Icelandic Journal of Engineering
Verktækni - Tímarit Verkfræðingafélags Íslands

https://doi.org/10.33112/ije.30

Use of Z-fuzzy numbers in the management of megaprojects

Dorota Kuchta
Wroclaw University of Science and Technology, Faculty of Management, Wybrzeże Wyspiańskiego 27, 50-370 Wroclaw, Poland.

Fyrirspurnir/Correspondence:
Dorota.kuchta@pwr.edu.pl

Greinin barst 1. febrúar 2023.
Samþykkt til birtingar 3. júlí 2024.
Birt á vef 9. júlí 2024

Abstract: The paper discusses the problem of the lack of credibility of estimates used for decision making in megaproject management. We present the high importance of this issue and discuss factors that negatively influence megaproject estimates. We propose a method to be used for the estimation of basic parameters of megaprojects based on Z-fuzzy numbers. Fuzzy numbers allow modelling the lack of full knowledge and the changeability, which are omnipresent in megaprojects. Z-fuzzy numbers allow for the adjustment of estimates based on the opinion about their authors' credibility among project stakeholders. As a result, the estimates will be much more realistic. Fuzzy numbers and Z-fuzzy numbers are described, and the estimation method is presented. Simple examples accompany the description, and the method proposed is illustrated using a real-world example. Limitations of the method are listed in the conclusions.

Keywords: *Megaproject, estimation reliability, estimation credibility, fuzzy number, Z-number.*

1. Introduction

Megaprojects are defined (Flyvbjerg, 2014) as "large-scale, complex ventures that typically cost a billion dollars or more, take many years to develop and build, involve public and private stakeholders, are transformational, and impact millions of people." Everybody is familiar with examples of megaprojects: high-speed rail lines, airports, seaports... There are several famous examples, like the Channel Tunnel or the Sydney Opera House. The market of megaprojects has constantly been growing, and the megaprojects themselves have been growing, too, up to a per project budget of more than 1 000 billion dollars (Flyvbjerg, 2014). New megaprojects are constantly initialized on all the continents, which might be good news: the megaprojects create positive value added for millions of people, facilitating mobility (aircraft, road, and rail projects), increasing well-being (sports, culture, urban projects), helping to return to normality after disasters (disaster cleanup projects), improving living conditions (water-related projects), increasing our knowledge, technological capacities, and life-saving possibilities (research projects), etc.

Certainly, the value that megaprojects have brought to our societies and communities cannot be overestimated. It is so because they usually come into being due to visionary ideas, enthusiastic views, and courageous decisions, and they are motivated by the real needs of both poor and wealthy societies. The problem is that for a project to be seen as successful, the value delivered by it should exceed its cost (Kerzner, 2013). And the performance of megaprojects has been very bad and has not improved (Flyvbjerg, 2008, 2014). Cost overruns of more than 100% or even by 1900 % (Suez Canal), 1600% (Scottish Parliament Building), or 1400 % (Sydney Opera House) are not uncommon, and delays are equally huge. There are many equally serious problems with quality and scope, as well as with the profitability of the project products (Flyvbjerg, Bruzelius, & Rothengatter, 2003). In 2010, about 70% of industrial megaprojects in the USA failed to meet their goals (Merrow, 2011). Therefore, the question about the actual value of these projects must be asked and ways to improve their performance should be searched.

D. Hulett (Hulett, 2011), who has been the manager of many projects, points to a very important problem linked to project planning (he talks mainly in terms of time and cost). He states that practically always projects are planned using schedules and budgets with fixed items (activities, cost objects), fixed numbers (durations, costs), fixed relationships (e.g., among activities), etc. Everyone seems to accept it, although everybody knows that, in reality, those allegedly fixed items will probably turn out to be unstable and completely different. Later, people complain about huge time and cost overruns in projects as if they were a surprise, although, according to D. Hulett, it should have been clear to all the parties involved that they would occur. There is simply so much inherent risk and uncertainty in projects and their environment (Cleden, 2017) that thinking in terms of fixed data in the phase of project planning is close to nonsense. Nonsense that most people seem to accept. Such an approach practically means no proper planning at all, because projects very soon spiral out of control and start living their own lives, totally detached from the plans. In (Hulett, 2011), but also in other literature items (Salkeld, 2013; Schuyler, 2001) we can find an answer to how projects should be planned; probabilistic modelling should be used, both while modelling numerical values, but also while modelling lists of project tasks, project networks, etc.

Megaprojects are not simply larger copies of other projects (Flyvbjerg, 2014). They require an almost completely different approach because their degree of complexity and uncertainty cannot be compared with that of other projects. Lots of unexpected and impossible-to-grasp factors (e.g., global factors like world politics and human factors like the behaviour of larger communities) influence megaprojects implementation (Flyvbjerg et al., 2003). Thus, planning megaprojects in terms of fixed values and generally fixed data is even much more nonsense than in the case of other projects. "The world of megaprojects planning and implementation is highly stochastic and rarely turns out as originally intended" (Flyvbjerg et al., 2003). This truth remains somehow unabsorbed by the megaproject world. Apart from fixed values, fixed task lists, and fixed relationships, in megaproject scheduling and planning fixed price contracts are generally used, which contradicts the very nature of megaprojects (Ajam, 2020). Research shows that governments, investors, and contractors make decisions about whether or not to start a megaproject based on fixed, biased numbers that cannot be trusted. They do not and cannot see the biases (that will be discussed later in the paper) because the numbers presented to them are fixed: a fixed, crisp number cannot show that, and how large deviations are possible. It is true that sensitivity analysis is used, but usually only for deviations under 20% (Flyvbjerg et al., 2003). As mentioned above, the actual deviations are much higher.

Introducing probabilistic distributions instead of fixed values into megaproject planning will not solve all the problems but may help in more realistic planning (and earlier, a more realistic selection) of such projects, as it does for smaller projects (Salkeld, 2013). However, practically no literature items exist on the application of probability theory to megaprojects. This may be due to the fact that statistical methods need reliable databases that do not exist because of the extreme uniqueness of megaprojects. Surveys (e.g., on passengers' preferences) are hardly helpful too because actual people's behaviour may and often does differ strongly from the survey declarations (Flyvbjerg et al., 2003).

