

Uncertainty treatment in stock market modeling

Stanislaw Raczynski
Universidad Panamericana
Augusto Rodin 498, 03910 Mexico City, Mexico
e-mail: stanracz@netservice.com.mx

Abstract

A model of stock market dynamics is analyzed using a new approach to uncertainty treatment. The problem of uncertainty is formulated in deterministic way, using the differential inclusions as the main modeling tool. This results in the shape of the reachable set for the model trajectory, namely the possible extreme values of the stock demand and price. The results of example analysis are shown, where the uncertainty consists in erroneous or false agent's information about the actual demand. It is pointed out that while treating the uncertain parameters as random ones we cannot obtain the real shape of the model reachable set. This may affect financial planning decisions and our knowledge about the dynamic properties of the stock market.

INTRODUCTION

The main topic of this article is an approach to the uncertainty problem rather than the detailed stock market modeling. The model used here is a simple system dynamics stock market model for the short time market behavior. Models of such kind can provide important hints for financial planning and strategic decisions. Recall that the systems dynamics models are mostly continuous models that reflect certain global and averaged trajectories of a possible changes of the system variables. In models of such kind the human factor is strongly simplified. The behavior of a stock market agent is difficult to model and to predict. Consequently, the application of system dynamics modeling methodology to this case, as well as to other systems with human factor is doubtful. The other possible approach is agent-based simulation (do not confuse with stock market agents). However, it has been proved that system dynamics models do reflect some properties of the modelled systems. Anyway we should remember that not all what happens in the real world is described by differential equations.

The stock market model used in this paper was taken from literature. It is a simple model of a market with only one stock type. For this and similar models consult *Andresen (1999)*, *Minsky (1982)* or *Goodwin (1967)*, to mention only some of thousands of publications on stock market behavior. Some more qualitative comments on stock modeling and the use of models can be found in less academic sources like *Glassman (1998)*.

The main problem in marketing, economic, social and similar (soft) systems modeling is lack of exact information. This **uncertainty** in model data (initial conditions, parameters, external signals etc.) and even in model structure need a special treatment. The simplest way to get some information about the behavior of a system with uncertainty is to assume some variables to be subject to random changes and to see the resulting model trajectories. The common opinion is that the uncertainty can be treated using stochastic models and probabilistic methods. Note, however, that the very essential definition of uncertainty has nothing to do with stochastic models. It is an error to identify an *uncertain* variable

with a *random* variable. An uncertain variable or parameter has an uncertain value that may belong to some interval or satisfy some restrictions. It may have no probability distribution and could not be random at all. The approach to uncertainty treatment proposed here is based on **differential inclusions** and is **deterministic**.

THE MODEL

We consider a simple model of the dynamics of one stock type only. This is an Ordinary Differential Equation (ODE) model like the models used in the System Dynamics approach.

Let p be the current current market price, and p_r be the real value of the stock. We will denote by n the current demand of the stock expressed in number of units. Suppose that this demand is the sum of the following components.

n_r - demand due to the agents being informed about the stock value.

n_b - demand due to the agents who observe the price increase/decrease rate and do their trading based on some kind of predictions. The subscript b stands for the "bandwagon effect". This means that the positive or negative price rate attracts increasing or decreasing numbers of agents, respectively.

n_e - demand due to erroneous information. This is the uncertain component of the demand n .

To find the model equations, observe the following facts. The demand n depends on the difference between the real and the current stock price. This difference should be expressed in relation to the price, so we assume that the demand can be calculated as follows.

$$n_r = A \frac{p_r - p}{p} \quad (1)$$

where A is a constant.

This demand, in turn, determines the price growth rate, so we have

$$\frac{dp/dt}{p} = r(t) = Bf(n) \quad (2)$$

where $f(n)$ is given as

$$f(n) = \begin{cases} n & \text{for } n > -I \\ -I & \text{for } n < -I \end{cases} \quad (3)$$

where B is a constant and I is the total number of stock issued. The function f is simply a saturation. This means that the surplus of stocks (which results in negative demand) cannot be greater than I .

