MODELING OPERATIONAL RISK IN DATA QUALITY

(Practice-oriented paper)

Anton Raneses
Oklahoma State
econ1992@yahoo.com

Camille Deyong, Ph.D.
Oklahoma State
camille.deyong@okstate.edu

Neil Glasstein
US Navy
nglasstein@hotmail.com

Abstract: In this paper, we address how data quality (DQ) is likely linked to failed business processes that pose operational risks to the Enterprise system. Operational value at risk (OPVAR), which is used in the finance literature to mean how much we might expect to lose if an event in the tail of the loss probability distribution *does not occur*, can be used to conduct Enterprise software reliability and damage function analysis. This paper explores (a) how to combine distributional assumptions for event frequency and severity to derive software loss cost estimates using the familiar example of software processing errors and (b) how to utilize the estimates of this distribution to estimate OPVAR-based losses. The empirical results show (a) that it is possible to fit DQ problems, such as the daily mishandling event data, to a distribution and to use maximum likelihood analysis to derive a consistent set of critical event count thresholds and (b) that the resulting OPVAR-based losses can be used by DQ managers to ascertain the real costs of mitigating DQ problems.

Key Words: Data Quality, Operational Value at Risk, Poisson and Weibull distributions, Expected Shortfall

INTRODUCTION

Organizations often find it difficult to see the distinction between internal risks (e.g., operational risks) and external risks (e.g., regulations, credit, and market conditions). This gap is particularly important for DQ managers because there is a grey area of risk disparity between products (e.g., DQ products) and processes (business requirements). Figure 1 illustrates the gap between product- and process-related risk types. From a product-related perspective, [1], [2], [3], and [4] investigate the extent to which DQ can be empirically estimated as a function of process risk and is demonstrated in the failure mode and effects analysis literature (FMEA). [5] inquires into market and non-market valuation types of DQ quality supply schedules that can be empirically estimated. These are all useful measures of risk; however, these models are comparatively static and lack the dynamic ability to quantify risk over a period of time.

Aspect	Type of risk		
Product related	Market risk	Credit risk	Other product risks
	Grey zone of risks hard to classify		
Process related	Operational risk		

Figure 1: Operational risks differ from market and credit risks (KPMG, 2001)

BACKGROUND, RATIONALE, AND PURPOSE

Since the publication of the Basel Committee's report on operational risk in September 1998 and the promulgation and execution of Section 13 of the Sarbanes-Oxley Act, functional IT specialists and DQ managers have started to speculate that perhaps the internal-external risk gap pertaining to DQ could be addressed by OPVAR. The Basel Committee report on operational risk analysis requires that financial institutions realize the benefits of measuring and streamlining the flow of capital, people, and information into and out of the organization through the use of TQM methods such as Six Sigma and Lean Manufacturing. In fact, "banks that have implemented efficiency programs such as Six Sigma are discovering significant overlaps between the data required by those programs and the data required to comply with Basel II [6]." OPVAR and operational risk indicators are currently utilized by global banks and financial institutions whose liquidity is ensured for the protection of public trust.

Operational risk indicators, which are the risk that the organization's operations or business processes will not generate the expected returns as a result of both external factors and internal factors, are random variables that are used to provide insight into future OR events [7]. Factors outside the organization can include government regulations: for example, the Data Quality Act, Clinger-Cohen Act, and Federal Financial Management Improvement Act. A change in regulatory policy could have an impact on the immediate business process requirements of the Enterprise. Mandated changes in an accounts payable system will tend to cause temporary sustained losses and earnings foregone as a result of mismatched requirements and transaction failures in the areas of processes, information systems, or people. Thus, operational risk indicators that are carefully tracked by organizations include transaction failures, outdated business processes, mismatched requirements, and new automation. In data quality, operational risk indicators tend to provide insight into future problems at their earliest stages so that preventive action can be undertaken to avoid or minimize a serious OR event.

The purpose of this paper, therefore, is to illustrate how OPVAR techniques that are currently utilized in the banking sector to estimate capital charge allocation could be utilized to solve data quality problems through software functionality reliability analysis and the cost of software upkeep.