There exists an alternative representation of not fully known numbers or other data that requires less input information and is generally less rigid and easier to use than probabilistic distributions: fuzzy numbers. There are pros and contras for both approaches to uncertainty or lack of knowledge modelling: the probabilistic and the fuzzy approach (Cheeseman, 1986). Generally, the fuzzy approach is seen as less formal and better suited to express information that has to be generated without an adequate history and data sets, using intuition and expertise instead. It may be no coincidence that, contrary to the probabilistic approach, the fuzzy approach has already been applied to megaprojects according to the scientific literature (Mojtahedi, Mousavi, & Aminian, 2008; Xie, Xu, Ju, & Xia, 2022; Zhang, Wei, Zhao, Han, & Chen, 2020).

For this reason, the subject of the present paper is the application of the fuzzy approach to megaproject planning. Fuzzy numbers (L. A. Zadeh, 1978) permit the representation of possible deviations and possible values of a not fully known magnitude (e.g. project cost), together with respective possibility degrees. The possibility degrees are similar to probabilities, but fuzzy numbers do not need to fulfil such strict conditions and are easier to use and interpret by non-mathematicians than random variables. We propose thus to use fuzzy numbers as estimations of costs and durations (and possibly other data, too) in megaprojects, instead of fixed, crisp numbers. The information on which crucial decisions regarding megaprojects will be based will then be more encompassing and thus more informative and objective. Methods of interviewing experts in order to obtain fuzzy (instead of fixed) values are described in the literature, e.g., in (Cheng, 2004).

Apart from the principal form of fuzzy numbers (L. A. Zadeh, 1978), several other forms have been developed over time (Bustince et al., 2016). Some of them are still better suited to express a lack of knowledge, uncertainty, hesitation, intuition, etc. Examples are intuitionistic fuzzy sets (Atanassov, 1986), hesitant (Torra, 2010), or type-2 (Mendel & John, 2002) fuzzy sets. The latter have been applied to megaprojects (Eshghi, Mousavi, & Mohagheghi, 2019). Here, however, we will propose, for the first time in the literature, the application to megaprojects of so-called Z-numbers or Z-fuzzy numbers (Lotfi A. Zadeh, 2011). This variant of fuzzy numbers has stemmed from the idea that a fuzzy estimation, even if it is not fixed, is still subject to biases, manipulations, or even falsifications, the diseases that plague megaprojects (Ajam, 2020; Flyvbjerg et al., 2003; Merrow, 2011). Z-numbers are composed of two parts: a fuzzy number representing possible values of a magnitude and its credibility degree. The degree of credibility allows us to adjust the fuzzy estimation and make it more realistic. The ultimate objective of this paper is thus to propose the application of Z-fuzzy numbers to megaproject management and, more exactly, to megaproject estimation.

The structure of the paper is as follows: In Section 2, we will present basic information about fuzzy numbers and Z-fuzzy numbers. In Section 3, we will present, based on existing research on megaprojects, the main reasons for incomplete (or even very small) credibility of estimates given by experts who are consulted in the stages before a megaproject is accepted for implementation and actually started. In Section 4, we present the method for using Z-numbers in the planning stage of megaprojects, and an example. The paper ends with some conclusions.

2. Fuzzy numbers and Z-numbers: Basic Information

We will consider here the simplest form of fuzzy numbers, the so-called triangular fuzzy numbers.

Definition 1: $\tilde{A} = (a_1, a_2, a_3)$ is a triangular fuzzy number if a triple (a_1, a_2, a_3) of real numbers exist such that $a_1 \leq a_2 \leq a_3$ and

$$\mu_a(x) = \begin{cases} 0 \text{ for } x \leq a_1 \\ \frac{x-a_1}{a_2-a_1} \text{ for } x\epsilon[a_1, a_2] \\ \frac{a_3-x}{a_3-a_2} \text{ for } x\epsilon[a_2, a_3] \\ 0 \text{ for } x \geq a_3 \end{cases} \quad (1)$$

μ_a in (1) is called the membership function of fuzzy number $\tilde{A} = (a_1, a_2, a_3)$. A positive triangular fuzzy number is one fulfilling the condition $a_1 > 0$.

Figure 1 shows the shape of a positive triangular fuzzy number membership function, which explains the name of this fuzzy number type:

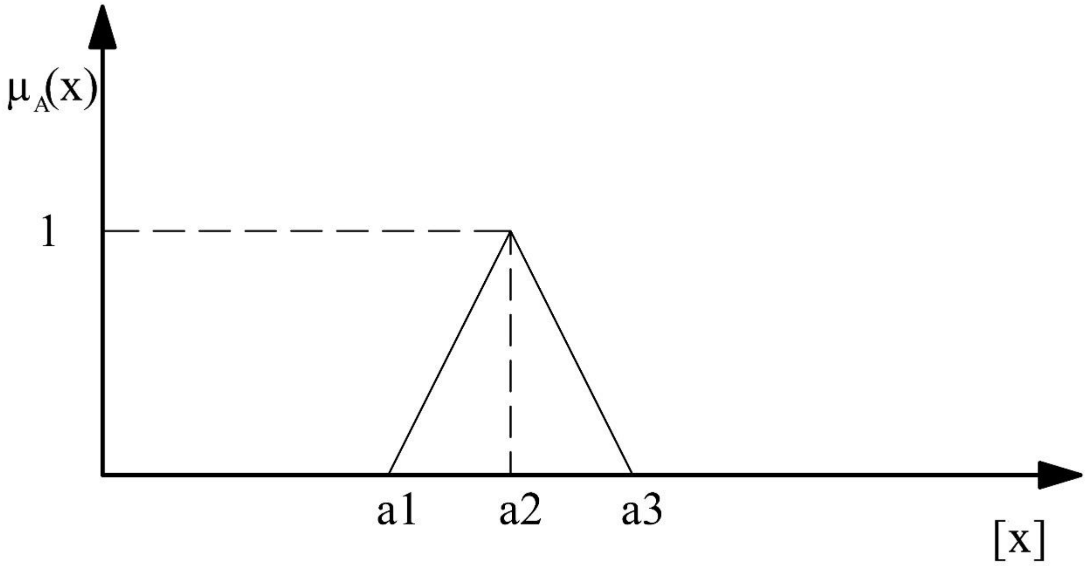

Figure 1: *The membership function of a positive fuzzy triangular number.*

Values smaller than a_1 and greater than a_3 are considered by the expert as impossible as actual values of the magnitude in consideration (e.g., the cost of an item or the duration of an activity); the occurrence of the other values in this role is seen as possible, although to various degrees. The most possible value of the item in question is, according to the expert, a_2. The possibility degree of a number *x* being the magnitude in question decreases with the distance of *x* from a_2.

