Interdependence between Structural Engineering and Construction Management
Edited by Ozevin, D., Ataei, H., Modares, M., Gurgun, A., Yazdani, S., and Singh, A.

ISBN: 978-0-9960437-6-2

QUANTITY UNCERTAINTIES IN REINFORCEMENT WORKS – COMPARISON OF PUBLIC VERSUS PRIVATE CLIENTS

MARKUS KUMMER

Institute of Construction Management and Economics, Graz University of Technology, Graz, Austria

Uncertainties are systematically considered and dealt with by applying probabilistic calculation methods, such as Monte Carlo simulations. When selecting appropriate distribution functions for input parameters, users are constantly faced with the issue of having to choose the “right” distribution function for the relevant parameter. Quantities of individual works play a crucial role for costing and pricing, but also for construction process and logistics planning purposes. Quantities stated by the client in its structural specifications are fraught with uncertainties owing to, for instance, incomplete plans at the time of specification, inaccurate calculations, or mere estimates. This is why actual quantities can either be greater or smaller than the specified quantities. This paper demonstrates how distribution functions can be derived from expert surveys delivering responses from actual construction practice. Specific reference is made to reinforcement works whilst distinguishing between public and private clients. The outcomes of the survey presented and discussed in this paper include descriptive data analyses as well as violin plots and fitted distribution functions.

Keywords: Distribution functions, Expert survey, Data fitting, Management of chances and risks, Monte Carlo simulation.

1 INTRODUCTION

Any probabilistic calculation requires the selection of distribution functions for uncertain input parameters. When trying to choose the “right” distribution, users are constantly confronted with the issue which of the many available theoretical (mathematically described) distributions would best reflect the characteristics of the relevant parameter. After selecting the suitable type or kind of distribution, the corresponding shape parameters must be defined to be able to perform a Monte Carlo simulation (Goulet and Smith 2011, Raaber 2003, Sander 2012, Mayer 2013).

Three methods and their combinations are generally available for selecting distribution functions: theoretical considerations based on known or assumed characteristics, gathering data from construction practice, and expert surveys.

In this context, theoretical considerations refer to the basic characteristics of the relevant distribution function, such as discrete vs. continuous, open vs. closed boundaries, mode, and skewness. The second basis for selecting an appropriate distribution function is to gather data directly from construction practice. In this case, an additional distinction is made between historical data (*ex post* – after the construction phase, for instance information taken from final costing or daily work reports) and data collected during the construction phase (*inter actio*[1] – such as from direct site observations). When continuously gathering data during the construction phase, appropriate boundaries or limits must be defined in terms of space, time, and content whilst also ensuring data quality. Furthermore, collected data must be assessed to find out if data can be applied directly or if there is any need for fitting. Only if data has been put in, or linked to, the appropriate context information can be generated from it for the purpose of making this data usable for additional analyses and calculations. Another key criterion that determines data quality is the number of data points, or number of observations. The smaller the sample and the greater the standard deviation of gathered data, the greater the inaccuracy of the derived estimation of the parameters of the underlying population (Hofstadler and Kummer 2017).

AbouRizk and Halpin (1992) state that it is not possible to categorize distributions on the basis of fewer than 20 observations. It should be noted, though, that 20 responses taken from expert surveys provide a higher quality of information than 20 data points taken from observations, which, for instance, can be extracted from various projects using related logs and records. This principle holds true particularly if special emphasis is placed on the selection of appropriate experts, the drafting and preparation of a standardized questionnaire, and the discussion of answers with respondents (similar to the Delphi method).

For the purpose of this research, expert surveys (*ex ante* surveys) were used to determine distribution functions for quantity deviations. Our preference for expert responses over individual measurements taken from construction practice is based on the fact that any expert statement already incorporates experiential evidence from several past projects. Moreover, respondents are capable of linking objective data to (known) complex contextual conditions in construction practice as well as to heuristics, which adds value to their responses. Any identification of quantity deviations between specification and invoicing will also involve the issue of defining appropriate boundaries with respect to contract addenda and/or cancelled works.

