SYSTEM DYNAMICS COMBINED WITH MONTE CARLO SIMULATION

Norodd Hagenson
Norwegian Defense Research Establishment
P O Box 25
N-2007 Kjeller Norway

ABSTRACT

The author suggests that a combination of System Dynamics (SD) thinking combined with Monte Carlo simulation models can yield new insight and be a useful tool. Systems with feedback loops often contain elements of uncertainty or randomness which can be modeled by Monte Carlo methods. On the other hand, feedback loop analysis could certainly benefit Monte Carlo simulation models. Studying single runs of SD models may yield considerable insight. But when a parameter is set to a constant or average value, variance is lost. Variance plays an important role in portraying any risk involved in a system.

These points will be illustrated by an example from an analysis performed at NDRE where SD thinking applied to a Monte Carlo model was the key to solving an important question. The example concerns dimensioning Airfield Damage Repair (ADR) capacity on Norwegian airbases subject to hostile attacks. One key question was: How long time must the runway be open per day in order to obtain acceptable operating conditions for air defense fighter aircraft? Does there exist some minimum threshold?

The main feedback loops concern damage on the runway and attrition between attacking aircraft, ground based air defense and defending air defense aircraft (depending on open runways). The elements of randomness concern the damage inflicted on the runway, and the repair time.

It is shown that under certain conditions (too low repair capacity) there is a risk of defending aircraft either being pinned in or winning the battle. The feedback loop between defending aircraft, attacking aircraft and the runway state plays a key role along with the randomness in the early damage. The statistical distribution of the fraction of day open may over time develop into having one peak close to 0 (closed), one peak close to 1 (open), and little in between. The average value is merely a weighted average between two extremes.

On the other hand, with sufficient repair capacity, the risk of being pinned in was eliminated. The effects were easily understood when thinking in terms of feedback loops, but the element of randomness was essential in order to recognize the threshold when the risk of being pinned in occured.

The author believes that a similar combination of techniques could benefit traditional SD models, too.

SYSTEM DYNAMICS COMBINED WITH MONTE CARLO SIMULATION

1. General ideas

Whether System Dynamics (SD) should be considered a technique or a lifestyle is a matter of opinion. To me SD is viewing problems from a different angle than a physical one, focusing on feedback, dynamic properties and simple structure. But I also consider it a simulation technique along with any ordinary simulation technique.

System Dynamics techniques and tools like Monte Carlo simulation models certainly have different properties. Choosing one or the other may gain in some respects while leaving something to be desired in others. However, quite often these different techniques may be combined. And maybe the differences appear largest on the surface and the similarities are stronger than we think.

SD has its force in gaining insight and investigate system behavior rather than accurate estimation. Also, SD can be considered as a way of viewing the world or perhaps a philosophical approach to describing virtually "anything". Surprisingly much can be viewed from the SD angle. However, SD alone may not always be the most fruitful approach.

SD typically starts with a feedback loop analysis of the problem (or system) followed by construction of a feedback loop model, including flow chart(s) and program.

In other kinds of simulation one starts by considering what elements to include and how to represent them in such a way that the relevant factors are properly described without undue detail. One may recognize the existence and importance of feedback and dynamic behavior, but there is usually no specific feedback loop analysis. Then a model is constructed based on either physical processes or other readily describable phenomena.

The difference between these approaches might be less than it appears. SD models could easily include technical calculations or even be programmed in any ordinary computer language. Similarly, feedback is often intentionally or unintentionally designed into ordinary simulation models.

Some systems contain natural variance in their behavior. These systems are suited for Monte Carlo simulation.

A strong point of Monte Carlo simulation is the repeated replications, yielding averages based on several factors varying randomly simultaneously. These averages may very well be different from a comparable run based on fixed expected values instead of statistical or probabilistic random values.

Systems with natural variance may also contain feedback which may cause some of the variance.

Especially when such feedback is involved, it becomes important to use Monte Carlo methods instead of expected values. But this is also the field where dynamic properties become important. Thus the combination of SD and Monte Carlo has a role to play here. It can be done in various ways.

It has been advocated to use two models, for instance one SD model and one ordinary simulation model. But if both models feature only one, but not both of the two essential elements, which are feedback and randomness, the benefit may be small. Or it may even happen that the models give quite different results, and it may be appearent that neither the individual models, nor the two of them together will do.