A triangular fuzzy number $\tilde{A} = (a_1, a_2, a_3)$ represents the opinion of an expert on the possible values of a magnitude. Other types of fuzzy numbers represent situations when the membership function has a different shape or is not continuous, but here we limit ourselves to the simplest case, which is sufficient in most applications.

Arithmetical operations on fuzzy numbers and relations among them and crisp numbers can be defined in different ways (see, e.g., (Lotfi A. Zadeh, 2011)) similarly as in the following definition:

Definition 2: Let us assume we have $\tilde{A} = (a_1, a_2, a_3)$ and $\tilde{B} = (b_1, b_2, b_3)$.

- Number $\tilde{A} + \tilde{B}$ is defined as the fuzzy number $(a_1 + b_1, a_2 + b_2, a_3 + b_3)$
- Number $s\tilde{A}$, for non-negative crisp numbers *s*, is defined as the triangular fuzzy number (sa_1, sa_2, sa_3)
- Number $s - \tilde{A}$, for crisp numbers *s*, is defined as the triangular fuzzy number $(s - a_3, s - a_2, s - a_1)$
- $\tilde{A} = (a_1, a_2, a_3) \leq (\geq) s$, where s is a crisp number, if $a_3 \leq s$ $(a_1 \geq s)$.

Fuzzy numbers can also be ranked in different ways, and the order among them is much less unambiguous than in the case of crisp numbers (Brunelli and Mezei, 2013).

The information conveyed by $\tilde{A} = (a_1, a_2, a_3)$ is more informative and more complete than that given by a crisp number alone. But it can still be strongly biased. That is why Z-numbers have been introduced. They are defined as follows (Lotfi A. Zadeh, 2011):

Definition 3: Z-fuzzy number $\hat{A} = (\tilde{A}, \tilde{Z})$ is a couple of fuzzy numbers $\tilde{A} = (a_1, a_2, a_3)$ and $\tilde{Z} = (z_1, z_2, z_3)$ such that $\tilde{Z}$ is a positive fuzzy number representing the credibility degree of estimation $\tilde{A}$.

Fuzzy numbers $\tilde{Z}$ are used to adjust the estimate expressed by $\tilde{A}$, taking into account the features of the authors of the estimation and the circumstances of the estimation process. Often these numbers are selected from a pre-set 'dictionary', where they are linked to linguistic descriptions of credibility. For example, in (Hendiani, Bagherpour, Mahmoudi, & Liao, 2020) the following dictionary is used: unlikely (0.1, 0.2, 0.3), fairly impossible (0.3, 0.4, 0.5), maybe (0.4, 0.5, 0.6), fairly possible (0.5, 0.6, 0.7), likely (0.7, 0.8, 0.9), most likely (0.8, 0.9, 1), certainly (1, 1, 1). 'Unlikely' means that the estimate given by $\tilde{A}$ should be heavily adjusted, 'certainly' stands for no adjustment need. The formulae used for the adjustment should take into account the nature of adjustment, and the bias type. If the estimation process concerns a magnitude that preferably should be large, and if are dealing with the most common bias type, the optimism bias (Prater, Kirytopoulos, & Tony, 2017), we are probably facing the situation where the bias has led the expert to increase the estimate with respect to the objective situation, and the estimate has to be reduced in order to be realistic. In fact, in project management, especially in megaproject management (Flyvbjerg, 2014), this type of bias frequently occurs with respect to revenue and profit estimation. Regarding cost and time, the optimism bias makes the expert decrease the estimates. In some cases, the bias will be "hesitant", and the method of dealing with Z-fuzzy numbers must fit the context.

Z-fuzzy number $\hat{A} = (\tilde{A}, \tilde{Z})$ is ultimately transformed into a single fuzzy number.

$$\tilde{A}^* = (a_1^*, a_2^*, a_3^*) = (Z_1(a_1), Z_2(a_2), Z_3(a_3)) \quad (2)$$

which is a corrected (adjusted) estimate, more credible than the original one. In the literature there are several approaches to this adjustment (Azadeh & Kokabi, 2016; Kang, Wei, Li, & Deng, 2012; Kuchta, Marchwicka, & Schneider, 2021; Marchwicka & Kuchta, 2022). Basically, they differ in the choice of functions Z_1, Z_2, Z_3. These functions are selected on the basis of $\tilde{Z}$, and the context (nature of the magnitude being estimated, bias type). The simplest form of those functions, assumed here for the sake of simplicity, can be the following.

$$Z_i(x) = z_i x,\ i = 1,2,3 \quad (3)$$

Example 1: Let us suppose that the expert, asked for cost estimate of a task of a megaproject, gave the answer (in US dollars): $\tilde{A}$ =(0,4 billion, 0,6 billion, 0,8 billion). It is known that the expert is strongly optimistic. The following dictionary for $\tilde{Z}$ is used: certainly (1, 1, 1), most likely (1.1, 1.2, 1.3), likely (1.3, 1.4, 1.5), fairly possible (1.4, 1.5, 1.6), maybe (1.5, 1.6, 1.7), fairly impossible (1.7, 1.8, 1.9), unlikely (1.8, 1.9, 2). The term "fairly impossible" has been selected as the one that best suits the expert. Thus, the corresponding Z-number would be ((0,4 billion, 0,6 billion, 0,8 billion), (1.7, 1.8, 1.9)), and the final estimate of project cost, calculated using (3), would be (0,68 billion, 1,08 billion, 1,52 billion). The most possible value of the project cost would thus be 1,08 billion. The cost may vary and decrease to 0,68 billion and increase to 1,52 billion, with decreasing possibility degrees, as shown in Fig.1. It is important to note that, because of the high optimism of the expert, the value seen by them as most possible (0,6 billion), after adjustment turns out to be impossible (it is smaller than 0,68 billion, the lower limit of the values seen, after adjustment, as possible).

