Uncertainty in Cost Performance as a Function of the Cusp Catastrophe in the NASA Program Performance Management System

Jonathan E. Butner\textsuperscript{1}, H. Lawrence Dyer\textsuperscript{2}, Thomas S. Malloy\textsuperscript{1}, Lucy V. Kranz\textsuperscript{2}

\textsuperscript{1} University of Utah, Department of Psychology
\textsuperscript{2} NASA, Multi-Purpose Crew Vehicle Program (MPCV) Program, Program Planning and Control (PP&C) Office

KEYWORDS: Cusp Catastrophe, Multilevel Model, Performance Management System, Cost Performance, Uncertainty

RUNNING HEAD: Uncertainty in the NASA Program Management System
Abstract

Cost performance by NASA programs can generate two stable states within the organization system used for program performance management. These states could be stable individually or show multistability with cusp-like behavior. To test this, we constructed a cusp catastrophe model in a multilevel modeling (HLM) procedure on the cost components of an eleven-month period of one NASA program. HLM was used for its ability to estimate random effects as replacements for unknown control parameters. The cusp catastrophe model was a good description of the data and there was some evidence that the overall budget size functioned as a control parameter. Results are presented in terms of three different forms of uncertainty: attractor strength, unistability/multistability, and control parameters. Recommendations and future directions are focused towards understanding the cause of uncertainty in complex management systems.
NASA is a highly complex organization with a hazardous mission: its core processes are tightly coupled; their interactions are not wholly predictable; and failure is enormously costly in lives, resources, and national stature (Mahler, 2009). Managing safety and cost are paramount, but the inherent uncertainty in the complexity of its technology systems, organization systems and missions has led to loss of life and notorious budget overruns and schedule delays (NRC, 2010).

The task for NASA Programs is to develop and manage technology systems that are technical, complicated, and push the boundaries of technology. This often involves a large number of interacting parts that must overcome unforeseen obstacles and function in both anticipated and unanticipated ways (Gehman, 2003). The components of a flight system, for example, fit together like a puzzle forming a tightly coupled system where no piece can afford to fail lest the entire project fails (Perrow, 1999).

Like many organizations, the organization systems that NASA has evolved for producing and operating its technology systems are emergent (Axelrod & Cohen, 1999). It is an organization of organizations that are hierarchical with many interrelated and interdependent components. When interrelated and interdependent components are integrated into a larger system, unanticipated interactions can occur; catastrophe can happen (Guastello, 1995). The emergent organization can be beneficial for NASA Program goals, but it can also be an obstacle.

The organization system NASA uses to control the production and operation of technology systems is termed the NASA Program Performance Management System. The system has a history of cost-overrun and schedule-delay performance escapes in building flight products (NRC, 2010; Young, 2001). In an effort to better understand uncertainty as a contributing cause of performance escapes within this system, we tested a cusp catastrophe
model on cost performance for one NASA program. We argue that the cusp catastrophe model is particularly appropriate due to its potential for multistability, and varying attractor strengths as a function of control parameters. In this case, the control parameters were unknown and random effects within multilevel modeling were utilized as a surrogate – a new approach for estimating catastrophe models. The results are discussed in terms of implications for the NASA Program Performance Management System providing the data, but also apply to any heterogeneous system where organizational complexity and performance uncertainty are involved.

The NASA Program Performance Management System

The NASA Program Performance Management system is complex. The chain of Agency, governmental, commercial and academic suppliers is extensive. Agency participants include Headquarters Offices for budget, finance, engineering and safety; Mission Directorate and Enterprise Offices for requirements and governance; Centers for infrastructure and governance; the Inspector General for oversight; Program Offices for acquisition development management; and participating organizations at multiple centers for providing government furnished equipment and services. The Government Accountability Office provides oversight. The Defense Contract Management Agency provides quality assurance. The system also includes representatives from other U.S. governmental agencies, from international governmental agencies, as well as academic and commercial entities. Internationally, space agencies from foreign governments such as in Europe, Russia, and Japan can be partners—each comprised of multiple constituents, stakeholders and contractors. Commercial entities include a hierarchy of aerospace contractors together with their subcontractors and vendors. Further, the performance management system adapts over time, generating additional complexity and
constraints on the system, as each newly elected Administration resets space policy and administers the Agency accordingly.

Program management coordinates the work of this interrelated and interdependent supply chain of participants to produce a technology system within cost and time constraints. And while the overarching goal of producing/managing flight systems is shared across the varying agencies, each agency also has their own, often conflicting, agendas. The resultant patterns of technology development emerge from the interactions of these many constituents within the limitations of some overarching, agreed upon, policies (that have also emerged over years of NASA management). For example, technology system development implements a multi-year plan operating over two life cycle phases: Formulation and Implementation. In the Formulation phase, system content is planned and designed, and technology is developed. In the Implementation phase, system content is developed, integrated into flight systems and elements, and tested. These life cycle phases constrain the goals of the various constituent agency participants in a synergetic fashion (Haken, 1980).

