Novel Schedule Forecasting for Low-Volume Highly-Complex New Product Development: Manufacturing Phase

Bruce Chehroudi, PhD	Jonathan Lam, PhD	Gus Benavides	Scott Morchower
ManTech International	US Space Force	Axient	ManTech International
El Segundo, CA	Los Angeles, CA	El Segundo, CA	El Segundo, CA
Bruce.chehroudi@mantech.com	jonathan.lam.1@spaceforce.mil	gustavo.benavides@axientcorp.com	scott.morchower@mantech.com

ABSTRACT

A novel schedule forecasting methodology (Nostradamus Objective) was developed to predict schedule risk of a unique, large-scale acquisition program that is characterized as low volume, highly complex, new product development (NPD) effort at its early stages. The goal was to produce objective evidence-based product-delivery schedule forecasts with high precision. Nostradamus was designed to model a manufacturer's accuracy in projecting it's product components completion dates of a recently-delivered unit (i.e., "*past information*"), and combining it with "*current information*" to make accurate/objective delivery date forecasts for subsequent units currently being manufactured. The approach, while does not use a Bayesian formulation, has a Bayesian-like strategy. The past information used conforms with the "reference class concept" described by Nobel Prize winning work of Daniel Kahneman and coworkers and possesses the highest similarity to products being manufactured. Additionally, the algorithm ranks a list of components that significantly affect product delivery dates. Hence, targeted measures can be taken to favorably affect the product delivery dates and reduce the overall project schedule risk. The product's major components, and associated estimated completion dates (ECDs), which are defined, determined, and provided by the manufacturer, constitute the current information. The current information and manufacturer's estimated product delivery date, collectively define the Line-of-Balance (LOB), which is provided at each Program Management Review. The "*Accuracy Level*" probability distribution function of the ECDs is defined and calculated for a most recently delivered product and utilized in a subsequent Monte Carlo simulation with the current information, to generate a product delivery failure probability (DFP). This DFP was used to generate the Nostradamus Objective's forecast of product delivery date at any probability level. Results of the tests indicated that, over a 2.5 year project duration, Nostradamus Objective achieved a time-averaged forecast imprecision value of -3%, as compared to 143% by the manufacturer, thereby providing reliable, precise, and consistent schedule assessments from which program risks can be identified and mitigated.

ABOUT THE AUTHORS

Dr. Chehroudi, accumulated years of technical/leadership experiences in different organizations. This includes, Mechanical Engineering Dept Head (Arkansas Tech Univ), Managing Director (Advanced Technology Consultants), Principal Scientist appointment at AFRL (ERCInc), Chief Scientist (RaytheonSTX), Visiting Technologist (Ford's Advanced Manufacturing Technology Development center), tenured Professor of Mechanical Engineering (Kettering University and University of Illinois), and Senior Research Fellow (Princeton University). He directed numerous interdisciplinary projects in chemically-reacting flows, combustion and pollutants emissions in engines, sustainable/alternative energy sources, distributed fuel ignition, material/fuel injection, advanced pollution-reduction technologies, propulsion concepts, gas turbine and liquid rocket engines, combustion instability, laser optical diagnostics, spectroscopy, supercritical fluids with applications in environmental and propulsion, advanced composites, and nanotechnology in propulsion. He has won many merit/leadership awards from SAE (*Arch. T. Colwell Merit Award*, *Forest R. McFarland Award*, *Outstanding Faculty Advisor*, *Sustained Leadership in Professional Service*), AIAA (*Best Publication Award of the Year*), AFRL (*Outstanding Technical Publication Award*, *STAR Team Award*), ILASS (*Marshall Award*), and has been a consultant for many organizations. He has PhD & MS in mechanical/aerospace engineering, MS in Finance/Management, senior AIAA Propellant/Combustion Committee member, and *AIAA Associate Fellow*. He delivered over 200 technical presentations, has over 150 publications with extensive experience in scientific, management, and finance areas.

Dr. Lam has a diverse technical background having started his academic career developing n-alkanethiol monolayers as chemical platforms for atomic force microscopy (University of California Los Angeles), and analysis of enzymatic activity in Catalina Kelp (University of Southern California). He continued his training in the biological sector with a B.S. in Biochemistry and minor in Economics (University of California San Diego, 2011), and a Ph.D. in Microbiology (The Ohio State University, 2018). During his academic career he taught multiple college courses, guiding and encouraging the next generation of STEM leaders to apply technical rigor to experimental design, to critically analyze data objectively, and to effectively present findings. His forte in project management and leadership led him to join the Space Systems Command at Los Angeles Air Force Base as an acquisition program manager (2019) and became responsible for the effective leadership of multi-disciplinary teams in high-visibility programs. As a servant leader, Dr. Lam encourages innovative thinking, adaptation to ambiguity, and application of objectives and key results to realize personal and professional growth of his teammates.

Gus Benavides has 15 years of technical & leadership experience in Aerospace and Software disciplines for United States (US) Space Force Satellite and Rocket programs for GPS and National Security Space Launch. On the software side, Gus has worked on a wide variety of projects. He designed, architected, and implemented a GPS signals monitoring and verification software system tool called Maverick, which is a key component to the SAASM/M-Code Mission Planning System deployed and in current operations at the Joint Space Operations Center at Vandenberg AFB. He also architected and implemented the Nostradamus Objective software application tool, which utilizes probability distribution to project delivery forecasts for any given system in development, and post-synthesizes projected forecast of a system with up to 95% percent accuracy. The tool is currently being utilized by US Space Force to help make better decisions for National Security Space Launch. On the Aerospace side, Gus has worked alongside Space Force, Launch Service Providers, and FFRDC's on launch-vehicle certification projects related to Avionics and Software design, development, testing, and flight of next generation Expendable Launch Vehicle Services. His efforts have included certification for SpaceX Falcon 9 & Heavy, Northrop Grumman's Omega, ULA's Vulcan Centaur, and Blue Origin's New Glenn.

Scott Morchower has extensive experience in the Aerospace industry encompassing Thermal Analysis, Stress Analysis, and Project Engineering. He spent 13 years in Thermal Analysis on the Delta II, Delta III, and Delta IV launch vehicles involved in all facets of thermal analysis from prelaunch to orbital environments along with factory, test, and launch support. He spent 9 years in Stress Analysis on the 737 MAX, 747-8 and 787 commercial airplane programs conducting stress analysis for aircraft refurbishment, product sustainment, factory tags, and new product development including a role as a technical lead for the 737 MAX Systems Stress team. His Project Engineering experience includes two years in commercial aircraft supporting the design, build, test, certification, and integration of supplier furnished equipment from program kickoff through customer acceptance. He is currently a Structures Project Engineer on the US Space Force New Entrant Launch Vehicle Certification program supporting new entrant compliance with Department of Defense (DoD) design, test, and performance requirements.