3. Factors determining the credibility of estimations in megaprojects

The credibility estimates of time, cost, and projected revenues, and other input information for project planning is a serious problem in all projects, not only megaprojects. In the literature, we read about various types of human estimators (optimists, pessimists, hesitant (Lock, 2013)), about the inherent incurable optimism of young project teams, ambitions, various types of pressure, and other external factors that substantially deteriorate estimation quality (Hulett, 2011). The type of biases is varied and difficult to identify: many researchers underline the optimism bias (Prater et al., 2017; Wang, Zhuang, Yang, & Sheng, 2014), but the well-known critical chain method (Goldratt , 1997) was born from the finding that often time estimations are too pessimistic. The primordial role plays, of course, also the mere fact of not knowing, the uncertainty, and risk.

For megaprojects, however, researchers have identified specific factors that deteriorate the credibility of the estimation to even greater degree. The following list of factors has been prepared on the basis of (Ajam, 2020; Flyvbjerg, 2014; Flyvbjerg et al., 2003; Merrow, 2011; Romestant, 2019; Xie et al., 2022). The list is by no means complete, in fact, it is merely the result of a literature survey, in which only those projects were taken into account that happened to be described in the scientific papers available in the scientific bases. Nevertheless, it can serve as a basis or example for an internal list of each organisation that would help to implement the method proposed in the next section. The list here has the form often used in checklists used in projects: It is composed of questions that can be answered with yes, no, or not applicable. The answer 'yes' indicates a negative situation, thus the presence of a problem. The list should be applied in the final stage of the planning phase of the project in question.

- PF1: Have data from the past, from other projects, been used without adapting them to the project specificity?
- PF2: Have survey results been used, without taking into account the fact that actual behaviour of people is often different from that stated in the survey?
- PF3: Do you see potential factors that may have an influence on the project but have not been taken into account (e.g., in railway projects: information, access to the stations, commercial supply, housing policy)?
- PF4: Do you judge it possible that problems in the implementation of other projects run in the organisation may influence the project in question, and this fact was not taken into account in the planning process?
- PF5: Do you think it is possible that political wishes or political pressure on the part of forecasters or consultants have influenced the planning results?
- PF6: Do you think that the interest in having the project started on the part of some consultants or experts has influenced the planning results?
- PF7: Do you think that the geological risk has been underestimated?

PF8: Do you consider it possible that safety and environmental demands have been underestimated?
PF9: Have there been significantly different opinions on the project estimates?
PF10: Have some important groups of experts not been consulted or ignored (e.g., the promoters of German high-speed rail gave the ministry very optimistic figures, which later turned out far too optimistic, while the ministry's scientific advisory council provided two negative reports);
PF11: Have regional and environmental groups been ignored?
PF12: Has environmental assessment and risk analysis been used only formally, merely as a necessary step in preparing the documents for the government?
PF13: Has the sensitivity analysis, if used at all, been performed considering only small range variation of the data, such as ±10% or ±20% percent?
PF14: Do you judge that the accountability for various aspects of the project has not been properly or fully assigned?
PF15: Can we say that in the project one entity plays two or more contradictory roles (e.g., a government is a promoter of a project which will bring a high profit to several businesses, and at the same time presents itself as a guardian of the interests of the whole local community)?
PF16: Has project managers not been sufficiently trained in project management?
PF17: Is there no established project management function in the organisation responsible for project implementation?
PF18: Is the project management function of the project in question being outsourced?
PF19: Are fixed-price contracts going to be used?
PF20: Are outdated procurement practices (such as culture of low bid) going to be used?
PF21: Do you judge that some of project promoters, owners, consultants, executives, and staff possess to a significant degree such features as arrogance, overconfidence, political involvement, etc.?
PF22: Has the distribution of the value the project should deliver among the different project stakeholders not been carried out fully and transparently?

Researchers investigating megaprojects have also identified the presence of additional factors that strengthen or weaken the effect of the above factors. Here, the results are rudimentary, but indicate that research in this direction is necessary. These factors are, e.g.:

AF1: Specific area of the project: Forecasts on the number of users are more biased in rail projects than in road projects;
AF2: Home country of the consultants: in the case of the high-speed rail link between Melbourne and Sydney, the forecast of the number of passengers stemming from Japanese consultants was much higher than those made by French consultants, which in turn substantially exceeded the forecast given by US consultants;
AF3: The underestimate of the cost of megaprojects appears to be more pronounced in developing countries than in Europe and North America;

AF4: Most mistakes are made by senior business managers in sponsoring firms.

The basic sense of the list of additional factors (of which AF1-AF4 are merely examples), is to make the decision maker carefully consider the specific situation, the specific stakeholders involved in the estimation process, and the specific circumstances of the very moment. This list will be used in the algorithm proposed further on to take into account the features of a specific estimate provider in their specific, present situation, and context. The list should aim to represent, and be based on, the current and general knowledge and expertise available at the moment. It will thus be a very dynamic and individual list.

One of the principles of good project management is the learning process. A substantial element of the learning process is learning from mistakes, which involves keeping track of factors that in past projects have contributed to problems. Each organisation or group that is a stakeholder of megaprojects should continue the type of research presented here, but with respect to the projects in which they have played a role: whether it was decision-making, estimating, forecasting, promoting, executing, using the final product, or simply being somehow influenced by the project. Each organisation or group should have a list of principal factors that deteriorate (such as the list PF1-PF21) or improve (here, we did not give examples for such factors) the reliability of forecasts or estimates. Additionally, information about reliability improving or deteriorating factors (such as AF1-AF5) that refine the main factors should be gathered. Both lists could be used in the estimation method proposed in the next section.

4. Estimation and forecasting in megaproject – fuzzy approach

Here, we propose to introduce the following procedure to the process of estimating and forecasting key values in megaprojects, on the basis of which decisions about implementing and pursuing megaprojects are taken (cost of the whole project and of basic project items, duration of the whole project, and of its principal phases, revenues and profits from the usage of the project product, etc.). The basic idea is to take into account the estimation of several stakeholders (all who feel or are seen as qualified to give an estimate of the magnitude in question), to let them use three-point estimates instead of expecting crisp numbers (as crisp numbers in the case of megaprojects rarely make sense), and to let them evaluate the credibility of other stakeholder estimates. The final estimate would then be the results of the opinion of all the stakeholders.

We assume low estimate quality to results from optimist bias, which seems to occur most often in practice (Flyvbjerg, 2014). The method has to be modified for other bias types. We consider here a specific, selected magnitude to be estimated for the project in question that preferably should be (relatively) small (e.g., the duration or cost of an activity). For other types of magnitudes, the method would have to be modified too.