A Formulation phase is characterized by change, as alternatives are identified, trades are analyzed, concepts are tested, and technology is developed. Change is frequent and a program has the ability to defer work product for budget, in response to internal and external factors, consistent with management approvals. The objective is to stabilize the design such that it can be manufactured within the cost and time resources planned for its production. In an Implementation phase, change is not good and flexibility to defer work product for budget is constrained to allocated reserves (e.g., liens for cost and margin for schedule): the objective is to accomplish the multi-year build within the cost and schedule resources agreed to, and made
available at the beginning of the phase. Deviations beyond reserves are cost overrun and schedule delay performance escapes.

Programs that manage the development of technology systems have a history of uncontrolled growth resulting in cost overruns and schedule delays (GAO, 2013). Performance management in any life cycle phase is comparing actual performance to planned performance and taking corrective action to stay on target. Periodically, measurements of cost, work product and risk are collected and analyzed to characterize performance and to support decision-making. The relationship between cost, work product, and risk is prescribed by management in the multi-year program plan and in each annual Execution plan. It is not mathematical; and the exact relationship is never known (Lewis, 2005). Program performance is obtained by ‘integrating’ individual cost and work product variance reports, risk metrics, and other information across the supply chain. However, because of uncertainty resulting from complexity, program performance cannot be accurately assessed simply by aggregating supplier performance: Uncertainty is a product of coupled, context-dependent, nonlinear interactions, and outcomes cannot be obtained by summing outcomes of its constituent parts (Holland 1998). A capability to account for uncertainty is needed.

Phase Portrait

Performance data produced by a management system can be used to establish characteristics of that system. At any point in time, for example, reported values for cost and work product could each be ahead, behind or on target for its planned values. Therefore, the state of some ongoing process can be characterized in terms of its cost/budget and work product. Six years of budget data along with cost and work product (schedule and technical) performance
data for the Multi-Purpose Crew Vehicle (MPCV) Program were qualitatively examined for patterns. The MPCV Program is in the Formulation phase of its life cycle.

Budgets are appropriated and allocated annually; and during a Formulation phase monthly costs are re-planned annually. Costs are easy to measure as they are defined numerically. Costs are reported monthly by supplier and will usually appear to be ‘on track’ with the matching budget. This creates a coupled system where the actual costs from month to month will closely match the budget.

Work product is a combination of technical and schedule data. Work product status and performance are hard to measure as it is difficult to assess how close an activity is to completion until it is actually completed; especially when schedule has been traded for budget or work product includes newly developed or recently matured technology.

Figure 1 illustrates the hypothetical phase portrait and its manifold for the MPCV Program Performance Management System for an activity generated from our qualitative examination. The monthly slice shown at the bottom of Figure 1 indicates that an activity may be in some combination of ahead (underspent), behind (overspent), or on budget target and ahead, behind, or on schedule. As we will show by walking through scenarios in this figure, actual cost vs. budget can match up by using work product as a sink to absorb problems. This makes it possible for a program to appear ‘on track’ for budgetary purposes while being behind in terms of work productivity under the guise that the work can be made up (which is probably sometimes true, but not always) and reports of work product being more ephemeral in comparison to budget. Thus, problems do not appear in budget until there is a catastrophic shift that can no longer be masked.
To make this more explicit, the central balance point in Figure 1 is labeled as a combination of on schedule and on budget; and the four quadrant triangular areas indicate different types of divergence from on schedule and on budget. The central balance point can be thought of as the ideal compromise between managers (who want the work done quickly and inexpensively) and suppliers (who want to complete the work while maximizing income) – an attractor. Each of the four quadrants depict different divergences from the balance point (due to perturbations of unexpected problems and timely solutions). Below we explore scenarios for each of the four quadrants. These will identify a second attractor at being on target for budget and behind in schedule. It will also identify a third attractor indicative of the budget overruns where the system moves beyond these two attractors. It is this third attractor that we believe characterizes performance escapes as overruns and delays.

In Figure 1, we have labeled different areas of the phase portrait to illustrate what might be occurring. Under complacency (ahead of schedule but having underspent; lower right quadrant) regular billing will bring the program costs to the central balance point. As slowing down work allows a program that has saved money to return to one that is on budget. In the underactive region (behind on schedule but having under spent; lower left quadrant) a program that begins the month behind schedule but with money to spare can easily transition to one that is behind schedule and apparently on budget; this is due to the ease of reporting accurate cost numbers combined with the difficulty of determining how far behind a program is in achieving scheduled work. We call this attractive deferment and it is indicated by a black oval (range of possible values for work product) on the horizontal axis of Figure 1 – the second attractor.

