Gallagher Re

Evaluation of Cyber Models

Being an industry centred on computer technology, it should be of no surprise that cyber catastrophe models have developed in complete tandem with the cyber insurance market. The early cyber models had to be built almost entirely on expert judgement, conjecture and speculation due to a scarcity of real data and an incomplete understanding of the emerging risk. The immaturity and limitations of these models were, for the most part, tolerated by the market and regulators while the materiality of the risk remained low, the class was evolving, and no better alternatives presented themselves.

The current cyber market has moved confidently away from being an emerging risk, following years of strong growth and expansion which is showing no signs of abatement,[1] and now sits as a material exposure for many of the largest multi-line insurance carriers and new tech-driven entrants to the market. The cyber models, in response, have done well to keep pace and have gone through rapid updates, improvements and re-builds that is as much an indication of the enhancement in model capabilities as it is an increase in understanding of the risk landscape; however, the ever-changing size and scope of the industry means that claims are often of a vector, frequency or magnitude that are still missing from the cyber models, such as the surge in ransomware attacks since 2018.

A combination of growing market exposure and continuing model development means that regulators, rating agencies and market associations are beginning to look more intensely at the adoption, application, adequacy and efficacy of cyber catastrophe models that are being used for policy pricing, exposure management and capital calculations by (re)insurers that fall under their scope or jurisdiction.

It can be seen that cyber catastrophe model development and maturation is following a similar path to that experienced by natural catastrophe models over the last 30 years, and we can use the lessons learnt by (re)insurers, from adopting and evaluating natural catastrophe models, to highlight nuances and preempt the challenges of evaluating cyber models. The maturation journey of natural catastrophe model usage in the (re)insurance industry can be summarised as a phase of rapid adoption and model development which was ended by a series of challenges to the accuracy and adequacy of the models following a series of large catastrophic events and market inertia. This was then followed by a period of coordinated efforts across the market to understand the model limitations, reduce systemic risk in the market, and improve quality of the available data that is used in the models, as well as an increase in vendor transparency and more measured development road maps. Currently, natural catastrophe models are in a period of maturation where their outputs for primarily perils are broadly relied upon for capital calculations and pricing or reinsurance, as well as companies beginning to adjust models to mitigate known limitations.

Cyber models are currently experiencing a period of rapid adoption, although mostly limited to accumulation management, and the market is seen to be entering a period of improving data quality and making efforts to understand the way cyber risk is simulated by the models, even though they remain unchallenged by a catastrophic event. Although the models still rarely dictate reinsurance pricing, some (re)insurers are beginning to incorporate the outputs into their capital models, and it is for this reason that model evaluation and regulation is becoming important not only for the cyber market but also for the insurance industry as a whole.

[1]Gallagher Re (2021) CY-FI: The Future of Cyber (Re)insurance

"There was also significant divergence on resulting losses among firms. This underlines the large uncertainty in cyber, the lack of reliable claims data and the immaturity of available models with potential links to capital adequacy."

Cyber Underwriting Risk: Follow-Up Survey Results Bank of England PRA 30 January 2019

BANK OF ENGLAND
PRUDENTIAL REGULATION AUTHORITY

Validation vs Evaluation

First and foremost, it is important to understand the distinction between model validation and model evaluation, which is often used interchangeably across the insurance industry.

Model validation is an intensive deep dive into the core elements of a model, stripping it back to the bare framework and assessing the correctness and the accuracy of the data, processes and assumptions used to construct the model. Alternatively, model evaluation focuses on considering the completeness and suitability of a model, and is "the process by which you determine whether the external catastrophe model provides a valid representation of catastrophe risk for your portfolio".[2]

This perhaps subtle distinction is important as the complexity of cyber models and the evolving nature of cyber risk means that the deep dive required for model validation is currently implausible or unachievable for most (re)insurers, and as such attempting to validate a cyber model is likely to be waste of company resources.

"At this stage, working with the models is predominantly about developing understanding—of the approaches, assumptions, limitations, and sensitivities—to enable them to be most appropriately deployed within a much wider framework that goes into forming an internal view of risk, and not as much about validating them. The distinction is subtle but reflects the balance of where we are on the journey of reliance on the models."

Alan Godfrey
Exposure Management Principal
Sompo International

Key Opportunities and Challenges of Model Evaluation

Through the subjective lens of a regulator, there is probably little negative to be said about enacting rigorous model controls upon the companies in the market it oversees.

Evaluation and validation, and the regulations governing them, encourages the proper use of models within a company, it drives understanding of the role they play in the internal operational model, defines processes and controls around their use, and in turn creates a more robust and resilient market. From the other side of the fold, the evaluation process can be a significant burden for risk managers, actuaries and modellers but for the most part, if done correctly, the outcomes can provide opportunity for a company to develop their own view of the risk giving them a competitive advantage and flexibility in model-driven markets.

The main challenge for any company undertaking a model evaluation will always be a case of resource both monetary and human. Many (re)insurance companies will by now have a well-established, and where possible partially automated, evaluation process but conducting a thorough and in-depth investigation of the models remains a time consuming exercise. Evaluation involves many hours of trawling through the details of documentation, holding meetings with model vendors and gathering research from independent sources, as well as additional time taken by internal and external committees and bodies to review and approve the final proposals and recommendations. The cost of evaluating a model is directly correlated with the complexity of the model and the depth of the evaluation; the deeper a company needs to dive into a complex model the more expensive the evaluation becomes due to the increased cost of unpicking and understanding the model components.

