The International Journal of Business Management and Technology, Volume 6 Issue 5 September-October 2022
ISSN: 2581-3889

Research Article Open Access

# Aggregating Operational Risks for Risk Reporting – Machine Learning Bayesian Networks

**Dr. Martin Leo, Dr. Suneel Sharma, Dr. Dhrupad Mathur**
*S P Jain School of Global Management*

**Abstract:** *Operational risk is managed through internal/external loss data, key risk indicators, risk and control self-assessments, scenario analysis, and requires to be measured at various organisational levels. Educated risk management decisions can be taken better when the framework allows for the integration of risk and control information, with aggregation happening across the bank. The use of a Bayesian Network (BN) provides such a model that allows for the integration of risk and control information to be delivered in a structured process. This article provides a practical framework for the aggregation of operational risk data for risk reporting through the learning of Bayesian networks. The parameters of the BN is initially learnt from an incident database and subsequently updated with expert opinion. This framework can be adopted and adapted by a risk manager enhancing their ability to deliver dynamic and timely risk intelligence for effective management of operational risk on a day-to-day basis.*

**Keywords**: Operational risk, Risk aggregation, Bayesian Networks, Machine Learning, risk reporting

## I. INTRODUCTION

Operational risk (OR) is defined by the Basel Committee on Banking Supervision (BCBS) as the risk of loss resulting from "inadequate or failed internal processes, people and systems or from external events" and is a "fundamental element of risk management" at banks[1]. In a review of Bank annual reports, OR was varyingly presented and included several sub risks and could be referred to more as a non-financial risk. [2].

Approaches typically followed in OR modelling are either top-down – based on macro data or bottom-up based on the identification of individual events or loss causes[3]. Risk measurement is a key tool in the capture and controlling of risks. Capital allocation and the establishment of criteria for objectivity and comparability in prioritising risk control for the improvement of the internal control environment are among the reasons for risk measurement. OR has, for a large part, been managed through qualitative risk management practices (e.g., checklist, operations manuals), which, while relatively easy to implement, have various limitations[4]. ORM should not be viewed as a set of disjointed tasks but a structured process that allows for educated risk management decisions when risk and control information is integrated[5].

Risk management traditionally focuses on risk identification, measurement, and monitoring, including designing mitigation strategies. Risk management typically seeks for risks to be managed holistically in a fully integrated framework across all the various risk types and functions within the enterprise[6]. Key decision-makers should be involved in the risk management process, and they must understand the interconnectivity of risks, thereby avoiding ignoring the risks that do not pertain to their domain [7]. Typically, to manage risks in banks, there are particularly common and widely used methodologies in the risk manager's toolkit (e.gRisk and Control Self-Assessment (RCSA), stress testing, scenario analysis, risk appetite frameworks, risk dashboards, limits, key risk indicators (KRI), exposures, VaR)[2]. The pertinent problem in this context is how a bank can capture and aggregate the data from the various techniques, and the risks, to enable internal stakeholders to determine and differentiate key influencers enhancing the bank's capability to manage enterprise-wide risks.

Probabilistic Graphical Methods (PGM) supports a data-driven approach to effective model construction[8]. We see that BNs have been widely explored for OR capital calculation and related risk exposure reporting and disclosures[9][3][10]. They have also been explored to analyse better and understand the OR related to specific process areas or sub risk areas[11][12]. Through machine learning a BN model from loss/incident data the ability to measure, monitor and communicate risk information can be enhanced [13]. In this paper, we extend the research done, exploring the use of BNs in the modelling of OR to enable risk managers and business managers better manage OR on a day to day basis, i.e., "business-as-usual" operational risk management (ORM).

Literature has predominantly focused on addressing the 'risk capital calculation' problem, aligned more with the regulatory requirements and directives from Basel. As a risk manager in the industry, one struggles to find practical references that could be adopted and adapted to enable and enhance the day-to-day ORM capability. While literature covers how risk should be measured and aggregated at the organisation level, it is not easy to find literature on how risks can be aggregated at various levels in the organisation to facilitate regular risk reporting to business operational or line managers for risk management. A risk manager does not have a robust methodology for aggregating the various risks being identified and monitored. While risk reporting is standard practice, much subjectivity is applied, for the aggregation of risk and risk information into a single metric or view or rating. The subjectivity is prevalent even in the presence of high-level rules. A consequence of this subjective method of aggregation is disjoint between the reported view and the underlying data and further the forward path.

The objective of this research is to evaluate and develop, through a practical approach, a framework appropriate for business-as-usual risk management, for the aggregation of operational risk data at multiple levels:

a) At the risk or risk event level – the aggregation of data from different risk techniques. Can data from the various risk management techniques – incidents, control testing, and RCSA – be aggregated to arrive at the probability of a risk or risk event?
b) At the Basel operational risk event type – the aggregation of risk or risk events. Can the probability of risk events for an operational risk type be aggregated to arrive at the probability of the Basel operational risk event type?
c) At the business line level – the aggregation of the operational risk event type. Can the probability of Basel operational risk event types be aggregated to obtain the business line operational risk?
d) At the group level – the aggregation of business line operational risk. Can the operational risk probabilities for each business line be aggregated to obtain the aggregated operational risk for the group?

Section 2 presents a review of the literature available and an analysis of the operational risk topics researched applying Bayesian networks. Section 3 elaborates on the research design and the methodology followed to build and analyse the BN using simulated data. Section 4 provides a discussion of how the BN can be applied for operational risk management, with section 5 providing a conclusion.

## II. Literature review

### 2.1 Operational Risk Management

The objective of an OR framework is to identify, assess, control and mitigate OR and to champion effective reporting of risk and emerging risk issues [14].[15] presents the main data building blocks of an OR framework in Figure 1. Risk identification and assessment are fundamental characteristics of an effective OR framework. The bank can better understand its risk profile and effectively target risk management resources and strategies through a sound risk assessment process.

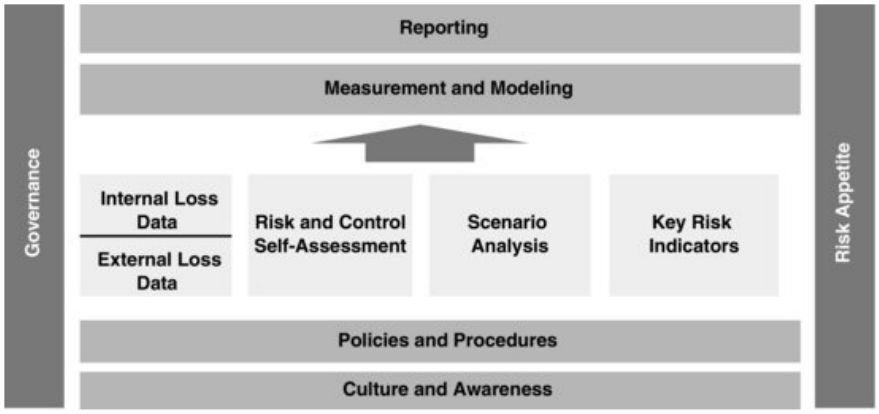


**Figure 1 Building blocks of an OR framework. Source: Girling (2013)**

Operational risks are diverse, originating from internal or external disruptions to business activities, with a level of unpredictability of their overall financial impact. This complicates the systematic and coherent measurement and regulation of operational risks making it quite distinct from that of other types of banking risks.

Losses are entered into a database and attributed to loss categories with sufficiently detailed classification for risk measures. The approach taken by BCBS in the classification of OR losses is to deal in terms of the causes of the loss, as is seen from the OR definition that focuses on the cause of the risk and the loss event type classifications. Modelling of the cause of OR losses is required for comprehensive analysis and is essential for understanding how risks arise within the organisation. It provides management with a basis on which management may intervene to achieve the desired alteration in the risk profile [16].

Events captured in the firm's OR event database provide a valuable backdrop and help identify the risks and control weaknesses to be addressed in the RCSA. They also demonstrate the possible impact and frequency of risk events and can be used to validate assessments made during the RCSA. The RCSA is designed to consider all possible risks, not just those that have occurred in the firm, which can be a difficult task. External events, events that have occurred in the industry are useful for this purpose and enable informed discussion about potential risks [15]. Internal operational loss data provides meaningful information for assessing a bank's exposure to OR and internal control effectiveness. Analysis of loss data can provide insight into the causes of losses and control failures. External loss data can be used to assess previously unidentified risk exposures [1].

The objective of a risk and control assessment (RCA) or risk and control self-assessment (RCSA) is to identify, measure and monitor the risks and controls of the firm. The RCA can be qualitative, based on value judgments, or quantitative, assessing the risks identified through percentages for likelihood and impact values [17].