Literature Review

When setting aside capital for OR events, banks and other financial institutions aim to ensure the availability of sufficient economic capital to allow continued operations in an adverse environment or when internal operational failures have generated large unexpected losses [9]. Accordingly, OPVAR is defined as "the operational risk capital sufficient, in most instances, to cover operational risk losses over a fixed time period at a given confidence level." OPVAR can be calculated if we know $F^{-1}(1-\alpha)$, where $F^{-1}(1-\alpha)$ is the percentile function in $F(x)$, where $F(x)$ is a cumulative distribution of a random variable. Thus, a 99 percent OPVAR implies α = 5 percent, which in this case is the 95th percentile, and thus $OPVAR_{0.95} = F^{-1}(1\text{-}0.95) = F^{-1}(0.05)$.

OPVAR captures both the economic and regulatory implications of risk in a financial institution, as shown in Figure 2 [10]. Basel II regulates banks on the basis of equity capital risk—the risk that the bank needs to set aside its reserves based on its financial operations. Figure 2 shows that operational risk is captured by the combined probabilities of event and loss given a certain event. Some of the measurement tools used to capture that information includes loss data pools and integration of external loss data reported in the areas of content analysis, Web-based coverage, and early warning functions. Operations exposure is manifested in the form of risk indicators, while the quality index can be portrayed by a risk inventory list. The quantitative methods used to model this type of framework include extreme value theory, cause and effect models, and other simulation models.

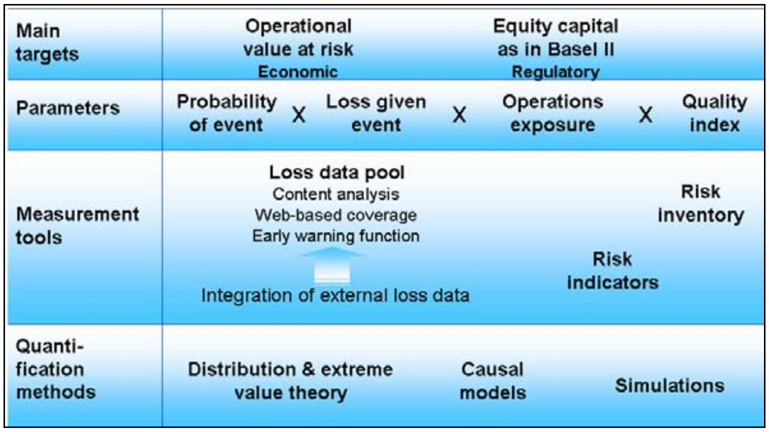

Figure 2: Calculation of Economic & Regulatory Capital (KPMG, 2001)

Figure 3 shows that financial institutions utilize OPVAR to measure, for any combination of business lines *i* and type of risk *j*, the combined probabilities of loss of event PE, loss of a given event LGE, operations exposure OE, and a quality index QI, where QI is independent from PE, LGE, and OE.

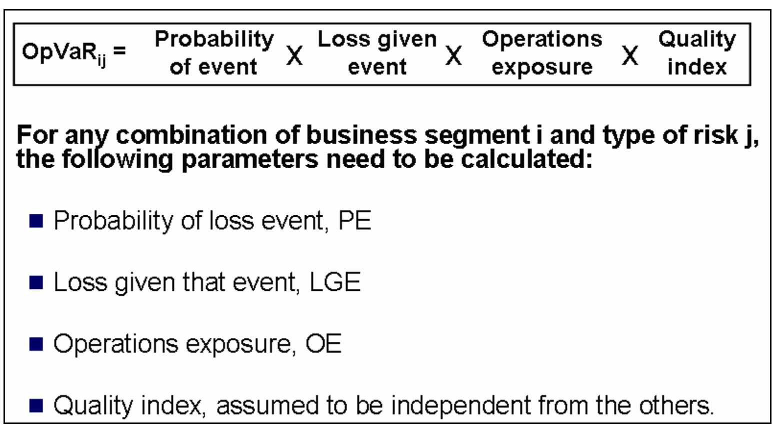

Figure 3: Financial Sector Definition of OPVAR (KPMG, 2001)

The Basel II Accord outlined in [9] accounts for three methods to calculate operational risk capital charge:

> "The first is the basic indicator approach, in which the required capital is determined by multiplying a financial indicator such as gross income by a fixed percentage. The second is the standardized approach, in which a bank divides its function into a number of business lines. For each business line, the required operational risk capital is calculated by multiplying an indicator, typically the gross income or asset size of the business line, by a fixed percentage. The total operational risk capital charge is the sum of the required capital across all the business lines. The third method is a type of internal measurement approach based on a bank's internal risk management system. In this approach, operational risk is categorized based on business lines and

event type determined by the regulators. The total capital charge is the sum of the required capital across each business line and each even-type combination."

