Investigating an automated method for the sensitivity analysis of functions

Sibel EKER s.eker@student.tudelft.nl

Jill SLINGER j.h.slinger@tudelft.nl

Gonenc YUCEL g.yucel@tudelft.nl

Delft University of Technology
2628 BX, Delft, the Netherlands

ABSTRACT

Automated sensitivity analysis approaches in system dynamics focus primarily on model parameters. Although table functions are often subjectively approximated, they do not form the focus of most sensitivity analyses. Recently, a promising approach that allows automation of sensitivity analysis on functions was proposed by Hearne (2010)*, but the applicability of this method to system dynamics table functions has not been studied, yet. In this study, the new method is applied to a simple system dynamics model. In the light of the observations a number of shortcomings are identified and a set of extensions to address these are proposed and then tested. The results of experiments with the original and the extended method demonstrate that the method can be used easily and efficiently for table functions. The extensions are shown to be valuable in creating a more comprehensive method, but they also raise the research issue of the tradeoff between their added value and the cost of dealing with increased complication. Apart from our experimental results, the article also puts forth a set of directions along which the approach can be improved further. Despite the issues requiring further research, the method holds promise for routine implementation.*

KEY WORDS: *function sensitivity analysis, automated sensitivity analysis, epidemic model, triangular functions, uncertainty, table functions, system dynamics*

1. INTRODUCTION

The problems modeled using system dynamics are characterized by uncertainty, arising from a lack of information on the system itself or due to conflicting opinions of the actors involved. This real life uncertainty is reflected in model building as uncertainty in the selection of parameter values or in expressions for variable interactions. Despite the difficulties associated with formulating them, models have to be robust/insensitive to such uncertainties to be considered valuable and useful in solving real life problems.

Sensitivity analysis is broadly defined as "the study of model responses to model changes" (Tank-Nielsen 1980), and provides us with a way to deal with uncertainty. Elaborately, the importance of sensitivity analysis stems mainly from four factors. Firstly, it enhances understanding about both the structure-behavior relationship in the model and real world. Secondly, and as the main reason for its common usage, it shows the effects of uncertainties in the model, e.g. estimated parameters, on the conclusions derived from this model. Thirdly, it determines the parameters which affect the behavior strongly and so facilitates the devotion of limited resources to estimate them. Lastly, the parameters which have a strong effect on the behavior may be the key points upon which new policies can be built (Sterman 2000, 830, Tank-Nielsen 1980).

However, system dynamics has been criticized in the past owing to the absence of a precise theory or method to conduct sensitivity analysis (Meadows 1980). Following the explication and initial refutation of this critique by Meadows, various approaches have been developed. Several authors addressed the importance of sensitivity analysis, and provided basic non-automated approaches to conduct it. In their early work, Ford, Amlin and Backus (1983) emphasized the difficulties of sensitivity analysis such as large number of parameters to change and large number of state variables whose responses are to be observed; and listed the approaches developed by that time. Later on, studies are focused on coping with these problems. One of the two main approaches was to use optimization to determine parameter values which cause maximum deviation from the original model behavior (J. Hearne 1987, Miller 1998). The second approach which attracted more attention was the use of statistics to determine combinations of parameter values in multivariate sensitivity analysis, and to interpret the responses of state variables (Kleijnen 1995, Clemson, et al. 1995). Another example of use of statistics was presented by Ford and Flynn (2005) where they screened the model and utilized the correlation coefficients to find the parameters most influential on the behavior. Meanwhile, the progress in software technology allowed system dynamics and simulation packages to include automated sensitivity analysis tools. This reduced the burden of a comprehensive sensitivity analysis to some extent, but the abovementioned statistical approaches are still needed. In addition to those, since the sensitivity is defined as behavioral sensitivity rather than numerical sensitivity in system dynamics, methods to recognize changes in the behavior patterns have been developed and incorporated with sensitivity analysis. Recent examples of such studies can be seen in (Yucel and Barlas 2011) and (Hekimoglu and Barlas 2010).

One common aspect of these studies is that they deal primarily with sensitivity of the model to changes in parameter values or to the changes in some model structure features such as the model boundary or the level of aggregation (Sterman 2000, 884)**.** In system dynamics, nonlinear relationships between two variables are usually specified by a lookup or *table function,* which shows how the dependent variable nonlinearly varies as the independent variable changes (Sterman 2000, 552). In the context of model building, table functions are sometimes considered as parameters (Tank-Nielsen 1980), but the attention paid to them even in parameter sensitivity analysis is very limited. However, model robustness can depend on the choices of table functions as well because the shapes of them are only approximations of real nonlinear relationships, and due to that sensitivity analysis is indeed the final step of formulating a table function (Sterman 2000, 558-562). Currently, sensitivity analysis activities, particularly automated analyses, focus on parameter sensitivity. Even if a modeler wonders about the outcome of his choices and carries out a manual sensitivity analysis for functions, this effort will be limited because varying the functions manually, especially in combinations of multiple functions, is cumbersome. Clearly, an automated method for function sensitivity analysis is required, before the analysis of the effects of uncertainty in functions can join parameter sensitivity analysis as a routinely applied tool.

Recently, Hearne (2010) proposed a novel method to conduct sensitivity analysis of model functions. This method is based on systematic and parametric perturbation of the graphical functions. Despite being a promising proposal, this approach has not yet been tested thoroughly nor has its wider applicability been established.

The objective of this study is two-fold. First, this study aims to explore the applicability of Hearne's recent work on the sensitivity analysis of system dynamics functions. This is undertaken by implementing the method on a simple model with a single table function, and identifying the concerns and issues requiring attention. Second, extensions to the method are proposed in order to overcome the identified shortcomings. The delineated extensions are then implemented and tested on the same model, and their usefulness is assessed.

With this intention, before diving into the sensitivity analysis method, the model that is used in this study is first introduced in the next section. Thereafter, the third section deals with Hearne's method and its application on the chosen model. In Section 4, the proposed extensions are described and implemented. The paper ends with a discussion of the results and a conclusion regarding the promise of the method.

2. THE SAMPLE MODEL

The model selected for testing the new method for the sensitivity analysis of functions is the basic form of the epidemics model (SIR model) with an alternative infection rate formulation. In the original model (Sterman 2000, 303) where S stands for *susceptible* population and I is the *infected* population. The infection rate (IR) is defined as

$$IR = \frac{dS}{dt} = S * \frac{I}{Total\ Population} * contact\ rate * infectivity$$

The behavior yielded by this formulation when the contact rate is 6 people per person per day, the infectivity is 0.25, the initial number of Infected is 1 and the Total Population is 10000 is depicted in Figure 1.

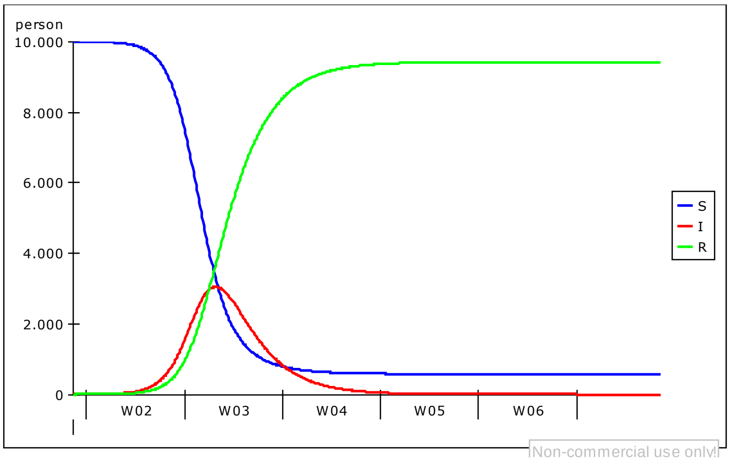


Figure 1: The behavior of original SIR model over time

In the alternative formulation the infection rate defined as

$$\frac{dS}{dt} = S * f\left(\frac{I}{Total\ Population}\right)$$

and *f* depicts the non-linear relationship between the ratio of I to the total population and the infection rate. This relationship is linear in the original formulation. To ensure similarity to the original formulation, the initial shape of this table function is specified as depicted in Figure 2, while the model behavior obtained with this function is depicted in Figure 3.