[2]LMA/Lloyds (2012) Validating External Catastrophe Models Under Solvency II

A dive into a model's components as part of the evaluation process is the first step for any company wanting to develop its own view of risk; where a model output is adjusted or augmented so that it reflects the company's understanding and underwriting of the risk rather than accepting the model vendor's standard view of the risk. Once a company has uncovered the elements of the model that drive the differences between model outputs and expected risk behaviour, it can focus efforts on closing these gaps by making adjustments to the model. Most model adjustments will require validation and approval if they impact capital calculations, and therefore overly simplistic or broad brush adjustments of a model's output, justified only because the outputs are felt to be too conservative, are unlikely to be viewed favourably by regulators. Conversely, a well-informed adjustment made to a specific element of a model, such as the frequency of large events or vulnerability of a specific industry class, is likely to be reflective of a deeper understanding of the model and risk, and therefore more appeasing to regulators. Model adjustments can give a (re)insurer advantages over its competitors by bringing a model's output in line with claims experience, which could be a reflection of company's superior underwriting practice, and allowing the company to avoid over-capitalisation due to overly penal modelling. Many of the leading Bermudian reinsurance companies, in the early 2010's, were quick to establish natural catastrophe research teams to conduct in-depth evaluations of the vendor models which ultimately lead to adjustments of the model so that the models better reflected their view of the risk. This investment in models and evaluation in part helped them free up capital to challenge the established reinsurance markets that were slower to adopt and adapt catastrophe models.

Unfortunately, these same internal controls have the potential to discourage rapid model development and innovation due to the prohibitive bureaucratic and economic burden of getting significant model updates or loss adjustments evaluated and approved. Even if a new or updated model is a better representation of the risk, the resources required to evaluate models may make (re)insurers, or the market as a whole, slow to upgrade unless the update was determined to have enough of a meaningful impact on the model outputs.

Cyber models will not experience the same period of unfettered model development that helped establish the first generation of natural catastrophe models and modellers.

"Nat cat models are on version 20 or so, with centuries of recorded history, frequent billion-dollar events, processes that change on geologic time, and multiple components of the model verifiable via computational physics and structural engineering. Even with all that, there remains much debate and competition. Cyber risk modelling has none of these advantages, and yet the size of the risk demands we still put our most rigorous foot forward. Model validation is not a binary process that outputs "yes, this model is valid." It is a process that reveals how to best work with well-thought-out models to deal with the uncertainty inherent in the world."

Cody Stumpo
Senior Director of Product Management – Portfolio Products
CyberCube

Brief Overview of Natural Catastrophe Model Regulation

The increasing scrutiny of cyber models is reminiscent of the mid-2000s where a series of large events challenged the adequacy of the natural catastrophe models and the 2008 economic crash lead to a series of directives that imposed more rigorous controls around the modelled outputs in capital adequacy calculations.

Natural catastrophe models have existed since the late 1980's and by the turn of the century their statistical outputs had become essential to large and technical reinsurance companies for treaty pricing and accumulation management. So as not to be left behind, there then followed a rush by other reinsurers and insurers to adopt catastrophe models even though the nuances and limitations of the outputs may not have been not fully understood by those using them. Technological advances and data improvements during the same period allowed for ever more detailed modelling to be completed, which added to the over confidence that some companies had in the accuracy and completeness of the catastrophe models. The first significant challenge to the validity of natural catastrophe models occurred during the highly active hurricane seasons in 2004 and 2005, where many loss driving event features, such as the levee breach during Hurricane Katrina, were not captured by the models and the frequency of extreme years had been under estimated due to similar seasons not appearing in the historical record used to build the models. Equally, as the model vendors were quick to highlight, the quality and quantity of the risk data used by (re)insurers in the model was often inadequate, leading to increased uncertainty and inaccuracies in the outputs. In the years that followed a lot of efforts were made by the vendors to improve the completeness and accuracy of the models, and by the (re)insurers to improve the quality of the modelled exposure data, spurred on by further large natural catastrophe loss years in 2011 and 2012.

Further justification for model regulation arose in 2011 when a major model vendor simultaneous overhauled their North Atlantic Hurricane and European Windstorm models. Although the reasoning for the updates was backed by better understanding of the risk, the compounded impact of the changes resulted in significant increases in modelled outputs for many portfolios and (re)insurers found themselves, almost overnight, being in breach of tolerances and unable to follow business plans despite there being no changes to underlying inforce exposure. This incident highlighted the operational and systemic risks of (re)insurers and the risk of markets being over reliant on model outputs without understanding their inherent uncertainty. In the aftermath of the model change, many companies moved to a blended modelling approach, using the average of multiple model outputs to reduce volatility of model change, but more recent trends have seen a return to single model views as the models become better understood and model updates better managed.

During the same period, there had been frequent calls by European regulators to standardise legislation for (re)insurers across the European economic market. After initial slow progress in developing a framework, the Solvency II directive was proposed in December 2009, pushed forward by the economic crash the previous year, and which was eventually adopted by the EEC in April 2014. Although the regulations covered many measures for insurers to meet, there were particular articles that pertained to the validation of catastrophe models that are to be used in assessing solvency capital requirements.