Novel Schedule Forecasting for Low-Volume Highly-Complex New Product Development: Manufacturing Phase

Bruce Chehroudi, PhD	Jonathan Lam, PhD	Gus Benavides	Scott Morchower
ManTech International	US Space Force	Axient	ManTech International
El Segundo, CA	Los Angeles, CA	El Segundo, CA	El Segundo, CA
Bruce.chehroudi@mantech.com	jonathan.lam.1@spaceforce.mil	gustavo.benavides@axientcorp.com	scott.morchower@mantech.com

INTRODUCTION

Project cost overruns are closely linked to schedule delays for major defense acquisition programs. The Government Accountability Office (GAO) previously found that manufacturing/integration and testing phases are most at risk of incurring cost and schedule growth. As of June 2022, NASA's portfolio of major projects experienced cost and schedule overruns while more projects were added. Out of 21 major projects in the development phase of NASA's acquisition process, 15 were responsible for a cumulative cost overrun of approximately $12 billion and cumulative schedule delays of 28 years. But just three projects—the James Webb Space Telescope, Space Launch System, and Orion—are responsible for more than three-quarters of the cost growth and almost half of the delays, see Russell, et al. (2022).

Based on an exploratory case study from one of the largest manufacturing plans of a global company, Koteshwar (2017) listed nine key challenges in the management of *new* product introduction projects: (1) designing/identifying the right resources (identifying the time for the right resources), (2) time-readiness and schedule, (3) stage-gate administration, (4) old ways of working, (5) poor communication and time-sharing, (6) missing learning opportunities (lack of action to spread the lesson, hence avoiding the same mistake), (7) defining business case, (8) poor coordination and alignment between different sub-projects, and (9) more projects with less competences. In an audit report, NASA identified the following challenges to meeting cost and schedule goals (see, Martin et al. (2012)): a culture of optimism estimating ability to overcome risks inherent in delivering projects within available funding constraints, technical complexity inherent in most projects, project managers' struggle to execute projects in the face of unstable funding, decrease in the number of small projects where aspiring managers can gain hands-on experience, and concerns regarding the decline in number of personnel with new product development experience and whether NASA can continue attracting technical talent. In *complex* new product introduction projects, there are larger levels of variations in scope, engineering, and late-stage changes that lead to late product deliveries. Additionally, with the involvement of a multitude of domestic and international suppliers, it is challenging to ensure that they work together to meet the pre-planned time schedule. Moreover, it is difficult to plan a specific target date for an entire program schedule involving multiple projects, investments, and suppliers. Indeed, a delay in equipment delivery from one supplier can affect the entire schedule for product delivery by the manufacturer. Finally, as noted by Javedi et al. (2013), there are challenges specific to a *low-volume* production system, such as: knowledge transfer from the sub-projects into manufacturing/production, development of the work instructions, the need for a higher level of training for the operators and production system design, and the required tailoring of new products to the existing production systems. All of the aforementioned factors make schedule forecasting for low-volume highly-complex new product development a challenging endeavor.

On-schedule availability of highly-reliable and complex new products (such as a gas turbine aircraft engine, space station, rocket engine, etc.) is quintessential for executing a program within cost and schedule constraints. Specifically, availability of key technical hardware is dependent on successful qualification and acceptance tests following product delivery by the primary manufacturing contractor, an event that commonly experiences schedule delays, but is difficult to predict or project. Recognition of the short-comings of existing schedule forecasting methodologies, prompted the investigation, development, and innovation of more objective approaches. Availability of such new powerful tools, further empowers customers or users of these tools, specifically the US Government (USG), with capabilities for future project schedule management in its acquisition programs. Therefore, objective estimates of the *product delivery date* and the subsequent *required-tests completion date* are of paramount importance.

The team members took two completely different approaches, addressing the challenging problem of schedule forecasting at two different manufacturing and testing phases of a new and highly complex product. Forecasting for testing phase will be submitted at a later NDIA conference. This is because of major differences in the nature of the tasks and processes involved in these two phases, which led to different schedule forecasting methodologies, using different principles, logical, analytic, and computational tools. In the following, only the manufacturing phase schedule forecasting is discussed.

BACKGROUND

In this section we concisely review relevant and required features of schedule forecasting methodology to generate objective project duration forecasts that minimizes or eliminates possible biases. The novel Nostradamus forecasting methodology described here contains all of the important features. This section should also facilitate an understanding, appreciation, and recognition of important features in the Nostradamus Objective forecasting approach (hereafter referred to Nostradamus). To the best of our knowledge, this is the first time that such a unique approach has been proposed and implemented in literature.

In project management, cost and resource distribution is influenced by *project duration*. We make distinction between the *planning or decision-making phase* and the actual *implementation phase* of a project. However, in both cases, an objective evidence-based forecast of the project duration is of paramount importance. During the planning stage of a project, the main interest is often on cost-benefit analysis. This is specifically a helpful tool in public investment policy and planning. During the implementation phase, one is interested in the project's progress, as compared to what was planned, cost overrun, and identification of possible measures to compensate for any delays in project completion date.

Unfortunately, as documented by Flyvbjerg and Bester (2021) and Flyvbjerg (2006), cost-benefit analysis, if not practiced carefully, is of less value because of psychological and political biases in its process that account for inaccuracies observed in forecasting. The psychological source of forecast inaccuracy stems from "optimism bias", a cognitive predisposition found with most people to judge future events in a more positive manner than is warranted by actual experience. Explanation of inaccuracy in terms of optimism bias has been developed by the work of Daniel Kahneman, see Kahneman (1994), Kahneman and Tversky (1979a, 1979b) and Lovallo and Kahneman (2003). They found that human judgement is generally optimistic because of "overconfidence" and "insufficient regard to distributional information." Hence, people will underestimate the project costs, completion times, and risks of planned actions, whereas they will overestimate the benefits of the same actions. Such behavior termed the "planning fallacy" and is reasoned that it originates from people or forecasters who take an "inside view" by focusing on the constituents of the specific planned actions or tasks rather than on the outcomes of similar actions or tasks that have already been completed. The latter source of forecast inaccuracy comes from strategic misrepresentation.

To compensate for the type of cognitive biases that Kahneman and Tversky found on decision making under uncertainty (which won Kahneman the 2002 Nobel Prize in economics), the "reference class forecasting (RCF)" approach was developed. They show that RCF can bypass human bias, including the ones mentioned earlier. In experimental research performed, RCF has been demonstrated to be more accurate than conventional forecasting methods. The study demonstrated that errors of judgement are often systematic and predictable than being random, suggesting bias rather than confusion. Interestingly, such errors in judgment are shared by both laymen and experts alike and that errors remain compelling even when the actor or forecaster is fully aware of their nature.

Traditionally, project managers focus on the specifics of the considered project (e.g., particular actions, tasks/subtasks) to produce estimations, as they attempt to forecast uncertain events that would influence the future course of the project. Such an "inside view" forecasting approach is based on human judgment. RCF, however, takes an "outside view" on planned actions. The outside view on a given project is based on knowledge about actual performance in a reference class of comparable projects. Therefore, RCF does not try to forecast the specific uncertain events that will affect the particular project, but instead places the project in a statistical distribution of outcomes from a "class of reference projects."