2 EXPERT SURVEY

Experts are defined as individuals who possess specific knowledge and intellectual skills and competencies in a clearly delineated field and who serve as a source of specific knowledge for the purpose of a survey. Expert knowledge usually comprises exceedingly large amounts of information, including simplifications, lesser-known facts, rules of thumb, and smart practices (i.e. heuristics) that enable efficient problem solving (Gläser and Laudel 2010). In a preselection process, a total of about 130 experts from Austria and Germany with experience in the fields of costing, process planning, construction, final costing, and invoicing were contacted in writing and asked to participate in the survey. In total, 27 experts were recruited for the survey. They provided 23 responses with respect to quantity deviations in reinforcement works.

The majority of respondents (i.e., approx. 63%) worked for large companies with more than 250 employees, about 26% worked for medium-sized businesses (50 to 249 employees), and about 11% came from small businesses (10 to 49 employees). In the survey presented in this paper, experts had an average professional experience of 17.7 years; this experience ranged from 5 to 41 years. Questions were designed and developed together with social researchers (i.e., sociologists) in several revision steps, applying the principles of simplicity, clarity, impartiality, and specificity. On average, each respondent was interviewed for about 45 minutes either on the phone or in a face-to-face session to overcome ambiguities, collect missing information, or obtain

[1] Latin for: during (work) execution

background information and justifications of responses (Hofstadler and Kummer 2017).

3 DATA FITTING

The issue of which distribution functions to apply to individual calculation parameters essentially determines the quality of the output of a Monte Carlo simulation. There is no uniform answer to this question. In the literature, both considerations regarding the "nature" and potential characteristics of individual parameters, the incorporation of historical data, as well as the use of expert opinions are discussed.

The transfer of such data records (e.g. from *ex ante*, *ex post* or *inter actio* surveys or observations) into distribution functions that can be described on a theoretical level is referred to as "data fitting". In this process, existing data is first analyzed graphically in histograms and then assessed using mathematical/statistical methods in order to identify the theoretical distribution function that best reflects existing data (i.e. with the smallest deviation). Since distributions need to be adjusted to project-specific circumstances or conditions, the shape or curve of distributions of uncertain calculation parameters is of greater interest than absolute values.

An expert survey is recommended as the method to select the underlying data because the *inter actio* and *ex post* gathering of data is likely to present definition and allocation problems in the course of data processing. Data processing is followed by actual data fitting, which involves a comparison and check for consistency of adjusted data with theoretical distributions that can be described mathematically. This theoretical distribution can then either be applied directly to the uncertain input parameter, or the input is optimized for the fitted distribution. This optimization step provides the advantage that the boundaries of the distribution can be modified on the basis of theoretical considerations and characteristics, which significantly simplifies the input (of minimum and expected values, for example) for the user. Furthermore, optimization makes it possible to define a conversion scheme to scale and adjust the derived distribution function to individual/project-specific conditions (Hofstadler and Kummer 2017). Fitting aims to determine the parameters of a distribution function (displacement and shape parameters) in a numerical process so that the distribution reflects gathered raw data as accurately as possible.

4 QUANTITY DEVIATIONS IN REINFORCEMENT WORKS

An expert survey conducted by Kummer and Hofstadler at Graz University of Technology in 2015/16 collected responses with respect to the magnitude of deviations from quantities stated in the specifications of reinforcement works expected for public and private clients (Kummer 2015, Hofstadler and Kummer 2017). No absolute quantities but percentage deviations from the specified quantity were requested, with the latter being defined as 100%. Experts determined the lower limit of quantity deviations (quantity reduction) by indicating a percentage value smaller than or equal to 100% and the upper limit (quantity increase) by indicating a percentage value greater than or equal to 100%.