In parallel to the European regulators, U.S. regulators began to modify their supervisory solvency framework in response to the 2008 economic crisis, and in October 2012 the National Association of Insurance Commissioners (NAIC) issued the 'Risk Management and Own Risk and Solvency Assessment Model Act (#505)' which intends to foster an effective level of ERM at all insurers, and provide a group-level perspective on risk and capital. Although not as prescriptive as Solvency II, as the Act is expected to be incorporated into State level legislation rather than be a regulatory law on its own, it contains the expectations that insurers must explain the level of validation efforts that have been conducted to allow reliance to be placed on external models (including natural catastrophe models).

As a result of the implementation of regulations, many (re)insurance companies that operate within the European, London and U.S. markets now have well-established, consistent and objective processes for validating natural catastrophe models as well as controls in place for managing the periodic changes and updates made by the model vendors.

Evaluation Process

The outcome of any evaluation process should be a document justifying the company's selection of an external model, the criteria by which it judged the model's fitness for use, and how the tests meet the regional regulatory requirements.

It is highly likely that many cyber modellers and actuaries have prior experience or ongoing involvement in natural catastrophe classes of business and therefore have encountered model evaluation previously, but before progressing further, it is best to recap on the fundamentals of external model evaluation.

Although there are many different regulatory regimes, they share a common underlying intent of reducing the risk of insolvency, pricing inadequacy, and risk management failure. This typically involves the need for a (re)insurance company to assess whether an external model fits with their internal view and understanding of the risk, and whether the correct systems, processes, and controls are in place for managing the ongoing use and application of the model.

Documentation Review and Fact Finding

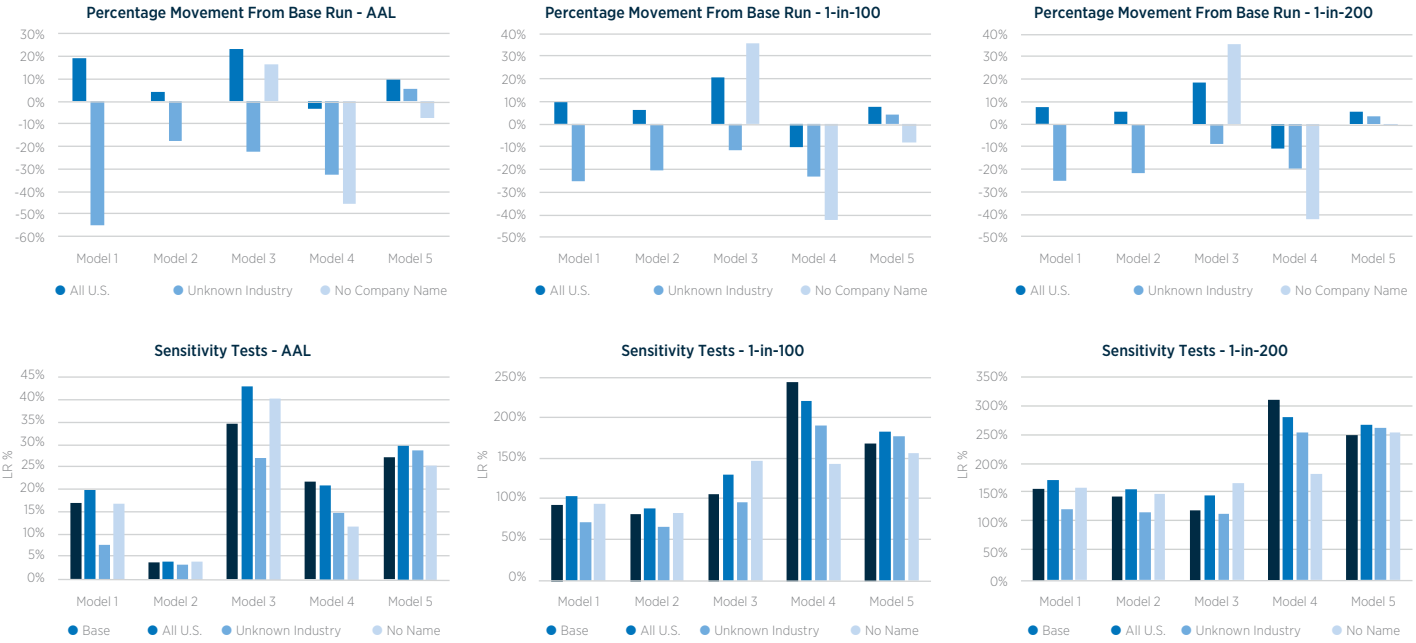
The initial stage of any evaluation process is to read and understand the documentation for the model under review. The aim of this process is to collect facts about the model, rather than the opinions of the reviewer. If enough information cannot be gleaned from the documentation then technical conversations should be had with the vendors. Key elements of the model should be identified and documented (such as attack vectors, coverages, data augmentation processes) as well as, in the case of a reevaluation due to version changes, updates to the model since the previous release; these key elements later inform the choice of validation tests. A useful question to pose to vendors is to ask them to list the three main limitations of their model, as this is can be seen as a test of true understanding and transparency. The reviewer should also record the methods and justifications used by the vendor to calibrate, validate and approve the model, as well as identify limitations in the model, such as non-modelled perils or risk types. The reviewer can provide an opinionated assessment of the quality, quantity and availability of the material and information.

