Risk publications

Model risk management

Special report 2019

Risk.net

Risk.net

Sponsored by

Management Solutions
Making things happen

pwc

sas

Management Solutions
Making things happen

PROFESSIONAL CONSULTING SERVICES

STRATEGY | SALES AND MARKETING MANAGEMENT | TRANSFORMATION: ORGANIZATION & PROCESSES | RISK MANAGEMENT & CONTROL | MANAGEMENT & FINANCIAL INFORMATION | NEW TECHNOLOGIES

Management Solutions is an international consulting firm whose core mission is to deliver business, risk, finance, organization, technology and process-related advisory services

GENERATING VALUE *THROUGH* RISK MANAGEMENT

MADRID BARCELONA BILBAO LONDON FRANKFURT PARIS AMSTERDAM OSLO WARSZAWA ZÜRICH MILANO ROMA LISBOA BEIJING NEW YORK BOSTON ATLANTA BIRMINGHAM SJ DE PUERTO RICO CIUDAD DE MÉXICO MEDELLÍN BOGOTÁ QUITO SÃO PAULO LIMA SANTIAGO DE CHILE BUENOS AIRES

Risk Management and Control

We provide the best solutions to effectively manage and control the risks associated with doing business: credit, market, operational, liquidity, structural, reputational, conduct, model, business, and other risks, and provide an integrated view of them (and of their impact on capital requirements or provisions).

AREAS

- ORGANIZATION, GOVERNANCE AND RISK FRAMEWORK
- RISK PROCESSES AND SYSTEMS
- RISK REPORTING, DATA AND MODELS

DIMENSIONS

- EWRM (APPETITE, CAPITAL MANAGEMENT, RAR, PRICING, ETC.)
- CREDIT
- MARKET AND COMMODITIES
- STRUCTURAL
- OPERATIONAL
- NON PRUDENTIAL

SOLUTIONS

- SIRO (INTEGRAL OPERATIONAL RISK SYSTEM)
- MIR (RISK INFORMATION MODEL)
- PALADIN (AML)
- GAMMA (MODEL GOVERNANCE SYSTEM)

Tom Osborn, **Editor, Risk Management**
tom.osborn@infopro-digital.com

Joanna Edwards
Senior Director, Marketing Solutions
joanna.edwards@infopro-digital.com

Stuart Willes, **Commercial Editorial Manager**
stuart.willes@infopro-digital.com

Alex Hurrell, **Commercial Subeditor**
alex.hurrell@infopro-digital.com

Antony Chambers, **Publisher, *Risk.net***
antony.chambers@infopro-digital.com

Philip Harding
Head of Content, Marketing Solutions
philip.harding@infopro-digital.com

Esha Khanna, **Commercial Manager, *Risk.net***
esha.khanna@infopro-digital.com

Nadia Pittorru, **Sales**
nadia.pittorru@infopro-digital.com

David Pagliaro, **Group Managing Director**
david.pagliaro@infopro-digital.com

Ben Wood, **Managing Director, *Risk.net***
ben.wood@infopro-digital.com

Ben Cornish, **Senior Production Executive**
ben.cornish@infopro-digital.com

Infopro Digital (UK)
Haymarket House, 28–29 Haymarket
London SW1Y 4RX
Tel: +44 (0)20 7316 9000

Infopro Digital (US)
55 Broad Street, 22nd Floor
New York, NY 10004-2501
Tel: +1 646 736 1888

Infopro Digital (Asia-Pacific)
Unit 1704-05, Berkshire House
Taikoo Place, 25 Westlands Road
Hong Kong, SAR China
Tel: +852 3411 488

Cover image
John M Lund Photography/Getty

Published by Infopro Digital

Model citizens

A spectre is haunting Europe – the spectre of model risk. Launched in 2016, the European Central Bank's (ECB's) Targeted Review of Internal Models (Trim) has forced a step-change in attitudes among European lenders towards ensuring their capital models are fit for purpose. In keeping with other regulators worldwide, the watchdog's team of inspectors is visiting banks to check everything from internal governance processes to the data inputs that underpin modelling assumptions.

If the early evidence from the review is anything to go by, banks still have significant work to do to get their houses in order. The latest set of findings, on the safety and soundness of banks' market risk models, landed in April – and made for grim reading. Of 30 banks that had been subjected to supervisory visits, the ECB found, on average, 32 issues with modelling practices – with, on average, nine issues deemed severe.

The review is already proving costly to lenders – and not just from a compliance point of view: ABN Amro cited changes made to its modelling practices as driving a €1.3 billion jump in credit risk-weighted assets during the first quarter of this year – implying the regulator thought its models were not adequately gauging the credit risk in its loan portfolios previously, necessitating a top-up.

For global lenders, Trim followed hot on the heels of the US Federal Reserve's SR 11-7 guidance on model risk management (MRM) – published in 2011, though not enacted until 2012. Where Trim is, as the name suggests, targeted in scope, SR 11-7 is broad enough to capture anything that looks like a model within a bank, from a value-at-risk model to a simple spreadsheet-based factor model.

In reality, of course, Trim was a politically motivated project – partly designed to keep pace with SR 11-7, but also to shore up confidence in the use of internal modelling among European watchdogs keen to have some collateral to back their pro-model stance during the final negotiations over Basel III. In the opposing camp were US regulators – distrustful of internal modelling practices in the wake of major failings revealed during the financial crisis, and preferring instead the use of revised standardised approaches where possible, as well as an output floor to bind internal model estimates to these.

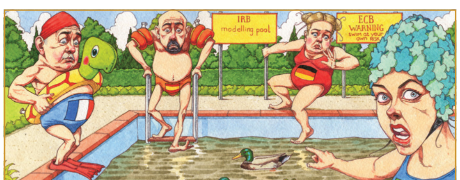
All of this has meant a compliance headache for banks, and a huge spend on hiring or redeploying quants from model development to risk management and validation teams. Quants don't come cheap, nor do the army of consultants brought in to oversee the process. Sources tell tales of one US bank that attempted to lower costs by cutting as many PhD model quants as it could, and replacing them with master's graduates – only to be red-flagged by its regulator.

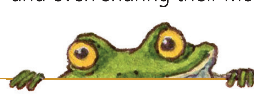
While some of the changes to validation practices have required quant upskilling, much of the change has been around people and processes – motherhood and apple-pie operational risk practices such as establishing independent oversight and effective challenge during the model development and deployment phases.

Anne-Cécile Krieg, deputy head of MRM at Societe Generale, notes that the mindset has shifted. All three lines of defence should be responsible for MRM; previously it tends to have been left to the second line of defence. Now there are specific roles allocated across the three lines and it is fully embraced and embedded. With MRM, there are a significant number of stakeholders in the first line of defence, including the designer of the model, the person implementing it, the users and those tasked with surveillance. Now, all of those roles are identified in the first line, with increasing emphasis on users and the model owner roles.

Tom Osborn
Editor, Risk management







Yields.io is the first FinTech to use AI to automate model risk management

- Spend more time on exceptions by automating periodic validation
- Allow for more robust analytics by monitoring data and model quality
- Start managing model risk using scalable AI technology

London Brussels www.yields.io info@yields.io

Machine learning governance

The ability of machine learning models to read great quantities of unstructured data, spot patterns and translate it into actionable information is driving a significant uptake in the technology. David Asermely, SAS MRM global lead, highlights the need for rigorous model governance as businesses expect to adopt artificial intelligence and machine learning models to support key risk business use cases

Today, there is great interest in harnessing machine learning to turn the massive volumes of data – including non-traditional data – into new insights and information. In contrast to traditional statistical models, which are limited in the number of dimensions they can effectively access, machine learning models overcome these limitations and can ingest vast amounts of unstructured data, identify patterns and translate them into actionable information.

It is therefore no surprise that machine learning modelling is being eagerly adopted. A recent survey conducted by SAS and the Global Association of Risk Professionals found that, over the next three to five years, businesses expect to significantly increase adoption of artificial intelligence (AI) and machine learning models to support key risk business use cases (see figure 1). Banks, for example, are using machine learning models in marketing, fraud detection and anti-money laundering.

However, the fact that machine learning models need more governance than other data models is often overlooked. While machine learning models offer the promise of better predictions, they may also introduce ethical biases and increased model risk.

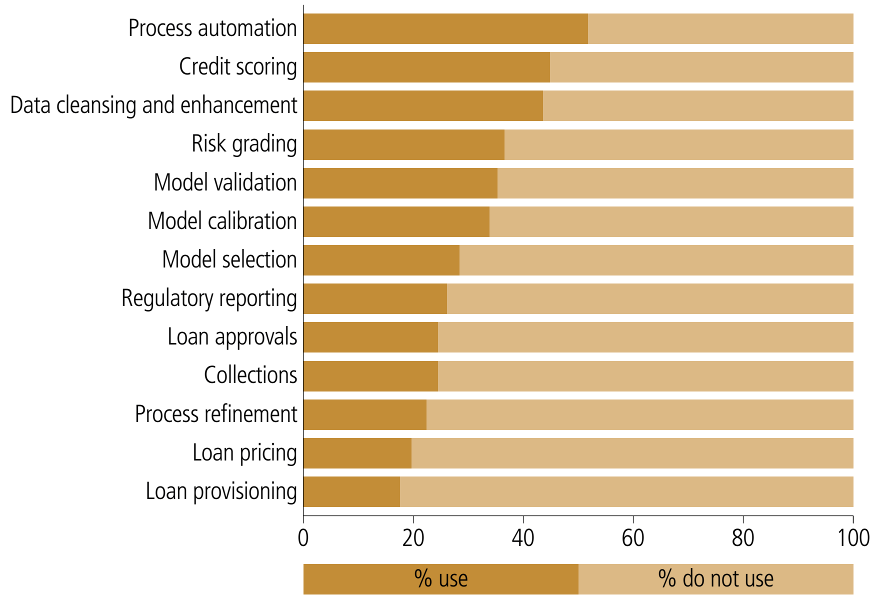
Machine learning models are designed to improve automatically through experience. This ability to 'learn' is what enables greater machine learning model accuracy and predictability. At the same time, it can heighten the need to quickly identify when a model begins to fail.

As a result, there's an increased need to define operating controls on machine learning inputs (the data) and outputs (the model results). The dynamic nature of machine learning models means they require more frequent performance monitoring, constant data review and benchmarking, better contextual model inventory understanding, and well thought out and actionable contingency plans.

Looking ahead, the need for effective governance for machine learning models will only increase. This is a result of:

- Growing complexities of the global, multidimensional marketplace
- An increasing volume and complexity of data
- Rapidly increasing model usage by industries
- Growing complexity of machine learning models.

1 Rates of AI adoption by risk use case



Source: SAS and the Global Association of Risk Professionals, *Artificial intelligence in banking and risk management: keeping pace and reaping benefits in a new age of analytics*

Implications of ineffective MRM on banks

Governance, risk and transparency concerns have introduced a major speed bump in machine learning model adoption. As noted by the US Federal Reserve: "Model risk increases with greater model complexity, higher uncertainty about inputs and assumptions, broader use, and larger potential impact."

The promise of analysing non-traditional data using less transparent machine learning models – for example, to make better predictions – has raised major financial, reputational and regulatory concerns. Banks must be able to clearly explain their own models and how outputs were achieved using machine learning techniques – and yet often struggle to do so with machine learning models. Without this, how can regulators measure the systemic risks of such models to the global banking community?

The 'explainability' limitations of machine learning have also stopped many banks taking advantage of new, non-traditional data sources such as social media. Many machine learning models therefore don't use new data sources, often resulting in them not meeting the expected lift in accuracy compared to historically tuned statistical models. Organisations must determine if the lift is worth the shift from well-understood and explainable models to more complex and less explainable machine learning models.

The solution: Robust, automated governance

Given concerns regarding transparency and the potential misuse of machine learning models, it is vital that organisations implement a robust and automated model governance system. The good news is that MRM teams are investing significant time and resources to determine how to best manage these models.

As AI/machine learning models become the norm, rigorous supporting model governance will be needed to classify machine learning models and introduce more frequent performance-monitoring quantitative data, benchmark comparisons, model usage, model interconnectedness, interpretability, variable sensitivity, modelling techniques, data metrics, model technique rational documentation and much more. ■

LEARN MORE

Complex artificial intelligence/machine learning black-box models are being considered to replace well-understood statistical models. This change offers the promise of better predictions but may introduce unknown ethical biases and increased model risk. Model risk professionals are grappling with how to best reduce this risk with automation, technology and best practices. By utilising best practices – rationale, mapping, data governance, performance monitoring, recalibration, interpretability, benchmarking and contingency planning – financial organisations can better satisfy increased regulatory demands when implementing a robust, reliable and automated machine learning model governance infrastructure.

To learn more about how SAS can help, visit *www.sas.com/mrm*

Branching out

Bayesian analysis can replace forest with a single, powerful tree, writes UBS's Giuseppe Nuti

The machine learning tidal wave is sweeping finance alongside most other industries. Our quants are busy applying new models to various (often old) problems: reinforcement learning for option pricing, deep neural networks for alpha generation, and so on.

Alas, colleagues in model validation – and possibly our regulators – are less enthusiastic, and likely with good reason: these models are often black boxes, making it close to impossible, for example, to explain why an algorithm was short in order to hedge an in-the-money put option on our books. More generally, how do we ensure our models are safe, fully compliant with current rules and regulation, and – an often-overlooked principle – act with common sense?

In a bid to tackle some of these explainability issues, we tried to transform one of the most commonly used black-box machine learning techniques into a white (or at least light-grey) box – something simpler and more transparent, producing results that are easier to interpret.

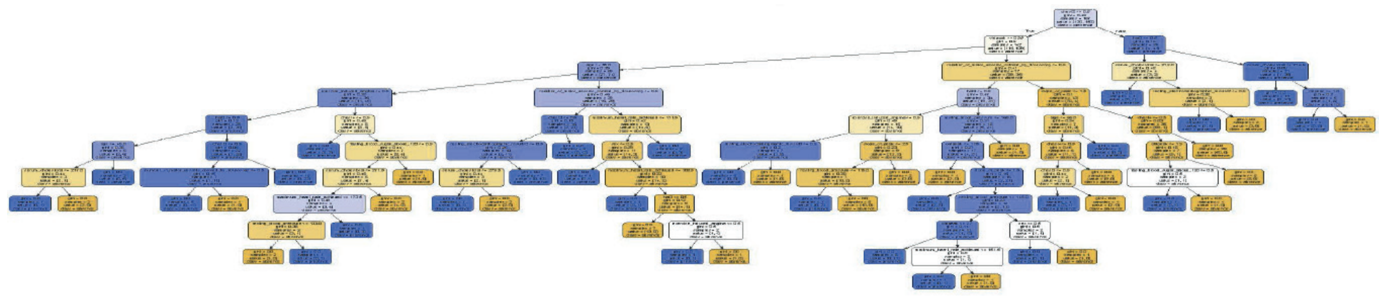
While most machine learning experts have their favourite technique, random forest is a ubiquitous choice for non-linear, multidimensional classification and regression problems. The technique is great at handling large datasets, in large dimensions. If there were a competition to select the technique that performs best without any prior, domain-specific knowledge of the problem, then random forest would be a contender for most datasets.[1]

Random forest works by trying various combinations of variables out of all of the data provided in order to build an extended family of decision trees (hence the name). For example, if we were looking to forecast the probability that a person will suffer from coronary heart disease, we might want to use all of the readily available patient information – such as age and gender – alongside medical data on resting blood pressure and electrocardiographic results.

