Providing Integrity for Satellite Navigation: Lessons Learned (Thus Far) from the Financial Collapse of 2008 – 2009

Sam Pullen
Stanford University

BIOGRAPHY

Sam Pullen is a Senior Research Engineer at Stanford University, where he is the director of the Local Area Augmentation System (LAAS) research effort. He has supported the FAA and others in developing GNSS system concepts, requirements, integrity algorithms, and performance models since obtaining his Ph.D. from Stanford in Aeronautics and Astronautics in 1996. Based on this work, he was awarded the ION Early Achievement Award in 1999. In addition to GNSS, his research involves the general topic of system design and performance verification under uncertainty.

ABSTRACT

This paper examines the failures in risk assessment and risk management that contributed directly to the worldwide financial crisis of 2008 - 2009 and identifies lessons that are applicable to risk management in general and to satellite navigation integrity verification in particular. The market-distorting events leading up to the financial crisis are reviewed, and the role of flawed risk-management concepts is revealed. In particular, over-exploitation of the erroneous Efficient Market Hypothesis (EMH) and the unquestioned assumption of Gaussian market state transition probabilities led to the formulation of risk-quantification models upon which the debt market that collapsed in 2008 were based.

This paper derives several lessons from the failures that led to the financial collapse, including the preference for probabilistic models in risk assessment and the need to highlight the assumptions behind deterministic models when they are used instead, particularly when one deterministic model is derived from another. These lessons serve as reminders that, despite the high level of conservatism applied to satellite navigation risk modeling; faulty assumptions can create serious problems if careful vigilance regarding the ongoing applicability of these models is not vigorously and continuously maintained.

1.0 INTRODUCTION

The causes of the financial crisis that has led to a severe global economic downturn since mid-2008 are many and complex. Beyond the combination of natural forces and human errors that led to the devastation of Hurricane Katrina in 2005 (see [1]), the less-desirable attributes of human nature – avarice, selfishness, narcissism – played a direct role in setting the stage for the financial crisis. However, the building blocks of the financial market structures that collapsed at the onset of the crisis were originally developed by academic economists and researchers who were not primarily motivated by personal greed. The framework of financial market modeling and risk management that economists developed over decades became nearly universally accepted by the 1980's, even though many of its flaws were readily apparent. As the decades progressed, quantitative models built atop this framework diverged further and further from reality. They provided the mathematical underpinnings to the complex securitized-debt investments that were extremely profitable for their originators in the investment-banking community but which collapsed into junk as their fundamental assumptions were shown to be invalid.

This paper examines the contribution of faulty risk assessment and risk management to the recent financial crisis. As with the paper on Hurricane Katrina from ION GNSS 2008 [1], the goal is to learn lessons from this experience that can improve the way that risk is assessed and managed in engineering projects, particularly those related to satellite navigation. These lessons are mostly indirect, as the context of risk management for safety-critical applications of satellite navigation is far more responsible and conservative than the approach taken by the financial community. Despite this, several connections can be drawn that relate to the details of how risk calculations are performed and the degree to which deterministic models of risk are dependent on underlying assumptions that cannot be proven.

This paper is organized into the following sections. Section 2.0 describes briefly explains the origins of the

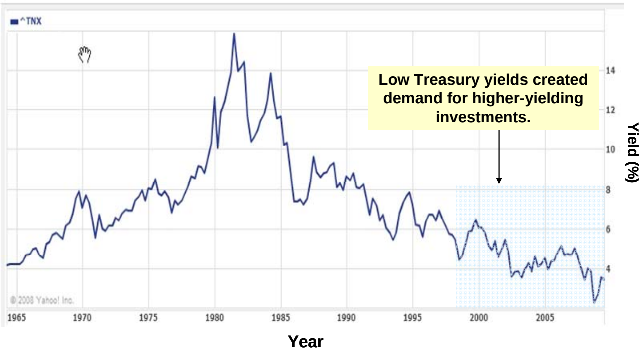


Figure 1: 10-Year Treasury Bond Yields, 1964-2009

financial crisis in terms of the development and sudden collapse of debt-based securities. Section 3.0 gives an overview of financial risk modeling and highlights specific theories that led to the construction of these securities. Section 4.0 summarizes the general lessons that can be learned from this disaster and which are relevant to engineering risk management as a whole. Section 5.0 examines two specific aspects of integrity risk modeling for satellite navigation that are linked to these lessons. Section 6.0 summarizes the paper and restates the lessons to be learned.

2.0 AN OVERVIEW OF THE FINANCIAL CRISIS

The financial crisis that began in late 2007 and gathered strength leading up to the collapse of Lehman Brothers in September 2008 was mostly a collapse of the market for debt (bonds) as opposed to equities (stocks). The debt market had changed dramatically in the decade or so leading up to the crisis. Several reasons have been cited for this change, but a primary motivation is shown in Figure 1, which plots the percentage yield on 10-year U.S. treasury bonds over the last 45 years [2]. The 10-year treasury bond is a benchmark in the debt markets because it is fully backed by the U.S. Government and is thus seen by most as "zero risk" or, at the very least, the least-risky form of debt. To obtain a higher yield in the debt markets implies taking on more risk of default – risk of not being paid back.

As shown in Figure 1, the yield on "risk-free" treasury bonds had dropped to historically low levels by the late 1990's and dropped further after the 2000-2001 stock-market decline as the Federal Reserve kept interest rates low to limit the damage to the broader economy. The maintenance of low interest rates during this decade fed the explosion of "easy money" that made it easy for consumers to borrow and spend past their means. The securitized-debt market that arose from this situation in the first half of this decade provided the mechanism and motivation for "easy money" to spread so easily throughout the broader economy.

Faced with the low yields on "safe" treasury bonds, and having a limited number of options to invest at higher yields with more risk, significant demand arose for debt investment vehicles that could do three things (see [4]):

1) Provide higher yields with only slightly higher risk than treasury bonds;

2) Provide multiple options tailored for different risk-versus-reward objectives;

3) Provide healthy profits ("fees") to loan originators and sellers (retail and investment banks).

The market in Collateralized Debt Obligations (CDOs) based on residential home mortgages grew so quickly because of its success in meeting these objectives. A simplified "wine glass pyramid" picture of how these securities were structured is shown in Figure 2 ([3], also see [4]). On the left-hand-side of the figure, collections of thousands of individual home mortgages are shown, and these combined mortgages pay principal and interest when the individual homeowners included in the collection pay their mortgages every month. The resulting outflow of money (shown at the top, flowing from the bottle exit) is sufficient to fill all of the glasses and trays shown below if all of these homeowners pay their mortgages in full. However, the CDO is structured with the knowledge that a few homeowners may stop paying and/or eventually default on their loans. Therefore, the first glass to be filled, at the top of the pyramid, is the least risky. Once it is filled, the "overflow" fills the next row of glasses, and so on until all rows are full or until the money runs out and the bottom row or rows go unfilled.