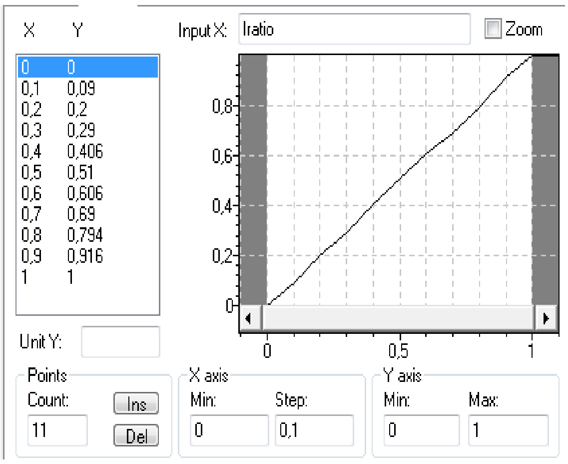


Figure 2: The table function f used in the alternative SIR model

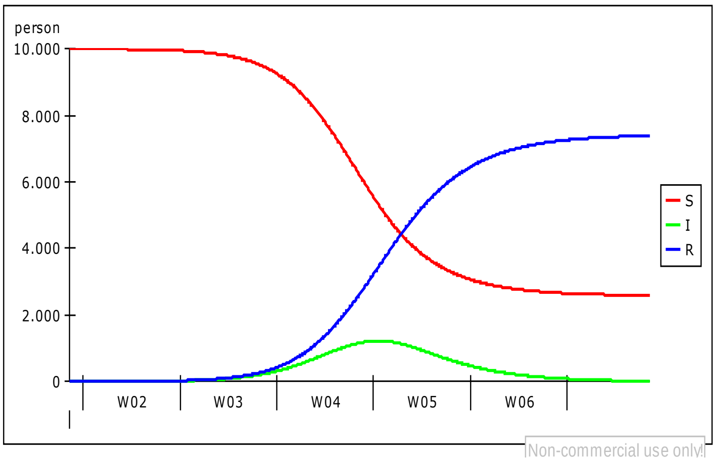


Figure 3: The behavior of alternative SIR model over time

Two assumptions embedded in this table function will be adopted in this study, too. First, the function starts at (0,0), which means that when there are no infected people in the community, then there is no infection. In other words, this disease is transmitted only by contacts between people, and this community is closed to interactions with the outside world. Secondly, the maximum effect of the infected fraction on the infection rate is 1. This means that, the infection rate cannot be higher than the susceptible population. It is important to note that this claim does not imply that the end point will be (1,1). When all or almost all people are infected, the effect of infected on infection rate may be smaller, and may even decline to zero.

3. APPLICATION OF HEARNE'S METHOD TO THE ALTERNATIVE SIR MODEL

In this section, the method of John Hearne will be explained, and then the results obtained by using it on the abovementioned SIR model will be presented.

3.1. Hearne's method for function sensitivity analysis

The basis of Hearne's method for function sensitivity analysis is the multiplication of a model function with another function of specific form but variable parameterization in order to distort the model function. In his study, Hearne used triangular functions as distortion functions. Technically speaking, these triangular functions have the analytical and graphical forms shown below. For a model function r(y) defined on the interval [a, b], perturbation function h(y, p, m) where m and p are parameters is given by:

$$h(y, p, m) = 1 + \frac{m(y - c)}{p - c}$$

$$c = \begin{cases} a \text{ if } y \leq p \\ b \text{ if } y > p \end{cases}$$

Here, it is important to note that *m* stands for the maximum deviation from 1, whereas *p* is the point where this deviation occurs. Since this study does not aim at exploring the effect of different end points of model functions on the behavior, in this paper the distortion function is formulated in such a way that no end point distortion occurs.

Below, Figure 4 shows an example table function which has an exponentially growing shape. If this table function could have an s-shape, almost linear or a less steep form, as R1, R2 and R3 in Figures 6, 8 and 10 respectively, then such shapes can be obtained by multiplying the original table function with the distortion functions shown in Figures 5, 7 and 9. With each possible combination of parameter values of the distortion function, a different table function shape can be obtained and a large span of possibilities can be explored.

