Interfaces
Vol. 39, No. 5, September–October 2009, pp. 415–427
ISSN 0092-2102 | EISSN 1526-551X | 09 | 3905 | 0415
DOI 10.1287/inte.1090.0452

Correcting Heterogeneous and Biased Forecast Error at Intel for Supply Chain Optimization

Matthew P. Manary
Server Platform Group, Intel Corporation, Hillsboro, Oregon 97124,
matthew.manary@intel.com

Sean P. Willems
School of Management, Boston University, Boston, Massachusetts 02215,
willems@bu.edu

Alison F. Shihata
Planning and Logistics Group, Intel Corporation, Hillsboro, Oregon 97124,
alison.f.shihata@intel.com

In 2007, Intel's Channel Supply Demand Operations launched an initiative to improve its supply chain performance. To ensure success, the process had to fit within the existing planning processes. In practice, this meant that setting service-level and inventory targets, which had previously been external inputs to the process, had to become part of the structured decision-making process. Although other Intel business units had achieved success implementing a multiechelon inventory optimization model, the boxed processor environment posed some unique challenges. The primary technical challenge required correcting for the impact of forecast bias, nonnormal forecast errors, and heterogeneous forecast errors. This paper documents the procedure and algorithms that Intel developed and implemented in 2008 to counter the impact of forecast imperfections. The process resulted in safety stock reductions of approximately 15 percent. At any given time, Intel applies this process to its 20–30 highest-volume boxed processors, determining an on-hand inventory commitment between $50 million and $75 million.

Key words: forecasting: applications; industries: computer/electronic; inventory/production: applications, multiechelon safety stock optimization.
History: Published online in *Articles in Advance* August 27, 2009.

Since 2005, Intel has used more rigor in its inventory-planning processes, as manifested in two ways. First, Intel has improved its processes by developing an integrated sales, inventory, and operations process (SIOP). Second, it has improved its decision-support technologies by integrating multiechelon inventory optimization with its supply chain planning solvers.

The SIOP process must operate within the constraints of Intel's existing organizational responsibility structure. In particular, the sales and marketing organization is responsible for entering the forecast data that are so critical to any SIOP process. At Intel, three forecasting problems—forecast bias, heterogeneity of errors, and nonparametric residuals—have nontrivial impacts on the output of its optimization models. Traditional forecasting textbooks, including Crum and Palmatier (2003), Franses (1998), and Montgomery et al. (1990), recommend addressing these problems prior to loading data into an optimization model. However, as Manary and Willems (2008) document, planners cannot always remove bias from the raw forecast data, especially when groups that do not reside in the supply chain organization have loaded the data.

From a decision-support perspective, it was necessary to address the three forecasting problems. Intel began with forecast bias. In an unbiased forecast, the forecast residuals, found by subtracting the period's forecast from actual demand, have a zero mean; in a biased forecast, the mean is nonzero. Based on the Intel experience that Manary and Willems (2008) cover, we expected bias to be the dominant error behavior to correct. Unfortunately, heterogeneity (i.e., the errors could not be characterized by a single probability distribution) and nonparametric behavior (i.e., errors could not be characterized by a standard probability distribution) error structures complicated solutions to Intel's boxed central processing unit (CPU) forecasts and required a more rigorous solution than bias alone would need. By addressing these forecasting-process problems, Intel could quantify the impact of service and inventory changes as part of its SIOP process. Our approach, which Intel uses today, addressed these issues.

Methodological Contribution

Manary and Willems (2008) proposed a way to correct for forecast bias in inventory optimization models without altering the raw demand signal. Their adjustment procedure provided the necessary improvement required for Intel management to implement multiechelon inventory optimization (MEIO) models for its embedded processor division; the result was a significant reduction in aggregate inventory in comparison to the prior policy, a weeks-of-inventory strategy with management selecting the target number of weeks. The adjustment procedure was conditioned on assumptions that forecast bias was stationary and service levels were predetermined. Intel's Channel Supply Demand Operations (hereafter referred to as Planning), a division with approximately $5 billion in annual boxed CPU sales to distributors, runs similar MEIO models; its business processes also do not allow the supply chain organization to alter the raw forecast data, which the Sales and Marketing Group controls. However, the boxed CPU procedure has one fundamental process difference: service levels are dynamic and optimized as a function of a single demand variance estimate for all service levels. This negates direct application of the embedded processor division's adjustment procedure, which determines a demand variance estimate for each service level.