Independent Validation and Suitability Assessment

A fundamental part of any evaluation process is being able to conduct an independent challenge to the underlying science, assumptions and parametrisation within a model. The reviewer should adopt a subjective stance and use alternative data sources to test the validity of the model.

The testing will often involve running the company's portfolio(s) of risks, and/or a synthetic or industry data set, through the catastrophe model and then assessing the appropriateness of the outputs. Back-testing the outputs against historical data and comparing outputs to industry expectations are the two simplest and most common approaches to assessing the model outputs. Where the outputs differ from expectations, reviewers should look to identify what is causing the differences through sensitivity and stress testing. This involves running multiple iterations of the data set through the model, varying key parametrised elements in each iteration, to test the impact each element has on the output. This allows the reviewer to identify where the key assumptions reside within the model. Data iterations may include making artificial changes to the geographic location of the risk, the industry classification, or the intentional reduction in data quality.

Where the parameterisation of the model is a result of expert judgement, the assessment should be a peer review of the elicitation process, and technical reviewers can also reference current academic papers and publications which may contain findings that challenge the underlying science of the model.



Example sensitivity tests conducted by Gallagher Re for major cyber models. Sensitivity testing is the process of making variations to the input data fields and then observing the impact the changes have to the model outputs. Typically one input field is varied at a time. For the sensitivity tests above, that were performed on an example dataset, the company names were removed/obfuscated, the country of all companies were set to U.S., and Industry codes were set to Unknown.

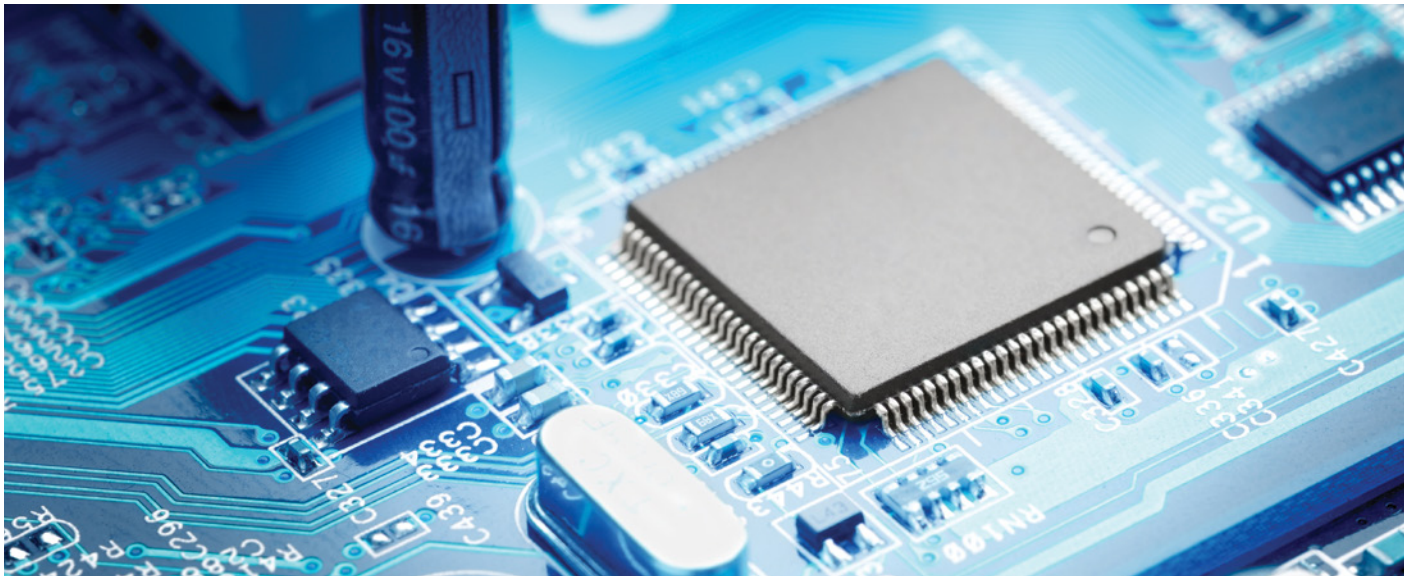
Once the independent validation process is complete, the reviewer should be able to draw conclusions on the strengths, weaknesses and limitations of the model, and be able to relate them specifically to the company's own portfolio.

Recommendations

The outcome of the evaluation process is a set of recommendations and justifications for adopting or rejecting the model that will be made to the relevant internal governance boards or committees. The decision should be supported by the findings made during the evaluation process and may be conditional on excluding specific classes of business or coverages from the modelling process, where outputs did not align with expectations, or making adjustments to the model outputs. The reviewer could also identify factors that would result in the falsification of their recommendations if they were to occur sometime in the future and therefore trigger a reevaluation of the model. Falsification factors include changes to the behaviour of the peril, such as a sudden and significant increase in event frequency, or new understanding of the risk that challenges or negates established expert judgement.

"We don't expect syndicates to try to convince us that the chosen cyber model is perfect. We would prefer that the syndicates are open and honest about what all the limitations are in the cyber model and what they intend to do to cope with the uncertainty."