The ability to effectively identify, assess, measure and manage risk is vital for OR excellence and robust risk management. For most organisations, the risk and control self-assessment (RCSA) process is central to the ORM framework [18]. The RCSA provides an excellent opportunity to integrate and coordinate the risk identification and risk management efforts and generally improve the understanding, control, and oversight of a bank's operational risks. RCSA provides a systematic means of identifying control gaps that could impact the business or process objectives. In this respect, RCSA promotes the analysis and monitoring of factors that affect the level of OR exposure. A key aspect of the process is the assessment of the impact and likelihood of risk. There are three main approaches to performing an RCSA, namely workshop, questionnaire or a hybrid [19], [20].

The loss database also demonstrates the possible impact and frequency of risk events and can be used to validate assessments made during the RCSA. RCSA forms an integral element of the overall OR framework, as it provides an excellent opportunity for a firm to integrate and coordinate its risk identification and risk management efforts and generally to improve the understanding, control, and oversight of its operational risks. The results and findings from the RCSA are used in conjunction with other components of the OR framework (e.g., loss data, scenario analysis) to enhance insight into the firm's OR profile [19].

Scenario analysis is a popular input in the OR measurement methodology and considered a successful, forward-looking technique. Scenario analysis is a process of obtaining an expert opinion from the business line and risk managers to identify potential OR events and assess potential outcomes [1].

Key Risk Indicators are used to monitor the main drivers of exposure associated with key risks. These are often paired with escalation triggers to warn when risk levels exceed acceptable thresholds [1]. KRIs are measurable metrics or indicators that track exposure or loss. Anything that can perform this function may be considered a risk indicator, and it becomes key when it tracks an especially important exposure. An underlying rationale is usually that a change in the value of the indicator is likely to be associated with a change in risk exposure or operational loss experience.

**Risk Aggregation**

A key element of risk management is a framework that allows for a consolidated view of all risks, enabling the board and senior management to take the appropriate decisions and actions [21]. Risk aggregation opens the door for 'Firmwide management of risk'. Risks can be aggregated either between risk types or between business units, with the former considered more amenable to the management of risk in an integrated fashion [22]. Financial institutions and their supervisors are placing increased emphasis on consolidated risk management, also called integrated or enterprise-wide risk management. The interdependent nature of risks within an organisation also serves as a motivation for developing a consolidated risk management system [23]. The enormous improvements in computing technologies have increased the practical possibility of risk aggregation [24].

A modern diversified financial institution engages in a broad set of activities (e.g., banking, brokerage, insurance or wealth management). Rapid financial innovation, developments in supervisory standards and the financial crisis has created a challenge for enterprise risk management in diversified financial institutions. Integrated risk management requires developing a coherent approach to aggregating different risk types. The goal of integrated risk management in

a financial institution, having a range of diverse business activities, is to measure and manage risk and capital across the institutions. Studies in this area have primarily focused on copulas application, estimating joint distributions risk and capital across a range of diverse business activities [25].

Risk aggregation provides valuable risk intelligence and good risk aggregation helps an institution in developing a better understanding of the breadth of individual risks and enables an efficient and effective risk management program [26], [27].

**Risk Reporting**

The ability to completely and accurately aggregate, analyse and report OR exposure is an essential capability of robust risk management. Accurate risk information provides 'intelligence' firms need to make informed, risk-based decisions on a day to day basis. Reporting through risk dashboards can be used to alert management of the changing risk conditions and thereby support decision making [28].

One of the ultimate challenges for banks is being able to take raw, granular, business-level data and aggregating them across levels in a manner that efficiently and effectively communicates risk intelligence. Such risk intelligence should be communicated within the bank both horizontally and vertically to enable executive management to assess the risks in running the organisation. The ability to do a bottom-up data aggregation and then across business lines, regions should exist within the organisation[29]. Gathering risk information from the different business units/departments and consolidating them into risk reports for senior managers and boards can be daunting for many organisations. The volume of risk data to be aggregated can be overwhelming, especially in organisations where risk is managed in silos within operational units. Lack of awareness of correlations and concentrations could allow risk to spread across the silos [26].

Improved reporting of risks enables managers to make informed decisions by allowing them to adequately consider the identified and measured risks leading to organisational success and increase in shareholder value. To address the needs of different stakeholders, each having their objectives and requirements, the reporting of risks should be at various levels. While operational level managers could require information relevant, accurate and periodic to their specific area, senior management's requirements may be of a different specificity creating reporting requirements specific to varying levels of decision making and control within the organisation. This would require risk reports to be comprehensive across all risks, with the level of reporting detailed or aggregated catering to the audience [30].

The goal of any OR framework should be to provide management with the information and resources required to take these types of decisions. Further, RCSA information must be aggregated to create a comprehensive view of operational risks across the organisation. The information must be tailored to support management decisions on prioritisation of risk mitigation programs [31].

## 2.2 ORM and Bayesian Networks

A graphical model is viewed as a probabilistic database, a machine to answer queries about the values of sets of random variables. The database is built applying probability theory, ensuring a consistent overall interpretation[32]. PGMs use a graph-based representation as the basis for compactly encoding a complex distribution allowing for a compact representation of a set of independencies and also defining a skeleton for representing a high-dimensional distribution. There are two families of graphical representations of distributions – BNs, uses a directed graph, and the second Markov networks use an undirected graph[8].

The development of OR models through a systems-oriented approach allows the models to be extended beyond risk capital allocation. Such a model requires a deeper understanding of causation in OR events incorporating antecedent events and causal pathways of operational losses. In a continually changing environment, the estimation of OR frequency and severity distributions cannot be done solely using historical data, given their limitations in predicting future losses. [11].

Any model that incorporates subjective prior beliefs is a form of the Bayesian model. OR modelling uses observable (objective) data and subjective choice (prior belief) parameters and more suited for BN modelling. BNs combine intuitive visual representation benefits with a sound mathematical base to allow for reasoning under uncertainity. BNs enable us to combine any quantitative data available with qualitative data, mirroring the causal structure underlying the process itself. BNs can successfully model dependencies between events and processes in

complex systems.[33]. BNs improve transparency for efficient risk management, as they are based on causal flows in an operational process[34].

In the area of OR, BN methodologies can be applied to build models that combine different sources of data (internal data, external data, and expert opinion), learning correlation structures, and incorporating dynamic information flows. BNs can be used for scenario analysis. [3], [9], [34]–[38].

BNs can be used for eliciting subjective components of risk forecast from experts by explicitly modelling scenarios involving different operational processes or threats to the business, with the likely outcome. Further on, they can model "long" tail distributions for the unexpected loss component of the total loss distribution [39]. They also allow for a domain's changing risk profile to be captured in an efficient, low cost, timely manner, supporting both individual and organisational learning. Risk communication, within the organisation, can be improved through a clear representation of OR and their subsequent losses and key drivers more explicit [40]. Possible correlations among different bank processes can also be captured [38].

Bayesian methodology can be used for estimating OR loss distributions using loss data and expert opinions. The elicitation process of expert opinion, through an RCSA, allows a better understanding and sharing of operational risk. RCSA data can be included in a suitable parametric prior distribution and through a Bayesian approach be merged with loss data to derive posterior distributions for deriving appropriate risk measures such as Value at Risk (VaR) or Expected Shortfall (ES) [9].

In a dynamic environment, such as the banking industry, there is a constant flow of information in the form of observed losses, new or changed policies, and controls. It is useful to incorporate this flow of information into the models. The Bayesian model allows for the efficient integration of these information flows to deliver timely benefits and model validation [35]–[37], [39]–[41]

A Bayesian approach is a viable option for the managing of OR in an environment of uncertainty and scarce information. While the application of BN to ORM has been explored, much of the focus has been around their application to risk measurement for capital calculation or regulatory disclosures, such as calculation of the Value at Risk (VaR) or loss distribution analysis [11], [42]–[44]. Several papers have extensively explored process and loss event analysis using BN modelling oriented more towards the better analysis and understanding of OR related to specific process areas or sub risk areas [11], [12].

Figure 2shows the distribution of papers reviewed (47) by the authors that covered both BN and operational risk, which addressed various terms related to operational risk management. Over 90% of the papers include loss data only, while very few cover KRIs or RCSAs. Figure 3 shows the number of papers that cover all four areas of operational risk management and how BN can be applied; this was found to be just one. Previous research has mostly focused on capital related calculations, with about 87% of the papers focused on VaR or capital calculation. This shows a gap in the literature covering the application of BNs beyond capital calculation and business-as-usual risk management.

