CLIMATE FINANCIAL RISK FORUM

A Risk Professional's Guide to Physical Risk Assessments

A GARP Benchmarking Study of 13 Vendors

October 2025

Contents

This chapter represents the output from the Climate Financial Resilience Working Group, part of the Climate Financial Risk Forum (CFRF).

This CFRF guide has been written by industry, for industry. The recommendations in this guide do not constitute financial or other professional advice and should not be relied upon as such. The PRA and FCA have convened and facilitated CFRF discussions but neither they, nor any contributors or organisations named in the chapter, accept liability for the views expressed in this guide which do not necessarily represent the view of the regulators or any contributor or organisation named and in any case do not constitute regulatory guidance. The views in this chapter reflect the individual participants and not necessarily the views of their employers and do not represent a commitment by any of them to a particular course of action. In some cases, where input has been provided it has only been to certain parts of the document and not to others. The material in the chapter 'A Risk Professional's Guide to Physical Risk Assessments' has not been independently verified by the CFRF Working Group members.

Acknowledgements and Disclaimers

This study was undertaken by Jo Paisley and Maxine Nelson, of the GARP Risk Institute, in 2025 on behalf of the Climate Financial Risk Forum (CFRF). We are grateful to support from the CFRF Financial Resilience Working Group, chaired by Billy Suid, Barclays. Thanks to Tom Clarke, KPMG, for support on survey design.

We are also very grateful for the participation and engagement of the following physical risk vendors that took park in the survey:

- Climate X
- Fathom
- First Street
- ICE
- JBA Risk Management
- Jupiter Intelligence
- Moody's
- MSCI
- Planetrics, a McKinsey & Company solution
- Riskthinking.AI
- S&P Global
- Twinn by Haskoning
- XDI

All data submitted by the firms remain anonymous throughout the study.

During an engagement, the relationship between a vendor and financial firm will include a great deal of conversation and nuance. Our study, in contrast, was a one-off questionnaire, which did not allow for the detailed interactions that would be part of a vendor relationship. So, readers should take this study as indicative, rather than definitive, of the scale of variation across vendor estimates.

Please also note the following disclaimers:

This report has been prepared by GARP using inputs from multiple providers, including JBA Risk Management Limited (JBA). JBA provided selected baseline and climate-change data only; GARP determined the scenarios, portfolio inputs, assumptions and all resulting findings and conclusions. JBA's climate-change data are derived from scientifically credible third-party climate models and JBA's own methodologies, which carry known limitations and deep uncertainties and cannot predict the future. The data illustrate possible outcomes under idealised scenarios and should be used with caution and full awareness of their limitations.

This report has been created by GARP drawing on selected data provided by Planetrics, a McKinsey & Company solution (which does not include investment advice). This report represents GARP's own selection of scenarios and portfolio data. GARP is solely responsible for, and this report represents, such scenario selection, all assumptions underlying such selection, and all resulting findings, and conclusions and decisions. McKinsey & Company is not an investment adviser and has not provided any investment advice.

1 Background and Context

Motivation and Overview of the Study

The physical risks associated with a changing climate are becoming increasingly evident[1]. For example, in 2024, the United States alone was hit by 27 weather/climate disaster events with losses for each exceeding USD 1 billion[2]. This is three times higher than the annual average for the period 1980–2024. While there is uncertainty over exactly how much physical risk we can expect over coming decades, it is indisputable that the frequency and severity of physical risk will continue to increase due to rising concentrations of anthropogenic greenhouse gas emissions.

Although we know that the overall threat level is rising, it is not straightforward to translate this into how physical risks might impact individual firms, including those in the financial sector, or even individual assets. For this, a granular spatial resolution of analysis is needed. The risk profiles for a bank, insurance company, asset owner or asset manager differ — as discussed in many CFRF documents over the years — but all firms share the common challenge of how to assess the physical impacts that they are exposed to, be that in their own operations or in the firms that they lend to, underwrite, insure or invest in. This study focusses on just that — how to assess physical risk at the level of the individual asset and how much estimates vary.

These is pressure on financial firms to have more expertise in this area as their supervisors are increasingly setting formal expectations that regulated firms embed climate risk considerations within their day-to-day risk management. The PRA, for example, was the first regulator to do so, with the release of SS3/19 in 2019 and a consultation to update it released in 2025 (Consultation Paper 10/25).

As this was a new discipline for most financial firms, many turned to third party providers (“vendors”) to help them with different areas of expertise. There are now many physical risk data vendors, which offer a variety of services to financial institutions. While vendor offerings often sound alike — providing projections of how physical risk could evolve for locations across a range of risks and climate scenarios — they can differ significantly in terms of features, approach, or suitability for specific needs, and the underlying models that these providers use differ in methodology and assumptions.

This benchmarking study, undertaken by the Global Association of Risk Professionals (GARP), examines the extent to which third party data vendors differ in their property-level risk assessments and projections. As far as we are aware, it is the first study of its kind, providing insights into why risk assessments might differ, as well as some quantification of that dispersion.

The benchmarking exercise comprises three main elements:

- A **methodological survey,** which provides a high-level view of the approach and capabilities of each participating vendor. This is covered in section 2.
- An **asset location survey** that explores how property location is allocated to

[1] See World Meteorological Organization’s supplement on ‘Significant Weather & Climate Events 2024’: https://wmo.int/files/significant-weather-climate-events-2024
[2] NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters (2025). https://www.ncei.noaa.gov/access/billions/, DOI: 10.25921/stkw-7w73

properties which do not have a known address, or which have only a vague address (such as the county or state). This is covered in section 3.

- A **physical risk quantification study,** which investigates how close different vendors' estimates of risk are for different types of physical risk, over time, at a range of locations. To do this, vendors provided estimates of the physical risk affecting a common portfolio of 100 properties, spread across the U.K., mainland Europe, the U.S., and Asia. The survey design is covered in section 4, with the results presented in section 5.

A key objective of the study is to help financial institutions better navigate the complex landscape of physical risk modelling. There are no right or wrong answers or single best vendors, but some might be more suitable for particular use cases. Throughout the document, we therefore highlight lessons as we go along. These are brought together in section 6, which contains a checklist of questions financial institutions need to ask themselves — and vendors — before undertaking a physical risk assessment. This will help firms better navigate the landscape of physical risk assessments, improving their risk management capabilities, and helping customers' outcomes as well. Conclusions are presented in section 7.

This study highlights a wide spread of results for modelling physical risk at the level of a particular asset. The study findings show that, as with any forward-looking exercise, there is a great deal of uncertainty about the modelling of physical risk, arising from a range of factors, such as methodological and modelling choices, downscaling approaches, and data used.

The inherent uncertainty within physical risk modelling underscores the importance of understanding the strengths and limitations of different vendors' approaches. The choice of climate scenarios, the granularity of hazard modelling, and the treatment of asset-level information can all substantially affect risk estimates and, consequently, strategic decision-making.

Given this context, the benchmarking exercise seeks not only to identify the range of vendor estimates but also to illuminate the drivers of divergence, providing valuable guidance for financial institutions navigating this evolving landscape. The findings highlight the necessity for transparency and robust due diligence when selecting and working with third-party data providers. As climate-related disclosures and regulatory expectations continue to grow, financial firms must enhance their internal expertise and foster open dialogue with vendors to ensure that physical risk assessments are both meaningful and actionable.

Key Takeaways

There is a large spread of hazard and physical risk estimates. For the 13 vendors in the survey, this dispersion was observed for both estimates of the physical hazards (such as flood depth or wind speed), as well as for the damages associated with those hazards.

Complex modelling underlies the dispersion. Although dispersion is not surprising, particularly given how complex the underlying modelling and methodological choices are, the scale of the dispersion is significant. For example, a property can be assessed as either highly exposed to flooding (or a different hazard) by one or more vendors, but not at all exposed by other vendors.

Estimates of vulnerability are highly variable. When we looked at the relationship between the mean estimates of hazards and the mean damage ratios, there was a sensible and intuitive link. However, this relationship was extremely variable at the level of the individual vendor. Estimates of damages vary considerably, with each vendor having their own perspective on vulnerability.

There is little standardisation of metrics. For some hazards, there is a degree of comparability: metrics for flooding were the most comparable, followed by cyclones and windstorms. However, there are several hazards where there is no standardisation, which makes comparisons across vendors challenging. For example, there were many different metrics for heat-related hazards. That said, different metrics may be more suitable for different use cases (although examining this is beyond the scope of this study).

Measuring how dispersion changes over time and with changes in severity is challenging. When measured by standard deviation, the dispersion of physical risk assessments increases for time horizons further in the future and by the severity of the return period. But when dispersion is measured by the coefficient of variation (standard deviation divided by the mean), we get the opposite result, since the mean of the distribution is rising faster than the standard deviation. Overall, this may simply indicate that all vendors are anticipating rising physical risks on average, driven by climate change.

Asset location accuracy can make a big difference to the results. The quality of asset location information impacts both the estimate of a hazard's severity and the associated physical risk assessment. Even small differences in location can make a difference to estimates for some perils — for example, flooding, but is far less important for perils like heat.

There are broader impacts of climate change that this study does not address (although some vendors do examine these). This study offers useful insights into vendor estimates of physical risk, but it's important to recognise broader uncertainties. Some arise from material assumptions — such as asset vulnerabilities, static flood defence standards, and simplified exposure pathways — which can drive significant differences in estimates. Beyond this, systemic climate impacts on food systems, health, and migration are difficult to model but increasingly relevant. Financial institutions will need to build internal capacity and collaborate with vendors and the broader research community to assess risks that extend beyond asset-level hazards and losses.

2 Physical Risk Assessments: Introduction and Overview

Introduction to Calculating Physical Risk

Before describing the survey, it's helpful to introduce the concept of physical risk measurement. First, let's define a few terms that are fundamental:

- **Peril** is a potential specific event or situation that could cause loss or damage, for example, a flood, wildfire, or hurricane.

- **Hazard** is the physical manifestation of a peril, described in terms of its intensity or magnitude at a given location, for example, flood depth, wind speed, peak temperature.

- **Exposure** is what is at risk if a peril occurs. For example, homes built in a flood zone are more exposed to flood than those built on a hill. The risk in the portfolio of a financial institution that lends to people buying property will be driven by the proportion of the mortgage book that is secured on properties that have exposure to the peril. In the U.K., for example, flooding is a major driver of exposure to physical risk.

- **Vulnerability** captures how susceptible something (e.g., a house or factory) is to being harmed when a peril occurs. For example, a house built on stilts is far less likely to be vulnerable to flooding than one that is not.

To illustrate the concepts that are explored in the survey, let's start with a physical structure, say a house. A bank that is considering providing a mortgage against that house, or an insurance company offering insurance on it, will be interested in several things from a physical risk perspective. For example:

- Where is it located? Is it in an area that is known to be affected by a hazard(s), such as flood, wildfire, or hurricanes?
- What are the physical characteristics of that property? For example, what is it built of? How old is it? What building standards does it adhere to?

Let's assume that this house is built near a river that has flooded a few times in the past five years but has not suffered great deal of damage to date, although properties nearby have been badly damaged.

- How is the changing climate likely to affect the likelihood and severity of future flooding?
- How could the flooding vary across different possible future emissions' pathways?

Now, let's assume that the property is a factory, with a large footprint in an area near a port. Some relevant considerations are:

- How do you measure the area that might be affected by a hazard? Do you assume the hazard impacts one point in the footprint? Part of the footprint? The whole footprint?
- If you ask for a physical risk assessment, does it matter if the vendor looks at the front entrance, the centre of the site, or the area most at risk?
- Will the estimate of risk include contents or just building damage? Will it also include business disruption?

Physical risk vendors will tackle these issues differently. Each stage of modelling introduces scope for differences in approach and methodology. Vendors also differ in their choice of metrics by which to measure hazards. Given the differences in vendors' physical risk assessments, the range of approaches is summarised before the variability in their outputs is discussed.

Vendors' Modelling: A Range of Approaches

It's helpful to look at some of the key approaches to modelling. Although this section simplifies the landscape, it is presented to illustrate the complexity of the modelling choices underpinning physical risk assessments which make vendor comparisons extremely challenging. We also look at how vendors deal with model uncertainty.

We start with **climate models**, the outputs from which many of the vendors use as an input to their risk calculations. Climate models are (often physics-based) models that simulate different parts of the climate system, including the atmosphere, oceans, land surface, and ice. They can also simulate the response of the global climate system to different drivers, including increasing greenhouse gas concentrations.[3] Some key points to note about these models are:

- There are a variety of climate models. Global climate models (GCMs) are complex and simulate climate processes globally. Earth system models are the most advanced GCMs and include biological and chemical processes (e.g., the carbon cycle and vegetation growth). The vendors tend to use outputs from climate model experiments conducted as part of the scientific community's Coupled Model Intercomparison Project (CMIP)[4].
- The models divide up the Earth into a series of grid cells. The size of these cells is called its spatial resolution, which is typically quite coarse for a global model, e.g., 100–250 km. As the grid cells become smaller, more detail is provided but this comes at the cost of needing more computational power to run the models and the need to parameterise processes where data or understanding may be limited.
- To make the output from these models appropriate for physical risk modelling for financial institutions, it needs to be downscaled — that is, outputs need to be refined to produce higher resolution climate information that can be used for regional or local analysis. Downscaling can be achieved using regional climate models, which have a higher resolution and are focussed over a specific region of the earth, as well as a variety of statistical methods.
- Climate models may also need bias corrections or calibration. This is to make the modelled output more closely align with available observations.

Some of the vendors use **catastrophe models** to calculate the behaviour of extreme events. There are a few insights on these models worth highlighting:

- Since historical observations alone do not provide enough data to model extreme events, catastrophe models typically use pre-generated stochastic event sets – large collections of statistically representative synthetic events based on the joint distribution of location, intensity, and spatial extent derived from historical records. These event sets allow models to estimate risk across a plausible range of event frequencies and severities.
- The benefit of this approach is that the model can assess the spatial correlation of these events, for example, the likelihood of two locations flooding at the same time during the event, and to what extent each location will be impacted.

[3] For more details about climate modelling, please see the briefing by Carbon Brief.
[4] CMIP is a standardised framework for making comparisons across the models to identify their relative strengths and weaknesses. For example, climate models might be run over common historical periods so that their predictions of the past climate can be compared with observations. They might also model the impact of an annual one percent increase in atmospheric CO_2 concentrations.

- Catastrophe models also include modules covering exposure data, vulnerability and financial modelling.
- Some catastrophe models have been created based on downscaled climate model data. Some are “climate conditioned” — so that their output better reflects the changes that will occur in the future as the climate changes. In practice, this conditioning can be done by adjusting the financial impacts, or climate-conditioning the underlying event set (e.g., by altering the likelihoods of various perils based on climate modelling).

Some vendors also use **integrated assessment models** (IAMs):

- These bring together societal characteristics and simple climate models; simulating how population changes, economic growth, and energy use can affect the climate.
- IAMs are often used to project future greenhouse gas emissions and can be used for a variety of purposes (for example, one vendor uses them to ensure consistency between greenhouse gas emissions and the potential impacts of climate change).

Some vendors use **structural models** that capture the physical characteristics of the buildings or physical structures that they are analysing (e.g., incorporating information from building codes and the physical construction of a building to assess its vulnerability).

This benchmarking exercise has asked the vendors to focus on just one climate scenario, namely RCP 8.5, a high emissions/high physical risk scenario. However, one vendor has developed a methodology which generates a probability-weighted synthesis of available climate science. Focusing on one scenario is not their preferred approach, as they would typically generate a single, comprehensive, probability distribution of potential outcomes for any particular physical asset.

