



UNIVERSITY OF TWENTE.

17-August-2017

Towards a continuous auditing philosophy

Ruiter, Bryan
PRICEWATERHOUSECOOPERS,
UNIVERSITY OF TWENTE

Document information

Title: Towards a continuous auditing philosophy
Type: Master thesis
Date: 17-08-2017
Version: Final

Candidate

Name: Bryan Ruiter
Student Number: s0141275
Study: Industrial Engineering and Management MSc.
Specialization: Financial Engineering and Management
E-mail: b.ruiter-2@student.utwente.nl
Bryan.ruiter@pwc.com
Telephone: +31645652264

University

University: University of Twente
Faculty: Faculty of behavioral, management and social sciences

Internal supervisors

Ir. drs. A.C.M. de Bakker University of Twente
Dr. B. Roorda University of Twente

External supervisor

Name: Wouter Weusthof MSc.
Organization: PwC
E-mail: Wouter.weusthof@nl.pwc.com
Telephone: +31653466792

Preface

This research project is conducted to complete the author's master study, Industrial Engineering and Management: Financial Engineering at the University of Twente.

I would like to thank the principal of this project PwC who offered me this opportunity and all the wonderful employees of the Risk Assurance PCPS department. Especially the employees at the Zwolle office who answered many of my questions. A special thanks goes out to my supervisor and coach Wouter Weusthof who was always there for me.

I would also like to thank my supervisor Toon de Bakker, you were very pleasant to work with and your feedback was very constructive. I could not have wished for a better supervisor. And thank you, Berend Roorda, for all your work as a teacher and for being my second supervisor.

Bryan Ruiter

Zwolle, 17-8-2017

Abstract

This research project is conducted at PwC, the Netherlands. PwC sees a trend of ongoing automation in the accountancy practice propelled by the growing adoption of IT solutions (PwC, 2015). Despite these advancements audits are conducted twice a year. Which results in current information not being assured and may lead to unidentified risk incorporation. The time gap between risk incorporation and identification is a period where an organization takes full risk exposure. While in hindsight an organization might have been willing to mitigate this risk. At this moment PwC does not have a cost-efficient practice to do continuous risk assessments. However if such practice exists PwC's value proposition will increase. The aim of this research is to design a framework that PwC could use to implement a continuous auditing (CA) philosophy in their practice. This research project does not only propose such framework but it also tests the feasibility of a continuous auditing solution by simulating the functioning of designed CA algorithms on a Purchase-to-Pay process with real data. This simulation showed that a CA solution is likely to be feasible. However it also showed that the success of a CA solution depends on the IT maturity of the organization including the fit between IT and formalized business processes.

Keywords: Audit, Control, Risk Management, Assurance, Continuous Auditing, Continuous Control Monitoring, Continuous Data Assurance, Continuous Risk Monitoring and Assessment, Simulation, Case Study, Purchase-to-Pay, P2P.

Contents

Document information	i
Preface	ii
Abstract	iii
List of figures	vi
List of Tables	vii
List of abbreviations	viii
1. Introduction	2
1.1. Background	2
1.2. Problem	2
1.3. Project scope	4
1.4. Problem statement	5
1.5. Research questions	5
1.6. Plan of approach	6
2. Measuring risk exposure	8
2.1. Measurable and unmeasurable risk	8
2.1.1. Risk definitions	8
2.1.2. Risk and knowledge	9
2.1.3. Knightian Risk	10
2.1.4. Knightian Uncertainty	11
2.1.5. Ignorance	12
2.2. Risk identification by external auditors	12
2.3. Contribution to the research questions	14
3. Continuous auditing	16
3.1. Definitions and philosophy	16
3.1.1. Continuous Data Assurance	18
3.1.2. Continuous Control Monitoring	20
3.1.3. Continuous Risk Monitoring and Assessment	22
3.2. Contribution to the research questions	24
4. Case study	26
4.1. Case description	26
4.2. Simulating a continuous audit	27
4.2.1. CDA algorithm	29
4.2.2. CCM algorithm	33
4.2.3. CRMA	34

4.3.	Results and lessons learned	35
4.3.1.	CDA algorithm results	35
4.3.2.	CCM algorithm results	39
4.3.3.	CRMA dashboard view	41
4.4.	Contribution to the research questions	42
5.	Continuous audit in a general context	44
5.1.	IT infrastructure	44
5.2.	Continuous audit implementation	45
5.2.1.	Continuous data obtainment: EAM vs. MCL	45
5.2.2.	Data analysis level: individual transactions vs. high level aggregation	46
5.2.3.	Determine the window size: growing windows vs. sliding window	46
5.3.	Contribution to the research questions	47
6.	Discussion	48
6.1.	Conclusions	48
6.2.	Limitations & Future Research	49
7.	References	52
	Appendix A	54

List of figures

Figure 1: Typical risk identification problem cluster.....	3
Figure 2: Risk dimensions (Aven & Renn, 2009)	9
Figure 3: Risk knowledge matrix	10
Figure 4: Fat tails vs. Normal distribution (Financial Times, n.d.).....	11
Figure 5: Relevant Continuous Auditing components (Vasarhelyi, Alles, & Williams, 2010)	17
Figure 6: Continuous Data Assurance.....	19
Figure 7: Continuous Control Monitoring	21
Figure 8: Continuous Risk Management and Assessment Scheme.....	23
Figure 9: Purchase-to-Pay process of client X (2016)	26
Figure 10: Dataset tables	28
Figure 11: Normal distribution with sigma bandwidth (Wikipedia, 2017).....	33
Figure 12: Real-time simulation anomaly detection example	38
Figure 13: CRMA Dashboard	41

List of Tables

Table 1: Research project’s plan of approach.....	6
Table 2: Continuous Data Assurance Results FY 2015 & FY 2016	36
Table 3: Vendor data	37
Table 4: Real-time simulation anomaly detection example (continued)	38
Table 5: Continuous Control Monitoring Results FY 2015 & FY 2016	39

List of abbreviations

AICPA	American Institute of CPAs
AP	Analytical Procedures
CA	Continuous Auditing
CCM	Continuous Control Monitoring
CDA	Continuous Data Assurance
COMO	Compliance Monitoring
CPA	Certified Public Accountant
CRMA	Continuous Risk Management and Assessment
EAM	Embedded Audit Module
ERPs	Enterprise Resource Planning systems
FK	Foreign Key
FY	Fiscal Year
IFI	Immediate Financial Impact
ISA	International Standards of Accounting
IT	Information Technology
KRIs	Key Risk Indicators
MCL	Monitoring Control Layer
P2P	Purchase-to-Pay
PK	Primary Key
PwC	PricewaterhouseCoopers
ROMM	Risk of Material Misstatement
RQ	Research Question
SOD	Segregation of Duties

1. Introduction

Organizations might run enormous risk exposures without even knowing it. And eventual risk manifestation may adversely affect the performance of an entity and could have disastrous effects like bankruptcy. Famous examples are the rogue trading of Nick Leeson at Barings Bank in 1995 and the interest derivatives of Vestia in 2012. These (operational) risks, could manifest because employees could operate in an insufficient controlled environment. This research project is about creating a controlled environment, in which operational risks are monitored and possibly mitigated, by implementing a continuous auditing philosophy in a Purchase-to-Pay context.

In this chapter we introduce the principal of this research project, the problem of interest and our research approach.

1.1. Background

This research project will be conducted at PwC the Netherlands at the Private Company and Public Sector team of the Business Unit Risk Assurance. PwC is a financial service company with lines of service in assurance, advisory and tax. The Business Unit Risk Assurance is part of the assurance line of service and helps clients to protect - and create value from - their processes, systems and people by identifying risks and by offering solutions to mitigate these risks when desired.

PwC acknowledges that there is a trend of ongoing automation of the traditional accountancy services propelled by technology advancements and the ongoing adoption of IT solutions by organizations (PwC, 2015). To be competitive in the future and to be able to meet the increasing quality standards PwC needs to expand their expertise about IT solutions that contributes to assurance of organizations.

1.2. Problem

The objective of financial reporting is to provide insightful information to relevant stakeholders to support their resource allocation decisions (Financial Accounting Standards Board, 2006). For information to be insightful it should be timely, free from material misstatements and a fair representation of the financial performance of the entity. Despite the current technology advancements that allow for almost near real-time monitoring of data, audits are traditionally conducted at periodic intervals. Audits occur twice a year typically, an interim and a final audit. These audits cover the historic data captured in that timeframe. Resulting in that current information is not assured. Consequently, management reliance on real-time information may result in possibly adverse resource allocation decisions (Chan & Vasarhelyi, 2011).

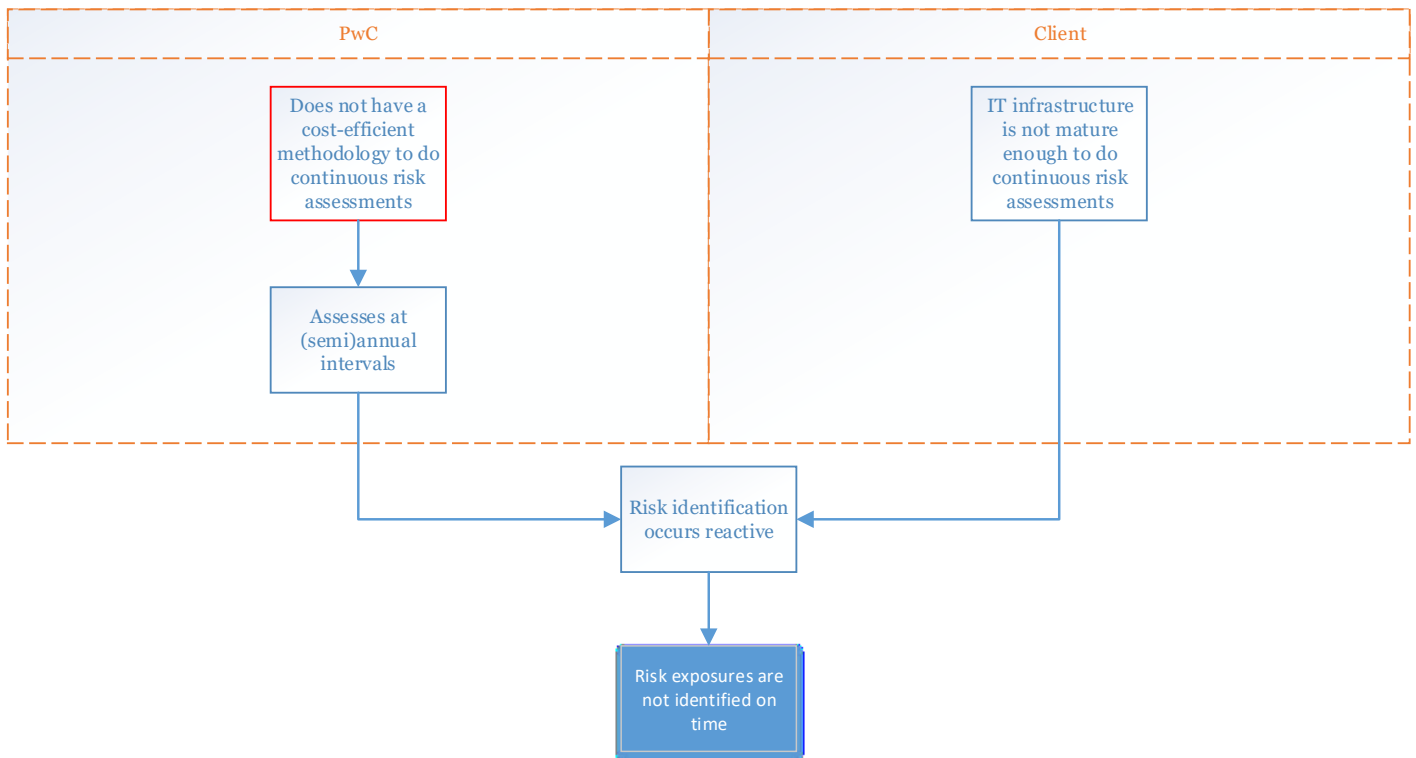


Figure 1: Typical risk identification problem cluster

Figure 1 describes a problem flow of a risk identification process at an hypothetical client of PwC. The studied problem is that risk exposures are not identified on time. Technology advances allow for a time norm that tends to shift to a more continuous approach. Every moment of elapsed time between risk identification and incorporation is a period where an organization takes full exposure. While in hindsight they might have been willing to mitigate this risk. Ideally organizations are able to identify risks as soon as they are embedded in the process. Therefore they can make a wise decision about whether they want to mitigate the risk or not. In order to make a fully informed decision current information needs to be assured.

Continuous auditing (CA) is a philosophy that helps organizations get closer to continuous assurance.

A continuous audit is a methodology that enables independent auditors to provide written assurance on a subject matter, for which an entity's management is responsible, using a series of auditor's reports issued virtually simultaneously with, or short period of time after, the occurrence of events underlying the subject matter (CICA / AICPA, 1999).