A much better approach is combining both randomness and feedback into a single model. This could be done using randomness inside an SD model, or apply feedback analysis to an ordinary simulation model, like in the example which follows.

The peculiarities of feedback and randomness will be further discussed in sec 8 : U-shaped distributions. First we shall look into an example where the combination of SD thinking and Monte Carlo simulation proved very useful, and was in fact the key to solving a specific problem.

2. An example

The example is based on an operations research study at the Norwegian Defence Research Establishment. The study was classified, which implies that not all assumptions and results can be presented in this paper. Therefore this paper will focus mainly on the modelling approach and general trends in the results. Results will be based on declassified data.

The example deals with airbase attack and operability, focusing on runway bombing with modern weapons.

Air warfare can be fought in more than one way. Offensive counter air (OCA) is an increasingly important aspect. OCA means attacking airbases and aircraft on the ground. OCA pays off best when used in the first phase of the air campaign. Such a tactic was used very effectively by Israel in the 1967 war. Since that time, defence against OCA has become a greater concern. Aircraft shelters have been built, protecting aircraft which is the most vulnerable asset. Repair and maintenance facilities can also be moved into protected sites. But the runway still lies unprotected and will be the main target. This paper will consider runway bombing and repair, focusing on airbase operability.

The traditional way of runway attack used to be medium bombers (MB) dropping unguided bombs from an altitude of 10000' - 20000'. Today rocket assisted bombs have changed this picture. These bombs can be effectively delivered by fighter bombers (FB) attacking low level. The low level attack will give the attacking aircraft better protection against ground based air defence as well as better aiming accuracy. The bomb's rocket motor ensures that the bomb penetrates the runway surface before detonation, resulting in a crater and upheaval around it. Both the visible crater and the upheaval must be repaired before aircraft can land and take off.

In our scenario, the attacker has a limited number of MBs, but many FBs. The old MB threat, which still exists of course, has been considered manageable by a rapid runway repair (RRR) organization. The FBs give the threat a new dimension. We must take into account that Orange has great flexibility using his FB force. Initially he may use a large fraction for OCA, and then revert to normal operations when he has achieved superiority. This initial period is critical.

The next generation of threat will be dispenser weapons. These are containers, mainly delivered stand off. They spread out either a large number of mines, or a number of small anti runway bombs in order to make many small craters, consisting mainly of upheaval, and supposedly a bit difficult to repair.

The question is now:

- can the airbase survive the FB threat against the runways?

This problem has been analyzed using a discrete event simulation model. The main measure of effectiveness was:

- the fraction of day that the airbase is open

By an open airbase we mean that there exists a minimum operating strip (MOS), say 1000m x 10m, for take off and landing.

A problem that had to be addressed by SD is this:

- what fraction of day open is acceptable?

3. Simulation modelling approach

Before this project started, there was made an attempt to demonstrate that SD could be applied in this field, too. This was done as an academic exercize with a very simple model. It proved the feasibility, but also showed that more detail and sophistication would be needed.

Thus a Monte Carlo simulation approach was selected for the problem, bearing in mind that dynamic behavior would be very important. This approach is somewhat similar to SD.

Like in SD, we model problems, rather than systems. This means that the problem in focus determines how to create the model which is best suited:

- what parts of the system should be included in the model
- what level of detail is appropriate
- how to make the necessary simplifications of the real system to allow modelling

A physical description would be appropriate to describe the runway bombing and repair in isolation. However, such a limited view would be inadequate. It was essential to include the right dynamic surroundings in order to achieve a reasonably correct hostile sortie stream against the airbases simulated. Orange sortie production capacity, Blue air defense aircraft, Blue ground based air defense and Orange attacks on Blue ground based air defense are essential elements. The impact of an open or closed runway will automatically be taken into account, also.

The system consists of Blue and Orange airbases, with aircraft which can be allocated to air defence, surface attack, and OCA. Blue is the defender and Orange the attacker. The problem is survivability and operability of Blue air bases. Runway bombing with rocket assisted bombs (or dispenser weapons) is the new factor to be assessed.