Beyond these broad categories of models, vendors differ in the way that they assess exposure, vulnerability, and the financial impacts of physical events:

- **Exposure data** captures information on the location and characteristics of the assets, such as property geo-location, the year it was built, its construction type, number of floors, basement type, value of contents, and types of potential losses that might arise from the asset not being usable for some time following the event.
- There will be differences in how vendors model the damages associated with different intensity levels of a hazard. Typically, vendors use **vulnerability curves**, which will vary by construction types and occupancy.
- Vendors also have a variety of ways of calculating the **financial impacts**, which bring together the hazard, exposure and vulnerability for each event and calculate a variety of loss metrics.

In practice, many vendors use a combination of climate model data, precomputed hazard maps, stochastic event sets, and vulnerability models — with methods tailored by peril, geography, and intended application. Understanding this layered architecture is essential to interpreting variation in vendor outputs.

In summary, this is an extremely complex landscape, with a large range of modelling approaches. It is inevitable that vendor assessments will differ, but the complexity of modelling choices makes it extremely challenging (if not impossible) to make any definitive insights on the reason that vendors’ assessments might differ.

Lesson 1. Financial institutions should engage with vendors to understand their approach to modelling so that the overall approach, the modelling assumptions used, and how climate change is accounted for in any outputs is understood.

Dealing With Uncertainty

Just outlining the range of modelling approaches illustrates very clearly how likely it is that vendors' modelling outcomes will differ. So, we asked the vendors whether and how they communicate uncertainty to their clients. Having some sort of measure of uncertainty can help clients understand the level of confidence they can have in a model's predictions.

Nine of the 13 vendors provide a measure of uncertainty for their customers — some provide multiple measures — and they varied quite a bit (four vendors do not provide a measure of uncertainty):

- The most common approach is to provide the standard deviation of modelled output (such as the damage ratio or average annual loss), followed by providing readings from the extremes of the distribution, such as the 5^{th} and 95^{th} percentiles (e.g., for the peril metrics, or cost distributions). One vendor noted that their approach would depend on the client's ability to digest any extra data.
- Another approach is to provide accuracy scores (indicating how well a model can explain observed historical hazards in a chosen region).
- One vendor provides distributions around the input data (such as hazard intensities, damage ratios), which can then be propagated through the model to see its impacts. In effect, this provides sensitivity analysis to help the user understand how different parameters are driving a portfolio's losses.
- One vendor provides information on model precision, and key uncertainty indicators.
- Another vendor provides full probabilistic distributions derived from a large ensemble of model experiments. These distributions capture the range and likelihood of projected outcomes.

Uncertainty in modelled outputs arises from multiple sources, including scenario choice, model structure, parameter selection, and input data quality. It's important to note that different vendors may characterise and quantify uncertainty in different ways, making direct comparisons challenging without understanding the context and method behind each estimate.

> **Lesson 2.** Financial institutions should consider how important it is to have measures of uncertainty provided with the modelled outputs. They should take steps to understand the uncertainties in the modelled outputs they are using, the ways this uncertainty may impact hazard and financial loss projections, and consider how to factor this into their decision making and risk management.

Differences in Vendors' Scope and Approaches

In this section, we highlight a range of other ways that the vendors differed in scope and approach when responding to this survey, including types of hazards and regions assessed, scenarios on offer, time horizons covered, choice of metrics, summarising options, model validation, and data delivery approaches.

Hazards and Regions Assessed

Beyond differences in the modelling approaches and ways of dealing with uncertainty, the vendors also cover different perils and regions (see Figure 1). In our sample, two vendors are flood specialists, while the others cover a range of hazards.

Figure 1 Vendors' Geographic Coverage of Perils

	North America	South America	Mainland Europe	U.K.	Africa	Middle East	Asia	Australia / Oceania
Fluvial Flooding	13	13	13	12	13	13	13	13
Pluvial Flooding	13	13	13	12	13	13	13	13
Coastal Flooding	13	13	13	12	13	13	13	13
Combined Flooding	6	6	6	6	6	6	6	6
Heat	10	10	10	9	10	10	10	10
Drought	10	10	10	9	10	10	10	10
Hurricanes/ Typhoons/Cyclone	10	10	10	10	10	11	11	11
Tornadoes	3	3	3	3	3	3	3	3
Windstorm	8	8	9	8	9	8	8	8
Coastal Erosion	3	3	3	4	3	3	3	3
Subsidence	4	4	4	4	4	4	4	4
Wildfires	11	11	11	10	11	11	11	11

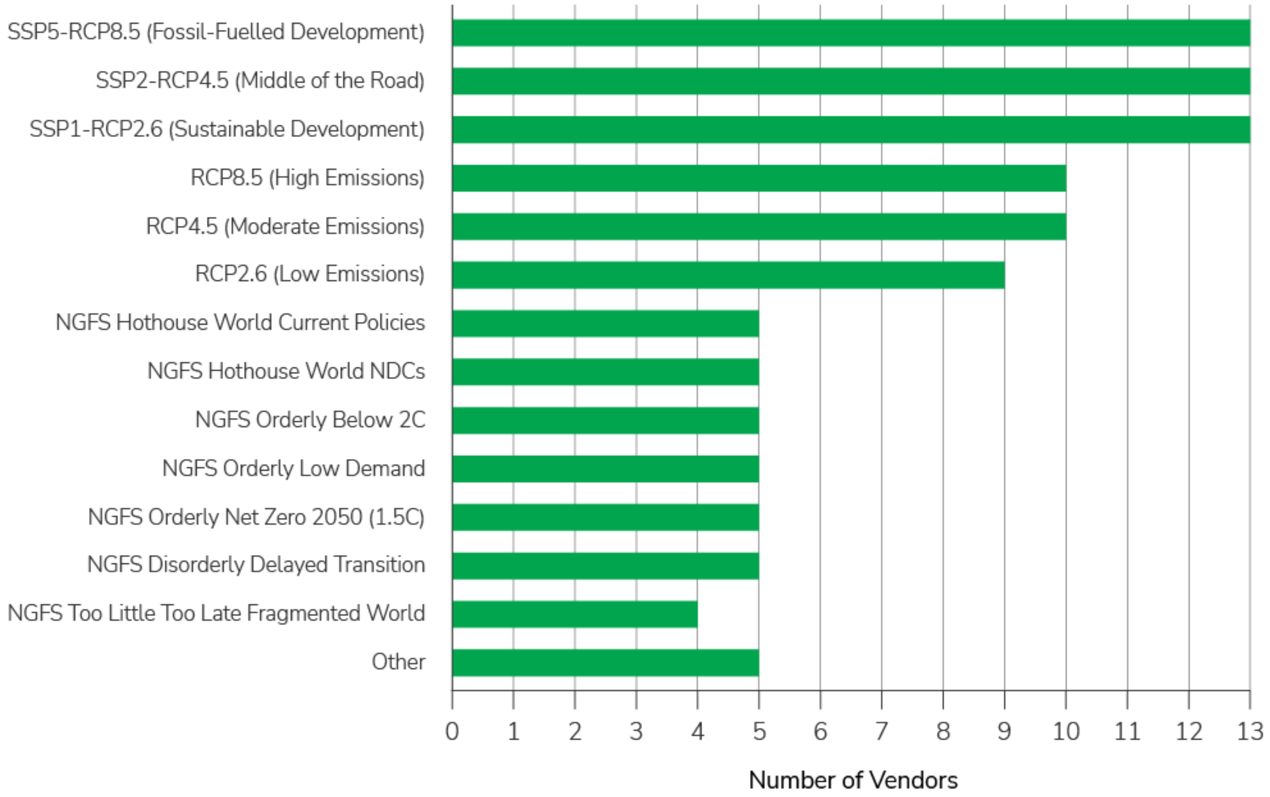
Note: Figures in the table are the number of vendors in our survey, out of a maximum of 13. Some vendors also assess other perils such as landslides, extreme cold, heavy snowstorms, heavy precipitation, hail,and freeze-thaw.

Lesson 3. Financial institutions need to consider which perils are most relevant and material to their portfolios, before choosing a suitable vendor(s).

Scenarios Assessed, Time Horizons and Intervals

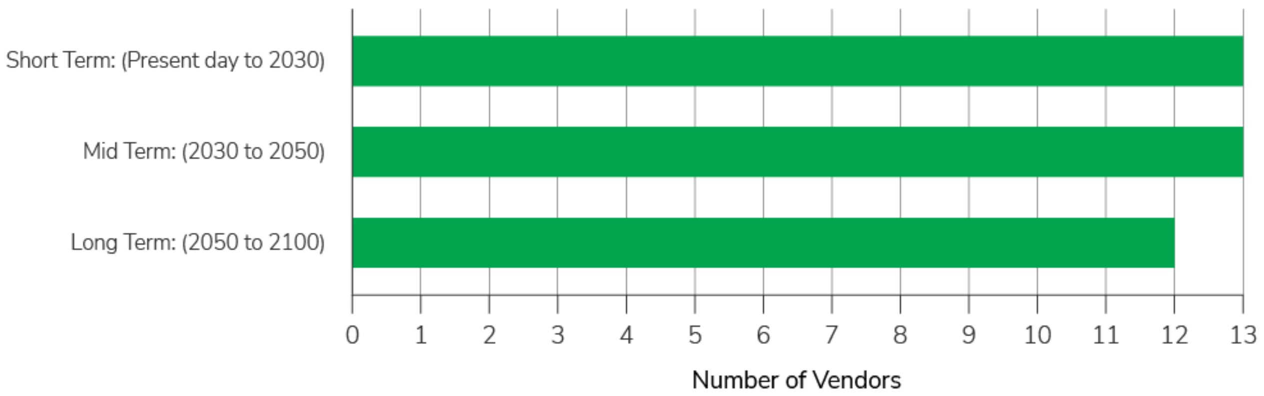
The vendors can model a range of "off-the-shelf" scenarios, as can be seen in Figure 2. The SSP-RCP scenarios developed in support of CMIP6 are the most popular. Those working closely with financial services firms have developed NGFS scenario capabilities, with some also mapping between these two broad families. Under 'other' are further scenarios which some vendors advised they can model, including Inevitable Policy Response scenarios, other RCP and SSP scenarios, as well as completely stochastic scenarios based on various emissions pathways and associated warming levels.

Figure 2 Climate Scenarios That Are Modelled



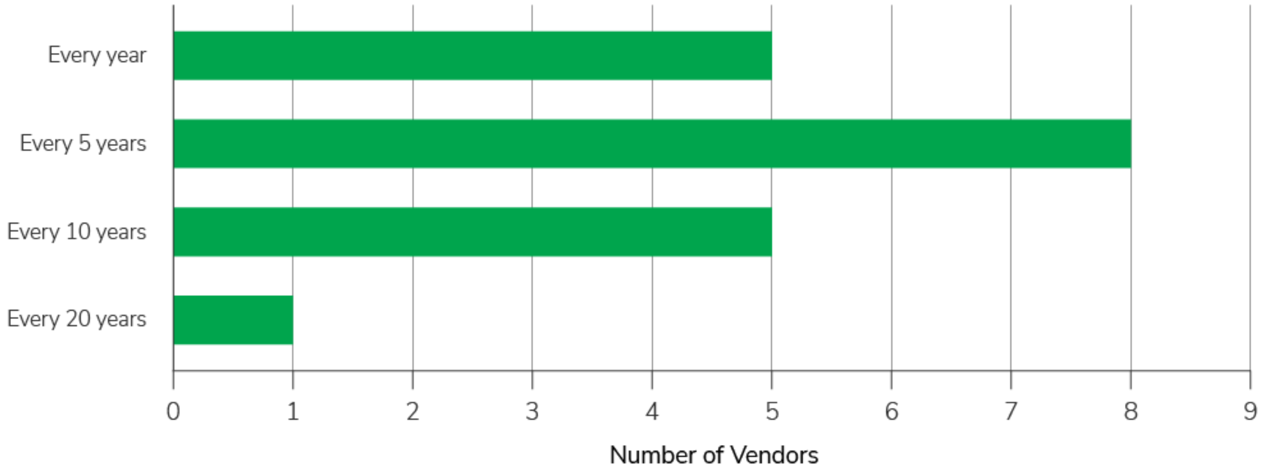
The vendors differ in the climate risk time horizons that they model, which will largely be a function of the scenarios that they choose. As Figure 3 shows, the most popular time horizons for modelling are the short (present day to 2030) and medium term (2030 to 2050), closely followed by long term.

Figure 3 Climate Risk Projection Time Horizons



Vendors differ in the time intervals that they model as well, as shown in Figure 4. The most popular option is five yearly. This figure shows the shortest time interval modelled, so, for example, a vendor that can model every five years can also model every 10 and 20 years.

Figure 4 Shortest Time Intervals Modelled



Some offered more frequent, annual, reporting, while others provided variants, such as having different frequencies for different time horizons. For many time horizons, twelve of the vendors can model every five years and all 13 can model every 10. Some vendors can model at bespoke time intervals.

The granularity of modelling also differs across vendors, with it also varying by peril (for example, more detailed modelling is needed for some perils such as flooding but is far less important for perils such as heat). For example, the granularity for different types of flooding ranged from 3-metre grids up to 1-kilometre grids. As one vendor pointed out, this interacts with the accuracy of geolocation data. For example, a flood map with 30 metres resolution assigns flood risk values to each 30-metre grid square; so, geolocation errors greater than 30 metres could result in a property being assessed using the flood risk from an entirely different grid square.

Of course, it's also important to consider how meaningful any very granular modelling is, or modelling a long way into the future.

> **Lesson 4.** Financial institutions should think carefully about which scenarios they need to be able to analyse, which time horizons they require, and the appropriate frequency of reporting.

Metrics Produced by Vendors

One of the most challenging aspects of undertaking this benchmarking survey was to decide on a common set of metrics, so that we could directly compare the vendors' hazard and physical risk assessments.

At a first stage, we asked vendors to supply us with details of the metrics that they commonly use with their clients for different hazards. These varied across vendors, for example, for hazard metrics:

- Those used for flooding were generally the most consistent and were typically flood depths expressed in metres.
- The most common metric for tropical cyclone was wind speed — either kilometres per hour or meters per second. However, that could be measuring different things, such as sustained 1- or 10-minute winds or 3-second gusts. But one vendor used the probability of a cyclone being a certain category on the Saffir-Simpson hurricane wind scale.