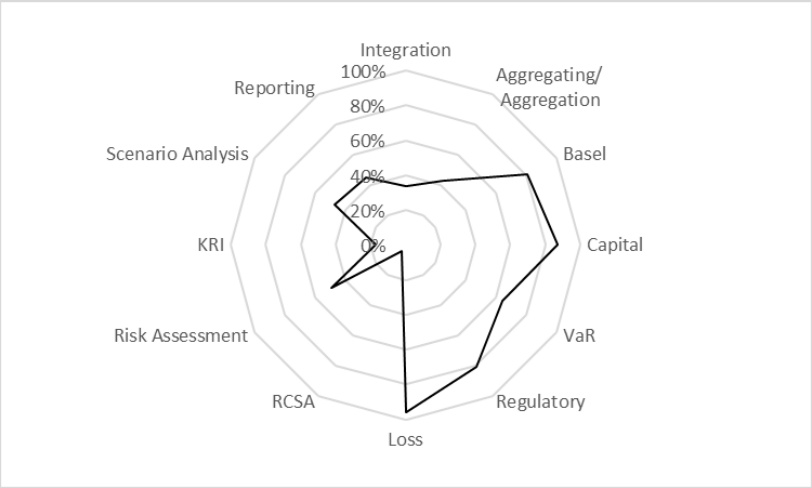


**Figure 2 Percentage of ORM-BN papers reviewed that covered the specific ORM topics**

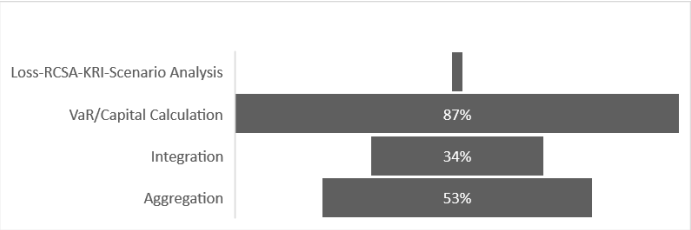


**Figure 3Percentage of ORM-BN papers that addressed the four ORM techniques together in the same paper**

Papers tend to either focus on the use of subjective data or objective data for structure and parameter learning. While the need to aggregate risks and have integrated risk reporting has been discussed there is limited reference to how this can be achieved within the institution, especially for regular reporting at the business unit or business line levels. The papers also refer to the integration of the various ORM techniques – RCSA, loss analysis, KRI, scenario analysis – but there is little in detailing how this can be achieved incorporating both a forward and backward-looking view. We see that further research is required to elaborate on how the BN models, can be constructed and learned, and then be applied in the effective management of OR at a bank or unit level. The pertinent and pressing question is how these models can be incorporated into the day to day risk management and not just for regulatory reporting or risk calculation purposes. Also, to understand how the models can enable risk management to be viewed at a bank-wide or bank unit-wide level and not just specific to a process or sub-risk type.

## III. Materials and methods

Figure 4 shows the research design through which the above objectives are to be achieved. Data from an operational risk incident database, including various types of operational risk incidents, will be used to construct and learn a Bayesian Network (BN) initially. This BN will then be updated with expert opinion, which in the business environment will be elicited through techniques such as risk assessments. The BN will then be evaluated to assess the aggregation capabilities and for purposes of operational risk management. The aggregation will be done at an operational risk type – aggregating all the risks related to that operational risk type, followed by at the business line level and then at the group level. The constructed networks will be evaluated for application to various business-as-usual operational risk management techniques such as – operational risk reporting, scenario analysis, loss analysis, key risk indicators.

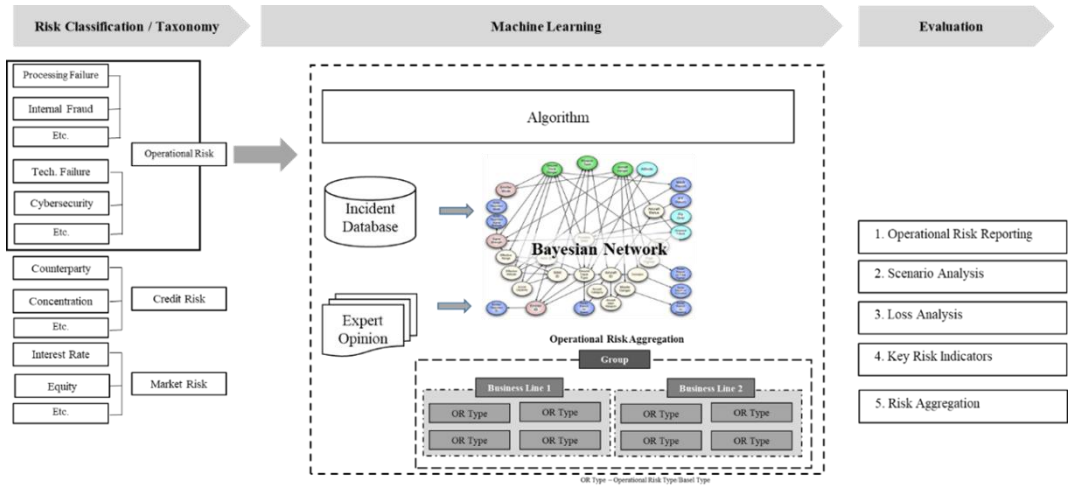


**Figure 4 Overview of the research design showing the construction of the BN and the application to evaluate the objectives of this research**

A banking group can have several business lines (retail banking, commercial banking, asset management). A Bank would have in place a taxonomy including all risk events the bank may be exposed to and also all the controls that is implied or should be implemented in the bank. A bank will have an OR incident database to capture all OR incidents

reported in the bank, in line with the internal policies and procedures. Additionally, the bank would have a risk register that captures all the current risks the bank is exposed to and the risk exposure level.

**Figure 5** shows the conceptual design of the research in the context of ORM at a bank.

The incident data available in an incident database, which includes the cause of the incident and the details of the control or control suite that has failed, can be used to determine the probabilities. From the incident data the probabilities of the cause being present, a control (or control suite) failing; consequently, a risk event occurring resultant from the cause and control failure and therefore probabilities of the risk type can be determined. Further, the data provides the probabilities of the risk impact for a risk event or risk type. The bank would seek to analyse the incident database to understand the backward looking probabilities of the various causes of a risk event, including:

- A control suite failing in the presence of a cause;
- A risk event materialising;
- A risk type occurring;
- The risk impact – high, medium, low – for a risk event or risk type.

The probabilities of the control failure can be updated based on the control testing that may be undertaken by the bank. This testing can include regular internal audits and activities to provide assurance on internal control systems (BCBS, 2012; Federal Deposit Insurance Corporation, 2015). A control found to be effective will have a lower probability of failure than one found to be not effective, which will have a higher probability of failure with the potential to result in risk events.

Probabilities can be updated with information elicited through the process of RCSA or scenario analysis. This allows for a forward view of cause occurrence or control failures to be incorporated by eliciting them from subject matter experts. The information from control testing and RCSAs are used to update the probabilities of the control (failure) and cause (occurrence). The probabilities of the risk event, risk type, and risk impact are propagated when the updates are made.

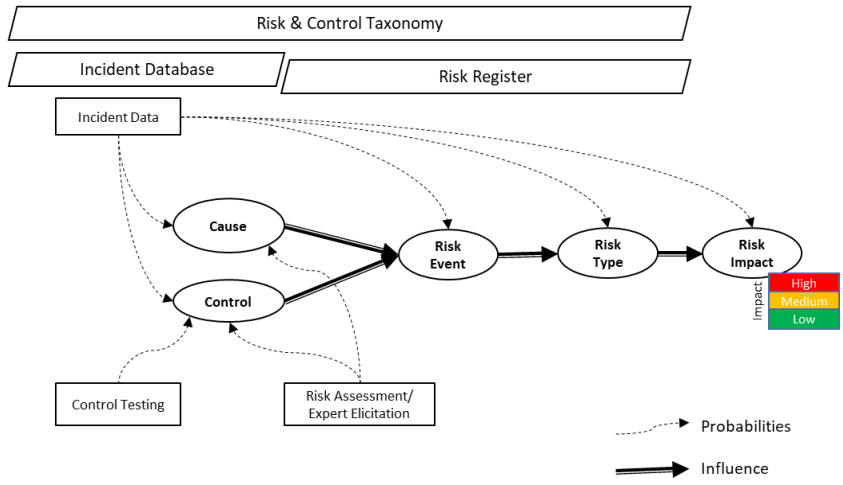


**Figure 5 Conceptual design of the research**

Figure 6 shows the process for this research for learning the structure and parameters of the BN and then applying it for operational risk management. A master database was synthetically created, to resemble a risk incident database of a bank. The database was synthetically created solely for modelling, to exhibit their application to use cases, a few limitations do apply. It will not be used for any empirical analysis, and relationships are only illustrative. Also, not all Basel OR Event Types have been considered. While a bank's OR incident database is likely to capture the actual loss amount ($ amount) in this case, we have assumed a discrete variable to indicate the risk impact. This allows for the capture of all potential impact types – monetary, regulatory, financial. Also, this allows for the database to be developed on for further application use cases. The master database was then randomly split into two databases representing the incident database of two business lines.