The steps to be taken are as follows:

A. All the principal stakeholders of the project (groups, organisations, communities) important for the magnitude in question are identified (see, e.g. (Manzoor, 2019) for adequate methods); Let us suppose that their number is M, the *j*-th stakeholder will be assigned a positive weight $v_j, j = 1, \dots, M$, where $\sum_{j=1}^{M} v_j = 1$;

B. All stakeholders who feel or are judged as qualified (e.g., the local community may not feel qualified to estimate the construction or electricity cost) should be listed in the first N positions ($N \leq M$). They should give the estimate of the magnitude in question in the form of a triangular fuzzy number $\tilde{A}_j, j = 1, \dots, N$, where N stands for the number of stakeholders who feel or are seen as qualified. The fuzzy estimations are based on the most possible value, the pessimistic one and the optimistic one. The three values should be justified by the respective stakeholder (see (Hulett, 2011) for details);

C. At this stage, all stakeholders can express their opinion on the credibility of the estimations of the other stakeholders. All stakeholders make use of their own lists of factors that may deteriorate the reliability of the estimate or influence this quality in another way. Therefore, they would have their list of all principal factors that, in their opinion, can deteriorate the quality of the estimate of the magnitude in question (composed of elements $PF_j^l, j = 1, \dots, M, l = 1, \dots, L_j$). The list should cover the whole spectrum of possible deteriorating situations. Thus, the worst case should be considered here: all the factors that may ever play a negative role should be included. The stakeholders should determine for each element $PF_j^l, j = 1, \dots, M, l = 1, \dots, L_j$, a positive weighting factor WF_j^l, such that $\sum_{l=1}^{L_j} WF_j^l$=1, expressing the relative maximal possible weight of the influence of PF_j^l. They should also have at their disposal their lists of additional factors that weaken the effect of the principal factors, thus, improve the estimation quality in specific cases, for specific estimates providers (these lists should be composed of elements denoted as $IF_j^k, j = 1, \dots, M, k = 1, \dots, K_j$).

D. For each $\tilde{A}_i, i = 1, \dots, N$, the *j*-th stakeholder determines the credibility degree $\tilde{C}_{ij}, i = 1, \dots, N, j = 1, \dots, M$ of the *i*-th stakeholder in the given context (for the magnitude being estimated). The degree of credibility can be defined in the following way:

   a. A dictionary of fuzzy terms should be chosen, expressing the effect of the accompanying list of additional factors and other knowledge on the principal factors, e.g. $\tilde{E}_1 = (0,0,0)$ – for "none", $\tilde{E}_2 = (0.1,0.2,0.3)$ – for "small", $\tilde{E}_3 = (0.3,0.5,0.7)$ – for "average", $\tilde{E}_4 = (0.7,0.8,0.9)$ – for "big", $\tilde{E}_5 = (1,1,1)$ – for "total". The elements $\tilde{E}_t, t = 1,2, \dots, T$ of the dictionary can be chosen arbitrarily, while two conditions should be satisfied: $\tilde{E}_1 \geq 0, \tilde{E}_T \leq T$. Additionally, the elements $\tilde{E}_t, t = 1,2, \dots, T$ should reflect various degrees of influence of additional factors on the principal factors and preferably should be assigned linguistic terms, such as in the above example.

   b. For each factor $PF_j^l, j = 1, \dots, M, l = 1, \dots, L_j$, the *j*-th stakeholder should consider the *i*-th estimate provider (*i*=1,…*N*), and assess the influence of the additional factors (and possibly facts resulting from other knowledge) on the weighting factor WF_j^l for the *i*-th estimate provider. It should be determined in the form of an element selected from among the terms $\tilde{E}_t, t = 1,2, \ldots, T$. The selected element will be denoted as $\tilde{E}_{ij}^{l}, j = 1, \ldots, M,, i = 1, \ldots, N, l = 1, \ldots, L_j$. A high value of $\tilde{E}_{ij}^{l}$ means that, in the opinion of the *j*-th stakeholder, the negative influence of the factor PF_j^l is strongly weakened in the case of the *i*-th stakeholder in the role of the estimation provider, a low value would mean that the factor PF_j^l has a high negative influence on the estimate provided by the *i*-th stakeholder. Often we will have $\tilde{E}_{ii}^{l}$ close to 1: here the estimation provider and the stakeholder assessing their credibility are identical. Also, we can assume that $\tilde{E}_{ij}^{l}$ will be close to 1 if the *j*-th stakeholder does not have enough knowledge with respect to the specific factor and the *i*-th stakeholder and does not want to underestimate the quality of the *i*-th stakeholder in the role of estimation provider.

c. $\tilde{C}_{ij}, i = 1, \ldots, N, j = 1, \ldots, M$ should be determined as:

$$\tilde{C}_{ij} = \sum_{l=1}^{L_j} WF_j^l \left(1 - \tilde{E}_{ij}^{l}\right) \quad (4)$$

From the above assumptions and Definition 1, it follows that $\tilde{C}_{ij}$ is a triangular fuzzy number, $0 \leq \tilde{C}_{ij} \leq 1$. The higher its value, the lower the credibility of stakeholder *i* as the estimation provider in the eyes of the *j*-th stakeholder. In the special case where *i* is equal to *j* (a stakeholder assesses their own credibility), $\tilde{C}_{ii}$ would be often close to zero.

E. We determine Z-fuzzy numbers

$$\hat{A}_{ij} = \left(\tilde{A}_i, 1 + \tilde{C}_{ij}\right), i = 1, \ldots, N \quad (5)$$

and we calculate

$$\tilde{A}_{ij}^{*} = \left(a_{1,ij}^{*}, a_{2,ij}^{*}, a_{3,ij}^{*}\right), i = 1, \ldots, N, j = 1, \ldots M \quad (6)$$

according to (3). Numbers $\tilde{A}_{ij}^{*}$ in (6) represent fuzzy estimations of the magnitude in question stemming from stakeholder *i*, adjusted according to the credibility that the *j*-th stakeholder assigns to stakeholder *i*.