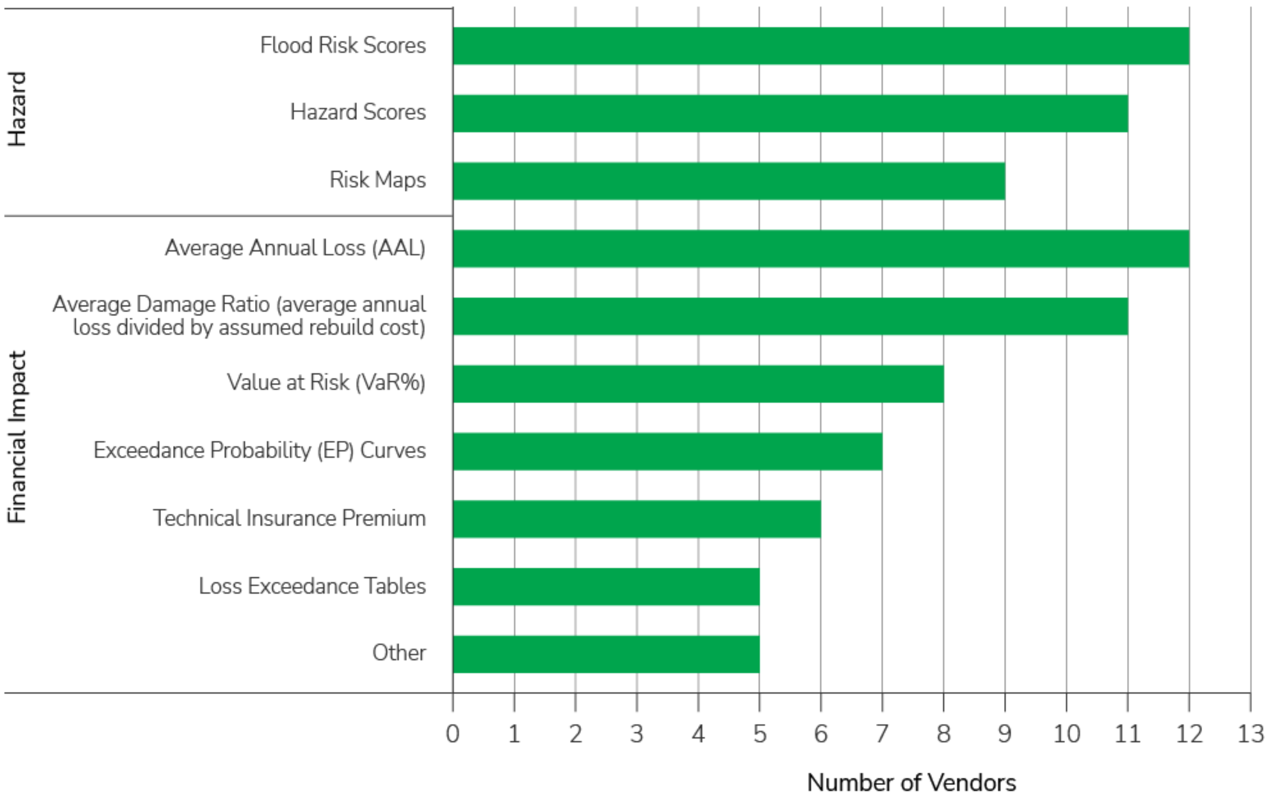
For other hazards, there was less common ground. Figure 5 provides a few of the metrics used for the three hazards on which there seemed to be least agreement, as illustrative of the broader issue of how difficult it is to compare physical risk data vendors. (Note that this study does not include an assessment of drought because the metrics used across vendors varied so much.)

Figure 5 Vendor Metrics Used for Selected Hazards

Fire	Drought	Heat
Fire weather index	Standardised precipitation indices	Various wet bulb measures
Annual wildfire probability	Water deficits	Cooling degree days
Susceptibility to fire	Water demand relative to supply	Location specific heat thresholds
Count of observed wildfires	Precipitation below a threshold	Days exceeding a particular temperature
Burn probability	Consecutive dry days	Heatwaves
Probability of catastrophic wildfire		

Vendors also use a variety of measures for both hazard and financial impacts of the physical risks associated with the hazards, as can be seen in Figure 6. Flood risk and hazard scores were more common hazard outputs than risk maps. Average annual loss (AAL) calculations were the most common financial impact output, followed by average damage ratio.

Figure 6 Vendors' Outputs Generated from Climate Risk Assessment Outputs

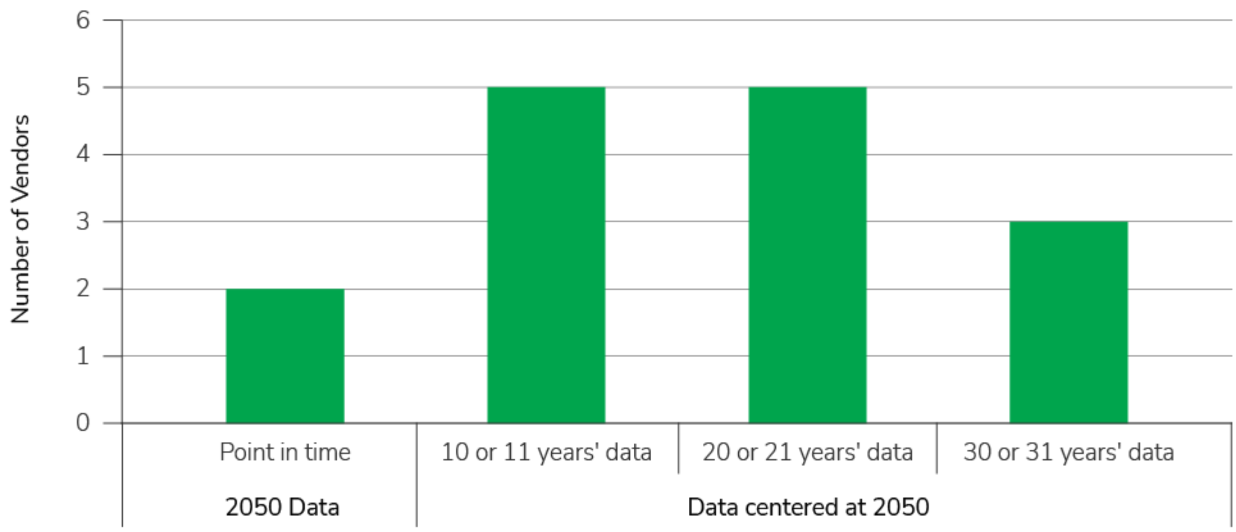


A wide range of indicators were reported under 'other', including business impacts (such as measures of business disruption or lost productivity), economic and supply chain impacts (such as impacts on critical infrastructure) and other measures of risk (e.g., damage ratio distributions).

Even if vendors had identical models and reported the same metrics, there might still be variation in responses simply because of the way that outputs are summarised. As

Figure **7** illustrates, when reporting a hazard at a point in time (2050 in this example) vendors will often present the data as a centred average (e.g., 10 or 20-year centred average) or use another method of summarising data either side of a point in time. Some vendors also choose different time periods over which to average, depending on the hazard.

Figure 7 Data Horizon For a 'Hazard in 2050'



In summary, even if a financial institution is clear on the hazards it needs to assess, a range of metrics is available for the hazard and its corresponding physical risk, and careful consideration is needed to identify the most appropriate one.

> **Lesson 5.** Financial institutions should consider which metrics for hazards and for financial losses are most suitable for their use cases and discuss options with the vendors. Financial institutions need to understand any summarising or aggregation of data that vendors apply when calculating outputs.

Model Validation and Data Delivery Approaches

Model validation approaches vary across vendors. Similarly, there are differences in how transparent they are regarding their methodologies: some publish their methods for all to see, some employ peer review processes, all provide documentation to clients, but with differing levels of detail. For a regulated firm using these services, understanding these approaches will be particularly important.

A robust model risk management (MRM) framework is needed to ensure that vendor models enable users to meet the requirements of IFRS9. Vendors varied considerably in their approach to MRM.

Financial institutions should also consider the format of the data that they will receive from vendors. Although not a central part of this benchmarking survey, we found that there was variation in how the vendors provided us with the data, with some not able to fill in the template that we provided.

> **Lesson 6.** Financial institutions need to consider model validation and MRM approaches, and the format in which the data will be received.

3 Asset Location and Characteristics

Asset Location Survey

Before looking at how vendors assess physical risk, it's helpful to understand how the location at which to assess the risk is identified. One challenge for financial firms and vendors alike is ensuring that they have solid asset location information. The degree of precision required will be more important for some hazards (e.g., flooding, soil movement) than others (e.g., heat, fire).

Many attributes are needed to be able to identify a location precisely. Since complete information might not always be available, we asked the vendors to explain what they did when they didn't receive full details of a property's location. Most vendors offer analysis at a less granular level such as a regional or county level, although the appropriateness of this will depend on the hazard. Some offer help cleaning data or working with the client to get the best possible information (e.g., by trawling the internet for company level data). One vendor stated that it offered geolocational services to their clients.

We wanted to see whether the way in which vendors locate properties could be a potential source of variation across the vendors. To test this, we provided them with information on 20 properties, across the world, with differing degrees of completeness. For example, some properties had a full postal address, but no latitude or longitude; for others we might just provide the name of a factory or house, together with the town and country. The vendors were asked to supply the longitude and latitude details for each property as a way of testing the consistency in their geocoding approaches.

A Quick Note on Geocoding

Latitude and longitude coordinates start with degrees and then narrow down to minutes and seconds to pinpoint the exact location.

- A degree is 60 nautical miles wide at the equator (nautical miles are slightly longer than the standard mile we use on land, at around 1.1 miles).
- Each degree is then broken down into 60 minutes: that is one minute of latitude at the equator corresponds to one nautical mile.
- Minutes are themselves broken down to 60 seconds. These are often given to several decimal points to be as precise as possible.

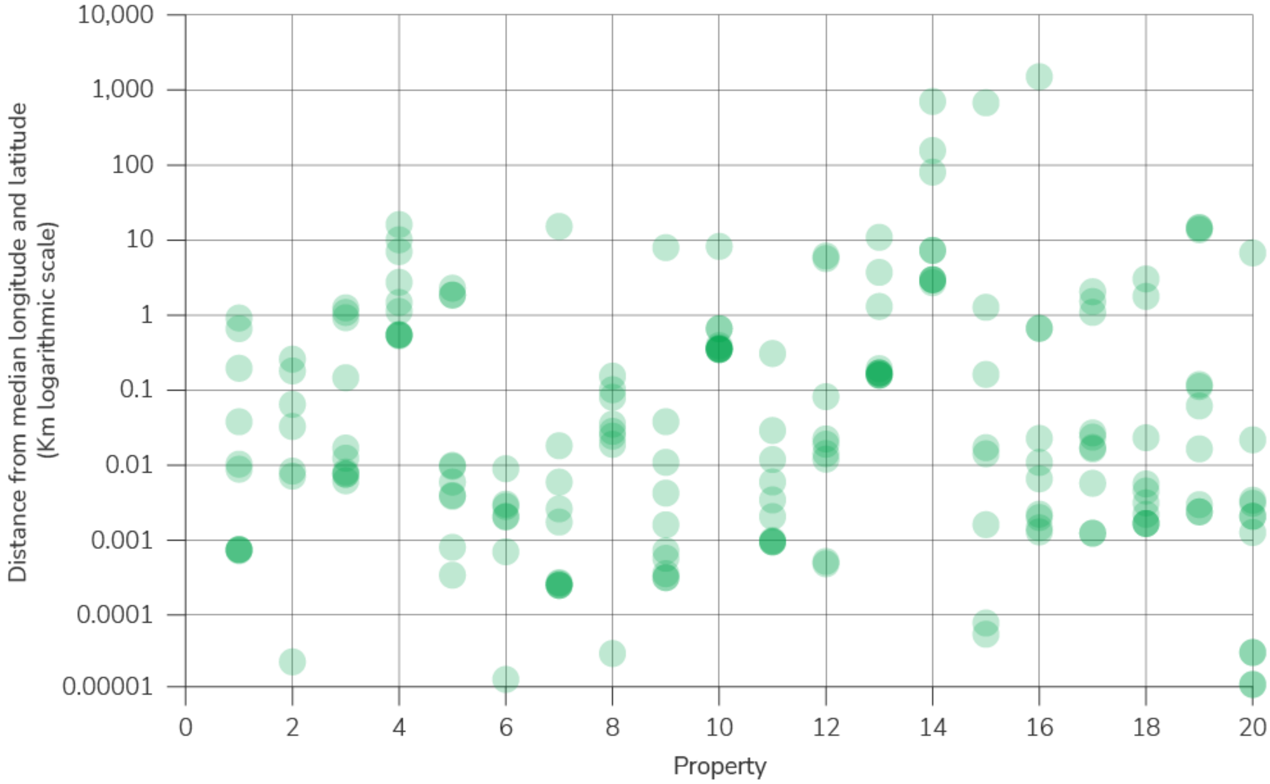
Our vendors provided their answers in decimal format, in which:

Decimal Degrees = degrees + (minutes/60) + (seconds/3600)

For example, 51.963048, 6.000372 converts to 51° 57' 46.9728", 6° 0' 1.3386"

Ten of the vendors provided location geocoding information. For each property we calculated the median longitude and latitude of the vendors' submissions. The distance (in kilometres) between each individual vendor's submission and the median value was then calculated. As Figure 8 shows, there is considerable dispersion across the vendors' geocoding (note the logarithmic scale needed because of the size of dispersion).

Figure 8 Dispersion in Vendors' Geocoding



Note: Each dot represents the distance between a vendor's geocoding and the median for all vendors' estimates. Since the dots are translucent, more observations at a given distance make the dots look darker.

There were a few extreme outliers including:

- The largest dispersion was for property 16 where one vendor submitted coordinates that were 1507 km away from the median of the others. The name of a well-known store in Boston was provided for the property, but the submission was for a similarly named road in Atlanta.
- Dispersion was particularly high for property 14 — an address in the Philippines.

Dispersion can be affected by the quality of the information provided. To explore this, we ranked the data quality from 2, where only two address items were provided (such as the first line of an address and the city; or the name of a factory and a city), to 6 (where we provided six items, such as asset name, asset type, address, city, county and postcode). Figure 9 illustrates this.

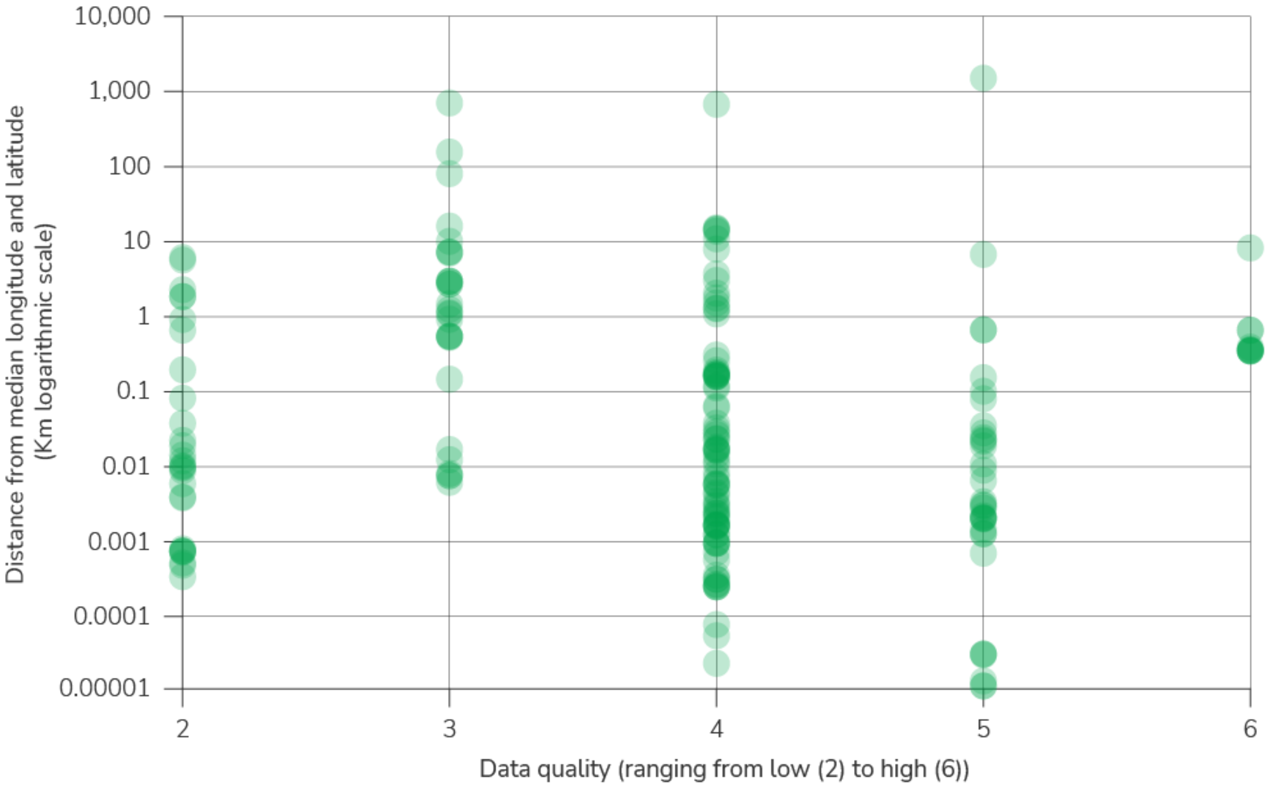
Figure 9 Measuring Quality of Address Information — Examples

Data quality	2	3	4	5	6
Asset type (e.g., factory, office, house)				X	X
Property number and street name or property name	X	X		X	X
City	X		X	X	X
County/State		X	X		X
Postcode/ZIP code		X	X	X	X
Country			X	X	X

Note: Crosses denote examples of the address line items that were provided to the vendors. Any items from the list could have been provided. For example, data quality 2 could correspond to any two items from this list.