Batselier and Vanhoucke (2016) compared the project duration forecasting performance of RCF with those by baseline estimate, task-duration-based classical Monte Carlo (MC) simulation (Gantt chart with symmetrical and asymmetrical triangular probability distribution for action or task durations), and Earned Value Method (EVM) for real-life construction projects. The study found that RCF was the most user-friendly because it does not require detailed information (such as distributional data on task/subtasks durations for MC simulation) or extensive calculations (such as periodical forecast updates for EVM). Although RCF produces pre-project forecasts that remain constant throughout project execution (just like baseline estimates and Monte Carlo simulation), it surpasses all the traditional techniques in terms of accuracy, stability, and timeliness. The dominance of RCF in accuracy is especially remarkable, considering that the competing EVM technique offers forecasts that are updated at tracking points during the project progress. Furthermore, the strong performance of RCF occurs for both cost and time forecasting in nearly equal measure. The key point to emphasize for the RCF approach is that a reference class (for schedule & cost forecasting) should consists of projects that are sufficiently similar to the considered project in order to attain the required level of accuracy in the forecasts.

The novel Nostradamus forecasting approach, while completely different than RCF (in formulation and execution details), maintains and implements key concept that a reference class as similar as possible to the project at hand, is to be used in some form, if an objective and accurate project duration is critical and required.

The Nostradamus forecasting methodology was developed in response to the realization that the existing conventional forecasting approaches were ill-equipped to be applied to low-volume, highly-complex new product development efforts practiced under USG acquisition process. The term "complex or complexity" is used in the sense defined by William and Hillson (2002). The William-Hillson's model of complexity is an extension of the Baccarini's model, in which uncertainty was added to two complexity dimensions, see Baccarini (1996). The two dimensions by Baccarini are the number of elements and the interdependencies of these elements. In essence, they attribute the increasing complexity of a given project to two compounding causes: the relationship between *product* complexity and *project* complexity, and the *length* of the project.

In addition to the complex nature of the product and project, an important feature of the scenario considers the fact that the forecasting of the project duration was to be accomplished for a new product that has never been built in USA. Hence, a significant departure from past experiences caused high uncertainty in so many processes, such as design, analysis, and particularly manufacturing. The following work focuses on forecasting the time duration needed for manufacturing and delivery of the first few units of a new product. It should be made clear that this method considers manufacturing and delivery of each unit of product as a single project for which product delivery-date forecasting is required. Additionally, the experimental data was limited to three identical product units, each at different stages of manufacturing, such that some or all learnings from a predecessor unit were transferred to subsequent units (learning curve effect). As explained later, Nostradamus does take into account such a manufacturing process improvement in its forecasting approach.

ACCURACY-LEVEL (AL) FORECASTING CONCEPT

In this work, the terms "task" and "component" are synonymous and refer to major tasks and major component, however, the word "major" is often omitted. A major "task" is defined as a set of all "activities" required to manufacture a major "component." Therefore, once a major task is successfully completed, an associated major component is manufactured. For example, a compressor-turbine unit is a major component for a turbocharged automotive engine, while nuts and bolts are not considered as major components. Here, a component can be a hardware or a software needed to control one or more hardware items. Nevertheless, these major components (or major tasks) are entirely defined by the primary manufacturer. A "product" is defined as an assembly of major components. The term "task" is more general, in a sense that its accomplishment may lead to a hardware, a software, or simply a service. The end of a project is determined at the time when all of the pre-defined major tasks are completed, and the final hardware, software or service is delivered. In what is presented here, a limited quantity of units of the same new products are being manufactured with different manufacturer-estimated delivery dates distributed into the future. The objective is to forecast delivery dates for each of these products.

Nostradamus uses two important pieces of information to forecast the product delivery date by the "primary manufacturer" (hereafter referred to as "manufacturer"). The *first* is the "*performance*" on a recently-delivered product (referred to as "*source product*"). Here, the "performance" is assessed and quantified by determining "*accuracy level (AL)*" the manufacturer was able to achieve in estimating the completion dates of different components of the product in question. Note that the methodology uses the essence of the "reference class" concept, and its validity is justified, because the information used by Nostradamus is as close as one can possibly reach to this concept. It should be clear that the Nostradamus forecasting methodology works for the second new product delivered by the manufacturer and beyond, relying on an performance assessment from a most recently delivered product. However, this limitation can, to a large extent, be addressed through use of a reference class product delivered in the past (not the same product), although for a completely new and complex product of the type, the efficacy of this approach was not investigated here. Further studies are examining this case and will be published.

The *second* important piece of information Nostradamus uses to provide product delivery date forecasting is the "*current information*" from a not-yet-delivered product (referred to as "*forecast product*") for which delivery forecast is needed. To understand how this is achieved, the term "line of balance (LOB)" as used in this work is introduced. The LOB consists of a list of major components or tasks of a product, along with the associated estimated completion dates (ECDs), and expected product delivery date. The LOB is provided by the manufacturer and updated at every periodically-held (2 to 4 weeks) Program Manufacturing Review (PMR) meeting. There is no need to specify any interdependencies between these components or tasks, as it is assumed that schedule interdependencies between major components are taken into account by the manufacturer and reflected in the ECDs. Nostradamus only requires dates the manufacturer thinks or estimates each of the components would be completed. Updated manufacturing schedules (ex. LOB, IMS) are often part of the contractual agreement between the USG and the manufacturer, and are required, if Nostradamus is to be used. It is assumed that at any PMR date, and to the best of their knowledge, the manufacturer's engineers and project managers considered all relevant aspects of engineering, manufacturing, and available resources pertaining to the project, in their published estimated component (or task) completion dates. This is reasonable, as manufacturers, especially those involved in medium/large scale projects, have vested interest in such a practice and estimates needed for their own internal project management purposes.

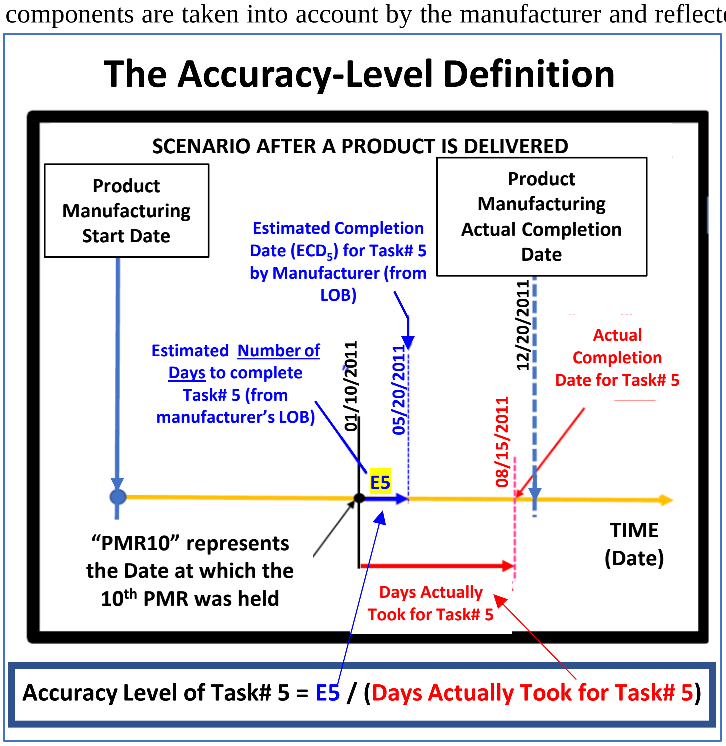

Figure 1. Shows an example as to how the Accuracy Level of an estimated completion date (ECD) for task #5 (or component #5), can be computed after when the product manufacturing is completed and it is delivered. The ECD is assessed by the manufacturer and, in this example, given on the 10th PMR. The logic applies to all other task ECDs provided in the LOB spreadsheet.