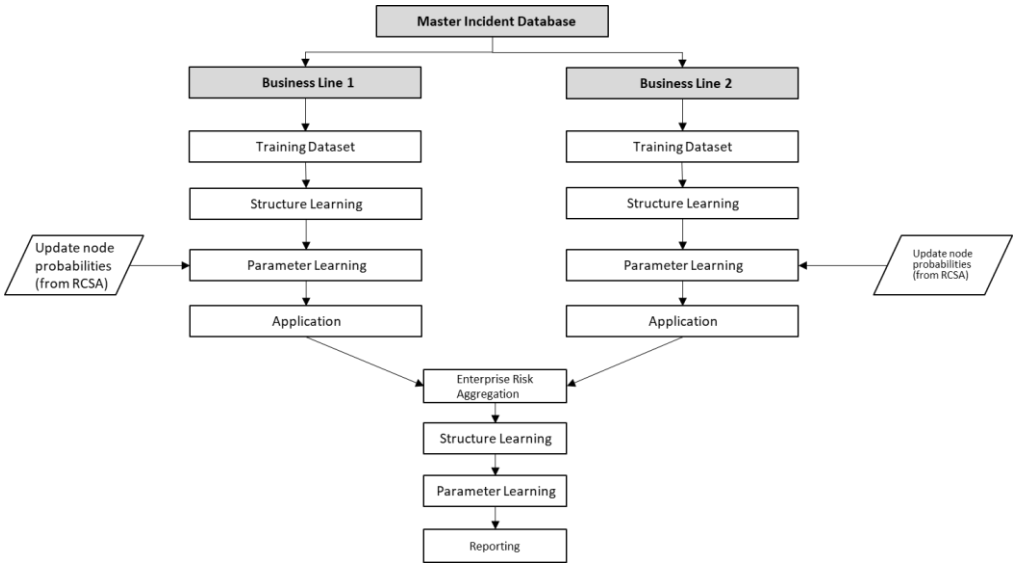


**Figure 6 Process to learn the BN and apply the BN**

Table 1 describes the groups of fields, the fields and shows a mapping between the fields. One group of fields in the control suite – to indicate a family of related controls that could fail – with the fields being CS1, CS2. Another field group is 'Cause' which relates to the cause of the risk event, namely Delivery Failure, Outsourcing. The other two field groups are 'Risk Event', with fields Service Disruption #1, #2 and 'Risk Type', capturing the Basel OR event types. The values for the fields are 'Y' or null, the former if the cause is present or control suite has failed, or the risk event has occurred. The mapping shows the relationship between the control suite and cause. The failure of a control suite, with the related cause being present, could result in the occurrence of the risk event. A mapping of the cause, risk event to the risk type is also shown below. An additional field is the 'Risk Impact' with values of High (H), Medium (M), Low (L). This is a discrete value to show the level of risk exposure. Each of the fields, as shown below, will represent a node when the BN is structured later.

| Control Suite | Cause | Risk Event | Risk Type |
|---|---|---|---|
| CS5 | Delivery Failure (DF) | Service Disruption (SD) #1 | Business Disruption and Systems Failure (BDSF) |
| CS9 | Outsourcing (Out) | Service Disruption (SD) #2 | Business Disruption and Systems Failure (BDSF) |
| CS12 | Hardware Failure (HW) | Technology Disruption (TD) #1 | Business Disruption and Systems Failure (BDSF) |
| CS8 | Telecommunications (Telecom) | Technology Disruption (TD) #2 | Business Disruption and Systems Failure (BDSF) |
| CS11 | Utility Outage/Disruption (UOD) | Technology Disruption (TD) #3 | Business Disruption and Systems Failure (BDSF) |
| CS7 | Hacking Damage (HD) | Asset Loss (AL) #1 | External Fraud (EF) |
| CS11 | Natural Disaster (ND) | Asset Loss (AL) #2 | Damage to Physical Assets (DPA) |
| CS2 | Product Defects (PD) | Product Loss (PL) #1 | Client, Products, Business Practices (CPBM) |
| CS1 | Model System Misoperation (MSM) | Processing Error (PE) #1 | Execution, Delivery, & Process Management (EDPM) |
| CS3 | Software (SW) | Processing Error (PE) #2 | Execution, Delivery, & Process Management (EDPM) |

**Table 1 Fields of the Incident Database and a mapping between the field groups**

Learning – structure and parameter – was done through packages available in R. Inference of probabilities were also through the relevant packages available in R.[13] elaborate the approach for learning BN models using R.

The structure of the BN was constructed per the model in **Figure 7**. The model is based on the Fenton and Neil risk event model explained through the causal view of risk[47]. The model has been modified, ignoring the mitigating variable in the original model while renaming consequence as risk impact. Additional variables were added to record the Basel operational risk type to which the risk event pertains this providing a structure more suited for aggregation of risk types. **Figure 7** (a) shows the risk event model, as explained through the causal view of risk. **Figure 7** (b) is a proposed model for this research modifying the original Fenton & Neil model.**Figure 7** (c) provides an example of the model showing the relevant nodes. Telecom failure is the cause and, if present, given the failure of the associated control, there would be a risk event – technology disruption. Technology disruption relates to the risk type business disruption and systems failure (BDSF). The risk event would have a risk impact that could be a random value – low, medium, or high. The incident data within the incident database was initially used to learn the probabilities for the BN; namely, the probability of a risk event resulting from a cause-control combination, the probability of risk types, and the probabilities of the risk impact.

**Figure 8** shows the BN structure that has been manually constructed using bnlearn and will be the basis for this study.

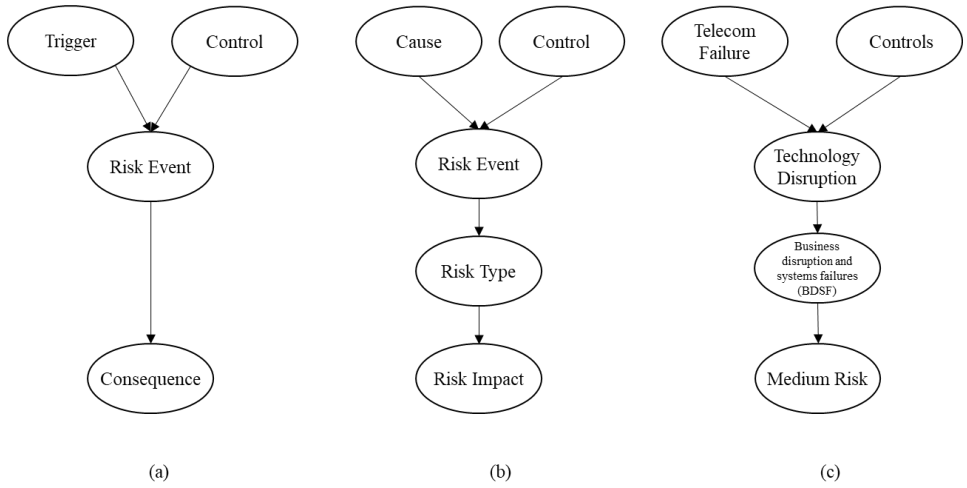


**Figure 7Structure for the construction of the BN (a) causal view of risk**[47]**, (b) modified structure of the causal view of risk, and (c) illustration of the model using a risk event as an example**