Apart from the extreme outlier with data quality 5 (which was the Boston store noted above), Figure 10 shows that higher quality address information generally reduces the dispersion of vendors' geocoding results, with the dots becoming more clustered and dispersion falling as the data quality increases. However, even with full information (data quality 6), there was still quite a bit of dispersion. And most the geocodes submitted for properties with the least information provided (data quality 2) were close together.

Figure 10 How Quality of Address Information Affects Asset Location



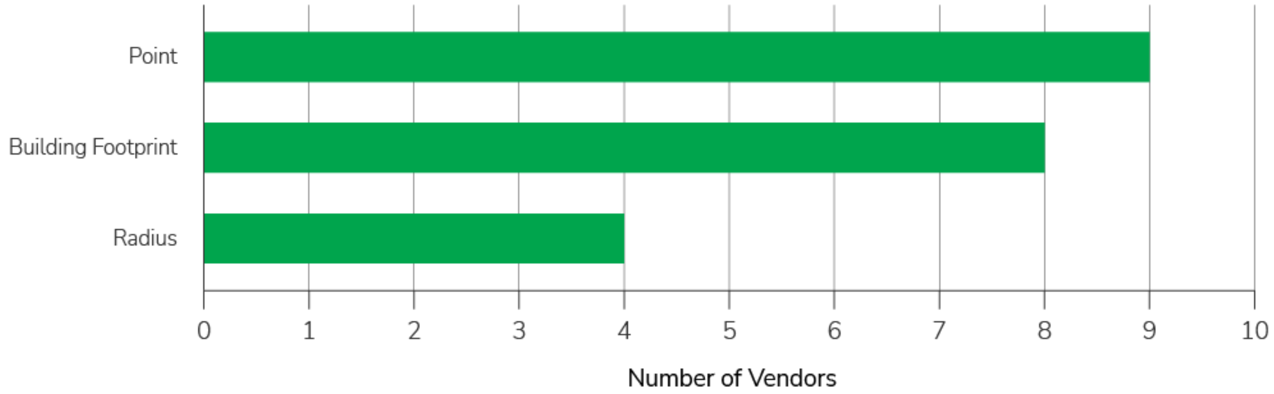
Note: Each dot represents the distance between a vendor's geocoding and the median for all vendors' estimates, at different levels of data quality. Since the dots are translucent, more observations at a given distance make the dots look darker.

On closer inspection, it seems that some types of information are more useful than others. For two "level 2 data quality" properties the dispersion was quite low. For one of these properties the vendors were given the asset name (a particular manufacturing plant), and the other was in Europe where addresses tend to be relatively straightforward to locate. And, as noted above, properties with level 5 data quality were distorted by the one highly erroneous property location (Atlanta rather than Boston).

While this was a small sample of properties, it's clear that the completeness and accuracy of asset location data provided to the vendors may have a material impact on the results.

Even if vendors have the same geocoding, there is still potential for differences in the way that they pick the point at which to assess the property's exposure to the hazard. For instance, on a sizable plot (such as a factory or shopping mall), the centre of the area might be used, or a radius, or a point that could be the most vulnerable. Figure 11 shows some different ways that vendors use to identify a property's location. Some vendors use multiple methods; others just use one.

Figure 11 How a Property's Location Is Identified



The range of options that are available include:

- Point-based analysis — a single point is placed for the property and all data is extracted for that point.
- Polygon-based analysis — a variety of comments were made about this:
    - one vendor noted that they allowed the client to draw a polygon on to a map to identify a site.
    - buffers around the point location are commonly used. Some use fixed buffer sizes while one offers multiple buffer sizes depending on the type of property and its use, (such as residential or different commercial classifications).
    - building polygons or land boundary datasets are also used where available.
- Site analysis — the whole area of a site is analysed as a polygon or mesh of points.
- Individual properties within a large site are analysed individually.
- Off-site analysis in which the area around an asset is analysed.
- Linear analysis — for long assets like roads, pipelines, or powerlines.

A firm should discuss pros and cons of the different approaches with their vendor(s), to determine which one best meets their needs.

Lesson 7. The quality of asset location information will impact the results of a physical risk assessment. Financial institutions should consider the quality of their asset location data and whether they need help from a vendor in improving data quality. They should ask vendors about the way in which they deal with incomplete data, and the method of identifying a property's location.

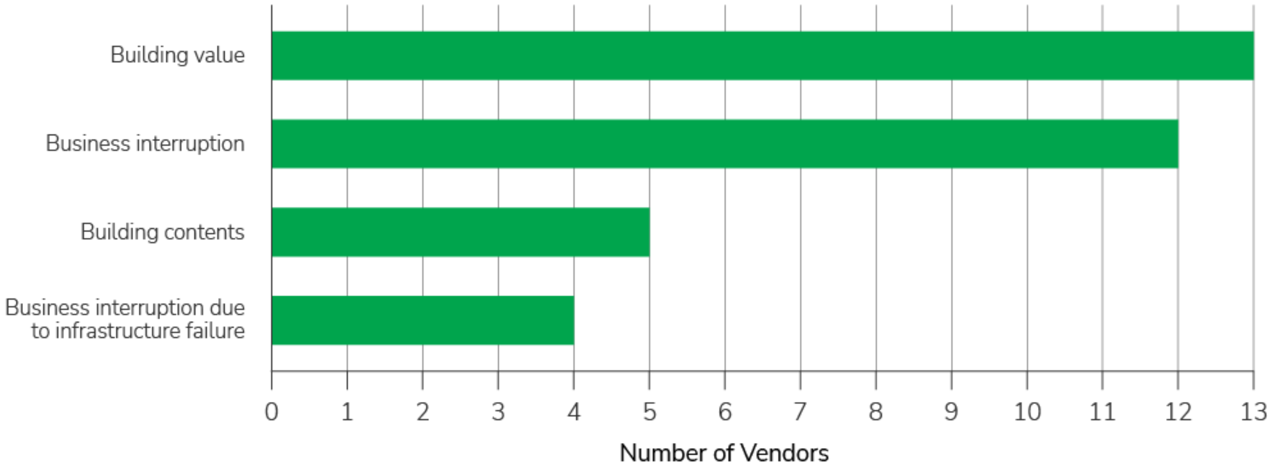
Property Characteristics

Beyond the location of a building, most vendors also take account of the physical characteristics of a property when assessing its resilience to various hazards. For example, a building made of wood is likely to be less resilient to fire than one made of concrete.

The vendors differed in how they incorporate the characteristics of buildings, such as age and construction material, when assessing a property's exposure. Many looked at the number of floors and presence of basements. Some took account of the age of the building and/or method of construction. Some vendors go into detail and analyse the different components of buildings, such as roof, walls, and foundations and assess their vulnerability to each hazard type.

Practically all of them reported looking at the asset use or occupancy type, such as whether they are commercial properties or houses. As Figure 12 shows, vendors assessed different aspects of risk for commercial properties, with building value being the most frequently assessed dimension, closely followed by business interruption. Building contents value and the cost of business interruption due to infrastructure failure are also assessed by some vendors.

Figure 12 Aspects of Commercial Property Risk Assessed by Vendors



Lesson 8. Financial institutions should be aware of the differences in vendors' approaches to incorporating the characteristics and uses of a building into their physical risk assessments. Some vendors have this data, and some rely on their customers providing it. Financial institutions that can capture data about property characteristics should be able to increase the quality of a physical risk assessment, if their vendor can process the data.

4 Physical Risk Quantification Study

Physical Risk Quantification Overview

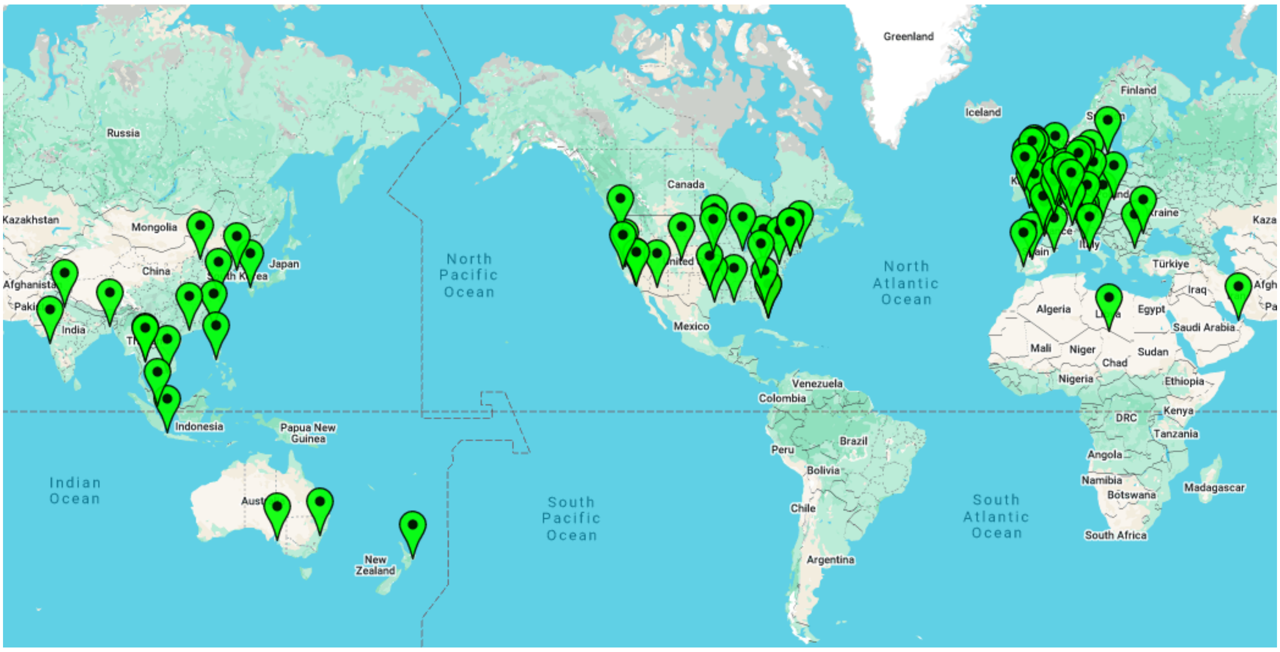
To assess how physical risk assessments vary we asked the vendors to assess a range of hazards for 100 properties located in the U.K, mainland Europe, Asia and the U.S. Figure 13 shows the information on each property that was provided as part of the benchmarking study.

Figure 13 Property Information Provided to Vendors

Asset type – for example, house, factory
Property number and street name or property name
City
County/State
Postcode/ZIP code
Country
Latitude
Longitude

Figure 14 is a map showing the property locations, worldwide and for the U.K.

Figure 14 Property Locations – Worldwide and in the U.K.



Note: The maps in Figure 14 were constructed using MAP: My First Map.

The negative sign was inadvertently left off the longitude for property 12, which placed it in the North Sea, rather than in Scotland. Some vendors identified this and asked us how they should treat it; some provided null returns as the property is in the sea; some provided information about hazards in the North Sea; and some used the address information to place it in the correct position on land. This exemplifies the variety of potential treatments for properties with incomplete or inaccurate address information.

Figure 15 gives details about the regional spread of the properties and the hazards that we asked vendors to assess in the different regions. For example, we chose not to assess tropical cyclones in the U.K. as this peril does not occur there. Both commercial and residential properties were provided.

Figure 15 Regional Spread of Properties and Hazards

Region	Number of properties	Coastal Flooding	Combined Flooding	Tropical Cyclone or Hurricane	Windstorm	Wildfire	Heat Temperature Increase
U.K.	30	X	X		X	X	X
Mainland Europe	30	X	X		X	X	X
U.S.	20	X	X	X		X	X
Asia	20	X	X	X		X	X
Total	100						

Combined flooding refers to pluvial and fluvial flooding, where:

- Pluvial flooding is caused by heavy rainfall that overwhelms urban drainage systems or the ground's capacity to absorb water, leading to surface water flooding.
- Fluvial flooding stems from rivers and streams that overflow, typically because of excessive

rainfall or rapid snowmelt.

By asking for data on combined flooding, we are not distinguishing between the two sources. Vendors were asked to provide us with their view on the damages from a combination of these types of events, supplying us with estimates of the worst case. (Note that this decision was made to simplify the analysis.)

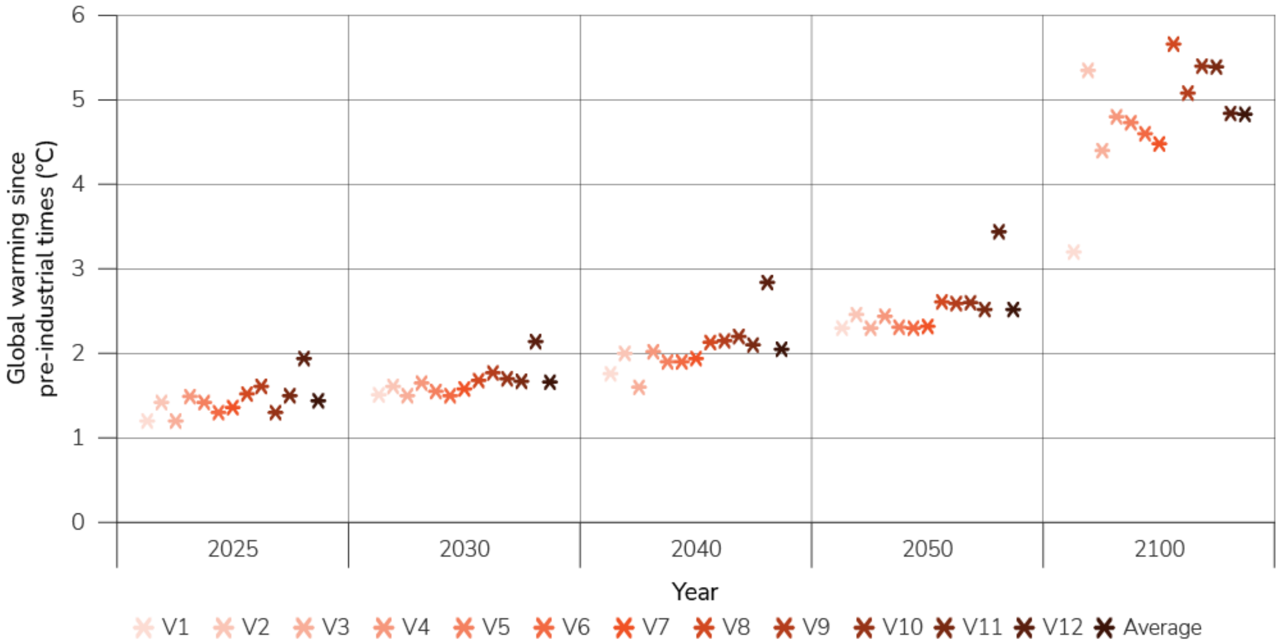
Because of the potential for flood defences to make a considerable difference to the calculation of physical risk affecting properties, we asked the vendors to assess both coastal and combined flood risk on two bases — defended and undefended.