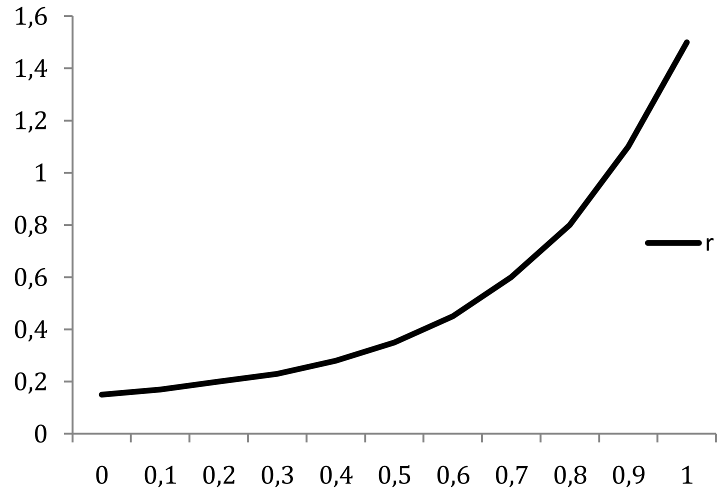


Figure 4: An example of a table function, r

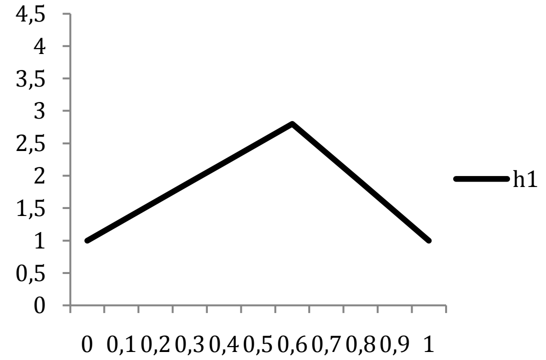


Figure 5: Distortion Function h1 with m=1,8 and p=0,6

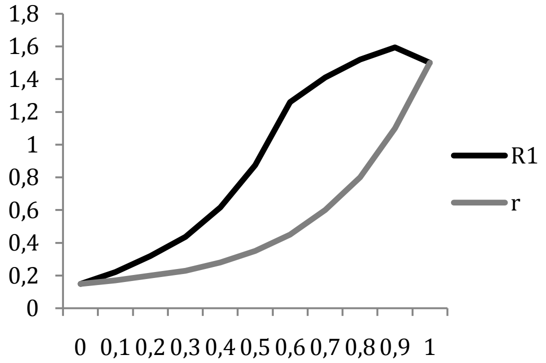


Figure 6: R1, the distorted form of the table function r, distorted by h1

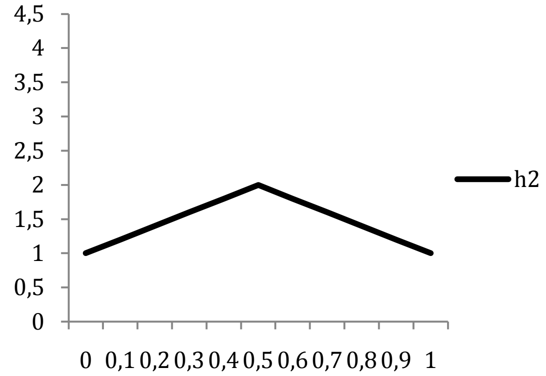


Figure 7: Distortion Function h2 with m=1 and p=0,5

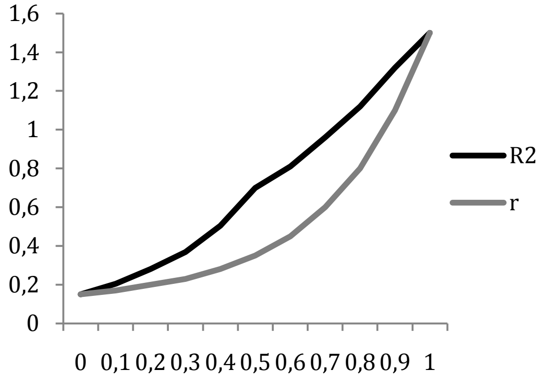


Figure 8: R2, the distorted form of the table function r, distorted by h2

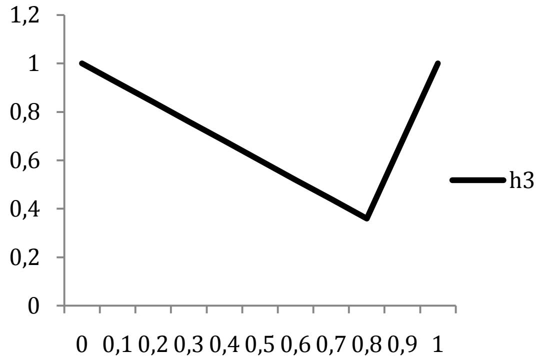


Figure 9: Distortion Function h3 with m=-0,64 and p=0,8

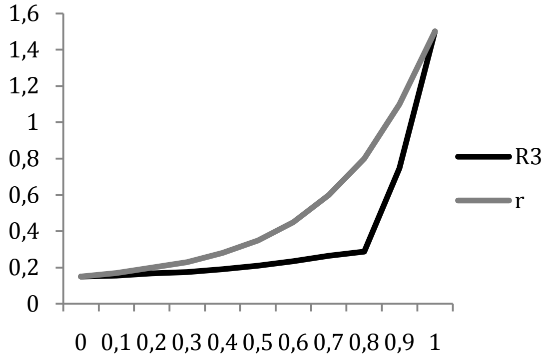


Figure 10: R3, the distorted form of the table function r, distorted by h3

To find the minimum distortion to the model function which causes the undesired model behavior, Hearne formulates the problem as follows:

$$\underline{\min_{p,m}}\, \underline{m}^2$$

$$\textit{constrained by the undesired model behavior condition}$$

In the solution *($\underline{m}^*$, $\underline{p}^*$) to this optimization problem,* m^* represents the magnitude of the distortion, and p^* is the point at which the function manifests most sensitivity to the changes.

This formulation provides the reason for terming this approach an automated method. Instead of making manual changes to the function, the shape of the function which gives rise to undesired behavior can easily be found with the aid of an optimization tool.

3.2. Experiments with the original approach

For the modified SIR model, three types of model behavior are specified as the criteria by which sensitivity will be determined. If these behaviors are generated by minor changes in the table function, then the model will be said to be highly sensitive to the shape of the table function. The three criteria and the rationale for choosing each of them are described subsequently.

i. Maximum number of Infected people greater than 3000 in two months

In the original SIR model, the infected population reaches its maximum value of around 3000 in the second week; whereas the peak value of around 1200 in the alternative model which includes the table function. The table function shape which will yield a similar maximum for infected people is sought by defining the criterion as the maximum number of infected people should exceed 3000 within two months.

ii. Maximum number of Infected people greater than 6000 in two months

Given the duration of infectivity specified in the model, recovery is quite rapid and the maximum number of people infected at any time over a two month period is not very high. The ambitious criterion of exceeding 6000 infected people in two months is chosen to figure out if any change in the table function can cause such extreme behavior.

iii. Number of susceptible people still greater than 9900 after 1 year

This model is intended to model the dynamics of an epidemic. Therefore, if there is no infection occurring, the model cannot be considered to achieve its purpose. This criterion is chosen in an attempt to understand if the shape of the table function could cause the model to be invalid.

Three optimization problems, one for each of these criteria, are formulated as proposed by Hearne. An additional constraint is added to ensure that the values of the new table function lie between 0 and 1, preventing infection rates that are negative or higher than the susceptible population. Theoretically, m can take any value; however, since the table function values cannot exceed 1, or go below 0, it is unnecessary to search for m in a large interval. Therefore, m values are assumed to be between -5 and 5.

The mathematical models specific to each optimization problem are included in Appendix I.

3.3. Results

To find the function shapes which create undesired model behaviors, the optimization problems described above have been solved using the optimization tool of Powersim Studio 7. Optimal solutions are presented in Table 1.

Table 1: Results of applying Hearne's method to the alternative SIR model

	CRITERION 1	CRITERION 2	CRITERION 3
m	0,690	**4,695**	-0,407
p	0,000	**0,238**	0,000

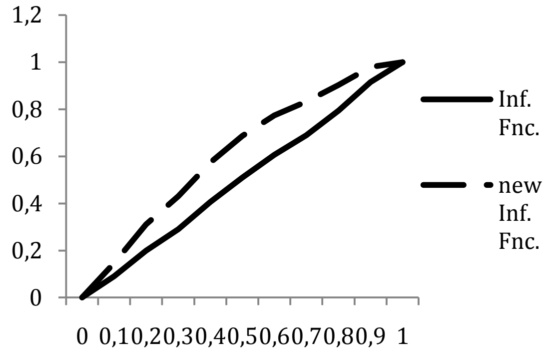


Figure 11: Table function distorted to meet criterion 1

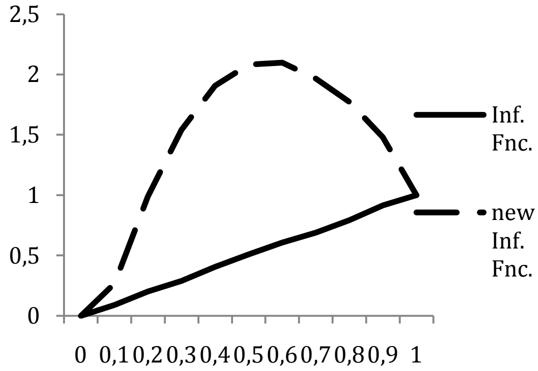


Figure 12: Table function distorted to meet criterion 2

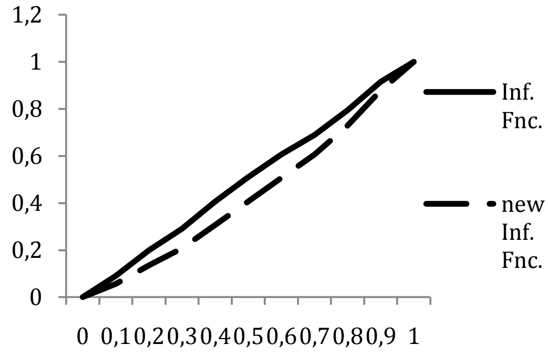


Figure 13: Table function distorted to meet criterion 3

Figures 11, 12 and 13 show how the original table function is distorted to meet each criterion with the function defined by the respective parameters in Table 1. In Figure 11, it can be seen that if the table function is changed to have a convex shape rather than a linearly increasing one, it then causes the maximum infected population to be greater than 3000. The relationship depicted in Figure 12 means that to have maximum infected population greater than 6000, infection rate should be more than, even twice as much as, the susceptible population when the infected people are between 15-100% of the total population. In Figure 13, we see that infection does not occur and the susceptible population remains high if the table function values are slightly decreased. The results generally indicate that the model behavior is more sensitive to the table function in terms of the third criterion, because the absolute maximum deviation from unity (m) is smaller in that case. A feasible solution for the second optimization problem couldn't be found; instead, the result obtained after removing the constraint on the new table function is given. Since such a distortion does not have a real life meaning, it can be said that the behavior is insensitive to the table function in terms of criterion 2.

3.4. Discussion on the observations and identifying potential improvements

The undesired behavior in the first and third cases was obtained by a distortion of the table function made using a triangular function. However, this does not mean that these are the only function shape changes which can generate the undesired behavior. For example, an s-shaped pattern which requires the original function to increase initially then to decrease, or vice versa, could potentially have created the same undesired behavior. As Hearne pointed out, triangular functions are not capable of making different changes on the model functions in two different intervals. With their end points at 1, they cannot decrease the function on one side, while increasing it on the other. However, in order to treat uncertainty in functional form thoroughly, such variations in perturbations are necessary. Therefore, an improvement to Hearne's method lies in determining alternative forms of distortion functions that are able to create such two-sided distortions.

Furthermore, it is observed that this method unnecessarily makes distortions on the regions of table function which are not used during the simulation. For the first criterion, the segment after 0.3 is not used, because the highest infected population is 3000. Similarly, only the part before 0.1 is used for the third criterion. This indicates that the distortions after those points proved unnecessary, but had to be made owing to the shape of the distortion function. A further improvement lies in avoiding these unnecessary distortions.

There is another reason for using such functions for the distortion of table functions: This table function ends at (1, 1) and single-extreme triangular functions with stable end points at 1 can be used to perturb them. On the other hand, for table functions passing through (1, 1) and going further, distorted functions should also pass through (1, 1) and to ensure this, distortion function has to pass through (1, 1) as well. It is not possible to achieve this by a partial function with two segments whose end points are 1. Therefore, it is inevitable to use functions which can provide different distortions in different intervals.