In summary, Nostradamus uses schedule information from a completed unit to forecast the product delivery dates of subsequent units using analysis of periodically released LOBs that contain a comprehensive list of major components/tasks. This concept appears similar to the Bayesian approach, where prior information is used along with new information to provide a posterior knowledge. However, details of the Nostradamus' approach, are quite different than the Bayesian approach.

The Nostradamus forecasting approach is based on the concept of "*Accuracy Level (AL)*" of the product manufacturer's task or component estimated completion date (ECD). The AL is defined as the number of days estimated by the manufacturer to complete a given task (Ei) divided by the "actual number of days" it took to complete this same task (ECD_i – PMR_j_Date). Note that this "actual number of days" is known by the time a product is completed and delivered. The AL values are often presented as percentages in which AL values less than, equal to, or greater than 100% indicate that the task completed later than, exactly on, or earlier than planned, respectively.

For a clarifying example, let us assume that a manufacturer is working on the <u>very first product</u>. At its 10th PMR event held on 1/10/2011, the manufacturer provided an updated LOB spreadsheet file consisting of 40 components, assembly of which constitutes a completed and deliverable product (see Fig. 1). On this LOB, one reads that the component #5 (or task #5) will be completed on 5/20/2011, as displayed in Fig. 1. On this PMR date of 1/10/2011, it is unknown how accurate this manufacturer's estimate is. However, the "*accuracy level*" can be calculated once this first product is delivered. Let us further assume that the completed product is delivered on 12/20/2011, but component #5 was completed earlier on 8/15/2011. Therefore, the "accuracy level" of the manufacturer's ECD for component #5, as stated on the PMR date of 1/10/2011, is equal to (number of days from 1/10/2011 to 5/20/2011) divided by the actual duration to complete (number of days from 1/10/2011 to 8/15/2011) resulting in an = AL = 130/217 = 0.599 (or 59.9%) for this single ECD pertaining to the component #5. Figure 1 explains how the AL is calculated after a product is delivered.

In general, at "jth" PMR for a given product presented on PMR_j_Date, the following relationships hold for any task #i (or component #i):

$$E_i = ECD_i - PMR_j\_Date \qquad (1)$$

$$AL_i = E_i / (ACD_i - PMR_j\_Date) \qquad (2)$$

Where, E_i (i.e., "E for task #i") is expressed in "number of <u>days</u>" between the two dates, ECD_i (Estimated Completion <u>Date</u> for task #i) and PMR_j_Date, AL_i is the Accuracy Level for task #i (or component #i) at PMR_j, and ACD_i is the date when task #i was actually completed. The term (ACD_i – PMR_j_Date) is the "number of days it actually took for task #i to be completed," which is only known when this task is actually completed.

As an example, a project which held 15 monthly PMR events for a product consisting of 40 components, will result in a maximum of 40x15 = 600 AL values once the product is completed and delivered. This dataset should be sufficient to construct a well-behaved probability histogram (PH) of the calculated AL values to determine the mean and standard deviation. The mean value of this PH is an assessment of an average accuracy or precision of the manufacturer's own forecasting methodology used in determining the task ECDs. The logic described here remains the same once a product is delivered, whether it is the first or the tenth unit. Once a product is delivered, Nostradamus applies the best statistical

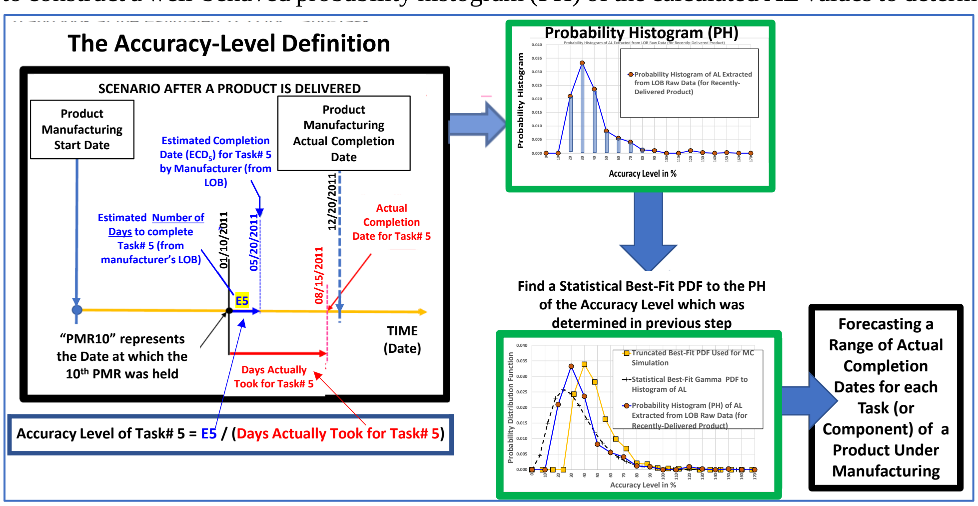

Figure 2. Describes steps needed to construct a statistical best-fit PDF function to the raw-date accuracy-level probability histogram (PH) from the LOB information of a recently-delivered product.

curve fit to the accuracy level PH in order to generate a Probability Distribution Function (PDF). This PDF is an analytic function that is subsequently used for computer Monte Carlo (MC) simulations. The PDF of the AL values, extracted from the LOB data of a recently-delivered product, is then used by Nostradamus (see Fig. 2) for simulation and schedule forecasting of subsequent products that are under manufacturing and not yet delivered. (See the algorithm and computational method section).

When a delivered product's PDF is used for simulation and forecasting, this product is referred to as "*source product*." The product for which a schedule forecast is desired, is called the "*forecast product*." It should be evident that this "forecast product" is under manufacturing and not yet delivered. If an accuracy-level PDF, constructed from the LOB information of a "source product" A, is used to provide a schedule forecast for a not-yet-delivered "forecast product" B, it is symbolically represented as "A → B."

UNDERLYING ASSUMPTIONS

As indicated earlier, the Nostradamus algorithm requires a best-fit PDF of the accuracy level constructed from the LOB data of a source product to provide schedule forecasts for all forecast products. There are three critical assumptions that should be kept in mind when utilizing the Nostradamus schedule forecasting approach:

1. Major components listed in the LOB spreadsheet are independent of each other; or if not, they are treated as independent because the ECDs for tasks are determined by a coordinated team of engineering, manufacturing, and project management experts who are cognizant of all inter-component dependencies (i.e., provide-need requirements) and use the manufacturer's own scheduling tools/system to generate these component ECDs. Therefore, schedule impacts of any dependencies or linkages between such components have already been incorporated in the manufacturer-assessed ECDs, and is inherently reflected in the LOB.
2. The Accuracy Level (AL) Probability Distribution Function (PDF) for a source product and its associated statistical parameters are largely applicable to all product units currently under manufacturing. The PDF derived from a source product is assumed to define and represent the underlying accuracy with which the manufacturer provides forecasts of the ECDs for any not-yet-delivered product.
3. The underlying probability distribution of the manufacturer's own forecasting accuracy for generating the ECDs pertaining to a given forecast product, is expected to improve during its manufacturing due to the "learning curve" effect. However, Nostradamus uses a fixed accuracy-level PDF (derived from the "*source product*") for schedule forecasting of a forecast product. This is justified, because impacts of any "learning curve" during the manufacturing of a given product are assumed as reflected in the manufacturer's updated LOB data disseminated at each subsequent PMR.