Because the risk of mortgage non-payment and default was thought to be well-understood, this structure allowed packages of mortgages to be "securitized" and sold to external investors in pieces known as "tranches". Investors aiming for a level of risk only slightly higher than treasury bonds but with significantly higher yields would purchase the "1st tranche", which was designed such that the agencies who rate debt instruments (Moody's, Standard & Poor's, etc.) would assign it the highest rating (i.e., lowest risk) of "AAA" or equivalent. Other investors with different risk vs. return preferences, such as mutual funds and hedge funds, purchased the lower tranches, taking on more risk but receiving a higher yield when their tranches filled up. The investors' risk could be lowered further by purchasing "credit default swaps" (CDSs), which were essentially insurance against default of the underlying tranche of the CDO. Because the 1st tranche, in particular, was rated "AAA", it was seen as having a very low default risk; thus a CDS against it could be bought relatively cheaply, even though a default would be enormously expensive to the firm selling the CDS.

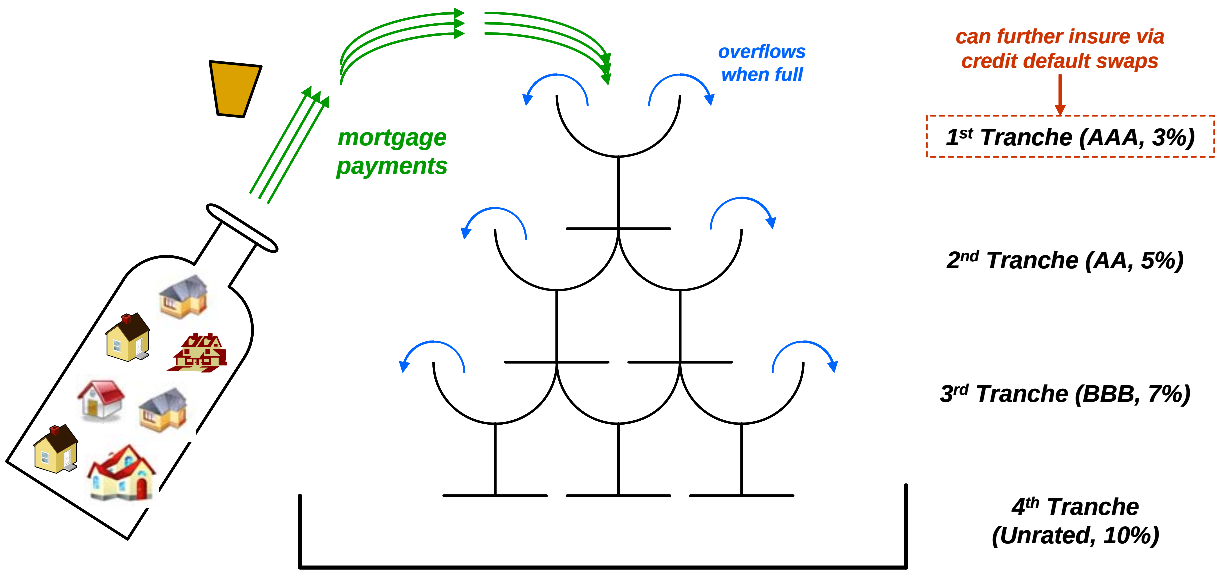


Figure 2: "Wine Glass Pyramid" Overview of Securitized Debt Market (from [3])

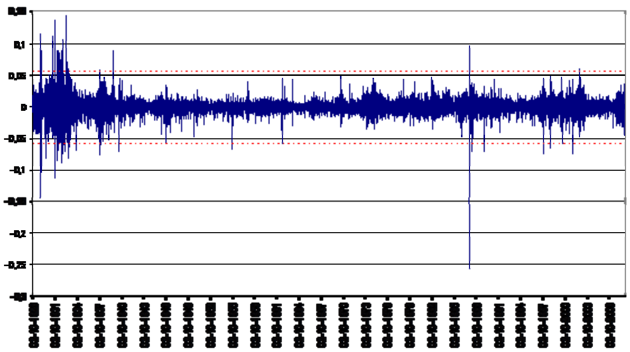
The collapse of these debt structures at the outset of the financial crisis was due to the fact that the models of default risk that they were built upon were seriously flawed. The key to modeling the risk of each tranche in the structure shown in Figure 2 is understanding the underpayment/default risk of each individual mortgage and the correlation of these risks across the thousands of mortgages comprising a given CDO. It was expected that any rise in defaults would be gradual and mostly uncorrelated from region to region of the U.S. Under these circumstances, the flow of money shown in Figure 2 would recede gracefully and would expose the lower tranches to the risk that they had agreed to assume. Instead, when homeowners across the nation began defaulting simultaneously at much higher rates than predicted, the flow of money seized up all at once, leaving even 1st-tranche investors unpaid. Institutions overly invested in CDOs and those who had sold credit default swaps that they could never pay off began to collapse in a domino pattern. The collapse of Lehman Brothers and the threatened collapse of AIG in mid-September 2008 led the U.S. and other governments, fearing a complete shutdown of the financial and credit markets, to step in and effectively guarantee the liabilities of the largest institutions – those deemed "too big to fail."

Section 3.0 will describe the financial-risk theories that led to the development of CDOs based on default-risk models that were so terribly wrong. It should be noted that excessive belief in these models led to behavior among the home lenders, investment banks, and debt purchasers that in retrospect looks fraudulent (and in many cases *was* fraudulent). As the popularity of CDOs grew, the demand for more and more of them and the fees generated by them led to compromises throughout the origination process. Home lenders relaxed their standards for issuing mortgages since they would be quickly sold to investment banks. The investment banks paid little attention to the declining quality of the mortgages they were buying because individual mortgage risk mattered little when thousands of mortgages were "averaged together" in a CDO. Rating agencies used the same logic to avoid looking closely at individual mortgages when rating tranches (in addition, they were motivated to provide a "AAA" rating to the 1st tranche, as they were paid by the banks originating the CDOs). These messy realities would have weakened the applicability of any risk models that did not take such changes in behavior into account, even if they were not as fundamentally flawed as the models that were actually used.

3.0 RISK MIS-MODELING UNDERPINNINGS OF THE FINANCIAL CRISIS

3.1 The Efficient Market Hypothesis (EMH)

The origin of the risk models upon which the structure of CDOs and CDSs was based is well-described in [5] (also see [6]). It starts with the observation of Louis Bachelier around 1900 that stock market prices tended to follow a "random walk", meaning that, given the observed present price, the following price (at the close of the following day, for example) appeared to represent a random change up or down. Bachelier examined the use of the "bell curve" (i.e., the Gaussian distribution) to model these changes but was cautious in doing so because he knew that the market prices were heavily influenced by human behavior and thus did not represent naturally-occurring combinations of smaller random factors. The mathematics he developed influenced Einstein's



A *non-Normal* October

Date	Percentage Change (1)	Average Frequency under Normal Law (2)
07/10/2008	-5.11%	Once in 5,345 Years
09/10/2008	-7.33%	Once in 3,373,629,757 Years
13/10/2008	11.08%	Once in 603,033,610,921,669,000,000,000 (3) Years
15/10/2008	-7.87%	Once in 171,265,623,633 Years
22/10/2008	-5.86%	Once in 117,103 Years
28/10/2008	10.88	Once 73,357,946,799,753,900,000,000 (3) Years

Figure 3: DJIA Percentage Change Compared to Normal (Gaussian) Distribution, 1928 – 2008 [9]

publication of similar tools a few years later, and the resulting concept became a popular means of modeling random variations in continuous time series, although it was not used in the financial world until the formulation of the Efficient Market Hypothesis (EMH) in the 1960's.

