Physica A 389 (2010) 1721–1728

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

# A Bayesian Networks approach to Operational Risk

V. Aquaro [a,b], M. Bardoscia [c,d,*], R. Bellotti [c,d], A. Consiglio [a], F. De Carlo [d], G. Ferri [e]

[a] CARMA, Research Consortium for Risk Management Automation, via Mitolo 17B, I-70124 Bari, Italy

[b] Formit Servizi S.p.A., via C.Conti Rossini 26, I-00147 Roma, Italy

[c] Dipartimento Interateneo di Fisica "M.Merlin", Università degli Studi di Bari e Politecnico di Bari, via Amendola 173, I-70126 Bari, Italy

[d] Istituto Nazionale di Fisica Nucleare, Sezione di Bari, via Amendola 173, I-70126 Bari, Italy

[e] Dipartimento di Scienze Economiche e Metodi Matematici, Università degli Studi di Bari, via C.Rosalba 53, I-70124, Italy

ARTICLE INFO

*Article history:*
Received 22 June 2009
Received in revised form 11 December 2009
Available online 5 January 2010

*Keywords:*
Operational Risk
Complex systems
Bayesian Networks
Time series
Value-at-risk
Different-times correlations

ABSTRACT

A system for Operational Risk management based on the computational paradigm of Bayesian Networks is presented. The algorithm allows the construction of a Bayesian Network targeted for each bank and takes into account in a simple and realistic way the correlations among different processes of the bank. The internal losses are averaged over a variable time horizon, so that the correlations at different times are removed, while the correlations at the same time are kept: the averaged losses are thus suitable to perform the learning of the network topology and parameters; since the main aim is to understand the role of the correlations among the losses, the assessments of domain experts are not used. The algorithm has been validated on synthetic time series. It should be stressed that the proposed algorithm has been thought for the practical implementation in a mid or small sized bank, since it has a small impact on the organizational structure of a bank and requires an investment in human resources which is limited to the computational area.



## 1. Introduction

In the past years a powerful set of tools to study complexity has been developed by physicists and applied to economic and social systems; among the several topics under investigation the quantitative estimation and management of several typologies of risks [1], like financial risk [2–6] and operational risk [7,8] has recently emerged.

*Operational Risk* (OR) is defined as "the risk of [money] loss resulting from inadequate or failed internal processes, people and systems or from external events" [9], including legal risk, but excluding strategic and reputation linked risks. Since it depends on a family of heterogeneous causes, in the past only few banks dealt with OR management. Starting from 2005 the approval of "*The New Basel Capital Accord*" (Basel II) has substantially changed this picture: in fact OR is now considered a critical risk factor and banks are prescribed to cope with it setting aside a certain capital charge.

Basel II proposes three methods to determine this capital: (i) the *Basic Indicator Approach* sets it to 15% of the bank's gross income; (ii) the *Standardized Approach* is a simple generalization of the Basic Indicator Approach: the percentage of the gross income is different for each Business Line and varies between 12% and 18%; (iii) the *Advanced Measurement Approach* (AMA) allows each bank to use an internally developed procedure to estimate the impact of OR. Both the Basic Indicator Approach and the Standardized Approach seems overly simplistic, since in some way they suppose that the exposure of a bank to operational losses is proportional to its size. On the other hand, an AMA not only helps a bank to set aside the required capital charge, but may even allow the *OR management*, in the prospect of limiting the amount of future losses.

* Corresponding author at: Dipartimento Interateneo di Fisica "M. Merlin", Università degli Studi di Bari e Politecnico di Bari, via Amendola 173, I-70126 Bari, Italy. Tel.: +39 0805442178.
*E-mail address:* marco.bardoscia@ba.infn.it (M. Bardoscia).

0378-4371/$ – see front matter 
doi:10.1016/j.physa.2009.12.043

Each AMA has to take into account two types of historical operational losses: the internal ones, collected by the bank itself, and the external ones which may belong to a database shared among several banks. Nevertheless, due to the recent interest for OR, only small and not adequately accurate historical databases exist and this is why each AMA is required to use also assessment data produced by experts. In addition, Basel II provides a classification of operational losses in 8 Business Lines and 7 Loss Event Types which has to be shared by all the AMAs. Finally, AMAs usually identify the capital charge with the Value-at-Risk (VaR) over the time horizon of 1 year and with a confidence level of 99.9%, defined as the maximum potential loss not to be exceeded in 1 year with confidence level of 99.9%, i.e. the 99.9 percentile of the yearly loss distribution; this implies that the probability of registering a loss being less than or equal to the value of the VaR in 1 year is equal to 0.999 or, equivalently, that a loss larger than the value of the VaR may occur on average every 1000 years.