CA requires a certain level of maturity from the IT infrastructure of an organization. When looking at Figure 1 we see that at a typical client the IT infrastructure is often not mature enough to do continuous risk assessments. This could have various underlying reasons and they differ greatly among the clients.

Moreover, investigation of these problems is not within the scope of this research project. Nevertheless, many organizations are increasingly implementing sophisticated IT solutions in their day in, day out business processes. Making CA a realistic organizational philosophy. The accompanying benefit of these solutions is the amount of data that become available about the daily business processes. Extensively exploring of this data could be interesting for both internal and external auditors.

CA creates a field of new opportunities in the area of risk identification and mitigation. In the traditional audit a backward looking perspective on dealing with risks is applied. From this perspective risks are identified with the help of a sample of historic data from that relevant audit period. The resulting risk advisory remains static until the next audit moment. But when CA is applied instead, the external auditor could theoretically provide their clients with dynamic risk assessments by means of identifying areas of potential risk exposures in near real-time and on the whole population of available data (PwC, 2015). However, the identification of these areas is not a trivial task and a practice to do this in an efficient and effective manner is currently missing at PwC.

The core problem in this research is formulated as follows:

External audit does not have a cost-efficient practice to do continuous risk assessments

The purpose of this research project is to design a framework that helps PwC expand their Risk Assurance practice. This framework aims to contribute to a better understanding of how risks could be assessed in an effective and efficient manner with the help of a continuous auditing environment from the perspective of the external auditor. The feasibility of a CA environment designed to identify risks will be tested as a case study on the Purchase-to-Pay (P2P) process of a well-known organization active in the utility industry (client X).

1.3. Project scope

The core problem involves many different aspects and thus is too broad for the timeframe of this research project. We will not evaluate the cost-efficiency of the proposed practice, but we will simply assume that a typical CA environment, and thus a high degree of automation, allows for a reasonable cost-efficient practice. Operational risks are the type of risks that are of interest to this research project. Operational risks are defined as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. We use the operational risk categorization prepared by the Basel II committee (BIS, 2011).

1. Internal fraud – misappropriation of assets, tax evasion, intentional mismarking of positions, bribery.
2. External fraud – theft of information, hacking damage, third-party theft and forgery.

3. Employment practices and workplace safety – discrimination, workers compensation, employee health and safety.
4. Clients, products and business practice – market manipulation, antitrust, improper trade, product defects, fiduciary breaches and account churning.
5. Damage to physical assets – natural disasters, terrorism and vandalism.
6. Business disruption and system failures – utility disruptions, software failures and hardware failures.
7. Execution, delivery and process management – data entry errors, accounting errors, failed mandatory reporting, negligent loss of client assets.

A CA environment could be especially useful for controlling risk categories 1 and 7.

1.4. Problem statement

To be able to continuously identify and assess operational risk exposures, knowledge needs to be obtained about techniques and/or methodologies that could assist with the quantification or qualification of risk exposure in a continuous audit environment. That leads to the following problem statement:

How can operational risk exposure be identified and assessed in a continuous audit environment from an external auditors perspective?

1.5. Research questions

We formulated the research questions in such way, that at the end of this project, we will be able to answer the problem statement and satisfy the objectives of this research project.

RQ 1. What is operational risk exposure and how is it measured in previous literature?

The goal of this research question is to get insight into what operational risk exposure exactly is and how it is typically measured in literature.

RQ 2. How does the external auditor assess operational risks in the current practice?

We answer this question by describing the current practice that PwC uses to assess risks. PwC practice is compliant with regulatory standards, and so we can assume that external auditors in general use similar approaches. This question aims to get insight into existing standards and make use of them as far as possible rather than constantly reinventing the wheel.

RQ 3. What is continuous auditing according to previous research?

The objective of this research question is to get an overview of what CA entails by providing definitions, philosophy and a description of individual components. This theoretical framework is an essential step towards an aligned research project.

RQ 4. Can we assess operational risks efficiently and effectively at the case study?

We will apply our obtained knowledge to a case study of Client X. We will investigate the feasibility of a continuous auditing approach in a purchase-to-pay context by simulating a real-time continuous audit by back-testing it on historic data.

RQ 5. What can we learn from the case study and what are its practical implications?

By answering this question we will review the test results and will discuss its practical implications.

RQ 6. How can we implement what we have learned into current audit practice?

Once we have answered RQ 1 – RQ 5 we are able to understand the problem context and obtained knowledge about the feasibility of a continuous auditing approach. We will apply this knowledge into a design of a framework that PwC could implement into their Risk Assurance practice.

1.6. Plan of approach

This remainder of this report will be structured as follows in Table 1.

Table 1: Research project's plan of approach

Chapter	Addresses research question	Methods
2 Measuring risk exposure	RQ 1, RQ 2	Literature review Desk research
3 Current state of continuous auditing	RQ 3	Literature review
4 Case study	RQ 4, RQ 5	Case study Interview Modelling Simulation
4 Design	RQ 6	Literature review Modelling
5 Discussion	Problem statement	Lessons learned

2. Measuring risk exposure

Socrates (469-399 BC) once said “I know that I am intelligent, because I know that I know nothing”. This quote from the Greek philosopher can be applied to the field of risk management. Stressing the importance of continuous knowledge obtainment to become more “intelligent” on the risks that an entity is running. And at the same time acknowledging that an entities risk management practice is constraint by the existence of immeasurable risks. In the remainder of this chapter we will discuss several risk definitions, distinguish measurable from immeasurable risk, discuss several analysis techniques and look at the current external auditors risk identification practice.

This chapter addresses the following research questions:

RQ 1. *What is operational risk exposure and how is it measured in previous literature?*

RQ 2. *How does the external auditor assess operational risks in the current practice?*

2.1. Measurable and unmeasurable risk

2.1.1. Risk definitions

Although humans have been dealing with risks forever, there is a wide variety of risk definitions used in literature (Ganegoda & Evans, 2014). Most definitions touch elements of probability, expected values, uncertainty and events (Aven & Renn, 2009). Some of the more common definitions are:

1. Risk is a measure of the probability and severity of adverse effects (Lowrance, 1976)
2. Risk is the probability of an adverse outcome (Graham & Weiner, 1995)
3. Risk is the combination of probability of an event and its consequences (ISO, 2002)
4. Risk equals the expected disutility (Campbell, 2005)
5. Risk refers to uncertainty (what is its likelihood?) about and severity of the events and consequences of an activity with respect to something that humans value (Aven & Renn, 2009)
6. Risk is the possibility of loss or injury (Merriam-Webster, n.d.).

For this project we prefer a definition that does not include probabilities. Because as Aven & Renn (2009) mentioned in their paper, uncertainties exist even without specifying their probabilities. Probability in this context is just a measure of uncertainty and the transformation from knowledge to probabilities is often not trivial. Consequently, we will not use a risk definition that contains expected values which is the multiplication of probabilities with consequences. Likewise, a measure of impact like loss, injury or disutility is useful for specifying the severity but it limits the scope when included in

a definition. The definition that we will use throughout this paper is definition 5. Risk according to this definition is two dimensional and can only be defined when both the likelihood and severity dimensions are considered (see Figure 2). The accompanying benefit of this definition is that it accommodates both desirable and undesirable outcomes, which is nice because a benefit to one stakeholder might not be beneficial for another stakeholder and stakeholder discussions are not especially relevant to this research project. Concluding, this definition combines two key dimensions of risk: uncertainties about and severity of events and consequences without limiting to specific measures of uncertainty and severity.

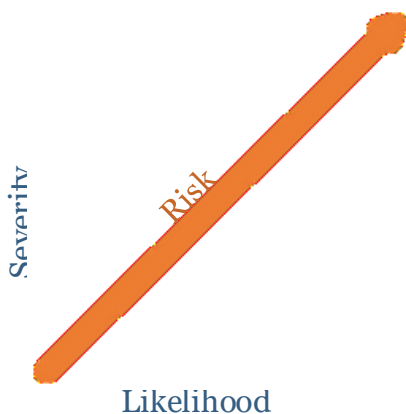


Figure 2: Risk dimensions (Aven & Renn, 2009)

2.1.2. Risk and knowledge

In the previous paragraph we mentioned that risks can only be defined when both the uncertainty (likelihood) and severity dimensions are considered. The amount of knowledge that is available about these dimensions determines the way how we identify and assess risks. The ultimate goal is to make risks measurable.

An early attempt to distinguish measurable from immeasurable risk is the work of Knight (1921). Knight claims that risk only applies to situations where knowledge about possible future states and their probability distributions are possessed. According to these properties, risk is quantifiable and measurable. Knight refers to situations where knowledge exists about possible future states but knowledge does not exist about underlying probability distribution as uncertainty (Knight, 1921). Since knowledge about probabilities is missing, Knight argues that uncertainty is immeasurable. By investing in knowledge uncertainties could be reduced and ultimately might become a risk (according to the definition of Knight (1921)).

When we refer to risk and knowledge in this paper we use the framework in Figure 3. As can be seen from this Figure, we use a much broader risk definition than Knight. But we do refer to Knight's (1921) classification with Knightian Risk (known/known), as the risk that we know exist and know how to

model, and Knightian Uncertainty (unknown/known) as a risk that we know exist but do not know how to model. Like Ganegoda & Evans (2014) we adopted the term Ignorance for risks that we are unaware of.

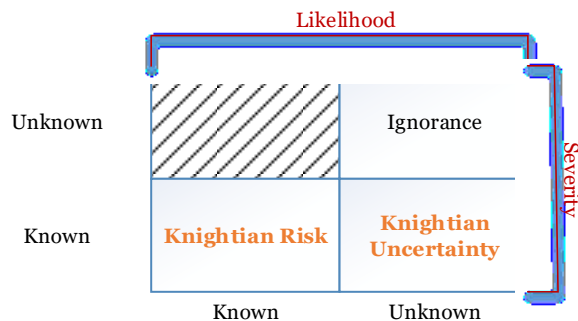


Figure 3: Risk knowledge matrix

Nowadays, sophisticated risk management strategies emphasize the importance of data-analysis to extract valuable knowledge. This is because knowledge contributes to the ability to make risks measurable. The quality of knowledge that is already available determines how risks assessments proceeds. We distinguished between several types of risk so that we can determine relevant analysis strategies for each type.

2.1.3. *Knightian Risk*

A Knightian Risk is an event for which we have knowledge about its consequences, and have great trust in the underlying models and theories that are used to quantify the probabilities of these consequences. Since probability distributions and consequences are known, the risk taker has no problem quantifying his risk exposure. As Knightian Risks are so easy to quantify an organization could easily find insurance against this kind of risks (Ganegoda & Evans, 2014).

Example: From the perspective of an insurer the total exposures on underwritten car insurance, life insurance and fire insurance is a Knightian Risk. For these types of risks, there is a lot of data available so it is relatively easy to model loss distributions with fairly high confidence. Once loss distributions are estimated, the entity could use methods like Monte-Carlo simulation combined with value-at-risk (VaR) and expected shortfall (ES) to determine capital buffers for the entity as a whole.

2.1.4. Knightian Uncertainty

When dealing with Knightian Uncertainty we have knowledge about possible future states, but we cannot estimate the underlying probability distributions accurately. One of the major reasons of not being able to do so is lack of relevant data. This is often the case at low-frequency/high-severity events like natural disasters. To overcome this lack of data an entity should pursue knowledge acquiring activities like for example drawing from external databases and expert opinions.

Another major reason of not being able to estimate probabilities might be that well-developed theories and models are not existent, which is regularly the case in the area of operational risk and market risk. One of the pitfalls of modelling Knightian Uncertainty is that an extreme observation can disproportionately impact the aggregate outcome. Consequently, the uncertainty of a risk measure based on these observations with a given fat-tailed risk is much greater than when modelling a light-tailed risk, e.g., a process that is modelled with a normal distribution (see Figure 4)¹.

Example: A risk that follows a fat-tailed distribution is the market risk of passively investing in the S&P 500 index from 1927 until 2006. When not being invested in the best 10 days of the market, the terminal wealth would decrease by 64%, whereas avoiding the worst 10 days would increase the terminal wealth by 202.5%. This particular example shows how a few observations impact the aggregate outcome drastically and that the uncertainty of a risk measure is for most part dependent on accurately estimating the probabilities of these drastic events.

The example above shows how hard it is to make a reliable model, but even if it is impossible to define probabilities, one could use methods such as stress testing and logic trees to quantify the possible impact of uncertain events (Ganegoda & Evans, 2014).