The EMH is normally credited to Eugene Fama of the University of Chicago School of Economics but is the child of many fathers at the University of Chicago and elsewhere, including such famous names as Paul Samuelson and Milton Friedman [5]. The basic idea of the EMH is that today's market price, being the product of a continual tug-of-war between buyers and sellers, is the outcome of all information available to a population of investors that is collectively (if not individually) "rational." Therefore, today's market price is the best possible estimate of the "true value" of the underlying security, and any change between today and tomorrow is due to changing information about this "true value." The EMH summarized in this manner is the "strong" version – in short: "*the price is right*." Weaker versions exist which place less reliance on the "perfection" of today's price and limit themselves to the concept of "*no free lunch*," meaning that the level of imperfection in today's prices is not large enough for individual investors to outperform the market, except by chance [7].

Over time, as described in [5], the stronger versions of the EMH became the foundation of both academic finance and investing as practiced by investment banks and large institutions. The dominant reason for this is the degree to which the EMH made tractable analytical, quantitative analysis of the financial markets. By stating that current prices contained all available information, the EMH denied that past price data (or other information) was relevant. Therefore, market price evolution could be modeled as a Markov process. In particular, the EMH justified the "random walk" model of prices, and reliance on the Central Limit Theorem allowed this random walk to be further approximated as a Wiener process with Gaussian transition probabilities ([6,8]). These factors came together to create an "elegant" analytical framework that could be manipulated on paper to generate useful results – a prime feature for most academic researchers. As the academic consensus grew and students of business schools flowed into Wall Street, these features became attractive to financial practitioners as well.

While the EMH has been tested with historical data many times since its formulation, the meaning of the results of these studies has been very much in the eye of the beholder. By and large, the random-walk price model justified by the EMH has been shown to approximately model price evolution most of the time. The key caveats are "most of the time" and the significance of historical market data in this context, as no objectively perfect measure of the relationship between market value and "real" value exists. Figure 3, taken from [9], is a simple example showing how a Gaussian model of percentage changes in the Dow Jones Industrial Average (DJIA) (with σ = 1.032% from 1928 – 2007) became grossly flawed as the financial crisis reached its peak (the exceptions shown in the table are for October 2008). It was also wildly incorrect during the Great Depression years of the 1930s. The more stable years in between are much closer to Gaussian but are not exactly Gaussian, illustrating the limits of the Gaussian assumption even under relatively benign market conditions.

Flaws in both the Gaussian model of market price changes and the assumption of "rational" behavior underlying it have become much better understood over the past 20 years, particularly due to the rise of so-called "Behavioral Economics" [5]. Why, then, did belief in the correctness of the EMH maintain such a hold on mainstream financial thought? One of the best-known experts in quantitative finance, Paul Wilmott, identified the reason quite bluntly in one of his recent books (when discussing the applicability of the Gaussian distribution) as follows ([6], pp. 33-35, emphasis added):

> In finance we often assume that equity returns are normally distributed. … We find ourselves using the normal distribution quite naturally for many financial processes.
>
> As often with mathematical 'laws' there is the 'legal' small print, in this case the conditions under which the Central Limit Theorem applies
>
> • The random numbers must all be drawn from the same distribution;
>
> • The draws must all be independent;
>
> • The distribution must have finite mean and standard deviation;

> Of course, financial data may not satisfy all of these, or indeed, any. In particular, it turns out that if you try to fit equity returns data with non-normal distributions you often find that the best distribution is one that has infinite variance. Not only does it complicate the nice mathematics of normal distributions and the Central Limit Theorem, it also results in infinite volatility. *This is appealing to those who want to produce the best models of financial reality but does rather spoil many decades of financial theory and practice based on volatility as a measure of risk for example.*
>
> However, you can get around these three restrictions to some extent and still get the Central Limit Theorem, or something very much like it. … *We tend to assume that equity returns are normally distributed, and equivalently, equities themselves are lognormally distributed.*

This statement makes the reality clear. It is obvious to those who retain an open mind that the normal (Gaussian) distribution component of the random-walk (strictly speaking, the Wiener-process) model does not apply to financial markets, but assuming that it does is too convenient to give up. "Many decades of financial theory" are based upon it, and these theories are so useful that it is better to force-fit the Gaussian distribution model to financial processes by "getting around" the restrictions of the Central Limit Theorem. It is fair to say that, in this case, academic and professional preference for tractable, analytical, easy-to-use models trumped the search for "truth" a long time ago.

3.2 The Gaussian Copula Default Correlation Model

As bad as the flaws in the EMH are, these flaws were grossly compounded by succeeding financial models that assumed the EMH was perfect true to the n^{th} degree. One of these successor models formed the key mathematical framework behind the CDO risk model described in Section 2.0. This was the so-called "Gaussian copula formula" for modeling correlations between loan default risks. As explained in [10], this theoretical correlation model was adopted for loan defaults in 1999 by David Li of RiskMetrics [11]. The simplification of correlation into a single parameter made it easier to fit historical data on default risk correlation to a tractable model. Despite this, the relative rarity of defaults prior to 2000 made any correlation model based on historical default data highly uncertain. Li got around this problem with a clever (and disastrous) insight: since the EMH stated that market prices were the best possible estimates of true value, statistical correlations inferred from these prices (in this case, prices of CDSs) were also the best possible estimates. Given the dearth of historical data on defaults themselves, this aggressive extrapolation could not really be tested, but no test was necessary, given the level of acceptance of the EMH by 2000.

The resulting market-driven copula model for default risk correlation became instantly popular, as it enabled the creation of complicated CDOs and other mortgage-backed derivatives without the in-depth analysis that would normally have been needed [10]. It should be noted that generating single-parameter correlation estimates (assumed to be constant over time) from CDS prices required a historical database of market prices for CDSs. CDS market data was far more extensive than data for actual defaults but was limited to the period since about 1998 when CDS trading volumes became significant [10]. This period did not see any significant declines in home prices and thus was not at all representative of the situation beginning in 2007, when home prices began falling simultaneously across much of the country. Under rising home-price conditions, defaults did occasionally occur due to personal misfortunes but were scattered and mostly uncorrelated. However, when home prices fell nationwide, similar financial stresses hit homeowners everywhere, many of whom had been extended credit beyond their means in the rush to satisfy investors' craving for more and more CDOs.

As noted in Section 2.0, the sudden and dramatic change in the correlations among home mortgage defaults completely destroyed the copula model and caused the collapse of the "tranching" system on which CDOs were based. Suddenly, the 1st tranche of CDOs was almost as vulnerable as the lower tranches, and some CDOs began to stop paying the 1st tranches, the resulting losses were far beyond what investors (and those carrying the risk of default through sales of CDSs to CDO investors) could handle. Reserves for losses were insufficient primarily because widespread, near-godlike faith in the EMH and the copula correlation model made losses of such magnitude seem impossible. But the lack of sufficient reserves was a broadly-shared problem throughout corporate America for reasons that will be discussed in the next section.