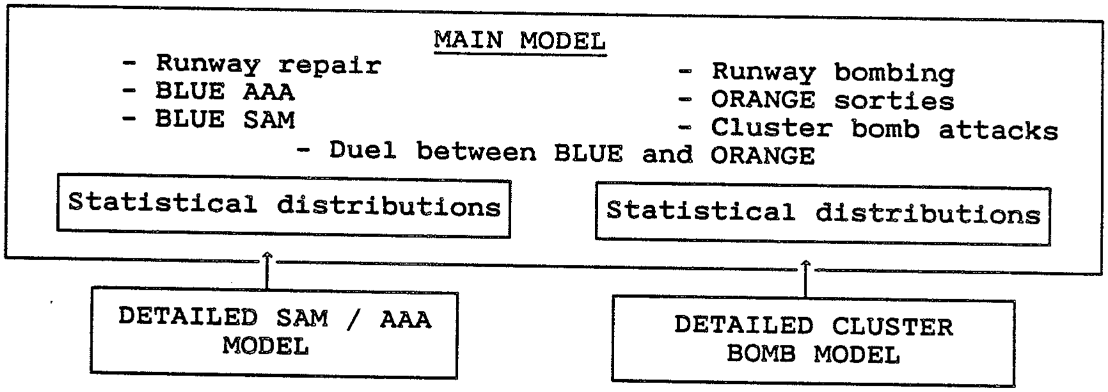

Fig 1 Model hierarchy

This problem has been addressed employing a hierarchy of Monte Carlo simulation models (fig 1). These models are:

- The Air Defense Model
- The Cluster Bomb Model
- The Airfield Attack Model (main model)

The two first models have been used separately in this (and other projects) to study limited problems in detail. Their strong point is their detailed realism. Because of their runtime and size it is not desirable and also somewhat unfeasible to try to integrate these models directly into the main model. Instead results from many runs with these two models have been aggregated into probability distributions which make sub models of the main model. Some technical characteristics of the models are listed in table 1.

	Airbase model (main model)	Cluster Model	Air Defense Model
Computer	ND-500	ND-500	CYBER 170 ND-500
Language	PASCAL	PASCAL	SIMULA
Size (lines)	6000	1500	16500
Typical run time (100 rep)	15 min	20 - 120 min	15 - 120 min

Table 1 Model characteristics

The Air Defense Model measures effectiveness of ground based air defense. The cluster bomb model computes damage from cluster bomb delivery. These models will not be discussed further in this paper.

The Airfield Attack Model has three main sections interacting:

- Orange sortie generation
- duel between Orange aircraft and Blue aircraft plus Blue ground based air defense
- runway bombing and repair

The runway bombing and repair section has most detail. It features:

- attack geometry
- runway geometry
- aim points
- MOS requirement
- delivery errors
- selection of repair site
- bomb functional probability
- repair logic

The bomb impact points are generated by random drawing of each delivery error from bivariate normal distributions. Each runway is divided into a grid of rectangles, and the number of craters is counted (accumulated) for each rectangle. The MOS will cover a number of rectangles in each direction. Finding the MOS with the least number of craters to repair is now a straighforward search process. With a suitable grid pattern, the method will be sufficiently accurate and have a quite acceptable runtime. An alternative method which is exact, requires storing of all crater positions plus extensive search through all possible MOS locations. That method requires longer run time, particularly when many craters are accumulated in repeated attacks. It is therefore suited for single attacks. The grid method was tested against the exact method with satisfactory results.

4. Dynamic surroundings

These elements are included based on previous experience and knowledge about air warfare simulation, as well as a general recognition of the vital importance of correct dynamic behavior. Feedback mechanisms are inherent in these elements. The inclusion of these elements could just as well be based on SD analysis, and perhaps some people would regard this as SD?

Attacks are generated from an Orange airbase which is considered as a queue system. Input is specification of attack times and desired number of aircraft against the airbases specified. If the specification cannot be met because there are too few aircraft ready, fewer aircraft will be sent out. An attack requires a minimum number of aircraft, and will be delayed if too few are available.

Attacking aircraft are engaged by:
- Blue air defense aircraft (if open runway)
- Blue surface-to-air missiles (SAM)
- Blue anti aircraft artillery (AAA)

Orange aircraft first attack the SAM units (optional), and then the runway. The number of aircraft shot down is drawn from the appropriate probability distribution. So is the number of SAM units destroyed, temporarily or permanently. A special model section integrates these numbers to a duel outcome, sorting out weapon delivery and "who kills first".