Our solution to this problem builds upon the adjustment procedure in Manary and Willems (2008) by combining it with Bartlett and Kendall's (1946) methods for testing heterogeneity of independent sample variances. This solution relies principally on applying the adjustment procedure to extract a random sample of variance estimates for a single product across multiple service levels. If we detect no significant difference in variability estimates, we apply a weighted proxy of variability through the root mean of the collective adjustment procedure at the sampled service levels. We treat stock-keeping units (SKUs) with significant differences in variability by using a kernel-smoothing technique that accounts for both heterogeneity and nonparametric error distributions, and generates a single estimate of variability.

This paper presents the solution that Planning developed to determine the joint SKU-location safety stock and service-level targets in the presence of forecast bias, nonnormal forecast errors, and forecast error heterogeneity by reparameterization of the estimate of the standard deviation of forecast errors. Since Intel implemented this process in early 2008, the process has maintained the high service levels affiliated with prior MEIO models; simultaneously, it has reduced total corresponding boxed CPU inventory levels by approximately 15 percent.

Intel's Channel Products Group

Intel's Channel Products Group (for which Planning manages supply and demand coordination) sells boxed CPUs in unique finished goods SKUs that are differentiated by CPU design, boxing material, and warehousing location. The data that this paper displays come from over a year of weekly forecasts of approximately 20 high-volume boxed CPUs that represent Intel's mobile, desktop, and server CPU businesses.

The boxed CPU supply chain is similar to a standard reseller's supply chain but includes no product manufacturing. Instead, the supply chain modeling focuses on boxing processes and distribution. Each SKU is modeled as a worldwide distribution network of packaged and unpackaged finished goods CPUs that consist of three echelons with multiple physical locations. CPU manufacturing echelons are beyond the scope of Planning's distribution network problem. The stages of the supply chain include component warehousing facilities; assembly facilities that box the CPUs and package them with other components, such as fans and heat sinks; and location-specific distribution centers that serve end customers. SKUs share a common supply chain in both echelon structure and

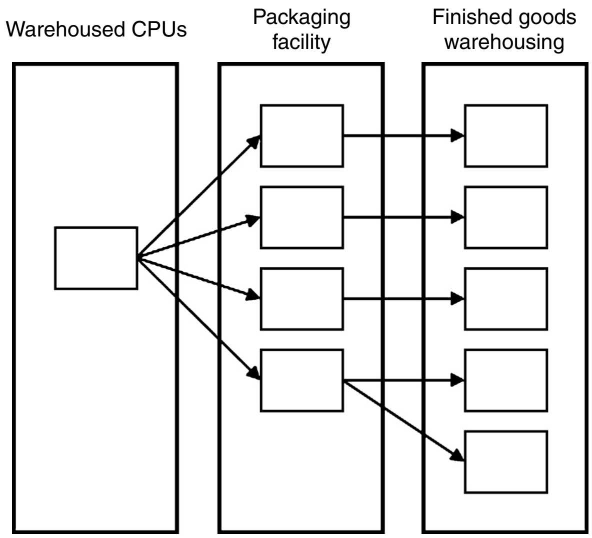

Figure 1: The graph depicts the three echelons in Intel's boxed CPU distribution network model.