3.3 "Value-at-Risk" (VaR) Modeling

Another important offshoot of the EMH and the analytical risk models that flow from it is what is known as "Value-at-Risk" (VaR) form of company-wide or investor-wide risk modeling. This approach was developed by JPMorgan's RiskMetrics group in the early 1990's with the express intent of providing a standardized risk-modeling and reporting method that would be used widely in the business and financial world (note that the RiskMetrics group, which was later spun off into an independent company, was where David Li worked when he developed the copula model) [12]. To encourage widespread adaptation, RiskMetrics published the basic methodology and made it available for free, while earning income from consulting and implementing the approach for individual companies. VaR uses simplified, straightforward extrapolations of the EMH to allow rapid, near-real-time calculation of a single output quantity known as the "Value at Risk" or "VaR". This single number is easy (perhaps too easy) for managers to understand – it expresses the maximum loss over a given interval at a predefined (and often unstated) confidence-interval percentile. For example, the weekly VaR for a given company aims to represent the maximum possible loss at the 95th or 99th percentile (typically one of these is used and is defined beforehand) over the next week of company-wide activity [12].

The flexibility, speed, ease of use, and ease of understanding of the single VaR output made VaR an attractive basis for risk reporting regulations issued by the U.S. Securities and Exchange Commission. By 2000, VaR had become the industry standard originally desired by JP Morgan [12]. Despite this, the limitations of VaR were obvious and should not have surprised anyone. By definition, VaR only models risk to the 95th or 99th percentile, which is the furthest that "nominal market" extrapolations of the EMH can be taken to even approximately hold. Therefore, it is of little use in predicting what might happen during events like the correlated home market decline of 2007-2008, although even this event should not have been so surprising given the evident extremes that home prices had reached (in comparison to incomes) by 2006. However, the very simplicity of the VaR output that encouraged its widespread adoption allowed managers to forget that VaR was only useful to, at most, the 99th percentile. It quickly became thought of as an actual "worst-case" bound on losses and treated as such in management and risk-based portfolio optimization. This led to the maintenance of loss reserves throughout the economy that were far short of what was needed, which allowed companies to boost reported earnings. Worse, as noted by the well-known risk-management skeptic Nassim N. Taleb, it actually encouraged foolhardy trades that would consistently pay off 99% of the time but which led to disaster under unlikely (and ignored by VaR) default scenarios [12,13].

As described in [12], a healthy academic debate exists as to what degree VaR is useful within the 95th – 99th percentile of risks and to what extent interpretations of VaR by risk professionals can be useful in estimating the threat posed by more extreme risks. But the fact that VaR calculations do not include the extreme risks that can destroy investors and companies should not have been so thoroughly ignored by managers and decision-makers. In retrospect, approaches like VaR that try to oversimplify risk to the point where managers think they fully understand it are worse than useless since they are so likely to be abused. VaR is far from useless for experts in risk modeling, but these experts should not need the crutches and simplifications of VaR. Their job should be to understand risk in all of its complexity and to communicate that risk to decision-makers (qualitatively, if necessary) as fully as possible.

3.4 The Fault of Mathematicians

This section has described several ways in which gross simplifications and distortions of financial reality contributed to the financial debt-market collapse. While the vast majority of the financial community was blind to these flaws (unwittingly or wittingly) prior to the financial crisis, some experts did warn of them repeatedly and loudly, only to be ignored. Prior to the crisis, many people warned of a bubble in the housing market that would be disastrous for the credit markets when it "popped." However, this inconvenient fact has been downplayed by those looking for bailouts, as they would prefer to pretend that the crisis "could not have been foreseen" and represented the equivalent of a magnitude-9 earthquake in a seismically-quiet region.

Several experts in and proponents of quantitative finance modeling have highlighted weaknesses in the underlying theory and means by which the models have been mis-applied by those who should have known better. Paul Wilmott, quoted regarding the Gaussian market-return assumption in Section 3.1, published a paper in 2000 in which he highlighted examples of misuse. The abstract to this paper is quoted in full below (emphasis added) [14]:

> The once 'gentlemanly' business of finance has become a game for 'players'. These players are increasingly technically sophisticated, typically having PhDs in a numerate discipline. The roots of this transformation have their foundation in the 1970s. Since then the financial world has become more and more complex. *Unfortunately, as the mathematics of finance reaches higher levels so the level of common sense seems to drop.* There have been some well-publicized cases of large losses sustained by companies because of their lack of understanding of financial instruments. In this article we look at the history of financial modelling, the current state of the subject and possible future directions. *It is clear that a major rethink is desperately required if the world is to avoid a mathematician-led market meltdown.*

This blunt, specific, and correct prediction of the crisis 8 years before it occurred has greatly enhanced Wilmott's reputation within the broader financial community. It has also motivated Wilmott to expand his efforts to mitigate the ongoing misuse of quantitative finance. He and another well-known expert, Dr. Emanuel Derman (formerly of Goldman-Sachs), have recently created "The Financial Modelers' Manifesto," which imitates the form of "The Communist Manifesto" but is far more useful. The full version of this document can be found online [15]; an extended excerpt follows (emphasis added):

> Our experience in the financial arena has taught us to be very humble in applying mathematics to markets, and to be extremely wary of ambitious theories, which are in the end trying to model human behavior. *We like simplicity, but we like to remember that it is our models that are simple, not the world.*

Building financial models is challenging and worthwhile: you need to combine the qualitative and the quantitative, imagination and observation, art and science, all in the service of finding approximate patterns in the behavior of markets and securities. The greatest danger is the age-old sin of idolatry. Financial markets are alive but a model, however beautiful, is an artifice. No matter how hard you try, you will not be able to breathe life into it. *To confuse the model with the world is to embrace a future disaster driven by the belief that humans obey mathematical rules.*

MODELERS OF ALL MARKETS, UNITE! You have nothing to lose but your illusions.

The Modelers' Hippocratic Oath

- I will remember that I didn't make the world, and it doesn't satisfy my equations.
- Though I will use models boldly to estimate value, I will not be overly impressed by mathematics.
- I will never sacrifice reality for elegance without explaining why I have done so.
- Nor will I give the people who use my model false comfort about its accuracy. Instead, I will make explicit its assumptions and oversights.
- I understand that my work may have enormous effects on society and the economy, many of them beyond my comprehension.

While this manifesto is focused on financial modeling, its lessons are broadly applicable to all aspects of mathematical modeling under uncertainty. Decades of experience in quantitative financial modeling suggests that, where mathematical models are concerned, the more useful they appear to be, the less-connected they are with reality. Even this is not a new observation – it goes back at least to Albert Einstein, who observed in 1921 that "As long as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality." [16] If this principle applies to physics, it certainly applies to all applications of mathematical risk modeling in engineering.