Emma Watkins
Head of Exposure Management & Aggregation
Lloyd's



Nuances and Complications of Evaluating a Cyber Catastrophe Model

As natural catastrophe model validation has become such an established part of risk management regimes, many (re)insurers are likely to feel that they will not struggle with applying those same processes to cyber models, but the nuances between natural catastrophe models and cyber models will present difficulties to those who are unprepared.

Prevalence of Expert Judgement and Lack of Market Consensus

An element of expert judgement exists in all types of peril models, whether natural or human-made, but the current cyber models place a heavier reliance upon expert judgement, than their natural peril counterparts, as the quantity and quality of relevant and meaningful data available to parameterise a model is limited in such a young and rapidly evolving industry.

The need to heavily rely on expert judgment is most notably highlighted by the absence of a broad catalogue of historic catastrophic cyber events upon which the tail of the models can be calibrated; NotPetya being the only notable event of in the last decade. This is further frustrated by the lack of consensus within the cyber industry around the definition of such a catastrophic cyber event, and although there is general agreement that a catastrophic event will be a systemic ransomware attack and/or the outage of major service provider, the likelihood and impact of such an attack is still a matter of discourse. Therefore, models and modelling firms are limited to using scenarios that have been elicited from expert judgement.

The prevalence of expert judgement in cyber models and lack of market consensus will decline as data is collected through market experience, but cyber, being a social science at its core, will always lack the same access to quantitative data that is available to natural perils as the data can be acquired for the latter through scientific empiricism. The use of expert judgement in the models, actually means that it is logically impossible to conduct a fully independent validation of a cyber model because validation requires objectivity but expert judgement is, and any alternative expert judgement will be, a subjective opinion of the risk. In contrast, it is possible to conduct an evaluation of a cyber model using expert judgement as this requires an assessment of the processes used by the modelling companies in reaching the expert conclusions and quantifications, ensuring that it is unbiased and appropriate, rather than an assessment of a set of parametrised values, which is the approach used to validate many elements of natural catastrophe models.

"A fully independent validation of a cyber model is not currently possible with current client experience and resource. Natural catastrophe models are highly reliant on statistical inference and therefore they are 'easy' to validate against experience. Cyber models should not be statistical as there is neither the data, nor the peril stability to use this method of model build; therefore, standard validation techniques are not meaningful. However, validation can be done on loss costs and/or comparing internally derived scenarios with model event sets."

Matt Harrison
Director, Product Management, Cyber RMS

Absence of Independent External Data Sources and Institutions

For natural peril models, insurers and modeller vendors can draw on information from well-respected and established institutions that are science led, politically autonomous, and economically independent. This data can provide an unbiased view of risk from where to build models and (re)insurers to validate the same, and the independence of the institutions helps to drive consensus of the risk across the insurance market. An example of such an institution is the United States Geological Survey whose datasets and scientific research forms the basis of nearly all major U.S. earthquake models.

An equivalent independent institution has yet to emerge for cyber risk but efforts to establish one are underway,[3] while other potential institutions currently focus on cybersecurity and incident response. Two such leading cybersecurity organisations being MITRE and FIRST. Although their data and products currently have limited application to model construction or evaluation, they may prove more relevant in the future.

Institution	MITRE	FIRST
Website	www.mitre.org	www.first.org
Description	MITRE was founded as a nonprofit company to advance U.S. national security and serve as objective advisors to government agencies, both military and civilian. They are a federally funded nonprofit organisation.	A Forum of Incident Response and Security Teams made up of global members. They are a U.S.-based nonprofit organisation.
Founded	1958	1989
Focus	Cybersecurity and Defence, and Vulnerability Identification	Vulnerability Classification and Incident Response
Key Products	Common Vulnerabilities and Exposures (CVE) programme to identify, define and catalogue publically disclosed cybersecurity vulnerabilities. ATT&CK— a curated knowledge base and model for cyber adversary behaviour.	Common Vulnerability Scoring System (CVSS) is an open framework for communicating the characteristics and severity of software vulnerabilities. Exploit Prediction Scoring System (EPSS) is an open data-driven effort for predicting when software vulnerabilities will be exploited, to assist with defenders with prioritisation of remediation efforts.
Relevant Model Component	Scenario Generation, Vulnerability	Vulnerability

Instead of relying on an external authoritative data source, a company may choose to invest in their own team of experts, but the pool of talent in the insurance market remains small in comparison to more established fields such as natural perils, and attracting cyber experts to the insurance industry is difficult as long as tech-focused sectors offer more exciting opportunities. Problems of holding knowledge within a company are further compounded by the rapidly evolving threat landscape, attack motivations, attack vectors, emerging system vulnerabilities, and current rates of attacks, which can give any cyber knowledge a short shelf life. Overcoming this problem will be a challenge for any insurance company.

[3]Insurance Insider (28 September 2022) "CFC spearheads cyber cat-declaration initiative to tackle systemic risk"

"A hindrance to current cyber risk management and model validation is the absence of a source of authority when considering the cyber hazard landscape. The lack of consistent baseline perspective (and presence) on hazard, that all interested parties could utilise—akin to USGS or NOAA in natural catastrophe terms—makes even the starting point for considering hazard opaque within the industry."