The component n_b that determines the "bandwagon effect" depends on the price increase rate. This reaction of the agents is not immediate and is subject to some inertia. We shall use here a simplest way to represent this, supposing that

$$n_b(s) = G(s)r(s) \quad (4)$$

where

$$G(s) = \frac{C}{1 + Ts}$$

Here s is the differentiation operator and $G(s)$ is a first order transfer function. The equation (4) implies the following.

$$dn_b / dt = (CBf(n) - n_b) / T \quad (5)$$

The equations (2) and (5) describe the dynamics of the model. It is a set of two ordinary, non-linear differential equations that can be easily solved using any continuous simulation tool. On figure 1 we can see an example of possible changes of the demand during two trading days. This trajectory was obtained using the PASION simulation system with the following parameters.

$$I = 10000, \quad T = 0.005, \quad p_r = 10, \quad A = 476.2, \quad B = 0.00007, \quad C = 14200.$$

The uncertain (erroneous) component n_e was supposed to belong to the interval $[-500,500]$. The above value of the "bandwagon" time-constant T can be less than assumed. The value of 0.005 was chosen to slow down the oscillations and make the the example trajectory more illustrative. The model initial conditions for the trajectories of figure 1 and for all other figures were $p(0)=8, n_b(0) = 0$ which means that we start with undervaluated stock, that generates a positive demand.

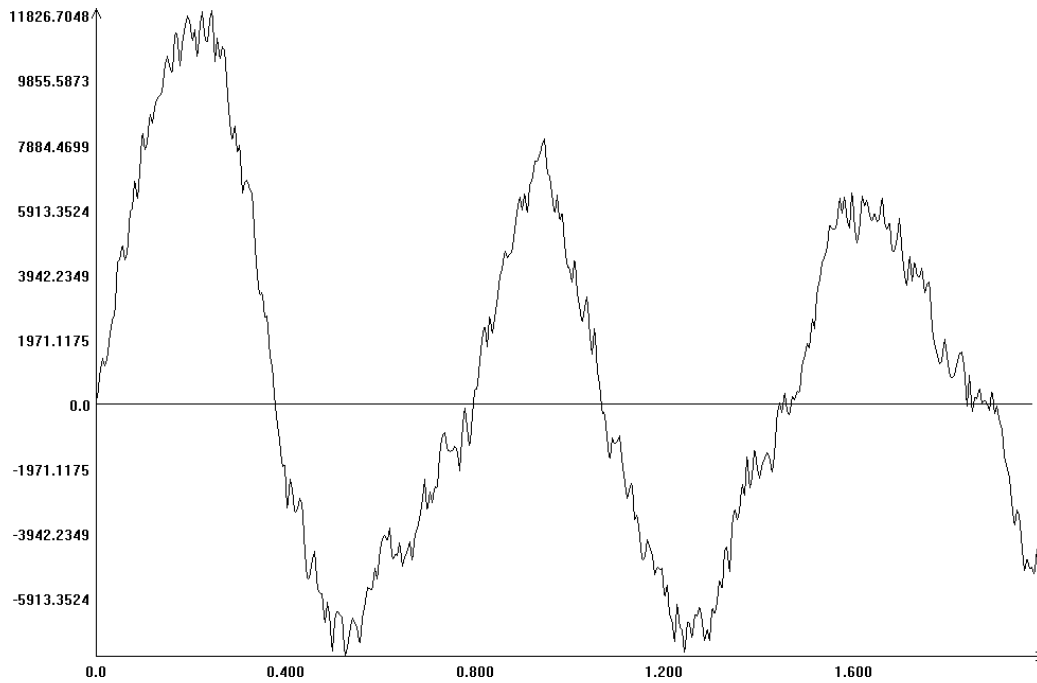


Figure 1. Possible changes in demand during two trading days.