4.0 GENERAL RISK-MANAGEMENT LESSONS FROM THE FINANCIAL CRISIS

Given all that has been described above, what lessons can be learned regarding the practice of risk modeling and risk management in engineering? While many books could be written about this subject, a couple of key points should be made clear. First is the fact that precise modeling of the unknown is simply not possible. This represents a fundamental limitation of *deterministic* mathematical modeling of risk, where "deterministic" refers to models in which "randomness" is limited to the assumption of specific, known probability distributions for a relatively small subset of model parameters. The fundamental cause-and-effect relationship is assumed to be known precisely, and unknowns that cannot easily be modeled as random variables (with known probability distributions) are handled by "assumptions" that transform actual uncertainty into apparent certainty. This framework describes almost all models used in practical risk modeling today, and all such models are vulnerable to "unknown unknowns" or "unforeseen extreme events" ("black swans", in the parlance of [13]). As described in Section 5.0 on satellite navigation, conservative assumptions can be made to limit this risk, but it cannot be completely eliminated.

For these reasons, *probabilistic* models are much more realistic when the level of uncertainty is significant. "Probabilistic" here refers to models that do not simplify or assume away uncertainty. Instead, uncertainty is modeled with as many probability distributions as are needed to represent the modelers' actual level of ignorance. For example, instead of assuming that a Gaussian distribution with known mean and standard deviation describes equity market returns (in the financial case), a truly probabilistic model would include multiple potential distributions (one of which might be Gaussian) and parameters with different probability weightings, and the weights themselves might be random variables. Choosing weightings and "probabilities of probabilities" must still be done by imperfect modelers, but the entire emphasis is different. Instead of attempting to produce specific numerical bounds on uncertainty, convolving the multitudes of probabilities in this approach (typically via Monte Carlo simulation) results in an output histogram that offers no such comfortable guarantees. Simply put, deterministic models provide precise quantification of uncertainty whose accuracy and precision are illusory because they depend wholly on the assumptions used to generate the results. Probabilistic models also produce imprecise outputs, but the imprecision is real, and the goal of these models is to identify this lack of precision as opposed to covering it up.

Because the probabilistic approach is so philosophically different from the deterministic one, it is likely that more-traditional deterministic risk models will remain dominant. As noted above, these models require multiple *assumptions* regarding uncertain behavior and *simplifications* to make the resulting model tractable and useful for analysis. As noted above, once useful models have been created through this process, the danger is forgetting how they were created and growing to believe in them too strongly while ignoring all contrary data, as happened with the EMH. To avoid this, the assumptions and simplifications on which deterministic risk models are based should be highlighted not only during the modeling process but also when results are presented. Since the results are, at best, only as good as the underlying assumptions, it is fair to say that the assumptions and simplifications remain more important than any results that may be achieved. If these shaky foundations are consistently emphasized, fewer people will be tempted to willfully or accidentally mis-interpret the results. Researchers will also be much less likely to extrapolate from one flawed model to another without consideration of these weaknesses. The copula model for default correlation described in Section 3.2 was a perfect example of an extrapolation from a flawed model (the EMH) into an absurd one, where mathematics ruled and common sense was completely thrown overboard.

5.0 LESSONS APPLICABLE TO SATELLITE NAVIGATION INTEGRITY ASSURANCE

The risk-modeling lessons of Section 4.0 have multiple applications to satellite navigation and, in particular, integrity assurance for satellite navigation. First, we must recognize that integrity or safety assurance is a unique application of risk assessment in which the aim is to protect passengers from the consequences of very rare but potentially hazardous threats. Accordingly, a great deal of conservatism is often applied when assessing these threats using deterministic risk models [17]. The following subsections describe specific characteristics that are shared between the risk models that led to the financial crisis and the much more conservative ones used in modeling satellite navigation integrity risk.

5.1 Use of the Gaussian Distribution

As in the financial world, the Gaussian probability distribution is used extensively to model nominal error behavior and to compute position-domain protection levels that are intended to bound worst-case user position errors at the integrity-risk probabilities required for user safety. The Gaussian model is also a convenient and efficient means of communicating ground-system errors to users of GNSS augmentations such as Space-Based and Ground-Based Augmentation Systems (SBAS and GBAS, respectively) in terms of a single parameter: the standard deviation or "sigma" of range-domain errors [17,18,19].

Fortunately, unlike the blanket assumption of the Gaussian distribution in the EMH and its progeny, great care is taken in using the tails of the Gaussian assumption to bound rare-event errors under nominal conditions (so-called "rare-normal" errors). Extensive data studies of GPS, SBAS, and GBAS data have shown that, while the Gaussian distribution approximately holds in many cases and is usually a good model within the 99th percentile of errors, it is not a good description of rare-event behavior. In particular, rare-event "tails" of actual data often considerably exceed what is predicted by the Gaussian distribution. Several reasons exist for this, but the dominant one is the phenomenon of "mixing" of errors with different underlying actual distributions [17,18]. This makes sense when one considers that "rare-normal" errors are not really "normal" but instead are combinations of various "off-nominal" conditions that have different causes.

Because use of the Gaussian distribution is built into the standards for SBAS and GBAS, the primary defense against its inapplicability at low probabilities is to inflate the sigmas broadcast by SBAS or GBAS (or assumed in user equipment) such that the assumed Gaussian distribution "overbounds" the actual, unknown (and likely very complex) error distribution at the probabilities that matter for user safety. Given the limited usefulness of theory and the limited amount of statistically-independent data that can be collected in an attempt to identify the "real" error distribution, this is a difficult problem. Several approaches to deriving bounding inflation factors from collected data have been published, e.g., see [20,21]. No matter which method is used, no means of "proving" rare-event error bounding by Gaussian distributions exists or can exist, given that the required assumptions cannot be proven. Despite this, the use of conservatism and common sense in deriving inflation factors (and then applying additional margin for "unknown unknowns") should sufficiently cover the underlying uncertainty.

Even after inflation has been applied, the reliance on Gaussian error models becomes much more critical when they are extrapolated to derive distributions for squares of errors, as is done in Receiver Autonomous Integrity Monitoring (RAIM) and in real-time monitoring of the broadcast sigma parameters. As with the derivation of the copula model for default correlation in Section 3.2, whatever errors exist in the Gaussian error model are greatly magnified when the errors are squared and then assumed to follow a chi-square distribution. For this reason, great care should be exercised when modeling any distribution that is derived from a starting assumption of a Gaussian distribution. This applies to results within the 99th percentile as well as to rare events.

5.2 Reliance upon Historical Data

As with financial risk quantification, historical data of GPS performance is used to build models of failure probabilities and anomaly behaviors, and this process suffers from a lack of data due to the recent heritage of GPS, which was not fully commissioned until 1995. For example, estimating the prior probability of sudden, unpredictable failures in GPS satellites is mostly based upon the observed failure history of GPS satellites in orbit since 1995, but such failures are quite rare and are not consistent across all GPS satellites. They occur more frequently as GPS satellites approach end-of-life, and they change as different GPS satellite types (Block II, IIA, R, etc.) are deployed over time. In general, there is no guarantee that future GPS satellite or GPS Operational Control Segment performance will correspond to what has been observed in the past. Thus, it is risky to estimate one failure rate across all GPS satellites [22].