Figure 4 shows the expected capital charge of a bank—it is determined by regulators as the expected operational risk loss EL_{ij} for each business segment and for each risk type $\gamma_{ij,}$ where EL is equal to the regulatory exposure per type of risk and business line segment EI_{ij}, probability of loss event PE_{ij} for type of risk *j* in the business segment *i* determined by banks, and the loss given event LGE_{ij} for type risk *j* in the *ith* business segment determined by banks.

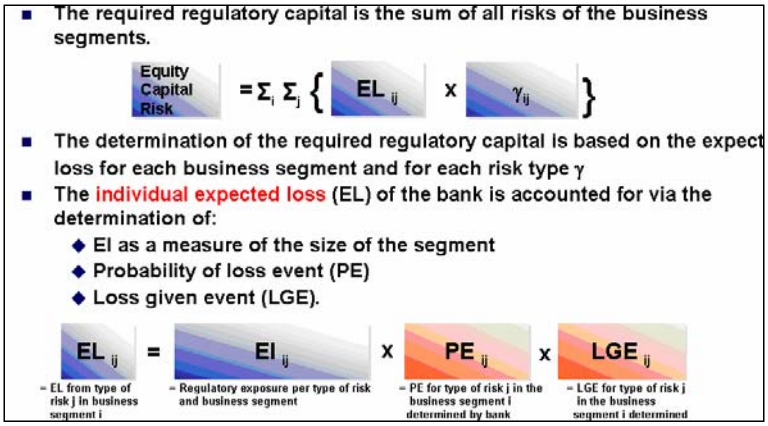

Figure 4: Capital Charge Derivation {Basel II Internal Measurement Framework (KPMG, 2001)}

OPVAR and Data Quality

Banks aim to set aside capital for operational risk events to ensure the availability of sufficient economic capital and to allow continued operation, especially when internal operational failures have generated large unexpected losses. Similarly, ERP-based organizations must mitigate software failures that usually originate from outdated business processes. OPVAR could be used by DQ managers as the operational risk capital sufficient to cover most DQ-related operational risk losses over a period of time at a certain risk confidence level. OPVAR measures the distribution percentile, disregarding data quality losses beyond a point called α. OPVAR therefore indicates the greatest amount an organization can expect to lose if that αth percentile occurs.

Bank capital risk was previously used to model the true costs of data quality by defining costs as a function of risk and multiplying by γ, the factor determined by the organization's DQ manager. [2] previously addressed the use of FMEA to approximate the true costs of data quality by quantifying the risk parameters severity, probability of failure, and detectability. This approach simply involves multiplying the nominal costs of IT manual workarounds by the combined probabilities of fault and escaping detection. Unfortunately, this modeling framework is not dynamic, and it relies heavily on risk priority numbers (RPNs) and on the costs of workarounds when the actual costs are not available. On the other hand, OPVAR does not require RPNs and allows simulation of costs over time.

Since data quality researchers developed clearer definitions of IQ, such as accuracy, believability, relevancy, and timeliness [10]—dimensions of IQ that determine whether the quality of information meets or exceeds the requirements needed to solve business problems—other researchers, such as [2], [4], [11], [12], [13], and [14], have provided a framework for the taxonomy, shape, and empirical dimensions of data quality cost structures. Modeling OPVAR in data quality contributes to the DQ literature by incorporating relatively new techniques to discover how internal factors—methodological failures in ERP transactions—affect software reliability and software maintenance costs. To obtain management visibility, data quality researchers can sample transaction failures periodically.