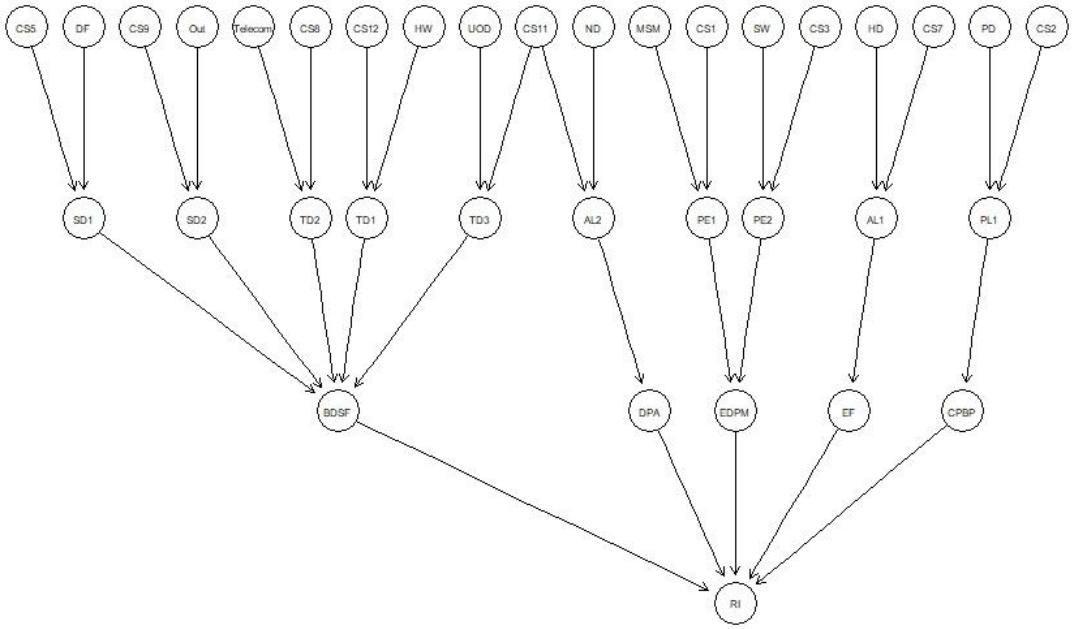

**Figure 8 BN structure constructed using bnlearn**

With the structure of the network defined, the parameters, the node probabilities, and the conditional probabilities can be learned from the database by the algorithm.

**Risk Assessment – Incorporating Expert Opinion**

BNs allow for a combination of both subjective and objective data in quantification. Risk assessment information can be incorporated into the risk model. Learning from the incident database provides the initial 'objective' data and risk assessments provide the 'subjective' data that can be incorporated into the BN. The parameters (node and conditional probabilities) learned from the incident database can be updated through expert opinion collected through the RCSA process.

The probabilities for the BN are learned initially from the incident database (**Figure 8**). The node probabilities for the parent nodes – cause and control suites are from the BNs that were learned using the incident database. The conditional probabilities for the child nodes in the BN are also learned and inferred from the incident database BN.

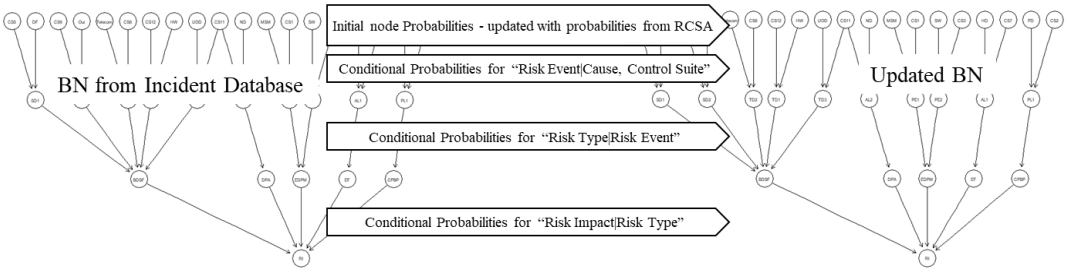


**Figure 9 Constructing the BN incorporating expert opinion**

Probability elicitation is the process of obtaining estimates from domain experts, and this can be done through the RCSA or RCA process for OR management. Probabilities comprise both marginal and conditional probabilities. This can be very challenging if the probability is conditioned on several states and nodes.

Table 2 illustrates how the information elicited through risk assessments for business line 1 can be incorporated into the BN. The values in the pre-assessment column refer to the probabilities as inferred from the BN parameters that were learnt from the incident database. This shows the probabilities of control suite #7 failing and also the probability of the cause being HD, based on the incident information collected. The probability of an AL1 risk event as inferred from the incident database could, therefore, be understood to be very low.

During an RCSA process, given the increase in cybersecurity threats [48], the risk manager and business managers could discuss and agree to increase the probability of HD being a cause for risk events to 0.4. Given the change in the threat and the need to enhance the control suite, the managers can agree that the probability of a failure in CS7 has increased to 0.2. This shows a higher probability of occurrence of the cause and failure of the control suite than in past incidents.

The BN is updated by learning the parameters after providing these updated node probabilities (* indicates the values elicited from the RCSA and input into the BN). Inferences are then made from the updated BN. The risk manager and business manager can infer, as shown under the Assessment column, the updated or new probabilities. Updating the node probabilities for HD and CS7 shows that the marginal probability of an AL1 risk event for the bank is now much higher (0.08 vs 0.00). The conditional probability of a high risk impact event occurring, with HD being the cause and CS7 failing, is now significantly higher, at 0.3960 (vs 0 earlier). This allows the risk managers and business managers to reset their priorities for risk remediation focus.

| | Pre-Assessment | Assessment |
|---|---|---|
| Pr(HD = Y) | 0.0019 | 0.4000* |
| Pr(CS7 = Fail) | 0.0019 | 0.2000* |
| Pr(AL1 = Y) | 0.0000 | 0.0800 |
| Pr(RI == "H" \| HD == "Y" & CS7 == "fail") | 0.0000 | 0.3960 |

| | | |
|---|---|---|
| Pr(EF == "Y" \| AL1 == "Y") | 0.0000 | 1.0000 |
| Pr(RI == "H" \| AL1 == "Y") | 0.0000 | 0.4071 |
| Pr(RI == "H" \| EF == "Y") | 0.0000 | 0.3949 |

**Table 2 Risk assessment business line 1**

In another case, the probabilities of the occurrence of the cause – model system misoperation - and its related control suite (CS1) was updated to lower probabilities 0.1 and 0.3 respectively from 0.3555 and 0.3555 earlier. This was to simulate an assessment where actions were taken to enhance the control environment and also reduce the causal factors. When the BN was rerun, the probabilities were automatically updated. The probability of the occurrence of the relevant risk event, processing error #1, showed a decrease to 0.0300 from the initial 0.1264. The probability of a high impact risk conditional the occurrence of a processing error #1 risk event had dropped to 0.2563 from 0.2721. While past incidents reflected a higher probability of the risk event happening and therefore a higher risk exposure – this could be reassessed and updated based on new information of improved control environment and a revised assessment of the probabilities.

The parent nodes – cause and control (suite) – probabilities can be updated based on the inputs from the RCSA and scenario analysis. This is then propagated through the BN automatically updating the probabilities for the nodes such as risk event, risk type and risk impact. In addition to the probabilities for the cause and control, the risk managers and business managers can also assess a need to update the conditional probability tables to reflect a changed view in light of environmental changes.

Risk assessment information can thus be incorporated into the risk model. The parameters (node and conditional probabilities) learned from the incident database can be updated through expert opinion collected through the RCSA process.

**Risk Aggregation**

One of the objectives of this research is also to address the problem of aggregation of risks at an enterprise or group-wide level. While in the above sections risk information can be aggregated and probabilities inferred at a business line level the next step is to aggregate at the group level to arrive at an aggregated risk probabilities and rating for the group.

A different BN structure is adopted here. **Figure 10** shows the BN structure to be adopted for the risk aggregation BN. The risk values for the two business lines – shown here as CB (Commercial Banking) for business line 1 and TS (Trading & Sales) for business line 2 - will be the parent nodes with "Group" being the child node. A conditional probability table (CPT) was manually created to facilitate the aggregation of the ratings.

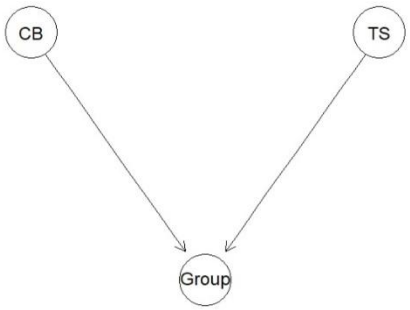


**Figure 10 Aggregation of risks of the business line**

The probabilities for the risk impact values for CB and TS are inferred from their respective BNs (BL1 and BL2 respectively), as learnt from the incident database and subsequently updated with risk assessment or scenario analysis information. The node Group shows the aggregated probabilities of the risk level, namely the probability of a high impact risk or medium impact or low impact risk impact from the group as a whole. This based on the probabilities for High/Medium/Low impact risks for each of the underlying business lines.

For this network, the conditional probabilities have been manually specified. These are subjective values and can be updated to reflect expert opinion and also factor in risk appetites defined or varied risk impact levels, if any and as applicable.