As it was said before, this table function assumes that the effect of infected population ratio on infection rate increases till 1. However, it may not be logical to expect all the susceptible people to become infected even if infected ratio is very high. Therefore, this table function may saturate below 1. This opinion in fact points out the general fact that end points of table functions may also be subject to uncertainty. Also, variable end points increase the variety of function shapes that can be obtained after perturbation. Hence, this function sensitivity analysis method should include the possible variations at the end points as well.

4. EXTENSIONS TO HEARNE'S METHOD

In this section, the formulation and implementation of the proposed extensions on the model will be explained. Then, the results of the implementation will be described.

4.1. Extensions

4.1.1. Double-extreme triangular functions

In selecting distortion functions there is a trade-off between their ability to provide the desired variety of perturbations and the number of parameters they have. The variety of perturbations is important for the reasons previously described. As for the number of parameters, the higher the number, the longer it takes to solve the optimization problem and the more difficult it is to interpret the impact of the parameters on the shape of the function.

Considering this trade-off, the simplest extension that can be made to triangular functions in order to obtain two-sided distortions is to use "double-extreme triangular functions", that is piece-wise functions with two extreme points instead of one. Such functions can be defined using four parameters with meanings similar to the parameters of triangular functions, and they yield two-sided distortions when m_1 and m_2 have opposite signs. If they have the same sign, then they just increase the variety. The definition and an example of these functions, for a model function *r(x)* defined in interval [0, 1] are given below (see Figures 14 and 15):

$$h(x, m_1, p_1, m_2, p_2) = \begin{cases} 1 + \dfrac{m_1}{p_1} x & , \quad 0 \le x < p_1 \\ 1 + m_1 - \dfrac{(m_1 - m_2)}{(p_2 - p_1)}(x - p_1) & , \quad p_1 \le x < p_2 \\ 1 + m_2 - \dfrac{m_2}{(1 - p_2)}(x - p_2) & , \quad p_2 \le x \le 1 \end{cases}$$

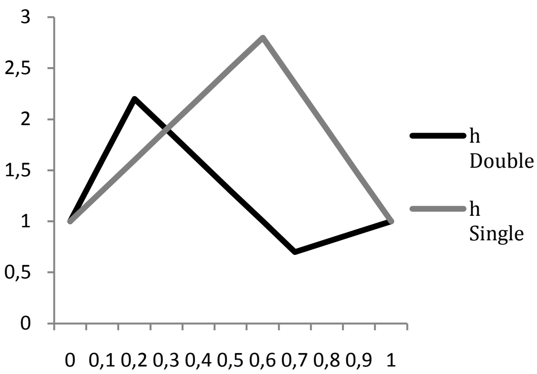


Figure 14: An example of a double-extreme triangular distortion function with m1=1.2, p1=0.2, m2=-0.3 and p2=0.7 (h Double is h1 in Figure 5)

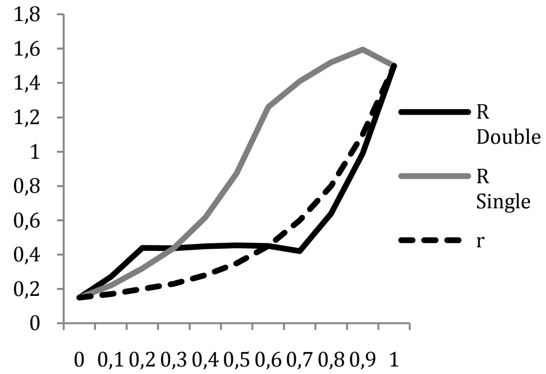


Figure 15: The function in Figure 4 distorted by the single and double-extreme functions in Figure 14

4.1.2. Variable end points

When the end points are allowed to vary, the distortion functions defined above will have to be modified as follows, where *l* and *u* are the values of the table function at 0 and 1, respectively:

For single extreme triangular functions:

$$h(x, p, m, l, u) = \begin{cases} l + \dfrac{m}{p} x & , \quad 0 \le x < p \\ l + m - \dfrac{l + m - u}{1 - p}(x - p) & , \quad p \le x \le 1 \end{cases}$$

For double extreme triangular functions:

$$h(x, m_1, p_1, m_2, p_2, l, u) = \begin{cases} l + \dfrac{m_1}{p_1} x & , \quad 0 \le x < p_1 \\ l + m_1 - \dfrac{(l + m_1 - u - m_2)}{(p_2 - p_1)}(x - p_1) & , \quad p_1 \le x < p_2 \\ u + m_2 - \dfrac{m_2}{(1 - p_2)}(x - p_2) & , \quad p_2 \le x \le 1 \end{cases}$$

It is worthwhile to mention that in case of variable end points, parameter *m* represents the maximum deviation from the lower or the closer end point, not the deviation from unity anymore.

4.2. Experiments with the extended method

The experimental setup for implementing the extensions is the same as explained in section 3.2. An optimization problem for each criterion and an extension is defined. For double-extreme distortion functions, both the cases in which the end points are stable and in which they are variable are considered.

Because the population is assumed isolated from external infection sources (see section 2), even the distorted table function is assumed to start at (0,0) and no change is made to this point. Yet, the lower end point of the distortion function is still assumed to be a variable and take values different from 1, because the beginning point also affects the shape of the function.

The mathematical models specific to each optimization problem are detailed in Appendix I.

To ensure two-sided distortion using double triangular functions, which was the reason for introducing them in the first place, m_1 and m_2 can be forced to have opposite signs.

4.3. Results

The optimization problems are solved as before. The optimal solution of each problem can be seen in Table 2 below:

Table 2: Results of applying Hearne's method with extensions to the alternative SIR model

<table>
<tr><th></th><th></th><th></th><th>CRITERION 1</th><th>CRITERION 2</th><th>CRITERION 3</th></tr>
<tr><td rowspan="6">SINGLE-EXTREME DISTORTION FUNCTION</td><td rowspan="2">END POINTS STABLE AT 1 (Original method)</td><td>m</td><td>0,690</td><td>4,695</td><td>-0,407</td></tr>
<tr><td>p</td><td>0,000</td><td>0,238</td><td>0,000</td></tr>
<tr><td rowspan="4">VARIABLE END POINTS</td><td>m</td><td>0,725</td><td>4,810</td><td>0,000</td></tr>
<tr><td>p</td><td>0,001</td><td>0,364</td><td>0,505</td></tr>
<tr><td>l</td><td>1,000</td><td>1,000</td><td>0,295</td></tr>
<tr><td>u</td><td>0,874</td><td>0,609</td><td>0,428</td></tr>
<tr><td rowspan="10">DOUBLE-EXTREME DISTORTION FUNCTION</td><td rowspan="4">END POINTS STABLE AT 1</td><td>m1</td><td>0,704</td><td>4,745</td><td>-0,446</td></tr>
<tr><td>p1</td><td>0,000</td><td>0,402</td><td>0,000</td></tr>
<tr><td>m2</td><td>-0,050</td><td>0,000</td><td>0,002</td></tr>
<tr><td>p2</td><td>1,000</td><td>1,000</td><td>0,366</td></tr>
<tr><td rowspan="6">VARIABLE END POINTS</td><td>m1</td><td>4,047</td><td>2,383</td><td>0,000</td></tr>
<tr><td>p1</td><td>0,484</td><td>0,000</td><td>0,230</td></tr>
<tr><td>m2</td><td>2,257</td><td>2,782</td><td>0,000</td></tr>
<tr><td>p2</td><td>0,957</td><td>0,729</td><td>0,556</td></tr>
<tr><td>l</td><td>0,736</td><td>1,000</td><td>0,031</td></tr>
<tr><td>u</td><td>0,131</td><td>1,000</td><td>0,869</td></tr>
</table>

There is no feasible solution which makes the maximum infected population greater than 6000 in any of the cases. The solutions above (in bold typeface) belong to the modified optimization problem without the constraint on the values of the distorted table function.

The graphs of the distortion functions with these parameters and how they perturb the table function can be seen in Appendix II.

