1 Introduction

"All models are wrong, but some models are useful."
G. Box

The Basel Accord, as well as the guidance issued by the Committee of European Banking Supervisors (CEBS) (CEBS 2006 and Basel Committee on Banking Supervision 2006a), state explicitly and unequivocally that an effective validation process is a *conditio sine qua non* for allowing financial institutions to adopt the Advanced Measurement Approach (AMA) for operational risk.

Furthermore, with the final agreement on Basel II and the subsequent issuing of the Capital Requirements Directive (CRD) attention has turned, at least outside the US, to the way national regulators will choose to implement it and to the timeline they will follow. For instance, the application form to be submitted to the Financial Services Authority (FSA 2006) in order to be allowed to use the AMA contains the following sections:

- high-level overview and implementation plan;
- overview of self-assessment against relevant standards;
- summary of the approach in a number of key areas;
  - governance,
  - requirements for the use of AMA,
  - data management and integrity – compliance with AMA standards,
  - validation,
  - documentation.
- details of the AMA model being used; and
- sign-off by CEO or equivalent.

The extensive documentation required may already be available in full or in part within the documents produced during the planning and development stage. For example, the implementation plan must have been produced before starting the process of establishing an AMA-compliant framework and compliance with standards should have been assessed at a very initial stage, while the details of the model should have been produced during and immediately after the selection and development stage. However, the validation process, which only makes sense for something that is truly up-and-running, is likely to be the last point of focus before AMA becomes enshrined in regular practice. Table 1.1 summarises the timeline the FSA will follow in granting permission to use AMA and illustrates the criticality of the next few months in the approval process for the main European financial market.

Table 1.1. Timeline the FSA will follow in granting permission to use AMA

Source: FSA website

Date of receipt of application	Decision due by
1 January, 2007 – 31 March, 2007	30 September, 2007
1 April, 2007 – 30 June, 2007	31 December, 2007
1 July, 2007 – 30 September, 2007	31 March, 2008
1 October, 2007 onwards	Six months from the date of receipt of a complete application

Definitions

According to the Merriam–Webster dictionary the word "validation" means:

- to make legally valid; or
- to support or corroborate on a sound or authoritative basis.

Whether we choose to emphasise the correspondence to formal requirements or the rigour and soundness of the underlying analysis, validation refers to the process of ascertaining that something satisfies a certain, predefined set of criteria.

The above definitions come from long-established practice in both hard and soft science where models are used as a key tool for the representation and analysis of systems and phenomena. In finance the development of increasingly complex transactions over the past three decades and the consequent need for sophisticated valuation and risk management models has generated a great emphasis on model validation as a means to reduce "model risk". Model risk in this context is the risk that "…decision makers either relied on erroneous price or exposure estimates, or on an overly broad interpretation of model results…" (Office of the Comptroller of the Currency 2000). But what does validation specifically mean in the context of AMA for operational risk?

The problem of model validation can be seen in the context of the so-called problem of induction, as formulated by philosopher David Hume "who argued that from the strict logical point of view we have no justification in generalising from instances we have experience of to those of which we have no experience" (O'Hear 1989), and of the ways different schools of thought have addressed it.

Logical positivists have stated that a theory can be proven true (verified) by cumulative observations consistent with the theory's predictions. Others, the most eminent of which is Popper (1959), have argued that, whereas a theory can be conclusively rejected on the basis of even a single experimental result contrary to its predictions, it can never be conclusively established as true (ie, verified) and as such will only be provisionally retained, until the first counter-instance will force us to discard it.

This approach solves the problem of induction by rejecting it outright as a means to establish the truth of a theory. In a recent book on risk management, Taleb (2004) has used a suggestive version of Hume's problem to highlight the pitfalls of allowing induction to guide risk assessment and predictions in finance, namely John Stuart Mill's reformulation of the problem of induction: "No amount of observations of white swans can allow the inference that all swans are white, but the observation of a single black swan is sufficient to refute that conclusion". However, although this approach seems eminently sensible when applied to predictions (and actions of consequence) based on historical observations, we may wonder whether it is in actuality a feasible one when dealing with operational risk models and methodologies.

