerlj, P., F., Lin, J., Twisdale, L., A., Young, M., A. and Lavelle, F., M., 2006. HAZUS-MH Hurricane Model Methodology. II: Damage and Loss Estimation. *Natural Hazards Review* **7**, 94-103.
- Ward, P. J., Blauhut, V., Bloemendaal, N., Daniell, J. E., de Ruiter, M. C., Duncan, M. J., Emberson, R., Jenkins, S. F., Kirschbaum, D., Kunz, M., Mohr, S., Muis, S., Riddell, G. A., Schäfer, A., Stanley, T., Veldkamp, T. I. E. and Winsemius,
 730 H. C., 2020. Natural hazard risk assessments at the global scale. *Nat. Hazards Earth Syst. Sci.* **20**, 1069-1096.
- Wehner, M. F. and Kossin, J. P., 2024. The growing inadequacy of an open-ended Saffir–Simpson hurricane wind scale in a warming world. *Proceedings of the National Academy of Sciences* **121**, e2308901121.



- Weinkle, J., Landsea, C., Collins, D., Musulin, R., Crompton, R. P., Klotzbach, P. J. and Pielke, R., 2018. Normalized hurricane damage in the continental United States 1900–2017. *Nature Sustainability* **1**, 808-813.
- 735 Willis Towers Watson. 2024. Natural Catastrophe Review 2024. Available: <https://www.wtwco.com/en-eg/insights/2025/01/wtw-natural-catastrophe-review-2024> [Accessed 18 May 2025].
- WorldPop 2018. Global High Resolution Population Denominators Project. In: www.worldpop.org — School of Geography and Environmental Science, U. o. & Southampton; Department of Geography and Geosciences, U. o. L. D. d. G., Universite de Namur; and Center for International Earth Science Information Network (CIESIN), Columbia University. Funded by The
- 740 Bill and Melinda Gates Foundation (OPP1134076). (eds.).
- Young, R. and Hsiang, S., 2024. Mortality caused by tropical cyclones in the United States. *Nature* **635**, 121-128.
- Zuzak, C., Goodenough, E., Stanton, C., Mowrer, M., Ranalli, N., Kealey, D. and Rozelle, J. 2021. National risk index technical documentation. 37. Available: <https://hazards.fema.gov/nri/>.

745