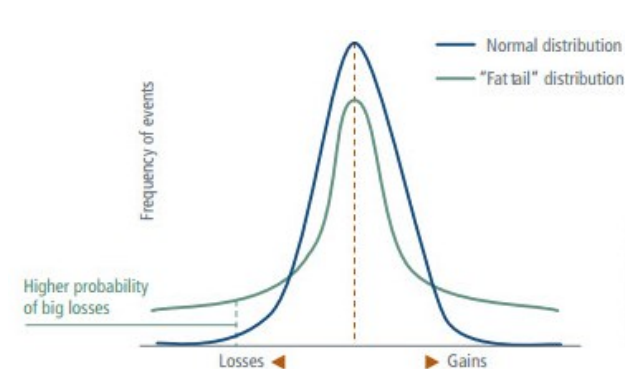


Figure 4: Fat tails vs. Normal distribution (Financial Times, n.d.)

¹ The fat tailed distribution that we used in this picture has a different sigma than the normal distribution that we compare it with. We used this distribution of a fat tailed risk to show the effect of more density in the tails.

2.1.5. Ignorance

In most cases, we are unaware of all the possible future states of the world. This ignorance makes it impossible to identify, let alone quantify, the risks that are in store for us. That makes these risks unavoidable (Ganegoda & Evans, 2014). Some of these unavoidable risks could have a high impact on an entity, the so called black swans in modern literature (Taleb, 2007). According to Taleb (2007), these black swans have three characteristics:

1. They are a complete surprise.
2. They have big consequences.
3. Once they have happened, people tend to believe the event is explainable in hindsight.

Although these black swans are unavoidable, that does not mean we cannot prepare for them. These events may have unique and unanticipated causes, but (Diebold, Doherty, & Herring, 2008) points out that the responses called for are often similar. That emphasizes the importance of sound crisis management.

2.2. Risk identification by external auditors

The external auditor is required to have a certain degree of understanding about the risks that an entity faces because it increases the likelihood that the auditor is able to identify risks of material misstatement (ROMM). This is important because most risks will eventually have financial consequences and thus effect the financial statements. According to the PwC Audit Guide (2017) the in scope risks are as followed:

A risk resulting from significant conditions, events, circumstances, actions, or inactions that could adversely affect an entity's ability to achieve its objectives and execute its strategies, or from the setting of inappropriate objectives and strategies (PwC, 2017).

Obtaining reasonable assurance about whether the financial statements are free from material misstatements, whether due to error or fraud, is one of the primary tasks of the auditor. In general misstatements, including omissions are considered to be material if, individually or in the aggregate, could lead to adverse economic decisions taken on the basis of financial statements. However, the auditor does not have the responsibility to assess all the risks that an entity faces, because not all risks give rise to ROMM (PwC, 2017). That said, the external auditor does have a natural advisory function. Clients and society expect from the external auditor that concerns are expressed, regarding for example IT security or continuity of the business, even when it does not lead to risk of material misstatement. Operational risks that arise from failed business processes, systems and/or people may lead to ROMM.

Examples of conditions or events are: (from International Standards of Accounting (ISA) 315 (IFAC, 2009))

- Lack of personnel with appropriate accounting and financial reporting skills.
- Developing or offering new products or services, or moving into new lines of business.
- Weaknesses in internal control.
- Significant transactions with related parties.
- Changes in key personnel including departure of key executives.
- Changes in the IT environment.
- Past misstatements, history of errors or a significant amount of adjustments at period end.

To identify potential areas of risk exposure the auditor may use so called analytical procedures. Analytical procedures (AP) are evaluations of financial information through analysis of plausible relationships among both financial and non-financial data. It includes the analysis whether data is consistent with other relevant information and whether data differs from expected values by a significant amount. When APs are used by the auditor, the auditor should first develop expectations about plausible relationships based on his knowledge about the business and industry, general economic conditions and prior audit experience. Quantitative or qualitative analysis could then be used to assess the client's financial information against expectations. To efficiently conduct AP they should be performed on data at an aggregated level or at a sufficiently disaggregated level. However, when data is used that is aggregated at a high level it may only give a broad indication whether material misstatement exist (PwC, 2017). The following list provides some sample questions that the auditor might consider:

- What are the trends in the entity's markets and how does the competitor compare with its peers?
- What are the key financial ratios that management focus on? Why are those ratios considered key? How do the entity's ratios compare with previous year and the industry?
- What are the trends in the entity's financial ratios measuring liquidity, solvency, leverage, and operating results? How is it compared with its peers.

To get insight into how PwC is identifying risks in the current situation we look at the PwC Audit Guide which states the practice that PwC adopts and functions as a baseline for all audits. PwC's practice is

based on compliance with the International Standards of Accounting. Of interest to this paper is the data-analysis techniques that PwC uses or the so called analytical procedures.

PwC Audit Guide (2017) mentions the use of five different APs:

1. Trend analysis – The analysis of changes in an account over time.
2. Ratio analysis – The comparison, a cross time or to a benchmark, of relationships between financial statement accounts or between an account and non-financial data.
3. Reasonableness testing – The comparison of accounts or changes within accounts with expectations from a developed model based on (non-)financial data.
4. Regression analysis – The use of statistical models to quantify expectations, with measurable risk and precision levels.
5. Scanning analytics – The identification of anomalies within account balances.

When we are answering RQ 2 in the next section we will be referencing to these APs.

2.3. Contribution to the research questions

RQ 1. *What is operational risk exposure and how is it measured in previous literature?*

By combining the project's scope, the definitions of Aven & Renn (2009) and BIS (2011) we obtain an operational risk definition that is used during this research project. Operational risk in this scope refers to uncertainty about, the likelihood, and severity of the events and consequences resulting from inadequate or failed internal processes and people. The ability to measure risk exposure depends on the amount of knowledge available about the risk's consequences and it's underlying probability distributions. Operational risks are often Knightian Uncertainties, which means knowledge about the possible risk events is within reach, but underlying probability distributions are unknown. Methods such as stress testing and logic trees are used to quantify possible impact.

RQ 2. *How does the external auditor assess operational risks in the current practice?*

The external auditor assesses operational risks that could lead to risks of material misstatement with the help of analytical procedures primarily. Trend analysis, ratio analysis, reasonableness testing, regression analysis and scanning analytics are the procedures used. Results of these procedures are benchmarked against expectations that are developed by the auditor based on knowledge about the business and industry, general economic conditions and prior audit experience.

3. Continuous auditing

Business models are rapidly changing with an increasing focus on digital and device. This development leads to a data explosion (PwC, 2015). These huge amounts of data might conceal risks that are not yet known. If implemented correctly, CA can help organizations bring these risks to the surface. And could give organizations real-time feedback in controlled environments to detect (potential) fraud, errors and abuse.

In this chapter we will explain what continuous auditing is, discuss several continuous auditing components and give a brief description of the history and recent developments of continuous auditing.

At the end of this chapter we will address the following research question:

RQ 3. *What is continuous auditing according to previous research?*

3.1. Definitions and philosophy

Continuous auditing is a philosophy that envisions that data flows of business process supporting systems is monitored and analyzed on a continuous basis in an automated fashion (Woodroof & DeWayne, 2001). A CA environment is most viable when continuous auditing principles significantly improve the reliability of information and when continuous information is vital to critical decision making (CICA / AICPA, 1999). In this environment systems scan the data for irregularities by checking whether it is in accordance with pre-defined set of rules, by the auditor, and algorithms in a continuous fashion. These irregularities are often called exceptions or anomalies and are treated in some accounting literature as synonyms (Woodroof & DeWayne, 2001). However, like Kogan, Alles, Vasarhelyi & Wu (2014) we will make a distinction between these two. We define exceptions as violations of pre-defined business process rules and anomalies as significant statistical deviations from expected process behavior. This definition of anomalies is slightly different from what is used by Kogan et al (2014). Their definition is as follows: “Anomalies are significant statistical deviations from the steady state of business process behavior”. We believe that expected behavior is more appropriate because a steady state implies a more limited scope, that of a system in which behavior is unchanging over time. A practical usage of the concepts exceptions and anomalies is given below:

Example: If a company requires from a procurement that it needs to be authorized by two different employees before submitting to a vendor, then an exception occurs when only a single employee has authorized the submitted order. If a company usually orders between 100 and 200 units of a certain good, then a correctly authorized and submitted order of 1000 units of that certain good is an anomaly.

In many research articles CA is divided between several functional components (Bumgarner & Vasarhelyi, 2015; Byrnes, Brennan, Vasarhelyi, Moon, & Ghosh, 2015; Vasarhelyi, Alles, & Williams,

2010). Although there exist different views on the exact nature of these components many specify at least a component that addresses the correctness of transactional data and a component that monitors the control compliance of business processes. The first is usually called Continuous Data Assurance (CDA) and the latter Continuous Control Monitoring (CCM). Of special interest to this research project, but mentioned less often in previous literature, is a component that addresses a continuous approach towards the management of risks. Continuous Risk Monitoring and Assessment (CRMA) is usually the name given to this component. Finally, but rarely, a functional component that is designed to monitor the business's regulatory compliance is mentioned in literature and can be found as Compliance Monitoring (COMO) (Bumgarner & Vasarhelyi, 2015). All of these components can operate individually but together they have synergetic effects.

Example: When CDA and CCM are implemented simultaneously and a certain business process triggers an alert because a certain control is not met, the probability of a true positive is higher because it is less likely that the incoming data is flawed.

We will, however, leave COMO out of the equation because it has less relevance to the goal of this research project. Figure 5 shows the CA components that we consider to be relevant for this research project. In the next subsections we will elaborate on these components.

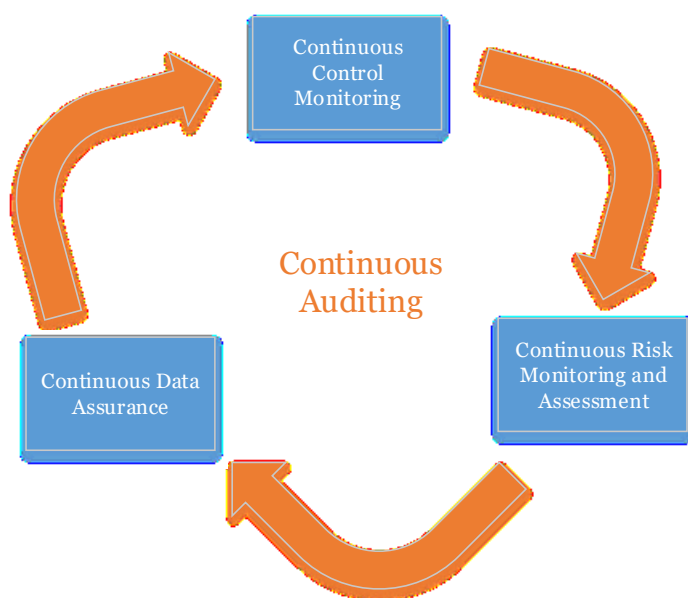


Figure 5: Relevant Continuous Auditing components (Vasarhelyi, Alles, & Williams, 2010)

3.1.1. Continuous Data Assurance

Because organizations adopted diverse IT systems fairly gradually over time, the data landscape of the modern corporate is complex. Data is stored in files for the legacy systems, in databases for the Enterprise Resource Planning systems (ERPs) and more recently in large (external) repositories that are called big data (Vasarhelyi, Kogan, & Tuttle, 2015). As data increase in size and complexity, there will be less room for the manual audit and it will be replaced with automated audit processes to keep sufficient levels of assurance (Vasarhelyi & Halper, 1991). CDA (see Figure 6) is such an automated technology, it extracts data from for example the above mentioned sources and puts it into a relational database, the data warehouse, and checks whether data from one source is conform the other. If not it generates an alert that should be reviewed by the auditor. An exception is not necessarily a violation and thus an action from the operator is not always a countermeasure. The operator could also decide to bypass the warning. These irregularities should be reviewed before it is suitable for automated testing. According to Vasarhelyi, Alles, & Williams (2010) there are two types of common errors that can be identified by the transaction verification component of the CDA module:

1. Data integrity violations include but are not limited to invalid purchase quantities, receiving quantities, and bank numbers.
2. Referential integrity violations are unmatched records among different business processes. For example, a purchasing order of raw materials cannot be matched with a production job.

Once the data is verified it can be analyzed at the transactional level. Analysis at the transactional level is quite different from the traditional accounting practices. The traditional accountings standards use sampling rather than examining the whole data set and use data at a more aggregated level. These traditional standards were more based upon cost and capability constraints than the ideal process for assurance. The assumption behind using data at the transactional, rather than aggregated, level is that auditors can compare the details of individual transactions against expected details in close to real time in an automated fashion. In order to do so, auditors need to establish benchmarks by using all kinds of data-analysis techniques. With the help of benchmarks the CDA can then bring anomalies and exceptions to the surface (Vasarhelyi, Alles, & Williams, 2010).