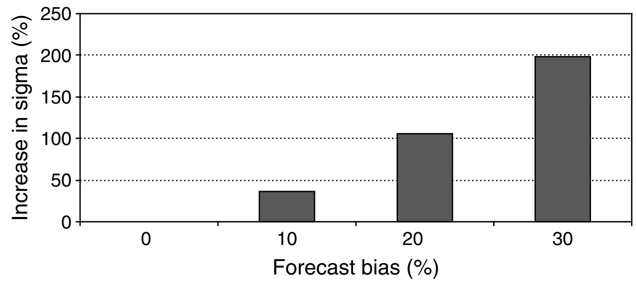

Figure 2: A squared error standard deviation estimate grows exponentially as we introduce bias into the forecast. At a 20 percent forecast bias, the sigma generated from a sum-of-squared-errors approach more than doubles the estimate generated when bias is absent.

number of stages (Figure 1). By echelon, the boxed CPU distribution network has (1) one virtual component stage that represents nonboxed, finished CPU inventory from Intel's factory network, (2) four stages in boxing and kitting assembly stages, and (3) five stages in geography-specific distribution centers. The MEIO model generates inventory targets for between 200 and 300 SKU locations every month.

The planning organization monitors inventory levels weekly; a major reset, which is driven by a new forecast from the sales and marketing organization, occurs monthly. Based on the new demand information, the monthly planning process resets the forward-looking demand at the SKU-location level. This reset modifies the current optimized safety stock targets; the targets feed an Intel-developed Systems Application and Products (SAP) advanced planning and scheduling (APS) optimizer that minimizes production costs, lost-sales costs, and costs for deviating from the inventory targets. The APS optimizer output is the constrained production and inventory plan, which constitutes the finalized factory schedule.

Addressing the Breakdown with Intel's Approach

Although operating parameters, such as transit times, are essentially identical across SKUs, forecast accuracy and forecast error homogeneity differ significantly. In the past, Planning's optimization process relied on the commonly employed standard deviation of forecast error (SDFE) calculation to provide the demand variability input,

$$\hat{\sigma}_{\mathrm{SDFE}} = \sqrt{\frac{\sum_{i=1}^{n} (F_i - A_i)^2}{n-1}}, \tag{1}$$

where F and A are the paired forecast and actual demand, respectively, for a period i. Equation (1) is influenced by any bias and (or) heterogeneity in the relationship between the forecast and actuals (i.e., the forecast errors). Figure 2 demonstrates how even small amounts of bias result in a significant increase in the calculated SDFE. The boxed CPU global forecast bias exceeded 20 percent for products whose safety stock targets were set using the optimization process.

Intel's optimization technique relies largely on the demand variance estimates to derive safety stock targets; therefore, the forecast bias naturally impacted Planning's modeling output. Figure 3 presents the bias problem that Planning faced using the sales and marketing forecasts. MEIO SKUs were collectively overforecast more than 60 percent of the time; the expected rate was 50 percent. Because Planning did not initially address the forecast bias (Figure 3), we can look at Figure 2 to understand how targeted inventory and service levels differed from the levels that Intel actually achieved.

Addressing the Boxed Processor Forecast Bias

In 2005, Intel's Communications Infrastructure Group (CIG) encountered a similar problem with forecast

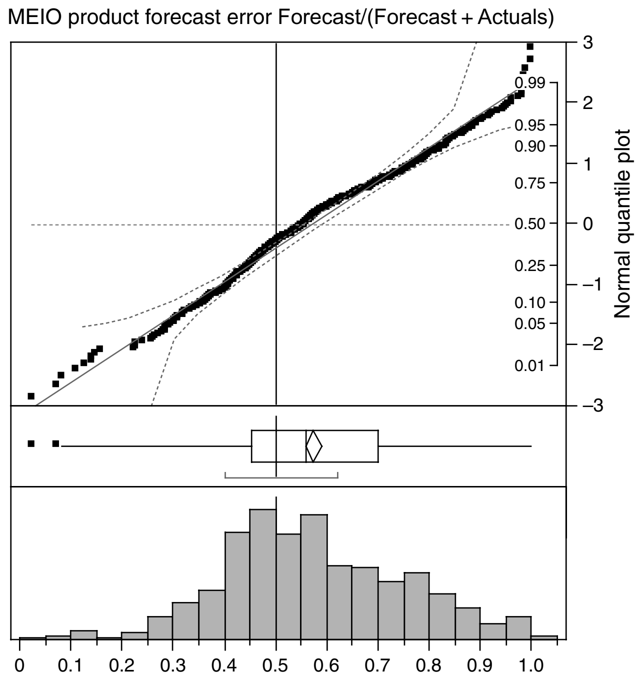

Figure 3: The graph shows the distribution of relative forecast errors, i.e., Equation (2), for boxed CPU MEIO products representing over 500 paired observations. The distribution's heavier right tail represents an overall propensity to overforecast. At the SKU level, products displayed overforecasting bias, underforecasting bias, and no bias.