The main criticism to the Popperian approach has been the observation that it is very difficult, if not impossible, to put into practice, and that, in the real world, a theory is not immediately rejected at the first counter-instance. Rather, the theory is revised and changes are made in one or more of the hypotheses made in order to accommodate the new results. While Popper maintained that this approach salvages a theory at the cost of lowering, or destroying, its scientific status, other thinkers, the most authoritative of which was Quine (1970), countered that, as experiments are fallible both ways, they can neither conclusively falsify nor verify a theory. Therefore scientists should be very sceptical revisers (rather than sceptical falsifiers) who try, wherever possible, to save theories rather to refute them. Moreover, a strict application of Popper's criterion would nullify any attempt at constructing theories based on probabilistic rather than deterministic statements, where individual instances of outcomes contrary to predictions are the norm rather than the exception.

Now, without in any way pretending to compare the task of operational risk assessment to the high pursuits of philosophy of science, it is undeniable that in risk measurement we feed a model with data and then compare results (predictions) with empirical observations (losses). In market risk, the area in which the most extensive experience has been accumulated over the past 10 years or so, there is specific regulatory guidance (Basel Committee on Banking Supervision 1996) on how to rigorously test and improve the quality of a model. However, even in this case, where models are strictly quantitative and statistical data is plentiful, contrary empirical evidence (backtesting exceptions) does not lead to rejection of the model, but rather to adjustments to hypotheses, parameters, data sets and so on (and to penalties on capital requirements).

It seems therefore all the more sensible to apply an essentially Quinean rather than a Popperian approach to model validation for operational risk. All such models are at various stages of development and are still far from the emergence of a proper industry standard and therefore in need of thorough testing and review (from basic principles to final outputs) precisely because, by and large, they are still far from being truly reliable.

Validation in AMA represents, therefore, not just a means to independently establish the accuracy of estimates, but also, and more importantly, a fundamental component of the development process of AMA models. By making validation an ongoing activity that will be performed contextually to risk measurement itself and by using its results to improve our understanding and ability to model and quantify operational risk, we may foster the evolution of more rigorous methodologies and the emergence of innovative and truly reliable models.

The Basel Committee on Banking Supervision (BCBS) in its January 2005 Newsletter addressed the issue of validation through a principle-based approach, but limited their remarks to the internal ratings based (IRB) approach to credit risk. The Committee of European Banking Supervisors (2006) has subsequently developed more articulated guidance on the subject of both IRB and AMA while still maintaining the BCBS principles as the key reference. This document has subsequently been amended, following industry feedback, and has become less specific and prescriptive in a number of areas. It remains, nevertheless, the most extensive and authoritative guide on the subject of implementation and validation of IRB and AMA.

First, although there is a lot of talk of "AMA models", we should remember that there is no such thing. AMA is an approach to measurement and, at best, a methodology based on a combination of tools, some of which may be mathematical models. This hybrid nature means that certain standard techniques may be used to validate the model components of AMA while more *ad hoc* methods may be needed to satisfy supervisory criteria, for instance on scenario analysis or on the business and control environment.

Another piece of regulatory literature that is worth reviewing is a relatively old issue of the OCC bulletin on model validation (OCC 2000). It focuses on financial products valuation and market value-at-risk (VaR) and, having been issued well before the second Basel Accord, does not even mention operational risk. However, it contains a number of key principles and suggestions that also apply to operational risk models as well as constituting a very useful discussion on the importance of a formal validation policy.

Inevitably, however, available regulatory guidance on validation is largely geared towards statistical model validation, leaving much to be decided as to what technique to use to validate scenario analysis or business and internal control environment factors.