Blue attacks on Orange air bases have not been considered.

Several simplifications have been made. Reinforcements are not considered, RRR capacity is constant over time, and so on.

In many ways the model is coarse, but it aims to address the essentials of the problem.

5. Model output and usage

The model output is statistical tables, usually based on 100 replications. For each attack or each day the most important output is:
- fraction of day that the airbase is open
- probability of closed runway, before and after attack
- minimum number of craters to repair to open a MOS
- remaining repair kits (expendables to repair one crater)

There are also several other tables concerning Orange aircraft, sortie production, allocation, Blue SAM and AAA, as well as histograms of fraction of day open and reopening times for the airbase. These tables are mainly used to get a better understanding of model results and also to check that the results make sense.

The main results of the study can be catgorized as follows:
- finding minimum requirements for RRR and mine clearance capacities
- establishing requirements for protection of SAM units
- comparing the new FB threat to the old MB threat
- answering the questions about airbase operability under varying assumptions

6. System Dynamics approach

SD is an alternative approach. In this particular case, the SD analysis was performed at a rather late stage primarily adressing the problem of acceptable fraction of day open. No separate SD model was made as the feedback was already built into the existing model.

The main feedback loops are positive ones. These portray the attrition, and are slightly different from feedback loops usually encountered in SD. No growth can occur due to feedback. Assets are given at a fixed initial level, and will be lost at varying rates. Therefore these attrition loops portray only various degree of decay. However it is important to note that such attrition loops can drive the results in different direction, representing one side or the other winning.

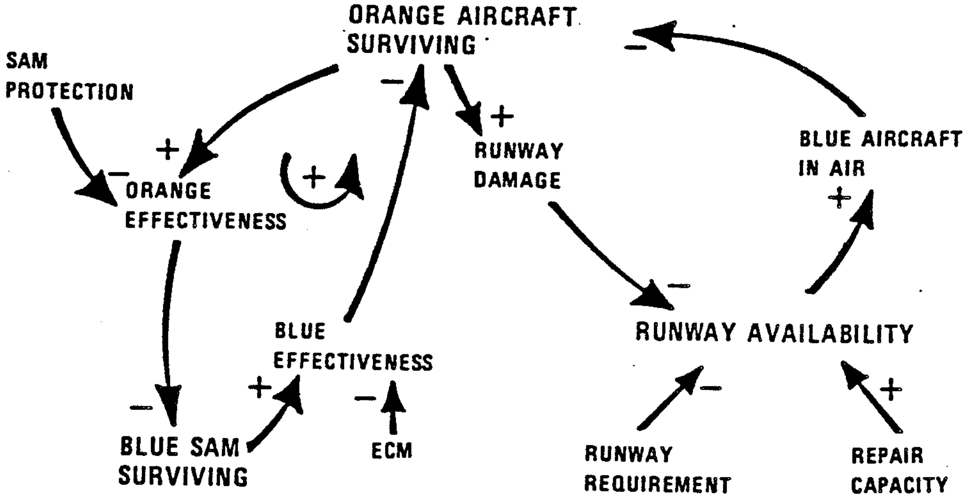

Fig 2 Main feedback loops

Feedback is strong, and the main loops are illustrated in fig 2.

There will of course also be negative loops which will slow the decay as assets are lost, eventually bringing the system into some stable state where one side may have the upper hand.

A typical attrition loop is positive and goes like this: When many aircraft attack the SAM systems, fewer SAMs survive, less aircraft are shot down, more aircraft survive to attack the SAMs next time. Should fewer aircraft attack, more SAMs would survive to shoot down more aircraft, and fewer aircraft would survive. This particular loop will stop playing after a while as the aircraft have completed the SAM neutralization phase.

The other loop concerns the Blue air defense aircraft. This is the feedback loop where stochastic phenomena may have a pronounced influence on the results.

When many Orange aircraft attack, the runway damage will be greater, the runway will be closed for a longer time, the chance of Blue air defense aircraft getting airborne is less, the more Orange attack aircraft survive. And of course similarly with fewer Orange aircraft attacking.

7. Some results

It often happens that results from model runs surprise you at first, but once you think it over, they are obvious. Here is a couple.