Stephen Gibson
Casualty and Cyber Exposure Manager
AXIS Capital

Stability in Cyber Risk Behaviour and Policies

Due to the changing nature of the cyber coverages, policy terms and risk categorisation, on-levelling a historic claim to current year, so that this can be used to validate the severity of model outputs, is complicated and introduces uncertainty. Equally, some types of attack vector, such as ransomware, have experienced a rapid change in frequency and severity that invalidate the relevance of historic claims, and the repeatability of some claims are questionable, especially where they exploited a vulnerability in a piece of technology or software that no longer exists.

These issues can be mitigated by both choosing a limited timeframe or selected attack vectors from which to draw historic claims, but this will reduce the size of the validation data set and potentially introduce bias into the evaluation process, and adopting a forward-looking approach to assessing risk which is more akin to validation of climate change models than of traditional catastrophe models. Expert judgement should be adopted to assess the appropriateness of the data set as part of a model evaluation.

Varied and Complex Models

If a company is truly intent on finding a cyber model that is the best representation of their view of risk, then they should consider evaluating multiple models, but as many modelling companies have moved away from allowing short-term licencing arrangements, this has become prohibitively expensive.

Unlike the natural catastrophe modelling market where model frameworks and methodologies have mostly converged over the years, there are still large variations in the methods used by different vendors to quantify cyber risk; particularly in relation to scenario definition, event generation, vulnerability indicators and single point of failures. This variation between the models means that knowledge and understanding of one cyber model is often not applicable to another, and more time must be spent understanding each cyber model than one would experience when evaluating multiple natural catastrophe models. Problems of understanding the models is compounded by the quantity and quality of model documentation provided by cyber model vendors and the transparency of the back-end of the models; this has shown improvements over the last few years but often remains below the standard set by natural catastrophe models and expected by regulation.

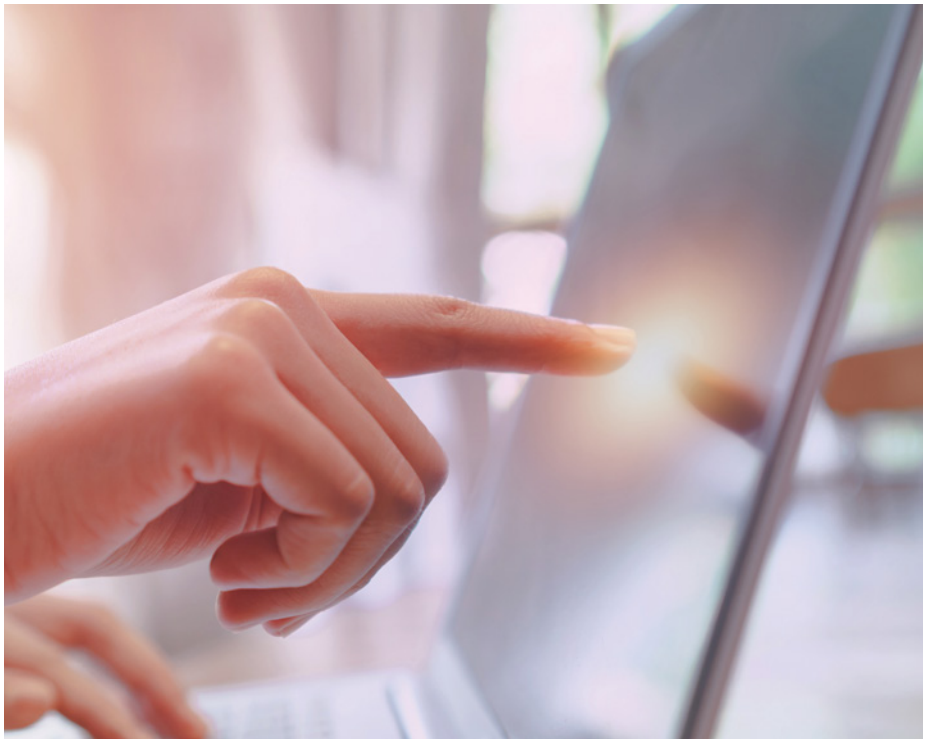
The differing methodologies between model vendors often also result in vastly differing output magnitudes and behaviours. This can render meaningful and deep comparisons between the model outputs meaningless if not conducted with care; yet another cost and burden that limits multiple model evaluations.

All that being said, brokers are typically able to assist if they have access to multiple models and can perform some of the analyses required by clients.

Rapidly Changing Models

Updates to models are typically in response to new understanding of the risk being modelled, a change in behaviour of the peril, or an improvement to the functionality and capabilities of the modelling systems and software. Given the dynamism of the cyber insurance market, the constantly shift threat landscape and the immaturity of the modelling systems and software, Cyber models undergo more frequent and significant updates than their natural catastrophe equivalents. It cannot be argued that these updates are not required to keep the models abreast of developments in cyber, but such severe and short cycles of model change is a burden on companies who must reevaluate the model and reconcile the output each time an update is released.

Many model vendors have stated that they are aiming for a reduction in the frequency and severity of their model updates in the coming years, which is in response to some (re)insurers voicing their desire to use the models as inputs into capital modelling and therefore require more annual stability.