bias in its MEIO models. To solve its problem, CIG implemented a modified estimate of the demand variation that Manary and Willems (2008) discuss. The premise was to calculate a measure of relative forecast accuracy,

$$\theta_i = \frac{\text{Forecast}_i}{\text{Forecast}_i + \text{Demand}_i}, \qquad (2)$$

and reconstruct from the pattern of error measurements, whether biased positively, negatively, or unbiased, a modified estimate of SDFE,

$$\widehat{\sigma}^{\text{Modified}} = \text{Max}\left[\left[\frac{((1-\theta_\beta)/\theta_\beta)-1}{t_{\beta,\,df}}\right]\mu, 0\right], \qquad (3)$$

where $\beta = 1-\alpha$, θ_β denotes the quantile point corresponding to β from the distribution of θs calculated in Equation (2), $t_{\beta,\,df}$ is the Student's t-distribution with a cumulative density of β and degrees of freedom coming from the number of historical points from which to draw, and μ is the average demand. By treating SDFE as a function of service level, Equation (3) factors out the MEIO model impact of bias in the forecast without changing the mean demand signal from sales and marketing. That is, Equations (3) and (6) later in the paper correlate negatively with bias. Because a product is more likely to be overforecast, based on its historical performance, the sigma estimates decrease, thus telling the planning system to plan for less safety stock (because the biased forecast already has "built-in" safety stock). In extreme cases in which a product has never been underforecast, the sigma estimate is zero, thus signaling the planning system to generate no inventory for demand variability. If the product has been underforecast historically, then the sigma estimate is larger than it would be had no bias been present. Products that do not display a bias either way converge to the sum-of-squared-errors estimate. It is worth noting that the method for calculating relative forecast error in Equation (2) is largely arbitrary; any number of forecast error measures would work. However, Equation (2) is a natural selection because Intel planners are familiar with it. If a different error measurement is used (e.g., centering the error on zero), Equation (3) would need to be updated to reflect any change in Equation (2).

Unfortunately, the modified sigma solution that CIG employed was incompatible with the boxed CPU optimization process because forecast errors were heterogeneous, and service levels were selected dynamically. CIG's modified sigma solution requires a stationary bias and was not robust to the heterogeneity present in the boxed CPU forecast error. In addition, boxed CPU service levels were determined using a proprietary selection process that balanced inventory costs, lost sales associated with different service levels, and strategic management guidance. Although the service-level selection process is beyond the scope of this paper, its key attribute was that it required a single constant estimate of demand variability to generate a single efficient inventory frontier prior to the service-level selection.

Determining a Constant Sigma Estimate

Similar to CIG, Planning needed a process to determine an SDFE estimate that would net out the forecast bias and heterogeneity without changing the sales and marketing-loaded forecast. Using Equation (3) was a natural first step; however, the equation's random-sampling approach across service levels to approximate demand variability did not allow for optimizing service-level selection under a stationary sigma assumption. The goal then was to approximate a single representation of demand variability from the random sample generated by sigma modified, which then allowed Intel's third-party software to calculate a single inventory efficient frontier; this aided Planning in ultimately selecting the optimized service level.

Under a normal and unbiased forecast error assumption, sigma modified will generate a random sampling of the true sigma across each service level. Therefore, it is a natural extension to treat samples independently and test for consistency of sigma across the service levels. One complication with this is that most variance equality tests are functions of the underlying data populations; however, because sigma modified is a calculation from one data point, it lacks individual observations to determine comparative statistics. This eliminated adapting a standard F-ratio method or borrowing from approaches developed by O'Brien (1979), Levene (1960), and Brown and Forsythe (1974), among others. However, Bartlett's F-test of homogeneity (Barlett and Kendall 1946) is not a function of observations; it is determined by the sigma estimates themselves. Therefore, Bartlett's approach was adapted to test homogeneity of the sigma-modified sampling (Appendix B) across service levels by simply replacing the variances from independent samples with an array of sigma-modified estimates drawn from different management-determined service levels.