In the expert survey, 23 responses were received with respect to quantity deviations in reinforcement work (see Table 1). The figures stated for public and private clients differed only to a minor extent, which permits the conclusion of a similar overall degree of accuracy of specifications for reinforcement works. Differences were primarily found for extreme values (i.e. minimum and maximum values). A mean quantity reduction to up to about 91 and 92% was observed for reinforcement works (see Table 1, line 2, cells B and D, and line 14, cells B and D).

The greatest specified quantity reduction declines to 70% in the case of public and to 60% in the case of private clients (line 10). Mean expected quantity increases amount to 7 to 8% (see Table 1, line 2, cells C and E, and line 14, cells C and E). A slightly larger quantity increase should be assumed for public-sector clients relative to the mean values and robust M-estimators (according to Huber). This tendency is also confirmed when considering the maximum values stated for quantity increases. Experts stated a quantity increase of up to 35% for private clients, whereas this increase amounted to up to 40% for public clients (line 11). The calculated range of M-estimators is roughly identical for public and private clients.

Table 1. Descriptive representation of data pertaining to quantity deviations in reinforcement work.

No.	Parameter	Reinforcement – steel bar – public clients		Reinforcement – steel bar – private clients	
		Negative deviation	Positve deviation	Negative deviation	Positve deviation
0	A	B	C	D	E
1	N	23	23	23	23
2	Mean	91.57	108.30	91.30	108.70
3	Standard error of mean	1.57	1.81	1.77	1.90
4	Standard deviation	7.53	8.66	8.48	9.10
5	Coefficient of variation	8.22 %	7.99 %	9.28 %	8.37 %
6	Median	95.00	110.00	95.00	110.00
7	Mean absolute deviation	5.00	5.00	5.00	6.00
8	Robust coefficient of variation	5.26 %	4.55 %	5.26 %	5.45 %
9	Mode	95.00	110.00	90.00	100.00
10	Minimum	70.00	100.00	60.00	100.00
11	Maximum	100.00	140.00	100.00	135.00
12	Spread	30.00	40.00	40.00	35.00
13	Skewness	-1.25	2.24	-2.45	1.26
14	M-estimator (H12)	92.51	107.26	92.65	107.54
15	Standard deviation (H12)	6.69	6.73	5.78	7.99
16	Robust coefficient of variation	7.24 %	6.27 %	6.24 %	7.43 %
17	Normal distribution (Shapiro-Wilk)	No	No	No	No

Figure 1 shows the ranges of expert responses but also the violin plots relating to them. Violin plots are an extension to the known box plot representation. They combine the information of a density function (or of a smoothed histogram) with that of a box plot in a single diagram. Their advantage is that the graphical representation includes not only box plot information but also distribution characteristics. In the diagram, the density curve is plotted symmetrically above and below the (vertical) box plot. The two halves of the violin plot are thus mirrored along the axis on which the relevant parameter is shown.

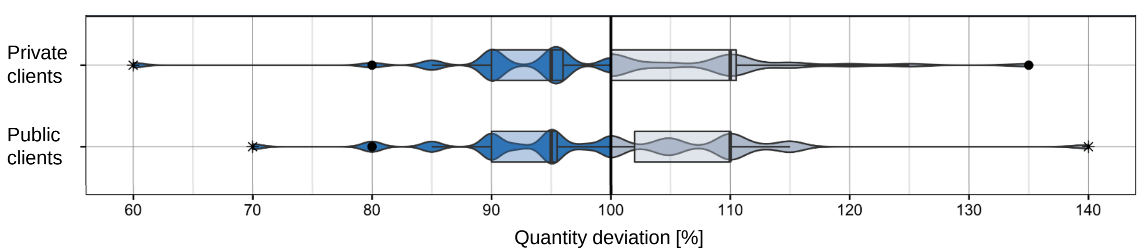

Figure 1. Violin plots of quantity deviations in reinforcement work – public vs. private clients.