In the overactive region (ahead of schedule and overspent; upper right quadrant), where the program is ahead of schedule but behind budget target, ceasing all work and billing will bring
it back to the central balance point (on budget, on schedule). Attractive deferment and balance point, as a set, create two stable regions (black oval and black circle) that are differentiable by work product, but not by cost numerically, since both are reported as on budget. In short, our examination of data and conversations with personnel indicate that monthly actual expenditure and monthly budgeted expenditure are nearly always the same during Formulation phase. We propose that work product (relative to schedule) is the sink that allows this to happen.

Figure 1. Theoretical Manifold for Budget and Work Product.

The fourth region labeled active repair (behind in work product and overspent; upper left quadrant) is of particular interest. Being behind schedule and behind in budget target differs from the others because no amount of billing or work inaction will bring the program back to the balance point if the activity has moved beyond some recoverable threshold. In essence, this is the
region of phase space where recovery may not be possible. This implies a tipping point where sudden large cost or budget and work product adjustments would need to be made to continue the planned trajectory of the system. Active repairs of one component of the program are indicative of catastrophic changes in the system producing that component, which will be apparent in both the schedule and costs. Moreover, active repairs on one component can potentially cascade to other components of the program given the inherent interdependencies of the components.

Figure 2 is a randomly selected sample of budget data for a series of activities from the NASA Orion Project (approximately 20% of the budget data available to us), each illustrated as a time series with month on the X axis and a difference between cost minus budget on the Y axis. To place all the time series on the same scale, each difference was divided by the total expected budget for the year for that activity. As one would expect from examining budget alone, a large portion of the time series hover around a value of zero representing the two stable regions from Figure 1. However, a small subset of budgets spiral off substantially from zero, consistent with the notion of being in the regime we called active repair. The key distinction we hope to identify is between those time series that will merely vary around zero vs those that spiral off vs. those that have high potential to spiral off even though they have not done so yet. To identify these scenarios from budget data, we turn to catastrophe models.
Uncertainty in the NASA Program Management System

Figure 2. Time series for a series of activities from roughly 20% of the sample mapping onto the different scenarios one might expect from the proposed model. Budget is scaled such that zero represents on budget, negative values are ahead budget target (below budget), and positive values are behind budget target. The units are in the total expected budget for that activity that year.

Cusp Catastrophe Model

The phase portrait shows the potential for the management system to suddenly change from one stable region to another. This is one of the many features of a cusp catastrophe model. The cusp catastrophe model has a relatively long history for depicting management systems (Dooley, 2009; Escartin, Ceja, Navarro, & Zapf, 2013; Guastello, 1981; Guastello, 1988; Guastello, 2013; Hanges, Braverman, & Rentch, 1991; Herbig, 1991; Sheridan & Abelson, 1983; Wagner & Huber, 2003; Yingluk & Tie-nan, 2011), and it seems particularly appropriate given the limitations in the program data available for analysis. The cusp catastrophe model is able to show both continuous and discontinuous change through time. It depicts a single system state under some circumstances (unistability) and two simultaneous states (multistability) under other circumstances (Gilmore, 1981; van der Mass & Molenaar, 1992).
Unistability and multistability are defined in the following way. Since we are analyzing cost only, attractive deferment and the balance point from Figure 1 are collapsed; they both, in terms of budget alone, are on-budget. This on-budget state creates one of the plausible states we will expect in our cusp catastrophe analyses. The other plausible state is the need for active repair (upper left quadrant), which, in terms of cost alone, is being behind in budget. Under multistability, a program can potentially be in either state at a given point in time. Under unistability, only one of the two states is plausible.

Referring back to Figure 2, most of these time series could be categorized as being in a unistable state around a value of zero – on budget. One time series (the one spiraling higher and higher in being behind in budget target) could also be described as unistable, but exceedingly off target. The multistable circumstance would be where one would expect to sometimes appear on target and other times appear extreme (as suggested by the time series which starts high but then suddenly comes back to being on target presumably because some repair action had been done such as revising the budget for the activity to account for its increased cost). Notably, some of the time series that appear unistable have the potential to be multistable, but have yet to display this behavior.

Estimating Catastrophe Models

Currently, there are two general approaches for examining the appropriateness of the cusp catastrophe model directly from data. The first approach is to utilize polynomial least squares regression equations for depicting a time series where change in outcome is treated as the dependent variable (See Guastello, 2011a). There are two advantages to the polynomial regression approach; (a) the equations have been fully realized for all the common catastrophe
models; (b) regression allows for simple comparisons between what would be predicted linearly vs. nonlinearly and even hierarchically between the various catastrophe models through effect size measure comparisons. The problem with the polynomial approach is that it requires between case data with only two time points to build a difference akin to pre-post logic and is not properly equipped for dealing with multiple measures within case unless there is only one time series.