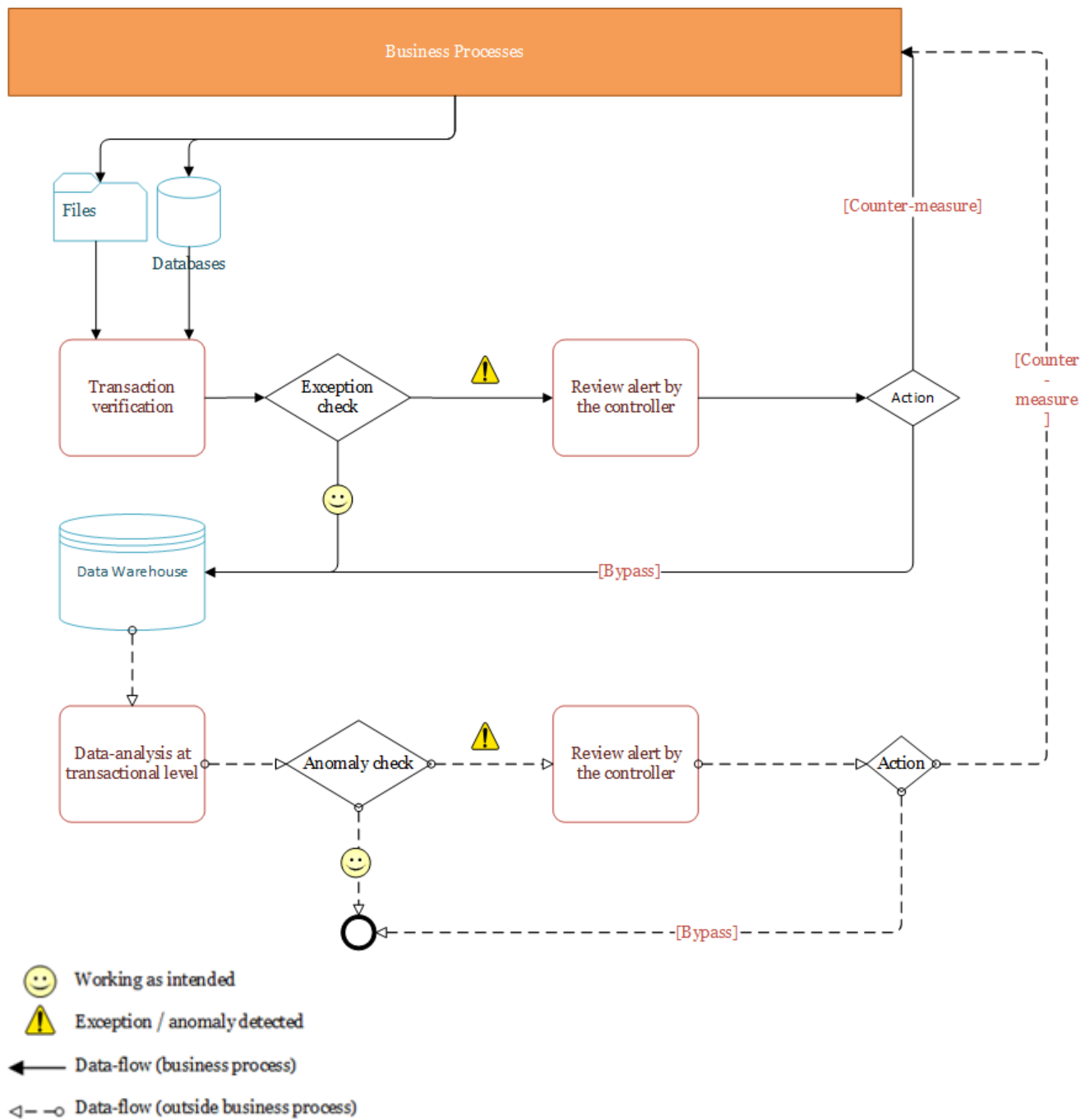


Figure 6: Continuous Data Assurance

3.1.2. Continuous Control Monitoring

Many large companies adopted ERP systems to support their business processes. Usually these systems contain extensive control settings to restrict prohibited behavior (Bumgarner & Vasarhelyi, 2015). CCM (see Figure 7) is a system that verifies whether business processes are proceeding in a controlled manner on a continuous basis. It presumes that both the controls and the monitoring procedures are definable. CCM is applicable to any kind of control environment, be it financial, quality or safety related (Hulstijn, et al., 2011). CCM consists of a control structure, which may or may not be implemented in ERP solutions, and a monitoring component that monitors control compliance, or in other words, whether these controls have the appropriate settings. For various number of reasons controls are not always satisfactory and may need to be temporarily re-parameterized. For example, the credit manager may decide that a customer is allowed to exceed the regular credit limit because the customer is very loyal and has a reputation of paying back on time. The operator could then decide to override the control settings or bypass the warning. This created a need to assure the control settings, and the nature of control overrides, by the auditor (Bumgarner & Vasarhelyi, 2015). CCM is a system that can fulfill this need.

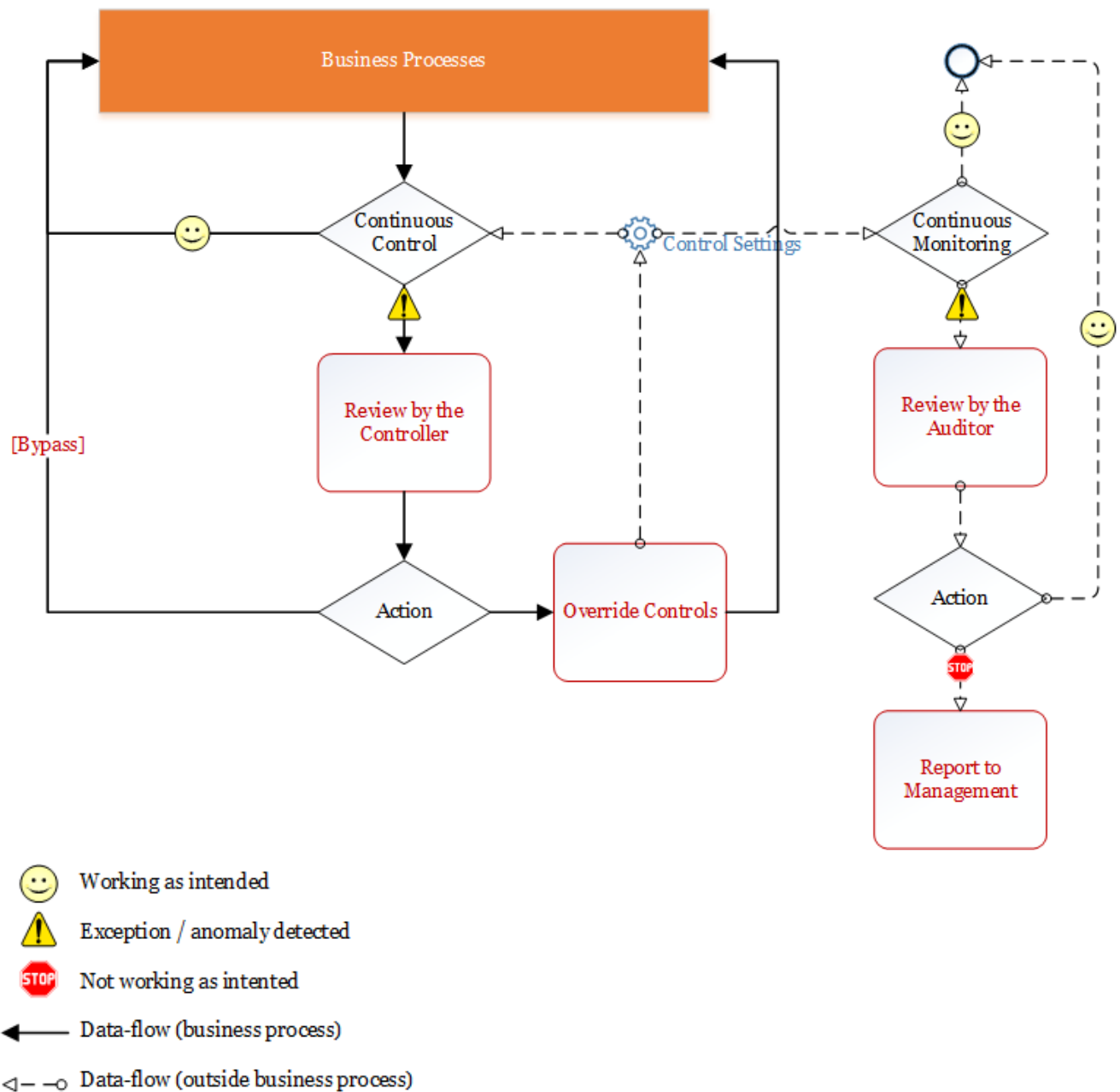


Figure 7: Continuous Control Monitoring

3.1.3. *Continuous Risk Monitoring and Assessment*

Unlike CDA and CCM, CRMA is an abstract concept that envisions the continuous monitoring of the risk landscape, by measuring leading risk indicators. Vasarhelyi, Alles, & Williams (2010) proposes a CRMA framework in which three types of risks are identified and assessed. These risk types are: operational risks, environmental risks, and Taleb's black swans (Taleb, 2007). The operational risks include several different risk types from risks in transaction processing, unauthorized activities and system risks to human, legal, informational and reputational risks (Durfee & Tselykh, 2011). Environmental risks are events caused by external forces like for example political turmoil, government regulations, enhanced competition, changes in customer preferences and changes in credit worthiness of external parties (Moon, 2016). A black swan is a low frequency, high severity event (Taleb, 2007) e.g., natural disasters, and the 9/11 terrorist attack. Although CRMA (see Figure 8) is useful to monitor several risk types, the focus of this research project is on the first risk category. In this concept these risk types are decomposed and examined to identify risk events and their pertinent risk drivers. Risk drivers are processes, people, systems and external dependencies. Once these risk drivers are mapped to associated risk events, leading Key Risk Indicators (KRIs) should be defined with a corresponding, and suitable, benchmark. These KRIs should be formalized and measured in such way that an entity is able to make a proactive response (Byrnes, Brennan, Vasarhelyi, Moon, & Ghosh, 2015).

Example: When an organization identifies potential liquidity risks caused by the inability of their clients to settle their liabilities, one might suggest the current ratio (ratio between current assets and current liabilities) as a suitable KRI. But once this current ratio is indicative of liquidity problems, an entity might already be incapable of paying their own liabilities. In this case, the current ratio is not a suitable indicator because it is reactive instead of leading. A more suitable indicator could be customer financial health trend information. If significant deterioration in the trend is detected, the entity could engage in preventive measures like increasing their cash positions

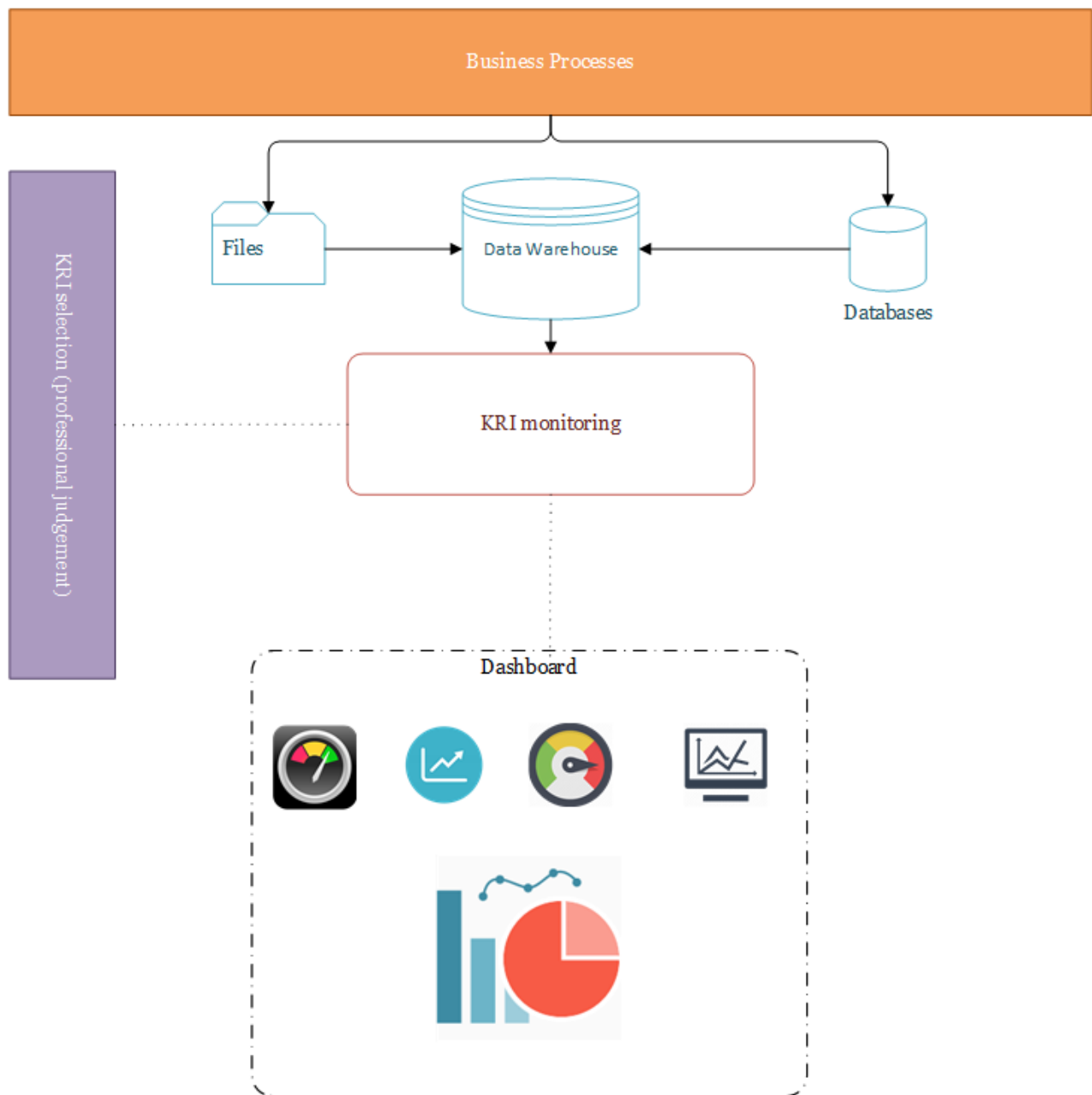


Figure 8: Continuous Risk Management and Assessment Scheme

3.2. Contribution to the research questions

RQ 3. *What is continuous auditing according to previous research?*

Continuous auditing is a philosophy that envisions a complete controlled environment in which business processes are monitored and analyzed on a continuous basis and in an automated fashion. Usually CA is divided into several functional components, each with their own unique capabilities. Continuous Data Assurance, addresses the correctness and reasonableness of transactional data. Continuous Control Monitoring monitors the control compliance of business processes. Continuous Risk Management and Assessment is a module that monitors and reports leading key risk indicators of an organization in real time. These components can work autonomously, but together they have synergetic effects.