A single decision tree would simply specify thresholds of these variables such that the probability of the disease is substantially different for the two sides of the threshold – for example, a tree would specify that anyone experiencing severe chest pain (on a level of three or more) and of age greater than 49 will have a 55% probability of heart disease. With a random forest, we try various

Giuseppe Nuti is a managing director with the Strategic Development Lab in UBS's foreign exchange, rates and credit business

1 Hard to read: a single, classical decision tree



Source: UBS

combinations of input variables; a new patient's data is then run over all the decision trees previously trained to determine the probability of the disease.

While this approach can result in accurate predictions, we often do not have the ability (and, in this case, the ability for a doctor) to look at the many decision trees in order to check if the algorithm makes medical sense or if, for example, it complies with all the relevant ethical guidelines (such as potential discrimination if used by an insurance company).

In our opinion, the main limitations to explainability arise from the challenge of summarising the overall behaviour of many trees, especially for more extended trees that concatenate numerous conditions to form a complex structure. We see these two problems (depth and number of trees) as a manifestation of the same issue: a complex tree can overfit the data, which is then mitigated by using an ensemble of trees such that, together, they offer a better generalisation of the data. In essence, each step of the tree generation lacks a measure of probabilistic significance.

In a somewhat separate branch of statistics, recent advancements in Bayesian analysis have allowed us to understand the concept of probabilistic significance of a model with respect to the observations available. This offers a way to solve the problem: by formulating the tree generation process as a Bayesian model selection process, we are able to specify a concise single tree that has a similar predictive performance to the forest of complex trees.

This article describes the results of our work in plain English. Having been submitted to an academic journal, it is currently subject to a peer review.[2]

To illustrate the results of this work with our heart disease example,[3] we compare one of the over 1,000 random forest trees against the single tree generated by the Bayesian process (see figure 1). Even one random forest tree – with 49 nodes – is likely to be unintelligible to a human, whereas the Bayesian Decision Tree has 16 nodes, making it easier for an expert to understand and verify.

As to the predictive power, the accuracy in assessing the presence of heart disease for random forests is 78.5%, compared with 83.0% for Bayesian decision trees.

2 Decision tree versus random forest versus Bayesian tree

Dataset	Dimensions	Samples count	Accuracy		
			Decision tree	Random forest	Bayesian decision tree
Credit	23	30,000	72.60%	78.10%	82.00%
Diabetic	19	1,151	62.60%	64.80%	63.50%
EEG	14	14,980	84.00%	88.60%	81.20%
Gamma	10	19,020	81.40%	85.60%	85.20%
Haberman	3	306	65.00%	68.30%	71.90%
Heart	20	270	76.30%	78.50%	83.00%
Ripley	2	250/1,000	83.80%	87.90%	87.60%
Seismic	18	2,584	87.70%	91.50%	93.20%

Accuracy of decision trees, random forest and Bayesian decision trees for several UCI datasets. Except for the Ripley dataset, we apply a tenfold cross-validation to each test.
Source: UBS

Beyond this example, we tested the predictive performance of Bayesian trees against both a single classical tree and a random forest – with generally favourable results – over a selection of eight standard machine learning datasets (see figure 2).

Our motivation to explore this topic is mainly driven by a need to have explainable, auditable models with similar predictive power to state-of-the-art black-box machine learning techniques.

In finance, and especially in algorithmic trading, this is of key importance to us as we develop new algorithms, but also to our internal control functions and our regulators (alongside the need to produce models capable of running queries with low latency).

As a practical example, we tried to tackle the problem of optimal order placement in our foreign exchange smart order router using a random forest to estimate the quality of each available trading venue, based not just on price, but also in terms of probability of execution and market impact. While the results were encouraging, execution decisions need to be as low-latency as possible and this was a drawback with the technique – querying many trees can be time-consuming. Just as importantly, we need to be able to inspect our routing decision and understand why we chose a specific venue for all our live orders. Bayesian decision trees allow us to solve both these issues without sacrificing accuracy.

Beyond finance, the machine learning world is also increasingly focused on explainability, which can be a prerequisite in fields such as medicine and self-driving, autonomous vehicles.

The machine learning community is built on both academic and industry contributions: here at UBS's Strategic Development Lab we opted to share our work, given that we believe its applicability goes beyond trading and financial applications. ■

Previously published on Risk.net

[1] *There are, indeed, plenty of such analyses – eg, this one: https://bit.ly/2WUAa2N*

[2] *G Nuti, L Rugama and A Cross,* A Bayesian Decision Tree Algorithm, *January 2019, https://bit.ly/2LUat1f*

[3] *The data for this example is taken from one of the standard problems in machine learning, available in the UCI ML repository found here: https://bit.ly/2HsLKgg*

Converging on sound MRM practices

Although most banks are progressing rapidly towards a certain standard in MRM practices, the rate of progress is uneven and so are the ambition levels. Management Solutions provides a summarised overview of the state of MRM evolution and how banks are striving to converge

When it comes to the level of maturity of MRM frameworks, a distinction can be made between at least three tiers of banks (see figure 1).

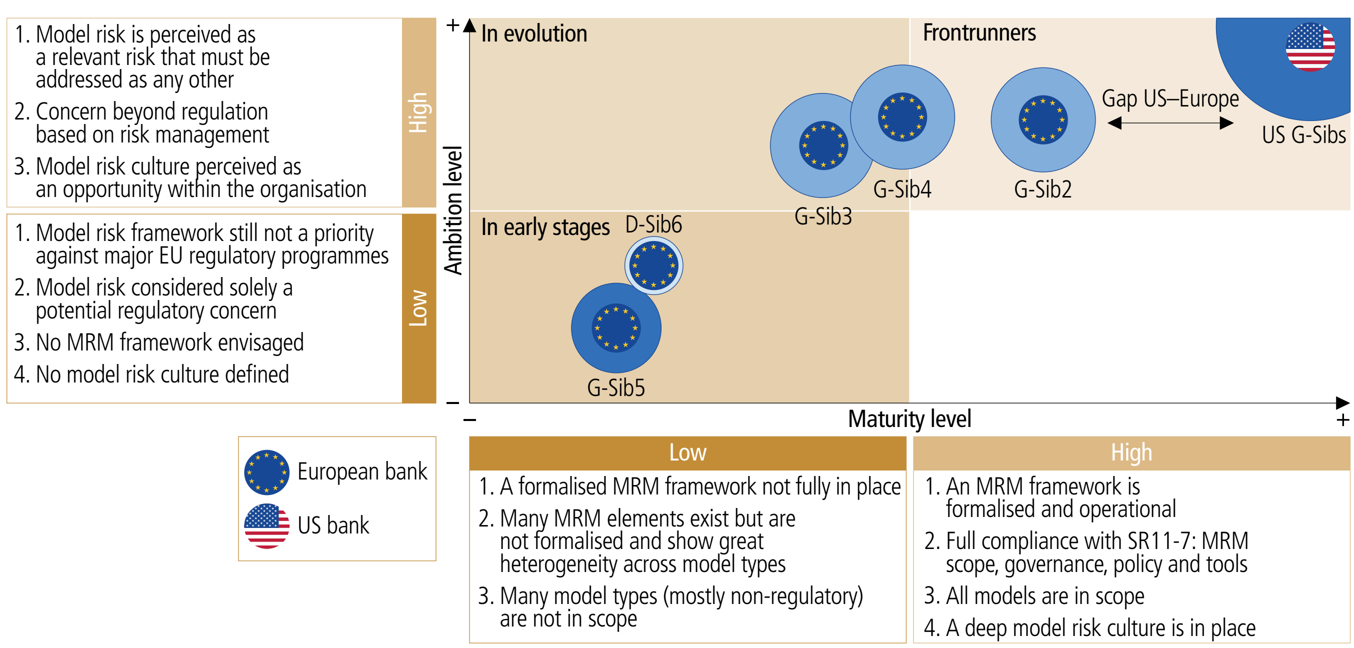
First, there are most major US banks, which have been subject to the US Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency's *Supervisory guidance on model risk management* (SR 11-7) for nearly eight years, and have been under major pressure to develop their MRM frameworks. This group's distinctive feature is their success in conceiving model risk as a concern beyond regulation. Their focus is to safeguard the bank, not merely to comply with regulation, although one leads to the other.

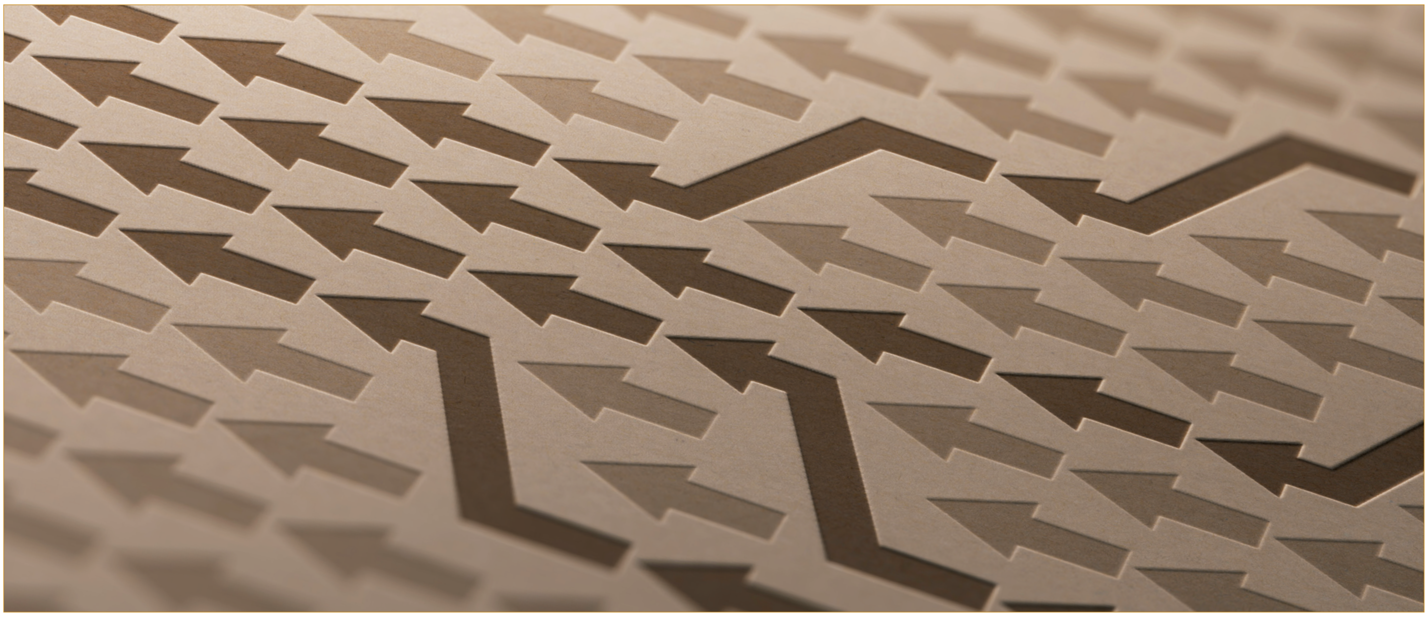
A second tier comprises major non-US banks with a significant presence in the US. They have also experienced the pressure to adapt to SR 11-7, but only in their US branches. They are bringing that experience upstream to their holding company and to other relevant subsidiaries. Even though their level of maturity is still behind their US peers – for example, their MRM scope tends to focus only on regulatory models – these banks are making rapid progress.

A third tier consists of all other banks, including medium-sized ones and major institutions that have prioritised other initiatives. Indeed, they are struggling with a tsunami of model-related regulatory requests – the Targeted Review of Internal Models (Trim), Fundamental Review of the Trading Book, International Financial Reporting Standard 9, new definition of default, and so on – and, since resources are limited, they have concentrated efforts on responding to these requests.

It is noteworthy that not all banks actually want to be on the exact same level of maturity regarding MRM. Indeed, with varying levels of regulatory pressure, organisational complexity and internal concern around model risk, the concept of ambition level regarding MRM becomes all the more important.

1 The MRM practices of global and local players





To this extent, with model risk being so transversal and MRM such a multidimensional activity, decision-making at a bank's desired level of ambition becomes a challenging task in itself. For that purpose, it may prove useful to have a set of MRM ambition indicators – for example, on MRM structure, policies, appetite – to determine how the bank relates to them (see figure 2). This is a high-level self-assessment tool, which is then adapted into a number of detailed statements that may be useful to periodically measure the level of evolution against an objective yardstick.

Finally, it is worth noting how banks are facing different challenges in their evolution towards a sound MRM framework, varying for the most part according to two axes: geographic footprint – and the applicable regulation – and size. Focusing on the latter, large institutions are concentrating efforts on ensuring consistent deployment of their MRM frameworks across countries and subsidiaries, appointing and empowering model owners, and fostering a model risk culture. Medium-sized institutions, meanwhile, are struggling to obtain sponsorship and the involvement of senior management despite reduced regulatory pressure. Additional challenges include inventorying all models in use or even defining what a model is.

Over the past few years, Management Solutions has observed a very positive increase in banks' awareness of model risk, which is rapidly translating into structured MRM frameworks. Although convergence is still distant and, to some degree, hampered by the heterogeneity of regulation across countries, the influence of SR 11-7, the European Central Bank's Trim project – which is already issuing findings on MRM through Pillar 2 – and US and European Union industry best practices seem to be leading the way for institutions worldwide.