The second method for testing the appropriateness of a catastrophe model directly from data involves estimating the probability density function (PDF) and relating it to the PDF one would expect based on catastrophes (Cobb, 1978; Cobb 1981; Cobb, Koppstein, & Chen, 1983; Cobb & Zacks, 1985; Guastello, 2002; Smerz & Guastello, 2008; van der Maas, Kolstein, & van der Plight, 2003). The PDF approaches have the advantage of being able to directly or indirectly test control parameters – key variables for a catastrophe model (more shortly). That is, the PDF method can either directly test the efficacy of certain variables that mitigate catastrophes, or assume they exist and then try multiple variables to see which (if any) function as these control parameters. However, the PDF approach has not been expanded to as many variations on the catastrophe model (Guastello, 2011b).

The problem we encounter in our study is that budget data has both within and between case qualities. Each activity has multiple measures giving it the properties of a time series. However, there are also multiple activities. So, any approach used needs to account for both within and between activity relationships. One option is to analyze each time series separately. This relies on each time series being long enough to provide an adequate test of the catastrophe model and could generate an underpowered and potentially biased estimation circumstance. Our approach presented here borrows from both procedures to estimate the viable nature of
catastrophe models through the utilization of multilevel modeling in combination with the polynomial regression method.

Multilevel modeling (HLM; also known as hierarchical linear modeling, mixed modeling and random coefficient modeling) is a maximum likelihood analytic procedure (some of the PDF approaches also utilize maximum likelihood) that expands regression to simultaneously account for both within and between case effects – it is the regression analog to repeated measures ANOVA (Raudenbush & Bryk, 2002), but it can also be thought of as being a form of the PDF approach, but where the PDF itself is a byproduct of the polynomial regression equation rather than directly expressed.

Our approach uses the equations from the polynomial regression method for catastrophe models and expands them to also account for between case differences. This logic stipulates that each time series (between cases) must conform to the same equation form. One can then choose to free up or limit additional parameters forcing the same regression equation terms across time series or allowing some to all parameters to vary across time series (between cases). Thus, our approach conforms to the polynomial method in that all the common catastrophe models can be directly expanded. However, it also conforms to the PDF approach through its utilization of maximum likelihood.

In the cusp catastrophe model in general, what differentiates multistability from unistability, as well as the asymmetry of strengths for these states, are a pair of control parameters (Guastello, 2011a) that are generally a-priori givens. However, in our case the control parameters are currently unknown. As already mentioned, some of the PDF methods can be used as a two-step process – first one tests for the conformity to the expected catastrophe model; then external variables can be added to test their adequacy as control parameters.
Uncertainty in the NASA Program Management System (Guastello, 2011b). Given that the HLM approach we present borrows from both methods, we will also be able to utilize this same two-step process – our control parameters are not known a-priori.

The cusp catastrophe model can be used to further differentiate three different types of uncertainty useful from a managerial perspective. First, each state (on-budget or overspent) implies a degree of variability around the state. For example, we can identify the set point for an attractor (the budgetary value the system seeks to settle towards on a monthly basis) and strength of that attractor (sometimes called the local Lyapunov or characteristic root; Abraham and Shaw, 1992). Second, when the system illustrates multistability, there can be uncertainty due to how frequently we might observe it switching between the two states (on-budget and overspent). This second form of uncertainty is characterized by the location and strength of the repeller between the states (sometimes called the tipping point, Gladwell, 2000). Third, control parameters could be changing over time within a component and between components. Thus, what might be unistable with strong attraction at one point in time (e.g., on-budget) might change to become multistable with weak attraction to each state (on-budget or overspent in budget). This approach is a direct byproduct for any dynamical model (can be employed for the other catastrophe testing methods too as pointed out by Guastello 2011a) and adds some useful information that is otherwise obscured in standard catastrophe testing.

To this end, we generated a time series of Formulation phase budget and cost data for multiple components of the MPCV Program Work Breakdown Structure (WBS). The multiple time series were tested against a cusp catastrophe model in a multilevel model to properly account for within and between component temporal relationships. Since the control parameters were unknown, we allowed for random effects across components where control parameters
would normally be included. We then tested the temporal proximity to the end of the budget cycle, and the total expected budget for a given system component as possible control parameters. We tested these as control parameters because timing in the budget cycle was necessary for us to complete the expected manifold and the total size the budget may over limit or depict unprecedented risks than smaller budget items.

Methods

We examined the monthly budget and cost data for 107 individual WBS line items for 11 months (excluding the last month of the budget cycle due to data availability). Selection of WBS line items was based on availability of recorded costs and planned budget for each month (though a month could be missing which was treated as missing data – HLM assumes missing at random, see Graham, 2009).