If Bartlett's test (Bartlett and Kendall 1946) indicated homogeneity across the service levels, then the root mean square error (RMSE) of the sigma-modified estimates was established as the point estimate of demand variability for all service levels, thus enabling Planning's SIOP process to determine feasible service-level and safety stock targets. In Figure 4, the variability in the sigma-modified sampling for product A was not significant enough to reject the null hypothesis of equal variances across all service levels; therefore, the pooled variance was used as the constant demand variability parameter in the optimization process.

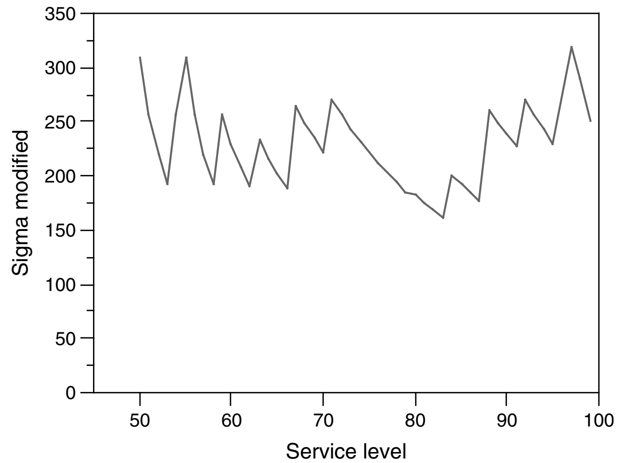

Figure 4: The graph shows sigma-modified estimates at given service levels for CPU product A. Although sigma modified is obviously different for each service level, per the adapted Bartlett's variance test, the estimates do not differ significantly.

We considered product A's sigma-modified variability sampling, and thus its underlying forecast error, to be well behaved (i.e., have no significant bias and be generally homogeneous). Unfortunately, the forecasting behavior of many boxed CPU products did not fit this description. Figure 5 demonstrates our challenge in working with boxed CPU forecast variability, as interpreted through sigma-modified and relative forecast-accuracy histograms. The graphs in Figure 5 illustrate both the lower 50 percent underforecast error distribution and modified sigma sampling at each service level for three products that represent the forecast error behavior across all boxed CPU products. For purposes of setting safety stock, our interest is the propensity to underforecast; therefore, we focus our forecast error analysis on the lower 50 percent of forecast error occurrences that represent the underforecast activity. In theory, each product's forecast error should follow a normal distribution with measurements independent of one another. In such a case, the histogram for the lower 50 percent forecast accuracy will also follow a one-sided (tail portion) normal distribution, in which all θ values from Equation (2) are 0.5 or less. For boxed CPUs, such as product A, some product forecast error follows approximately a normal distribution for the lower 50 percent. Evidence of that lies with the sigma-modified estimates at each service level that appear