As in FMEA, the frequency of errors in the functionality of IT systems should be assessed, as well as their severity, measured in its relative economic value. In this manner, sampled transactions can be checked and corrected where indicated, then revalued. Comparing this result to the original valuation provides an estimate of the severity (model errors) for IT transactions. Extrapolating the empirical results from the sample to the overall population will provide an estimate of the frequency, severity, software functionality mismatch, and business process requirements of the organization. Using the OPVAR method, it will be possible to fit the frequency and severity distributions and define critical values using confidence intervals. Ultimately, these results can be used to control the sampling process, and depending on the point in the lifecycle where the error occurs, a causal model can be used to translate the error into a true DQ-driven loss estimate.

METHODS

Modeling Data Quality Losses

The methodology depicted in this section is adopted from [7]. Accordingly, OPVAR analysis shows that by studying transaction processing errors, distributional assumptions for event frequency and severity can be combined to derive loss estimates. Transaction handling losses can be handled as a single distribution. However, combining separate distributions for the data mishandling event process and the severity is preferable, as this provides insight into the root causes of DQ losses. The event process for handling transaction errors is likely best approximated as a Poisson process. Its frequency of error distribution is that of a Poisson variable as well. In the example below, Mondays and Fridays tend to have a higher proportion of mishandled transactions than other days (Figure 5). The numbers of transaction mishandling events on the various days of the week follow different Poisson processes with respective parameters. On Mondays, for instance, the number of mishandled transactions is distributed as $P(\lambda_{mo})$, while that for Tuesday is $P(\lambda_{tu})$, and so on.

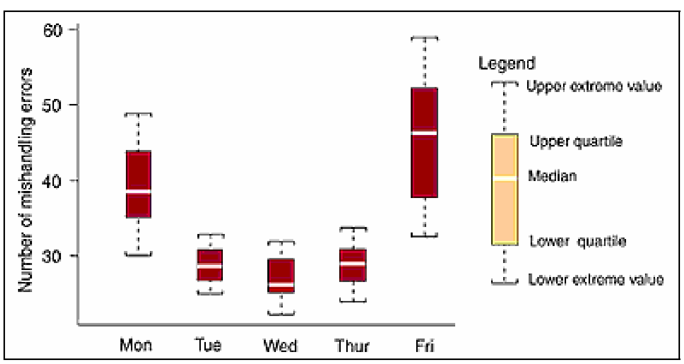

Figure 5: Mishandled Errors (O'Brien, 2006)

After fitting the daily mishandling event data to a distribution, a consistent set of critical event count thresholds are derived for each day of the week using Maximum Likelihood Analysis. By using the same confidence intervals applied to the daily distributions, the appropriate warning signs regarding the source of the mishandled error are identified. When estimating the continuous variable that describes the severity of transaction mishandling events (e.g., penalty payments in the case of mismatched supply order and delivery terms), the usual choice is the Weibull distribution (Figure 6).

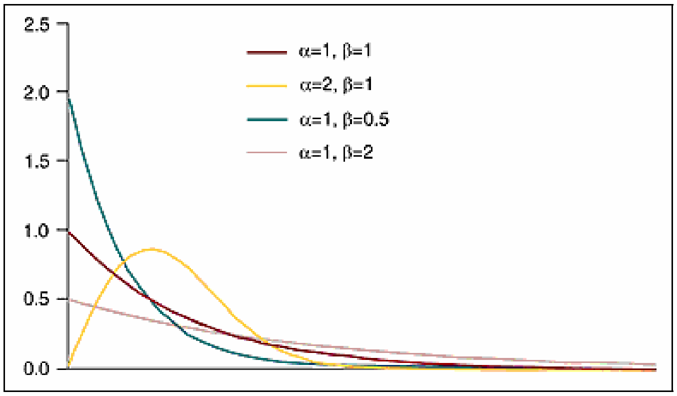

Figure 6: Weibull distribution (O'brien, 2006)

The shape of the Weibull distribution is primarily a function of its parameters α and ß and its probability density function, which is given by:

$$f(x) = \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$$

where $0 \leq x$, $0 < \alpha$, and $0 < \beta$.

To model total losses, a mixture of the Poisson and Weibull distributions is formulated. For simplicity, these mixed distributions are modeled as a compound Poisson process. The total loss due to mishandled DQ transactions, or severity amount, S (t), for some time interval (0, t), forms a compound Poisson process if (a) the frequency of mishandled transaction events forms a Poisson process; (b) the individual loss amounts are independent and identically distributed; and (c) the individual loss amounts are independent of the number of events N(t).