Choice of Scenario, Time Periods and Severity

We chose to use just one climate scenario for this exercise, namely RCP 8.5. This was simply to make the exercise more tractable and eliminate an obvious source of variation. We asked vendors to provide us with the estimate of global warming, from the climate models they use, for RCP 8.5 for various periods out to 2100. This was to act as a check on assumptions that were underlying the risk assessments.

Figure 16 shows that even for 2025 estimates, there is some dispersion across the vendors' assumptions, with a 0.7°C difference between the highest and lowest. By 2100, there is considerable dispersion, with vendors assuming between 3.2°C and 5.7°C of warming.

Figure 16 Degrees of Warming Assumed by Vendors in RCP 8.5 (High Emissions Scenario)



When we investigated further, there was a variety of reasons for the dispersion of estimates of warming under this one scenario. For instance, one vendor's modelling platform did not use RCP 8.5 and so their 2100 figure was derived from the NGFS Current Policies scenario (providing us with a 90th percentile temperature, adjusted to align with baseline warming, and extending to 2080 rather than 2100). Also, there was variation in the baseline used: most vendors reported warming relative to the average temperature during pre-industrial years (1850 to 1900), but some initially reported it relative to a different baseline, such as 1986-2005. While these discrepancies do not reflect differences in underlying models per se, they do indicate that if a financial institution asks different vendors for an RCP 8.5 projection, one driver of differences could be that different levels of warming are assumed.

Lesson 9. Financial institutions should request scenarios which meet their purpose and check which underlying assumptions vendors are using about key variables of interest, such as global warming.

We asked the vendors to assess a range of measures at different time horizons and for different degrees of severity (or "return periods"), as shown in Figure 17. For example, a return period of 20 corresponds to a 1-in-20-year risk, equivalent to a risk at the 95th percentile; a return period of 1000 is far more severe, corresponding to a 99.9th percentile.

Figure 17 Information Requested for Different Return Periods

	Average Annual Damage	Return Period RP 20 (95th percentile)	Return Period RP 200 (99.5th percentile)	Return Period RP 1000 (99.9th percentile)	Maximum Value
2025	x	x	x	x	x
2030	x	x	x	x	x
2040	x	x	x	x	x
2050	x	x	x	x	x
2100	x	x	x	x	x

Choice of Metrics to Measure Hazards

This section outlines the metrics that were chosen for the hazards focused on in this study. As we saw in Section 2, vendors use a wide range of metrics to measure physical risk. For this exercise, it was important to standardise those as much as possible to allow for a valid comparison across vendors. For each hazard, we asked for a measurement of the severity of the hazard itself, as well as a measure of the damage it would cause on the selected property (Figure 18).

Figure 18 Metrics Used to Measure Hazards and Financial Impacts

	Flood	Tropical Cyclone/Hurricane and Windstorms	Wildfire	Heat Temperature Increases
Measure of hazard	Depth in meters	Wind Speed in km per hour	Fire Weather Index	Annual Days Above 35°C
			Count or Probability-based Measure*	
Measure of damage	Damage Ratio for Annual Maximum Loss (%)			Productivity Loss Due to Heat
	Average Annual Damage (in USD)**			
Measure of uncertainty	Standard Deviation of Annual Damage			

* Completed if a count or probability-based measure of wildfire is used
** Also referred to as Average Annual Loss or Expected Annual Loss

5 Physical Risk Benchmarking Results

In this section, we present the results of the quantification survey, starting with how the risk assessments vary by hazard. To start with, we fix the time period at 2030 and return period at one-in-200 year (99.5th percentile) and look at the dispersion of the projection of the different hazards and associated damages.[5]

For all perils, the number of vendors whose hazard data is shown may not be the same as for the damage ratios, as some vendors could provide one or the other and some could provide both.

Physical Risk Assessment Dispersion: RP 200 in 2030

We start with flooding, which is the peril for which we received the most comprehensive data from the vendors and which the most vendors could provide data for.

Combined Flooding (RP200 in 2030)

Combined flooding is both pluvial and fluvial flooding. Some vendors model these together, and some model them separately and gave us the maximum of the two. Some vendors also modelled a combination of pluvial, fluvial and coastal flooding.

It should be noted that adaptation and mitigation measures for pluvial and fluvial flooding may be quite different, which is a limitation of this approach. For example, fluvial flooding could be reduced by using barriers along the river itself, while pluvial flooding will depend heavily on drainage and building-level defences. Also, vendors' differing assumptions about defences and their efficacy could contribute to the variation in the outputs.

Flooding was estimated to affect practically all the properties. However, there was considerable dispersion in estimates of both flood depth and damage ratio. The maximum, minimum and mean of the vendors' estimates of flood depth — when flood defences are in place in suitable locations — are shown in Figure 19. Dispersion was similar for undefended flooding.

[5] Note that dispersion measures in this section show the range of estimates *across* vendors; uncertainty measures discussed in Section 2 are provided by a single vendor as estimates of uncertainty for their own modelled outputs.

Figure 19 Defended Flooding Depths Across Properties

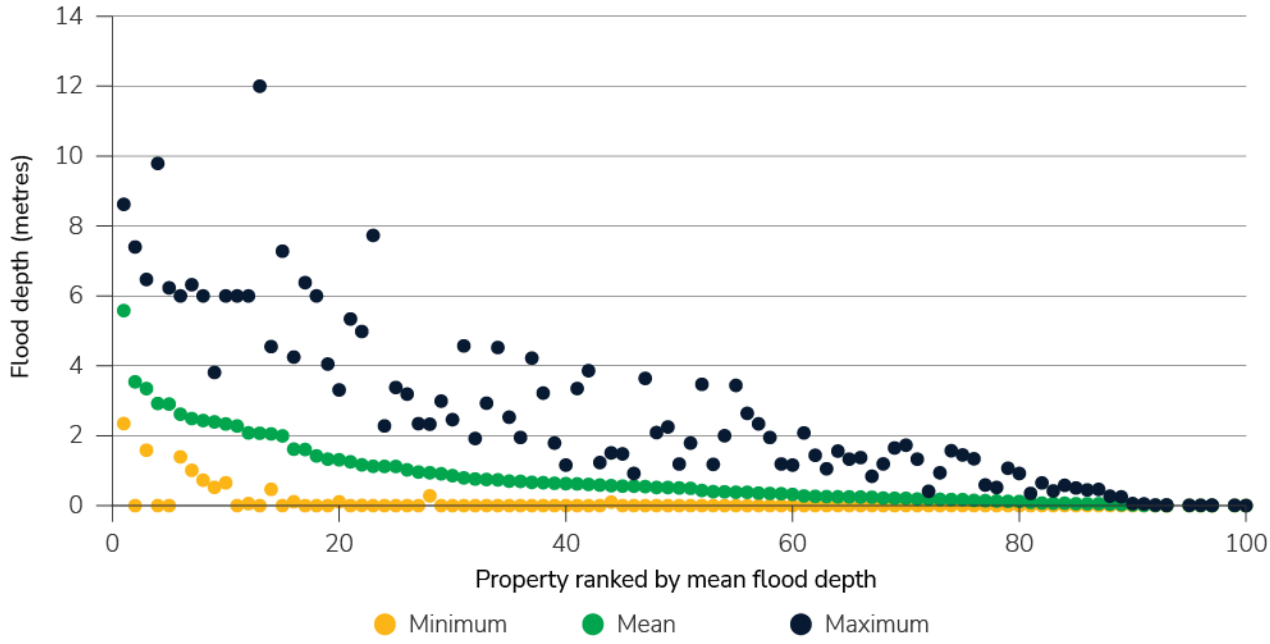
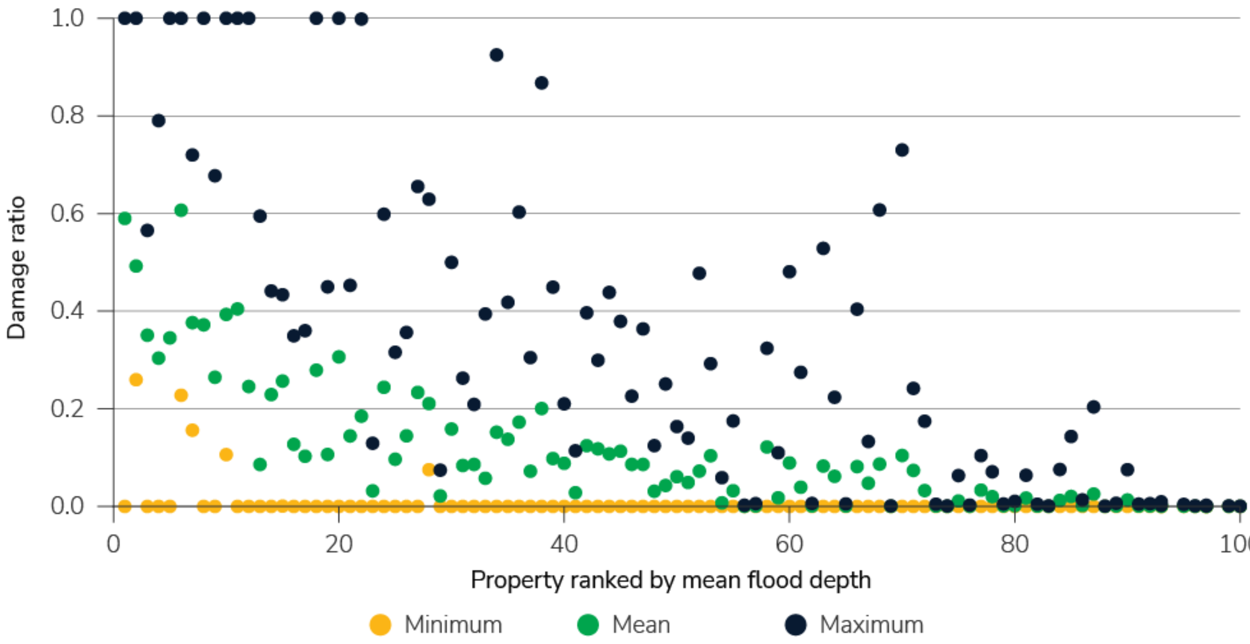


Figure 20 shows the corresponding distribution of damage ratios across these properties.

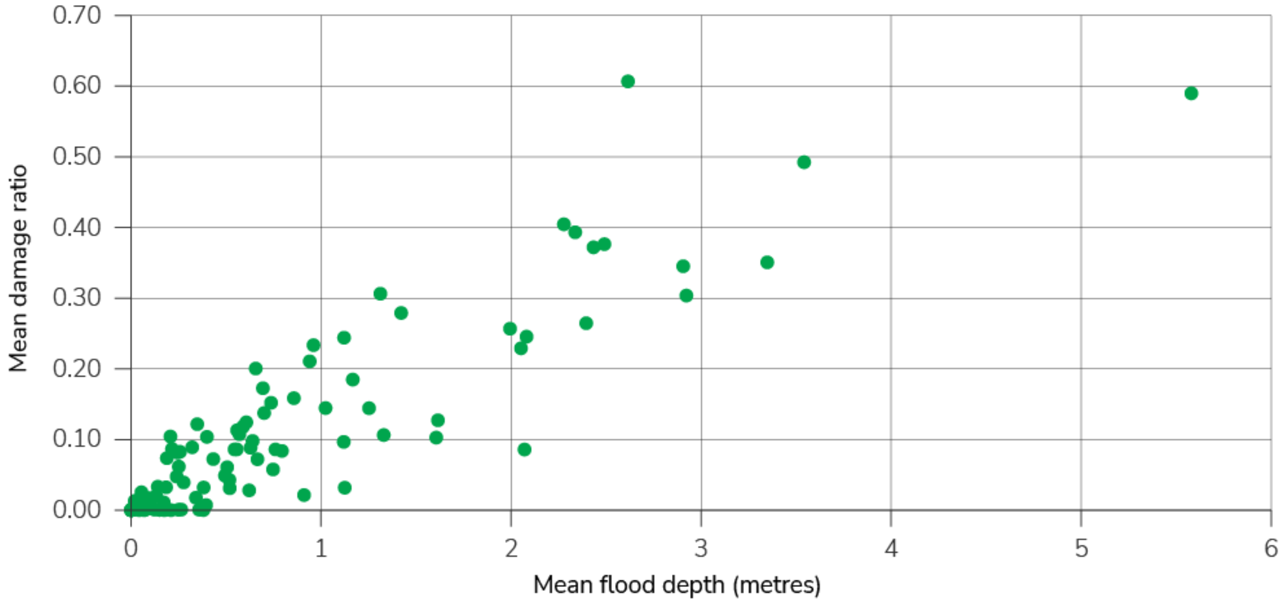
Figure 20 Defended Flooding Damage Ratios Across Properties



Note: Figure 19 shows dispersion of flood depth estimates for defended combined flooding for RP200 in 2030. Properties were ordered by the mean, so that property 1 experienced the most flooding according to the mean of the vendor estimates for a 1-in-200-year flood in 2030. Figure 20 shows the corresponding damage ratios for properties in the same order.

For most vendors, the deeper the flooding, the higher the damage ratio for the property in question, as illustrated in Figure 21 which shows the mean of the vendors' estimates for depth and damage, for defended flooding.

Figure 21 Relationship Between Mean Flood Depth and Damage Ratio



However, vendors do vary considerably in their views on the amount of damage that a given level of flooding will have. The correlation between different vendors damage ratio projections ranged from 0.2 to 0.9, as shown in Figure 22.

Figure 22 Correlation Matrix of Vendors' Damage Ratio

	v1	v2	v3	v4	v5	v6	v7	v8
v1	1	0.20	0.52	0.53	0.32	0.46	0.56	0.53
v2		1	0.72	0.32	0.23	0.45	0.38	0.36
v3			1	0.31	0.46	0.32	0.63	0.73
v4				1	0.38	0.42	0.89	0.66
v5					1	0.28	0.37	0.37
v6						1	0.53	0.42
v7							1	0.76
v8								1

Note: This table shows the correlation between eight of the vendors' projections of damage ratios for the 100 properties in our portfolio for a one-in-200-year, defended, combined pluvial and fluvial flood in 2030.