## IV. Evaluation of the BN for operational risk management

The BN that was constructed and learned could be applied to the various objectives that the research intended to achieve.

a) **At the risk or risk event level – aggregation of data from different risk techniques**: The BN allows for the aggregation of data from the different risk techniques at the risk event level as shown in Table 3. Risk event information can be learned from an incident database providing an aggregated view of risk events. This information can be enriched, incorporating information from RCA and scenario analysis, providing expert opinion akin to forward looking views of the risk environment. As seen the probability of a risk event such as Asset Loss #1 can be determined by inferring from the BN, created from the incident database initially, and subsequently with probabilities updated after incorporating expert opinion from RCA or scenario analysis. This allows for a better and more dynamic view of the probability of a risk event through the integration of different data types. Table 3 provides an overview of the sources of data, the various ORM techniques, and the types of data that can be aggregated through the BN. The probability of risk events can be directly sourced from the incident/loss data. The BN allows for the data from incident/loss data, data from RCSA and scenario analysis to be aggregated, showing that data from different sources of risk information can be aggregated. Loss data is backward looking and is relatively objective, given the quantitative nature, while RCA/RCSA and scenario analysis are quite forward looking with much subjectivity given the nature of expert assessment. Through the aggregation of the data from incident/loss and RCA/RCSA/Scenario analysis – it is seen that forward looking and backward looking data can be aggregated and so can subjective and objective data sets.

| | Incident /Loss data | RCA / RCSA | Scenario Analysis |
|---|---|---|---|
| Risk events | √ | | |
| Different sources of risk information | √ | √ | √ |
| Forward Looking and backward looking | √ | √ | √ |
| Subjective and objective data | √ | √ | √ |

**Table 3 Types of data and the sources of data that can be aggregated through the BN to get the probabilities of a risk event**

b) **At the Basel OR event type – aggregation of risk or risk events**: Probabilities of the risk events can be aggregated to arrive at the probability of the related OR event type. The probability of the risk type Business Disruption and System Failure (BDSF) can be inferred by aggregating the related risk events, namely SD#1 and SD#2. Similarly, the probability of risk type Execution, Delivery and Process Management (EDPM) can be inferred from the BN given the aggregation of the related risk events (PE#1, PE#2). This allows for a richer understanding of the likelihood of the different risk types based on the probability of underlying risk events, facilitating a more focused approach to risk mitigation by the risk owners namely the process owners or line managers.

c) **At the business line level – aggregation of OR event type**: The BN incorporates information at the risk event and risk type level, therefore, integrating all the low-level risk information required for providing the business line risk view. As shown the risk impact probabilities for the business line can be inferred from the BN. The likelihood of a high impact OR for the BL1 or BL2 can be inferred from the BN, thus evidencing the aggregation capability. This allows the BL managers to determine the areas requiring risk prioritisation and monitor the risk accordingly. The aggregation allows for efficient capture in risk reports avoiding the 'disjoint' problem by allowing for a drill down through to the underlying factors namely the controls and causes that have the highest probability of failures and occurrence. Figure 11 shows the BN for the business line 2.

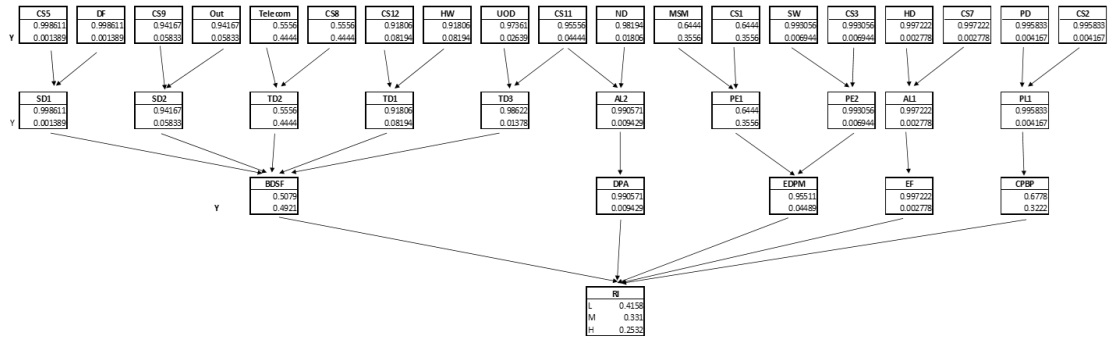


**Figure 11 Bayesian network with marginal probabilities for business line 2**

d) **At the group level – aggregation of business line operational risk:** As shown the aggregation BN allows for the risk of the business lines to be aggregated to arrive at the group risk. This enables better risk reporting at the group level to the managers interested in understanding group level exposure. They can further drill down – moving to the business line BNs to determine the top probable risk areas. This allows managers, at corporate or group level, understand the overall OR exposure and the exposure of each underlying business lines.

The BN could be applied to the various ORM techniques:

**Loss analysis**: It was possible to query the BN to determine the causes of risk events that had the highest probabilities of occurrence or, the controls that had the highest probability of failures or, risk events with the highest probability of occurrence. Figure 12 shows how the probabilities of the risk events for a business line can be presented, allowing for a focus on the risk events that have the highest probabilities of occurrence.

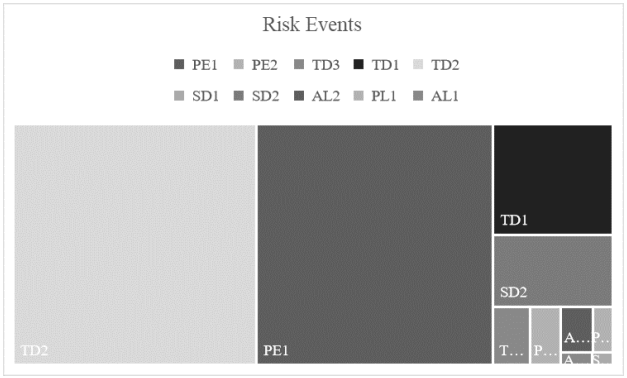


**Figure 12 Chart showing the probabilities of the various risk events for a business line**

Also, we could condition the BN for only high impact risks and then query to see which controls, causes and risk events have the highest probabilities. This provides a more holistic view across all operational risks and risk types for a given business. Figure 13 shows the marginal probabilities of the control (=Fail) and cause nodes (=Y) when the BN is conditioned for the risk impact being High.

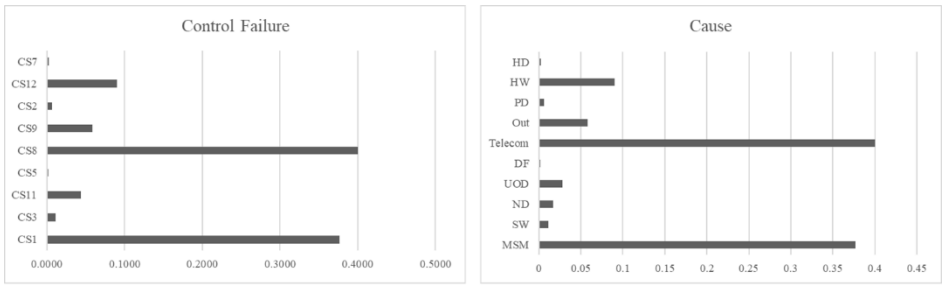


**Figure 13 Marginal probabilities for the control (=Fail) and cause nodes (=Y) conditioned on the risk impact (=High)**

**Key Risk Indicators (KRI)**: We could evaluate the BN to design KRIs that related to the specific risks in a clear and consistent way. If we wanted to focus only on events related to the Business Disruption and System Failures (BDSF) risk type, we could condition the BN to this risk type and then query the BN for the related causes and controls. It allows us to define KRIs for specific risk types or risk events that we see a high probability of occurrence. By tracing back, we can define KRIs to monitor the related causes and controls.

We could condition the BN for High risk impact (Figure 13) and then query to determine the risk types and risk events that had the highest probabilities. KRIs could be defined to monitor the controls and causes related to these risk events. This allows for the targeted monitoring of high impact risk events. KRIs could be designed to monitor the failure of these controls that could result in a higher number of risk events.

**Scenario Analysis**: The current state of a given characteristic could be inferred from the BN providing the basis for further analysis to determine whether this is commensurate with the potential future state or if it requires to be updated. The focus of this research is not the computation of loss distributions or OR capital calculation to facilitate the understanding of a given scenario.

We could, separately, specify scenarios related to 2 risks that are considered among the top risks for 2019 (Risk.net, 2019) and infer the probabilities of there being a high impact risk occurring.