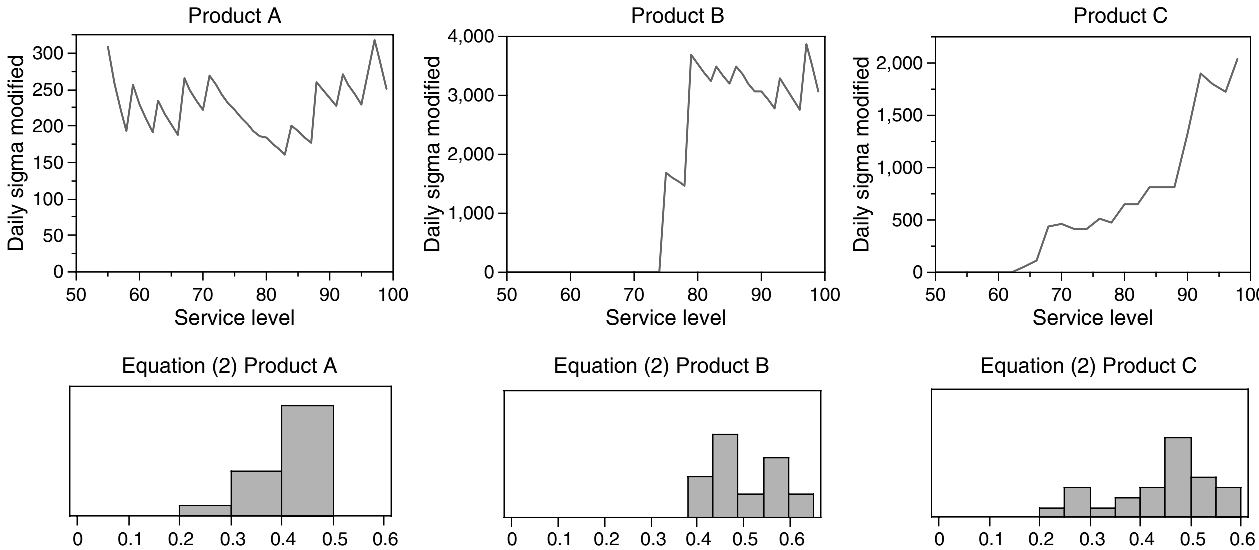

Figure 5: The graphs show sigma-modified sampling for errors and represent the worst 50 percent of underforecasts across service levels for CPU products A, B, and C. Intel describes the forecast error for product A as homogeneous unbiased, product B as homogeneous biased, and product C as heterogeneous biased.

randomly distributed and pass a homogeneity test, as we saw earlier. In contrast, product B displays a bias to overforecast, as evidenced by the histogram of forecast errors, which show that about half of what should have been underforecasts were overforecasts. The calculated sigma-modified values also reflect this because the approximated sigma is zero for service levels up through 75 percent. However, when forecast errors begin to represent underforecasts, the modified sigma values appear randomly distributed and stable, and the product is considered to have a biased forecast; otherwise, the error structure is normal beyond the bias. Product C also displays some level of overforecast bias; however, because its extremely long tail to underforecast breaks with normality, it drives an increasing estimate of sigma as higher service levels are desired. In this case, the product is considered biased, heterogeneous, and nonparametric, as we will see later.

Deriving a Normalized Point Estimate of Sigma with Bias and Heterogeneity

Sigma modified is designed such that when bias to overforecast is demonstrated, the estimate of sigma is zeroed out (because a product that is habitually overforecast will not need safety stock for demand variability). Because of Bartlett's natural log approach (Bartlett and Kendall 1946), testing for homogeneity requires strictly nonzero estimates of sigma. Figure 5 shows how products B and C would clearly fail the equal variance test on just the 0-sigma frequency that bias causes. Figure 6 shows a product count breakdown of the full boxed CPU portfolio behavior. Approximately 40 percent of the boxed CPU SKUs displayed a significant degree of overforecast or underforecast bias. However, Planning management established a minimum allowable service level; thus, we tested only sigma from service levels in the top quartile. This allowed some SKUs with moderate

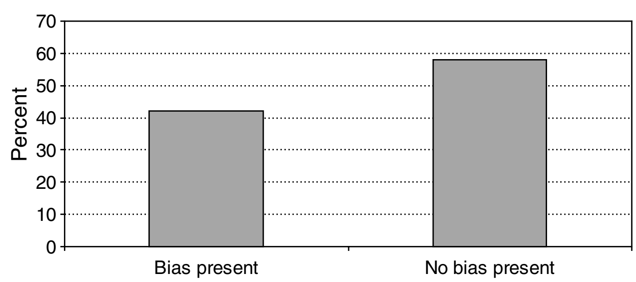

Figure 6: More than 40 percent of Intel's boxed CPU SKUs demonstrated a significant bias based on a binomial test at the $\alpha = 5$ percent level.

bias to pass the homogeneity test. Some products for which bias was not a factor failed because of heterogeneity or nonparametric error.