Our theory is based on the difference between budget (planned costs) and actual costs (reported expenditures). We therefore created a new variable at each time point for each WBS line item by subtracting the budget from the costs and dividing the difference by the budget for the last month prior to the end of the budget cycle (same as shown in Figure 2). This resulted in a number that was zero when actual costs and budget were the same, a negative number when costs were underspent and a positive number when costs were overspent, where the units were in number of total year budgets for that item. For example, a value of 1 indicated that the expenditure was twice the total budget for the year spent in a single month (a substantial overspending). This rescaling placed all the of the budget data on the same metric accounting for scaling differences similar to standardizing the data. We therefore utilize unstandardized estimates for our model.
Analytic Strategy

We conducted the analyses in both SPSS Mixed 19 and HLM 7 procedures (verifying the results across programs but also taking advantage of techniques that are easier to produce in one or the other program) using a Restricted Maximum Likelihood estimator. For each time series separately, we created a difference score of consecutive months of the rescaled budget data. These ten differences for each line item were then the dependent variables in a multilevel model accounting for differences within and across WBS line items resulting in 1070 unique instances of data.

Multilevel modeling is analogous to conducting separate regression analyses within each WBS line item over time, saving out the coefficients and then conducting regression analyses across line items using the coefficients as the dependent variables, though no individual analyses are actually done. Thus, it is commonly expressed in terms of levels of equations that parallel this analog. The level one (within a WBS line item) equation for a given line item, i, at a given point in time, t, was:

$$\Delta \text{Budget}_{it} = \beta_0 + \beta_1(Budget_{it}) + \beta_2(Budget_{it}^2) + \beta_3(Budget_{it}^3) + e_{it}$$ (1)

Equation (1) is a form of the polynomial regression approach to the cusp catastrophe excluding the terms relating to control parameters. The cusp catastrophe model includes two control parameters. One is a main effect while the other interacts with the linear effect of budget. To test a control parameter varying within a line item over time, it would be included as additional predictors or interactions with budget in the level one equation – the level one
equation would be identical to the equation used in the polynomial regression methodology.

Unfortunately, we were unsure of the control parameters within our model and thus generated an unconditional model initially with no expansion on the level one equation. Instead, we included random effects in the level two equations where control parameters were appropriate. The level 2 equations were:

\[
\begin{align*}
\beta_{0i} &= \gamma_{00} + \omega_{0i} \\
\beta_{1i} &= \gamma_{10} + \omega_{1i} \\
\beta_{2i} &= \gamma_{20} \\
\beta_{3i} &= \gamma_{30}
\end{align*}
\]  

How this reproduces the entire cusp catastrophe model (with some assumptions) can best be identified in the nested equation where the level one and two equations are combined:

\[
\Delta \text{Budget}_{it} = \gamma_{00} + \gamma_{10}(\text{Budget}_{it}) + \gamma_{20}(\text{Budget}^2_{it}) + \gamma_{30}(\text{Budget}^3_{it}) + \omega_{0i} + \omega_{1i}(\text{Budget}_{it}) + e_{it}
\]  

The gammas are the average coefficients across WBS line items (fixed effects). The omegas represent variability (random effects) in the coefficients across WBS line items. Thus, the omegas capture how variability in control parameters generate a main effect on change in planned expenditures (budget) and interact with the actual expenditure linear effect of budget. Further, the random effects represent variability in the control parameters across WBS line items, but not within them. This allows for a cusp catastrophe model where WBS line items are allowed to have hypothetically different values of control parameters, but are assumed to be constant within line items over time. Variation in control parameters within line items over time.
comes out as error in estimation unless the actual control parameter is added to the level one equation.

To help facilitate understanding how a random effect can be used to capture a control parameter, let’s visit a form of the cusp catastrophe in polynomial regression (Guastello, 2011a) expanded to include all lower order effects, as is appropriate for ordinary polynomial regression (Cohen, Cohen, West, and Aiken, 2003).

\[ \Delta Y_t = b_0 + b_1 Y_t + b_2 Y_t^2 + b_3 Y_t^3 + b_4(a)_t + b_5(b)_t + b_6(b)_t Y_t + e_t \]  

(4)

The two control parameters, are (a) the asymmetry parameter, and (b) the bifurcation parameter. Note that this equation is different from the polynomial and PDF catastrophe methodologies in that they always suppress the linear effect of \( Y \) (\( b_1 \)) for topological reasons. They also frequently suppress the quadratic relationship (\( b_2 \)) because it allows for gradient effects. We include this linear effect of \( Y \) and all lower order relationships because our approach is designed to function in raw metric. In regression (of which multilevel modeling uses the same interpretation rules) exclusion of lower order terms implies a nested relationship, which does not actually exist here. Thus, excluding these terms would enhance the estimation of the cubic relationship, but make the effects of the control parameters contingent on their scaling. Change the zero value of a control parameter and it would alter its predictive nature. Guastello (2011a) and others resolve this by standardizing all variables involved.