2 Nine key indicators of MRM ambition

	Domain	Top MRM ambition indicators	Example of ambition level
1	MRM structure	• There is an MRM function in place: • Responsible for managing, controlling and reporting model risk • Reporting directly to the chief risk officer • Absorbing some/all of the internal validation functions • There is an MRM committee responsible for MRM-related decisions, besides the model approval committee(s)	
2	MRM policy	• There is an MRM policy defined: • Containing MRM definitions, principles, roles and governance, adapted into instructions • Approved by the board	
3	Roles and culture	• Every model has an appointed owner responsible for ensuring control of the model throughout its entire lifecycle • Senior management and the board receive regular model risk reports	
4	Model validation and approval	• Model validation covers all model-related components: conceptual soundness, implementation and use • Every new model use undergoes internal validation and approved at the relevant model approval instance	
5	Model risk appetite	• There is a clear model risk appetite defined, approved by the board and monitored through periodic risk reporting	
6	Tiering	• There is a group-wide model tiering procedure with clear drivers; all models in the group have an assigned tier, and tiers impact all model-related processes (such as the frequency and intensity of validation)	
7	MRM tool	• There is one global MRM tool implemented, including: • Global model inventory • Model workflow • Model documentation repository • Model risk reporting	
8	MRM scope	• The scope of MRM is all the models in the group, including regulatory and non-regulatory model types, with a phased-in approach	
9	Non-models	• There is a non-model policy, defining the management of those methods defined as non-models (such as expert judgement)	

Longer datasets to handle economic cycles

Decades, not years, of credit losses are required for accurate risk modelling, argues expert. By Alexander Campbell

Macroeconomic models used for forecasting and stress-testing frequently rely on too short datasets, and should be backtested against much longer sets of economic data – measured not in months or years but in economic cycles, according to upcoming research.

In a paper to be published in the *Journal of Risk Model Validation* in 2020, Joseph Breeden, the founder and CEO of risk modelling consultancy Prescient Models, argues risk modellers use methods that often underestimate the level of correlation in data. The higher the correlation, the less useful the dataset for model validation.

The paper puts forward a method of determining the amount of "structure" in a dataset; in other words how robust the data is for financial models. Breeden suggests using economic cycles as a unit of measurement, weighting the cycles by severity to give an objective yardstick.

An alternative way is to use economic recessions to test data, but this approach is flawed for two reasons, argues Breeden. There is little universal agreement on a definition for a recession, and some countries, such as Australia and China, have not suffered a recession in recent decades, which may lead analysts to wrongly conclude that data from these countries is invalid for modelling.

Models for risks linked to macroeconomic factors – such as credit risk, where default rates track the state of the economy as a whole – could be trained on as little as five years of monthly data and tested on one to three years of out-of-sample data. But this barely allows for a single economic cycle to be covered in the training data – when model builders should be aiming to cover several.

Breeden uses the example of US mortgage charge-off rates from 1990 to show that the correlation in data points is high, only declining to zero at 30 months. It continues to a strongly negative correlation after 60 months (see figure 1). Non-correlation of data points – what statisticians call "independence" – is crucial in creating and testing models, and shorter time series appear to undermine this independence.

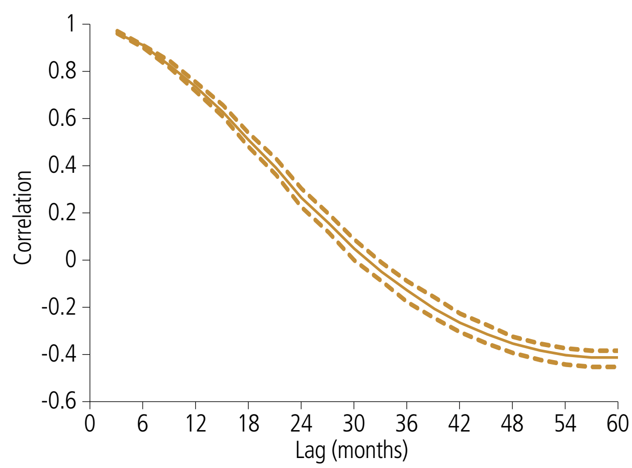
"If the sets of monthly data have been treated as independent when they are not, this seems like a serious problem," says Breeden.

In fact, he argues, autocorrelation of monthly data within an economic cycle means that each cycle only contains four truly independent data points: the high and low points, and the midpoints of the rise and fall.

High autocorrelation can make the p-values – a measure of the level of confidence in estimates – misleadingly optimistic. Some academics have tried to address this problem by using "robust" p-value estimation techniques, designed to take account of the greater uncertainty. But Breeden claims these "unstable" methods "don't address the bigger question of how good the dataset is for modelling".

And the same applies to attempts to segregate data geographically: moving from national to state-level US information, for example, will not produce 50 independent datasets, as there is strong state-to-state correlation in almost all cases.

1 Autocorrelation for US mortgage charge-offs



Note: Autocorrelation includes upper and lower 95% confidence bands. Charge-off data is from US Federal Reserve FRED website, 1990 to date.

Even the statistical technique of measuring effective sample size (ESS), which discounts the number of points to take account of autocorrelation, is found to be unsuited for cyclical data, since ESS was intended for use with datasets whose autocorrelation tracks to zero and doesn't veer into the negative.

To demonstrate the effect of the new approach, the paper takes a hypothetical US loan portfolio with five years of data back to January 2013. Breeden calculates the weighted number of cycles for the five-year period to December 2017 as 0.23, or one-quarter of an economic cycle. But extending the data back a further five years, to January 2008, causes the weighted number of cycles to leap to 2.01.

"Increasing the number of months of history by two times increases the amount of information in the data by 8.7 times," Breeden writes.

The bottom line for risk modellers, Breeden says, is that validating their models will require far more data than they are using at present – enough to cover multiple economic cycles for training and at least one for testing, with each cycle lasting several years (6.4 cycles since 1980 in US non-farm employment data, for instance, giving an average period of six years).

The findings won't necessarily undermine all model validation; operational risk, for example, is driven more by the tail of the loss distribution than by behaviour over time, Breeden says, so any cyclical behaviour will be less important. But for any risk linked with the macroeconomic cycle, the implication is that decades, not years, of reliable data will be needed for a properly robust model validation process. ■

Previously published on Risk.net

MRM transformation

Financial institutions have been maturing their approaches to MRM and – as models become more complex and pervasive, and regulatory expectations continue to increase – leading financial institutions seek faster and further movement. Ashutosh Nawani, head of financial risk management, and George Stylianides, global risk lead, financial services, at PwC Risk Consulting, explore why, by revising their approaches to MRM, institutions are targeting flexible, adaptable and efficient MRM functions that are fit for the future and will deliver real business value

Future-proofing MRM

Within the next three years, leading MRM functions will comprise three pillars as high-performing, enterprise-wide business enablers:

- Holistic mandate-encompassing diverse skills, and embracing technology, regulation and analytics
- A seamless venturing between modelling, reporting and monitoring
- An agile, integrated and responsive framework to organisational and regulatory change.

To achieve such mature framework, transformation enablers are required across a number of areas. Firstly, processes and policies need to change to reflect models' complexity, materiality and breadth, which will determine the degree of oversight and governance required. Secondly, more automation will be introduced to standardise and industrialise model validation procedures and libraries to drive rapid assessment and deployment. Finally, approaches such as robotic process automation (RPA) can execute routine and repetitive tasks far more efficiently.

It is imperative for institutions adopting new technology approaches to support an enterprise-level centralised data platform. This will act as a single source of truth for model risk activities such as validation, monitoring and reporting. Furthermore, integrated solution tools can link MRM components to enhance the automation, effectiveness and transparency of governance and oversight. And improvements to MRM operating models should secure cost effectiveness and efficiency by moving from a solely in-house function to a hybrid model that leverages strategic outsourcing partnerships.

Financial institutions expect to gain significant value from accelerating their journeys to a next-generation, more simplified, automated MRM function. There are a number of key opportunities to take action and start making progress, embracing policies, processes, technology and people. Clarity of roles and responsibilities and the alignment of key model management activities is as important as the appropriate use of digital technologies such as RPA, machine learning and advanced analytics. Cultivation of the right talent mix and culture is vital, as is the development of a common data platform.

Ashutosh Nawani

George Stylianides

Model risk for a new wave of models

Following the recent Prudential Regulation Authority supervisory statement on algorithmic trading models, banks are going to face challenges related to the MRM of this new wave of models.

Algorithmic trading, machine learning and artificial intelligence (AI) models allow financial institutions to process a broad variety of data types – both structured and unstructured – and it can be hard to test the integrity and appropriateness of the data used by these algorithms. Consequently, the first challenge is that identification of risks arising from errors in data, its use and its quantification is still in an early stage, but rising very quickly to the top of MRM agendas.

Another challenge MRM managers face is driven by the assessment of the conceptual soundness of these models. This is more complicated than that of standard models, since algorithmic trading/AI/machine learning models are not yet thoroughly understood by practitioners. Therefore, it is hard to evaluate how fit for purpose a given model is for a particular task.

If a model has had a strong performance in predicting future outcomes of a certain quantity up until today, going forward one can never be sure that the model will continue to be fit for purpose when the market enters a new phase.

Consolidation of MRM framework is therefore key for algorithmic trading/AI/machine learning models. The first step is to empower governance framework. The next step is to reassess the definitions of models within the bank to ensure all new model types are covered. These new categories of models should be listed in a dedicated model inventory, including associated controls and other mitigants – kill-switch procedures, for example.

Furthermore, new and existing machine learning and AI models – or other new types in general – should be forced to follow the same steps of testing as other more established model types. Subsequently, controls should be implemented to cap exposure to a counterparty, order attribution, message rate, frequency of order, stale data, and order and position size. Lastly, specific quantitative tests to these models should be implemented. For instance, additional stress-testing of IT systems in dynamic testing environments may be required to understand the impact of extreme events and avoid future disaster in the event of IT system errors.

In conclusion, the ultimate goal is an MRM framework fully embedded in business decision-making that enables proactive risk management for any type of model. While every financial institution will embark on different journeys based on different strategic visions, in the context of current capabilities and regulations the desired final destination remains common to them all. ■

Tech-driven MRM

Driving value and efficiency

Next-generation MRM tools such as automated model documentation, performance monitoring and powerful network visualisations can improve efficiency by minimising time-consuming tasks and allowing organisations to focus more on high-value activities. A forum of industry leaders discusses the considerations and opportunities of technology-driven MRM, how emerging technologies are changing banks' approaches to MRM, and the impact of regulation going forward

David Asermely
Global Lead
SAS Model Risk Management
www.sas.com

What impact is growth in financial models having on banks?

David Asermely, SAS: It's an increasingly competitive analytical world, and having better information differentiates top banks from the pack. As a result, financial institutions are relying more on models to make critical business decisions – from assessing risk and capital planning, to investment management and financial reporting. Models also play a critical role in regulatory and accounting compliance processes, such as performing stress tests, calculating expected credit loss – International Financial Reporting Standard 9 and Current Expected Credit Loss – and conducting anti-money laundering investigations. Banks are responsible for developing, supporting and using their valuable models appropriately, so they are spending more resources on modelling activities.

Ashutosh Nawani, PwC: In the case of trading algorithms that result in fast-paced, high-frequency market activity, this can lead to equally swift accumulation of losses. With multiple market participants utilising algorithms that may not have undergone sufficient testing and validation, there is a potential systemic model risk that could result in large-scale market abuse and a financial disaster.

Where these new types of models sneak into the model inventory – and as a result face rigorous model risk policies – there may not be relevant expertise within the model risk function to perform a sufficient review. Or, more often, there may not be capacity or resource to bring them the governance framework as efficiently and as quickly as possible.

Regulators have certainly shown they are keen to try to solve the problem, with the UK Financial Conduct Authority and the Prudential Regulation Authority (PRA) conducting reviews across firms and publishing guidance and supervisory statements earlier this year regarding algorithmic trading.

First, a firm is expected to explicitly approve the governance framework, bodies, controls and policies for algorithmic trading. This includes the identification and empowerment of the specific management body to manage the model risk of such models with resources with suitable expertise. The next step should be to reassess the definition of models within the bank to ensure these new model types are covered. These new categories of models should be listed in a dedicated model inventory, including associated controls and other mitigants – for example, kill-switch procedures.

Slava Obraztsov, Nomura: Defining the scope of the MRM process has always been quite a difficult challenge. Definition in the US Federal Reserve's Supervision and Regulation Letter (SR) 11-7 is quite broad and, for example, the PRA's SS3/18 MRM principles for stress-testing provide only some directions as to how models should be defined. It is therefore important for financial institutions to find a balance between how wide the applied scope should be, while preventing MRM from becoming a box-ticking exercise. Nomura concentrates on models where the definition includes at least one model assumption that could be an expert judgement. Similar approaches are employed by others where banks try to distinguish between models and tools, and different processes are applied in those situations. Recently, some regulators, such as the PRA, have made no significant distinction between models and tools by requiring the risks associated with the implementation of all types of calculations to be adequately understood, controlled and documented. As a result, financial institutions need to adjust their MRM approaches accordingly.

Anne-Cécile Krieg, Societe Generale: The main change I can see regarding traditional financial models such as asset-liability management, valuation or credit is a cultural shift in the way they are approached. First, models are not left mostly to quants any more, as the shift now involves all the stakeholders in the model value chain, including IT personnel and bankers; second, all phases of the model lifecycle are under scrutiny, with an increased focus on robustness in stressed conditions. Most of the growth in models is coming from non-traditional areas such as back-office processes, which increasingly rely on automation and compliance.

Konstantina Armata, Barclays: Banks typically use a few thousand different models as part of their operations. Examples include models used in the front office to price transactions or to determine borrowers' creditworthiness; models used to calculate risk-weighted assets (RWAs), liquidity, stress or provisions by risk and finance; models used by compliance to detect fraudulent activity; and even models used by researchers to assist stock evaluation. The sheer number of models and their varied nature requires extensive MRM capability and skill set, as well as a pragmatic approach to determine prioritisation and depth of coverage. The end-goal is the development of a holistic and coherent risk management framework that includes a view of model risk across the organisation that can be clearly communicated to senior stakeholders and drive decisions.

pwc

Ashutosh Nawani
Head of Financial Risk Management
www.pwc.co.uk

What are the main emerging risks in this area?

Ashutosh Nawani: The emerging risks underpinning model risk in the next two years are a direct function of the reshaping of the regulatory landscape – the Fundamental Review of the Trading Book, interbank offered rates and the standardised approach to counterparty credit risk regulation – and the adoption of new technologies.

Institutions in the early stages of defining their model risk framework must overcome the primary hurdle of poor clarity on lifecycle roles and responsibilities across the three lines of defence. This increases the risk of incorrect design and consequently a lack of (full) compliance with regulatory requirements.

Once the framework matures into a foundational stage, institutions seek increased compliance and operational costs due to cross-border governance as well as levelling out inefficiency and the cost of human capital sourcing. A key risk to address to avoid stagnation on this suboptimal front is found in the correctness of the implementation of a 'model validation factory'. This risk will be mitigated by developing standardised model validation procedures and libraries for rapid assessment and model deployment.

Institutions with more robust and established model risk frameworks are being challenged by the risk of a lack of fit technology supporting an enterprise-level centralised data platform as a single source of truth for modelling, validation, inventory, monitoring and reporting.