Both Blue aircraft and AAA can reduce the threat considerably over time. It is interesting to see how the mechanisms differ (fig 3). AAA engages only the aircraft bombing the runway, not the aircraft attacking the SAM units. This is due to the short AAA range and close-in deployment. Blue aircraft have therefore been able to reduce the threat considerably during the SAM neutralization phase. The result is higher runway availability during the most critical phase.

Simulation of dispenser attacks gave an amusing result (fig 4). Comparing a low and a higher RRR capacity (in terms of parallell repairs), we saw that the low one gave shorter RRR endurance (in terms of filling materials/top cover used). Why? Looking back at the feedback loops, it is not so surprising. With low RRR capacity, Blue aircraft might sometimes be trapped at a closed airbase, allowing Orange to deliver more bombs against the runway. This gave more craters to repair, higher chance of Blue aircraft being trapped, and even more craters through positive feedback. Therefore less endurance.

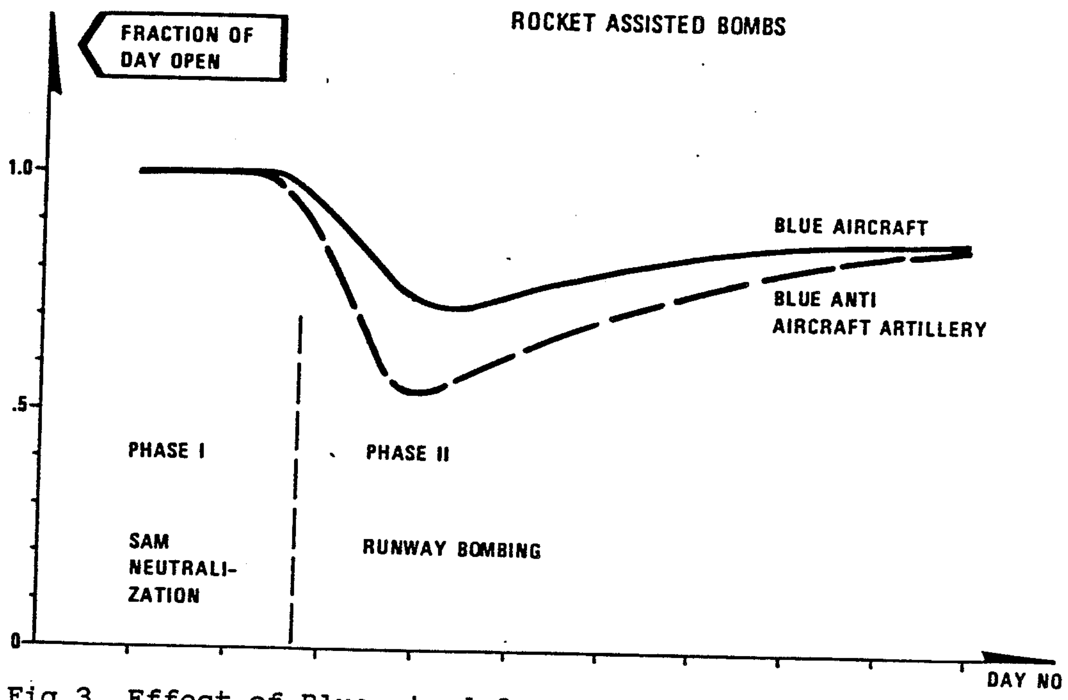

Fig 3 Effect of Blue air defense aircraft versus AAA

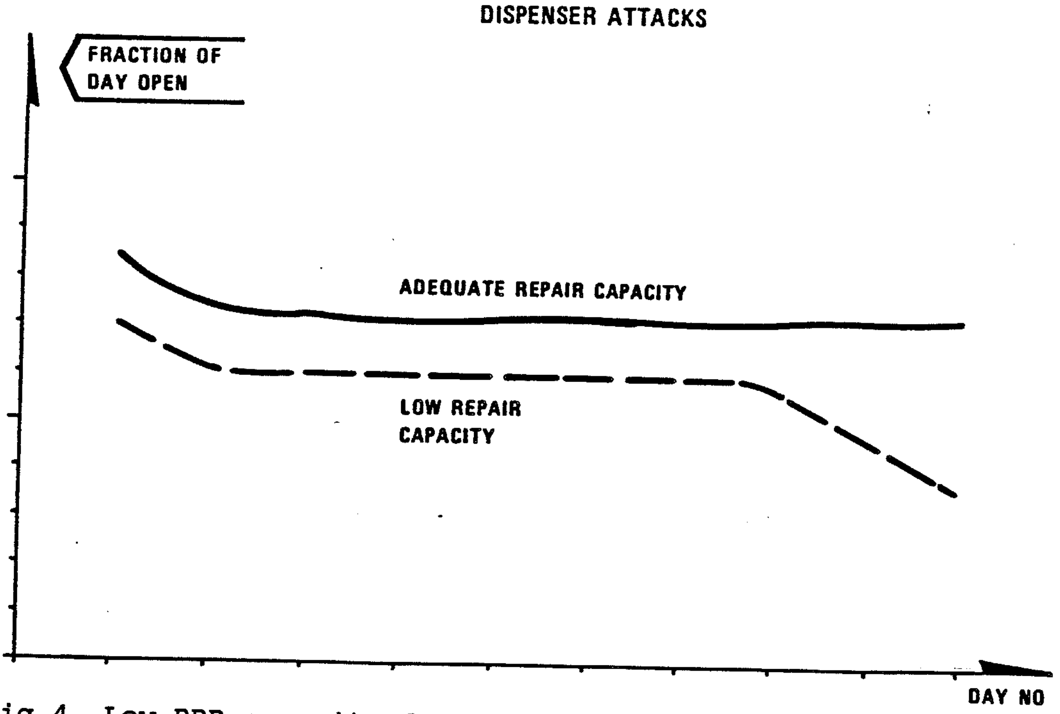

Fig 4 Low RRR capacity depletes crater repair materials faster

8. U-shaped distributions

When stochastic phenomena are introduced in attrition feedback loops, two different replications of the same simulation may have very different outcomes. This is the case with the runway/ Blue aircraft loop in fig 2. Suppose that RRR capacity is low, and Blue aircraft are very efficient. Now there is a risk that Blue aircraft may be pinned in so that the Blue aircraft loop does not play. This has a dramatic effect. If Orange happens to get the upper hand initially, Blue aircraft will be grounded, and Orange has won. On the other hand, if Blue initially gets the upper hand, many Orange aircraft will be shot down, so that Blue airbases will remain open, and Blue wins.

This phenomenon was observed at times in our results. And it created some peculiar problems for the interpretation of the results. Our simulation results were based on runs of 100 replications. In special situations we could observe the value 0 (closed all day) in many of the replications and the value 1 (open all day) in many others. And this would take place in the very same run. The main output figure, "fraction of day open", would be an average where 0 and 1 might dominate. In such situations, the average would have little or no significance. We have used the name "U-shaped distribution" for this extreme two peak distribution.