If the mishandled transaction events occur in accordance with a Poisson process with rate λ and the moment generating function (MGF) of the individual loss amounts (random variable x) is Mx(u), then the MGF of S(t), the mixture distribution, is:

$$M_{S(t)} = e^{\lambda t \left[M_x(u) - 1\right]}$$

It can be shown that

$E(S(t)) = \lambda t m_1$

and

V(s(t))=λtm2

where m1 is the mean of the Weibull distribution and m2 its variance.

$$m_1 = \beta\Gamma\left(1+\frac{1}{\alpha}\right)$$
$$m_2 = \beta^2\left[\Gamma\left(1+\frac{2}{\alpha}\right)-\Gamma^2\left(1+\frac{1}{\alpha}\right)\right]$$

Taking the overall loss distribution and using it to attribute risk capital to the overall transaction process will lead to mean and variance estimation of the total loss from the DQ mishandling. Thus, just as banks have to scale the loss distribution from daily totals to the confidence interval level as imposed by the financial institution's capital risk allocation policy, DQ managers must measure the loss distribution based on the organization's risk aversion policy to DQ errors. DQ managers must therefore have a DQ "risk capital policy," a set of rules that allocates a portion of the firm's IT budget to fix the functional distortions that are attributed to detrimental business requirements in check. Simulation can then be used to aggregate loss distributions across multiple operational risk categories.

In spite of the quantitative analysis aspect of OPVAR, reducing internal business process risk is ultimately the goal. This can be addressed by either modeling data flows or capturing the breakdown in business process requirements and making appropriate improvements in the Enterprise. The greatest gains from OPVAR modeling, therefore, will involve streamlining the organization's core workflows.

The Data

Time series data were obtained for Physical Movement ID, Action, Source, Type of Transaction, Operation, Volume Summary, Transaction Date, Material, Mode, Load, Discharge, Document ID, Contract Number, Order Number, Batch, Ship Number, Last, Funding, Signal Code, Customer Supply DODAAC, Suffix, Reason Code, Deemed Date, Create Date Time, Last Update Date Time, Movement Level, Original Movement ID, Freight, Error, and Physical Movement ID. The data stem from an Oracle Constellar Report from a Department of Defense database for the period of January 27 through July 29, 2005 (Figure 7). We used a sample dataset of 396 to simplify the scope of the analysis and to amplify the results of the suggested modeling framework.

The Value of Damage Function is a variable that originates from multiplying the Volume Summary, the quantity of fuel procured by the Department of Defense for its customers (e.g., Army, Navy, Marine Corps, Air Force), by the DoD standard price of $1.78 in fiscal year 2006, the rate DoD charges its customers for the administrative costs of procurement for the year. These administrative costs are used as proxies for the estimated premium used by the Defense Finance & Accounting Service (DFAS) to charge an overhead for administrative actions stemming from data quality errors. The data quality errors are deeply rooted in the business requirements that define the software functionality. For instance, the error "Using the supply order, the delivery terms and delivery point were not found" stems from a business rule that the delivery terms (e.g., pipe, barge, truck) and delivery point (e.g., origin, FOB destination) were not placed in the supply order. The Value of Damage Function variable will be the foundation of the Weibull analysis: (a) reliability table, (b) Weibull parameters, and (c) survival graphs. These Weibull results will be used to obtain both operational value current loss and expected shortfall.