F. Decision makers and all principal stakeholders will be presented with the following pieces of information.

- *i*-th stakeholder, $i = 1, \ldots, N$ with the weight v_i,
- original estimation $\tilde{A}_i, i = 1, \ldots, N$
- adjusted estimations $\tilde{A}_{ij}^{*}$, taking into account the opinions of all the stakeholders and their weights
- an aggregation $\tilde{A}_i^{*}$ of the adjusted estimations for the *i*-th stakeholder, calculated. e.g., as the triangular fuzzy number.

$$\tilde{A}_i^{*} = \sum_{j=1}^{M} v_j \tilde{A}_{ij}^{*}, i = 1, \ldots, N \quad (7)$$

$\tilde{A}_i^{*}$ represents the estimation of the magnitude in question given by the *i*-th stakeholder, corrected by the other stakeholders. The influence of the opinion of the other stakeholders is tuned by their respective weights;

- an envelope $\widetilde{AE}$ of the adjusted estimations calculated. e.g., as the triangular fuzzy number

$$\widetilde{AE} = \left(\min_{i=1,\ldots N} a^*_{1,ij}, \sum_{i=1}^{N} v_j a^*_{2,ij}, \max_{i=1,\ldots N} a^*_{3,ij}\right) \quad (8)$$

The envelope $\widetilde{AE}$ aggregates all the adjusted estimations of all the stakeholders that have participated in the estimations. Its lowest parameter represents the optimistic estimate (the lowest one that occurs after the adjustment), the highest parameter, the pessimistic one, and its middle parameter is the weighted average of all the middle parameters from the estimations of all the estimators. The envelope permits to gain an insight on the spectrum of the opinions of all the experts on the magnitude in question and can serve as a basis for project planning.

Example 2: Let us consider Berlin Brandenburg Airport (BER) in Germany (Sedlin et al., 2020). Originally planned to open in 2011, the airport faced numerous delays and did not open until October 2020, nearly a decade later than planned. The cost overruns were substantial, with initial estimates of €2.83 billion finally reaching over €7 billion. The factors that contributed to this discrepancy were technical, legal, and managerial; some of them were related to the relationships with the local community. The details of the estimation process cannot be reconstructed, but it is known that there were constant changes in the project caused by the above groups of factors. If the initial estimate of €2.83 billion had been subject to the evaluation of the respective credibility from the different stakeholders: technical experts, layers, and experienced project managers, as well as the local community representatives, it is likely that it would have been adjusted, represented as interval or a three-point estimate, and the risk of exceeding it would have been more visible. Let us hypothetically simulate the possible application of the method to the project to the estimation of the total project cost.

Let us assume that we have 5 principal stakeholders (*M*=5), with weights being assigned by the project investor (Step A):

Table 1: *Stakeholder groups in the example.*

Stakeholder number (*j*)	**Stakeholder name**	**Stakeholder weight (v_j)**
1	Project manager	0.6
2	Technical experts	0.1
3	Lawyers	0.1
4	Experienced project managers (of other projects)	0.1
5	Local community	0.1

Exclusively the project manager was responsible for the estimation process (N=1). They provided the fuzzy estimation $\tilde{A}_1 = (2.8, 2.83, 2.85)$ (in billion €) (Step B).

In Step C all the stakeholders consider their lists of principal factors $PF_j^l, j = 1, \dots, M, l = 1, \dots, L_j$), and of additional factors $IF_j^k, j = 1, \dots, M, k = 1, \dots, K_j$. For the sake of simplicity, let us assume here that the lists of factors $PF_j^l, j = 1, \dots, M, l = 1, \dots, L_j$ are identical for all the stakeholders ($PF_j^l = PF^l, j = 1, \dots, M, l = 1, \dots, L$), the weights WF^l, $l = 1, \dots, L$ included, and have the form given in Table 2. This is of course an oversimplification, but as the actual data is unknown, this simplified version should suffice to illustrate the algorithm.

Table 2: *List of principal factors that potentially deteriorate the estimation quality in the example.*

Factor number *l*	Factor PF^l and its notation in Section ..	Weight WF^l
1	Do you think it is possible that political wishes or political pressure on the part of forecasters or consultants have influenced the planning results? (PF5)	0.4
2	Have some important groups of experts not been consulted or ignored? (PF10)	0.1
3	Have regional and environmental groups been ignored in the planning process? (PF11)	0.1
4	Has project managers not been sufficiently trained in project management? (PF16)	0.2
5	Do you judge that some of project promoters, owners, consultants, executives, and staff possess to a significant degree such features as arrogance, overconfidence, political involvement, etc.? (PF21)	0.2

In step D, values $\tilde{E}_{ij}^l, j = 1, \dots, M, i = 1, \dots, N, l = 1, \dots, L_j$ should be determined. $\tilde{E}_{ij}^l$ should represent the positive (thus reducing) influence, for the i-th stakeholder (i=1,...N), of additional factors $IF_j^k, j = 1, \dots, M, k = 1, \dots, K_j$ on the weights WF^l, as seen by the j-th stakeholder. Here, N=1, the additional factors will therefore concern only the 1.stakeholder, the project manager, as seen by all stakeholders, the project manager themselves included.

As explained earlier, lists $IF_j^k, j = 1, \dots, M, k = 1, \dots, K_j$ should represent the available knowledge related to the specific case, here to the specific stakeholder. In the following, we will not present the specific lists, but simply sketch the general knowledge each stakeholder might have at the very moment with respect to the 1. stakeholder, the project manager.

Table 3: *Positive influence $\tilde{E}_{1j}^{l}$ of additional factors on the weights WF^{l}*

	l=1	*l*=2	*l*=3	*l*=4	*l*=5
j=1	average (0.3,0.5,0.7)	total (1,1,1)	large (0.7,0.8,0.9)	total (1,1,1)	total (1,1,1)
j=2	small (0.1, 0.2, 0.3)	none (0,0,0)	total (1,1,1)	large (1,1,1)	small (0.1, 0.2, 0.3)
j=3	none (0,0,0)	none (0,0,0)	total (1,1,1)	small (0.1,0.2,0.3)	total (1,1,1)
j=4	small (0.1,0.2,0.3)	small (0.1,0.2,0.3)	total (1,1,1)	total (1,1,1)	total (1,1,1)
j=5	small (0.1,0.2,0.3)	total (1,1,1)	none (0,0,0)	total (1,1,1)	none (0,0,0)

The first stakeholder, the project manager, is here at the same time the estimate provider, and their additional factors $IF_1^k, k = 1, \dots, K_1$ represent the knowledge or opinion they have about themselves. The first row of Table 3 represents their own evaluations, for the different factors $WF^{l},\ l = 1, \dots, 5$, of the reduction of the deteriorating influence of those factors due to their specific features, in their own opinion. The evaluations are rather high, which means that this stakeholder thinks that the specific situation, i.e., their competences, the way they performed the evaluation, etc., renders most of the principal factors deteriorating the estimation quality practically irrelevant. For example, by setting $\tilde{E}_{11}^{2}$ as 'total', they want to convey the message that in their opinion, no important groups of experts have been ignored. However, they did not put 'total' everywhere in the first row of Table 3. This means that they acknowledge that their estimate quality was deteriorated by some factors. In particular, by setting $\tilde{E}_{11}^{2}$ as 'average', they admit that their estimation process was to some extent influenced by political interests and pressure.