1. Cyber-attack – a scenario where the cause is "HD" (HD=Y)
2. Outsourcing and third-party risk – a scenario where the cause is "Out" (Out = Y)

For analysis of the scenarios, hard evidence could be set to simulate the cause and failure of the control suites associated with these causes and infer the probabilities of a high impact risk occurring. Managers can simulate the outcome if controls have failed or a causal factor has materialised.

| Business line | Scenario | Node | State | Probability of High impact event conditional the node states |
|---|---|---|---|---|
| 1 | 1 | HD<br>CS7 | Y<br>Fail | 0.3876 |
| | 2 | Out<br>CS9 | Y<br>Fail | 0.2497 |
| 2 | 1 | HD<br>CS7 | Y<br>Fail | 0.38 |
| | 2 | Out<br>CS9 | Y<br>Fail | 0.2400 |

**Risk Reporting:** When evaluating a model for risk reporting it is essential to evaluate if it allows for risk information to be communicated in a manner that would cater to different senior managers and risk managers allowing senior managers to be aware of the risk types, probability of the occurrence and risk impact and accordingly prioritise their risk remediation efforts

We have seen that BN structure and the parameter learned can give us much insight into the key controls, causes, risk event types, and the impact. These allow us to measure the operational risk exposure and also present them either as KRIs or an overview of the probabilities of the risk events and risk types occurring catering to different groups of

senior managers. Business managers can view the risk types and events that have the highest probability of occurring. They are then able to work with the risk manager to understand better the specific control and causal factors related to their area, those on which they need to prioritise their remediation efforts. Process owners are also able to better manage their process by gaining visibility into the controls that have a higher probability of causing a risk event and then prioritise their remediation efforts there. KRI focused on these areas will help them efficiently manage this risk while understanding the alignment with the risk event and the risk type.

Banks typically use risk grids for showing the risk levels aligning the risk levels with the risk appetite that has been set. To evaluate the use of the BN for risk reporting within the bank, a risk grid has been constructed (Figure 14). The risk grid is a matrix of the likelihood (of risk event occurrence) and the impact of the risk event. For this research, a 3x3 matrix has been constructed with the likelihood divided into three regions, and the impact also divided into three ranges of $ impact. The impact has been divided as Low risk (less than $ 10,000), Medium risk (between $10,000 and $100,000) and High risk (greater than $100,000). A risk can be plotted on the risk grid, taking into account the likelihood of the event and the risk impact. Each cell in the grid can be numbered for quick reference and refers to a specific loss amount range. As an example, a risk event that had a likelihood of occurrence at 0.5 with a risk impact of $50,000 would be plotted in the cell number 5 indicating it is a medium risk and can be reported thus. We took that values for the risk types from the BN and plotted it on the risk grid. For business line 1 (Figure 15 a) the risk type External Fraud had a 0.38 probability of a high impact risk and is plotted accordingly on the chart. For business line 2 (Figure 15 b) the risk type EDPM, there is a 0.25 probability of a high impact exposure. The grids also shows the aggregated business line rating. For business line 1, there is a 0.25 probability of a high impacting risk exposure.

Figure 16 shows how a group level manager can receive an aggregated view of the risks on the grid. In this case, it shows the business line level rating – for business line 1 and 2 - and the aggregated group level rating.

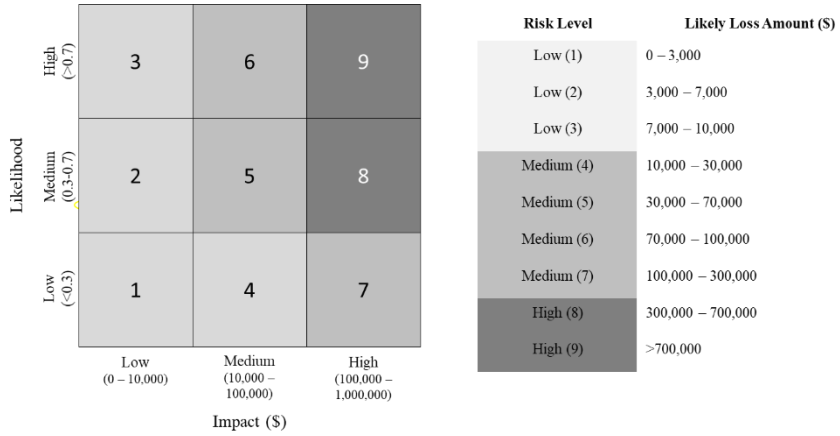


**Figure 14 Risk grid to plot the risks for management reporting**

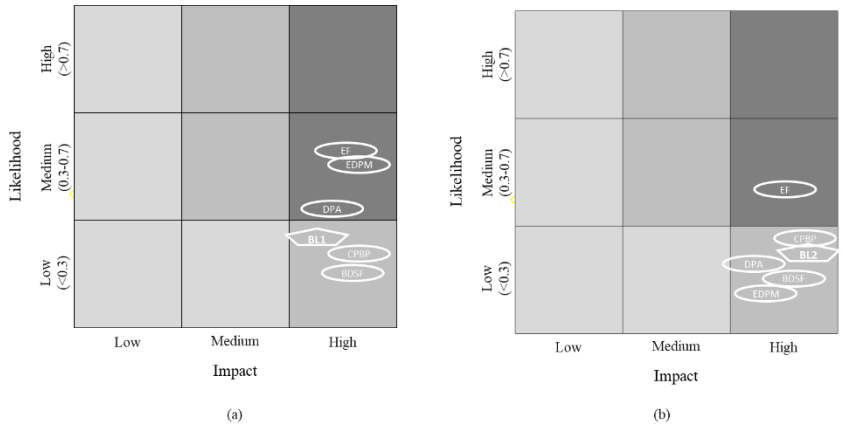


**Figure 15Reporting on the risk types and the business line aggregated risk (a) reporting on risks for business line 1 (b) reporting of risks for business line 2**

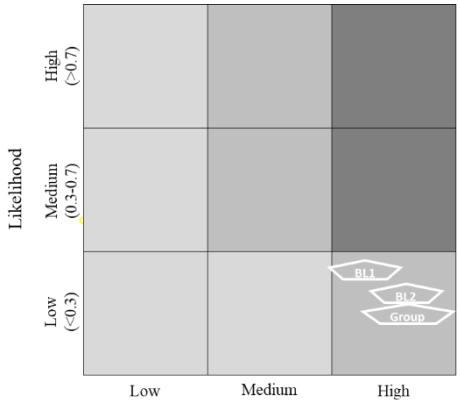


**Figure 16Risk reporting at the group level**

## V. Conclusion

The framework presented in the research provides, as a first, a practical approach to using machine learning, specifically Bayesian networks, for the aggregation of operational risk data for risk reporting. While BNs have been used to measure operational risk for capital calculation (VaR, loss distributions) and in analysing risks in specific process areas, this research has addressed a gap showing that machine learning BN models can significantly enable the process of aggregation of operational risk data for operational risk reporting. The BN model allows for all the building blocks of an operational risk framework – loss data, RCSA, scenario analysis, KRIs - to be integrated through risk aggregation enabling better measurement and reporting. This research has also addressed how aspects of a pure data-driven model can be preserved while incorporating expert knowledge. It shows how expert knowledge can be incorporated into the BN and the operational risk measurement model. In a dynamic environment, such as the banking industry, there is a constant flow of information in the form of observed losses, new or changed policies, and controls. It is useful to incorporate this flow of information into the models. The framework in this research can be adopted and adapted at any bank to deliver dynamic and timely risk intelligence for the effective management of operational risk on a day-to-day basis. The framework can also be extended to aggregate more operational risk sub-types, including where data is limited and additional business lines.

## Appendix A List of abbreviations

| | |
|---|---|
| BN | Bayesian Networks |
| ORM | Operational Risk Management |
| OR | Operational Risk |
| RCA | Risk and Control Assessment |
| RCSA | Risk and Control Self Assessment |
| CB | Commercial Banking |
| TS | Trading & Sales |
| BDSF | Business disruption and systems failures |
| EF | External Fraud |
| DPA | Damage to Physical Assets |
| CPBP | Client, Products, Business Practices |
| EDPM | Execution, delivery, & process management |
| KRI | Key Risk Indicator |
| BL | Business Line |
| VaR | Value at Risk |
| PGM | Probabilistic Graphical Methods |
| ERM | Enterprise Risk Management |
| | |

## References


[1] Basel Committee on Banking Supervision, "Principles for the Sound Management of Operational Risk," *Bank Int. Settlements*, no. June 2011, pp. 1–27, 2011.