For SBAS and GBAS, conservatism and common sense must again be applied to limit the impact of these uncertainties. Failure-rate estimates are made from data where different satellites are combined, but significant margin is applied to account for the differences between satellites. The resulting prior probabilities for failures that might threaten user integrity are conservative for all fault types and are extremely conservative for others where limited or no data exists. The problems that come from reliance on limited historical data are even more severe when "threat models" are created to represent possible system behaviors when a particular fault or anomaly (e.g., satellite signal deformation, ionospheric storms) occurs. In the case of satellite signal deformation, deterministic threat models have been extrapolated from a single observed event, the fault on SVN 19 discovered in 1993 [23]. As described in Section 5.4, worst-case assumptions are typically used to handle uncertainty within the resulting threat models.

5.3 Correlation of Errors and Failures

The failure of the Gaussian copula model described in Section 3.2 to properly handle the correlations between loan default risks highlights the problem of modeling uncertain and potentially time-changing correlations in general. For satellite navigation, this problem breaks down into two categories: *error correlation* and *anomaly correlation*. Correlation among nominal errors is relatively easy to deal with because significant data exists – one does not have to wait for anomalous conditions. However, even when truly uncorrelated data is present, the statistical "noise" inherent in correlation coefficients estimated from data is always non-zero. Since the designer cannot tell whether real correlation exists or not, the resulting error sigmas must conservatively allow for significant non-zero correlations.

One example of this in GBAS is the siting of ground-system reference-receiver antennas far enough apart (~ 100-200 meters) such that diffuse multipath (and most specular multipath) should be statistically independent from receiver to receiver. However, this cannot be guaranteed, and even if it is true at a given site, statistical correlation estimates will be non-zero. Therefore, the assumption that nominal errors in the resulting pseudo-range corrections are reduced by a factor of two when averaging measurements across four reference receivers is not valid. Conservative handling of the estimated correlation at a given site can properly "de-weight" the assumed credit given for averaging [24], or the designer can choose to take no averaging credit at all.

On the other hand, modeling correlations among rare-event anomalies mirrors the problem faced in modeling correlations among loan defaults and is very difficult. Unlike loan defaults, which are clearly correlated to some degree, GNSS satellite failure correlations are very hard to foresee because of our limited understanding of their causes [22]. The temptation to ignore correlations and to treat all failures as statistically independent is very high, as this allows simplified probability models to be used and results in probabilities of multiple failures that are usually small enough to be ignored. But this is a dangerous trap that could result in neglecting important sources of integrity risk. Avoiding this pitfall requires assuming some non-zero degree of failure correlation, but without detailed failure cause-and-effect information, it is very difficult to know how much correlation is sufficiently conservative in a deterministic risk model. This is one place where probabilistic models are far superior, as our degree of uncertainty regarding actual failure correlations can be handled directly by representing different correlation scenarios, or possible states of reality, and assigning probability weights (which can themselves be random variables) to each one.

*5.4 Worst-Case Failure Approximation*s

Since the degree of uncertainty inherent in the development of deterministic failure threat models is well-understood, the resulting threat models are usually applied in terms of the worst-case fault within the bounds of the threat model. Once one agrees to ignore the possibility of faults exceeding the threat model bounds, this worst-case-fault assumption is the most conservative one possible. Note that the "worst-case fault" is judged from the user's point of view, rather than that of the GNSS or service provider. For example, the worst-case C/A-code signal-deformation on a GPS satellite depends upon the design of the reference receiver providing differential corrections (if any) and the design of the user receiver. SBAS and GBAS users are allowed a pre-specified "receiver design space," and given the reference receiver chosen by a given SBAS or GBAS installation, finding the worst-case signal-deformation fault requires error maximization over all possible deformations in the threat model and all possible user receiver design parameters [23,25].

The GBAS (LAAS) threat model for anomalous iono-spheric spatial gradients in CONUS is a good example of the effects of worst-case fault modeling. Figure 4 shows a simplified, linear model of a large, "wedge-shaped" spatial gradient affecting a GBAS installation, and Figure 5 shows a graphical summary of the parameter bounds of this threat model [26,27]. The geometry assumed in Figure 4 is a simplification of reality and cannot be assumed to hold precisely, even though the threat model assumes that it does. Fortunately the resulting risk assessment is not very sensitive to small deviations from a

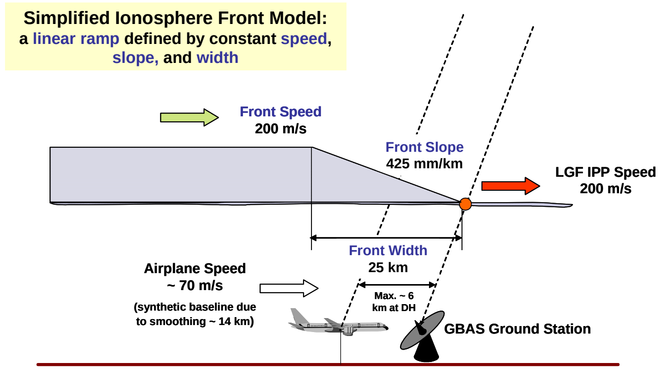


Figure 4: Geometry of GBAS (LAAS) Ionospheric Threat Model for CONUS

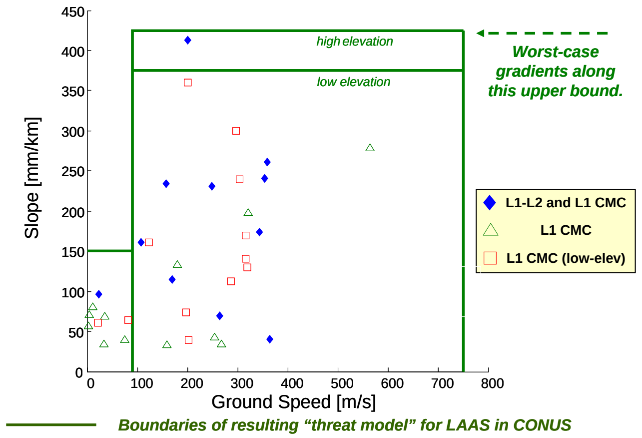


Figure 5: Parameter Bounds on GBAS (LAAS) Ionospheric Threat Model for CONUS

perfectly linear front slope. This kind of sensitivity analysis is what is required to test our vulnerability to violations of deterministic models whose underlying assumptions cannot be verified.