Slava Obraztsov: There have been only a few new product developments in the past several years. The model universe is changing, but models are applied to broadly the same product sets, and the same products continue to cause the greatest challenges. For example, constant maturity swap (CMS) spread corridor options, which caused significant losses in 2008 due to euro curve inversion, are becoming popular again. This time, the activity is related to the flattening of the US dollar curve, where the swaps and options market is much bigger, but challenges with managing CMS spread digital risk remain the same as in 2008.

Other examples include collateralised loan obligations, which are difficult to manage from a modelling perspective, and more complexity in so-called 'simple' products. Increasing regulatory and industry attention continues to be paid to the treatment of the general wrong-way risk and the development of valuation adjustments – collectively known as XVAs.

Model interconnectedness is another issue where model risk in one area could be propagated to models in other areas. It is especially relevant to so-called 'feeder models' – scenario-generation models in stress testing and proxy time series used in capital and exposure calculations. For a proper analysis of model interconnectedness, it is important to have a robust model inventory, including a detailed record of the applications of all model outputs.

David Asermely: Regulations and business requirements are driving banks to develop and support more models. As their model inventories grow, banks may struggle to understand the context within which their models interact; likewise, the increasing level of interconnectedness between models makes it difficult for banks to assess their true model risk. Banks are also deploying more complex models – including machine learning techniques – that are less transparent. Model volume, interconnectedness and complexity are stressing many MRM teams and systems.

Anne-Cécile Krieg: Many actors are focusing on artificial intelligence (AI) and machine learning. These are key technologies that should be specifically managed, but consistently, within the overarching framework. Another emerging risk is the potential for 'rogue modellers' and the difficulty of identifying all types of models that should be brought under an MRM framework. It's becoming easier to develop models, especially with open-source code. Even a non-specialist can develop one, so there is a risk these models remain outside the reach of the MRM process. This is mitigated by strict controls of the model inventory process.

ManagementSolutions
Making things happen

Javier Calvo Martín
Partner, Management Solutions
www.managementsolutions.com

Whose responsibility is MRM?

Javier Calvo Martín, Management Solutions: It is a common misconception in some banks that MRM is exclusively the responsibility of internal validation (IV) or, in a slightly broader view, risk areas.

Indeed, risk and IV carry a relevant share of the responsibility, but it is essential to acknowledge that model risk is very transversal – just like operational risk, for example – in terms of stakeholders and model lifecycle coverage.

This means the entire organisation should be involved in the appropriate management of model risk. In best-practice banks, this is achieved through the allocation of responsibilities across three lines of defence, where the first line of defence (1LoD) contains model ownership, development, monitoring, implementation and use; the second line of defence (2LoD) holds model governance and model validation; and the third line of defence is model audit, controlling the entire model lifecycle. Interestingly, in some countries, banking supervisors – such as the European Central Bank (ECB) or the Fed – are referred to as the fourth line of defence and external auditors as the fifth line of defence.

This implies a deep cultural change, especially when it comes to the model owner role. Model owners are typically close to or within the business, and may sometimes be model users. In many cases, it is not at all straightforward to raise their awareness of how critical their role is for an appropriate MRM, especially in terms of responsibility and accountability. So, while the 2LoD is usually more mature in this respect, there is typically more work to be done in the 1LoD.

Management Solutions is already observing a steady increase in awareness of the transversal quality of model risk in banks, notably in non-US countries, and alongside a consistent increase in responsibility and accountability regarding MRM.

Anne-Cécile Krieg: In theory, this hasn't changed: all three lines of defence are responsible for MRM, but operationally the mindset has shifted. In the past, it tended to be left to the 2LoD, but now there are specific roles allocated across the three lines, which are fully embraced and embedded.

In particular, within the 1LoD, all roles are better identified – from the designer of the model and the person implementing it, to the users and those tasked with surveillance, with increasing emphasis on the model owner role.

Slava Obraztsov
Global Head of Model Risk
Nomura
www.nomura.com

Slava Obraztsov: The aim of MRM is to develop a framework around the full model lifecycle, rather than concentrating on validation, as was previously the case. It's a fundamental change, but who is responsible for it now?

MRM has introduced a relatively new concept of 'model owners', but allocating that ownership and its associated responsibilities could be a difficult task. In my experience, nobody volunteers to be a model owner, but it's an important part of the process. There's also a shift towards widening the responsibilities of model development areas, especially around model documentation, standards of testing and model risk analysis. The current requirements are very different from what was happening 10–15 years ago.

There is more pressure on validation teams in terms of the completeness of the model validation process, including documentation and additional analysis, attention to governance processes and involvement of senior management. Firms' leaders don't need to be modelling experts, but they need to be aware of major model risks. Well-designed model risk quantitative and qualitative metrics would help senior management make strategic business decisions or prioritise some model developments if model risk becomes excessive.

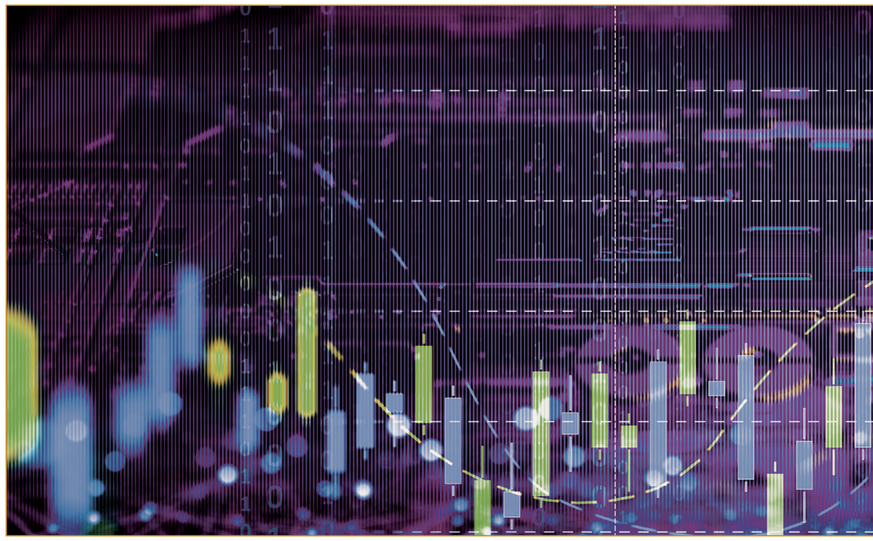
David Asermely: Most banks employ a multilevel approach to MRM. Model owners, developers and validators comprise the 1LoD and are responsible for producing and supporting well-functioning models. Typically, the 2LoD is the MRM group, which is responsible for crafting and enforcing policy, aggregate reporting and assessing model interconnectedness risk. Internal auditors provide an independent check to confirm the firm's policy enforcement. Ultimately, the board of directors is responsible for understanding and properly managing model risk. The Fed's SR 15-18 states: "The board should direct senior management to provide information about a firm's estimation approaches, model overlays and assessments of model performance" for all models within the capital planning process. Other regulations make similar assertions.

Konstantina Armata: As with all risk stripes, responsibility for MRM lies both with the 1LoD – mode users/developers – and the 2LoD – MRM.

How are banks' approaches to MRM evolving?

Javier Calvo Martín: There are different starting points and varying regulatory pressure depending on region – the US versus the European Union, for example. Acknowledging this as common ground, Management Solutions has found that banks' MRM frameworks are evolving along several dimensions:

- Formalising MRM governance, organisation and policies – specific model risk committees are being set up, MRM areas are being created and are absorbing IV units, and the whole MRM framework is being formalised in policies and procedures approved at the appropriate levels up to board level.
- Inventorying all models and deploying a model governance tool – best-practice banks have already inventoried all models, not only regulatory ones, and have deployed an MRM tool – whether a vendor system or internally developed – which acts as the backbone of the MRM process.



- Widening scope of the MRM framework – banks are trying to avoid a 'big bang' and are thus progressively incorporating all model types under the MRM scope (starting with regulatory models) and all business units (starting with holding and major subsidiaries). Lastly, and especially in the US, major banks incorporate non-models within the MRM scope.
- Developing model risk quantification and appetite – while still open in terms of methodology, most banks agree that measuring model risk is necessary, both in terms of model governance key performance indicators (KPIs) (for example, the number of non-validated Tier 1 models) and inherent model risk KPIs (for example, confidence interval on model outcome). This links strongly with model risk appetite, which is commonly based on these KPIs.
- Deepening the scope of the IV review – IV reviews the entire model lifecycle (data, development, implementation, monitoring and usage), and its reviewing techniques and tests are wider and deeper.
- Strengthening the link between models and data – the interdependency of models and data leads to a parallel development of the MRM framework and the data governance framework, including the Basel Committee on Banking Supervision's standard 239. In practice, this results in parallel governance structures and fluent communication between them.
- Increasing concern about advanced machine learning and AI models – due to their difficult-to-interpret nature, these models pose unique challenges to their MRM, especially regarding IV activity. Banks are developing specific procedures and techniques such as enhanced sensitivity analyses to address them.

Above and beyond this, probably the major achievement in the evolution of MRM frameworks is deploying a model risk culture across the organisation, involving the whole range of stakeholders across the three lines of defence, the board and senior management. If we were to choose a single acid test of MRM evolution, the effective implementation of a sound model risk culture across the organisation would be it.

Anne-Cécile Krieg: Banks want to address both their internal risk management needs and supervisory requirements, so there is a clear need for an industrialised approach with a focus on addressing these expectations in a modern way. What is striking is the increasing awareness of model risk among senior management.

Konstantina Armata
Global Head, Model Risk Management
Barclays
www.barclays.co.uk

Konstantina Armata: Most banks have established a holistic MRM function as the bank's 2LoD risk management capability for this type of risk that covers all models in the bank, although the level of maturity of this function may differ across the industry. This is a long way ahead of the starting point, where banks viewed MRM simply as one-off model validation, and even that was only for a couple of model types such as valuation models and, later, capital models. Further evolution will involve model risk being assessed in a holistic manner and ideally quantified in metrics that are well understood and, thus, enable decision-making.

Slava Obraztsov: Currently, one of the biggest challenges is quantification of model risk – at an individual level and in aggregate. For individual models, this technical assessment should include dependence on modelling assumptions and uncertainty around opaque parameters, model stability, and so on. For senior management reporting, there should be some type of model risk aggregation across different model types. How that could be achieved is hotly debated, but aggregation shouldn't be applied to all model types. For example, aggregation of model risk across pricing models and retail models may not be very meaningful. Instead, model risk should be aggregated across relevant types of models, such as pricing models or capital models, with a set of numbers to be checked against model risk appetite and presented to senior management.

Another trend is for building a robust process for model performance monitoring. It's important to have the right set of tools and indicators to monitor model performance under current market conditions and a portfolio composition to proactively identify and mitigate model limitations.

David Asermely: Historically, firms have undertaken MRM at an individual team level – often with tools such as Excel spreadsheets, PowerPoints, and governance, risk and compliance solutions. We are seeing, as the number of models increases, inventories growing more complex and organisations dedicating expensive resources to the development and support of these high-value assets, a shift to modernised MRM solutions. These platforms offer many efficiencies that reduce the burden on modelling teams while enforcing best-practice governance. A good example of this is performance monitoring. Regulators require financial organisations to understand how well a model is performing. With models and data changing more frequently – especially with some machine learning techniques – organisations must conduct performance monitoring more frequently. Previously, MRM teams did these types of backtesting manually. Now, however, MRM solutions provide automated performance monitoring with threshold alerts to streamline these activities and allow for better transparency.

Ashutosh Nawani: Model risk's fast-paced journey from a tertiary and seldom-considered risk to a strategic one of equal importance to credit, market, liquidity and other strategic risks has only just begun.

Long-term focused initiatives are under way at leading institutions to transform the MRM's structure and strategy to create a flexible, adaptable and efficient MRM function.

A constantly evolving area in model risk is the process and policy changes space. The aim is twofold. First, to more efficiently manage models based on their complexity and materiality, and avoid excessive oversight and governance driven by models' risk profile and characteristics. Second, to simplify through optimisation and centralisation model development and validation activities such as group-level refined standard methodologies, templates and documentation.

Depth focus is currently put on automation as institutions need to move towards a 'model validation factory' and to industrialise standard routine and repetitive tasks for model calibration, monitoring and reporting.

Last but not least, it is critical for financial institutions to select the most appropriate framework to employ fitting technology supporting an enterprise-level centralised data platform as a single source of truth for the functions I previously mentioned and to implement workflow management tools to connect MRM components for enhanced and transparent governance and oversight.

How effective have regulatory initiatives such as SR 11-7 and the Targeted Review of Internal Models (Trim) been in improving standards?

David Asermely: MRM regulations have forced many organisations to define and execute comprehensive policies that encompass robust model development, validation, implementation and usage. To comply with regulators, banks need a reliable MRM structure to ensure all risk categories related to models are identified, monitored and controlled. Regulators want to know whether organisations can answer these simple questions about each of their models:

- Has it been reviewed for conceptual soundness?
- What was it designed to answer?
- What are its limitations and assumptions?
- How is it being used?
- Is it performing well?
- Does the organisation have sound and appropriate documentation?

Konstantina Armata: SR 11-7 established model risk as equivalent to other risk types such as market and credit risk, requiring a comprehensive framework for ongoing management of the risk against defined limits and risk appetite. This is very different to the pre-SR 11-7 era.

Javier Calvo Martín: Although some banks have seen beyond regulation when launching internal MRM programmes, it is fair to say SR 11-7 has triggered most of the banks' initial efforts in this domain in the US and also abroad. From Management Solutions' perspective, the Fed and the Office of the Comptroller of the Currency (OCC) are carrying out essential work in responding to industry concerns – through industry conferences, for example. Furthermore, the influence of SR 11-7 and SR 15-18 in other countries' regulation – for example, the Polish Financial Supervision Authority's Recommendation W – and in banks' best practices is remarkable.

Meanwhile, in the EU, MRM is being approached in a different manner, though not as specifically as in the US. SR 11-7 has no equivalent and the ECB's guide to internal models only states that "institutions should have an MRM framework in place that allows them to identify, understand and manage their model risk for internal models across the group". The guide provides seven items that need to be covered, and has so far proceeded no further. In practice, the Trim project reviews many areas belonging to MRM frameworks – for example, model governance, organisation, documentation and policies – for Pillar 1 models.

Also noteworthy is the PRA's success in enhancing MRM standards in the UK through a principles-based regulation on MRM, for now aimed at stress-testing models.