Is there some minimum acceptable limit for the average fraction of day open? The U-shaped distributions were used to analyze this problem. Fraction of day open was plotted against probability of the runway not being reopened before the next attack. The plot was based on a variety of parameter combinations from many different runs. It showed:

- with too low RRR capacity, Blue aircraft would frequently find themselves trapped when the airbase was not reopened before the next attack. This would occur at any value of (average) fraction of day open.
- with standard RRR capacity there was a minimum limit for fraction of day open. Above this limit there were few problems, below there were increasing problems of aircraft being trapped

It is important to recognize situations of this kind where the combination of feedback and randomness can make runs diverge in oposite directions. Such recognizing is only possible by SD thinking, but the phenomena can only occur when randomness is combined with feedback.

9. Risk of undesired development

In the example shown, the U-shaped distributions concern the risk of Blue aircraft being pinned in. Risk is often the key word when some initial randomness gets propagated by positive feedback. It could be the risk of a disaster, or it could merely be the risk of some undesirable development.

SD is well suited for as an instrument for policy making decisions. Utilizing randomness in SD models could certainly improve the ability to recognize potentially unwanted development, not only as the expected result, but also as a potential effect. It would also be possible to study conditions for avoiding or minimizing the likelihood of such undesirable development.

10. Conclusion

The conclusion of the project in the example was that RRR capacity in combination with air defense provides protection against the new threat. Correctly dimensioned RRR greatly reduces the uncertainty caused by feedback loops and randomness.

The combination of Monte Carlo simulation and System Dynamics is of mutual benefit. Feedback and randomness can lead to results with extreme two peak distributions rendering averages of little value. Using randomness in SD models can improve the ability to recognize the risk of potentially undesirable development.