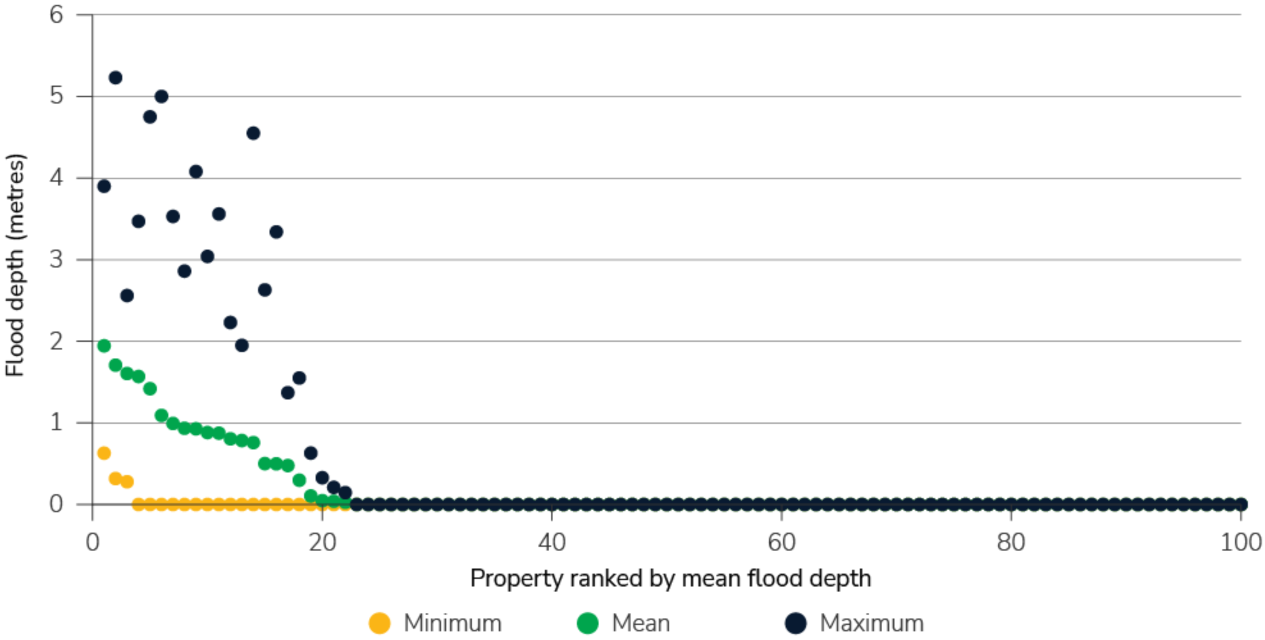
This variation might reflect differences in the vendors' insights about the severity of the hazard, the property's exposure to it, or the vulnerability of a property (e.g., they might be accounting for different attributes of buildings, or they may have different views about the elevation, age or value of the properties). It is yet another dimension of the variation across results which makes it hard for financial institutions to compare vendors. In practice, strong communication with vendors can help improve the analysis.

Coastal Flooding (RP200 in 2030)

Vendors had different definitions for flood risk: for example, some measured rising sea levels, some measured rising sea levels and storm-surges, and some incorporated astronomical tides. Whereas most vendors can calculate coastal flooding for various return periods, one vendor treated it as a chronic hazard (effectively giving it a return period of one in one).

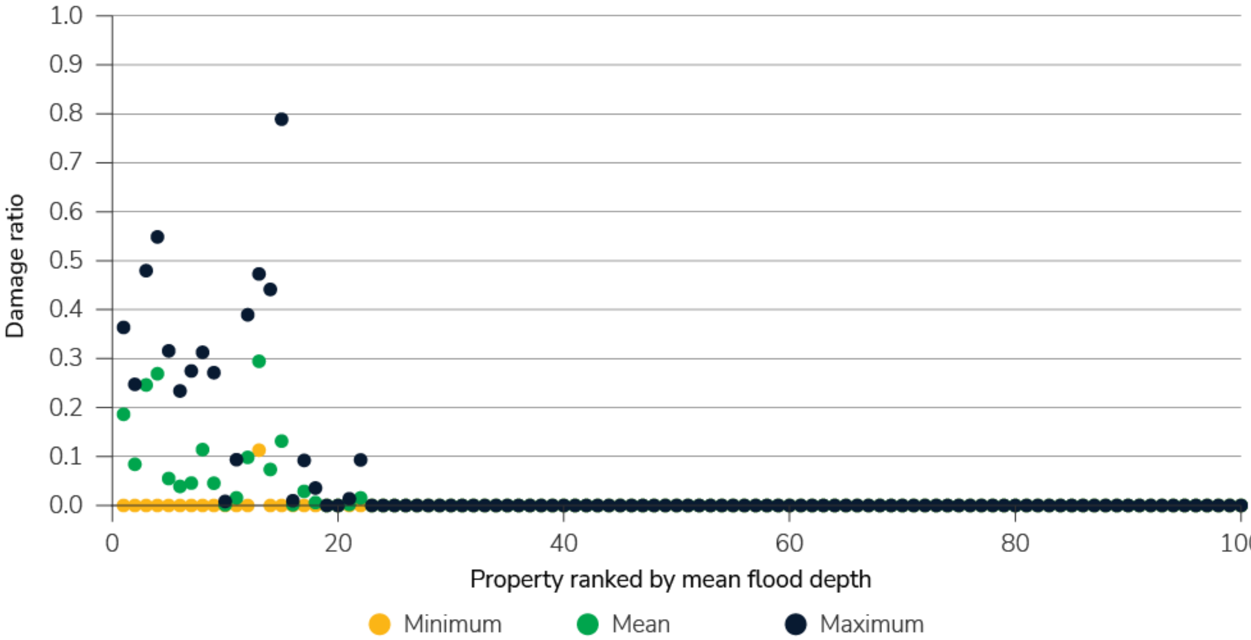
Only 22 of the sample of 100 properties were affected by coastal flooding, after taking account of flood defences, as Figure 23 shows. The pattern of dispersion is similar to combined flooding, but with a much smaller flood depth.

Figure 23 Defended Coastal Flooding Depths Across Properties



As with combined flooding, there is also similar dispersion across estimates of damage as can be seen in Figure 24.

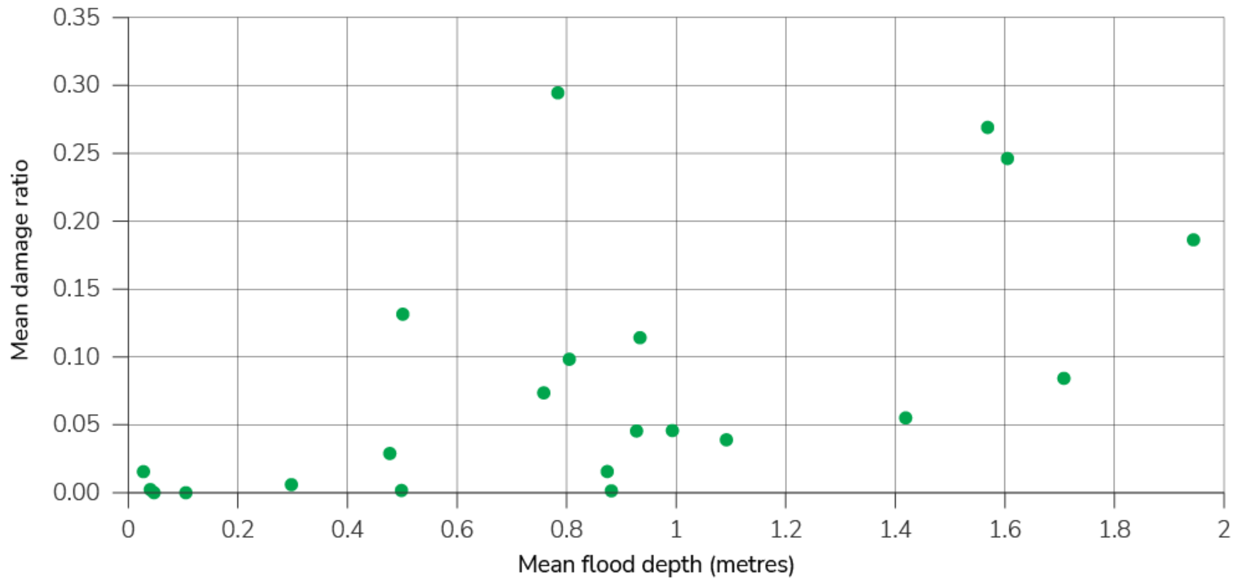
Figure 24 Defended Coastal Flooding Damage Ratios



Note: Figure 23 shows dispersion of estimates for defended coastal flooding for RP200 in 2030. Properties were ordered by the mean, so that property one experienced the most flooding according to the mean of the vendors for a one-in-200-year flood in 2030. Figure 24 shows the corresponding damage ratios for properties in the same order.

Just focusing on the 22 properties affected by coastal flooding, Figure 25 shows some relationship between flood depth and damage though a less pronounced relationship than for combined flooding, perhaps indicative of better flood defences for these properties.

Figure 25 Relationship Between Coastal Flood Depth and Property Damage



Note: Figure 25 shows the relationship between average coastal defended flood depth and average damage ratio.

Cyclone (RP200 in 2030)

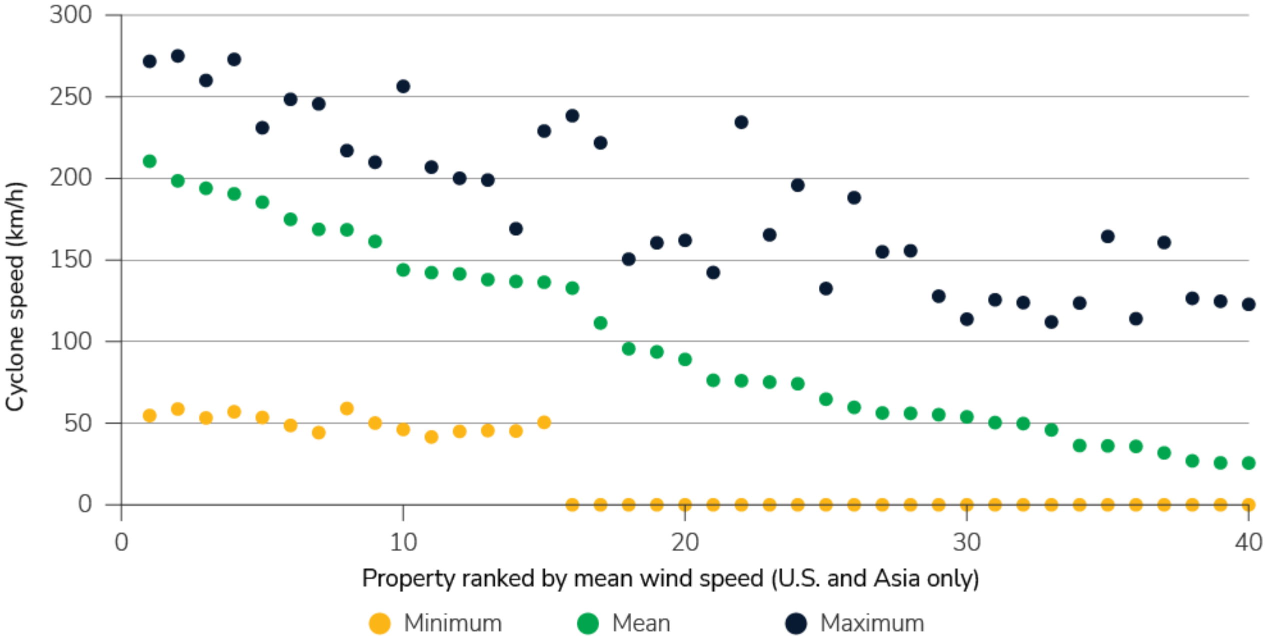
A cyclone is a large low-pressure system. The most severe type is called a [tropical cyclone](), which is referred to differently in different regions:

- Tropical cyclones above a certain windspeed that originate in the West (mainly over the Atlantic Ocean and Gulf of Mexico) are called hurricanes.
- Tropical cyclones above a certain windspeed that originate in the East (mostly over the western Pacific and northern Indian Ocean) are called typhoons.

In this study, we refer to all of these as simply cyclones and focus on their effect in the U.S. and Asia. (In the next section we look at how windstorms affect properties in the U.K. and Europe.)

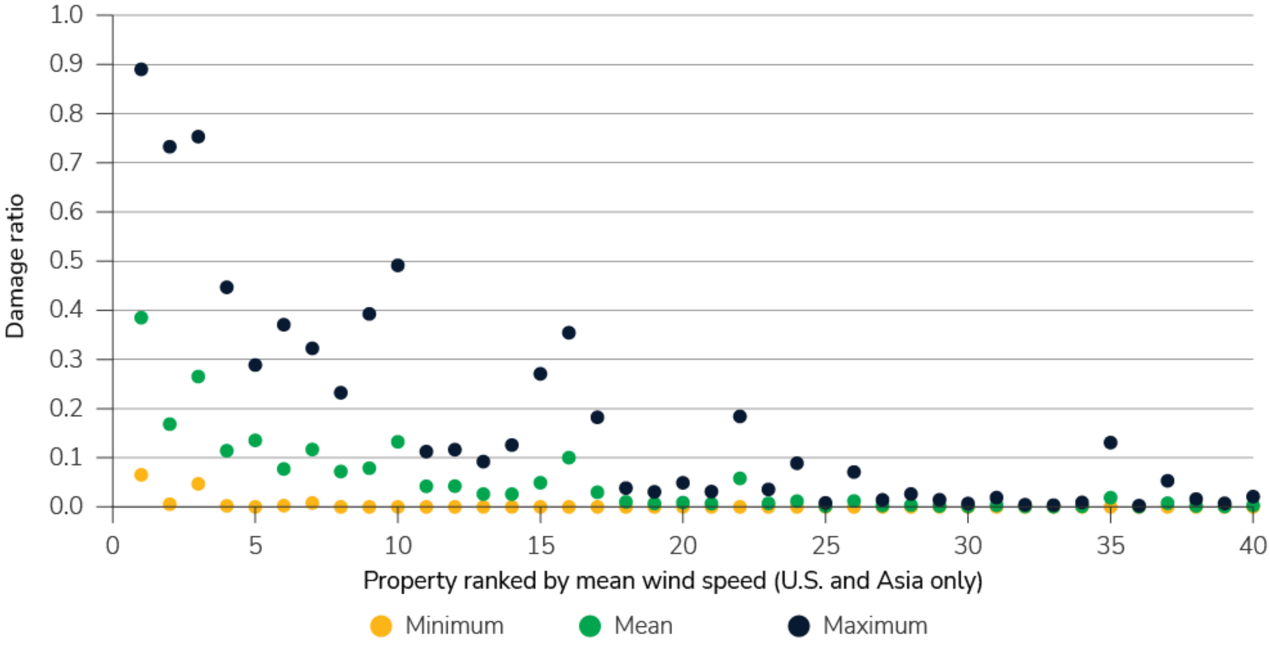
For cyclones, the hazard is typically measured by windspeed. Most vendors expressed this as km per hour. However, a few different things were being measured, such as km per hour for a 3-second gust, or an average windspeed over 1 or 10 minutes. Figure 26, which shows cyclone speeds, therefore contains a mix of gust and sustained wind speeds for RP 200 in 2030. (Although this does mix up concepts, the sample was too small to look at the same metric across vendors.) As with flooding, there was considerable dispersion in the projected hazard from cyclones, for the 40 properties in the U.S. and Asia.

Figure 26 Cyclone Speeds Across Properties



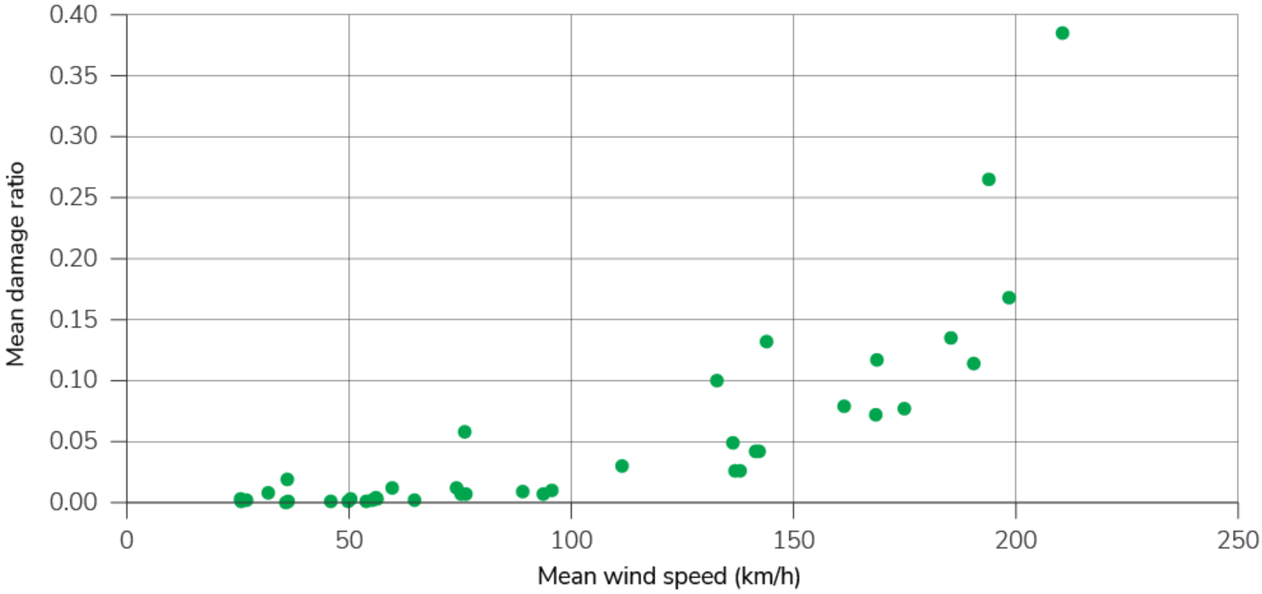
The damage ratios are also dispersed (Figure 27).