THE UNCERTAINTY

As stated before, the component n_e of the estimated demand represents the erroneous data. To obtain the trajectory of figure 1 this component was assumed to be a random variable taken from the interval $[-500,500]$. This is a common approach in uncertainty treatment. Treating the uncertain parameters as random ones we can obtain a possible model trajectories, or carry out various statistical

analysis over a sets of hundreds or thousands of integrated trajectories. However, note that **uncertainty** should not be confused with **randomness**. First of all, to consider a variable as random, you must know that it is really random. If so, you must know something about its probabilistic properties, to be able to generate it. In the case of dynamic systems it is not only the probability distribution, but also the spectral density, mutual correlations and other stuff hardly known in practical applications. On the other hand, if a model variable is said to be uncertain, we only assume some interval (maybe infinite), where it belongs and nothing more. The result of the uncertainty analysis should be the reachable set for the model trajectories. Note that **such uncertainty treatment is deterministic**. Other reason to treat the uncertain variables in a non-probabilistic way is that such analysis gives us information about possible extreme values (recall the "law of Murphy"). This also may be useful if we expect that the uncertain variables could be intensionally generated to move the system to the extreme values (manipulated and false information).

For example, looking at our model, given by the equations (2) and (5) we can see one uncertain variable, namely n_e . In the vector form our model is described by the equation

$$\begin{aligned} dx/dt &= f(x, n_e) \\ n_e &\in [-500, 500] \end{aligned} \tag{6}$$

where x is the state vector $x = (p, n_b)$ and f is a vector-valued fuction that includes the two right-hand sides of the equations. We do not indicate constant model parameters in the arguments of f . However, n_e appears in the right-hand side of (6) because it is a variable and not a fixed parameter. The equation (6) can be written as follows.

$$dx/dx = F(x)$$

where F is a set defined by f , when n_e takes the values from the interval $[-500, 500]$. What we obtained is a *differential inclusion* instead of a differential equation. This is the proper way to treat the dynamic uncertainty. The solution to a differential inclusion is the reachable set, that is the set where all model trajectories must belong. This is exactly what we need as the result of the uncertainty analysis, and not particular model trajectories. Note that this problem statement is completely deterministic.

The solution to a DI is the reachable set for the possible system trajectories that is exactly the solution to our uncertainty problem. In this very natural way the uncertainty in dynamic system modeling leads to differential inclusions as a corresponding mathematical tool. Note that this tool is known for about 70 years and that there is wide literature available on the DIs theory and applications. The first works have been published in 1931-32 by Marchaud and Zaremba. They used the terms "contingent" or "paratingent" equations. Later, in 1960-70, T. Wazewski and his collaborators published a series of works, referring to the DIs as orientor conditions and orientor fields. As always occurs with new theories, their works received sever criticism, mainly from some physicists who claimed that it is a stupid way of wasting time while dealing with so abstract an useless theories. Fortunately, the authors did not abandon the idea and developed the elemental theory of differential inclusions. In the decade 1930-40 such problems as the existence and properties of the solutions to the DIs have been solved in the finite-dimensional space. After this, many works appear on DIs in more abstract, infinite-dimensional spaces. Within few years after the first publications, the DIs resulted to be the basic tool in the optimal control theory. Recall that optimal trajectories of a dynamic system are

those that lay on the boundary of the system reachable set. In the works of Pontragin, Markus and Lee, Bellman and many others, one of the fundamental problems are the properties of the reachable sets.

DIFFERENTIAL INCLUSION SOLVER

A differential inclusion (DI) is a generalization of an ordinary differential equation (ODE). In fact, an ODE is a special case of a DI, where the right-hand F is a one-point set. One could expect that a solution algorithm for a DI might be obtained as some extension of a known algorithms for the ODEs. Unfortunately, this is not the case. First of all, the solution to a DI is a set. Namely, it is a set in the time-state space, where the graphs of all possible trajectories of a DI are included. Finding the boundary of such set (named *reachable set*, or *emission zone* as in the works of Zaremba and Wazewski) is not an easy task. I will not discuss here more theoretical details about the DIs. A more extended survey can be found in *Raczynski, (1996)*. An excellent book on theoretic background was written by *Aubin and Cellina, (1984)*. Other fundamental publications are those of *Zaremba, (1936)* and *Wazewski, (1961)*.