In other words, it is unquestionable that regulatory initiatives are effectively improving – or even triggering – MRM standards in banks. However, this is being undertaken from different perspectives: a top-down approach in the US and UK, starting with the need for an overall MRM framework and drilling down to specific components; and a bottom-up approach in continental Europe, with the Trim project assessing many MRM-related components, departing from each individual model under inspection, which should ultimately lead to building up a holistic MRM framework.

Slava Obraztsov: Since the introduction of SR 11-7 in 2011, significant progress has been made in the development of MRM processes. This includes extending the scope of covered models, strengthening requirements for the quality of model development and validation documentations, introducing firm-wide model governance frameworks – such as committees, policies, and so on – establishing model inventories, and supporting model management workflows.

Initially, MRM requirements mainly affected large firms regulated by the Fed and the OCC. However, other regulators – such as the ECB with Trim, and the PRA with SS3/18 – have gradually aligned their model requirements to the major elements of SR 11-7. In some cases, regulators have adopted even stricter requirements.

I've found SR 11-7 and Trim requirements slightly different. While the former is based on high-level principles around the three lines of defence and major MRM requirements, ECB requirements are more specific. The scope is narrower and applied only to different types of risk models, but there are more specific expectations of what types of analysis must be performed or what reporting provided. Despite those requirements being not yet fully finalised, some European regulators have started to apply them in their inspections of financial institutions.

Global financial organisations face quite a challenge to satisfy the requirements of different regulators, which are not always fully consistent. It is clear regulators would often like to see more region-specific approaches as this would allow them to have more control over applications of MRM framework. On the other hand, if MRM approaches become too fragmented, it would become impossible for international institutions to run them efficiently and consistently.

Anne-Cécile Krieg
Deputy Head of Model Risk Management
Societe Generale
www.societegenerale.com

Anne-Cécile Krieg: SR 11-7 is a historical cornerstone of MRM and Trim is a relevant exercise, so both have proved to be the basis for internal review and improvement of our set-ups.

Which MRM areas offer the greatest potential to drive value and efficiency?

Konstantina Armata: The well-established ones such as MRM for valuation and RWA models. The know-how of these areas can be shared across other model types and testing automation can enable better utilisation of resources and cost reduction.

Ashutosh Nawani: As with any other strategic risk, the lifecycle of MRM offers opportunities for creation, enhancements and optimisation. Breaking it into granular components, organisations are thinking of the following key elements:

1. Globally, institutions are looking to re-examine their MRM function; the current operating model is deemed unsustainable, considering the ever-growing number of models, their increasing complexity and heightened regulatory expectations. In particular, the greatest areas to drive value and efficiency can be grouped into the following three categories. Banks want to embrace an intelligent risk-based approach to MRM operations. A risk-based MRM framework and related methodological approaches enable tailored oversight, rigour and levels of effort linked to the model materiality, complexity and the relative impact on the business. This forms the foundation for simplifying processes, eliminating excessive governance and enabling a more cost-effective, scalable operation.
2. Institutions aim to leverage technology to maximise automation opportunities. Industrialised model management processes that are conducive to automation and strive towards centralisation by implementing comprehensive workflow systems to manage MRM, model validation and development, and interactions across the model lifecycle to enable efficient alignment of MRM standards and practices across the organisation.
3. Institutions' goals are shifting to create a sourcing talent strategy for capacity and scalability. They aim to create a talent strategy that minimises irregular workloads, scarcity of resources and mismatch of skill sets by building centres of excellence or shared services teams, complemented by managed services, to allow resource allocation synergies while maintaining consistent MRM practices.

Anne-Cécile Krieg: Basically, managing model risk avoids losses. More importantly, managing model risk paves the way for the future: with the acceleration of the digital transformation comes the management of the risks involved with new technologies.

David Asermely: SAS recently held an MRM customer connection event. Clients identified automated model documentation assistance as a major opportunity to improve efficiency. Implementing automated documentation requires a single source of truth and consistency across an organisation. Such an effort takes considerable time and resources. Reducing those manual efforts – performed by many PhD-level individuals – immediately increases value and efficiency. Other potential areas for improvement include robust performance monitoring, automated model usage capture, flexible reporting and powerful network visualisations. These next-generation MRM tools minimise repetitive mundane tasks and allow organisations to dedicate more resources to high-value activities. They also provide important insight on how models are used and in what context. This information is becoming increasingly valuable as many model risk practitioners and regulators shift their focus from individual models to model ecosystems.

Slava Obraztsov: A proper MRM process around the full model lifecycle – including model design, development, validation, approval and reporting – should be able to significantly optimise model management and use, and should thus be considered a benefit, not a cost.

In particular, model quantification and model interconnectedness can drive value. In addition to my previous comments on these topics, I'd add that model interconnectedness should impact the whole model approval process. For example, consider a change of pricing model or a model introduced for a new product type: in this case it is not sufficient simply to understand how it would affect the value of the product or its hedging, but also the impact on other areas such as capital and exposure calculations, independent price verification processes in the finance area or liquidity risk management in the treasury. It should therefore be properly understood how changes to or the limitations of one model affect others. All of these elements should be combined into a single comprehensive product approval process. As previously discussed, robust model inventory and product taxonomy are key to the management of model interconnectedness.

To what extent have banks been able to take advantage of new technology and analytics such as machine learning?

David Asermely: Machine learning models are proving very valuable in model validation, feature selection and benchmarking. Yes, models validating models, or even models validating themselves. Of course, this complexity and lack of transparency adds another layer of risk, including the potential of survivor bias to propagate models that 'lie' about everything being fine. However, there are opportunities to better automate model validation and identify risks that humans would not be able to see.

Slava Obraztsov: Technological development as a whole has been very important for MRM. Some areas are becoming quite intensive and simultaneously customised – some elements of validation, reporting or monitoring, for example. With the proliferation of models, more tools need to be put in place just to run proper processes and controls. It includes appropriate inventory, building tools for performance monitoring and model risk analysis, and automation of generating model development and validation documentation.

There are many applications in MRM where machine learning can be deployed, typically for processing large and less-structured datasets. In particular, machine learning can improve the quality and consistency of input data. It could also be an important tool for stress-testing, defining stress scenarios and propagating those scenarios across all underlying risk factors. If propagation is undertaken in a simplistic way, it may lead to inconsistent building of interest rate curves and volatility surfaces, which would lead to the breakdown of many underlying pricing models and an unreliable stress test output.

Anne-Cécile Krieg: New technology and analytics currently have two main applications. First, to develop models to optimise internal processes and some compliance-related topics. Second, to challenge existing models, particularly in terms of performance.

Ashutosh Nawani: Technology is at the heart of MRM. This is widely observed in different activities varying from the aim of a universal data system for enterprise MRM activities – development, validation and governance – to the consolidation of model risk assessment and issues management systems. The ultimate objective is optimised key processes across model risk capability areas with integrated data taxonomies for process, products, risks and controls. In other words, robotics process automation in monitoring, validation and testing, and reporting – and banks leveraging technologies and analytics in remediation processing.

Konstantina Armata: This is a very promising area, but is still in its infancy.

What are the main considerations in adopting these technologies?

Anne-Cécile Krieg: First and foremost, the MRM framework needs to be applied to all models, regardless of the technique – from classic statistical or stochastic approaches to new AI methodologies. This all begins with identifying the models and registering them within the model inventory. Regarding machine learning, specific considerations such as 'over-fitting' and black-box risks must be overseen and carefully monitored over time.

David Asermely: Machine learning models need governance just like any other models – only more so. Some machine learning models improve automatically through experience. This is a desirable feature, but it also exacerbates an organisation's model risk. Thus, there is an increased need to define operating controls on inputs (data) and outputs (model results). The dynamic nature of these technologies requires more frequent performance monitoring, constant data review and benchmarking, better contextual model inventory understanding, and well thought out and actionable contingency plans. Machine learning models are rapidly moving toward broad financial industry adoption. These models offer better predictability potential in the face of a global, multidimensional marketplace and the increasing volume and complexity of data, but this complexity and lack of transparency threaten to overload traditional model risk guardrails and governance practices.

Konstantina Armata: It's a paradigm shift. New skill-set and mind-set adaptation is required, as well as a functioning and efficient IT infrastructure.

Slava Obraztsov: It is often a matter of funding and resources. But these new technologies are unavoidable and will play an increasingly important role in financial modelling, including capital optimisation and algorithmic trading.

There are currently no industry standards for model validation and performance monitoring of machine learning models, and even regulatory requirements are not specific and are usually covered by general concepts such as 'effective challenge'. Undoubtedly, we are going to witness active application of machine learning approaches to MRM in future, so I think more refined regulatory standards – but not necessarily new regulation – will need to be developed soon. ■

>> The questionnaire respondents were speaking in a personal capacity. The views expressed by the panel do not necessarily reflect or represent the views of their respective institutions.

Valuation model risk on the rise at EU banks

More than two-thirds of fair value assets are priced using banks' models. By Louie Woodall

The share of European bank portfolios subject to model risk is on the rise, partly due to the new system of accounting for credit assets introduced at the start of last year.

At end-June 2018, €5.2 trillion ($6 trillion) of assets were classified as either level 2 or level 3 under International Financial Reporting Standard (IFRS) 13 fair value measurement accounting rules, according to data published by the European Banking Authority (EBA).[1] This means they lacked quoted prices in active markets to determine their valuation, and were priced using banks' own models instead. This is up by €453 billion (8%) from end-2017.

Level 2 and 3 instruments accounted for 67% of total bank assets at end-June, up from 63% at end-2017. Their share of the total had previously been trending downwards from end-June 2016 to end-December 2017.

The EBA said the increase was likely due to the rollout of IFRS 9 accounting standards in January 2018, which required banks to recognise certain assets that may previously have been valued at amortised cost at fair value instead.[1] Those newly re-designated assets without active markets would then have been classified as level 2 or 3 as a result.

Norwegian banks had the highest percentage of fair value assets designated as level 2 or 3, at 90% of their aggregate portfolios. Romanian lenders had the lowest, at just 1%.

The majority of European bank level 2 and 3 assets at end-June 2018 were assets held for trading, making up 66% of the total. Fair value assets measured through other comprehensive income increased as a share of total level 2 and 3 assets from 17% at end-2017 to 23% at end-June 2018.

What is it?

Financial assets are typically valued at either fair value or amortised cost in banks' accounts. Fair value assets are priced at their estimated potential market price. Those calculated in reference to actual quoted prices are designated level 1, those to other observable inputs level 2, and those to unobservable inputs level 3.

Those assets whose price fluctuations feed through into a bank's income statement, such as assets held for trading, are classified as fair value through profit or loss (FVTPL). Other instruments, like those designated as available-for-sale, have their valuation shifts recorded in other comprehensive income and are classified as fair value through other comprehensive income (FVTOCI).

Assets held at amortised cost are accounted for at their initial purchase price, minus principal payments, any impairment losses, and foreign exchange differences. Assets held-to-maturity, like loans and certain bonds, are typically valued at amortised cost.

Data in figures 1–3 is taken from the EBA's 2018 Risk Assessment Report, based on supervisory data submitted to the EBA on a quarterly basis for a sample of 187 banks from 25 European Economic Area countries.

Why it matters

IFRS 9 was founded on the concept that most assets should be measured at fair value, and therefore tightened restrictions on those instruments that could be measured at amortised cost.

Now, assets held to maturity whose cashflows are not related solely to payments of principal or interest (SPPI) must be valued at FVTPL or FVTOCI. Clearly, a huge number of EU bank assets failed this SPPI test on IFRS 9's rollout, and a hefty chunk of these could only be fair valued using non-market inputs.

On the one hand, that more assets are being valued at the price a third party would be willing to pay for them is a positive development, as an asset held at amortised cost may not fetch anything like its book amount – especially in the midst of a crisis – if put up for sale. Such assets may prove the source of huge losses in a fire-sale scenario.

On the other hand, if a large share of these newly designated fair value assets are priced according to banks' own models, estimates and judgements, then they are susceptible to gaming. An inadequately calibrated and monitored valuation model could inflate the fair value of an asset far above the price it could fetch at market, leaving a bank in the same spot they would have been in a crisis had the assets remained at amortised cost.

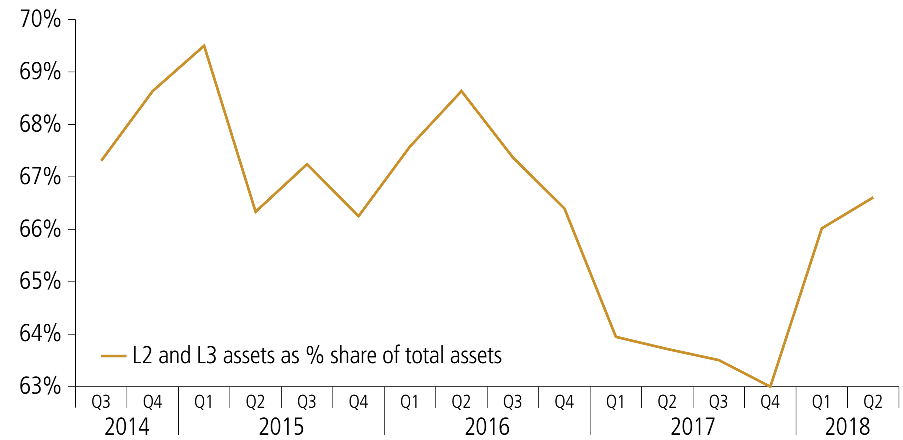
Previously published on Risk.net

[1] *EBA*, Risk assessment of the European banking system, *December 2018, https://bit.ly/2ECngAW*

Get in touch

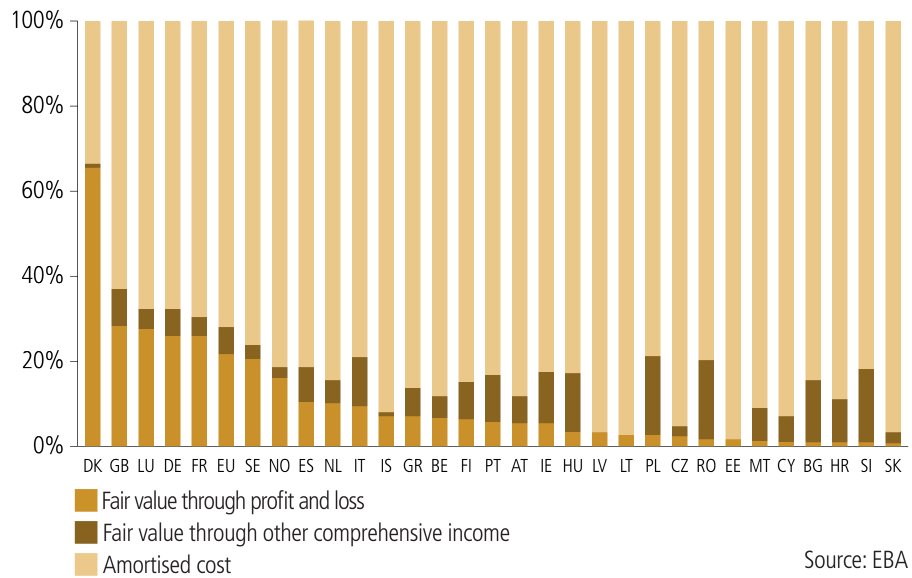
If a larger share of banks' portfolios are valued using their own models, the greater the amount of assets that could experience violent price swings in a crisis. If this concerns you, get in touch by emailing *louie.woodall@infopro-digital.com*, tweeting *@LouieWoodall*, or messaging on LinkedIn. Keep up with the Quantum team by following *@RiskQuantum*.