Another benefit of violin plots is that the shape of the distribution or the presence of any accumulations or clusters within certain value ranges is easier to visualize. This type of information can get lost in conventional box plots. The greater the frequency of tapered zones in the violin plot, the larger the likelihood of a greater frequency of expert responses within certain value ranges. For instance, responses got clustered at quantity reductions of 90% and 95% in relation to reinforcement works both for public and private clients. Between these two percentages, the violin plot exhibits only a small width, which permits the conclusion that fewer responses lie within this range. If no responses exist within a defined value range, the associated area of the violin plot reduces to zero and appears as a line (see, for example, quantity increase in reinforcement works – public clients – near the 130% mark).

Figure 1 uses circles/dots for outliers and stars for extreme values. The number of responses cannot be derived directly from the violin plot, which is why descriptive data analyses are listed in Table 1. The @Risk software suite was used to perform data fitting on the basis of expert responses for public and private clients (see Figure 2) to enable realistic recommendations for potential distribution functions that would appropriately represent quantities in reinforcement works. A LogLogistic distribution for public clients and a Logistic distribution for private clients reflect the expert statements very accurately, as confirmed by a Kolmogorov-Smirnov test.

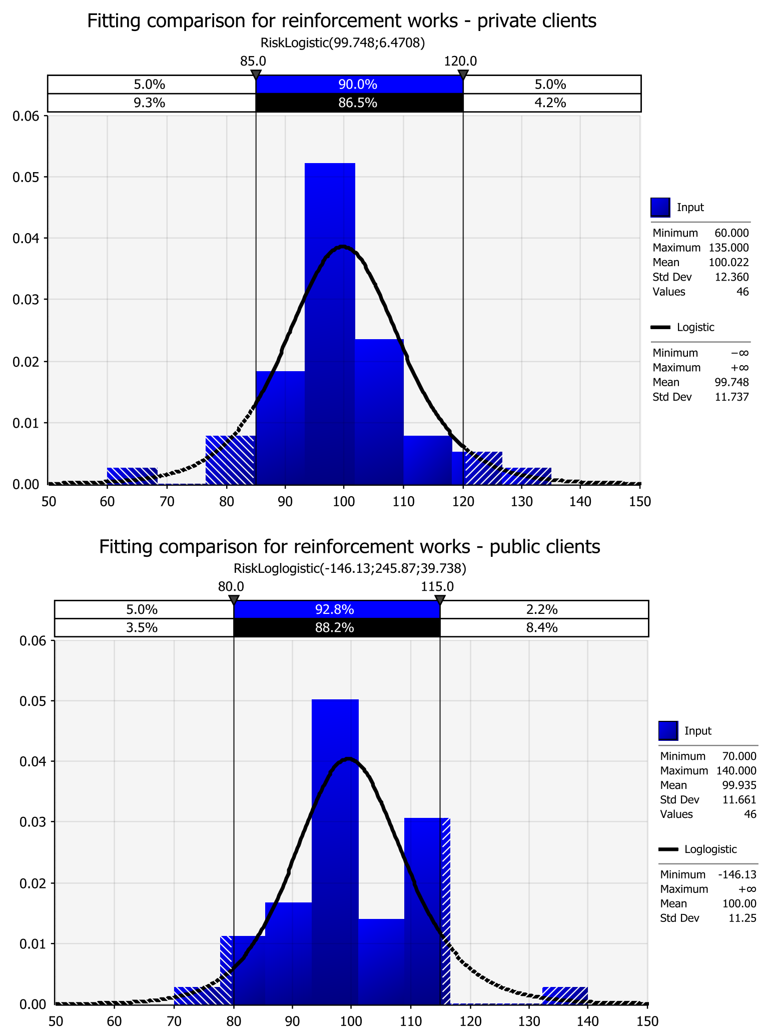

Figure 2. Data fitting – quantity deviations in reinforcement works.