One of the properties of the reachable set is the fact that if a trajectory reaches a point on the boundary of the RS at the final time, then its entire graph must belong to the RS. This fact is well known and used in the optimal control theory. Observe that any trajectory that reaches a point on the boundary of the RS is optimal in some sense. Such trajectories can be calculated using several methods, the main one being the *Maximum Principle of Pontriagin* (consult *Lee and Markus, (1967)*). This can be used to construct an algorithm for RS determination. If we can calculate a sufficient number of trajectories that scan the RS boundary, then we can see its shape. We will not discuss here the theoretical issues related to the DIs, which are complicated and need rather a book than a short article. In few words, the solving algorithm generates trajectories that scan the boundary of the reachable set. On each of such trajectories an expression known as *hamiltonian* is maximized with respect to a *control variable* (in our case n_e). Integrating a sufficient number of such trajectories we can obtain a good approximation of the reachable set. The DI solver based on the above principles has been presented in several articles; consult *Raczynski, (1996)* for more detail.

One could expect that the reachable set of a DI can be obtained by a simple random shooting, that is, by generating n_e randomly and then looking for the boundary of the resulting points reached by the trajectories. Unfortunately, this is not the case, except perhaps some very simple and trivial cases. What we obtain by such *primitive random shooting* is a cluster of trajectories in a small region that has little to do with the true shape of the reachable set, even if with great number of calculated trajectories.

SOME RESULTS

Figure 2 shows the solution to our differential inclusion at the end of a one-day trading. The dotted contour shows the boundary of the reachable set, that is the boundary of the set where the model trajectories must belong, on the price-demand plane. This contour was obtained by storing about 500 model trajectories. To see how useless is a primitive random shooting method mentioned before, the figure also shows the result of such shooting with 10000 trajectories integrated (a small cluster of pixels inside the reachable set). The random values of n_e in this primitive shooting were generated on the boundary of the allowed interval $[-500,500]$. While generating n_e randomly from inside of this interval, the cluster is even smaller. This does not mean that the computing time needed to solve the DI is 20 times smaller compared to primitive shooting. The point is that the primitive shooting provides no solution at all. On the other hand, the DI solver is rather slow, because of the complexity of the

algorithm that needs the hamiltonian to be maximized on each integration step. In the presented case, about 10 minutes of computing time was necessary to get the solution, using a 450Mhz PC. Figure 3 shows a 3D image of the reachable set.