1 Level 2 and 3 fair value assets as share of total assets at EU banks



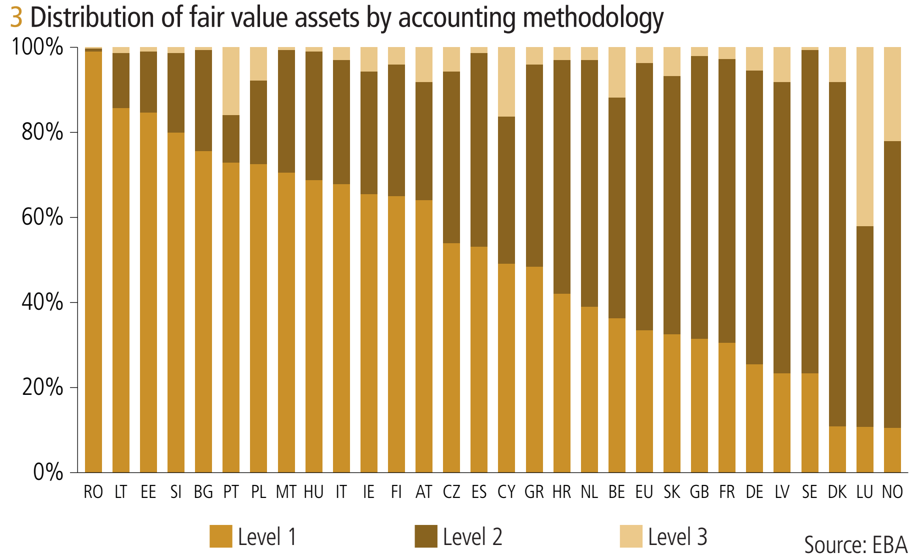
Source: EBA

2 Distribution of financial assets by valuation approach



Source: EBA

3 Distribution of fair value assets by accounting methodology



Source: EBA

Pool party

A new collaborative mood is taking hold of the credit modelling industry, as tougher rules and shrinking benefits prompt banks to consider outsourcing the work, pooling their data – and even sharing their models. By Philip Alexander

Banks that model their own credit risk capital requirements are being squeezed on both sides. Supervisors want them to work harder, while the rewards available – in terms of capital relief relative to the cruder standardised approach – are shrinking.

This is a problem not only for banks but also for European regulators who fought to save the internal ratings-based (IRB) approach from their sceptical peers. As a result, a new attitude is starting to emerge on both sides of the divide. Instead of a sharp-elbowed, rule-bending attempt to stay ahead of the chasing pack, some now see credit modelling as ripe for collaboration. Data could be shared, the models could be developed by a group of banks, and the testing and validation could be handed off to a common platform.

Need to know

- Credit risk capital models may have escaped tough earlier proposals when Basel III was finalised in late 2017, but costs are still set to rise for banks while capital savings slip.
- Despite that, banks and regulators alike want to keep the models running. The industry's attention has turned to cost-efficiency and some regulators are looking for ways to help.
- The European Central Bank, for example, has endorsed the practice of using pooled data to fill individual banks' modelling gaps. It also entertains the idea of jointly developed models.
- Credit risk quants, meanwhile, suggest taking it further – via an inventory of 'pre-approved' models, an industry designed standard model or a validation utility.
- "Whereas in the past, analytics were a source of competitive advantage from a capital perspective, they are now becoming less so," says one model development head.

"Whereas in the past, analytics were a source of competitive advantage from a capital perspective, they are now becoming less so," says the head of model development at a Tier 2 eurozone bank. "The models are moving to shared data for corporate defaults, so jumping to shared models is a logical next step."

Regulators are doing their bit by making it clear that banks do not have to do everything on their own. Unveiling draft chapters of its new modelling guide in September, the head of internal models at the European Central Bank (ECB), Robert Lauter, told *Risk.net* that banks are "generally allowed" to draw data from external pools when trying to model portfolios where there is a shortage of data.

Regulators have been ambiguous on this point in the past. And the ECB guide goes further, raising the possibility of banks pooling or outsourcing some of the modelling capacity itself.

Some banks are said to be benefiting from this kind of policy already. Two sources who spoke to *Risk.net* for this story share hard-to-substantiate rumours of banks that have recently received a green light to hand off their model testing and validation to a third party (see box: Taskmasters).

From a vendor perspective, of course, this looks like a big opportunity. In fact, the more constraints and rules applied by regulators to internal modelling, the more the convergence between banks lends itself to a pooled model approach.

"For us, when there are clear rules, this is much better. Then we can calibrate and make adjustments to what we develop to make sure the models are aligned with what banks do," says Cristiano Zazzara, head of credit analytics at S&P Global Market Intelligence.

Mix and match

The advent of the ECB as a single supervisor for 118 banks in the eurozone in 2014 accelerated the process of scrutinising and harmonising credit risk model inputs and techniques in Europe, but there was still a long way to go.

In 2017, a European Banking Authority survey of IRB banks found 102 participants had a total of 252 probability-of-default (PD) models between them, while 95 banks had a total of 202 loss-given-default (LGD) models – two of the key ingredients for credit risk capital modelling. That is roughly 4.6 PD and LGD models at each bank, and the problem is compounded by the latitude given to lenders in the past.

Theo van Drunen, head of portfolio management within the corporate and institutional banking division at ABN Amro, says national supervisors were willing to approve first-generation models in the 2000s under Basel II that compensated for an absence of data with plenty of expert judgement – as long as the model could be explained, and the judgement incorporated sufficient margins of conservatism.

"When the ECB became the supervising body, a more standardised approach was needed – they cannot look at expert models from every bank, [so] their analysis is much more based on data provided by the banks. Partly, that's because they have a larger number of banks under their supervision, and also because they have to ensure a level playing field when comparing the models of all those 118 banks," says van Drunen.

The ECB has begun to enforce this convergence via the Targeted Review of Internal Models (Trim). This process has generated reports to banks on where the most frequent shortfalls in modelling standards lie. The accumulated wisdom is being distilled into the guide on internal models to establish common best practices across the eurozone.

"The regulator has been very clear on expectations – on setting the bar – and I quite like that. All of the historic workarounds are just not acceptable. In the new rules, wherever there is a weakness, you need to call out a margin of conservatism and add something to the outcome," says the Tier 2 bank's head of model development.

The guide's draft chapters on specific risk types were proposed for comment by the ECB in September, including a long section on credit risk modelling, which includes the words "representativeness" or "representative" 20 times when referring to the use of data. Hardly catchy, but the meaning is clear: to avoid those unwanted margins of conservatism, banks must be able to demonstrate that any pooled or external data is a good match for their own portfolios and policies.

"The ECB document clearly says if you are using pooled data in addition to your own loss experience, you need a separate calculation highlighting what the differences are. Before and up to now, banks were using external data without necessarily disclosing this information to regulators, so it is an additional requirement to make sure banks are aware [of] what they are doing and using," says S&P's Zazzara.

This is where there are a few niggling worries, which are likely to show up in the consultation responses, about how the ECB will interpret its own draft guide (see box: What the ECB says).

Some 11 European banks established a not-for-profit credit data pool, now called Global Credit Data, way back in 2004, to prepare for Basel II data needs. Now numbering 53 members, GCD has spent the subsequent decade and a half collecting a database containing 30,000 large corporate defaults and 6,000 defaults of financial institutions, marshalled with around 400 validation rules to ensure data quality. Members have also refined how best to use the default and LGD data in the pool.

At the heart of the process is the concept of a reference dataset: a selection of data from the pool chosen by each bank to be directly relevant to its own portfolio. That is why Richard Crecel, executive director of GCD and a former risk modeller for a global bank, emphasises the pool itself should not be seen as either representative or unrepresentative.

"It is not a native given; it is something that results from how you select the data from the external pool. The external pool is an aggregation of all the banks' portfolios, and the banks create representative samples from the pooled data by enriching their own small portfolios with our large aggregation of portfolios," he says.

Van Drunen at ABN Amro gives the example of a ship finance portfolio. ABN Amro might submit between 40 and 50 defaults over a 15-year period and receive back data on as many as 600 defaults from the pool, but this needs to be filtered according to the firm's own lending policies – including the maximum age of a ship it is prepared to finance and the maximum loan-to-value at the point of origination.

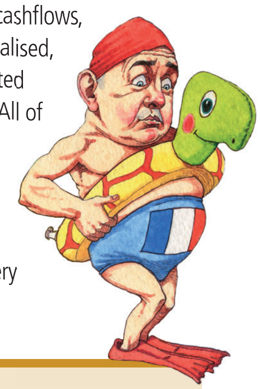
"That could reduce it by 100–150 observations, but that is still 450–500 observations – 10 times what we have in our own observations – and therefore a much more powerful set on which to develop a statistical model. Of course, it is important the data model is extensive enough to contain the risk drivers needed to distinguish the low and high risks," says van Drunen.

Follow the policies

In fact, GCD's Crecel suggests that judiciously used pool data could even be more representative than internal data. He expands on the example of ship finance. Lacking sufficient internal data points, a bank might estimate PDs for shipping loans by throwing both oil tankers and container ships into the same risk bucket. But, with the additional data contained in GCD, a bank could potentially make those estimates separately for each ship type, producing results that are arguably more accurate reflections of the risks specific to each class of borrower.

GCD also incorporates a very detailed default history to help banks align their sample from the pool with their own recovery policies in order to demonstrate representativeness when calculating LGDs.

"We are collecting the precise date of default, the reason for default, the cashflows and recoveries in default, the date of those cashflows, the collateral if it has been realised, when and how, and if defaulted loans have been sold or not. All of this very detailed information is precisely describing the recovery process – this is how we can understand the differences in terms of recovery strategies," says Crecel.



TASKMASTERS

While IRB banks must prove they have an adequate governance framework to manage models internally, there are specific tasks that could be outsourced because they are not intrinsic to the development of the models themselves.

Algosave's Botbol says the banks that the firm has worked with generally regard its product as a challenger model to help identify internal models that might need closer examination if there is a wide divergence between the two sets of outputs.

This is also the explicit purpose of pooled data provider Credit Benchmark, which aggregates corporate PD and some unsecured exposure LGD model outputs from around 30 participating IRB banks.

"All the estimates banks give us come from models that have an audit element to them, so they are more likely to review a model based on what we do than review the PD for an individual name. They also compare the data with their own default experience, and they look at trends to see if there are changes going on in some sub-sectors that they should focus on to prioritise model recalibration," says David Carruthers, head of research at Credit Benchmark.

The benchmarking approach explicitly mirrors the techniques used by the ECB to identify banks whose models are producing outlier results that might merit attention during an on-site supervisory visit. In fact, Carruthers says the supervisors themselves have shown interest in the monthly data from Credit Benchmark, which is timelier than the year-old data often used in regulatory benchmarking exercises.

The model validation process could also be a natural target for outsourcing, because banks must in any case prove to regulators that validators are independent from the model development function. Two sources say banks are exploring this with consultancies, and one believes the UK Prudential Regulatory Authority has already signed off on such an arrangement in at least one instance.

"There is a real thirst for a common platform to outsource backtesting… a more industrial offer. The standard [practice] is to offer a consultant working for you, but there is a more platform way of doing things – a service centre shared between multiple companies," says the model risk manager at a large eurozone bank.

Beyond GCD, there are commercial credit loss databases available, particularly from the solutions arms of credit risk firms such as Moody's and Standard & Poor's. These will not necessarily have access to the granular details of individual banks' underwriting and recovery policies, but they can still use available data to reflect the portfolio of a client as accurately as possible.

For instance, says Leonardo Checchi, a solutions specialist at Moody's Analytics, recoveries through the courts are known to take longer in Italy than in the UK, so data can be selected accordingly. The sheer size of commercial databases should also allow detailed segregation of borrower characteristics to help match bank portfolios, including both PD and LGD inputs.

S&P holds data on 600,000 companies, while Moody's has data on more than 55,000 listed and almost 20 million unlisted firms, including more than three million defaults by the latter.

"It is not just about saying we need to create a synthetic portfolio of data that will enhance banks' data, [chosen by] country and industry; it is also to understand what kind of clients there were, what kind of activities the corporates were involved in. Without going into that kind of granularity, we can lose value for the client's model estimation process," says Checchi.

It's your responsibility

The ECB's requirements for the use of pooled models are even more strenuous than those on pooled data. Fundamentally, an IRB bank is an IRB bank – model approval depends on showing in-house capacity to manage models. That includes providing supervisors with timely access to model development information, says ABN Amro's van Drunen, and proving the bank can cope if, for any reason, the external vendor falls away. He therefore questions the business case for a pooled model provider.

"Since the bank – not the vendor – is always facing the regulator, actually it's a kind of double work. You need to have the capacity and do the work as a bank yourself anyway, while you have to pay for the pooled model as well," he says.

Showing you have the capability to run the model could be harder if it was developed by a group of banks or a vendor. GCD's Crecel cites the example of the standardised initial margin model (Simm), which was created to calculate margin on non-cleared swaps ahead of mandatory requirements introduced in September 2016. In essence, this is a pooled model developed by an industry committee at the International Swaps and Derivatives Association.

Crecel was involved in Simm implementation during his time as a risk modeller on a trading desk in New York. He says he observed all banks facing the same reality that model validation for the pooled

"[Banks] look at trends to see if there are changes going on in some sub-sectors that they should focus on to prioritise model recalibration"
David Carruthers, Credit Benchmark

model was in some ways even more complex than for a pure internal model; for instance, supervisors asked for evidence that the bank's processes and portfolios complied with the way the model was originally developed. The linkage between the model's development and its practical use needed to be addressed even more specifically for those developed externally.

"The experience was that implementing a standardised model doesn't save you from implementing all the expectations regarding MRM. There is a marginal economy, which is the development cost, but the highest costs when running models are collecting and managing good-quality data, as well as monitoring the model for its performance, and for sound implementation and usage by end-users," says Crecel.