Let us consider some other examples:

- $\tilde{E}_{12}^{2}$ equal to 'none' means that the second stakeholder, the technical experts, thinks that the experts have been totally ignored in the estimation process.
- $\tilde{E}_{13}^{5}$ equal to 'none', conveys the message that the local community feels that it has also been ignored.
- $\tilde{E}_{14}^{4}$ equal to 'total' means that experienced project managers either do not have information on project manager training (and do not want to underestimate the competencies of the project manager in

the situation of not possessing the relevant information), or know that the project manager has been well trained in project management.

The other values in Table 3 have an analogous meaning. Formulae (4), (5), and (6) will lead to the following results:

- $\tilde{C}_{11} = (0.13,0.22,0.31)$, which means that the first stakeholder, the project manager, who has provided the estimate $\tilde{A}_1 = (2.8, 2.83,2.85)$, thinks this estimate should be slightly adjusted, to $\tilde{A}^*_{11} = (3.16, 3.45, 3,73)$;
- $\tilde{C}_{12} = (0.53,0.58,0.64)$ – technical experts are of the opinion that the original estimate of the project manager should be increased to a higher degree and become $\tilde{A}^*_{12} = (4.28, 4.5, 4.7)$;
- Furthermore, the other 3 stakeholders would also adjust the original estimate of the project manager, according to the following values (the higher $\tilde{C}_{1j}$, the higher the adjustment, thus the lower the quality of the initial estimate of the project manager, according to the *j*-th stakeholder)
  - $\tilde{C}_{13} = (0.64,0.66,0.68)$ and $\tilde{A}^*_{13} = (4.6, 4.7, 4.8)$
  - $\tilde{C}_{14} = (0.35,0.4,0.47)$ and $\tilde{A}^*_{14} = (3.78, 3.86, 4.2)$
  - $\tilde{C}_{15} = (0.31,0.35,0.39)$ and $\tilde{A}^*_{15} = (3.67, 3.82, 3.96)$

According to formula (7), the aggregation $\tilde{A}^*_1$ will be equal to (3.52, 3.76, 4). This value represents the weighted adjustment of the initial estimate, given by all the stakeholders. As the author of the estimate, the project manager, has the highest weight, their opinion has a high influence on $\tilde{A}^*_1$.

The envelope $\widetilde{AE}$ (formula (8)) will be equal to (3.16, 3.76, 4.8). This number has a wider range than $\tilde{A}^*_1$, because here no weights of the stakeholders are taken into account and the stakeholders providing the estimate are not treated individually. The envelope simply shows the widest range of the values adjusted by all the stakeholders.

The two fuzzy values, € (3.52, 3.76, 4) billion and € (3.16, 3.76, 4.8) billion, provide a better insight into the risk of cost overruns in the project than the original estimate $\tilde{A}_1 = (2.8, 2.83,2.85)$. Although the actual value of project cost, €7 billion, is still beyond those ranges obtained (it exceeds both 4 and 4.8), the cost overrun would have been lower than it actually was if the approach proposed here had been used. And of course, we were not able to use the actual knowledge of the real-world stakeholders. The example was fundamentally illustrative.

Conclusions

We propose to use the fuzzy approach in the estimation of key parameters of megaprojects. To be more exact, we suggest the usage of a relatively modern form of fuzzy numbers, the so-called Z-numbers. Fuzzy numbers themselves have been used in project management for years to model the lack of full knowledge and the variability inherent to projects. Megaprojects are even more susceptible to the lack of knowledge and variability; it should be systematically required to avoid crisp numbers in megaprojects estimation. At least three point estimates, modelled as probability distributions (Hulett, 2011) or fuzzy numbers, should be used. Fuzzy numbers have the advantage of being more flexible and allowing more subjectivity than probability distributions.

Z-numbers add an important element to fuzzy estimation: that of credibility assessment. Lack of credibility, due to many factors discussed in this paper and present in megaprojects to a much higher degree than in projects generally (because of the size of the megaprojects, the amount of financial means involved, and the large groups of stakeholders impacted), means that the estimates of megaprojects cannot be trusted, because they are often formulated based on political or financial interests or by incompetent experts.

The method we propose, which makes use of Z-numbers, allows all key stakeholders to express both their opinion on the estimate and their opinion on the credibility of the experts. Such an approach will exclude cases as the one quoted in this paper, when a ministry ignored negative reports on a megaproject prepared by a highly reliable scientific advisory board.

The method is merely a proposal that has to be tested in practice and calibrated after each application. Its efficiency depends on the quality of the list of factors that negatively or positively influence the quality of the estimate. This list per se is difficult to work out, and each potential project stakeholder has to prepare their own list. In this paper, we propose a starting point: a list following from the literature.

We hope that our proposal, even if the road to its full implementation is long, will ultimately help to spend money on megaprojects in a much more efficient way than is currently the case.

Funding: National Science Centre (Poland), grant number 394311, 2017/27/B/HS4/01881 Grant title: Selected methods supporting project management, taking into consideration various stakeholder groups and using type-2 fuzzy numbers.