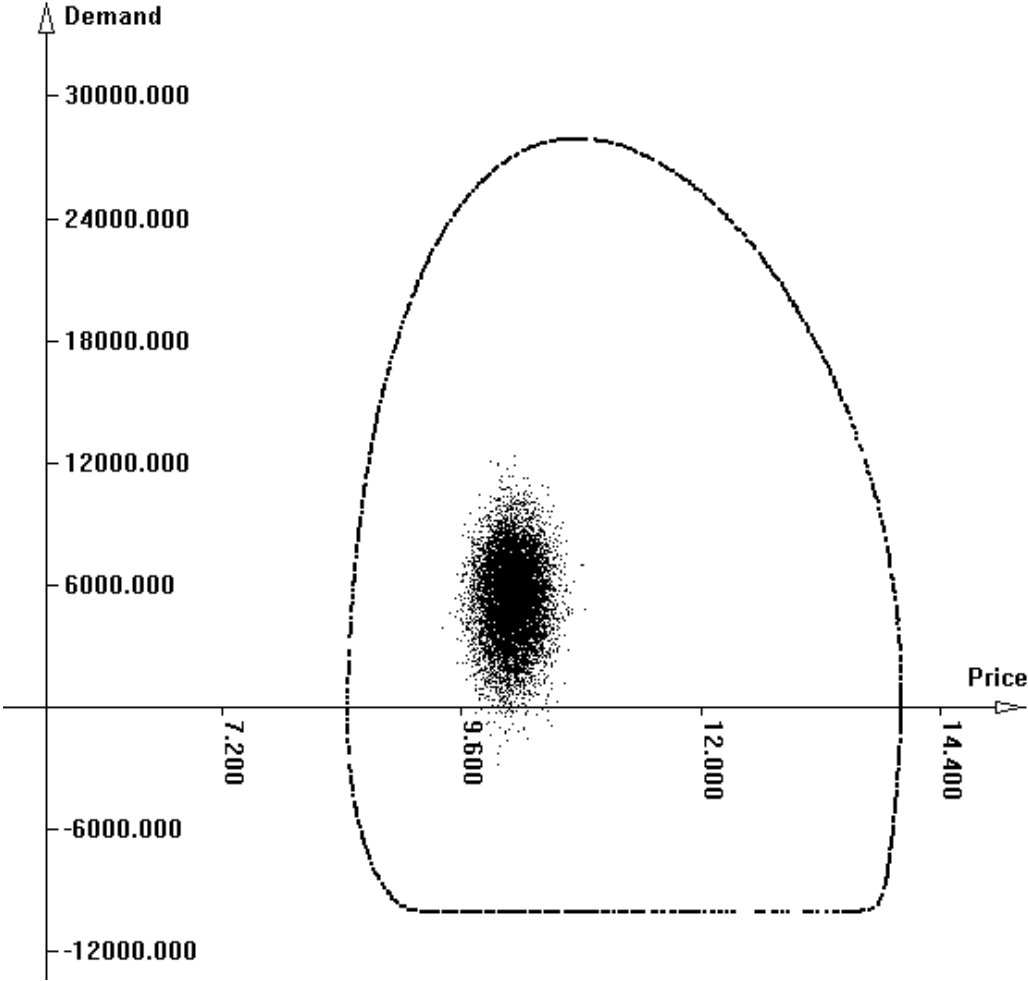


Figure 2. The reachable set at the price-demand plane. A small cluster of points inside the set was obtained with a primitive random shooting.

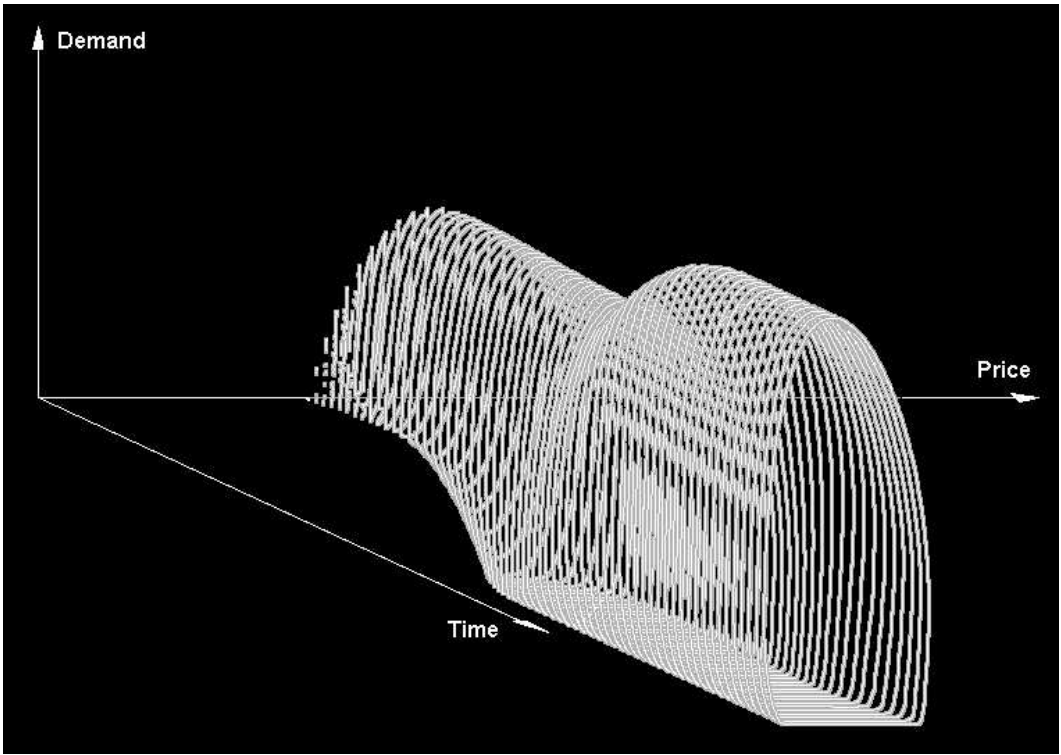


Figure 3. The 3D image of the reachable set.