Firmographic Data Quality and Data Augmentation

Although Gallagher Re have seen a steady improvement in the quality and quantity of the bordereaux data, the most complete data fields remain focused around policy terms and conditions. Although these are essential for both exposure management purposes and modelling, to be able to identify the insured company and quantify the risk many of the models require additional firmographic information that is often not supplied or recorded by the insured/insurer.

This disparity between the availability of data and the model's expectations of data is something that should be considered by (re)insurers when designing evaluation tests. An example of this disparity is 'Record Count' which some of the main cyber models are highly sensitive to, but as the number of records the insured holds is a dynamic factor it is quite unlikely that they will accurately know true quantity of records, let alone the split between personal health, personal credit card, and personal identification information. The insured therefore often makes an estimate or does not supply it at all. Before conducting a model evaluation (re)insurers should consider improving the quality of data in their portfolio to ensure that what they hold is adequate for conducting appropriate sensitivity tests.

A further complication of poor data, is that many models augment the firmographic data in the considered portfolio by drawing information and risk exposure from an actively maintained database of information compiled from various data sources or scrapped from the internet. Parallels have been drawn with the Hazard Look-up process that natural catastrophe models utilise.

The key limitation with the data augmentation process is that it relies on first being able to identify and match the company name, which each model does with varying degrees of accuracy and rapidly diminishes as the size of the company and quality of the initial firmographic data decreases. If the models are unable to match a company name then a variety of methods are adopted to either link the unmatched company to a synthetic one with similar characteristics or to use an aggregate data/market share approach. Many of these matching approaches are non-repeatable and/or affected by the portfolio mix—such that two near identical portfolios may have the same unmatchable companies being linked to different synthetic companies—and these processes limit the ability to easily and accurately conduct deep dive sensitivity testing. The inaccuracy of the company match also hampers model to model comparisons because the level of difference between the two outputs that is a result of differing data augmentation methods must be determined before the quantum between event losses can be determined.

	Data Field	Model Importance	Market Sample Average 2021	Market Sample Average 2022	Change
Firmographic	Insured name	High			–
	Revenue	High			↑
	Trade industry/SIC/NAIC code, etc.	High			–
	Website	Low			↑
	Region/Country	High			↑
	Number of records	Low			↑
	Employee count	Low			↑
Policy Terms	Risk ID/Policy ref	Low			–
	Inception date	Low			↑
	Expiry date	Low			↑
	Currency/FX rate	Mid			↑
	100% limit	High			↑
	Limit to insurer	High			–
	Attachment of insurer	High			↑
	BI wait period	Mid			↑
	SIR and/or deductible	Mid			–
	Share	High			–
	Gross gross premium	Low			–
	Gross net premium	Low			–
Sublimits	Third-party sublimit/Deductible	High			↑
	Breach response sublimit/Deductible	Mid			↑
	BI sublimit/Deductible	Mid			↑

Gallagher Re have been monitoring the data quality of bordereaux in the London market and have seen a steady improvement over the last few years but the focus of data remains on policy terms and not firmographic data which models tend to rely upon, particularly for company matching and data augmentation processes.

Risk Scoring

Additional functionality that limits the ability to conduct model evaluation through model comparison, is the reliance of some models on risk scores to quantify the vulnerability of a company. The value of these scores, which are typically provided by a third-party outside-in scanning agency, differ significantly between security firms[4] and the scores can change many times or without notice or justification within the span of a year. Thus, validating a model on a fixed portfolio at different times of the year can result in different outputs simply because the risk scores have been updated. The noise this causes in the outputs must be accommodated for when comparing outputs between models or when conducting a programme of sensitivity testing that could stretch over a long period of time.

Immature Models

When evaluating a mature natural catastrophe model, it is often possible to spend only a brief time testing that the core elements of a model are functioning correctly because many of them, such as financial engines, have already been repeatedly tested and broadly accepted across the market and have remained unchanged for many years. In contrast, these same elements in cyber models are still developing and may not accurately reflect real world behaviour even if they intend to follow accepted actuarial methods, and therefore when (re)insurers conduct a cyber model evaluation they must plan to spend more time focused on the basic elements of a model than they would for a mature natural catastrophe model.

Recent examples of financial engine errors found in various cyber models, and which have since been rectified by the vendors, include time based deductibles being effectively ignored for business interruption losses, ransomware losses maxing out layer limits even if ground-up loss is below excess, and double counting of revenue across single-points-of-failure thus resulting in overstated losses estimates.

Operational Risks

Most regulations that govern model usage, require an assessment of external factors that could impact the ability of the company to operate effectively. One such consideration should be the sudden loss of access to a model upon which capital and pricing calculations are dependent.

In the brief history of cyber models, and their start-up companies, there have already been some notable casualties where either a company has ceased to exist or a model is decommissioned. Recently, there was also nervousness and gossip around the market as a major cyber vendor was appearing to have difficulties raising financial support from the venture capitalist market. Therefore the fiscal and organisational health of a cyber vendor should be a large consideration for a (re)insurer when conducting model evaluation, as the continuing reliability of the model is critical for future business stability.

[4]Gallagher Re (2021) Looking from the Outside-In: Can taking the threat actors' viewpoint help insurers?

"Cyber-market growth necessitates a comprehensive framework for understanding and quantifying correlated large cyber losses, one which delivers validated models that the market can have confidence in as being realistic and capturing the full range of catastrophe experience."