Among the AMA methods, the most widely used is the *Loss Distribution Approach* (LDA). In LDA the distribution of frequency and the distribution of impact (severity) modeling the operational losses are separately studied for each of the 56 pairs (Business Line, Loss Event Type). LDA makes two crucial assumptions: (i) frequency and severity distributions are independent for each pair; (ii) the distributions of each pair are independent from the distributions of *all the other* pairs. In other words LDA neglects the correlations possibly existing between the frequency or the severity of the losses occurring in different pairs.

The idea of exploiting BNs to study OR has already been proposed in Refs. [10–14], and various approaches are possible. The main advantages offered by BNs are two:

- the possible correlations among different bank processes can be captured;
- the information contained into both assessments and historical loss data can be merged in a natural way.

One approach [15,16] may be to design a completely different network for each bank process, trying to determine the relevant variables (in the context of each process) and the causal relationship among them; this kind of network has only one output node which typically represents the loss distribution for the process under investigation; the correlations among different processes can be captured by building a "super-network" which contains all the networks built for the single processes and in which the nodes representing the loss distributions of the processes may be connected by links. Since it deals with the variables governing the underlying dynamics of the bank, this approach seems to be the most convincing; nevertheless it suffers from some drawbacks as regards the practical implementation inside a bank: (i) domain experts are needed for each process, in order to properly identify the variables and to define the topology of each network; (ii) if the historical data need to be used, a system monitoring all the included variables with an acceptable frequency and accuracy has to be built; since this kind of network can easily reach large sizes (tens of variables), managing such systems is quite challenging and resource demanding for a mid or small sized bank.

A simpler approach [17] is to design a unique network composed by a node for each process which represents its loss distribution; all nodes are output nodes and the operational losses are sufficient to build a historical database, so that collecting the data and managing them is much easier; in comparison with the previous approach even the experts' task becomes simpler since their assessment reduces to an estimate of some parameters of the loss distributions; the correlations among different processes are captured through the topology of the network. This approach resembles a way of reasoning typical of the field of Complex Systems: the information carried by the "microscopic" degrees of freedom (the relevant variables identified in the first approach) is integrated out and the state of the system is represented by some "macroscopic" quantity (the loss distribution in the second approach).

Let us remark that, as regards the practical implementation inside a bank, the difference between the two approaches is huge: in the first approach tens of variables for each process need to be monitored, while in the second approach only one variable per process (the registered losses) has to; considering that an AMA-oriented bank has to track its own internal losses in any case, the cost of the proposed implementation is minimum.

## 2. Bayesian Networks

Before defining a Bayesian Network [18–20] let us introduce some general definitions about graphs. A directed graph is defined by a set of nodes $\{X_1, X_2, \ldots, X_N\}$ and by a set of directed links between couples of nodes; a node $X_1$ is said to be a *parent* of the node $X_2$ if a directed link $X_1 \to X_2$ exists; a node $X_2$ is said to be a *descendent* of the node $X_1$ if a directed path which starts at $X_1$ and ends at $X_2$ exists; if no such a path exists the node $X_2$ is said to be a *non-descendent* of the node $X_1$; a directed path from a node to itself is called a *directed cycle*; a directed graph containing no directed cycles is called *Directed Acyclic Graph* (DAG).

In order to define a Bayesian Network two elements are necessary: a set of random variables $V = \{X_1, X_2, \ldots, X_N\}$ and a DAG whose nodes correspond to the random variables in $V$; note that, since there is no risk of ambiguity, we use the symbol $X_1$ to refer both to the random variable and to the corresponding node. Moreover, the joint *Probability Distribution Function* (PDF) $P(X_1, X_2, \ldots, X_N)$ must satisfy the Markov condition, i.e. each random variable $X_i$ and the set of all its non-descendents must be conditionally independent, given the set of all its parents. It can be proved for discrete variables (which turns out to be our case) that the Markov condition easily allows one to calculate the joint PDF as:

$$P(X_1, X_2, \ldots, X_N) = \prod_{i=1}^{N} P(X_i|\mathrm{Pa}_i), \tag{1}$$

where $\mathrm{Pa}_i$ is the set of random variables whose corresponding nodes are parents of the node $X_i$.