The trajectories that determine the reachable set are oscillating around the boundary of the reachable set. These non-linear oscillations suggest that the extreme points of the reachable set boundary are reached when the model enters in some kind of "resonance". This is hardly possible with a random excitation, but quite possible when the uncertain parameter is changed intentionally to reach the boundary or extreme points. Figure 4 shows some of such trajectories randomly selected. Note that those are not random trajectories; only their selection is random. Each of the trajectories of figure 4 is a 2D projection of a trajectory that lies on the reachable set boundary.

Finally, figure 5 shows the projection of the reachable set on the time-demand plane.

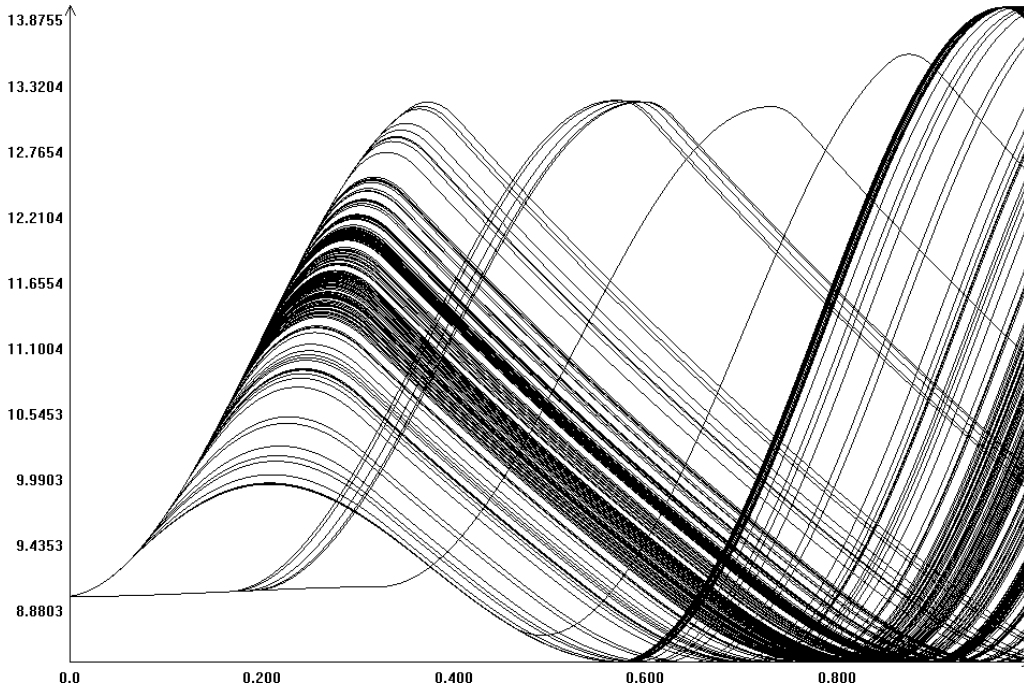


Figure 4. Some randomly selected model trajectories that scan the boundary of the reachable set.