This is compounded by the fact that there is also no regulatory guidance, let alone prescription as to what kind of model, or models, should be used under AMA. Aside from knowing that it must use statistical analysis of internal and external data as well as scenario analysis and business environment and internal control factors, bankers are left to their own devices. Academic literature is also of little help as on one side it is scarce in its own right, and on the other it lacks an acceptable body of empirical evidence unequivocally pointing at the effectiveness of one or more model. On this topic it is worth quoting in full the Basel Committee on Banking Supervision (2006b) in their study on existing AMA practices.

> "The combination and weighting of individual elements varies widely across banks. Some banks base their operational risk capital estimate largely – or even solely – on scenario analysis, and incorporate internal and external data only indirectly as inputs to the scenario generation process. Other banks rely heavily on internal data, using external data and scenario analysis only where there are gaps in their own loss experience. Others use internal data to model the frequency of operational risk losses and external data to model loss severity, especially in the tail.

> Most banks, however, incorporate more than one element directly in their AMA model and some incorporate all four, albeit with varying weight."

The picture that is emerging points to various hybrid methodologies for distribution fitting and parameter estimation, as well as to different ways to integrate hard and soft data. A few common solutions (analytical or numerical) are presumably going to be used to generate an aggregate loss distribution, and a VaR figure, starting either from historical data or from scenario analysis.

AMA validation in this context shall therefore answer one key question: Does it work? To put it more explicitly, does the approach address the right problem and provide accurate information, and is it actually used for the intended purpose? The last point is especially important as, both in the spirit of the Accord and in the regulators' intention, AMA is not just a measurement methodology, but a risk management framework. Using AMA means not only that capital figures are calculated in line with certain quality criteria, but that both the input and the output of AMA are key elements in the risk management process and that such processes rely substantially on those two in order for key decisions to be made.

Validation Policy

Any effective policy should clearly state its objectives, identify roles and responsibilities, and specify, to the necessary level of detail (but, ideally, no further), the principles and the key processes governing the bank's behaviour towards its subject.

Objectives

The key objective of validation is to assess the predictive ability of the risk estimates produced through AMA. Risk estimates, and related capital requirements, add value to the institution to the extent that they are forward-looking and insofar as they are both able to adequately differentiate and appraise operational risks. As a consequence, the validation process should be aimed at assessing the adequacy of an AMA model by covering the following points:

1. the predictive power (accuracy) of risk estimates;
2. the processes and procedures governing the production of such estimates; and
3. the control framework in place around an AMA process.

Validation should be a key component of an AMA process, in particular in ensuring its systematic review and its reassessment whenever material differences are observed between its estimates and the actual operational losses. This suggests that validation should be an ongoing process, inextricable from the actual risk assessment process.

Roles and Responsibilities

Supervisory guidance makes it clear that banks are expected to validate their own AMA models and that supervisors, and possibly auditors, review this process and its results in order to obtain the necessary assurance about the soundness of the banks' risk estimates and, ultimately, capital at risk. It is therefore essential that banks clearly set up the responsibilities for performing and controlling the validation process within their own organisational structures, keeping in mind the objectives described above.

Board of Directors

As the top responsibility for operational risk management resides with the board of directors, it is this body that has to ensure that validation is regularly performed, its results carefully analysed and the implications of these results effectively acted upon.

Therefore it should be the responsibility of the board of directors to ensure that managers perform the validation process in accordance with the bank's policy. The board should be responsible for approving and periodically reviewing the bank's AMA validation policy. The member of the board responsible for overseeing the operational risk framework should take direct responsibility for all aspects of AMA validation.

Senior Management

Senior management is responsible for exercising effective oversight on the validation process. This translates into ensuring on an ongoing basis that validation and the related controls are performed in accordance with the validation policy. Senior management should have a good understanding of an AMA framework and of the control processes around it (including validation) and should ensure that all the issues highlighted through the validation process are promptly addressed.

Operational Risk Function

Validation is a complex task requiring both quantitative skills as well as knowledge of operational risk management and of the internal control environment of the bank. These are the same specialised skills required to develop and implement an AMA successfully and they are usually scarce in every institution. There is, therefore, a trade-off between this scarcity and the requirement for an independent review, and between the ability to perform the task properly and full independence of those who perform it. This trade-off must be balanced by an effort, on the part of those building and implementing an AMA, to develop clear and informative documentation about all the stages of the measurement approach and to effectively communicate all aspects of the approach to senior managers and all relevant decision makers.