To allow Planning to run the SIOP process, an alternative to the pooled variance method needed to be developed for the products that failed the homogeneity test. Any estimate developed needed to consider four major factors: bias, heterogeneity, the potential for nonparametric error, and the requirement that it be a single variance estimate translated back to an assumption of normality for the optimization software.

The initial approach, which Intel applied in production in early 2008, was a minimum function of the modified sigma based on both a relative error and an absolute volume error, as Manary and Willems (2008) outline. We knew in advance that this approach would likely overstate a sigma estimate; however, we were comfortable with a conservative estimate for what we thought would be largely an exceptions process. During the spring of 2008, the data showed that almost half the boxed CPU products were failing the homogeneity test. This meant that the sigma estimate for failed tests would play a prominent role in establishing overall inventory targets and thus increase the necessity for a more robust approach. Although the initial approach applied in early 2008 typically delivered a 30 percent reduction in a product's variance estimate (versus a SDFE calculation in approximately 80 percent of the cases), the technique was not an unbiased estimate of sigma and was susceptible to significant overstatement because of heterogeneity and nonparametric error structures. Our biggest concern with the initial technique used in early 2008 was that the sigma estimate was often driven by errors that occurred at forecast levels significantly higher or lower than the current forecast level (i.e., the forecast driving the MEIO model). To correct for this, we wanted to minimize the heterogeneity and (or) nonparametric error impact by isolating a sigma estimate close to the current forecast level. One challenge we faced in doing this was a relatively low number of observations around the current forecast level to draw on as a population. To compensate for this, we decided to apply a weighting system to forecast error observations with decreasing influence as error observations were further away from the current forecast. Ultimately, we found that the most robust technique for calculating sigma under heterogeneous and nonparametric error was to apply a nonparametric approximation of the product error-density function across the full range of prior forecasts, allowing a kernel-smoothing technique to help determine a weighted, localized error pattern. We selected the kernel standard deviation based on Bowman and Foster's (1993) recommendation and then performed a bivariate normal kernel smoother. This allowed us to weight the influences of prior forecasts that were similar in volume to the current forward-looking estimate and helped us to address the problem that the previous uniform weighting approach had caused. Figure 7 presents the completed nonparametric error-probability density function across the history of prior forecasts for product C, a product whose error pattern was seen to be biased, heterogeneous, and nonparametric.

The next step was to identify the error density localized to the current forward-looking forecast. That portion of the probability density function (PDF) usually defined at the forward-looking forecast becomes the curve under which the new estimate of demand variability is weighted. Figures 8(a) and 8(b) represent

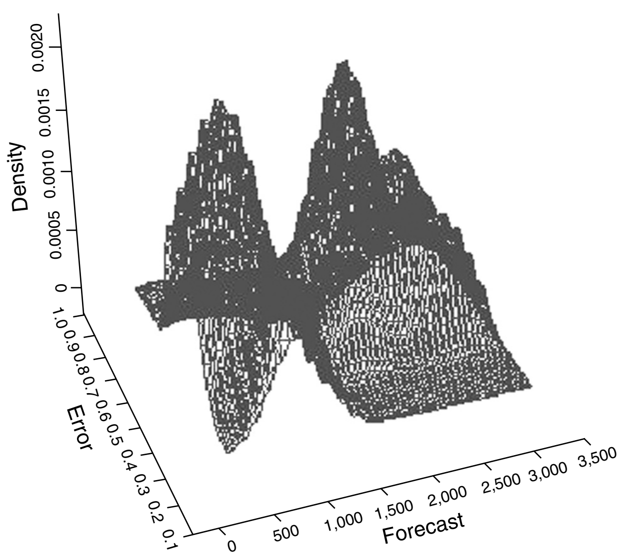

Figure 7: The graph shows CPU product C's nonparametric error-density field across all prior forecast levels. If the relative forecast error was normally distributed with no bias, we would see a bell-shaped "tunnel" lying on the 0.5 error line. However, the error density for product C indicates heterogeneity and nonparametrics.