As with all analytic procedures predictors must show variation to be detected. Therefore within/across case values of (a) and (b) must be varying (hence the subscript of \( t \) for time in the polynomial regression circumstance). Focusing on the interaction between (b) and \( Y \) for a
moment – the bifurcation coefficient – if we follow the logic of simple slopes \( b_6 \) represents how changes in \( b \) alter \( b_1 \) – the main effect for \( Y \). Thus, identifying consistent variation in \( b_1 \) is indicative of the bifurcation effect even though \( b \) is not known. The random effect of \( \omega_{1i} \) captures consistent variation in the main effect of budget representing the bifurcation control parameter.

The same logic can be applied to the asymmetry parameter coefficient \( b_4 \) in regards to the intercept, \( b_0 \). As \( a \) varies, the intercept changes as a function of \( b_4 \), which would result in variation of the intercept. In our budget model \( \omega_{0i} \) is the extent to which there is variation in the intercept and thus represents the asymmetry control parameter even though \( a \) is not known.

Common practice for testing catastrophe models is to compare the incremental improvement in prediction from the linear equivalent model. Given that random effects can be used to capture control parameters, the choice for the comparative model may seem complex. We utilize two different linear comparisons to extricate possible effects. The first linear model is called a random intercept model where each individual gets an estimated mean over time and that mean has a random effect allowing it to differ across individuals. Equation 5 shows the random intercept in nested equation form.

\[
\Delta \text{Budget}_{it} = \gamma_{00} + \omega_{0i} + e_{it} \tag{5}
\]

Since change is the outcome in our equations, this model is equivalent to linear growth modeling.

Our second baseline model adds a linear effect of budget as a predictor, but no random effect of budget.
ΔBudget_{it} = γ_0 + γ_1(Budget_{it}) + ω_{it} + e_{it} \quad (6)

This is the equivalent of allowing for a single attractor state, but allowing the location of the attractor to vary across WBS line items. Note that this model already begins to impinge on some of the properties associated with cusp models, though lacks the notion of allowing for both continuous and sudden changes associated with catastrophes – the key nonlinear contribution.

Results

Table 1 contains the unconditional multilevel model results. All coefficients were significantly different from zero at alpha = .05, two-tailed. The fixed effects show the average equation while the random effects show how the intercept and coefficient for the main effect of budget change as a function of control parameter differences across line items. As mentioned earlier, HLM is a maximum likelihood procedure and thus would normally be limited in its ability to compare models. Since we are predicting changes akin to the polynomial regression methods, however, we can generate estimates for each data instance (both predicted and residuals). As an effect size measure, we calculated a form of pseudo-$R^2$ by saving out predicted values and residuals for each instance and calculating a sum of squares predicted over the sum of squares of predicted and residuals combined (Snijders & Bosker 2012). This approach is analogous to an $R^2$ from OLS regression in terms of its interpretation and metric, though imperfect because the predicted values generated from the multilevel model are known to not perfectly follow the parameter estimates when including random effects. Specifically, OLS estimates of the predicted values generate more variability than the random effects. So, most
HLM procedures utilize a Bayesian method for creating predicted values that are a weighted combination of the HLM fixed effects model and what would be observed if a regression was run within a specific case (line item here). These Bayes estimates tend to reproduce less variability than indicated by the random effects. Thus, the predicted model is not perfectly replicated in the predicted values. Our model accounted for 32.5% of the variance in budget changes. As a comparison, both the random intercept model and the model with the linear effect of budget and random intercept only account for 6% of the variance in budget changes.

<table>
<thead>
<tr>
<th>Table 1. Results from Unconditional Cusp Catastrophe Model with Random Effects for Control Parameters</th>
<th>95% CI Lower Bound</th>
<th>Estimate</th>
<th>95% CI Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.056</td>
<td>-0.035</td>
<td>-0.015</td>
</tr>
<tr>
<td>Budget</td>
<td>-0.197</td>
<td>-0.122</td>
<td>-0.046</td>
</tr>
<tr>
<td>Budget Squared</td>
<td>0.065</td>
<td>0.133</td>
<td>0.201</td>
</tr>
<tr>
<td>Budget Cubed</td>
<td>-0.042</td>
<td>-0.029</td>
<td>-0.016</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.004</td>
<td>0.007</td>
<td>0.011</td>
</tr>
<tr>
<td>Budget</td>
<td>0.032</td>
<td>0.054</td>
<td>0.092</td>
</tr>
<tr>
<td>Residual</td>
<td>0.029</td>
<td>0.032</td>
<td>0.035</td>
</tr>
</tbody>
</table>

1 All random effects are displayed as variances.

Figure 3 illustrates three different model results for three of the 107 WBS line items chosen as exemplars for each of the possible system states we would expect to observe. These graphs were generated by saving out the estimated Bayesian coefficients for each line item and graphing the equations. Figure 3a shows a case where there is a relatively strong single attractor at slightly below budget (the set point of an attractor is defined by the point at which change is zero with the slope at that point defining the strength of attraction as the characteristic root;
attractors always have negative slopes at the set point). Figure 3b shows two attractors, one at substantially below budget and one at over 4 times over the entire year’s budget. The repeller (or tipping point) in this case is slightly less than being over by the entire expected budget in a single month. Figure 3c shows a single attractor at more than 3 times over budget. Figures 3a and 3c represent the known forms of uncertainty where the line item generally appears on-budget (3a) and generally appears as over-budget (3c). Figure 3b on the other hand represents the potential for catastrophic uncertainty.