In practice, the size and sophistication of the bank's structure may pose substantial constraints on the fulfilment of this requirement. In a large bank, active in most business lines, with many groups specialising in modelling of various kinds (market risk, credit risk, product valuation and so on), independent expertise may be readily available. In this case the operational risk function should be responsible, besides of course the development, documentation and implementation of AMA, for the development and maintenance of the validation policy, ideally in cooperation with the independent unit performing the validation itself. Such unit would then be responsible for performing the validation process according to the provisions of the policy and for providing the related analysis, reports and recommendations.

Their conclusions, whether positive or negative, on the validity of the model should be clearly motivated. The validation unit should rely on AMA documentation provided by the operational risk function and should be able to challenge assumptions and modelling choices.

A number of possible solutions may be sought whenever a fully independent unit holding the necessary expertise and competence cannot be found within the institution. As an example, another modelling unit within risk management (market- or credit-risk-related, for instance) may perform the validation, even if they share the same reporting line with the operational risk function. When such an arrangement is not practicable, and the only unit available is the operational risk function itself, another possibility relies on the segregation between staff responsible for AMA development and staff responsible for AMA validation under the supervision of the function's head. These solutions will of course be balanced by a growing emphasis on clear documentation, senior management oversight and audit review (see below).

Business Units

Although business units are not directly involved in the validation process, they may nevertheless be required to support it by making internal data available for revision, audit or backtesting purposes, as well as by providing information on self-assessment and scenario analysis during the testing and reviewing of procedures. Although their availability should not be abused, business units should recognise the importance of their role in an AMA process and provide the necessary time and resources for its successful fulfilment.

Internal Audit

Internal audit, as a rule, should not perform the validation process, but regularly review it within the context of its inspection of the operational risk management framework. Such independent review should provide further assurance, to the board, external auditors and regulatory authorities, of the reliability and accuracy of the operational risk estimates. However, as mentioned above, in cases where the independence of the validation unit is less complete or even virtually absent, internal audit could extend the scope of its tests in order to perform a more complete revision and obtain a comprehensive coverage of the validation process.

Validation Methodologies

There is no single validation methodology which would fit all models and which would be clearly superior to others. Furthermore, AMA requires the validation process to cover diverse items both within the input and the actual modelling process. Different methodologies can be applied for validating statistical data, data from self-assessment and data on the business and control environment given the different sources, nature (qualitative and/or quantitative) and role within an AMA of these kinds of information. Likewise, different validation steps can be taken when validating a statistical model (for instance on the selection of distribution and on the estimate of parameters) and when validating the process of scenario analysis.

It is difficult to provide a comprehensive list of validation methodologies and techniques, but we can try to develop a classification within the context of AMA. We will do this along two key dimensions: scope of the validation (input, methodology and output) and kind of validation (from the more qualitative to the more quantitative). Table 1.2 shows a summary of the key methodologies for AMA arranged accordingly.

Table 1.2. Validation methodologies

	Input	Model and output
Qualitative	Review and tests of standards, procedures and controls	Independent review of logical and conceptual soundness
	Peer group comparative analysis	Peer group comparison and benchmarking
↓	Reconciliations with the general ledger	Analysis of distribution fitting and parameter estimation process
	Statistical tests, scaling methodologies, calibration	Comparative analysis of other models' results
Quantitative	Regression analysis	Sensitivity analysis Backtesting