In the context of the modified SIR model, the results above, supported by the visualizations in Appendix II, are very similar to the results of the original method. They indicate that the model is moderately sensitive to changes in the shape of table function f when the sensitivity criterion is that the maximum value of infected population exceeds 30% of the total population. If this threshold is increased, that is, if the model is expected to demonstrate even less desirable behavior, then it can be said to be totally insensitive, because the distortions required to produce such behavior no longer have real life interpretations. However, if the undesired behavior is that the final susceptible population is greater than 99% of the total population, then the behavior can be said to be highly sensitive to the changes in the shape of f, because slight decreases at the initial points of the table function are sufficient to keep the infection rate very low.

In the context of method extension, it is seen that on this model and for these criteria, the single-extreme triangular function with stable end points was sufficient to carry out the function sensitivity analyses. The double-extreme functions chosen by the optimization algorithm act very similarly to the single-extreme ones, because two of their four intervals are negligibly narrow. When the end points are allowed to vary, they are chosen different from 1, but these changes are not really effective, because these parts of the table function are not used. Still, the extensions proposed here are theoretically necessary because they enlarge the solution space and so facilitate a fuller consideration of uncertainty. Furthermore, these experiments are useful in determining the utility of the method, delineating necessary improvements and demonstrating their effects.

4.4. Discussion on the observations

For some of the optimal solutions, the end point deviates from unity. However, this final segment of the table function is never even used. A *u* value lower than one is chosen by the optimization algorithm either arbitrarily since it has no effect, or due to the ease of creating lower values for other parts of the function with a small end point. Therefore, before deriving the conclusion that the behavior is sensitive to the changes in the end point of the table function if it falls below u, one needs to verify whether this part of the function is used or whether there is a significant effect on the part actually used.

Further, the variable end points add another level of complication to the problem. It is indeed straightforward to include them in the problem and they are expected to stay at 1 if variations don’t cause a significant change in the model behavior. However, as in the case of the double-extreme distortion function and the first criterion, although it is known that there is a feasible solution when the end points are 1, the search algorithm does not find this. To overcome this difficulty, either the search process can be extended, which brings the dilemma of quality versus time and effort; or another optimization tool can be used, which decreases the efficiency of the analysis procedure by separating the environments in which the optimization and simulation are executed.

Moreover, although double-extreme triangular functions are able to generate two-sided distortions, their perturbation variety is still limited. To increase this variety, further functional forms, such as cubic polynomials or sinusoidal functions, can be tried out; and the added value of having this variety against the complications and difficulty of interpretation caused by the higher number of parameters, can be assessed.

5. DISCUSSION AND CONCLUSION

In this study, the applicability of Hearne's automated sensitivity analysis method for system dynamics functions has been tested by applying the method to a simple model containing a table function. Extensions to the method have been developed and implemented. Specific shortcomings of both the original method and the extended version are delineated in sections 3.4 and 4.4. In addition to these, there are some general issues that need to be addressed.

First, the claims made about the table function's range in use should be generalized, because it is not unusual for models to use only a limited part of the table functions under given initial conditions and parameter values. However, it is not possible to capture the exact end points of the domain interval used in runtime, and to make distortions only in this interval. Therefore, it is necessary to check the active domain of the table functions, and how much of the distortions are really used in the simulation before deriving conclusions from the results obtained with this method.

Second, in this study, undesired behaviors were generated by single-extreme functions in only one direction. To fully understand the added value of double-extreme functions, different undesired behavior types should be determined and tested. Also, one may argue that sensitivity in this study is defined more in terms of numerical changes here, rather than in terms of significant behavioral changes. This claim is undeniably true; but with an optimization tool in which the constraints can be applied only once during the simulation, not continuously, it is not straightforward to capture behavioral changes. Indeed, automated behavioral change analysis methods such as pattern recognition would need to be incorporated in the extended Hearne method if the goal of creating an automated sensitivity analysis method for model functions is to be attained.

Third, in this paper the method is applied on a simple model with a single table function, and undesired behaviors are selected on the basis of the behavior of a single stock variable. If the method is used on a larger model to create simultaneous distortions on multiple table functions and criteria dependent on combinations of several stock variables are defined, then the value of this automated method can be established.

Both the original and the extended version of Hearne's new function sensitivity analysis method were applied using the optimization tools available in system dynamics software, demonstrating the promise of the method to deliver automated sensitivity analysis of functions in system dynamics models. A number of points which need attention if this method is to be used routinely for table functions sensitivity analysis provide the points of departure in future work on this issue.

REFERENCES

Clemson, Barry, Yongming Tang, James Pyne, and Resit Unal. "Efficient methods for sensitivity analysis." *System Dynamics Review* 11, no. 1 (1995).

Ford, A, J.S. Amlin, and G. Backus. "A Practical Approach to Sensitivity Testing of System Dynamics Models." *Proceedings of the 1983 International System Dynamics Conference. Chestnut Hill, MA*, 1983: 261-280.

Ford, A., and H Flynn. "Statistical Screening of System Dynamics Models." *System Dynamics Review*, 2005.

Hearne, J, W. "An Automated Method for Extending Sensitivity Analysis to Model Functions." *Natural Resource Modleing*, 2010.

Hearne, J.W. "An approach to resolving parameter sensitivity problem is system dynamics methodology." *Applied Mathematical Modelling*, 1987.

Hekimoglu, Mustafa, and Yaman Barlas. "Sensitivity Analysis of System Dynamics Models by Behavior Pattern Measures." *Proceedings of the 28th International Conference of the System Dynamics Society.* 2010.

Kleijnen, J.P.C. "Sensitivity analysis and optimization of system dynamics models: regression analysis and statistical design of experiments." *System Dynamics Review*, 1995.

Meadows, D. "The Unavoidable a Priori." In *Elements of the System Dynamics Method*, by Jørgen Randers. Cambridge, Massachusetts, USA: Productivity Press, 1980.

Miller, J.H. "Active Nonlinear Tests (ANTs) of Complex Simulation Models." *Management Science*, 1998.

Sterman, J. *Business Dynamics: Systems Thinking and Modeling for a Complex World* . Irwin/McGraw-Hill, 2000.

Tank-Nielsen, T. "Sensitivity Analysis in System Dynamics." In *Elements of the System Dynamics Method*, by Jørgen Randers. Cambridge, Massachusetts, USA: Productivity Press, 1980.

Yucel, Gonenc, and Yaman Barlas. "Automated Parameter Specification in Dynamic Feedback Models Based on Behavior Pattern Features." *System Dynamics Review* 27, no. 2 (2011): 195-215.

APPENDIX I: OPTIMIZATION PROBLEMS

1. Single extreme, stable end points, model behavior 1

(1) $$\min_{p,m} m^2$$

s.t.

(2) $$\max_t I \geq 3000$$

(3) $$0 \leq new\ Infection\ Function[i] \leq 1 \quad ; \quad i = 0, 0.1, 0.2, \ldots, 1$$

(4) $$0 \leq p \leq 1$$

(5) $$-5 \leq m \leq 5$$

where *new InfectionFunction[i] = InfectionFunction[i] * h[i]*

InfectionFunction[] = {0, 0.09, 0.2, 0.29, 0.406, 0.51, 0.606, 0.69, 0.794, 0.916, 1}

$$h[i] = 1 + \frac{m(i-c)}{p-c} \text{ where } c = \begin{cases} a \text{ if } i \leq p \\ b \text{ if } i > p \end{cases}$$

(1) *Objective function:* The optimization problem for sensitivity analysis is stated as "finding the minimum distortion which creates the undesired model behavior". Therefore, the distortion is measured by the square of maximum deviation of the distortion function from unity, to ensure that absolute distortion is taken into account.

(2) *Undesired model behavior constraint:* The first undesired model behavior criterion is formulated as a constraint which restricts the maximum value of the Infected population over the simulation time to be greater than 3000. In Powersim, this constraint was defined by using RUNMAX() function.

(3) *Feasibility constraint:* It is known that the values below 0 and above 1 have no real life meaning for the table function used in this study. However, it is not possible to set a constraint to check the distorted values at each time point. Therefore, an array of discrete values of the distorted table function, namely newInfectionFunction[], is created and restricted to the interval of [0, 1], although it is known that the distortions are continuous. Each element of newInfectionFunction[] is formed by multiplying each element of the original distortion function array (InfectionFunction[]) with the corresponding value of the distortion function(h[i]).

(4) *Search range of p:* As it can be recalled, parameter p represents the point where the maximum distortion from unity occurs. Therefore, it has to within the domain of table function, which is [0, 1] in this case.