To characterize the commonality of each state, we calculated the set points and the slope at the set points from the estimated Bayesian coefficients. Figure 4 is a scatterplot of these values with different symbols for those set points when they were unistable (one set point) vs. multistable (three set points). On the x-axis is the set point expressed in units of yearly budgeted cost from -2.0 to 8.0. On the y-axis is the strength of the set point in terms of the characteristic root, where values above zero indicate a repeller and values below zero indicate an attractor. To
distinguish repellers from attractors, we added a horizontal line at a characteristic root of zero. Distance from this horizontal line indicates the strength where being close to zero is indicative of greater uncertainty or greater variability around the set point.

The unstable patterns (open circles) occur in two clusters. One cluster (toward the left of the graph) had set points near on-budget (zero on the x-axis); this cluster all had negative characteristic roots indicating they were all attractors, though the strengths of their attraction varied. The varying strengths of this cluster of near-budget attractors indicated varying degrees of uncertainty around those set points. There was also a second cluster of set points showing unistability (open circles) farther to the right on the x-axis; again these were always attractors. This second cluster was spread out between three and seven times the expected annual budget (spent in a single month) with stronger attraction (more negative characteristic root) as the attractor deviated more extremely from on budget (0 on the x-axis).

The dark triangles on the scatterplot demark the attractors and repellers for the multistable line items. Those line items with multistability (dark triangles) show one set of attractors around or below budget (between -2.0 and 0.0 on the x-axis), repellers (those dark triangle above the horizontal line) between 0 and 2 times the expected budget for the year on the x-axis, and another set of attractors between around 2 to 5 times annual budget.
Figure 4. Characteristic Roots (Y-axis) and Set Point values (X-axis) from each of the 107 WBS line items. Each is differentiated in terms of unistability vs. multistability. There are two clusters of attractors for the unistable line items, on budget and substantially over budget. The multistable line items have similar attractors with repellers between 1 and 3 times the expected budget for the entire year spent in a single month.

Returning to the over-budget unistable attractors (open circles), the near linear association between the characteristic root (strength of attraction) and the x-axis set point for the extreme attractors is particularly intriguing. This was unexpected. It is unknown as to the extent to which this is a function of the model itself (e.g. a complex byproduct of model assumptions and restrictions), or truly a function of budget occurrences. If this was an accurate description of cost performance, it suggests that risk characterizations of those line items known to be overspent might also be related to the multistability circumstance.
Control Parameters

In terms of the manifold illustrated in Figure 1, we considered that proximity of the analyzed-month to the end of year may function as one of the control parameters. The argument is that the budget may become more or less constrained as the end of the budget year approaches. We therefore expanded the level one equation to include month as a main effect and interaction with budget. Both were non-significant [main effect of month, F(1,1052.102) = 2.138, p=.144; interaction with budget, F(1,957.129) = 0.068, p=.795] providing no support for this claim. The rest of the model was unchanged.

We also tested the notion that larger or smaller budget items could be more (or less) at risk for the different types of uncertainty. We therefore re-ran the analyses using the total expected budget at the end of the cycle as a level 2 main effect and as an interaction with budget. The main effect was significant while the interaction was not [main effect of total expected budget, F(1,106.958)=8.222, p=.005; interaction with budget, F(1, 34.558)=1.838, p=.184]. Figure 4 illustrates that those WBS line items with larger expected budgets had greater potential for multi-stability (random effects are ignored in the figure). That is, items with larger expected budgets were more likely to have behavior that appeared on budget and then suddenly stabilized at appearing substantially overspent in budget. However, random effects remained (Variance for intercept after controlling for total budget expenditure = 0.006, Z=3.696, p<.001; Variance for main effect of budget after controlling for total budget expenditure = 0.053, Z=3.621, p<.001) suggesting that this may either be an indirect representative of the bifurcation control parameter (Chen, Stanton, Chen & Li, 2013) or other parameters remain. Something else must account for the remaining variability.
Figure 5. Predicted model of Change in Budget (Y-axis) as a function of budget (X-axis) and total Annual Expected Expenditures. At or below the mean in annual expected expenditures only crosses zero change (generating set points) once while 1 Sd above the mean generates a second attractor (generating three set points).