4. Case study

In this chapter we will put the things that we have learned from previous chapters into practice. We will investigate the feasibility of a continuous audit approach by implementing it in a purchase-to-pay process of a client. We do this by simulating a continuous audit by creating, and implementing, CA algorithms that we back-test on historic data.

We start the chapter with a case description where we introduce client X. Next, we will elaborate on the algorithms behind the continuous auditing modules that we have built from scratch. We will end with results and lessons learned.

At the end of the chapter we will answer the following research questions:

RQ 4. *Can we assess operational risks efficiently and effectively at the case study?*

RQ 5. *What can we learn from the case study and what are its practical implications?*

4.1. Case description

The case study is done at Client X, a typical client of PwC regarding revenue and IT maturity. The client is a big player in the utilities industry with hundreds of millions in revenue. The dataset contains three fiscal years of P2P SAP export. Fiscal years 2014, 2015 and 2016. SAP is the number one ERP software package (Apps Run the World, 2016). The P2P process consists of multiple steps:

1. Purchase requisition (not mandatory)

Employees can file a purchase request, that reflects their desire for a certain good/service. Before a request becomes an actual purchase it needs to be evaluated by authorized personnel, and it might be declined or altered before it becomes an actual purchase document.

2. Purchase transaction

Authorized personnel can file an actual purchase order. A purchase document has a unique ID, and each purchase document can consist of multiple lines, and each line is usually a different good/service.

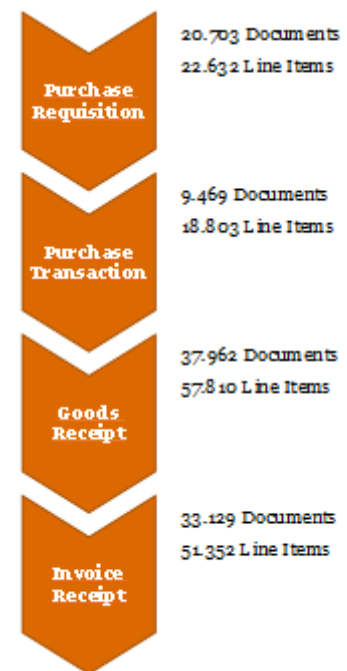


Figure 9: Purchase-to-Pay process of client X (2016)

3. Goods receipt

A document that proofs that good/service is actually received. Every good receipt document should have a corresponding purchase document. Multiple goods receipt documents can refer to the same purchase document when partial deliveries are made.

4. Invoice receipt

An invoice in exchange of a received service or good. According to best practice invoices should have corresponding goods receipt and purchase documents. Invoice receipts usually have multiple lines, and every line refers to a single type of product. An invoice receipt corresponds to a single vendor, but every line of an invoice receipt can refer to a different purchase document.

Typical purchases by client X are e.g., materials that are necessary to maintain or further develop the infrastructure that is necessary to provide the service of Client X, and services from external parties like for example a constructor.

The process that we described above is the typical example of a procurement process according to the three-way match principle. A three-way match occurs when there is a match between purchase document – goods receipt – invoice receipt, the purchase requisition is not a mandatory step. The process is designed in such way to minimize the occurrences of unexpected invoices, and to control the risk of procurement errors and fraud. An additional benefit is that this procedure can save an organization a lot of paperwork. If for example, a supplier of an ordered good delivers according to agreement the price of the invoice will not be a surprise and therefore should not be send back to the budget owner. A condition for a properly working three-way match procedure is segregation of duties (SOD). The steps of the procurement procedure should be executed by different employees (Groenewold & Rijn, 2004).

A P2P process can be easily formalized and is of a repetitive nature and we believe that this creates opportunities to implement a CA solution. The risks of this process are mainly due to failing business processes, systems and human error and thus are operational risks. The dataset is extensive, it contains vendor information and keeps track of changes that are made to the documents over time and therefore offers ample opportunities for this study.

4.2. *Simulating a continuous audit*

As our dataset contains several years of data and each year consists of ten thousands of documents, we are able to simulate a real time environment. Figure 10 gives an overview of the tables and their attributes that are of use and available to us. Every table starts with a table header, followed by at least one attribute that is a primary key (PK). A (set of) PK(s) can be used to access a specific row of a table.

For example, as mentioned before every purchase document has an ID and can consist of multiple lines, so if we want to access a specific line of a purchase document, we need the combination of both PurchaseDocID and PurchaseDocLineItem. A table can also have a (set of) foreign key(s) (FK). FK(s) are the access keys of an external table and are indicative of a relationship between these tables. When you want to join two tables, the FKs can be used as the join condition.














Invoices  PK InvoiceID  PK InvoiceLineItem CreationDate UserID VendorID ItemCat Quantity Amount  FK PurchaseDocID  FK PurchaseDocLineItem	PurchaseDocuments  PK PurchaseDocID  PK PurchaseDocLineItem Description MaterialGroup MaterialNumber CreationDate UserID VendorID Quantity Amount MaterialGroupDescription
GoodsReceipt  PK GRDocID CreationDate UserID VendorID MaterialNumber Quantity Amount  FK PurchaseDocID  FK PurchaseDocLineItem	Creditor  PK VendorID CreationDate UserID
	Creditor Changes to IBAN  PK VendorID  PK CreationDate  PK CreationTime UserID

Figure 10: Dataset tables

We can go through the data, sorted on date and treat it as it was supposed to happen in real time with knowledge limited to up to that point in time. We do this by creating an algorithm that is importing the invoices line by line and matches the invoice through the PurchaseDocID and PurchaseDocLineItem keys to the corresponding purchase documents and goods receipt documents if possible. We make sure that only the corresponding purchase and goods receipt documents that are created prior to the invoices are imported.

We learned from previous chapter that a continuous auditing approach consists of three different modules:

1. A CDA module that is able to identify exceptions and anomalies from transactional data.
2. A CCM module that is able to identify restricted behavior from a previous set of rules.
3. A CRMA module that keeps track of leading risk indicators.

We created algorithms to simulate the functioning of each individual module in practice. The algorithms are not readily available and should be designed in accordance with specific business processes. The algorithms we created are designed for a P2P process and are tailored to the specific case study of client X. The algorithms have underlying data analysis techniques similar to the auditors analytical procedures that we identified in Chapter 2.

For this research project we chose to focus on the detection of risky invoices as these are the risks with the most financial impact as invoicing leads to actual payment. The risks can be categorized as Knightian Uncertainties as underlying probability distributions are not known. But the financial impact, the price of the invoice, is known. We learned from Chapter 2 that by investing in data analysis more knowledge can be obtained about these risks and eventually they might become easily quantifiable Knightian Risks. Although the algorithms are in itself built for this specific case study, the underlying ideas are likely to be generalizable to other situations as well.

In the next subsections we will discuss the specific design of these algorithms.

4.2.1. CDA algorithm

We identified in the previous chapter that a CDA algorithm needs to be able to verify transactions for its correctness regarding data-entry and when correct it needs to be able to analyze this data by comparing it against previous organizational behavior. This means that it should analyze transactions against for example previously ordered, but comparable, goods/services. And that it should warn the controller when exceptions and anomalies occur. When they do occur, the invoice is more likely to be risky. The controller can then decide whether to block the invoice for further investigation, or to whitelist the transaction. When a transaction is whitelisted, the algorithm learns that this behavior is

normal and will treat similar transactions in the future as legitimate. When a transaction is blocked it will enter a different workflow which means that it will not proceed for payment until further authorization is given.

The CDA algorithm that we have designed identifies anomalies and exceptions from an invoice perspective and is designed to detect risks that have direct financial impact for the organization. After reviewing the data model for its capabilities we implemented the following features in the CDA algorithm:

Exceptions:

- Invoices with corresponding purchase documents that are entered into the system after the invoice
- Invoices with other vendorIDs than the vendorID on the purchase document.
- Invoices with higher quantities than is received (not received what will be paid for).

Anomalies:

- First time ordered good/service
- Invoices that have higher unit prices than comparable invoices from the past (also for services)

Next we will go into specific details about what the exception/anomaly means, why it is important and how the algorithm determines these exceptions and anomalies with Pseudo-code. Pseudo-code is informal and is often used in educational texts to describe algorithms, so that readers can understand it even without specific knowledge of a certain programming language. Pseudo-code does not have a systematic standard form and is often combined with natural language and mathematical expressions (Oda, et al., 2015). We will not go into detail about, all of the many, but very specific data manipulations and handling of missing values.

Exception: Invoices with corresponding purchase documents that are entered into the system after the invoice

This exception occurs when an invoice has a corresponding purchase document that is created after the invoice is entered into the system. A purchase document corresponds with a purchase order and a purchase order should always be previous to an invoice. When a purchase document is filed after an invoice it bypasses the principles of the three-way match. This can be risky, because it circumvents controls embedded in the three-way match process like mandatory authorizations of different employees, a match of quantity and amount between ordered/received/invoiced could be made fit afterwards.

```
Join Table(Invoices,PurchaseDocuments) By Keys(PurchaseDocID,PurchaseDocLineItem)
```

```
For every new invoice_line
```

```
Find Date(InvoiceCreation)<Date(Invoice_PurchaseDocCreation)
```

```
If TRUE Do Log(Exception_details), Warning(Exception_details)
```

Exception: Invoices with other vendorIDs than the purchase document

Client X orders goods/services from different kind of external parties (vendors). Vendor master details like name, address, IBAN and chamber of commerce number are stored in a table. Every vendor has a unique ID which functions as a key to access the master data of a specific vendor. When an invoice document has a different vendor ID from the purchase document it means that it might happen that eventually money is (unintentionally) transferred to a wrong bank account.

```
Join Table(Invoices&PurchaseDocuments) BY Keys(PurchaseDocID,PurchaseDocLineItem)
```

```
For every new invoice_line
```

```
Find String(InvoiceVendorID)≠String(Invoice_PurchaseVendorID)
```

```
If TRUE DO Log(Exception_details), Warning(Exception_details)
```

Exception: Invoices with higher quantities than is received

When an invoice has a higher quantity of what is actually received it means that there is a chance that the vendor invoices more than is fair. An invoice can correspond to multiple partial deliveries, so we have to sum over the individual goods receipts.

```
Join Table(Invoices&GoodsReceipt) By Keys(PurchaseDocID,PurchaseDocLineItem)
```

```
For every new invoice_line
```

```
Find Double(InvoiceQuantity)>Sum (Double(Invoice_GoodsReceiptQuantity))
```

```
If TRUE Do Log(Exception_details), Warning(Exception_details)
```

Anomaly detection

The detection of first time ordered goods/services are per definition an anomaly because they cannot be compared with other previously ordered of the same type. This type of anomaly is detected with the same algorithm that compares new invoices with historic invoices of the same type. The algorithm works as follows:

1. Make an item list which states all unique goods/services that are ordered. If corresponding purchase document info is missing come up with new values by replacing MaterialGroup for InvoiceItemCategory and by replacing MaterialNumber for a number representation of the vendor so that they are comparable with other similar type of invoices in the future.