Physical Mvt Id	10657		10630	1499552	1499564
Action	NAAN		NAAN	NAAN	NAAN
Source	PCSPORTS		PCSPORTS	FCCFAS	FCCFAS
Typ	PURCH		PURCH	PURCH	PURCH
Op	Issue		Receipt	Receipt	Receipt
	Value of Damage Function	$178.00	$1,185.48	$138.84	$117.48
Vol SUM	100	100	666	78	66
Tranx Date	01-APR-04	0	01-AUG-04	01-DEC-04	01-DEC-04
Matl	9130012720983		9140013980697	9130010315816	9130010315816
Mode	TRUCK (A)		TRUCK (A)	PIPELINE (8)	PIPELINE (8)
Load					
Dschg	UCNTOE		N68805	UCNTOE	UCNTOE
Document Id	N4523A409200 01		N6880542100001	UY700243368J0B	W90JA043368J0B
Contract	02D4542		02D4536	02D0532	02D0532
Order	GE26		GE74	Z188	Z176
Batch	PCS4092		PCS4214	GBL00001	GBL00001
Ship No	PCS4092		PCS4214		
Last					
Fund	XX		XX		
Sig Code	X		A		
Cust Sup Dodaac			FP2500		
Suffix					
Reason Cd					
Deemed Dt	01-APR-04		01-AUG-04		
Crt Date Time	25-Apr-2005		20-Mar-2005	25-Jul-2005	25-Jul-2005
Last Updt Date Time	13-Jul-2005		11-Apr-2005	25-Jul-2005	25-Jul-2005
Esp Mvt Level	0		0	0	0
Original Mvt Id					
Freight					
Error	Process flag is 77, submit request for technical support to re-process movement.	**Process flag is 77, submit request for technical support to re-process movement. Count**	Using the supply order, the delivery terms and delivery point were not found.	Using the supply order, the delivery terms and delivery point were not found.	Using the supply order, the delivery terms and delivery point were not found.
Phy Mvt Id	10657	1	10630	1499552	1499564

Figure 7: DoD Constellar Report (Department of Defense)

RESULTS

To obtain OPVAR estimates, we utilize a reliability engineering analysis to determine Weibull parameters α and β and Poisson parameter λ. These parameters are needed to fit the daily mishandling event data to a distribution such that a consistent set of critical event count thresholds are derived for each day of the week using Maximum Likelihood Analysis. To estimate the spread or shape of the distribution of failure times, we conducted a Weibull cumulative distribution function transformation so that it appears in the familiar form of the straight line $Y = mX+b$. By applying Ordinary Least Squares to the logarithmic transformation of (1/(1-Median Ranks)) on the intercept and the logarithmic transformation of (damage costs), this provides a straight line once the predicted Y and actual Y are placed on scatter plot. We obtain parameter estimates that will enable us to make inferences about the software design's reliability.

The α and β parameters are 638.7 and 0.980210544, respectively. Figure 8 below shows that based on the current software design, about 29.5 percent of the software functionality should survive at least at the 121st cycle or threshold of $638.7. Since β borderlines between less than 1.0 and close to 1.0, there are two possibilities: (a) 1.0 indicates a constant failure rate, which means that the software functionality components that have survived burn-in will subsequently exhibit a constant failure rate; (b) less than 1.0 indicates that the product has a decreasing failure rate, a feature of "infant mortality" indicating that the product is failing during its ”burn-in“ period.

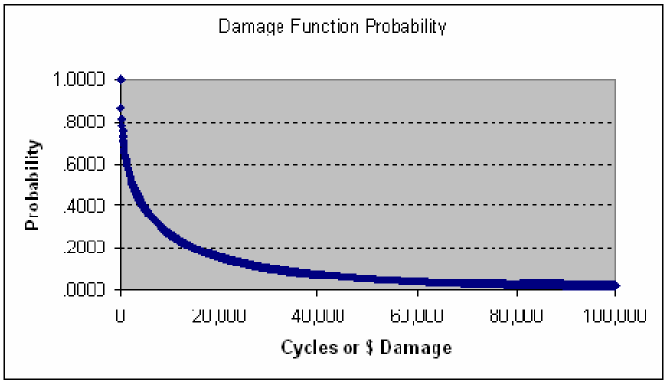

Figure 8: Approximately 30 percent of the software functionality should survive at least 121 cycles or at the $638.7 threshold.

To estimate λ, consider the frequency of loss daily data for software shown in Figure 9. The first column shows the potential number of errors of this category per day, beginning at 0 and ending at 10. For each data quality error category, the number of observations is recorded in column i and the number of events per day is logged in column n. The product of both columns is placed in the third column, i * n. Summing the third column and dividing it by the number of events equates to an estimate of λ, which is 49.45.