Discussion

The cusp catastrophe model provided promising results for depicting uncertainty in the NASA Program Performance Management System. As expected, there was evidence (see Figure 4) of three patterns: two unistable cases and one multistable case. The first pattern (near-budget attractor) is the ideal circumstance from a managerial viewpoint. Uncertainty regarding costs would involve actual expenditures varying around expected expenditures. At any given month, cost might go above or below what was expected, but any uncertainty in the state would be manageable because the costs would be attracted to return to the near on-budget set point. The
issue with this first state, as we will discuss below, is how to distinguish it from the multistability circumstance.

The second pattern (extremely overspent budget attractor) is less than ideal, but it is a known situation from a managerial standpoint (which came out in our interviews of managers). In this state, costs are attracted to a set point that is a multiple value of annual budget. Managerial uncertainty is due to variation around this overspent budget set point. This could occur when a program is underfunded and decisions about priority and timing have to be made.

The third pattern represents a form of uncertainty that is both problematic and challenging. In this state program costs are attracted to two stable set points, one on budget and the other well overspent in budget. There is a repeller between these two attractors so that the program costs tend to stay at whichever attractor they currently inhabit (on-budget or overspent in budget). This means that a program that appears to be on-track may suddenly switch to being extremely costly. This is a manager’s nightmare.

Knowing which of these three patterns the system is functioning is invaluable to managers; but being able to do so requires more extensive knowledge of control parameters. Mathematically, in a cusp catastrophe model, control parameters determine which system state will be expressed at a given point in time. For example, in pattern three above, if the control parameter that moved a system from on budget to overspent in budget were known and could be measured, then prediction would be possible and the effects of uncertainty could be managed. Ideally, a manager could not only predict occurrence but also manage the effects of uncertainty by adjusting control parameter values.

Since we did not have a-priori knowledge of control parameters, we attempted to explore some plausible candidates for control parameters by using random effects in the multilevel
model. We assumed random effects are normally distributed. Assumptions of normality are less than ideal in a nonlinear representation of phenomena akin to systems theory. However, it is not uncommon to have to make such assumptions for model estimation. Generally speaking, assuming normality when phenomena are not creates underpowered circumstances (Cohen, Cohen, West, & Aiken, 2003). That is, this random effects approach is less powerful than actually having a measure of the control parameters themselves.

Of the two variables tested as possible control parameters, only the total planned expenditure had any role. A larger budget led to more cases of multistability. This suggests that particularly big budget items may have the greatest risk for producing unexpected outcomes. This is directly contrary to the notion that a larger budget gives more cushion, but may be a function of the types of activities that get larger budgets within NASA. For example, larger budget activities may be ones that involve more technological risk and complexity where a greater portion of the activity involves untested advances necessary for the completion of the NASA project. To this end, total expected budget may be more of an indirect representation for one of the control parameters. This is consistent with our results of continuing to find a significant random effect after including total expected budget in the model. Thus, it is appropriate for managers to pay particularly close attention to the larger budget items, which is what one would expect from a manager. However, this approach will not necessarily identify when there is the potential for multistability due to its distal relationship.

Conclusion

Organization systems are complex because of the number and diversity of participants and the extent of interrelationships and interdependencies between them. The NASA Program
Performance Management System is a complex system. Because of its complexity, management outcomes are subject to uncertainty. Evidence of complexity and uncertainty were confirmed by analyzing budget and cost performance data for the MPCV Program. A phase portrait of the system was developed. It was used to map regions of performance outcomes, and to suggest a model for explaining and predicting outcomes so that the effects of uncertainty could be managed. A cusp catastrophe model was developed using Program data and making assumptions about control parameters. The model showed that program performance, influenced by the values of a control parameter, can oscillate between times of on-plan and extremely over plan. Performance escapes in the form of cost overruns and schedule delays result from off-plan performance.

The idea of control parameters is crucial. From a managerial standpoint, knowing the value of applicable control parameters in the cusp catastrophe model would help identify the conditions under which a program might be both currently stable yet vulnerable to extreme change in the near future. Knowledge of conditions would permit identification and corrective action. Two possible control parameters were tested with marginal success. Further work is required as understanding of control parameters is fundamental to being able to predict performance outcomes and manage effects.
References


GAO. (2013) *Assessment of Selected Large-Scale Projects*. GAO-13-276SP.


Author Note

All correspondence should be sent to the first author, Jonathan Butner at Department of Psychology, 380 S. 1530 E., Room 502, University of Utah. Salt Lake City UT 84112 or via email at jonathan.butner@psych.utah.edu.

This study was supported by grant NNX11AR71A from NASA Shared Services Center awarded to Jonathan Butner (PI). We thank Joel Cooper and Jake Jensen for their assistance with this project and Mike Geuss for providing comments on an early version of the manuscript.