Join Table(Invoices&PurchaseDocuments) By Keys(PurchaseDocID,PurchaseDocLineItem)

If Ismissing(purchaseMaterialGroup) Do Replace BY InvoiceItemCategory

If Ismissing(purchaseMaterialNumber) Do Replace BY VendorID

Create

MaterialList=Unique(Table(purchaseMaterialGroup,purchaseMaterialNumber,purchasegroupDescription))

AVGList=ZerosMatrix(Height(MaterialList),2000)

SDList=ZerosMatrix(Height(MaterialList),2000)

2. Calculate and remember moving averages and moving standard deviations of MaterialList unit prices of the historic invoices.

For every historic invoice_line

Find RowNumber=MaterialList.purchaseMaterialGroup&purchaseMaterialNumber And find
columnNumber=first empty column of AVGList & SDList

If columnNumber == 1

AVGList(rowNumber,columnNumber)=invoiceAmount

SDList(rowNumber,columnNumber)=0

Else

Calculate new moving average

Calculate new moving standard deviation.

3. Compare the current invoice details against historic details and determine if invoice is an anomaly, if so log it and create warning.

anomalyThreshold=2 (sigma)

Join Table(Invoices&PurchaseDocuments) By Keys(PurchaseDocID,PurchaseDocLineItem)

For every new invoice_line

If Ismissing(purchaseMaterialGroup) Do Replace BY InvoiceItemCategory

If Ismissing(purchaseMaterialNumber) Do Replace BY VendorID

Find RowNumber=MaterialList.purchaseMaterialGroup&purchaseMaterialNumber And find
columnNumber=first empty column of AVGList & SDList

If columnNumber == 1 Do Log(First_vendor), Warning(First_vendor)

AVGList(rowNumber,columnNumber)=invoiceAmount

SDList(rowNumber,columnNumber)=0

Else

If invoiceAmount>AVGList(rowNumber,columnNumber-

1)+anomalyThreshold*SDList(rowNumber,columnNumber-1) Do Log(Anomaly
Detected), Warning(Anomaly Detected)

Calculate new moving average, calculate new moving standard deviation

The anomaly threshold that is used is based on the normal assumption (see Figure 11). The zero point in this figure corresponds to the calculated moving average. An anomaly is detected when the invoice unit price exceeds the 2σ bandwidth. Under the normal assumption 2.2% of the observations are beyond this threshold. The observations beyond this threshold have a direct negative financial impact on the business. Observations below this threshold are not considered.

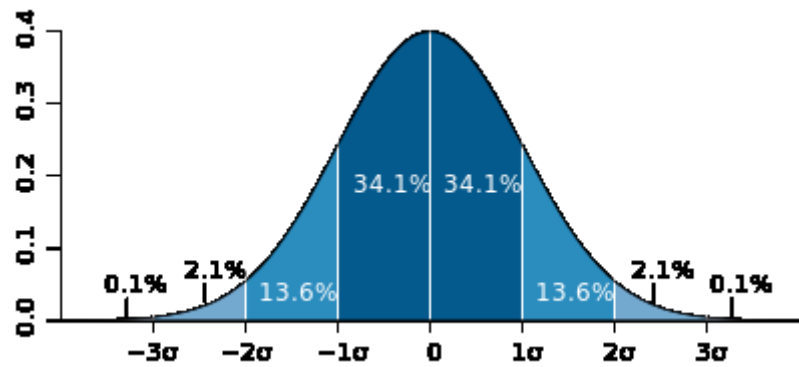


Figure 11: Normal distribution with sigma bandwidth (Wikipedia, 2017)

We will test The CDA algorithm that we have created on the case study. Results can be found in section 4.3.1.

4.2.2. CCM algorithm

In the previous chapter we identified that a CCM module consists of a control structure that may or may not be implemented in an ERP. At Client X this is the case, Client X uses SAP for their P2P process, SAP is known for its extensive control structure. As we do not have direct access to the current nor historic control settings of Client X's ERP we cannot monitor these settings. But that does not mean we cannot simulate a functioning CCM module. We can define controls ourselves, and we can monitor the whole dataset whether it behaves as allowed by the controls. The scanning of the control settings itself becomes redundant because we can directly deduct from the exceptions that the control settings are not in place.

We verified with an experienced IT auditor that was part of the client X audit team that the controls below are indeed controls that could be expected.

1. Invoices need to have corresponding purchase documents and goods receipt documents.
2. The three-way-match needs to be preserved with SOD requirements
3. Invoices may not be booked by the same user that also entered the vendor into the system.

4. Critical fields of vendor master data (IBAN) may not be changed by the same employee who did the invoicing for that vendor.
5. IBAN master vendor data may not be changed after office hours (07:00-19:00)
6. Invoices may not be more expensive than budgeted in the purchase document by 5%
7. Individual invoice line items may not be more expensive than budgeted unit prices by 5%

The above mentioned controls are easy to translate into very specific rules. We need to program the algorithm in such way that it checks whether these rules are obeyed. We can achieve this by matching the Invoice, purchase and good receipt documents through a database join. Once this is done we can test the relevant fields according to the business rules that we mentioned above. We will describe the complete functioning of the CCM algorithm in Pseudo-code in Appendix A.

We will test The CCM algorithm that we have created on the case study. Results can be found in section 4.3.2.

4.2.3. CRMA

Although CRMA is a more abstract concept in literature and implementations can differ greatly among business cases, three components are always present: the continuous monitoring of, the KRIs, and the dashboard. How the dashboards looks like, and which specific KRIs are used depends on to whom the information should be presented. Internal management of the client, might be interested in for example: third party unit price trend analysis of regular ordered goods/services or third party delivery times trend analysis. While the external auditor might be more interested in trend information to determine whether the specific anomalies and exceptions are structural or are isolated risks. To make a dashboard that is useful to the external auditor, we decided to make a dashboard that presents (trend) information about the exceptions and anomalies detected by the CDA and CCM algorithms. The dashboard that we used for this case study will be presented in section 4.3.3.

4.3. Results and lessons learned

We have three years of available data. The anomaly testing component of the algorithm (CDA) needs to be fed with data in order to learn what is normal behavior for the organization. The more data the algorithm is fed, the more accurate its anomaly detection capabilities. We could test a continuous audit on two different years. We simulated a continuous audit on fiscal year 2015, with the knowledge of fiscal year 2014. And we also simulated a continuous audit on fiscal year 2016, with the knowledge of both fiscal year 2014 and fiscal year 2015. The performance results of the CDA and CCM module are given below.

4.3.1. CDA algorithm results

In Table 2 we can find the plain results of the CDA algorithm of fiscal year (FY) 2015 and by plain we mean without any model parameterization (except for a necessary outlier threshold, set to 2σ). We programmed the algorithm in such way that a controller/auditor can fine tune the risk appetite of the algorithm. Above the already mentioned outlier threshold we can also set a minimum invoice line price, limit to certain types of invoices and set the minimum number of invoices the algorithm needs to detect anomalies. By loosening these restrictions, the amount of detected anomalies will consequently decrease. The algorithm detected for FY 2015 around 5,000 exceptions and anomalies which is a lot to review, it comes down to around 20 invoice reviews per day. In reality, when a controller actively monitors the P2P process, the amount of exceptions and anomalies will be reduced. This is because the controller will decide to whitelist certain risky transactions, and when he does, the algorithm will learn that this is acceptable behavior and will not mark similar transactions as anomalies anymore. From an external auditors perspective, the resulting subset of exceptions and anomalies is an efficiently and effectively obtained target to do substantial tests on.

Table 2: Continuous Data Assurance Results FY 2015 & FY 2016

CDA Results		2015	2016
General info	Total number of invoices_lineItems booked	49,550	51,390
	Total number of exceptions found	1,489	3,455
	Total number of anomalies found	3,448	3,379
Exceptions	Purchase documents entered after the invoice	41	45
	Invoices with mismatching vendorIDs with the PD	1,438	3,393
	Invoices with higher quantities than is received	10	17
Anomalies	First time ordered goods/services	427	114
	Invoices with significantly higher prices than comparable orders from the past	3,021	3,265

When we compare FY 2015 with FY 2016 (see Table 2) there are a few things to note. The first thing to note is that we observe that the number of first time ordered goods/services (first time from our algorithms point of view) in 2015 is a lot higher than in 2016, which we expected. Second we see that in 2016 there are many more mismatches between the vendorID on the purchase document with the vendorID on the invoice document. At first we thought that the algorithm made mistakes, but after taking some samples we saw that this was not the case (see Case study example 1). Last we are observing that there are more invoices with significantly higher prices detected than expected in FY 2016, while we expected that the algorithm would detect less because it can predict more accurately (it is fed more data) what anomalies are and what not. We suspect that this could be explained partly by the consequences of what was mentioned in Case study example 1 and maybe partly because unit prices are shifting more drastically than is accounted for by the algorithm.

Case study example 1.

Client X has a huge list of vendors (well above 10,000), although every vendor has its unique ID it does not necessarily mean that it is also a unique vendor / entity in reality. We observed the following typical example in Table 3. We see four times the appearance of an organization called Jan Smit, Jan Smit could have multiple entities but at least three of the entries have the same Tax Code so they must be the same entity. It seems that somewhere, either in the purchase department or the invoicing department (or both), the picking of the right vendor happens very sloppy. We see constantly small mistakes where almost similar in name, but different vendors in the system are chosen on both documents. This could have serious consequences when for example a vendor decides to change his banking details and not all the appropriate vendors in the list are changed accordingly.

Table 3: Vendor data

Vendor ID	Creation Date	Vendor Name	Address	Postal Code	Due days condition 1	Tax Code
1038945	1-1-2001	Jan Smit Energy b.v.		1000AA	30	NL952623043B01
1087854	28-2-2002	Jan Smit Energymanagement BV	Amsterdamsestraat 1	1000AA	30	NL757653025B01
2158756	2-2-2010	Jan Smit installatieadviseurs BV 1	Amsterdamsestraat 1	1000AA	14	NL757653025B01
4187859	8-12-2016	Jan Smit installatieadviseurs BV 1	Amsterdamsestraat 1	1000AA	30	NL757653025B01

Considering the CDA results and in particular the big number of detected exceptions and anomalies, you might question whether this algorithm, or a continuous auditing philosophy at all actually works. When we sampled some of these anomaly observations we actually saw that some of these anomalies were indeed obvious mistakes (see Case study example 2).

Case study example 2.

Figure 12 is an example of a real-time anomaly detection warning that we have built as a feature of our CA simulation. This particular example warns us about an invoice that is created for total price of €200 with a unit price of €2.00. It compares the ordered unit Houtdraadbout RVS 10x50 against 4 historic transactions and it warns us that the €2.00 unit price exceeds the 2-sigma threshold of €0.14 (about 14 times more expensive than allowed). Because this transaction is so different from the previous transactions the controller of the process should review this transaction. The controller can then decide

to trust this transaction, or to block the transaction. When a controller choses to trust a particular invoice, the algorithm will learns that this is allowed behavior and it will increase its future anomaly detection threshold.

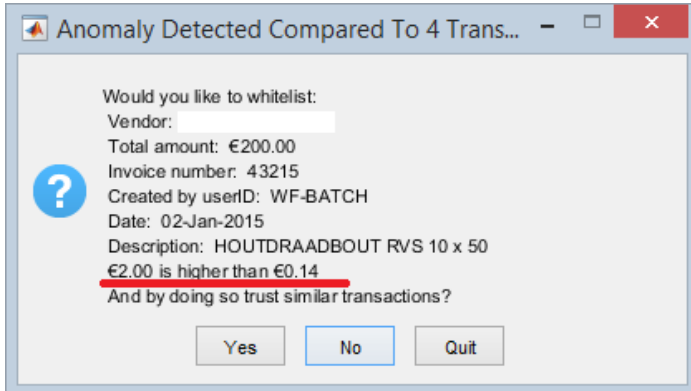


Figure 12: Real-time simulation anomaly detection example

As we were in the shoes of the controller, we looked into this suspicious transaction and observed the following in Table 4. We can see from this table, that this transaction indeed deviates from previous transactions by a lot. We actually can observe that this transaction was indeed a mistake and it is corrected 10 days later, something our algorithm could have prevented.

Table 4: Real-time simulation anomaly detection example (continued)

<i>Invoice NR.</i>	<i>Creation Date</i>	<i>Vendor</i>	<i>Material Number</i>	<i>Description</i>	<i>Quantity</i>	<i>UoM</i>	<i>Invoice price</i>	<i>Price per unit (calculated)</i>
462248	18-4-2013	Y B.V.	201739	HOUTDRAADBOUT RVS 10 x 50	100	ST	€13.65	€0.1365
475471	22-7-2013	Y B.V.	201739	HOUTDRAADBOUT RVS 10 x 50	100	ST	€13.65	€0.1365
41856	22-12-2014	Y B.V.	201739	HOUTDRAADBOUT RVS 10 x 50	100	ST	€13.65	€0.1365
43104	2-1-2015	Y B.V.	201739	HOUTDRAADBOUT RVS 10 x 50	100	ST	€13.65	€0.1365
43215	2-1-2015	Y B.V.	201739	HOUTDRAADBOUT RVS 10 x 50	100	ST	€200	€2
44495	12-1-2015	Y B.V.	201739	HOUTDRAADBOUT RVS 10 x 50	-100	ST	-€200	Correction

4.3.2. CCM algorithm results

In Table 5 we stated the results of the CA algorithm that monitors the control compliance.