A head of MRM at a large eurozone bank is more optimistic about the development of shared utility models, but agrees they might not work for all portfolios at all banks. Each bank will need to know and monitor the development methodology for the model, and assess whether it is appropriate in each particular case. For example, a bank with corporate lending concentrated in central and eastern Europe would need to think twice before using a model developed using western European credit data, he says.

Advocates of pooled model resources also mention another potential obstacle going beyond regulatory and governance concerns. While senior executives are always on the lookout for ways to cut operating costs, procurement is likely to involve people much closer to the risk models: the modellers themselves, who are understandably reluctant to outsource their own jobs.

"The difficulty is not to be felt as a threat. What we say to bankers is you have four C's in credit risk: capacity, collateral, covenants and character. Character, we cannot cover – this is your expert-driven judgement; for instance, who the management is, who are the competitors, what is the location [of the borrower], what does the future look like in this area? The other three C's we can do for you," says David Botbol, chief executive of start-up risk-modelling firm Algosave and a former corporate bond portfolio manager.

This points to the nature of the opportunity for external modelling firms – to process data that is not internal to the bank's own policies or risk appetite, and to undertake modelling with information or techniques that go beyond the bank's own capacity.

Checchi at Moody's provides an example from one of the key components of IRB: the need to model PDs and LGDs through the cycle to avoid a procyclical situation where bank capital levels prove inadequate to absorb losses during a downturn. Clearly, the economic and business cycle is not tied to a bank's internal policies, and can safely be modelled by an external provider using large quantities of data, as long as the knowledge is shared with the bank for the purposes of supervisory inspection.

Another activity that could be undertaken on a pooled basis because it is not dependent on bank policies is to model collateral values on secured loans, with average recovery times in each jurisdiction as a key input.

Pre-approved models?

There is a cloud on the horizon for all IRB models in the form of the Basel III capital floor. While the Basel Committee on Banking Supervision ultimately rowed back on March 2016 proposals that would have scrapped modelling altogether for financial institution and large corporate exposures, the flooring of aggregate model-based outputs at 72.5% of the RWAs produced by standardised approaches could still have a significant impact.

Credit risk accounts for the vast majority of bank portfolios, so it will be very difficult for overall IRB exposures to drop much below the 72.5% floor. But individual portfolios that are particularly penalised under standardised approaches could still pierce the floor, which increases demand from banks to better distinguish between different business lines in their modelling capacity.

Moreover, models help banks demonstrate they are adequately managing and capitalising risks and

> "…when there are clear rules, this is much better. Then we can calibrate and make adjustments to what we develop to make sure the models are aligned with what banks do" Cristiano Zazzara, S&P Global Market Intelligence

correlations between exposures during supervisory discussions over possible Pillar 2 capital add-ons – if a watchdog decides the prescriptive rules for models or standardised capital charges have a blind spot.

"The more options you have on the table to focus on specific niche businesses and better measure the risk differentiation, the better it is. When it comes to capital allocation, Pillar 2 strategies, it is also important to consider concentration and diversification across different asset classes at a granular level," says Checchi at Moody's.

More broadly, banks still want to use internal models to manage their economic capital, even if the regulatory capital is floored by the standardised approaches. Several sources say the larger banks are conscious that Basel floors built on standardised RWAs could prompt herd behaviour, as lenders forsake those assets facing the most punitive standardised capital charges and flock to those that may be treated more lightly.

Banks want to avoid this by having a sharper internal view of risk, even if regulators will not allow capital relief based on those models. Which puts the onus back on to cost-efficiency – and hopes that regulators could offer further support.

The large eurozone bank's head of model development says the ECB is building what he believes amounts to "a kind of standardised inventory" of regulatory capital models as part of its efforts to cull the estimated 7,000 credit risk models in use under its jurisdiction.

Although intended as a supervisory aid, he thinks it could also be used by banks as a "common repository" of approved models against which to assess their own operations and model selection: "This could be a standardised tool you deploy in every institution."

He even makes the radical suggestion that regulators should consider some kind of supervised pooled industrywide model – a credit risk Simm, if you like, as an alternative to the standardised floor.

While the Basel Committee maintains that standardised RWA calibration is derived from real-world loss experience, anyone who has followed the formulation of Basel rules knows the calibration often changes substantially between first consultation and final draft.

That is partly the product of haggling between member states that are focused on the potential impact on specific asset classes in each of their jurisdictions.

A shared model could be a healthier alternative.

"In a utility way, [Basel] could rely on a standard model developed on standard data, where they believe those items are not bank-specific… The aspect of more risk-sensitive [RWAs] and less arbitrage around calibration levels would be closer to some form of reality," says the model risk manager. ■

Previously published on Risk.net

>> Further reading on www.risk.net

- Basel to allow IRB models for low-default portfolios *www.risk.net/2470641*
- Risk managers defend IRB against Tarullo criticism *www.risk.net/2355248*
- Banks diving into credit data pools as official support grows *www.risk.net/5263276*
- A third of eurozone banks fall short on IRB model standards *www.risk.net/5938536*
- Basel plans modelling curb for billions in credit RWAs *www.risk.net/2452449*

WHAT THE ECB SAYS

Although banks have dabbled with pooling credit loss data and been able to purchase credit risk modelling tools from vendors since the advent of Basel II back in the 2000s, the ECB's draft guide on internal risk models in September 2018 marked the first time a regulator had given such explicit criteria for using these techniques. This provides reassurance to banks, and helps them devise policies that should allow them to use external data or models without incurring supervisory capital add-ons.

For data, the key requirement is that external or pooled sources should be subject to the same rigorous criteria as internal data. Specifically, "institutions should have sound policies, processes and methods in place… for assessing and improving the quality and representativeness of the data used in the modelling and risk quantification process".

"Proving representativeness in cases where an institution uses external data is generally more difficult as internal data are scarce," the guide says. "If an institution cannot provide sufficient proof that the external data are representative, in the ECB's view it may still use external data if it shows (by quantitative analysis and/or qualitative argumentation) that the information gained from the use of the external data outweighs any drawbacks stemming from the deficiencies identified and an appropriate margin of conservatism is applied."

Where data is drawn from a pool, "the rating systems and criteria of other institutions in the pool must be similar to its own". To comply with this rule, pool participants must "ensure that there is a common definition of the key drivers and processes" and ensure procedures involving human judgement "can be applied in a consistent and comparable manner across all participating institutions".

Where a model purchased from a third-party vendor is used, its usage must comply with Article 144 of the Capital Requirements Regulation. The requirements of this article include meaningful risk differentiation, proper linkage to credit risk management systems, collection and storage of all relevant data, validation of the model, and documentation on the rationale for the model's design.

The guide also offers guidance for a model that has been developed jointly – an approach that has been used within Germany's savings bank network, but is not thought to be in wide use elsewhere.

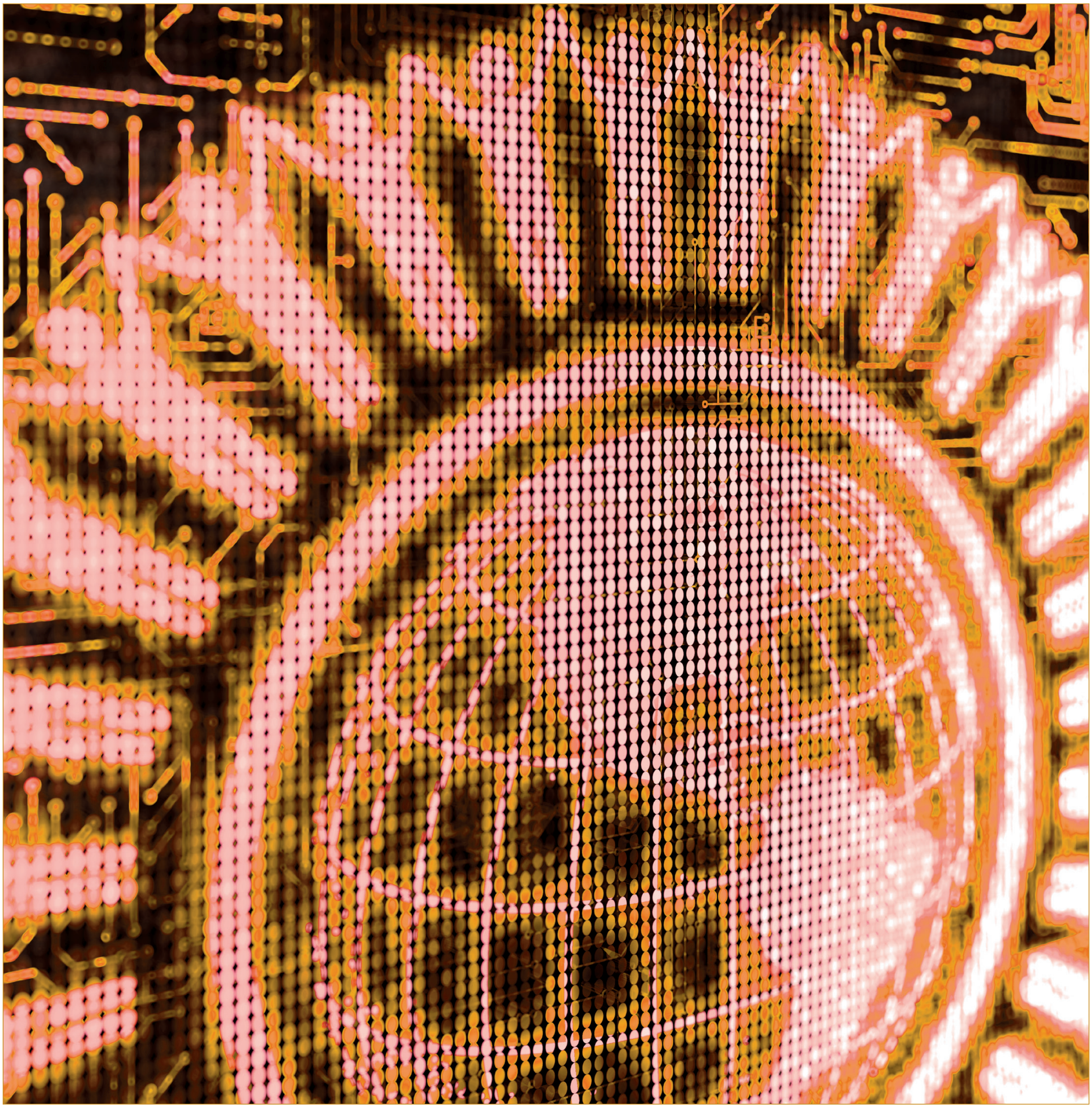
"Where several institutions use a common pool model, each should ensure that its rating process is aligned to the extent that all input risk drivers are defined in the same way across all participating institutions," the guide states.

If that pooled model is used to compute PD or LGD estimates, the ECB requires alignment of policies for managing distressed creditors both before and after default. The pool must also update information on each obligor it contains in a timely fashion.

Crucially, each member institution is responsible for the validation and performance of the model on its own portfolios.

A professional group for model risk managers

As models of all stripes crowd into finance, the people who screen them form an association. By Steve Marlin



"All models are wrong, but some are useful," said statistician George Box in 1978.

And now there's an association for the people who figure out which are which.

With models virtually flooding the financial world, those who vet them are banding together to try to create a standards-setting body. Founded in March last year, the Model Risk Managers' International Association (MRMIA) has already drawn 1,800 professionals via LinkedIn, and is forming chapters in New York and London.

The group – open to practitioners, academics, consultants, regulators and vendors – wants to share knowledge, advance and promote model filtering as the pivotal job it has become. The association will offer courses on model validation and governance, and professional certification in MRM. Networking and what the arc of a career might look like will also be part of the brief.


Dennis Bennett, MRMIA

"The accountants have their professional associations, so do the internal auditors," says Dennis Bennett, the group's chief and founder, speaking of other trade bodies. "Model risk managers need their own professional association."

Bennett was previously head of MRM at the Federal Home Loan Bank of New York, and has been in the field since the 1980s.

Multiple facets

It's a complicated time to be a model risk manager, he says. Models now run the gamut from traditional risk management tools to shields from cyber attack to artificial intelligence programs to read the mood of consumers.

In 2011, the US Federal Reserve addressed the creeping role of models in banking, and asked banks for an inventory of every model they used, from those used in trading strategies to the ones that decide who gets a credit card.

The Fed's guidance, SR 11-7, is widely considered the gold standard in model risk governance. In Europe, the Targeted Review of Internal Models, set up in 2016, fills a similar role. An interim assessment of Trim by the European Central Bank in September found nearly one-third of the banks examined did not meet the European Union's internal ratings-based credit-modelling standards.

At the same time, modellers are being asked to do more. The lifetime expected credit-loss accounting standards – International Financial Reporting Standard 9 and its US counterpart, Current Expected Credit Loss – have entailed massive development efforts and placed credit risk models at the centre of a debate over the extent to which those standards would affect capital and lending.

Yet another aspect is the evolving digital component. As banks slowly turn into tech companies, model risk managers will have even more overlap with their coding kindred. For instance, banks are increasingly using models to anticipate consumer attitudes, and developers are relying on predictive tools created by social-networking companies.

"A lot of big data and machine learning models are for forecasting customer behaviour. Somebody's going to have to validate and govern those models," said Bennett. "Increasingly, companies need to hire people with expertise in those areas."

Indeed, MRMIA is already attracting members from the Faang (Facebook, Apple, Amazon, Netflix and Google) giants, other Silicon Valley companies and even bio-engineering firms, he said.

Besides Bennett, the group's founding members include Joseph Breeden, chief of Prescient Models; at DZ Bank, quantitative analyst Christian Meyer and Peter Quell, head of portfolio analytics for market and credit risk; Sergio Scandizzo, head of model validation at the European Investment Bank; and Peter Carr, chair of the finance and risk engineering department at New York University's Tandon School of Engineering.

The model risk story parallels that of internal audit. Forty years ago, most internal audit departments reported to the chief financial officer, a built-in conflict of interest: the executive who prepared the financial statements was the same person attesting to their veracity. Over time, internal audit began to report to the board of directors, sometimes through a chief audit officer.

Conflict of interest

MRM is in a parallel situation. Most banks' model risk managers report to the chief risk officer (CRO) – even though the CRO ultimately owns many of the models that are being questioned by model risk managers.

"You've got an inherent conflict of interest. It's not just one or two banks, it's the majority of MRM groups that report to the CRO," said Bennett. "If you want MRM to be independent, you've got to have it report to the risk committee of the board or to a senior executive that doesn't own models."

And getting buy-in from that senior management is crucial. In an article in the *Journal of Risk Model Validation*,[1] Bennett noted MRM requires technical skills, but, crucially, also the people skills to get that backing: "MRM must be organisationally and politically supported by the risk committee and board of directors, so that everyone in the enterprise understands that judgements, recommendations and actions taken by MRM are supported from the top."