Methodologies for Input Validation

- *Review and tests of standards, procedures and controls*: the internal control framework around the collection of loss data, around the computation of key risk indicators (KRI) and around the execution of self-assessment and risk scoring should ensure the quality of input data. Checking that this control process is up and running and is closely followed is a first important step towards AMA validation.
- *Peer group comparative analysis*: comparing data from different business units or, where available, from comparable business units in different corporate entities and/or jurisdictions can be useful to highlight inconsistencies and in some cases under-reporting of losses.
- *Reconciliations with the general ledger*: albeit this practice has been much criticised by industry practitioners as unreliable and even impossible to perform properly, nevertheless it could be of use in the limited number of cases where so-called hard losses can be clearly identified (see Chapter 2 for more details).
- *Statistical tests, scaling methodologies, calibration*: there are a number of quantitative techniques that can be applied to loss data, both internal and external, to enhance its quality and ensure its validity for AMA purposes. Tests for instance can be run to decide whether external data can be considered as drawn from the same distribution as internal data, and thus can be used to supplement it. Scaling and calibration can be applied to adapt external data according to the size of the institution and to correct for reporting bias (see Chapter 2 for more details).
- *Regression analysis*: regression analysis can be used to identify relationships among variables. For example certain quantities can be identified as KRIs on the basis of regression analysis showing their relationship with operational events or losses.

Methodologies for Model and Output Validation

- *Independent review of logical and conceptual soundness*: although the most basic of validation tools, the importance of having an independent eye cast a critical look on key assumptions, as well as perform a sanity check on results remains a simple, but essential test that should not be overlooked and that can help reveal inconsistencies and drawbacks.

- *Peer group comparison and benchmarking*: identifying a peer group, according to size, business and market mix, geographic situation and so on, can be very helpful for the establishment of a benchmark with which to compare capital requirements and other key figures. Any material discrepancy from such a benchmark should be convincingly explained.
- *Analysis of distribution fitting and parameter estimation process*: statistical theory provides a panoply of tests and techniques to guide the choice of probability distributions and the estimation of parameters from sample data. The execution of such tests already should be part of the implementation of the LDA component of AMA. The validation process should verify that this is the case and that they have been applied consistently and correctly. It should also supplement them whenever needed in order to ensure the solidity of the approach.
- *Comparative analysis of other models' results*: another important technique that can be used to validate a model's methodology and results consists in building either the same model independently, for instance through a spreadsheet, or developing a different, perhaps simpler, mathematical model that should, however, give comparable results. This is a rather time-consuming approach, as it requires, to an extent, doubling part of the effort of model developing, but has the advantage of providing results that are directly comparable and give the opportunity to compare models' behaviour under different circumstances.
- *Sensitivity analysis*: altering one or more of the components (variables, parameters, inputs) of an AMA estimation and examining the variations in the results can help ensure the stability of the system or highlight otherwise unpredictable model behaviours.
- *Backtesting*: the term backtesting refers to the test of the performance of a VaR model through a comparison between the risk estimates produced by the model and the actual losses experienced by the institution. Backtesting is a formally required validation technique for trading market risk VaR under the first Basel Accord and, although difficult to translate *tel quel* to operational risk, remains a fundamental validation technique. Chapter 2 presents a more detailed discussion of the use of backtesting in the context of AMA.

Validation Process

The validation policy should not limit itself to objectives and principles, responsibilities and methods, but should specifically define all the steps of the validation process including frequency and reporting obligations. It should detail, and possibly also justify, the choice of validation techniques to be used and, if relevant, their sequence and relative weight towards the final result.

For input validation the policy should prescribe which method should be applied to each category of data (review and tests of procedure for internal loss data, scaling and calibration for external loss data, KRI reviews and so on as appropriate) and what criteria should be used to arrive at a final decision.

For model and output validation the policy should detail the choice of tests to be performed, tolerance levels and relevance of each technique adopted. It should give enough freedom to the staff performing the validation to adapt the methods to the realities of the estimation process and its evolution while at the same time maintaining a set of objective and auditable criteria for determining the final accuracies of the results obtained from the use of the model.