Table 5: Continuous Control Monitoring Results FY 2015 & FY 2016

CCM Results		2015	2016
General info	Total number of invoices_lineItems booked	49,550	51,390
	Total number of exceptions found	14,250 (29%)	8,492 (17%)
Invoices need to have corresponding purchase documents and goods receipt documents	Missing Goods Receipt documents	4,985	6,846
	Missing Purchase Documents	8,702	1,014
The three-way-match needs to be preserved with SOD requirement	SOD requirements not met	32	11
Invoices may not be booked by the same user that also entered the vendor into the system and no critical master data may have been changed by that user.	Booked by same user	9	14
	Changed critical data	1	9
IBAN master vendor data may not be changed outside office hours (07:00-19:00)	After office hours changes	0	1
Invoices may not be more expensive than budgeted in the purchase document Individual invoice line items may not be more expensive than budgeted unit prices	Invoices over budget (>5%)	206	214
	Invoice unit prices over budget (>5%)	315	383

From Table 5 we can immediately see that most exceptions are caused by missing purchase documents and/or goods received documents. It is striking that there are a lot more missing purchase documents in 2015 than 2016. However, this can be easily explained because many invoices correspond to ongoing contracts and purchase documents that are prior to FY 2014 and not within our dataset. It is not surprising that in FY 2016 there are a lot less missing purchase documents. When the algorithm is fed with more data it will address these exceptions more accurately. When a purchase document is missing, the corresponding goods receipt document cannot be found due to the design of the data model (see 4.2). This is however not a new exception. So the rise in reported missing goods receipt documents in FY 2016 compared to FY 2015 can be solely explained because of this purchase document dependency. There are simply more goods receipt documents to monitor. We conclude that the P2P-process improved in terms of SOD requirements regarding the three-way match. Nevertheless, when looking at the result, there is still room for improvement at the following procedures:

- Vendor creation – Invoice booking
- Changing of IBAN vendor details – Invoice booking
- Invoices over budget
- Invoice unit prices over budget

4.3.3. CRMA dashboard view

We developed a dashboard that gives an overview of all the identified exceptions and anomalies by the CA algorithm over time during the invoicing process. These should be reviewed before actual payment is made. Best practice would be that the CA algorithm produces notifications before the invoicing is finalized. When doing so this CA algorithm is a preventive measure instead of a detective measure. The different type of exceptions and anomalies are therefore leading KRIs and could be monitored over time with the help of a dashboard such as Figure 13. The dashboard gives real time insight into what kind of errors are produced, which users are responsible, how much money is at stake, and if possible it calculates the immediate financial impact (IFI) for the organization. This is the total amount of money that is over the expected invoice amount². This is helpful for both the controller and the auditor. The controller can use this overview to manage the process in a more effective and efficient manner. The auditor can immediately see trend information and observe whether the anomalies and exceptions are isolated or structural risks. In addition he can filter the dashboards by date, users, and type of errors to get a better understanding of what is going on.

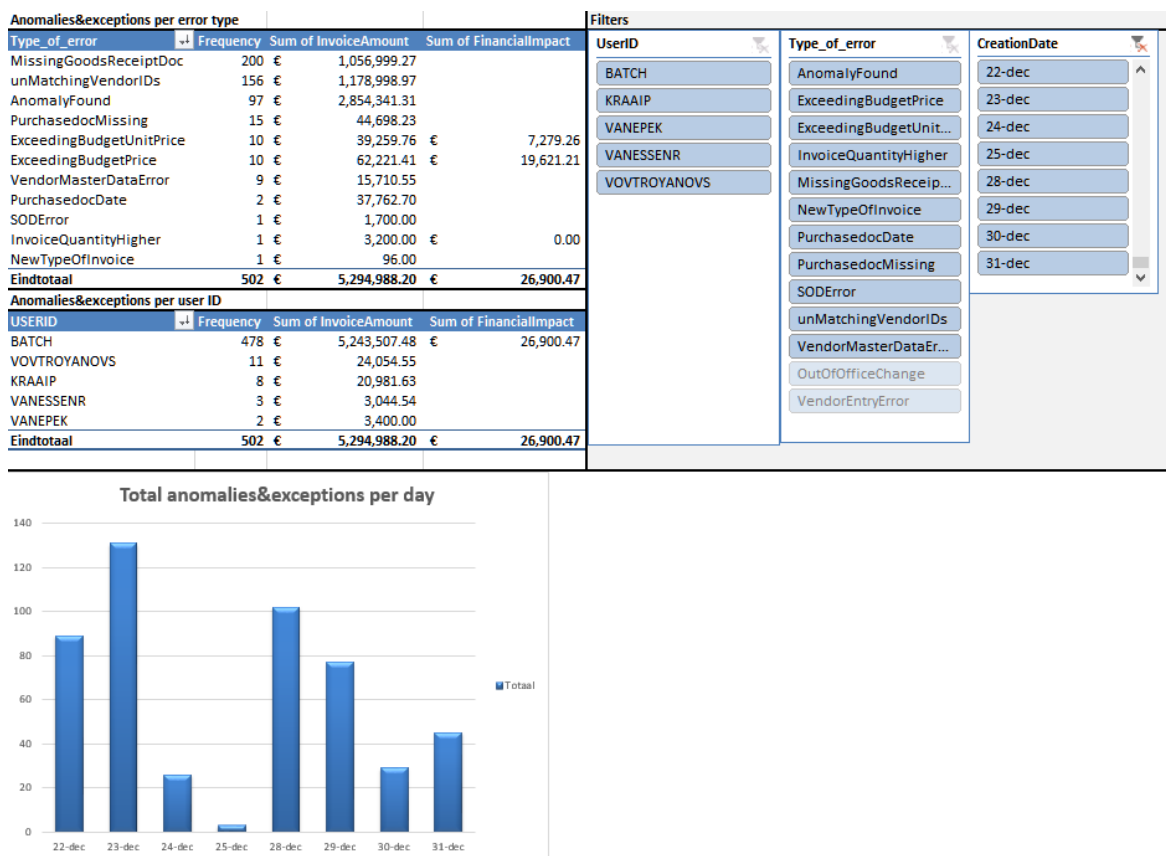


Figure 13: CRMA Dashboard

² IFI is only calculated for the ExceedingBudgetUnitPrice, ExceedingBudgetPrice and InvoiceQuantityHigher type of errors. For these types it is easy to calculate the IFI.

4.4. Contribution to the research questions

RQ 4. *Can we assess operational risks efficiently and effectively at the case study?*

The CA solution that we implemented at the case study of Client X is able to assess operational risks in an efficient and effective manner. The simulation proved that we could identify anomalies and exceptions with the help of the algorithms in an automated fashion. Moreover, numerous invoicing mistakes were detected that our algorithm could have prevented.

RQ 5. *What can we learn from the case study and what are its practical implications?*

A CA solution is capable of bringing process flaws the surface. CA is a helpful tool to support the auditor with determining whether the occurring risks are isolated or structural risks and with his sampling decisions. Moreover, a CA solution proves to be helpful for the clients internal control, it can help the client directly by functioning as an extra line of defense and it can assist the controller with the identification of potential process improvements. However, we also realized that the CA algorithm at client X triggered so many alerts that an auditor / controller is unlikely to review them all. The feasibility of a CA solution depends on the IT maturity of an organization. When an organization is fairly mature, CA is a solution that helps organizations to be more in control of their risky business processes.

5. *Continuous audit in a general context*

We have successfully implemented a CA solution for identifying risks in a P2P context at the case study. This created CA solution was built for a specific business case and is not mature enough to be of general use. This chapter aims to provide the reader with guidance on how to implement CA successfully in a more general context at other organizations than the case study of this research project. We will do this by defining minimum infrastructure requirements, and by designing a general implementation approach that is applicable to any CA business case.

At the end of this chapter we will answer the following research question:

RQ 6. How can we implement what we have learned into current audit practice?

5.1. *IT infrastructure*

The most important thing to consider when thinking about implementing a CA solution is IT infrastructure. The IT infrastructure of the client needs to be mature. We identified in Chapter 1 that the IT infrastructure of the clients is often not mature enough (see 1.2). Even at major companies that have fairly sophisticated systems, a seamless fit between systems and business processes might still be lacking. This is what we also observed at the case study, who used the sophisticated ERP SAP for their P2P process. There were still thousands of exceptions and anomalies that you could trace back to business process flaws. During this research project, and especially during the case study, we had to overcome several problems and saw things that influenced the results of the CA algorithm. When thinking about implementing a CA solution into a certain business process you should at least consider the following:

<i>Things to consider</i>	Choose one	
The business process is:	Repetitive	Not repetitive
The business process proceeds through:	Centralized systems, like ERPs	Many legacy systems
The supporting systems and databases:	Well-maintained	Ill-maintained
The span of control of employees within organization:	Well-defined	Ill-defined
There exists a seamless fit between defined span of control of employees and systems:	Seamless	Not seamless
Control structure of business process is:	Formalized	Not-formalized

When the answer to one of the questions from the previous page is the second option then a CA solution is probably not very feasible for that business process.

5.2. Continuous audit implementation

Given that the client has a mature IT infrastructure, the following design steps should be undertaken to implement a CA solution (successfully):

1. Obtain the data model of the studied process
2. Select the relevant variables for the analysis
3. Obtain relationships between the variables
4. Continuous data obtainment (store data into a data warehouse)
5. Choose a data aggregation level
6. Determine the window size
7. Built a(n) expectation model(s)
8. Design an exception and anomaly handling workflow
9. Test, implement and improve

The purpose of this chapter is to generalize the CA implementation of the case study towards other implementation areas. Steps 1-3 and 7 are very case dependent so we will not elaborate on these steps. Steps 4-6 are more easy to generalize into a few options, we will go into detail about these below. Steps 8 and 9 are very important, but go beyond the scope of this research project.

5.2.1. Continuous data obtainment: EAM vs. MCL

The design of how data flows between client and auditor is generalizable into two distinct approaches each with their own advantages. First you have the Embedded Audit Module (EAM) approach, in this approach the CA module is developed and implemented inside the walls of the ERP using its native programming language. One of the major advantage of this approach, is that incoming transactions can be evaluated as they happen in the application (in real-time). When these transactions trigger alerts, these could be pushed directly to the desired users within the ERP system itself (Kuhn & Sutton, 2010). There are also downsides however, there are sources in literature that report that non-vanilla ERP options are dramatically decreasing the performance of the enterprises its systems (Henrickson, 2009) (Debreceeny, Gray, Tham, Goh, & Tang, 2005). Even when using the ERPs native programming language, small background processes have rather adverse effects on the transaction processing

performance (Kuhn & Sutton, 2010). Second you have the Monitoring Control Layer approach (MCL), in this approach the CA module is developed and implemented outside the walls of the ERP. But instead the CA module is external software that is hooked into the clients databases (Kuhn & Sutton, 2010). A major advantages of this approach is that you can develop CA software independently from the distributor of the ERP software and you could develop CA software that is compatible with different types of ERPs. There are however also a few disadvantages. The obtainment of data is not real-time but happens through scheduled database exports (snapshots, pseudo real-time). These exports are then stored in a data warehouse that the CA software uses for its analysis. Consequently, when data is analyzed, alerts are not pushed in real-time to the end-user and more importantly are not pushed directly to the ERP system.

All things considered, it seems that EAM is the more favorable option for internal control because of the real-time monitoring push notification capabilities to the ERP itself. Eventually the ERP developers might overcome the performance issues, making it a feasible option. MCL is the favorable option for the external auditor, using that approach they can develop CA software that is scalable towards many different kinds of clients. Pseudo real-time is sufficient enough for the external auditor.

5.2.2. Data analysis level: individual transactions vs. high level aggregation

A critical decision that needs to be made during the design of the CA algorithm is the level of data aggregation. The use of data aggregation can reduce the variability that is observed among individual transactions and it facilitates the construction of a more stable model. Unstable models trigger too many alerts and are therefore of less practical use. The main argument against the use of data aggregation, is the loss of information about individual transactions. And when alerts are triggered, the review process is usually more expensive because a set of transactions, instead of a single transaction should be investigated (Kogan, Alles, Vasarhelyi, & Wu, 2014).

For this research project we chose to review transactions individually and compare each of them against comparable type of transactions in the past. Kogan Alles, Vasarhelyi & Wu (2014) proved to be successful with a CDA implementation of intermediate aggregation, they used daily and weekly aggregations at a case study of a major health institution. But, their data set consisted of at least 500,000 invoices per year which is around 10 times of what we had available. When datasets are not that extensive, we suggest to use individual transactions as the baseline for the expectation models.