5 DISCUSSION AND CONCLUSIONS

From the bidder's/contractor's point of view, the above-mentioned quantity uncertainties constitute a risk (for instance with respect to the contribution margin), but also a chance (such as in unit-priced contracts in the case of quantity increases at a high unit price) and should thus be reflected by distribution functions of relevant quantities in the probabilistic costing exercise. Considering quantity uncertainties is also of interest to clients in order to investigate, during the phase of selecting the best bidder, how prices quoted by bidders would move when specified quantities change in the case of unit-priced contracts. Overall, the results of the expert survey show that, quantity uncertainties in reinforcement works are considered to be slightly greater for private clients than those assumed for public clients. A LogLogistic distribution for public clients and a Logistic distribution for private clients reflect the best approximation. However, the open ends of these theoretical types of distribution should be viewed critically. If probabilistic calculations are performed using such distributions (with no upper and/or lower limit), this may lead to exceedingly implausible individual results within the simulation. It is thus useful to combine the findings of the survey and the results of mathematical/statistical data fitting with theoretical considerations, particularly with respect to possible limits to individual parameters. This broader analysis includes both data fitting and theoretical considerations and can be used to establish a polygonal representation of the distribution's shape and to apply limit values, thus making it possible to retain the shape of the distribution taken from the expert survey and to adjust the value range to specific project conditions.

Quantity uncertainties will not only affect costing and pricing. Even more importantly, they will also have an influence on construction process planning and logistics as well as on the scheduling of resources. Knowledge of potential/expected ranges thus helps project stakeholders to include outliers and extremes in their project considerations and planning exercises and to develop a variety of solutions for certain scenarios from the outset. Generally speaking, knowing the possible ranges of specified quantities contributes to achieving greater transparency in process planning and project organization and/or management.

References

AbouRizk, S., and Halpin, D., Statistical Properties of Construction Duration Data, *Journal of Construction Engineering and Management,* 118(3), 525-544, 1992.

Gläser, J., and Laudel, G., *Experteninterviews und Qualitative Inhaltsanalyse als Instrumente Rekonstruierender Untersuchungen,* Springer Fachmedien GmbH, Wiesbaden, 2010.

Goulet, J., and Smith, F., *Extended Uniform Distribution Accounting for Uncertainty of Uncertainty,* Proceedings of the First International Conference on Vulnerability and Risk Analysis and Management (ICVRAM 2011), Hyattsville, Maryland, ASCE, April 11-13, 2011.

Hofstadler, C., and Kummer, M., *Chancen- und Risikomanagement in der Bauwirtschaft – Für Auftraggeber und Auftragnehmer in Projektmanagement, Baubetrieb und Bauwirtschaft,* Springer-Verlag, Berlin, Heidelberg, 2017.

Kummer, M., Aggregierte Berücksichtigung von Produktivitätsverlusten bei der Ermittlung von Baukosten und Bauzeiten – deterministische und probabilistische Betrachtungen, Doctoral Thesis, Graz University of Technology, Graz, 2015.

Mayer, F., *Kostensicherheit zum Zeitpunkt der Realisierungsentscheidung – Entwicklung eines Kosten-Prognose-Modells für Bauwerkskosten im Hochbau,* Ph.D. Thesis, Lehrstuhl für Bauprozessmanagement und Immobilienentwicklung der Technischen Universität, München, 2013.

Raaber, N., *Beitrag zu Ausschreibung und Vergabe von Bauleistungen. Schriftenreihe des Instituts für Baubetrieb und Bauwirtschaft der Technischen Universität Graz,* Verlag der TU Graz, Graz, 2003.

Sander, P., *Probabilistische Risiko-Analyse für Bauprojekte – Entwicklung eines branchenorientierten softwaregestützten Risiko-Analyse-Systems*, Ph.D. Thesis, Leopold-Franzens-Universität Innsbruck, 2012.