(5) *Search range of m:* There is no limit on the values of m, but it is anticipated that high values cause infeasible distortions. Therefore, the search range is specified as [-5, 5] to include downward distortions as well.

2. Single extreme, stable end points, model behavior 2

Problem 1, except the first constraint which is replaced by

(2) $$\max_t I \geq 6000$$

3. Single extreme, stable end points, model behavior 3

Problem 1, except the first constraint which is replaced by

(2) $$S(t) \geq 9900 \quad ; \quad t = 1\,year$$

4. Single extreme, variable end points, model behavior 1

(1) $$\min_{p,m,l,u} m^2$$

s.t.

(2) $$\max_t I \geq 3000$$

(3) $$0 \leq new\ Infection\ Function(i) \leq 1 \quad ; \quad i = 0, 0.1, 0.2, \dots, 1$$

(4) $$0 \leq p \leq 1$$

(5) $$-5 \leq m \leq 5$$

(6) $$0 \leq l \leq 2$$

(7) $$0 \leq u \leq 1$$

where *new InfectionFunction[i] = InfectionFunction[i] * h[i]*

InfectionFunction[] = {0, 0.09, 0.2, 0.29, 0.406, 0.51, 0.606, 0.69, 0.794, 0.916, 1}

$$h[i] = \begin{cases} l + \frac{m}{p} i & , \quad 0 \leq i < p \\ l + m - \frac{l + m - u}{1 - p}(i - p) & , \quad p \leq i \leq 1 \end{cases}$$

(6) *Search range of l:* In addition to the problem 1, search ranges for the end point parameters are specified in this case. For l, the end point at 0, this range is [0, 2] because it is thought that the values higher than 2 would easily cause infeasibility.

(7) *Search range of u:* The end point of the table function at 1 is already 1, and since multiplying it with values higher than 1 would cause infeasible values, the search range of u is kept between 0 and 1.

5. Single extreme, variable end points, model behavior 2

Problem 4, except the first constraint which is replaced by

(2) $$\max_t I \geq 6000$$

6. Single extreme, variable end points, model behavior 3

Problem 4, except the first constraint which is replaced by

(2) $$S(t) \geq 9900 \quad ; \quad t = 1\,year$$

7. Double extreme, stable end points, model behavior 1

(1) $$\min_{p_1, m_1, p_2, m_2} {m_1}^2 + {m_2}^2$$

s.t.

(2) $$\max_t I \geq 3000$$

(3) $$0 \leq new\ Infection\ Function(i) \leq 1 \quad ; \quad i = 0, 0.1, 0.2, \dots, 1$$

(4) $$0 \leq p_j \leq 1 \quad ; \quad j = 1,2$$

(5) $$-5 \leq m_k \leq 5 \quad ; \quad k = 1,2$$

(6) $$p_1 - p_2 \leq 0$$

where *new InfectionFunction[i] = InfectionFunction[i] * h[i]*

InfectionFunction[] = {0, 0.09, 0.2, 0.29, 0.406, 0.51, 0.606, 0.69, 0.794, 0.916, 1}

$$h[i] = \begin{cases} 1 + \frac{m_1}{p_1} i & , \quad 0 \le i < p_1 \\ 1 + m_1 - \frac{(m_1 - m_2)}{(p_2 - p_1)}(i - p_1) & , \quad p_1 \le i < p_2 \\ 1 + m_2 - \frac{m_2}{(1 - p_2)}(i - p_2) & , \quad p_2 \le i \le 1 \end{cases}$$

(1) *Objective function:* This time, two parameters, both m_1 and m_2 contribute to the maximum distortion, hence they are both included in the objective function.

(4) *Search range of p_1 and p_2:* The search range [0,1] applies to both p parameters.

(5) *Search range of m_1 and m_2:* As specified in the first problem, the search range [-5, 5] applies to both m parameters.

(6) *Difference of p's:* The distortion function is defined based on the assumption that p_1 precedes p_2. With this constraint, it is guaranteed that p_1 is smaller than p_2.

8. Double extreme, stable end points, model behavior 2

Problem 7, except the first constraint which is replaced by

(2) $$\max_t I \ge 6000$$

9. Double extreme, stable end points, model behavior 3

Problem 5, except the first constraint which is replaced by

(2) $$S(t) \ge 9900 \quad ; \quad t = 1\, year$$

10. Double extreme, variable end points, model behavior 1

(1) $$\min_{p_1, m_1, p_2, m_2, l, u} m_1^2 + m_2^2$$

s.t.

(2) $$\max_t I \ge 3000$$

(3) $$0 \le new\ Infection\ Function(i) \le 1 \quad ; \quad i = 0, 0.1, 0.2, \dots, 1$$

(4) $$0 \le p_j \le 1 \quad ; \quad j = 1,2$$

(5) $$-5 \le m_k \le 5 \quad ; \quad k = 1,2$$

(6) $$p_1 - p_2 \le 0$$

(7) $$0 \le l \le 1$$

(8) $$0 \le u \le 1$$

where *new InfectionFunction[i] = InfectionFunction[i] * h[i]*

InfectionFunction[] = {0, 0.09, 0.2, 0.29, 0.406, 0.51, 0.606, 0.69, 0.794, 0.916, 1}

$$h[i] = \begin{cases} l + \frac{m_1}{p_1} i & , \quad 0 \le i < p_1 \\ l + m_1 - \frac{(l + m_1 - u - m_2)}{(p_2 - p_1)}(i - p_1) & , \quad p_1 \le i < p_2 \\ u + m_2 - \frac{m_2}{(1 - p_2)}(i - p_2) & , \quad p_2 \le i \le 1 \end{cases}$$

(7) and (8) Search ranges of end points: The search range of both of the end points is specified as the interval [0,1] because the values out of this interval are infeasible.

11. Double extreme, variable end points, model behavior 2

Problem 10, except the first constraint which is replaced by

(2) $$\max_t I \geq 6000$$

12. Double extreme, variable end points, model behavior 3

Problem 10, except the first constraint which is replaced by

(2) $$S(t) \geq 9900 \quad ; \quad t = 1\ year$$

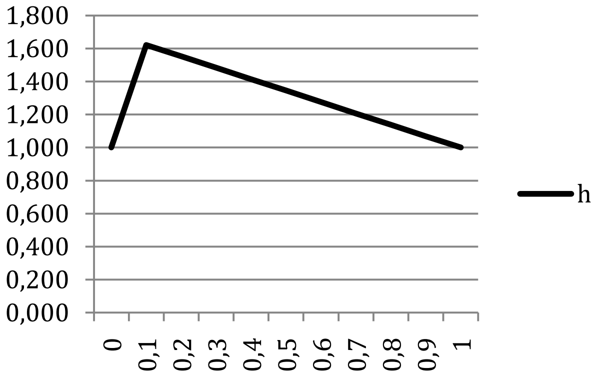
APPENDIX II: EXPERIMENT RESULTS

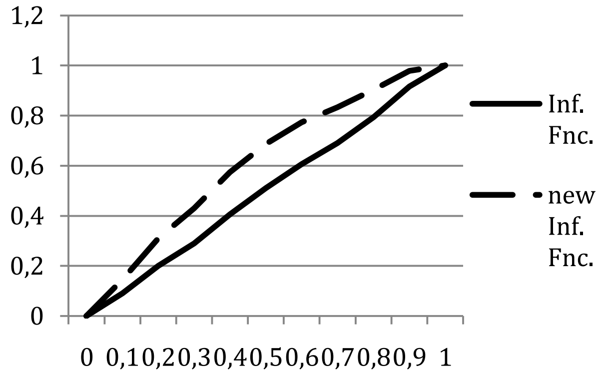
1. Single-Extreme Triangular Distortion Function

a. Stable End Points

i. Criterion 1 (Model behavior: Max(I)>3000)