Category	Obs (i)	No. events per day (n)	i * n
Neither a stock point nor delivery point were found for the DoDAAC in the location table	3	0	0
Neither a stock point nor delivery point were found for the DoDAAC	8	1	8
Neither a stock point nor delivery point were found for the DoDAAC in the location table	8	2	16
Neither a stock point nor delivery point were found for the DoDAAC in the location table	3	3	9
Neither a stock point nor delivery point were found for the DoDAAC in the location table	1	4	4
This transaction was a UBAN and Multiple records were found to match the movement id	12	5	60
This transaction was a UBAN and No records were found to match the movement id	9	6	54
PW2, FOB water purchase (carrier) uploaded from supplier suspended; no matching cargo number	16	7	112
The reason code was invalid	15	8	120
TW8 - Requires an existing AU movement with this batch number	18	9	162
Process flag is 77, submit request for technical support to re-process movement	0	10	0
Estimated Lamda = **49.54545455**			

Figure 9: Poisson model parameter estimation based on frequency of loss data

As explained in the methods section, we obtain aggregate operational risk losses by collecting data on frequency and severity of losses for a particular operational risk type and then fitting a frequency and severity of loss model to the data. The simplest technique to obtain this result is Monte Carlo simulation through the following steps: (a) choose a severity of loss and frequency of loss probability model—this was accomplished through a Poisson frequency of loss model and a Weibull severity of loss model; (b) simulate the number of losses and individual loss amounts and then calculate the corresponding aggregate loss; (c) repeat many times (at least 5000 times) to obtain an empirical aggregate loss distribution. The empirical results are displayed in Figure 10. The Monte Carlo simulation shows that given the previously estimated Weibull parameters α and β and Poisson parameter λ, the current loss is approximately $43 million. At 95 percent confidence, the operational value at risk is equal to $59.62 million. Given the frequency and severity of loss model, we compute the aggregate operational loss distribution of each risk type and business line—data quality errors stemming from underlying software functional failure and business process inconsistencies. This is similar to the way financial institutions determine the cost of capital; the estimated OPVAR at $59.72 million is the "capital" required by data quality managers that could be used to mitigate operational loss experience across business lines in the Enterprise.

Number Simulations	
5000	
Weibull	Parameters
α	683.7
β	0.980211
Poisson	Parameter
λ	49.5
Current loss	$43.09
Count	5000

Percentile	OpVaR (Millions $'s)
0.95	$59.72
Weibull	Parameters
α	683.7
β	0.980211
Poisson	Parameter
λ	49.5

Figure 10: Monte Carlo simulation results

DISCUSSION

Thus far, the paper has focused on controlling operational risk; to do so, the risk must first be measured. Unfortunately, operational value at risk is not the optimal measure of risk because aggregating individual risks does not increase overall risk, or is not a coherent risk measure [8, p. 110]. A consistent alternative to OPVAR is expected shortfall (ES), which is the average value of losses that can be expected if a loss in excess of OPVAR is observed. [8] states that "ES informs data quality managers how much might the organization expect to lose if an event in the tail of the distribution does occur." Figure 11 shows the results of the probability weighted average loss beyond OPVAR, which is done by slicing the tail of the aggregate loss distribution above the OPVAR confidence level into *N* slices and then calculating the percentile of each slice. The results show that the average of these slices provides an estimated ES of $63.5 million. This value is larger than the estimated OPVAR value of $59.72 million because "it reflects what one can expect to lose on average if an event beyond OPVAR occurs."

Confidence Level		Slice	Percentile	Value
95.0%				
		1	95.5%	$59.7423
		2	96.0%	$59.7522
OpVar	$59.7227	3	96.5%	$60.7077
N	ES	4	97.0%	$60.7262
10	$63.4666	5	97.5%	$61.6811
25	$62.7961	6	98.0%	$61.7107
50	$62.5057	7	98.5%	$62.6838
100	$62.4192	8	99.0%	$64.6309
500	$62.3434	9	99.5%	$66.6052
1000	$62.3382	10	99.99%	$76.4256
5000	$62.3328		Average of values	$63.4666

Figure 11: Expected Shortfall results

LIMITATIONS

Other risk-measuring techniques, such as the probability weighted average loss beyond OPVAR, and other algorithms with properties that a good risk metric should possess could be more optimal than OPVAR itself. When aggregating individual risks does not increase overall risk due to a non-coherence risk measure violation, OPVAR cannot become the optimal risk measuring technique. While the proof of this coherent risk measure violation is beyond the subject of this paper ([8], p. 110), readers will only need to be aware of the context of two random uncorrelated losses X and Y and a risk measure that is denoted by ρ(). If ρ() is an optimal risk measure, it will have to satisfy the subadditivity criteria: For all X and Y, $\rho(X+Y) \leq \rho(X) + \rho(Y)$, which implies that aggregating individual risks does not increase overall risk. Thus ES is used as the average value of losses that can be expected from data quality problems if a loss is observed in excess of OPVAR.