ALGORITHM AND COMPUTATIONAL METHOD

Once the probability histogram (PH) of the accuracy level is determined from the LOB of a source product, the Nostradamus algorithm applies a curve-fit using a suitable analytic probability distribution function (PDF) commonly used in statistics. The "Maximum Likelihood" approach is applied to the accuracy-level raw data from which the PH was constructed, see Montgomery and Runger (2003). Steps in this process are depicted in Fig. 2. While several different analytical PDFs were tested, it was found that the Gamma function best conformed to the dataset. Note that the Gamma function value of zero is meaningful when its independent variable (i.e., accuracy-level) approaches zero. Additionally, the Gamma function is only defined for zero and positive variable values, which is consistent with the fact that "negative" accuracy level values are neither defined, nor meaningful. Recall that the accuracy level range is from 0 to greater than100%.

Once the best-fit Gamma PDF parameters are determined from the LOB data of a source product, it is used for a Monte Carlo (MC) simulation. Figure 3 shows how the statistical best-fit PDF is used to produce 3,000 randomly-generated accuracy level values for each task (or component) of the forecast product, for which schedule forecasting is needed. In this figure, the comments below the "*accuracy level formula*" are for when a source product LOB information is used to determine the statistical best-fit PDF. Once the PDF and associated parameters are determined, the comments above the "*accuracy level formula*" explain how this PDF is used to produce 3,000 randomly-generated accuracy level values (MC simulation) for each task "i". This is done to calculate 3,000 corresponding "forecasts" of the actual completion date for each task "i" of a forecast product (Fig. 3). Essentially, a single ECD_i (and its corresponding E_i – see, for example, E_5 in Fig. 1) which the manufacturer provides for a task "i" at a given PMR, is "revised/corrected" by the algorithm, 3,000 times to generate 3,000 "forecasts" of the actual completion date for this specific task. For each task "i" (or component "i"), a probability histogram can be constructed, using these 3,000 "forecasts," from which the Nostradamus' forecast completion dates at the 5% and 95% cumulative probabilities (Confidence levels) are determined. These same steps are repeated for each task (or component) listed in the LOB from each PMR.

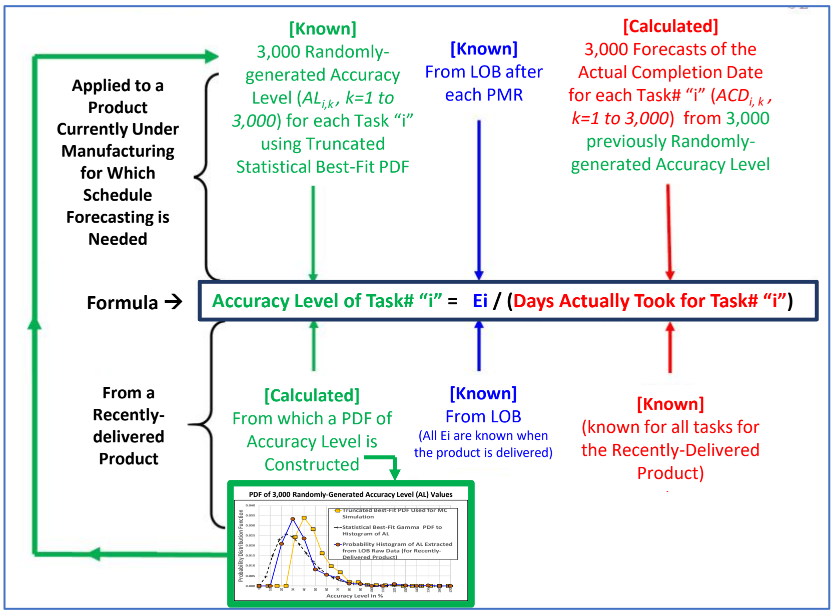

Figure 3. Depiction of how the statistical best-fit PDF, determined from LOB of a source product, is used to produce 3,000 randomly-generated accuracy level values ($AL_{i,k}$, $k=1$ to 3,000) for each task "*i*" (or component "*i*") of a forecast product, for which schedule forecasting is needed.

Figure 4 illustrates the Nostradamus schedule forecasting logic flow required to generate the delivery date forecast for a forecast product from a specified PMR date. For each task "i" (or component "i") of a forecast product, the algorithm randomly generates 3,000 AL values ($AL_{i,k}$, $k=1$ *to* 3,000), from which 3,000 "forecasts" of the actual completion date ($ACD_{i,k}$, $k=1$ *to* 3,000) are calculated using MC simulation (see comments above the AL "formula" in Fig. 3). This simulation is based on a truncated best-fit PDF of the accuracy level values (see Fig. 4(a)). Details on proper selections of truncation cutoff values (i.e., Lo_Limit and Hi_Limit) are discussed later in the calibration and anchoring section. From these 3,000 data, a probability histogram of the Nostradamus' "forecast" of the ACDi is built for each task "i", (Fig. 4(b) shows this probability histogram for task "i"). From this Fig. 4(b), two forecasts of the task completion dates at the 5% and 90% cumulative probability values are determined for task "i". Once these two Nostradamus' forecast dates are determined for task "i", they are connected by a vertical segment of line, as depicted in the Fig. 4(c). These steps are repeated for all other tasks in a software loop to complete the left-middle graph. The Fig. 4(c) simultaneously shows the Nostradamus' forecast ranges (i.e., dates from 5% to 90% cumulative probability values) for all the tasks (or components) of the product which is currently under manufacturing. Additionally, this same graph displays the PMR date (dashed horizontal line) and the manufacturer's expected product delivery date (solid green horizontal line) given at this PMR date. Moreover, the ECDs for all tasks given by the manufacturer (as hollow rectangular symbols) are shown for this same PMR event. The graph in Fig. 4(c) is important, because it provides a visual snapshot in time of the manufacturer's estimates and results of the Nostradamus forecast for a not-yet-delivered product at a given PMR date.