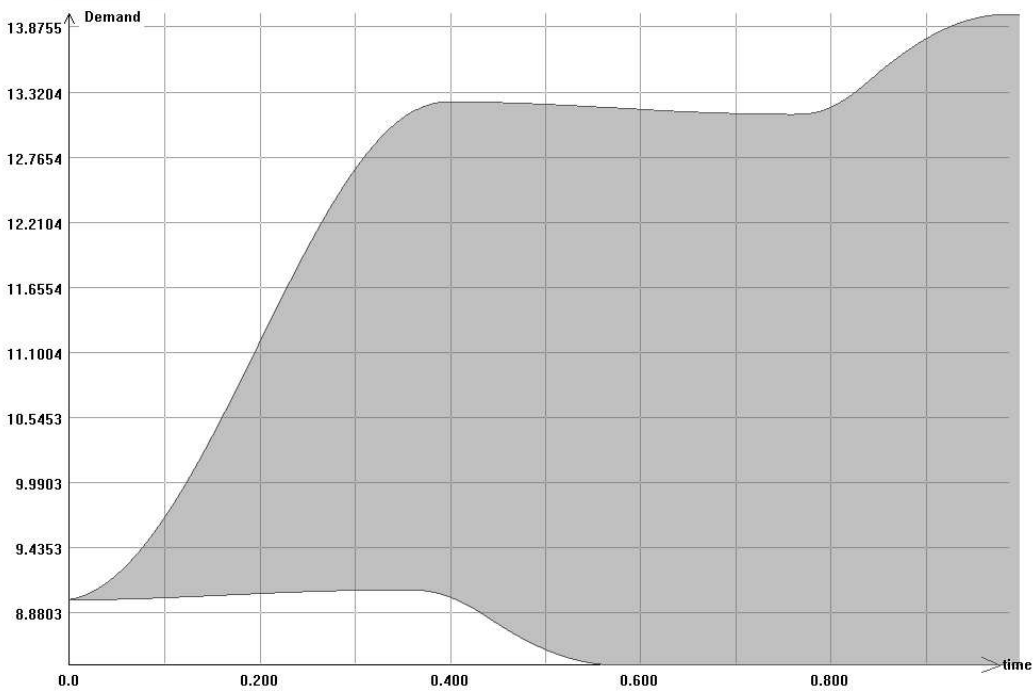


Figure 5. The projection of the reachable set on the time-demand plane

CONCLUSIONS

The main point of this article is the application of the differential inclusions to the problem of uncertainty in dynamical systems. The uncertainty problem is deterministic. If we treat the uncertain parameters as random ones, we could obtain very poor estimates of the possible extreme values that the model variable can reach due to the uncertainty. The presented differential inclusion solver works quite well, though it is still under construction. The model of the stock market dynamics is a good example of uncertainty and may provide interesting information about possible stock market behavior and can be used in financial planning.

REFERENCES

- Aubin J.P., Cellina A.**, 1984, *Differential Inclusions*, Springer-Verlag.
- Andresen, T.**, 1999, *The Dynamics of Long-range Financial Accumulation and Crisis*, Nonlinear Dynamics, Psychology, and Life Sciences, vol.3 no.2.
- Glassman K,J.**, *Trying to Crack the Code; Computer Models Lag as Stock Pickers*, The Washington Post, October 15,1998
- Goodwin R.M.**, 1967, *A Growth Cycle*, in Capitalism and Economic Growth, Cambridge University Press
- Lee E.B., Markus L.**, 1967, *Foundations of Optimal Control Theory*, Wiley, New York.
- Minsky H.**, 1982, *The Financial Instability Hypothesis: An Interpretation of Keynes and Alternative 'Standard Theory'*, Inflation, in Recession and Economic Policy, Wheatsheaf, Sussex
- Raczynski,S.**, 1974, *On the determination of the reachable sets and optimal control by the random method*, Proceedings of the IFAC Symposium on Optimization Methods, Varna, Bulgaria, 1974.
- Raczynski S.**, 1996, *Differential Inclusions in System Simulation*, Transactions of The Society for Computer Simulation, vol. 13, no.1, pp 47-54.
- Wazewski T.**, 1961, *Sur une condition equivalente a l'equation au contingent*, Bull. Acad. Polon. Sci. Ser. Math. Astronom. Phys., 9 no. 12, Warszawa.
- Zaremba S.K.**, 1936, *Sur les equations au paratingent*, Bull. Sci. Math., 60 (1936)