Aidan Flynn
Head of London and International Underwriting Management, Cyber and Tech
Beazley

Closing Opinion

Despite a lot of companies showing cautiousness around relying on cyber models that remain untested by large and significant catastrophe events, or are displaying hesitancy about burdening themselves with the task of validating models that are still so uncertain and prone to large updates, the next stages for in the cyber model development journey will see them being generally accepted for use in solvency calculations, and for that reason there is a need for evaluation and a need for understanding the nuances around evaluating cyber models as the cyber risk modelling industry is still in a period of maturation and change that makes validation and evaluation of the models more complicated than natural catastrophe models.

While the cyber models remain outside of the capital calculations and reinsurance pricing is not driven by model outputs, cyber model vendors should continue to push for ambitious and broad model development road maps, and the models should be seen as trying to help the market understand a risk not dictate the losses.

"Even though, result stability is important for e.g., business steering, Munich Re sees the model quality as the most important aspect. The cyber market is still in its infancy and therefore, changes in the results are expected and (for MR) also accepted. Munich Re is working on its own models and also facing the challenges with volatility in model results."

Dr. Stephan Brunner
Senior Cyber Actuary
Munich Re

Gallagher Re Evaluation Framework

Gallagher Re have built a cyber model evaluation framework to assist with our clients model evaluation exercises. The framework is aligned to solvency II requirements and relates to Lloyds' Catastrophe Exposure Principles. It also goes beyond typical evaluation frameworks by considering market opinions and operational risks.

Vendor
Assessment of the vendor's choice of partnerships, company solvency client service levels, market opinion, licencing costs and regulator relationships.

Documentation
Review of the quality, quantity and availability of model documentation.

Development
Assessment of the processes and data used by the vendor in developing, parametrising and calibrating the model.

Scope
Review of the model scope, completeness, and the appropriateness and plausibility of the event scenarios.

Infrastructure
Focusing on the hardware requirements, and software performance, functionality and user experience.

Change Management
Review of the frequency of model updates, the model development road map and level of support provided by vendor.

Exposure
Assessment of the quantity and availability of data required by the model and effectiveness of data augmentation and company match processes.

Outputs
Benchmarking of model outputs, sensitivity testing, validation of actuarial methods and review of available dimensions for loss outputs.

The evaluation framework is structured as a questionnaire that covers eight key evaluation components, and 22 sub-components, with both qualitative and quantitative considerations. Based on the findings from the questionnaire, a model is given a principle-based score for each evaluation component, as well as a weighted overall score. The questionnaire and principal-based approach gives the framework flexibility, enables rapid assess of new models and allows for simple comparison models across different vendors.

Model Validation Under Solvency II

Solvency II is a European legislative programme that was implemented in the EU member states, including the UK, on 1 January 2016. Due to the geographical scope of Solvency II and the interconnectivity between the European, Bermudian, Swiss and London (re)insurance markets, it is likely that most markets have been affected by its regime and is therefore a fitting example of model validation being enforced through regulation.

The aim of Solvency II is to proscribe a deeper understanding of a company's risk management framework and capital calculations than previous regulatory requirements, and to harmonise the approach of supervisory regimes across EU member states. There is also additional guidance to allow companies of non-EU 'third countries' to comply with local regulation in respect of EU activities.

The principals based framework consists of three pillars:

Pillar 1: Financial Requirements	Pillar 2: Governance and Supervision	Pillar 3: Reporting and Disclosure
Principals laying out quantitative methods used for calculating a company's capital requirements, and assessment of the risks around these calculations.	Qualitative assessment of a company's risk management and governance structure, and how the viability of the company would be threatened if they failed.	Requirements setting out the continuing reporting of company performance and ongoing compliance with regulations.

Pillar 1 is the most important in terms of model validation, as it contains six articles that detail how a catastrophe model must be assessed before its outputs can be used for capital calculations purposes. The six areas of assessment are:

Use Test	Assessment of how catastrophe models are used in decision-making, how well understood is the science and functionality of the models within the company, and why the models fit to the business.
Statistical Quality Standards	Assessment of the completeness of the models (in terms of material risk), how the risk data is used within the model, and validate the quantitative methods used within the models, and assumptions made.
Calibration Standards	Assessment of the measures and methods used to calibrate the models.
Profit and Loss Attribution	Assessment of the degree to which the financial profits and losses may be explained by the model.
Validation Standards	Assessment of the methods used by the company to validate the models, detailing independence of validation process, tools and methods used to validation, and key assumptions made.
Documentation Standards	Assessment of quality and quantity of external model documentation and data, standards for the recording of model changes, and minimum documentation requirements.

Further information about Solvency 2 can be found on the Lloyd's website: https://www.lloyds.com/conducting-business/regulatory-information/solvency-ii/about.

Contacts

Simon Heather
Head of Cyber Catastrophe Modelling
E: Simon_Heather@gallagherre.com

Justyna Pikinska
Head of Specialty Analytics
E: Justyna_Pikinska@gallagherre.com

Michael Georgiou
Senior Cyber Actuary
E: Michael_Georgiou@gallagherre.com

Ed Pocock
Senior Cyber Security Consultant
E: Ed_Pocock@gallagherre.com

Gallagher Re