One point that the policy should not overlook is the fact that virtually all models are nowadays implemented through computer programs, be it in the form of internally developed applications or, more and more frequently, of commercially available software. Of course, no matter how sound and theoretically correct the mathematical description of the model, an incorrect computer coding will spoil all the efforts made towards ensuring robustness and accuracy. While proprietary computer code can be professionally tested and proofread with relatively little cost, computer code from commercial software packages would not normally be available to the user. This fact, however, should not deter banks from requiring vendors to document, to the extent that this can be done without revealing proprietary information, how the product has been built and validated and its mathematical formulation. Another validation tool that should not be overlooked is the construction, for instance through a spreadsheet, of a simpler, "equivalent" benchmark model, on the basis of the information provided by the vendor. Using the same input, such models should give comparable results and behave coherently in a sensitivity analysis. Differences will be inevitable, as neither theory nor practical implementation will be exactly the same, but the staff performing the validation should formulate a reasoned judgement as to whether such differences are expected or are due to a processing error in the application examined.

The policy should also specify what actions should be taken and under whose responsibility, whenever the validation process highlights mistakes, inconsistencies or other problems in an AMA process. It should enforce the concept that validation is an iterative process directly supporting the ongoing update and improvement of AMA.

Documentation

The policy should specify how each step should be documented, in electronic or paper-based form, and how the results should be reported to management, including the appropriate level of segregation of duties between staff developing the model, providing input and validating the model. The policy should also mandate that model documentation should be complete and precise to the point of allowing replication of the actual model by the validation unit.

A real-life validation policy will be presented in Chapter 3.

Validation of AMA Models

If operational VaR is used to determine regulatory capital, the accuracy of the VaR estimate becomes, of course, a critical concern. However, as it is already the case for market and credit risk, we should recognise that VaR measures are estimates; ie, they result from calculations based on the observation of sample data and on the computation of parameters (quantiles as well as means, variances, standard deviations and so on) on the basis of the same samples. As a consequence, accuracy is undermined both by the sampling error (the extent to which the distribution of a sample is not the same as the distribution of the underlying population) and by the hypothesis made in calculating parameters from historical data, the most common kind being parameters that remain constant over extended time periods.

Several studies have been conducted on the accuracy of VaR models, first and foremost on market risk. The interested reader can consult, for instance, Styblo Beder (1995), Chappel and Dowd (1999), Dowd (2000), Berkowitz and O'Brien (2001) and Contreras and Satchell (2003). These authors review strengths and weaknesses in VaR estimation methods and results, and suggest several critical points and potential improvements. We will report here some selected key concepts, which will be looked at again in Chapter 2 when discussing the accuracy of operational VaR estimates.

- A VaR estimate with 99.9% confidence does not mean the certainty that a larger loss will not occur more than once out of the next 1,000 time periods, but merely that the same estimate, repeated 1,000 times on samples of similar size, will fall 999 times within the stated confidence interval.
- The size of such a confidence interval, which represents the error made in the estimate of VaR, is largely dependent on the size of the sample. Other factors, such as confidence level and holding period, are much less relevant in comparison.
- The results of a VaR estimate can change dramatically if some key assumptions are changed.
- A number of risk variables that do not normally figure in VaR computations, such as political risk and liquidity risk, can have nevertheless a substantial impact on the bank's losses.
- The last two items point to the need for risk estimates to be embedded in a framework comprising stress testing, as well as appropriate control procedures, one of the most important of which – in order to support the accuracy of such estimates – is certainly validation.

As we will discuss in more detail in Chapter 3, similar results apply to the use of VaR in operational risk, albeit, unsurprisingly, with the not marginal difference of having to work in most cases with smaller samples and more severe sampling errors (owing to reporting bias and other problems in data collection). However, this is also somehow an advantage, as the recognition of the limitations of a pure loss distribution approach has forced practitioners and regulators alike to develop around operational VaR both a more complete control framework and an awareness of the importance of qualitative and managerial information. The result is the current wave of AMA models that although still far from having the desired long-tested reliability, nevertheless represent a substantial innovation to traditional VaR models in that they aim at rigorously integrating quantitative hard data with expert judgement and qualitative information. As we will see in this book, this aspect is at the same time the most challenging and the most potentially valuable in AMA validation.