Figure 27 Cyclone Damage Ratios Across Properties



Not surprisingly, the relationship between the hazard and property damage shows a more non-linear relationship, with little damage below a certain threshold, but then rising increasingly sharply as wind speeds pick up. This can be seen clearly in Figure 28, which shows the mean damage ratio versus mean wind speed.

Figure 28 Relationship Between Cyclone Speed and Property Damage

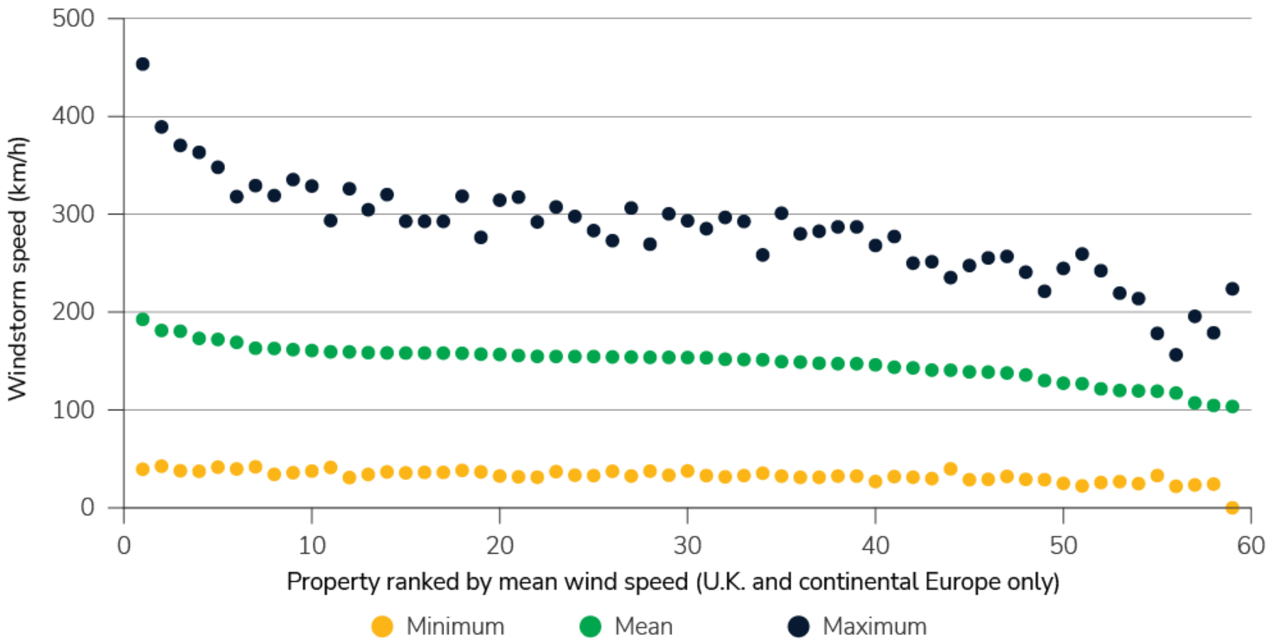


Windstorm (RP200 in 2030)

For the U.K. and continental European properties, we looked at windstorms, rather than tropical cyclones. Windstorms are winds that are strong enough to cause at least light damage to trees and buildings and may or may not be accompanied by precipitation. Wind speeds during a windstorm typically exceed 55 km per hour.

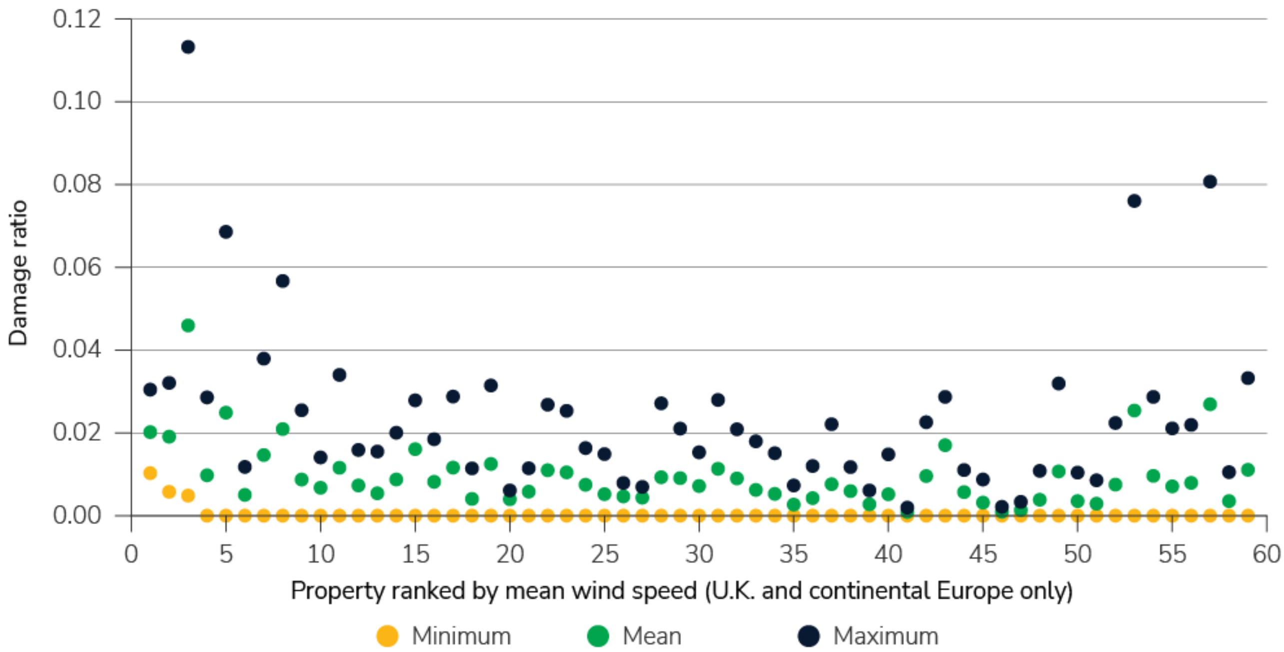
Windstorms are usually measured by their windspeed in km per hour. In this study six vendors used that measure, generally measuring the speed of a three-second gust. Others used a variety of measures which had to be excluded as they were not comparable (e.g., counting the annual number of days with peak wind gusts exceeding a certain threshold). As with the other hazards, there was considerable dispersion across the vendors (Figure 29).

Figure 29 Wind Speeds for U.K. and European Properties



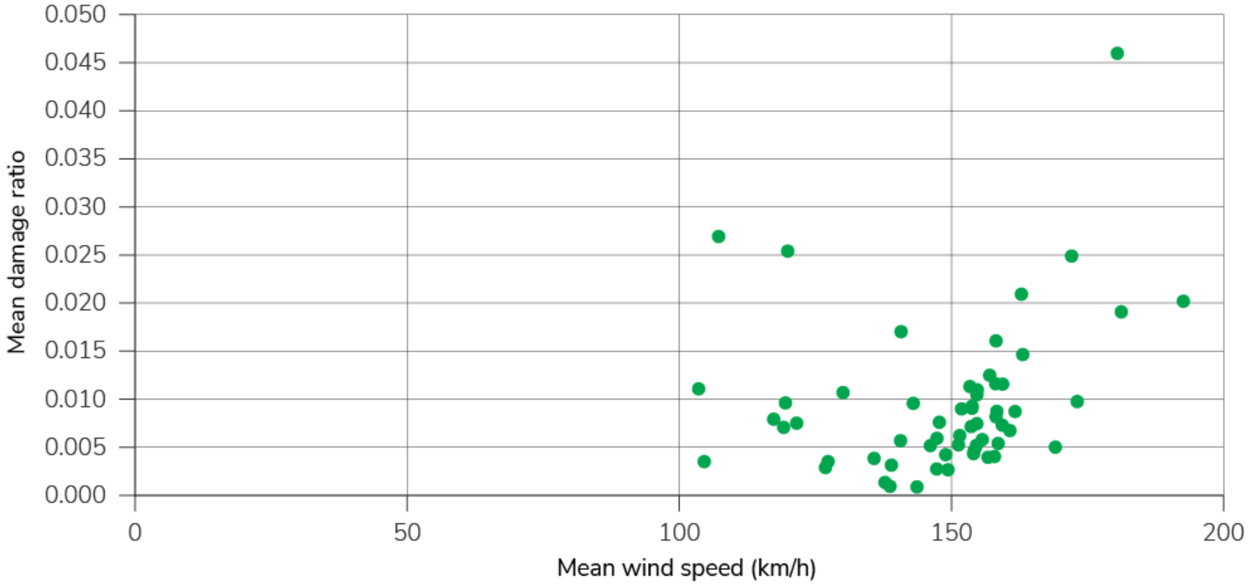
The damage ratios are also dispersed (Figure 30).

Figure 30 Windstorms Damage Ratios for U.K. and European Properties



As with cyclones, there is a non-linear relationship between wind speed and property damage, although much less pronounced and with the level of damage much lower than for cyclones (not surprisingly given the lower wind speeds involved) (Figure 31).

Figure 31 Relationship Between Windstorm Speed and Property Damage



Heat (RP200 in 2030)

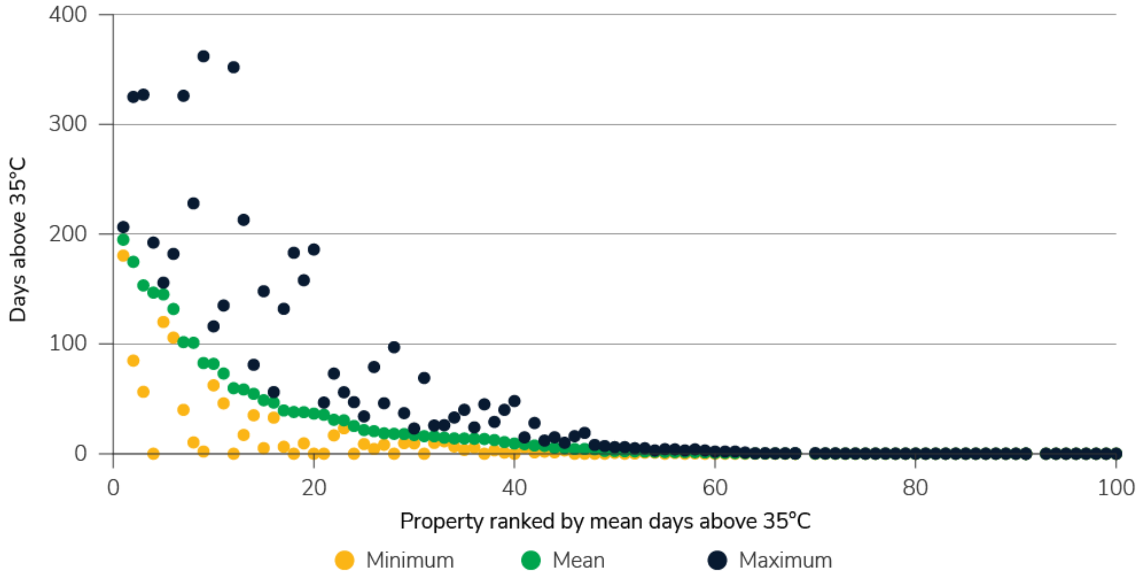
Heat is different to the other hazards that we have looked at because the reference measure/definition of heat can be very location specific. For example, it can be measured by the number of days the temperature breaches a certain relative threshold (such as the 99th percentile historical temperature for a location). Since the underlying temperature distributions differ according to location, what counts as heat, and as dangerous, may also differ according to location. In other words, a very warm day in London would barely be noticed in Dubai, which is a lot hotter than London all year round. Consequently, local temperature distributions are often used to assess heat and its impacts.

How does heat affect buildings? Some vendors examine the probability of failure of elements of a building structure (e.g., electricity outages). Heat also contributes to air conditioning expenses. But the impacts are less straightforward to measure than for other hazards. For instance, it is likely to affect work productivity, but this is outside the scope of this exercise. (Note that extreme cold days can also have economic impacts, but we chose to concentrate on heat for this exercise.)

Vendors were asked to measure this hazard by the number of days above 35^{o}C (Figure 32). Seven vendors provided data on this basis, two used local percentile-based measures, and one provided an annualised damage rate only.

Interestingly, for some properties there is enormous dispersion — for example, property four, which is in Hong Kong, where the minimum is zero days and the maximum is 192. On closer inspection, the minimum corresponds to an assessment by one vendor that this particular property is not exposed to this hazard. In other words, the zero corresponds more to zero exposure, rather than indicating that there are no days where the temperature is above 35^{o}C. This is a typical example of the type of issues that arise when comparing data across vendors.

Figure 32 Days Above 35°C Across Properties



Wildfire

Wildfire proved to be the most challenging hazard to get comparable metrics across the vendors. As Figure 18 showed, we had hoped to narrow the metrics to two possibilities, namely a fire weather index or a count/probability measure (such as an annual fire probability). In the end, we had a variety of measures, including fire weather indices (FWI), counts of days exceeding different FWI thresholds, probabilities of wildfire, and flame length. Given this variety, we were not confident that we could draw any meaningful benchmark conclusions.

Hazards and Dispersion by Year, Return Period, and Region

Having focused on RP200 for 2030, in this section we look more broadly at other patterns in the vendors' submissions. First, how does hazard measurement vary across different time periods and return periods?

Figure 33 shows the mean levels of undefended flooding predicted across vendors for different return periods in 2030. Focusing on undefended flooding eliminates any differences that might arise from variation in the way that vendors model flood defences. Lower flooding is predicted for RP20 than RP200, and RP200 flooding is expected to be lower than RP1000 flooding, which aligns with expectations that flooding is much higher in more extreme return periods. (Note, this figure shows higher flood depth than Figure 21 due to different vendors providing defended and undefended flood data.)