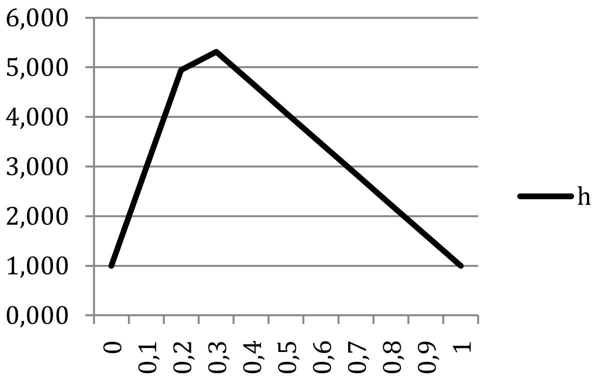
m	0,690	p	0,000

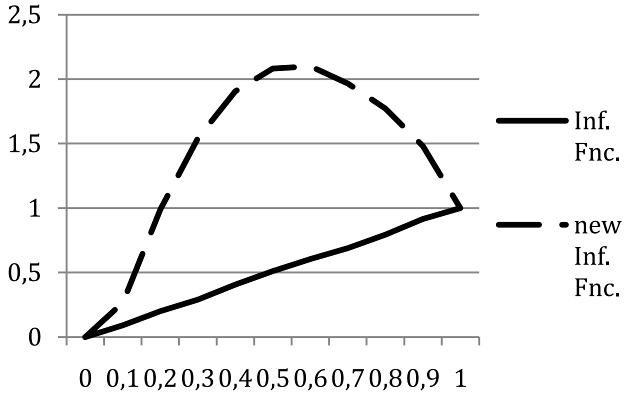




ii. Criterion 2 (Model behavior: Max(I)>6000)

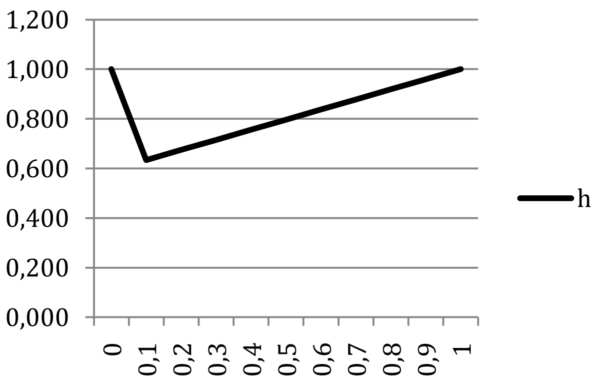
m	4,695	p	0,238

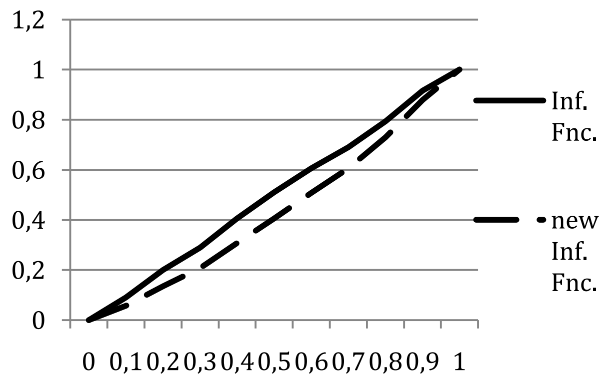




iii. Criterion 3 (Model behavior: Final S>9900)

m	-0,407	p	0,000

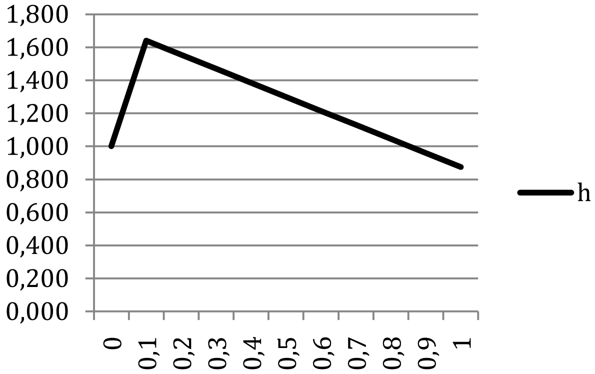


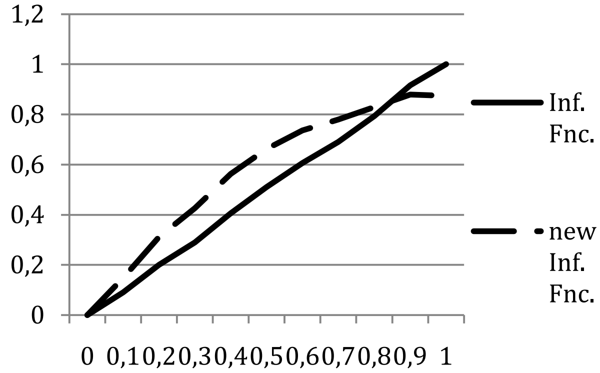


b. Variable End Points

i. Criterion 1 (Model behavior: Max(I)>3000)

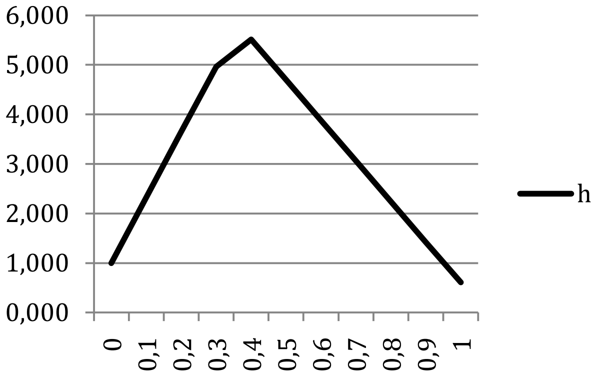
m	0,725	p	0,001
l	1,000	u	0,874

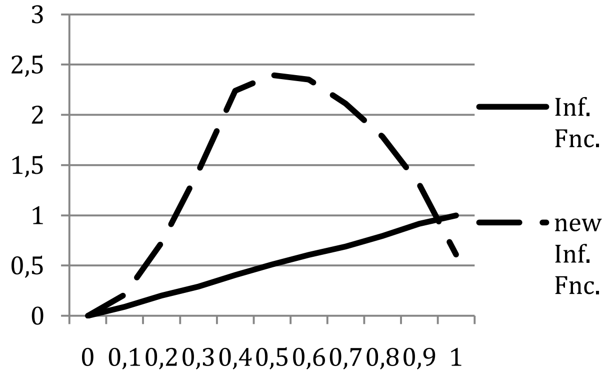




ii. Criterion 2 (Model behavior: Max(I)>6000)

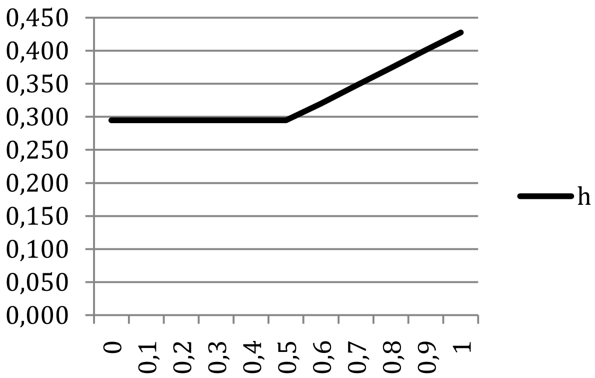
m	4,810	p	0,364
l	1,000	u	0,609

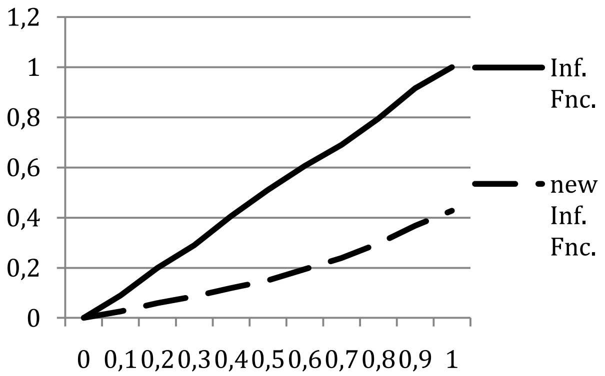




iii. Criterion 3 (Model behavior: Final S>9900)

m	0,000	p	0,505
l	0,295	u	0,428



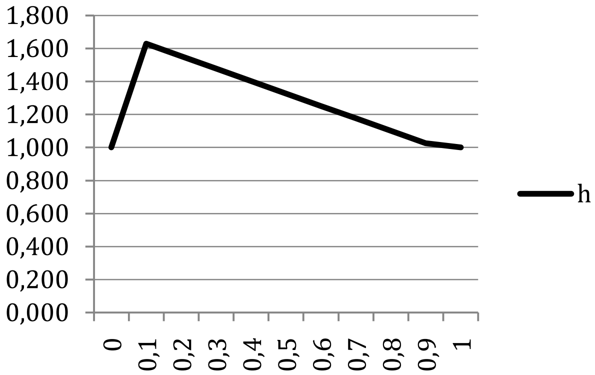


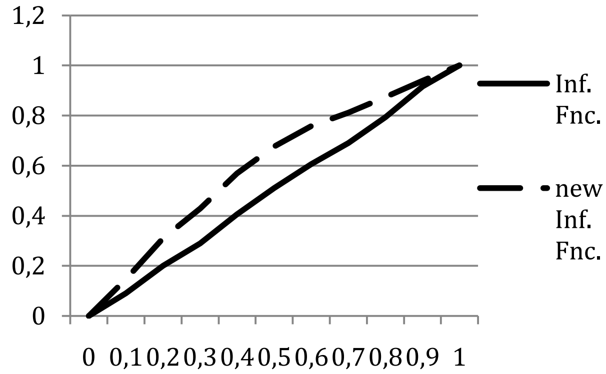
2. Double-Extreme Triangular Distortion Function

a. Stable End Points

i. Criterion 1 (Model behavior: Max(I)>3000)