The last step in the Nostradamus' schedule forecasting logic flow diagram is the construction of the product *Delivery Failure Probability* (DFP) as a function of time (i.e., date) measured from the PMR date. The DFP is determined because the project risk of not being completed by the manufacturer's projected delivery date is directly proportional to the "failure" probability. The DFP is a curve that ranges from 100% failure to deliver to 0% failure to deliver. The Fig. 4(d) shows an example of the DFP curve. For example, the DFP value at a 01/20/2012 date (see the pink horizontal line in Fig. 4(c) and pink vertical line in the DFP graphs shown in Fig. 4(d)), is determined by the number of cases, out of 3,000 MC simulations, where at least one task's forecasted delivery date exceeds this 01/20/2012 date. Once the DFP curve is determined, the date at any DFP value can be ascertained. Additionally, the interim data matrix used by the algorithm to construct this DFP curve is further analyzed to quantify and rank the contribution of each task towards the product delivery failure probability value. Knowing this, programmatic measures can be taken to favorably affect the product delivery dates and reduce the schedule risk.

CALIBRATION AND ANCHORING PROCEDURE

The Nostradamus schedule forecasting algorithm requires calibration and anchoring to a reference and relevant dataset from a source product before it is able to provide acceptable and objective forecasts for forecast products on the same manufacturing line. Once a completed product is delivered by the manufacturer, ECDs at every PMR event pertaining to this source product are known. Additionally, because this product is already delivered, the *actual* completion date of each task (or component) of the product and when the completed product was delivered are known.

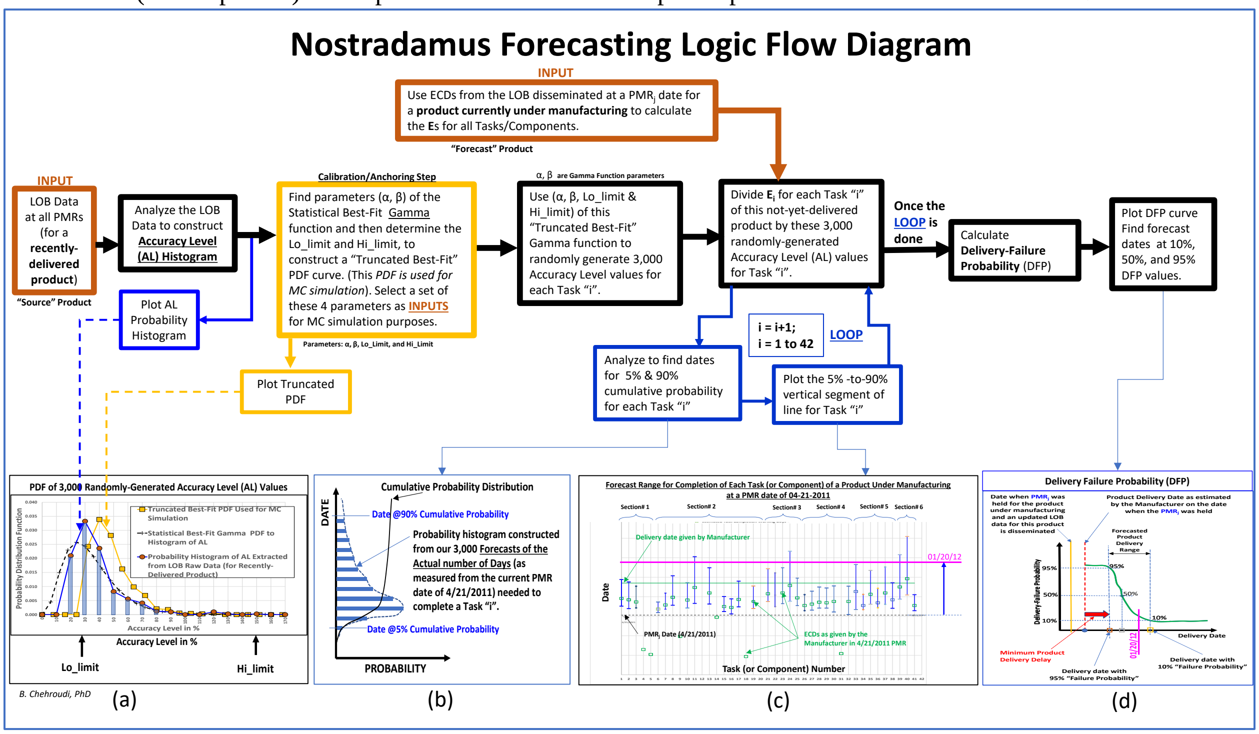

Figure 4. Shows the Nostradamus' logic flow diagram indicating major inputs and outputs. First, a statistical best-fit PDF to the Accuracy Level (AL) is determined for a recently-delivered product, shown in 4(a). For each task (or component) of a product currently under manufacturing, the program generates 3,000 "forecasts" of the actual completion date from which a probability histogram is built, see 4(b), and the dates at 5% and 90% cumulative probability values are identified and plotted (see vertical segments of line in 4(c)). Finally, a product Delivery Failure Probability (DFP) is constructed in 4(d).

As indicated before, a statistical best-fit Gamma function is used for the MC simulation. However, in practice, this Gamma PDF is truncated to reflect realistic forecasts. For example, an accuracy level value of 5% implies that the estimated number of days needed to complete a task ("E") was under estimated by a factor of 20 (see the formula in Fig. 3) thereby indicating that if the manufacturer estimated that this task would take 3 months to be completed, the forecasted actual duration to complete this task would be 60 months. This forecast is an unrealistic assessment. Hence, a Lo_Limit value (i.e., the lower cutoff for the truncated PDF) should be higher than 5%, (Fig. 4(a)). As shown in Fig 4(a), low and high limit cutoffs were incorporated to truncate the full Gamma function, that is, the Lo_Limit and Hi_Limit values. It is worth emphasizing that the accuracy level of 100% means that the manufacturer correctly estimated completion date for this task. while an accuracy level value >100% is plausible and implies that the manufacturer completed the task earlier than estimated. This scenario is realistic and has been observed in this analysis.

Note that the calibration process and the forecast outcomes are more sensitive to selection of the Lo_Limit value rather than the Hi_Limit. The Hi_limit value also needs some fine tuning. In practice, however, a Hi-Limit value of 150% was adequate for all cases investigated here and appeared realistic from actual data analysis.

The calibration process begins by the construction of a best-fit Gamma PDF to the accuracy level probability histogram of the source product. Using the best-fit Gamma PDF, it is pretended as if this source product is not delivered and attempts are made to construct a delivery failure probability (DFP) graph for each of the PMR events held for this specific product. This calibration/anchoring step follows the same logic shown in Fig. 4, but replaces the "forecast product" with the same source product. For a given fixed Lo_limit value, the product-delivery forecast date at a DFP=50% is determined for all of the PMR events pertaining to this same delivered source product. For each of these forecast dates, which is associated to a specific PMR event, a "*forecast imprecision*" is defined. A product's delivery-date *forecast imprecision* at a given PMR is the difference in number of days between the forecasted delivery date and the known actual product delivery date, expressed as a percentage of the number of days from that same PMR event to the actual product delivery date. The imprecision value indicates how close or far (in percentages) a forecast date was to the actual product delivery date. It should be evident that the product delivery-date *forecast imprecision* can only be calculated when a product is completed and its actual delivery date is known. Once the *forecast imprecision* is calculated for the 50% DFP date for all PMR events, a "*time-averaged forecast imprecision*" is computed. Note that this time-averaged forecast imprecision value is determined at the previously-given fixed Lo_Limit value. The primary goal of the calibration and anchoring procedure is to find an optimum Lo_Limit value that produces a "time-averaged forecast imprecision" of equal to zero.