Bibliography

Ajam, M. A. (2020). *Leading megaprojects : a tailored approach.*

Atanassov, K. T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, *20*(1), 87–96. https://doi.org/https://doi.org/10.1016/S0165-0114(86)80034-3

Azadeh, A., & Kokabi, R. (2016). Z-number DEA: A new possibilistic DEA in the context of Z-numbers. *Advanced Engineering Informatics*, *30*, 604–617. https://doi.org/10.1016/j.aei.2016.07.005

Bustince, H., Barrenechea, E., Pagola, M., Fernandez, J., Xu, Z., Bedregal, B., Baets, B. De. (2016). A Historical Account of Types of Fuzzy Sets and Their Relationships. *IEEE*

Transactions on Fuzzy Systems, *24*(1), 179–194. https://doi.org/10.1109/TFUZZ.2015.2451692

Brunelli M., Mezei J. (2013). How different are ranking methods for fuzzy numbers? A numerical study, *International Journal of Approximate Reasoning*, 54(5)

Cheeseman, P. (1986). Probabilistic versus Fuzzy Reasoning. In L. N. KANAL & J. F. B. T.-M. I. and P. R. LEMMER (Eds.), *Uncertainty in Artificial Intelligence* (Vol. 4, pp. 85–102). https://doi.org/https://doi.org/10.1016/B978-0-444-70058-2.50011-5

Cheng, C.-B. (2004). Group opinion aggregationbased on a grading process: A method for constructing triangular fuzzy numbers. *Computers & Mathematics with Applications*, *48*(10), 1619–1632. https://doi.org/https://doi.org/10.1016/j.camwa.2004.03.008

Cleden, D. (2017). Managing project uncertainty. In *Managing Project Uncertainty*. https://doi.org/10.4324/9781315249896

Eshghi, A., Mousavi, S. M., & Mohagheghi, V. (2019). A new interval type-2 fuzzy approach for analyzing and monitoring the performance of megaprojects based on earned value analysis (with a case study). *Neural Computing and Applications*, *31*(9), 5109–5133. https://doi.org/10.1007/s00521-018-04002-x

Flyvbjerg, B. (2008). Public Planning of Megaprojects: Overestimation of Demand and Underestimation of Costs. In *Decision-Making on Megaprojects: Cost-Benefit Analysis, Planning and Innovation* (pp. 120–144). https://doi.org/10.4337/9781848440173.00014

Flyvbjerg, B. (2014). What You Should Know About Megaprojects and Why: An Overview. *Project Management Journal*, *45*. https://doi.org/10.1002/pmj.21409.

Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003). MegaProjects and Risk: An Anatomy of Ambition. In *Megaprojects and Risk: An Anatomy of Ambition*. https://doi.org/10.1108/09513550410530199

Goldratt , E. M. (1997). *Critical chain*. Great Barrington, MA: North River Press.

Hendiani, S., Bagherpour, M., Mahmoudi, A., & Liao, H. (2020). Z-number based earned value management (ZEVM): A novel pragmatic contribution towards a possibilistic cost-duration assessment. *Computers and Industrial Engineering*. https://doi.org/10.1016/j.cie.2020.106430

Hulett, D. T. (2011). *Integrated cost-schedule risk analysis* (P. (Firm), Ed.). Farnham, Surrey : Gower.

Kang, B., Wei, D., Li, Y., & Deng, Y. (2012). A method of converting Z-number to classical fuzzy number. *Journal of Information and Computational Science*, *9*, 703–709.

Kerzner, H. R. (2013). *Project Management Metrics, KPIs, and Dashboards*. John Wiley & Sons, Inc.

Kuchta, D., Marchwicka, E. D., & Schneider, J. P. (2021). Sustainability-oriented project scheduling based on Z-fuzzy numbers for public institutions. *Sustainability*, *13*(5), 1–17. https://doi.org/10.3390/su13052801

Lock, D. (2013). *Project management* (10th ed.). Routledge.

Manzoor, E. D. A. (2019). *Project Stakeholder Management*.

Marchwicka, E. D., & Kuchta, D. (2022). Critical Path Method for Z-fuzzy numbers. *Intelligent and Fuzzy Techniques for Emerging Conditions and Digital Transformation : Proceedings of the INFUS 2021 Conference, Held August 24-26, 2021. Vol. 2*, 871–878. Retrieved from https://link.springer.com/chapter/10.1007/978-3-030-85577-2_100

Mendel, J. M., & John, R. I. B. (2002). Type-2 fuzzy sets made simple. *IEEE Transactions on Fuzzy Systems*. https://doi.org/10.1109/91.995115

Merrow, E. W. (2011). *Industrial megaprojects : concepts, strategies, and practices for success*. Retrieved from http://www.books24x7.com/marc.asp?bookid=41723

Mojtahedi, S. M. H., Mousavi, S. M., & Aminian, A. (2008). Fuzzy Group Decision Making: A case using FTOPSIS in mega project risk identification and analysis concurrently. *IEEM*, 1769–1773. https://doi.org/10.1109/IEEM.2008.4738176

Prater, J., Kirytopoulos, K., & Tony, M. (2017). Optimism bias within the project management context: A systematic quantitative literature review. *International Journal of Managing Projects in Business*, *10*, 370–385. https://doi.org/10.1108/IJMPB-07-2016-0063

Romestant, F. (2019). Sustainability agencing: The involvement of stakeholder networks in megaprojects. *Industrial Marketing Management*. https://doi.org/https://doi.org/10.1016/j.indmarman.2019.09.005

Salkeld, D. (2013). *Project Risk Analysis*. Gower Publishing, Ltd.

Schuyler, J. (2001). *Risk and Decision Analysis in Projects / J. Schuyler.*

Sedlin, T., Beckmann, V., & Tan, R. (2020). *Public Participation and Airport Development: The Case of the Site Selection for Berlin Brandenburg Airport (BER) in Germany*. Sustainability (Basel, Switzerland), 12(24), 105-35,. https://doi.org/10.3390/su122410535

Torra, V. (2010). Hesitant fuzzy sets. *International Journal of Intelligent Systems*. https://doi.org/10.1002/int.20418

Wang, J., Zhuang, Yang, J., & Sheng, Z. (2014). The effects of optimism bias in teams. *Applied Economics*, *46*. https://doi.org/10.1080/00036846.2014.948678

Xie, L., Xu, T., Ju, T., & Xia, B. (2022). Explaining the alienation of megaproject environmental responsibility behavior: a fuzzy set qualitative comparative analysis study in China. *Engineering, Construction, and Architectural Management*. https://doi.org/10.1108/ECAM-10-2021-0919

Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*. https://doi.org/10.1016/0165-0114(78)90029-5

Zadeh, Lotfi A. (2011). A Note on Z-numbers. *Information Sciences*. https://doi.org/10.1016/j.ins.2011.02.022

Zhang, Y., Wei, H.-H., Zhao, D., Han, Y., & Chen, J. (2020). Understanding innovation diffusion and adoption strategies in megaproject networks through a fuzzy system dynamic model. *Frontiers of Engineering Management*, *8*(1), 32–47. https://doi.org/10.1007/s42524-019-0082-8