5.2.3. Determine the window size: growing windows vs. sliding window

The next design decision is whether to use a growing window or a sliding window for anomaly transaction comparison. When a growing window is used, transactions are compared with an increasingly number of transactions from the past, old data is never discarded. Using a sliding window on the other hand, the total amount of data is kept constant, which means that an evenly amount of old data is discarded when new data comes in (Kogan, Alles, Vasarhelyi, & Wu, 2014).

When a business process is fairly stable a growing window offers more reliability, because the more data it is fed the more accurate the comparison. When a business process is changing over time and there is plenty data available to compare with a sliding window offers a better fit. There are various studies about the optimal window size, and how to optimize this trade-off between bias and forecast error variance (Pesaran & Allan, 2007).

5.3. Contribution to the research questions

RQ 6. *How can we implement what we have learned into current audit practice?*

CA is a viable option to use into current audit practice given that the IT infrastructure of the relevant client is mature and there is a seamless fit between the formalized business processes and the IT systems. When these constraints are met CA software needs to be developed that hooks into the databases of the clients according to the Monitoring Control Layer concept. The data model of the client should be studied and expectation models needs be built for the to be audited variables. To be useful this software needs be capable to perform the APs mentioned in RQ 2 in an automated fashion. To account for the existence of steady and changing environments several data aggregation levels and multiple window sizes should be considered.

6. Discussion

Now that we have dealt with all the research questions we should have a good idea what CA is and how it could help us measure operational risk. The aim of this project was to design a Continuous Auditing framework that helps PwC expand their Risk Assurance practice. We will start this chapter by answering the problem statement. Next we will discuss this research project's implications. What its shortcomings are and some directions for future research.

6.1. Conclusions

Problem statement: *“How can operational risk exposure be identified and assessed in a continuous audit environment from an external auditors perspective?”*

The CA implementation at the case study was successful. It proved that operational risk exposure can be identified and assessed on a continuous basis and in an automated fashion with the help of continuous auditing software that hooks into the databases of the client. This software features three Continuous Auditing components with synergetic effects. The Continuous Data Assurance and Continuous Control Monitoring components have underlying algorithms that automate analytical procedures that are used to identify exceptions and anomalies. The Continuous Risk Monitoring and Assessment component on its turn is used to assess these exceptions and anomalies for its nature and financial impact. Together they can bring the (intentionally) mismarking of positions to the surface which creates an opportunity to mitigate the risk of internal fraud and data entry errors.

When the external auditor wants to implement a CA philosophy into its current Risk Assurance practice for more clients than the case study, CA software should be developed. This research project contributes to the to be developed CA software by providing the underlying theory and inspiration for the design of the algorithms, which was the aim of this research project. The implementation of CA software into the current Risk Assurance practice is the solution to the core problem of this research project.

The core problem: *“external audit does not have a cost-efficient practice to do continuous risk assessments”*

The to be developed CA software can disrupt the current audit practice. Rather than analyzing aggregate data to plan audit targeting, and then conduct analytical procedures based on target samples, the whole data set can be scanned in an automated fashion for exceptions and anomalies. The resulting exceptions and anomalies lists can then be sampled, with a lower volume, to provide targets for additional field work. This new way of working not only increases the efficiency of the auditor, but also the effectiveness of the field work, since the target is already identified as being an exception or anomaly. When the auditor implements CA software in its current practice, it increases its value

proposition by providing more assurance to the client at an increased audit interval. However, not for all clients CA is a feasible solution. The IT infrastructure of the client needs to be mature, and there needs to be a seamless fit between the formalized business processes and the IT systems. If not, the CA software will prompt too many false alerts.

6.2. Limitations & Future Research

We assumed a fairly stable and unchanging business process during the design of the CDA anomaly algorithm. We did not investigate the appropriateness of other forecasting techniques such as simple exponential smoothing, Winter' seasonal forecasting method, time series decomposition, and linear regression. Nor did we experiment with a sliding window instead of a growing window and higher levels of data aggregation.

We used three fiscal years of data consisting of around 30,000 invoices each, which is still a rather small dataset. We could have had better results when more data was available to us.

There seems to be a consensus in literature on what CA entails. After implementing this philosophy at the case study, we concluded however that the use case of the external auditor is different from the use case of the internal auditor/controller³. Consequently we believe that the designed CA software during the case study has a better fit with internal audit than with external audit. This is because it is unlikely that the external auditor will review the generated alerts on a (near) real time basis. When the external auditor would be actively involved in this process, the role of the external auditor would change drastically. Their responsibility of business performance will increase as they become responsible for both CA software and the monitoring of alerts itself. We feel like the perspective of the external auditor is neglected in literature, and that most literature has a better fit with the perspective of the internal auditor. The applicability of CA for the external auditor should receive attention in future research.

In order to successfully implement CA software into current practice, research needs to be done about for example software design, IT security, IT communication protocols, hosting and ownership of data. Moreover, many governing bodies like the AFM need to approve the new way of working.

We proved that we could identify exceptions and anomalies, but once identified, we did not elaborate on how to proceed. In this area there remains opportunities for future research. For example the following questions are still open:

- How should communication proceed between the external auditor and the client?
- How often will the auditor sample the anomaly and exception list for additional field work?

³ Depends on the organizational structure of the subject could be useful to both

- How do we implement an audit trail into a CA solution?
- Who audits the audit trail?
- How will a CA solution impact our current audit practice?
- Will the auditor be held responsible when fraudulent behavior is not detected by the CA solution of the external auditor?

7. References

- Apps Run the World. (2016). *Top 10 ERP Software Vendors and Market Forecast 2015-2020*. Retrieved from Apps Run the World: <https://www.appsruntheworld.com/top-10-erp-software-vendors-and-market-forecast-2015-2020/>
- Aven, T., & Renn, O. (2009). On risk defined as an event where the outcome is uncertain. *Journal of Risk Research* 12, 1-11.
- BIS. (2011). Principles for the Sound Management of Operational Risk.
- Bumgarner, N., & Vasarhelyi, M. A. (2015). Continuous Auditing - A New View. *Audit Analytics and Continuous Audit: Looking Toward the Future*, 3-52.
- Byrnes, P., Brennan, G., Vasarhelyi, M., Moon, D., & Ghosh, S. (2015). Managing Risk and the Audit Process in a World of Instantaneous Change. *Audit Analytics and Continuous Audit: Looking Toward The Future*, 129-143.
- Campbell, S. (2005). Determining overall risk. *Journal of Risk Research* 8, 569-81.
- Chan, D. Y., & Vasarhelyi, M. A. (2011). Innovation and practice of continuous auditing. *International Journal of Accounting Information Systems* 12, 152-160.
- CICA / AICPA. (1999). *Continuous Auditing Research Report*. Toronto.
- Debreceeny, R., Gray, G., Tham, W., Goh, K., & Tang, P. (2005). Embedded Audit Modules in Enterprise Resource Planning Systems: Implementation and Functionality. *Journal of Information Systems* 19, 7-27.
- Diebold, F. X., Doherty, N. A., & Herring, R. J. (2008). *The Known, the Unknown, and the Uknowable in Financial Risk Management*. University of Pennsylvania.
- Durfee, A., & Tselykh, A. (2011). Evaluating Operational Risk Exposure Using Fuzzy Number Approach to Scenario Analysis.
- Financial Accounting Standards Board. (2006). Conceptual Framework for Financial Reporting: Objective of Financial Reporting and Qualitative Characteristics of Decision-Useful Financial Reporting. *Financial Accounting Series*, Vol. 1260-001.
- Financial Times. (n.d.). *Fat Tails Definition From Financial Times Lexicon*. Retrieved from Financial Times: <http://lexicon.ft.com/Term?term=fat-tails>
- Ganegoda, A., & Evans, J. (2014). A Framework to manage, the measurable, immeasurable and the unidentifiable financial risk. *Australian Journal of Management* 39, 5-34.
- Graham, J., & Weiner, J. (1995). *Risk versus risk: Tradeoffs in protecting health and the environment*. Cambridge: Harvard University Press.
- Groenewold, J., & Rijn, R. (2004). Het Principe Van Three Way Matching. *Controlling*, 6.
- Henrickson, R. (2009). Practioner Discussion of Principles and Problems of Audit Automation as a Precursor for Continuous Auditing. *University of Waterloo Centre for Information Integrity and Information Systems Assurance* 6th.
- Hulstijn, J., Christiaanse, R., Bharosa, N., Schmid, F., van Wijk, R., Janssen, M., & Tan, Y.-H. (2011). Continuous Control Monitoring-Based Regulation: A Case in the Meat Processing Industry. *Lecture Notes in Business Information Processing* 83, 238-248.

- IFAC. (2009). International Standard on Auditing 315.
- ISO. (2002). *Risk management vocabulary. ISO/IEC Guide 73*. Geneva.
- Knight, F. (1921). *Risk, Uncertainty and Profit*. Boston and New York: Houghton Mifflin Company.
- Kogan, A., Alles, M. G., Vasarhelyi, M. A., & Wu, J. (2014). Design and Evaluation of a Continuous Data Level Auditing System. *Auditing: A Journal of Practice & Theory*, 221-245.
- Kuhn, J. R., & Sutton, S. G. (2010). Continuous Auditing in ERP System Environments: The Current State and Future Directions. *Journal of Information Systems* 24, No. 1, 91-112.
- Lowrance, W. (1976). *Of acceptable risk - science and the determination of safety*. Los Altos, CA: William Kaufmann Inc.
- Merriam-Webster. (n.d.). *Definition of Risk*. Retrieved from Merriam-Webster: <https://www.merriam-webster.com/dictionary/risk>
- Moon, D. (2016). Continuous risk monitoring and assessment: CRMA.
- Oda, Y., Fudaba, H., Hata, H., Sakti, S., Toda, T., & Nakamura, S. (2015). Learning to Generate Pseudo-code from Source Code using Statistical Machine Translation. *IEEE/ACM International conference on Automated Software Engineering*, 574-584.
- Pesaran, M. H., & Allan, T. (2007). Selection of estimation window in the presence of breaks. *Journal of econometrics*, 134-161.
- PwC. (2015). *Evolving internal audit - Starting your journey to continuous assurance*.
- PwC. (2017). *Global PwC Audit Guide 2017*.
- Taleb, N. (2007). *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.
- Vasarhelyi, M. A., & Halper, F. B. (1991). The continuous audit of online systems. *A Journal of practice & theory* 10, 110-125.
- Vasarhelyi, M. A., Alles, M., & Williams, K. T. (2010). *Continuous Assurance for the Now Economy*.
- Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big Data in Accounting: An Overview. *Accounting Horizons* 29, 381-396.
- Wikipedia. (2017, June 14). *Standard deviation*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Standard_deviation
- Woodroof, J., & DeWayne, S. (2001). Continuous audit model development and implementation within a debt covenant compliance domain. *International journal of accounting information systems*, 169-191.

Appendix A

Join Table(Invoices&PurchaseDocuments&GoodsReceipt) By
Keys(PurchaseDocID,PurchaseDocLineItem)

Join Resulting Table(Invoices_PurchaseDocuments_GoodsReceipt&Creditor&IBANChanges) By
Keys(vendorID)

Exception=FALSE

budgetTreshold=1.05

While Exception=FALSE

For every new invoice_line

%%Control 1: Invoices need to have corresponding purchase documents and goods receipt documents.

If Ismissing(Invoice_PurchaseDoc.purchaseDocID) OR
Ismissing(Invoice_Goodsreceipt.GRDocID)

Exception=TRUE

%%Control 2: The three-way-match needs to be preserved with SOD requirements

If StringCompare(invoiceUserID,purchaseUserID,GoodsReceiptUserID)

Exception=TRUE

%%Control 3: Invoices may not be booked by the same user that entered the vendor into the system.

If StringCompare(invoiceUserID,CreditorUserID)

Exception=TRUE

%%Control 4: Critical fields of vendor master data may not be changed by the same employee who did the invoicing

If StringCompare(invoiceUserID,IBANChangesUserID)

Exception=TRUE

%%Control 5: IBAN master data may not be changed after office hours

If IBANChangesCreationTime<09:00 OR IBANChangesCreationTime>19:00

Exception=TRUE

%%Control 6: Invoices may not be more expensive than budgeted in the purchase document by 5%

Find InvoiceID AND PurchaseDocID

Calculate for these documents

invoicePrice=SUM(invoiceAmount)

purchasePrice= SUM(purchaseAmount)

If InvoicePrice>budgetTreshold*purchasePrice

Exception=TRUE

%%Control 7: Individual invoice line items may not be more expensive than budgeted unit price by 5%

If InvoiceAmount/InvoiceQuantity >

budgetTreshold*purchaseAmount/purchaseQuantity

Exception=TRUE

If Exception=TRUE

Do Log(Exception_details), Warning(Exception_details)