Once the optimum Lo-Limit value is determined in the calibration process, along with its associated Hi_Limit, the MC simulations for each forecast product which is currently under manufacturing are conducted. This is done by generating 3,000 randomly-generated numbers from the previously-determined statistical best fit Gamma PDF, but only accepting numbers in between the pre-determined Lo-Limit and Hi_Limit values.

TEST RESULTS AND DISCUSSIONS

The Nostradamus forecasting concept described here has been tested extensively using realistic data from past USG acquisition programs. In this section, typical results observed in our tests are presented and discussed. The results presented here utilize LOB data from the third already-delivered product (called "A" i.e. source product) to calibrate the PDF parameters, which is subsequently used in the MC simulation to provide delivery date forecasts for the fourth product (named "B" i.e. forecast product) at each of its respective PMR dates. It is important to indicate that when these forecasts were made, product B was not yet delivered. Conclusions reached here are also similar for any relevant product pair (e.g., 1st manufactured product, as a source product, and 2nd one, as a forecast product, etc).

Figure 5 is the Historical Forecast Curves (HFCs), and shows the results of the product-delivery forecasts (DFP = 10%, 50%, 95%) for product B from LOB data from each PMR. For comparison purposes, Fig. 5 also shows the manufacturer's estimated delivery date for the same product B (Blue line). Each symbol on these curves is a product-delivery forecast made at a given PMR date, either by Nostradamus (three forecast dates at 10%, purple; 50%, yellow; and 95%, red DFP values) or by the manufacture (a single blue forecast date). The two solid green lines represent product B's actual delivery date, which was known only when it was actually delivered. The number of days from the first PMR shown in Fig. 5 to the actual product delivery date was about 873 days. A 45-degree line is also shown as a visual reference line. A linear fit to the manufacturer's HFC (Blue) (not shown) confirmed that the project progression was at an average slope of less than 45 degrees. The three HFCs derived by the Nostradamus algorithm converge to a point towards the end of the project, which appears to be on the 45-degree line. In fact, it can be reasoned that all four HFCs must converge to a point on the 45-degree line, despite the fact that no PMR was held past the last one shown in Fig. 5. However, it is assumed that the manufacturer continued holding PMRs and the last PMR was on the day when product B was actually delivered. On this date, the manufacturer's forecasted delivery date would have been reported the same as the PMR date, corresponding to a point on the 45-degree line. This point is the actual delivery date of the product B under investigation.

At first glance, all three Nostradamus HFCs in Fig. 5 provide delivery-date forecasts that are much closer to the actual delivery when compared to the manufacturer's forecasts. However, the most important observation is that the 50% DFP forecast date hovers around and closely follows the actual product-delivery date as shown by the green horizontal line (Fig. 5). Considering the manner in which the calibration/anchoring was conducted, the date at the 50% DFP value is Nostradamus' *nominal delivery-date forecast* with variations within the bounds defined by dates at 10% and 95% DFP values.

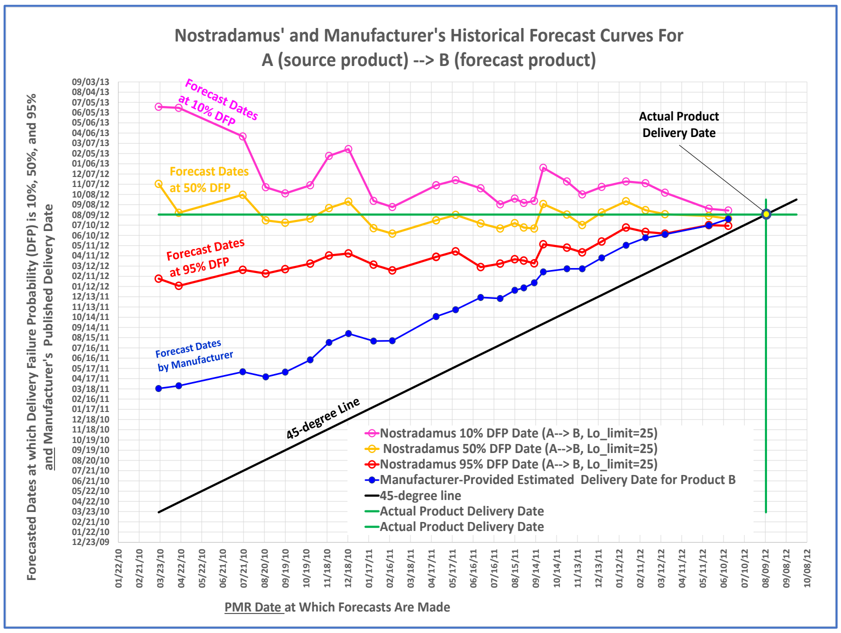

Figure 5. Shows the Nostradamus' product-delivery Historical Forecast Curves (HCFs) at three DFP values of 10%, 50%, and 95% for the "forecast product B." A single product-delivery HFC by the manufacturer for this same product is also shown. The recently-delivered "source product" in this case was named "A" and the "forecast product" was "B". All delivery-date forecasts are made when this product (i.e., "B") was not yet delivered.

Figure 6 depicts the delivery-date forecast imprecision curves for product B, calculated after the product was actually delivered. For the Nostradamus, the nominal forecast date was used to calculate the forecast imprecision values. The imprecision for product B is defined in the same manner as already described for the "source product." Starting from the beginning of the project for product B, the manufacturer's imprecision (orange) ranges from 110% to 222% over the course of 2 years, and only improves in the last 6 months of the project. Towards the end of the project, the manufacturer's imprecision curve trends closer to zero. In comparison, the imprecision for Nostradamus' product-delivery nominal forecast closely hovers around zero throughout the duration of the project. These results validate the calibration and anchoring process, and exhibit an accurate forecast capability 873 days ahead of the project end date and before the product was delivered (i.e., at the beginning of the project).

The maximum, average, and minimum imprecision values for Nostradamus' forecast during the entire project period were -29%, -3%, and 16%, respectively, whereas those by the manufacturer registered at 222%, 143%, and 32% values. The promising and impressive forecasting performance of the Nostradamus should be apparent.