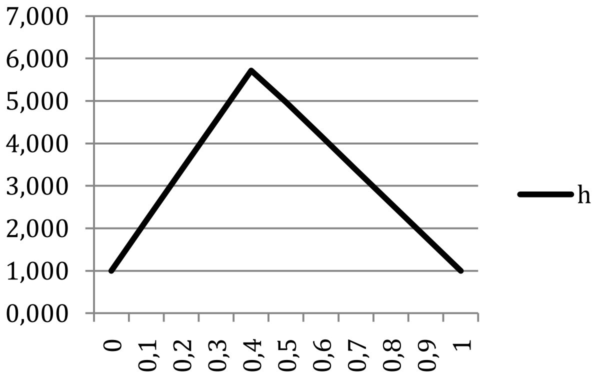
m1	0,704	p1	0,000
m2	-0,050	p2	1,000

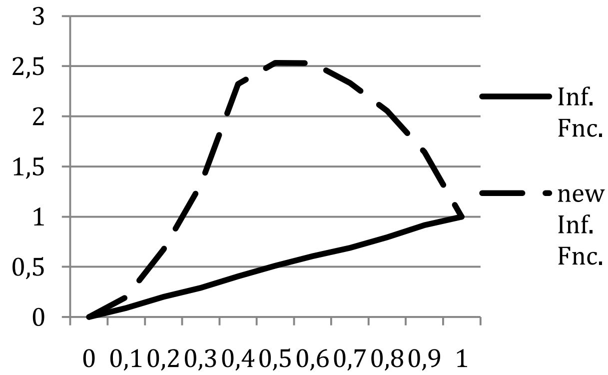




ii. Criterion 2 (Model behavior: Max(I)>6000)

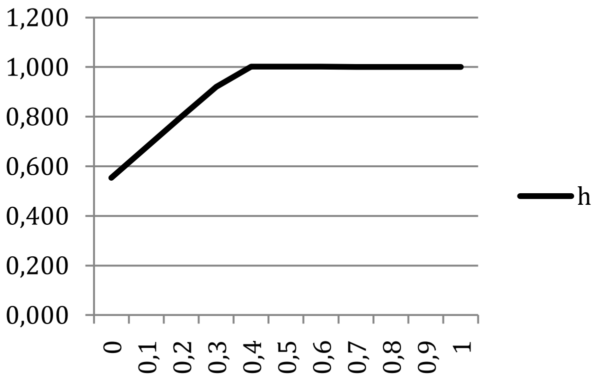
m1	4,745	p1	0,402
m2	0,000	p2	1,000

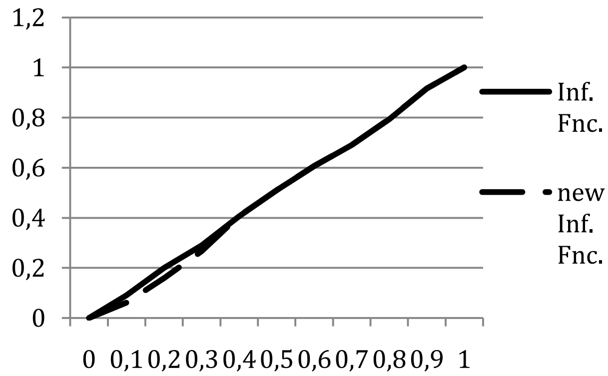




iii. Criterion 3 (Model behavior: Final S>9900)

m1	-0,446	p1	0,000
m2	0,002	p2	0,366

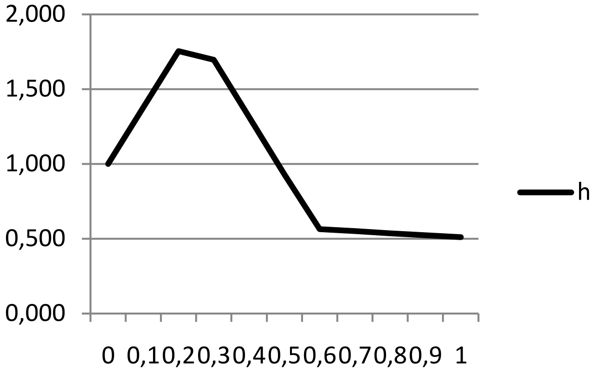


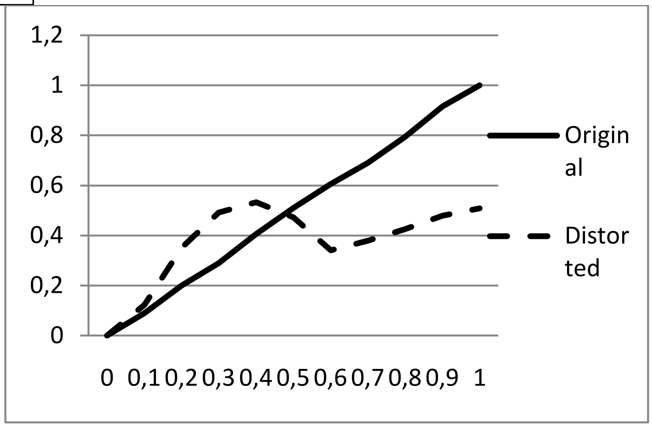


b. Variable End Points

i. Criterion 1 (Model behavior: Max(I)>3000)

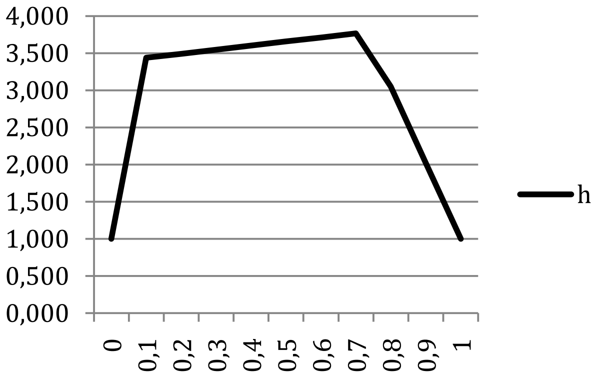
m1	4,047	p1	0,484
m2	2,257	p2	0,957
l	0,736	u	0,131

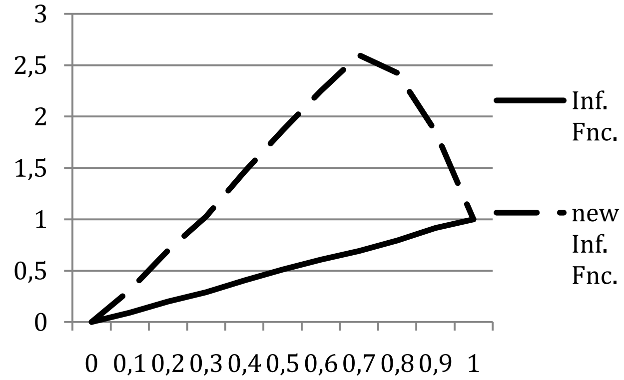




ii. Criterion 2 (Model behavior: Max(I)>6000)

m1	2,383	p1	0,000
m2	2,782	p2	0,729
l	1,000	u	1,000





iii. Criterion 3 (Model behavior: Final S>9900)

m1	0,000	p1	0,230
m2	0,000	p2	0,556
l	0,031	u	0,869