Figure 33 How (Undefended) Flooding in 2030 Changes by Return Period

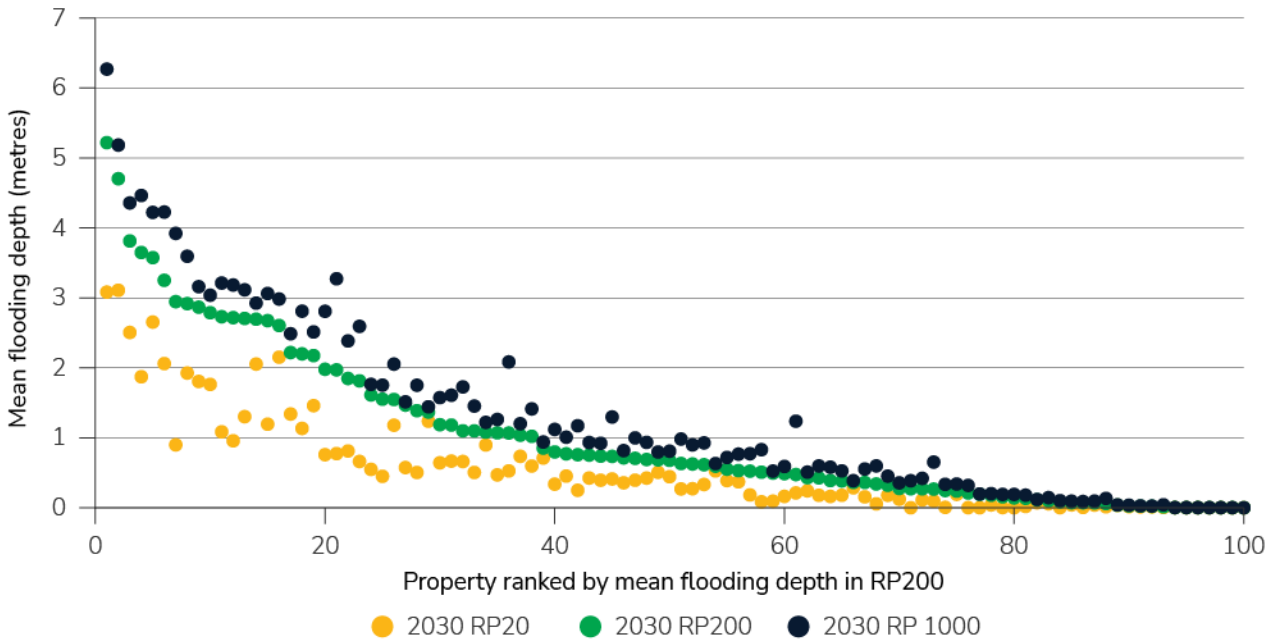
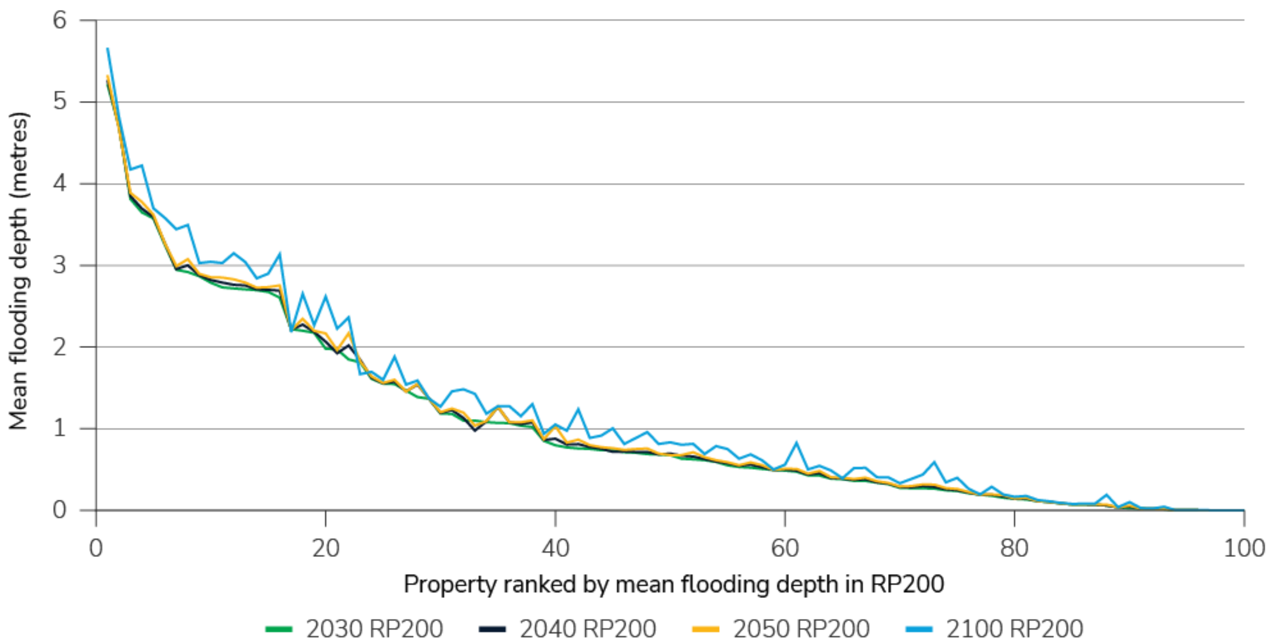


Figure 34 shows how much flooding we might expect in a one-in-200-year flood in different decades, again without flood defences.

Figure 34 How a One-in-200-Year Flood (RP200) Changes by Decade



Note: This figure is a line chart rather than a scatter plot simply for clarity, given how close the observations are.

By decade, for a constant return period, the increase in flooding depth is quite small out to 2050 but then jumps considerably by 2100.

We wanted to extend this analysis over more time horizons and return periods. Our working hypothesis was that dispersion would increase the further into the future we went, and the further into the tails of the distributions, reflecting increasing levels of uncertainty. To understand whether that happened, we need a measure of dispersion. One measure is a simple standard deviation of the vendors' submissions. On this measure, for flooding, dispersion generally increases as we look further into the future, for a given return period. It increases more strongly as the return period increases for a given year (i.e., the flooding becomes more severe as we move further into the tail of the hazard distribution). And, by region, dispersion is the lowest for properties in the U.K., followed by mainland Europe, U.S., and Asia (Figure 35).

Figure 35 Dispersion of Flood Depth Results (as Measured by Average Standard Deviation)

		UK	Europe	US	Asia	Portfolio
RP20	2025	0.49	0.52	0.57	0.79	0.57
	2030	0.47	0.51	0.54	0.78	0.56
	2040	0.50	0.54	0.57	0.83	0.59
	2050	0.48	0.53	0.55	0.84	0.58
	2100	0.60	0.62	0.61	1.03	0.69
RP200	2025	0.71	0.75	0.84	1.26	0.86
	2030	0.72	0.78	0.83	1.23	0.86
	2040	0.78	0.84	0.88	1.29	0.92
	2050	0.75	0.82	0.88	1.32	0.91
	2100	0.84	0.86	0.97	1.43	0.99
RP1000	2025	0.91	0.92	1.24	1.57	1.11
	2030	0.86	0.94	1.21	1.49	1.08
	2040	0.94	0.99	1.26	1.58	1.15
	2050	0.88	0.97	1.23	1.60	1.12
	2100	0.99	1.03	1.34	1.72	1.22

Note: This table shows the average standard deviation of estimates of defended flood depth across different return periods (RP), for a range of time horizons and across regions.

However, the standard deviation is not a good measure of dispersion when series have significantly different averages. For this reason, we also looked at the coefficient of variation, which is the standard deviation divided by the mean. This is a standardised, unitless, measure that allows for better comparisons across series with different means.

And curiously, dispersion on this measure generally behaves in the opposite way, *decreasing* as we look further into the future. For example, dispersion of flooding estimates (as measured by the coefficient of variation) tend to decrease as we look further into the future, for a given return period. They also decrease as the return period increases for a given year (i.e., the flooding becomes more severe). And by region, the relative dispersion is different, with the highest for properties generally in the U.K., followed by mainland Europe and Asia, with the U.S. the least dispersed (Figure 36).

How do we interpret these results? It basically reflects the fact that the mean of the distribution is increasing more rapidly than the standard deviation. This may simply indicate that all vendors are anticipating rising physical risks on average, driven by climate change.

Figure 36 Dispersion of Flood Depth Results (as Measured by Coefficient of Variation)

		UK	Europe	US	Asia	Portfolio
RP20	2025	2.12	2.09	1.94	1.99	2.05
	2030	2.21	2.23	2.03	2.08	2.15
	2040	2.05	2.09	1.90	1.96	2.01
	2050	2.15	2.18	1.98	2.04	2.10
	2100	2.03	2.17	1.87	1.91	2.02
RP200	2025	1.62	1.65	1.40	1.66	1.60
	2030	1.67	1.73	1.47	1.71	1.66
	2040	1.62	1.65	1.39	1.64	1.59
	2050	1.64	1.68	1.47	1.70	1.63
	2100	1.62	1.67	1.45	1.68	1.61
RP1000	2025	1.47	1.48	1.25	1.48	1.43
	2030	1.50	1.55	1.31	1.52	1.48
	2040	1.48	1.49	1.25	1.46	1.43
	2050	1.50	1.52	1.31	1.49	1.47
	2100	1.47	1.49	1.34	1.41	1.44

Note: This table shows the average coefficient of variation of estimates of defended flood depth across different return periods (RP), for a range of time horizons and across regions.

Rankings and Temperature Increase

Finally, we wanted to investigate how the differences in estimates of global warming (as shown in Figure 16) influenced the results. In other words, was the dispersion in estimates of flooding or heat simply a reflection of a difference in assumptions about how much global warming there was?

To do this, we looked at the rank of the vendors' estimates, and whether this was correlated with their estimates of global heating. For example, was the vendor that consistently thought there was going to be more global warming per decade more inclined to forecast higher levels of physical risk?

For flooding, higher temperature assumptions did not seem to be driving the differences in ranking, suggesting that other issues were more of a consideration (such as modelling and methodology).

For windstorm and cyclone, the rankings of vendors were less variable across properties, but again not strongly correlated with their assumptions about global warming.

6 Checklist for Financial Institutions

There is no single best vendor or model to use for physical risk assessments. But, drawing on the lessons raised through this document, we have drawn together a checklist of questions financial institutions should consider — both for themselves and for the vendors — when undertaking this type of analysis.

Checklist of Questions

1. **Use Case**
   - What is your particular use case for physical risk assessments? For example, is it concerned with risk identification, valuation of adaptation measures, vulnerability assessments for clients?
   - Do you need to understand a hazard (for example, flooding) in detail at a particular site or to understand flood risk at portfolio level and how this might change over time?
   - If a very detailed risk assessment is needed for a high value location, this might need an assessment that is carried out on-site. The vendor data could be used as an initial risk assessment that leads to more detailed analysis.

2. **Perils and Hazards Coverage**
   - Does the vendor cover the geographies that you need?
   - Given the geographic spread of your portfolio, does the vendor have the capability to model the key perils that will affect those regions?

3. **Asset Data Quality**
   - What is the quality of your asset data (such as building location and characteristics)? Will you need help improving its quality?
   - How does the vendor undertake its geocoding?
   - Do you understand the way in which vendors deal with incomplete data, and how they identify a property's location?
   - Does the vendor treat a property as a point, a predefined shape, or does it use the footprint of the property?
   - How does the vendor incorporate information about the characteristics of a property (e.g., height, building material) into their physical risk assessments? Do you have access to this type of data?
   - How does the vendor specify the area which is to be assessed? Do they use a point, a predefined shape around a point, the footprint of the property, or some other approach? How do they assess the risk affecting a large building or factory site?

4. **Scenarios**
   - Which scenarios do you want to be able to model?
   - Do you want the vendor to be able to produce their own scenarios?
   - What severity of impact are you interested in?
   - What time periods do you need the modelling for? (E.g., are 5-year intervals out to 2100 required, or is less frequent or a shorter timescale suitable for your needs?)

5. **Granularity of Portfolio and Resolution of Modelling Required**
    - What granularity of modelled output do you need? For example, do you need modelling at the individual asset level, or will portfolio-level information suffice?
        - The size of your portfolio may impact your requirements for granularity. For example, larger portfolios might need less granular modelling, although you might want more in-depth modelling for hotspots.
        - The resolution of modelling might also vary by peril. For example, flood results would need a higher modelling resolution, in contrast to that needed for analysis of the impact of heat. The usefulness of granular hazard information also depends on how granular asset information is.

6. **Methodology and Outputs**
    - Which metrics for hazards and financial losses are most suitable for your firm's use case?
    - Do you require a particular approach to modelling — for example, stochastic scenario analysis, hazard maps?
    - Do you understand how the vendor models physical risk and how any outputs are "climate conditioned"?
    - Which climate models does the vendor use, and how are the outputs downscaled?
    - Do you understand and agree with the vendor's underlying assumptions on issues of most importance to you?
    - What metrics do you need — for example, annual average losses, damage ratios, flood depths, productivity lost through excessive heat?
    - Do you understand the vendor's summarizing options applied to reported outputs? For example, do they use 10- or 20-year centred averages?
    - Do you understand the vendor's approach to model creation and validation?
    - Are you comfortable with the vendor's model risk management governance processes?
    - What format do you need the data in, and can the vendor provide that?
    - How much expert help do you need to make the most use of the data from the providers?
    - What is the turnaround time and speed to delivery, to support daily usage for transactional services such as due diligence processes?

7. **Estimates of Uncertainty**
    - Do you want estimates of uncertainty around any aspects of the outputs?
    - Does the vendor calculate uncertainty of the modelled results?

Although this is by no means an exhaustive list, it is a good starting point for establishing a productive relationship with a vendor.

7 Conclusions

Against the backdrop of rising physical risk levels from a changing climate, it will become increasingly important for firms to be able to identify, measure and manage them. This study has looked at how physical risk assessments vary across vendors at the level of individual assets. As this is a relatively new area for many financial institutions, it is useful to understand just how complex the subject matter is and the types of reasons that physical risk assessments might differ. Given the complexity of the choices that must be made, it is no wonder that vendor assessments will differ.

From modelling choices to the interpretation of climate data, from scenario and hazard choices to geocoding, and from metrics to averaging options, the scope for differences of interpretation and methodological divergences is significant. It is therefore not surprising that the study found a substantial divergence in hazard estimates and their corresponding damage. But it's important to recognise that this is a relatively new discipline and it will take time for norms to emerge and become established.

While there is uncertainty over exactly how much physical risk we can expect over coming decades, it is indisputable that the frequency and severity of physical risk will continue to increase due to rising concentrations of anthropogenic greenhouse gas emissions. It is therefore important for financial institutions to understand how much climate-related risk is in their portfolios. To help financial institutions navigate through this complexity we have included a checklist of practical questions that can be used when assessing vendors.

The study should also be put into context. It has focused on the direct impacts of a range of perils on a selected portfolio of properties to get a sense of the dispersion of estimates of financial loss. But this is quite a limited view of the physical risks that might affect a firm or a household at a particular location. For example, the estimates do not include losses from impacts on the contents of a property. Nor do they include any indirect impacts or second-order effects, such as those arising from disruptions to infrastructure services (e.g., power, water and roads), although some impacts can be mitigated through suitable adaptation. For firms, there may also be significant indirect impacts through their supply chain or trade relationships.

The losses from physical risks in the real world might become amplified. For example, repeated severe flooding can have consequences for the price and/or availability of insurance, leading to knock-on consequences on asset valuations. Extreme weather events might also make a location less desirable for investment. Similarly, physical risk events which cause disruption for businesses may impact their valuations.

There is also the possibly of cascading and complex impacts from climate change. For example, as global warming continues to rise, water and food availability will be increasingly impacted by physical events.

The amplifying and cascading effects will impact business continuity, human health, and even give rise to the potential for social conflict and migration. As global warming increases, these cascading influences are likely to dominate and might cause disruption in non-linear and complex ways.

In summary, physical risk assessments are likely to become an increasing influence on day-to-day risk management, and it's important for all financial institutions to grow their understanding and insight into how their portfolios are likely to be impacted. Understanding the nuances and influences that affect vendors' approaches is a good first step. We are pleased that so many vendors gave their time and information to improve this knowledge.