CONCLUSION

In this paper, we address how data quality is likely linked to failed business processes that pose operational risks to the Enterprise system. Operational value at risk is at least one of the proposed tools to estimate appropriate costs required to mitigate data-quality-driven IT functionality problems. We estimate OPVAR by constructing the aggregate loss distribution, assuming that simulated losses are a result of the combined frequency and severity of loss distributions through the Weibull and Poisson models. For each frequency/severity of loss model, the parameters α, β, and λ must be estimated before loss distribution can be simulated. DQ managers will benefit from knowing that risk measures such as operational value at risk or expected shortfall (approximately $63.5 million in this case study) can be estimated by fitting the loss distribution of the organization.

REFERENCES

[1] Raneses, Anton; Thompson, Mik; and Winston, Gregory. "Business Process Re-Engineering at DESC: Saving Time and Money through more Efficient Work Methods." Defense Energy Support Center, Defense Logistics Agency. 17 October, 2003.

[2] Raneses, Anton. "Beyond Business Process Reengineering: Data Quality Engineering." 9th International Conference on Information Quality. The 9th International Conference on Information Quality, MIT, Cambridge, Massachusetts, 2004.

[3] Eppler, Martin J. and Markus Helfert. "A Framework for the Classification of Data Quality Costs and an Analysis of Their Progression." The 9th International Conference on Information Quality, MIT, Cambridge, Massachusetts, 2004. 311-325.

[4] Tallberg A., *An Economic Framework for Information Integrity.* Forskningsrapporter Research Reports, Swedish School of Economics and Business Administration, Finland, 1999.

[5] Raneses, Anton and Mielke, Michael. "Information Quality Cost Curves." 10th International Conference on Information Quality. 3-4 November 2005. Massachusetts Institute of Technology, Cambridge, MA.

[6] Accenture. "Mastering Bank Mechanics: Making operational risk management a value driver." *The Point*. 2004.

[7] O'Brien, Niall, Smith, Barry and Allen, Morton. "The case for quantification." http://www.financewise.com/public/edit/riskm/oprisk/opr-modelsp.htm

[8] Da Costa Lewis, Nigel (2004). *Operational Risk with Excel and VBA: Applied statistical methods for risk management*. New York: John Wiley.

[9] KPMG. "Solutions for Managing Operational Risk." http://www.garp.com/library/Papers/OperationalRisk/ papers/Garp%20KPMG%20OpRisk.ppt

[10] Richard Y. Wang, Yang W. Lee, Leo L. Pipino and Diane M. Strong. "Manage Your Information as a Product." *Sloan Management Review*, summer 1998, Vol. 39, No. 4, pp. 95–105.

[11] Smith, Nancy Duvergne. "Is Corporate IQ (Information Quality) High Enough? The Perils of Dirty Data Inhibit Strategic Decisions: Interview with Richard Wang, Director of MIT Information Quality (IQ)." *Impact*, winter 2003: Faces of Research. http://web.mit.edu/ctpid/www/i11/impact11_prof.html

[12] Shankaranarayan, G., Ziad, Mostapha, and Richard Y. Wang. "Managing Data Quality in Dynamic Decision Environments: An Information Product Approach." *Data Management*, Feb 2003.

[13] Mandke, Vijay and Nayar, Madhavan, "Modeling Information Flow for Integrity Analysis." The 3rd International Conference on Information Quality, MIT, Cambridge, Massachusetts, 1999.

[14] Carbone, Thomas and Donald D. Tippet. "Project Risk Management Using the Project Risk Failure Mode and Effects Analysis." *Engineering Management Journal*, volume 16, no. 4, December 2004.