It is important to note that for forecasting delivery-dates of any on-going project, Nostradamus does not require the dates when tasks (or components) started. All that is needed, are the manufacturer-provided estimated completion dates for these tasks. Although biases exist (such as overconfidence, anchoring, strategic optimism, and availability, (see Kahneman and Tversky (1979a)), Nostradamus compensates using the actual performance on a most-recently-delivered source product. Also, requiring the manufacturer to disseminate its best estimates of task (or component) completion dates in each PMR, not only substantially reduces forecasting complexity, but also provides an incentive for offering its best forecasting efforts in order to minimize manufacturer's reputational risk. If a new unexpected technical issue arises after a PMR date that causes a substantial schedule delay, the Nostradamus algorithm will utilize updated information that the manufacturer incorporates in the next LOB that takes into account the technical issue's

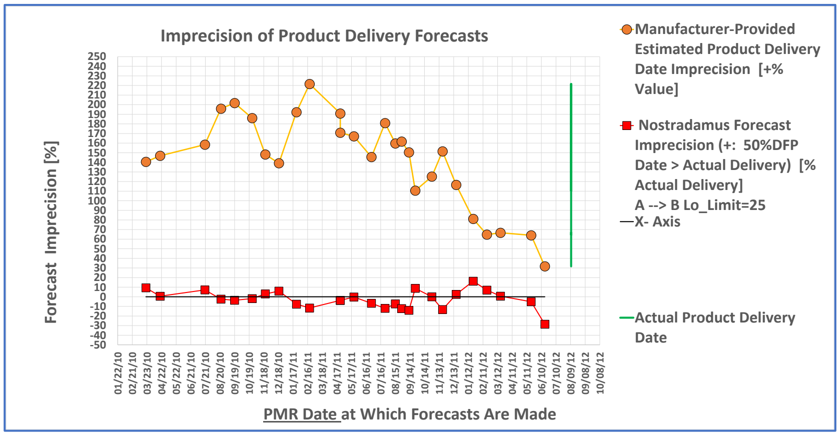

Figure 6. Shows the imprecision of the forecasted product delivery dates for Nostradamus and the primary manufacturer. Zero imprecision value implies 100% accurate forecast. All delivery date forecasts are made when this product was not yet delivered. Imprecisions are then calculated when the product is completed and delivered at the end of its project.

schedule impact (Assumption #1). Indeed, this Bayesian-like approach is inherently incorporated into the Nostradamus program logic (i.e. using the past established performance and combining it with any newly-provided information to update and correct its product delivery-date forecasts) pays tribute to the adaptable and agile nature of the algorithm. Additionally, no information regarding the inter-task dependencies or linkages (or inter-activity dependencies/linkages within a task) are needed (assumption #1) because their impacts are captured in the task ECDs provided by the manufacturer. This is in stark contrast to forecasting methodologies requiring Gantt charts with begin-end dates and inter-task relationships to function. When compared to other statistically-based approaches, the use of the most recent performance information substantially reduces complexity in the forecasting methodology proposed here. Lastly, and most notably, Nostradamus can be used at any time within a progressing project regardless of the project/product type.

CONCLUSIONS

A novel accuracy-level LOB-based schedule forecasting methodology (Nostradamus Objective) is proposed for low-volume highly-complex new product development. While Nostradamus is expected to work reliably and objectively for any product manufacturing (or project management) effort, it is especially shown valuable and useful for new product development (NPD) and during the early stage of the project when uncertainties are high. During this early stage, one or a very few product units are delivered (defining the "*past or most recent performance*"), but forecasts are highly needed for all not-yet-delivered products using LOB data from recurring PMRs (i.e., "*current information*"). The Nostradamus logic conforms with the "reference class concept" described by the Nobel Prize laureate Daniel Kahneman and coworkers, and is able to reliably and accurately forecast project completion dates of products on the same manufacturing line. Additionally, the Nostradamus software program ranks a list of top components (or project tasks) of the product that heavily affect the delivery date or delivery delays. To the best of our knowledge, this is the first time in literature that the LOB data is used for schedule forecasting in such a unique approach described here.

The raw data analyzed here is extracted from actual new product development in the context of US Government acquisition process. The algorithm was entirely developed in-house as no such commercially-developed program was available or able to provide reliable forecasts. The algorithm has gone through extensive verification and validation testing to build confidence on its implementation and results. Typical forecasting results are presented here, where a full LOB data from a recently-delivered product "A" is used to construct a probability distribution function (PDF) of the "accuracy level" values. This PDF is then used to perform MC simulations and produce 3,000 possible delivery date forecasts for each component of a forecast product "B". Once the product "B" is later delivered, forecast imprecisions are calculated to see how close the Nostradamus forecasts were to the actual delivery date of product "B," and compared them with those calculated for the manufacturer's forecasts.

Results of these tests indicated that, over the project duration, a time-averaged forecast imprecision value of -3% was achieved using Nostradamus, as compared to 143% by the manufacturer. The imprecision of Nostradamus' product-delivery forecast, remained near zero even from the beginning of the project (nearly 2.5 years ahead of project completion date). The maximum, average, and minimum imprecision values for Nostradamus' forecasts during the entire project period were -29%, -3%, and 16%, whereas those by the manufacturer registered at 222%, 143%, and 32%. The promising and impressively-high forecasting performance and accuracy of Nostradamus Objective will continue contributing to high-visibility programs with strict cost and schedule constraints pertinent to delivering needed capabilities to the USG.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of the following organizations: US Air Force (for original authorization and support of the work), ManTech International, and Axient (both, for support of this activity). In addition, we are grateful for continuous support and encouragements by Maj. Marissa Reabe and Capt. Peter Cho from US Space Force. Mr. Michael Sanders from ManTech is specifically thanked for his support and marketing, leading to a recommendation for this conference publication.

REFERENCES

Batselier, J. and Vanhoucke, M., 2016. Practical application and empirical evaluation of reference class forecasting for project management, *Project Management Journal*, Vol 47, No 5, 35-51.

Baccarini, D., 1996. The concept of project complexity—a review, International Journal of Project Management, vol. 14, no. 4, pp. 201–204.

Flyvbjerg, B. and Bester, D. W., 2021. The cost-benefit fallacy: why cost-benefit analysis is broken and how to fix it, *Journal of Benefit-Cost Analysis*, October, pp. 1-25.

Flyvbjerg, B., 2006. From Nobel Prize to project management: getting risks right, *Project Management Journal*, vol 37, no 3, 2006, pp.5-15.

Javadi, S., Bruch, J., Bellgran, M., and Hallemark, P. (2013), Challenges in the industrialization process of low volume production systems, 11th Int. Conference on Manufacturing Research, Cranfield University.

Koteshwar, C., 2017. Challenges in managing new product introduction projects: an explorative case study. 21st International Conference on Engineering Design, ICED17, 21-25 August.

Kahneman, D., 1994. New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics*, 150, pp 18-36 .

Kahneman, D. and Tversky, A., 1979a. Prospect theory: An analysis of decisions under risk, *Econometrica*, 47, pp 313-327.

Kahneman, D. and Tversky, A., 1979b. Intuitive Prediction: Biases and Corrective Procedures, *Management Science*, vol 12, pp 313-327.

Martin, K. P., et al., 2012. NASA's challenges to meetings cost, schedule, and performance goals, Report No. IG-12-21 (assignment No. A-11-009-00), Sept 27, 2012.

Lovallo, Dan and Kahneman, D., 2003. Delusions of Success: How Optimism Undermines Executives' Decisions, *Harvard Business Review*, July, pp 56-63.

Montgomary, D., C. and Runger, G. D., 2003. Applied Statistics and probability for engineers, 3ed Edition, John Wiley & Sons, Inc.

Russell, W. W., et al., 2022. United States Government Accountability Office (GAO) Assessment of major projects, Report No. GAO-22-105212.

Williams, T. and Hillson, D., 2002. Editorial–PMI Europe 2001, *International Journal of Project Management*, vol. 20, no. 3, pp. 183-184.