An enforcement action against US affiliates of Dutch asset manager Aegon showed what happens when models go wrong.[2] According to the complaint, models used for investment decisions had been developed by an inexperienced junior analyst, and management failed to alert investors even after they learned the models weren't working and had stopped using them.

In August last year, the Securities and Exchange Commission (SEC) took the unusual step of fining not only the company, but also the two managers who it said had failed to properly oversee the models. Non-bank financial companies, it seems, would be held to the same risk modelling standards as banks.

"It appears the SEC is taking a very serious view toward model risk controls, model risk governance and failure to disclose model errors," said Bennett.

Model risk is so new a profession it's still borrowing skills from other risk disciplines and adapting them to its own purposes. Bennett says he is unaware of any graduate degree in MRM – although many quant finance master's programmes recognise its growing importance.

But the field is moving fast. "I've been building, installing, using and improving models since the very beginning of my career," said Bennett, who's had a ringside seat for decades. "I've seen the best – and worst – that models and model users have to offer."

The world over those years has become a more intertwined place, reliant on overlapping systems. Bennett commented: "MRM is an important part of that topography." ■

Previously published on Risk.net

[1] *D Bennett,* Journal of Risk Model Validation, *Governance and organizational requirements for effective model risk management, January 2018, https://www.risk.net/5379046*

[2] *US Securities and Exchange Commission,* Transamerica entities to pay $97 million to investors relating to errors in quantitative investment models, *August 2018, https://bit.ly/2UGbhuv*

The disputed terrain of model risk scoring

There is no concord on how banks should police their model risk. But two Fed economists have an idea. By Steve Marlin

Need to know

- The Federal Reserve says the riskiness of all the models banks use should be measured in the aggregate. But it has given no guidance, so banks are coming up with their own ideas.
- Two Fed economists – acting independently of the central bank – have put forward a way to measure aggregate model risk for 'families' of models.
- Experts are sceptical of certain aspects of the plan. But neither the industry nor regulators can agree on a better idea.
- Some banks claim to have created ways of tallying up aggregate risk for groups of models; the Fed researchers and others say no bank has actually succeeded.
- Grading a bank's model risk with a single, overarching score is a holy grail for the field. This idea, at present, is very close to fantasy.

As banks lean ever more heavily on models – for pricing, risk, capital and other vitals – their boards are demanding a clear view of exactly how much risk those models entail, and how they may be abetting or denting the bank's financial position. But there is no clear path on how to deliver that.

"We all have different techniques," says the head of MRM at a US global systemically important bank (G-Sib). "Everybody is not using the same approach, but everyone has an approach."

Into this wilderness have stepped two US Federal Reserve economists, with a method that includes both numerical, or quantitative, measures and more subjective, or qualitative, ones to create an aggregate risk score for 'families' of models.

And their proposal – not yet a paper – has become a bit of a punching bag for other experts with their own ideas on the subject.

"I don't mean to trash them," says a senior modelling expert at a large US bank. "It's well meaning, but I'm doubtful how practical it is."

The job of assessing models – used for things like capital planning, balance sheet management, and measuring exposure to market risk, credit risk, and operational risk like cyber defence – is a many-tentacled affair. A variety of things come into play: fallible human opinion, the degrading of models over time, new datasets, and much else. Models can work together, interlocking like inverted staircases in an Escher drawing, making their assessment even trickier.

"A majority of banks are struggling with how to quantify model risk," says the head of model risk at a European bank.

In the US, the Federal Reserve stipulates in SR 11-7 that model risk should be understood "not just for individual models, but also in the aggregate".

But the Fed has remained quiet about how this should be done. So companies are feeling their way blind, gathering statistics, like the number of models being reviewed or approved, or the number that perform poorly, as crude measures of aggregate model risk.

Although some banks claim to have solved model risk aggregation, there is no agreed-upon standard, either in the industry or among regulators.

"The banks aren't there," says the senior modelling expert at the US bank, speaking of aggregate model risk. "The supervisory guidance talks a lot about model risk, but the industry is still trying to get there."

Two Fed economists, one big idea

Stepping into this fray, two Fed economists, Ray Brastow of the Federal Reserve Bank of Richmond and Liming Brotcke of the Chicago Fed, have come up with a way to assign a risk score to model families. They've based their approach on observations gleaned during supervisory reviews – a bird's eye view of industry-wide risk-modelling efforts. (The two underscore that the approach is their own, and does not represent Fed policy.)

They use two quantitative measures: a "model robustness index", which measures the risk of a model at its inception – that is, the risk that the model just won't work – and a "model stability index", which measures model risk once it's in operation. The indexes can be added up for a numerical measure of risk within families of models.

In using a number to size risk they hope to move away from opinion, or qualitative, measurements in favour of something firmer – an effort they haven't seen much in their supervisory forays.

Brastow says it's clear risk cannot simply be added up across businesses to come up with a single encompassing number. "But it's also difficult using non-numerical ways to assess risk," he adds. "All we're suggesting is there are ways to add discipline to the aggregation of model risk."

Everyone's a critic…

Model risk experts at banks briefly praised the Fed economists' approach as a stalwart try at addressing a thorny problem. Then some of them sandbagged it.

"An oversimplification," the US G-Sib executive calls it, adding that its failure to emphasise data quality is "a fundamental flaw".

For instance, he and others say the example used in the Fed presentation was a model used to predict the likelihood of default on home mortgages. But that approach isn't suitable for more sophisticated applications such as derivatives pricing, they say.

The choice of weights is crucial to coming up with model risk scores; some experts would have liked to see more attention paid to how these very subjective elements are selected.

"Before you can aggregate model risk, you need to look at individual model risk components," says Peter Quell, head of portfolio analytics for market and credit risk at DZ Bank in Frankfurt. "The only problem is the weights you are assigning to the different indicators are artificial.

"There is no rational way to derive these weights. It is always up to judgement. But if you want to condense everything into one single number, somehow you need to come up with these weights."

"I don't believe in aggregating model risk across different model types. For example, aggregating the risks associated with pricing models, risk models and retail models might not make sense" Slava Obraztsov, Nomura

To construct indexes for model robustness and stability, the Fed researchers propose assigning weights to the various factors that determine model risk. The right selection of statistics and weights is crucial to building good indexes, they note.

Once weights and risk factors have been selected, models can be graded numerically on robustness (at the beginning of deployment) and stability (after they've been put into production). Models can be further classified as low, medium or high in both robustness and stability.

For example, on a scale of zero to 100, models whose robustness or stability scores are under 60 could be classified high-risk, those between 60 and 85 as medium-risk, and those from 85 to 100, low-risk.

Within a model family, the scores could be aggregated along with other factors such as model complexity and financial impact to derive an overall aggregate model risk measure.

Another issue was data quality – or the lack thereof – in measuring model risk, especially in pricing illiquid assets. A good model blighted by bad data is no better than a flimsy model with good data.

"Data is not considered here," says the head of model risk at the US G-Sib, of the Fed researchers' approach. "You can have all the other parameters like statistical significance and stability, but terrible data. That isn't captured."

The Fed researchers note that although some banks have taken approaches similar to theirs, those attempts have fallen flat, says Brastow. When pressed about their approaches during supervisory reviews, banks say they err on the side of conservatism by adding a 'capital buffer', a fudge factor intended to compensate for any errors in the models. How these amounts are determined is necessarily arbitrary.

"We're not aware of any banks that have gotten very far," says Brastow. "When we press banks about their capital charges for model risk, they're evasive.

"A lot of banks hold extra capital for model risk, a buffer. But if you ask them how they size that, they say: 'We're trying to be conservative.' They don't have a good way of knowing when it increases or decreases."

Home brews

Left to their own devices, banks have been confecting their own model risk tests. The aggregation efforts that are furthest along have employed a mix of qualitative and quantitative components.

A US subsidiary of a large international bank, for instance, is developing a process for aggregating model risk that identifies the relevant parameters for individual models, and then extrapolates those to account for linkages between models.

At the individual model level, the bank's independent validation team tracks qualitative characteristics: complexity, for instance, whether on a traditional model in a mature modelling area or one that uses advanced techniques, like machine learning.

These and other metrics are tallied in the initial validation review and used to develop a 'model risk scorecard' for each model. Each item in the scorecard is weighted to arrive at a total score that can then be added up for all the models in a family to come up with an aggregate risk score.

The bank's US head of MRM says that he looks "at all the individual parameters that are highlighted in individual scores to create a total score for families of models or similar types of models."

The bank says the model aggregation project is progressing well and should be live across all models by Q4 2019. The system will inform senior management on the residual risks the bank is taking, and will help ensure compliance with SR 11-7, though this is not the main objective, he says.

"The bigger benefit is not adherence to SR 11-7," says the head of MRM. "It's more about getting a better handle at articulating risk and reflecting that back into a business-as-usual assessment process."

The scorecard concept used by the Fed researchers is also similar to one outlined in a 2015 paper by Michael Jacobs Jr, then a consultant at Accenture and currently a quantitative analytics expert at PNC Financial Services. Brastow says that of all the theoretical approaches he's aware of, the Jacobs paper comes closest to his and Brotcke's approach, although it doesn't go as far in developing a quantitative method.

Jacobs suggests dividing model risk into two categories: inherent model risk and risk mitigation.

The first category, inherent model risk, is assigned a score based on the model's complexity, uncertainty, availability of data and other factors. The second, risk mitigation, is scored based on criteria such as model validation, performance monitoring, benchmarking and backtesting.

An aggregate score is derived by multiplying, for each model in a family, its risk score by a 'risk weight' and totalling the results.

The Jacobs model has advantages: It's relatively easy to put into motion and it's comprehensible to senior management. On the minus side, it is primarily qualitative, and hence its scores, weights and overall results are subjective. It also doesn't capture model interdependencies.

Why aggregate at all?

Yet others reject outright the idea of trying to peg risk across different types of models.

"I don't believe in aggregating model risk across different model types," says Slava Obraztsov, global head of model validation at Nomura in London. "For example, aggregating the risks associated with pricing models, risk models and retail models might not make sense."

Nomura evaluates its models during an initial validation stage and, after the model goes into operation, the bank refines these numbers based on model performance and aggregates model risk narrowly, within specific model types, such as pricing or risk models.

"It is more appropriate to take large groups of models and try to aggregate model risk across the individual models in those groups," says Obraztsov.

Obraztsov, among others, suggested the Fed researchers, in their zeal to come up with a way to aggregate risk, ignored or oversimplified model risk at the individual level.

"Model risk quantification on an individual model basis is a difficult exercise," says Obraztsov. "Nomura has implemented well-developed methodologies, and due to the complex nature of what we are trying to achieve, the approach is not simple."

Obraztsov says the problem with the Fed researchers' methodology is that it spends too much time on how to aggregate model risk, without addressing its purpose. The approach comes up with an isolated measure without tying that into the bank's overall risk appetite and capital, he says.

"Quantification of model risk is what could be a firm's loss because of model limitations. I'm not sure they're addressing this point," he says.

For example, the output of a model risk aggregation could be expressed in numerical terms as an amount at risk, say $100 million. But that figure needs to be viewed in the context of the firm's economic capital: for some firms, that might be a huge amount; for others, it could be insignificant.

"In our reporting of quantitative measures of model risk, we don't just report potential losses in terms of specific numbers," says Obraztsov. "We always compare those numbers with available capital, whatever threshold is determined by the board of directors. That metric is much more meaningful than actual potential loss."

whose accuracy the livelihood of the business turns.

Indeed, SR 11-7 stipulates that senior management is responsible for reporting on model risk to the board, both at the individual and aggregate levels.

Given the limitations of the Fed model, banks are not likely to settle on a quantitative measure of aggregate model risk anytime soon, nor are model risk teams likely to devote the resources needed to nail one down. Indeed, the senior modelling expert notes that independent model risk teams are having trouble retaining talent, face strained budgets and are under regulatory pressure.

And in the end, model risk is a thankless job. Much of it comes down to saying no to models a company has sunk large amounts of money into. Against those hopes, it can be hard for a person to remain impartial.

"When you're talking about independent model validation, it's tricky to maintain independence while getting buy-in from the rest of the organisation," says the risk modelling expert. "People want to do the right thing, but it's a challenge organisationally."

As for the dreamed-of, shining single score of risk, it's for another day. Coming up with a single score for risk for all models within a 'model family', though, should be a high priority for banks, says Brastow.

"Aggregating model risk is difficult. We don't think anybody does it well, but firms need to do it," he says. On his own approach, he is grounded. "Adding these indices has value, but we're not suggesting this is the only way to do it." ■

> "We all have different techniques. Everybody is not using the same approach, but everyone has an approach"
> Head of MRM at a US G-Sib

A work in progress

To be fair, the Fed researchers acknowledge the limitations of their approach. They note, for example, the subjectivity in selecting appropriate statistics and weights in constructing the indexes.

They also note that a single measure of model risk across families was never the objective, and indeed is not even expected by SR 11-7. Instead, supervisors expect banks to employ multiple measures of model risk such as performance, robustness and stability.

The researchers have a number of other caveats on their approach. The number and types of models used by large institutions make it difficult to establish a uniform approach for aggregating model risk. The authors admit that they haven't yet tried to simulate an actual bank's use of models, and that the approach needs to be tested in real-world situations. In addition, it does not take into account the interdependence of models and networking effects on model risk.

"We're not suggesting that this is a panacea to measure model risk," says Brastow. "These are imperfect measures.

"However, by quantifying, you're forced to think more deeply about where model risk comes from. If you're adding up the number of criticised models, that's inherently an *ad hoc* judgement. What we're suggesting is this approach might create some consistency."

The two researchers have presented their idea at conferences, and published earlier this year.[1] Still, the fact that the approach isn't quite ready for prime time does not negate the need to quantify model risk both at the individual and aggregate levels. Boards and risk committees need to have insight into the risks presented by models, upon

Previously published on Risk.net

[1] Journal of Risk Management in Financial Institutions 2019, *Volume 12 Number 2*

>> Further reading on www.risk.net

- National supervisors put pressure on global risk models *www.risk.net/6063461*
- Fed's Brainard wary of black-box AI models in consumer credit *www.risk.net/6125691*
- Finally, a professional group for model risk managers *www.risk.net/6008561*

Mastering the evolution of the Model Risk Management Transformation

Our ambition is to deliver real value for our clients.

And we know that value means different things to different people. That's why we'll work with you to understand what's important to you.

pwc.co.uk

Empower your business with intelligent risk analytics

Meeting risk and regulatory compliance challenges head-on while growing business value takes innovative technology. With a single, integrated platform, SAS® Risk Management is the proven choice to modernize your organization. Whether tackling stress testing requirements, meeting accounting standards or making strategic business decisions, SAS solutions give you a complete enterprise view of risk and compliance.

sas.com/risk