The parameter bounds shown in Figure 5 cover the worst "validated" ionospheric gradients observed in CONUS since 1999 [27]. They cannot be guaranteed to cover future anomalies; thus ongoing monitoring of ionospheric anomalies is required to see if these bounds need updating in the future. However, the outer bounds of the existing threat model appear to be very conservative because they come from a single ionospheric storm on a single day (20 November 2003) in a small region of CONUS (Ohio). The other observations shown on Figure 5 represent examples of other anomalous conditions and do not include the vast majority of otherwise-anomalous gradients with slopes under 200 mm/km, which are generally not threatening to GBAS users. Therefore, in a probabilistic model, the vast majority of the weighting (given that an anomaly condition exists) would go toward non-threatening gradients with tolerable slopes, a small fraction would go to the 200 – 300 mm/km slope range, a much smaller fraction to the 300 – 425 mm/km range, and

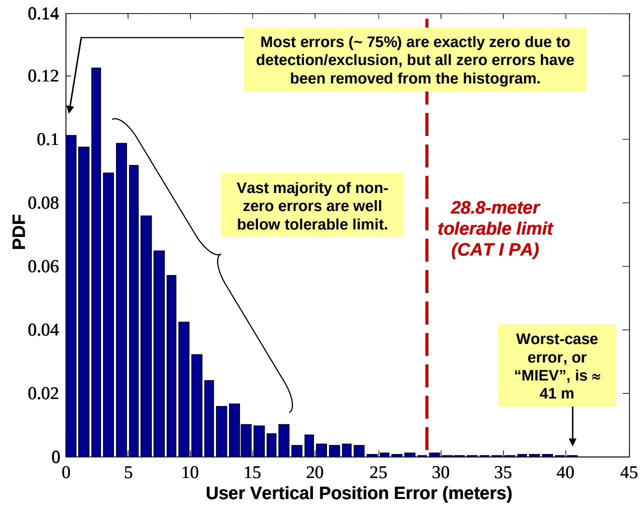


Figure 6: Near-Worst-Case Ionosphere-Induced Vertical Position Errors at Memphis

then a very small but non-zero fraction to gradients above 425 mm/km (the upper bound in Figure 5) that have not been observed to date but cannot be ruled out.

Given this uncertainty within a deterministic model, the worst-case gradient of 425 mm/km (for high-elevation satellites) is assumed to be present at all times, and its hypothetical presence is simulated, with the worst possible approach geometry and timing relative to a single approaching aircraft, on all pairs of satellites otherwise approved by a LAAS Ground Facility, or LGF. The largest resulting vertical position error over all potential user satellite geometries represents the Maximum Ionospheric Error in Vertical (Position), or MIEV, that must be protected against. Before mitigation by LGF geometry screening (see [26,29]), this worst-case error can be as large as 40 – 45 meters. Figure 6 illustrates the potential magnitude of vertical errors under "near-worst-case" ionospheric anomaly conditions based on a limited probabilistic model that varies front slope (above 350 mm/km), speed, satellites impacted, and approach direction relative to that of the aircraft for a user approaching the LAAS facility at Memphis International Airport with the SPS-Standard 24-satellite GPS constellation [28] (only subset geometries with two or fewer satellites removed are considered) [26]. Note that the worst-case position error or "MIEV" prior to LGF geometry screening is about 41 meters, but the relative likelihood of this result is very low. Much more common are errors in the 5 – 15 meter range (note that the majority of cases where the LGF detects the anomaly before significant error occurs are not included in the histogram). LGF geometry screening acts to remove potential subset geometries (i.e., make them unavailable by inflating the broadcast parameters) whose worst-case error exceeds 28.8 meters, but the price of this is substantially lower availability for CAT I precision approaches [29].

Figure 6 shows the extreme level of conservatism that typically results from deterministic worst-case threat model impact analysis. This level of conservatism is so great that it is hard to imagine that the actual user integrity risk is somehow worse than what is modeled in this manner. However, "hard to imagine" does not equate to "is guaranteed not to happen." The goal of worst-case analysis is to eliminate uncertainty (by assuming the worst possible outcome of the uncertain variables) and thus "prove" that a given probabilistic integrity risk requirement is met. However, the limited knowledge upon which threat models are based means that such "proof" is illusory at best and dangerously misleading at worst. Meanwhile, a great deal of performance (in terms of user availability and continuity) is sacrificed. As shown by the limited example of Figure 6, probabilistic analysis makes it possible to trade off risk reduction and performance benefit in a coordinated manner. The illusion of guaranteed bounds on risk is abandoned, but as the financial crisis illustrates, it is just that – an illusion.

6.0 SUMMARY

This paper provides an overview of the events leading up to the financial crisis of 2008 – 2009 and examines the failures of risk modeling in quantitative finance that led directly to the debt-market collapse that precipitated the crisis. The Efficient Market Hypothesis (EMH) that lies at the core of quantitative finance was known to be fundamentally flawed, but its elegance and convenience blinded most researchers and practitioners from the growing evidence of its weaknesses. Worse, the near-complete acceptance of the EMH motivated researchers to build models that dramatically accentuated its flaws and led to absurd (but eagerly accepted) conclusions, such as the copula model for loan default risk. These models proved dramatically vulnerable to changes in the housing market in 2007 – 2008 and led directly to the ensuing crash.

Fortunately, the gross inattention to potential anomalies and violations of "nominal" behavior that characterize quantitative finance do not apply to satellite navigation integrity assurance. Similar techniques and probability distributions are used, but the understanding of what can go wrong leads to detailed emphasis on modeling and mitigating rare events. Where significant uncertainty exists, conservative assumptions are made in an attempt to be robust to it. As a result, the certification of SBAS and GBAS likely demonstrates that these systems meet their integrity risk requirements with substantial margin.

Despite this, the predominant use of deterministic models for risk assessment is potentially dangerous because it purports to provide "guaranteed" bounds on uncertainty that do not apply in practice. The conservative nature of satellite navigation risk assessment greatly reduces the underlying integrity risk but cannot eliminate it, while it leads to performance losses that potentially have unmeasured safety impacts. Given the degree of uncertainty that is present, probabilistic models are much better suited to providing "illusion-free" risk assessments that enable realistic system level design trade-offs.

ACKNOWLEDGEMENTS

While this paper represents independent research, the author would like to thank the FAA Satellite Navigation Program Office for its support of his research on LAAS and WAAS. However, the opinions expressed in this paper are solely those of the author. The author would like to thank several colleagues who helped him refine the arguments and thought processes that went into this paper, including Per Enge, J. David Powell, Brad Parkinson, Todd Walter, and Sherman Lo. The author would also like to thank the many friends, colleagues, and relatives who have shared their thoughts and experiences of the recent financial crisis with him. The author is particularly grateful to his wife, Peggy Brister, for her ever-present help and support.

REFERENCES

[1] S. Pullen, "Satellite Navigation Integrity Assurance: Lessons Learned from Hurricane Katrina," *Proceedings of ION GNSS 2008*. Savannah, GA., Sept. 16-19, 2008 (presentation). http://www-leland.stanford.edu/~spullen/Katrina%20Lessons%20Learned%20pres%20GNSS-08%20fin.ppt

[2] 10-Year Treasury Note (^TNX) Chart from http://finance.yahoo.com/, Sept. 23, 2009.