[2] M. Leo, S. Sharma, and K. Maddulety, "Machine Learning in Banking Risk Management: A Literature Review," *Risks*, vol. 7, no. 1, p. 29, Mar. 2019, doi: 10.3390/risks7010029.

[3] M. Pergler and A. Freeman, "Probabilistic modeling as an exploratory decision-making tool," 2008.

[4] Y. Yasuda, "Application of Bayesian Inference to Operational Risk Management," no. January, 2003.

[5] A. Samad-Khan, "Modern operational risk management," *Emphasis*, vol. 2, pp. 26–29, 2008.

[6] N. Allan, N. Cantle, P. Godfrey, and Y. Yin, "A review of the use of complex systems applied to risk appetite and emerging risks in ERM practice," 2011.

[7] N. Cantle, Y. Yin, and N. Allan, "Modeling the Interconnectivity of Risks in ERM," in *Risk Management Symposium*, 2008, pp. 1–18.

[8] D. Koller and N. Friedman, *Probabilistic graphical models : principles and techniques*. 2009.

[9] S. Figini, L. Gao, and P. Giudici, "Bayesian operational risk models," *J. Oper. Risk*, vol. 10, no. 2, pp. 45–60, 2015, doi: 10.21314/jop.2015.155.

[10] K. P. Svensson, "A Bayesian Approach to Modeling Operational Risk When Data is Scarce," 2015. [Online]. Available: http://lup.lub.lu.se/student-papers/record/5205925/file/5205926.pdf. [Accessed: 25-Jun-2019].

[11] A. Sanford and I. Moosa, "Operational risk modelling and organizational learning in structured finance operations: A Bayesian network approach," *J. Oper. Res. Soc.*, vol. 66, no. 1, pp. 86–115, 2015, doi: 10.1057/jors.2013.49.

[12] S. Ibrahimovic and U. Franke, "A probabilistic approach to IT risk management in the Basel regulatory framework: A case study," *J. Financ. Regul. Compliance*, vol. 25, no. 2, pp. 176–195, 2017, doi: 10.1108/JFRC-06-2016-0050.

[13] M. Leo, S. Sharma, and K. Maddulety, "Managing Operational Risk using Bayesian Networks: A practical approach for the risk manager," vol. 4, no. 6, pp. 54–69, 2020.

[14] The Institute of Operational Risk, "Risk Control Self Assessment," 2010.

[15] P. X. Girling, *Operational risk management: a complete guide to a successful operational risk framework*. John Wiley & Sons, 2013.

[16] Y. K. Yoon, "Modelling Operational Risk in Financial Institutions using Bayesian Networks," Cass Business School, London, 2003.

[17] T. Blunden and J. Thirlwell, *Mastering Operational Risk: A practical guide to understanding operational risk and how to manage it*. Pearson UK, 2012.

[18] A. Samad-Khan, "Modern operational risk management," *Emphasis*, vol. 2, pp. 26–29, 2008.

[19] Institute of Operational Risk, "Operational Risk Sound Practice Guidance Risk Control Self Assessment," 2010.

[20] F. Harmantzis, "Operational Risk Management in Financial Services and the New Basel Accord," *SSRN Electron. J.*, 2011, doi: 10.2139/ssrn.579321.

[21] X. Zhou, R. Giacometti, F. J. Fabozzi, and A. H. Tucker, "Bayesian estimation of truncated data with applications to operational risk measurement," *Quant. Financ.*, vol. 14, no. 5, pp. 863–888, 2014, doi: 10.1080/14697688.2012.752103.

[22] S. Mathur, "Aggregation of Risk : Issues and Perspectives," *J. Compliance Risk Oppor.*, no. May, pp. 1–11, 2010.

[23] J. Spivack, "The Challenges of Risk Management in Diversified Financial Companies," *CFA Dig.*, vol. 31, no. 4, pp. 82–84, 2001, doi: 10.2469/dig.v31.n4.984.
[24] L. K. Meulbroek, "Integrated Risk Management for the Firm: A Senior Manager's Guide," 2002.
[25] H. Inanoglu and M. Jacobs, *Models for Risk Aggregation and Sensitivity Analysis: An Application to Bank Economic Capital*, vol. 2, no. 1. 2009.
[26] L. Narvaez and K. Warner, "How to Aggregate Risks Across Your Organization," *Cfo.Com*, 2013. [Online]. Available: http://ww2.cfo.com/risk-management/2013/07/how-to-aggregate-risks-across-your-organization/. [Accessed: 11-Sep-2018].
[27] J. V. Rosenberg and T. Schuermann, "A General Approach to Integrated Risk Management," 2004.
[28] K. RMA, "Operational Risk Management Excellence – Get to Strong Survey Executive Report," 2014.
[29] C. Stewart and T. O'Connor, "Embedding Risk Data Aggregation and Reporting Principles Across the Organization," 2013.
[30] M. J. Epstein and A. R. Buhovac, *The reporting of organizational risks for internal and external decision making*. CMA Canada, 2006.
[31] B. Ellis, I. Kristensen, A. Krivkovich, and H. P. Singh, "Driving value from postcrisis operational risk management," 2012.
[32] M. Jordan, "Conditional Independence and Factorization," in *An introduction to probabilistic graphical models*, 2003.
[33] M. Neil, D. Marquez, and N. Fenton, "Using Bayesian networks to model the operational risk to information technology infrastructure in financial institutions," *J. Financ. Transform.*, vol. 22, pp. 131–138, 2008.
[34] C. Alexander, "Bayesian Methods for Measuring Operational Risk," *SSRN Electron. J.*, vol. 44, no. 0, pp. 1–15, 2000, doi: 10.2139/ssrn.248148.
[35] R. G. Cowell, R. J. Verrall, and Y. K. Yoon, "Modeling Operational Risk With Bayesian Networks," *J. Risk Insur.*, vol. 74, no. 4, pp. 795–827, 2007.
[36] G. W. Peters, P. V. Shevchenko, and M. V. Wüthrich, "Dynamic operational risk: modeling dependence and combining different sources of information," vol. 4, no. 2, pp. 69–105, 2009.
[37] P. V. Shevchenko and M. V. Wüthrich, "The Structural Modelling of Operational Risk via Bayesian inference: Combining Loss Data with Expert Opinions," *J. Oper. Risk*, vol. 1, no. 3, pp. 3–26, 2006.
[38] V. Aquaro, M. Bardoscia, R. Bellotti, A. Consiglio, F. De Carlo, and G. Ferri, "A Bayesian Networks approach to Operational Risk," *Phys. A Stat. Mech. its Appl.*, vol. 389, no. 8, pp. 1721–1728, 2010, doi: 10.1016/j.physa.2009.12.043.
[39] M. Neil, N. Fenton, and M. Tailor, "Using Bayesian networks to model expected and unexpected operational losses," *Risk Anal.*, vol. 25, no. 4, pp. 963–972, 2005, doi: 10.1111/j.1539-6924.2005.00641.x.
[40] A. D. Sanford and I. A. Moosa, "A Bayesian network structure for operational risk modelling in structured finance operations," *J. Oper. Res. Soc.*, vol. 63, no. 4, pp. 431–444, 2012, doi: 10.1057/jors.2011.7.
[41] J. F. Martínez-Sánchez, M. T. V. Martínez-Palacios, and F. Venegas-Martínez, "An analysis on operational risk in international banking: A Bayesian approach (2007–2011)," *Estud. Gerenciales*, vol. 32, no. 140, pp. 208–220, 2016, doi: 10.1016/j.estger.2016.06.004.
[42] C. Alexander, "Bayesian Methods for Measuring Operational Risk," *Ssrn*, pp. 1–15, 2000, doi: 10.2139/ssrn.248148.
[43] N. Fenton and M. Neil, "The use of Bayes and causal modelling in decision making, uncertainty and risk," *Risk Inf. Manag. Res. Gr.*, no. June, pp. 1–19, 2011.
[44] M. Neil, D. Häger, and L. Andersen, "Modeling operational risk in financial institutions using hybrid dynamic Bayesian networks," *J. Oper. Risk*, vol. 4, no. 1, pp. 3–33, 2016, doi: 10.21314/jop.2009.057.
[45] Basel Committee on Banking Supervision, "The internal audit function in banks," 2012.
[46] Federal Deposit Insurance Corporation, "Internal Routine And Controls," 2015.
[47] N. E. Fenton and M. (Martin D. . Neil, *Risk assessment and decision analysis with bayesian networks*. CRC Press, 2013.
[48] Risk.net, "Top 10 operational risks for 2019 - Risk.net," *Risk.net*, 2019. [Online]. Available: https://www.risk.net/risk-management/6470126/top-10-op-risks-2019. [Accessed: 11-Nov-2019].