[3] P. Hirsch, "Crisis Explainer: Uncorking CDOs," *Marketplace*, Minnesota Public Radio, 2009. http://soundlearning.publicradio.org/subjects/economics_finance/financial_crisis/uncorking_cdos.shtml

[4] J. Jarvis, "The Crisis of Credit Visualized," 2009. http://www.vimeo.com/3261363

[5] J. Fox, *The Myth of the Rational Market: A History of Risk, Reward, and Delusion on Wall Street*. New York: HarperCollins, 2009. http://www.amazon.com/Myth-Rational-Market-History-Delusion/dp/0060598999

[6] P. Wilmott, *Frequently Asked Questions in Quantitative Finance (1st Ed)*. West Sussex, England: Wiley, 2007. http://www.amazon.co.uk/Frequently-Questions-Quantitative-Financial-Engineering/dp/0470058269

[7] "Efficiency and beyond," *The Economist*, July 18th, 2009. http://www.economist.com/displaystory.cfm?story_id=14030296

[8] "Random walk," *Wikipedia, The Free Encyclopedia* (accessed 10/1/09). http://en.wikipedia.org/wiki/Random_walk

[9] P. DeGrauwe, *et al*, "How Abnormal was the Stock Market in October 2008?" *Euro Intelligence*, Nov. 11, 2008. http://www.eurointelligence.com/article.581+M5f21b8d26a3.0.html

[10] F. Salmon, "Recipe for Disaster: The Formula that Killed Wall Street," *Wired Magazine*, March 2009. http://www.wired.com/techbiz/it/magazine/17-03/wp_quant

[11] D. Li, "On Default Correlation: A Copula Function Approach," The RiskMetrics Group, Working Paper 99-07, April 2000. http://www.defaultrisk.com/_pdf6j4/On%20Default%20Correlation-%20A%20Copula%20Function%20Approach.pdf

[12] J. Nocera, "Risk Mismanagement," *The New York Times Magazine*, Jan. 4, 2009. http://www.nytimes.com/2009/01/04/magazine/04risk-t.html

[13] N. Taleb, *The Black Swan: The Impact of the Highly Improbable*. New York: Random House, 2007. http://www.amazon.com/exec/obidos/ASIN/1400063515/nassimtalebsfavo/002-8533486-7104820

[14] P. Wilmott, "The Use, Misuse and Abuse of Mathematics in Finance," *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*. The Royal Society, Vol. 358, No. 1765, Jan. 2000, pp. 63-73. http://www.jstor.org/stable/2666776

[15] E. Derman and P. Wilmott, "The Financial Modelers' Manifesto," Jan. 7, 2009. http://www.wilmott.com/blogs/paul/index.cfm/2009/1/8/Financial-Modelers-Manifesto

[16] A. Einstein, "Geometry and Experience," Lecture Before the Prussian Academy of Sciences, Berlin, Germany, Jan. 27, 1921. English translation at: http://pascal.iseg.utl.pt/~ncrato/Math/Einstein.htm

[17] T. Walter, P. Enge, and B. DeCleene, "Integrity Lessons from the WAAS Integrity Performance Panel (WIPP)," *Proceedings of ION NTM 2003*, Anaheim, CA., Jan. 22-24, 2003, pp. 183-194. http://waas.stanford.edu/~wwu/papers/gps/PDF/WalterIONNTM03.pdf

[18] S. Pullen, "System Overview, Recent Developments, and Future Outlook for WAAS and LAAS," *Proceedings of the Tokyo University of Mercantile Marine GPS Symposium*, Tokyo, Japan, Nov. 11-13, 2002. http://waas.stanford.edu/~wwu/papers/gps/PDF/PullenTokyo02.pdf

[19] S. Pullen and J. Rife, "Differential GNSS: Accuracy and Integrity," in S. Gleason and D. Gebre-Egziabher, Eds., *GNSS Applications and Methods*. Norwood, MA: Artech House, 2009, Ch. 4. http://www.artechhouse.com/Detail.aspx?strIsbn=978-1-59693-329-3

[20] B. Pervan, I. Sayim and B. Khanafseh, "Synthesis of Empirical and Theoretical Approaches Toward the Establishment of GBAS σ_{pr_gnd}," *Journal of Navigation* (The Royal Institute of Navigation), Vol. 58, No. 2, May 2005, pp. 297-313. http://journals.cambridge.org/action/displayIssue?jid=NAV&volumeId=58&issueId=02#

[21] J. Rife, S. Pullen, and B. Pervan, "Core Overbounding and its Implications for LAAS Integrity," *Proceedings of ION GNSS 2004*, Long Beach, CA.,Sept. 21-24, 2004, pp. 2810-2821. http://waas.stanford.edu/~wwu/papers/gps/PDF/RifeIONGNSS04.pdf

[22] S. Pullen, J. Rife, and P. Enge, "Prior Probability Model Development to Support System Safety Verification in the Presence of Anomalies," *Proceedings of IEEE/ION PLANS 2006*, San Diego, CA., April 24-27, 2006, pp. 1127-1136. http://waas.stanford.edu/~wwu/papers/gps/PDF/PullenIONPLANS06.pdf

[23] A. Mitelman, *Signal Quality Monitoring for GPS Augmentation Systems*. Ph.D. Dissertation, Stanford University, Dept. of Aero/Astronautics, Dec. 2004. http://waas.stanford.edu/~wwu/papers/gps/PDF/Thesis/AlexanderMitelmanThesis04.pdf.zip

[24] M. Luo, S. Pullen, *et.al*., "Development and Testing of the Stanford LAAS Ground Facility Prototype," *Proceedings of the ION 2000 National Technical Meeting*, Anaheim, CA., Jan. 26-28, 2000, pp. 210-219. http://www.ion.org/search/view_abstract.cfm?jp=p&idno=25

[25] R.E. Phelts, *Multicorrelator Techniques for Robust Mitigation of Threats to GPS Signal Quality*. Ph.D. Dissertation, Stanford University, Dept. of Aero/Astronautics, June 2001. http://waas.stanford.edu/~wwu/papers/gps/PDF/Thesis/EricPheltsThesis01.pdf

[26] S. Pullen, Y.S. Park, and P. Enge, "Impact and mitigation of ionospheric anomalies on ground-based augmentation of GNSS," *Radio Science*, Vol. 44, Aug. 8, 2009. http://www.agu.org/journals/rs/rs0904/2008RS004084/

[27] S. Datta-Barua, J. Lee, *et al*, "Ionospheric Threat Parameterization for Local Area GPS-Based Aircraft Landing Systems," Submitted to *AIAA Journal of Guidance, Control, and Dynamics*, Aug. 2009. http://www.aiaa.org/content.cfm?pageid=322&lupubid=23

[28] *Global Positioning System Standard Positioning Service Performance Standard*. Washington, D.C., U.S. Department of Defense, October 2001. Internet URL: http://www.navcen.uscg.gov/gps/geninfo/2001SPSPerformanceStandardFINAL.pdf

[29] S. Ramakrishnan, J. Lee, *et al*, "Targeted Ephemeris Decorrelation Parameter Inflation for Improved LAAS Availability during Severe Ionosphere Anomalies," *Proceedings of ION 2008 National Technical Meeting*, San Diego, CA., Jan. 28-30, 2008, pp. 354-366. http://waas.stanford.edu/~wwu/papers/gps/PDF/RamakrishnanIONNTM08.pdf