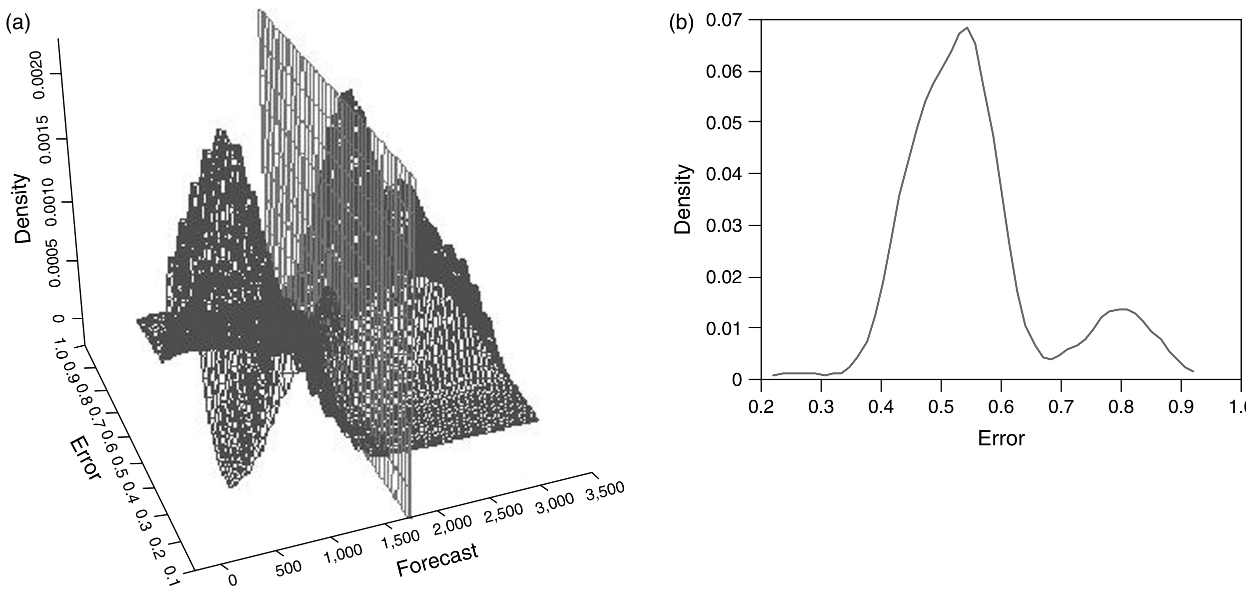

Figure 8: CPU product C's forecast error density field (panel a) is localized at the current forward-looking forecast of 1,708 units (panel b). The distribution localized at the current forecast level determines the product-level sigma estimate.

CPU product C's error density localized to the new forward-looking forecast—1,708 units.

Having obtained an estimate of the forecast error likelihood conditioned on a new forward-looking forecast, the next step is to estimate the expected underforecast and translate it into a singular normalized parameter for the optimization software. In our nonparametric case, we considered the underforecast tail of the conditioned PDF (error values <0.5) and computed the weighted average error in Equation (4).

$$\bar{\theta} = \frac{\int_0^{0.5} P_{\theta|\text{Forecast}}(\theta) * \theta \, d\theta}{\text{CDF}_{\text{Forecast}}(0.5)}. \tag{4}$$

The integrated approach allowed us to weight across the potential heterogeneity of the forecast error. Figure 9 illustrates product C's tail zone for underforecast error conditioned on the forward-looking forecast.

$\bar{\theta}$ now represents the weighted-average underforecast error for product C conditioned on the current forward-looking forecast. We can now reintroduce Equation (3) to normalize the new weighted-demand underforecast. Replacing θ_β with $\bar{\theta}$, and replacing the t-distribution with the average tail value from the standard Normal distribution, generates a normalized point estimate of demand variability back in the units of the forecast,

$$\hat{\hat{\sigma}}^{\text{Modified}} = \left[\frac{((1-\bar{\theta})/\bar{\theta}) - 1}{\sim 0.8}\right]\mu, \tag{5}$$

where μ is the forward-looking forecast. This approach is similar to Bowman and Azzalini's (1997) sugges-

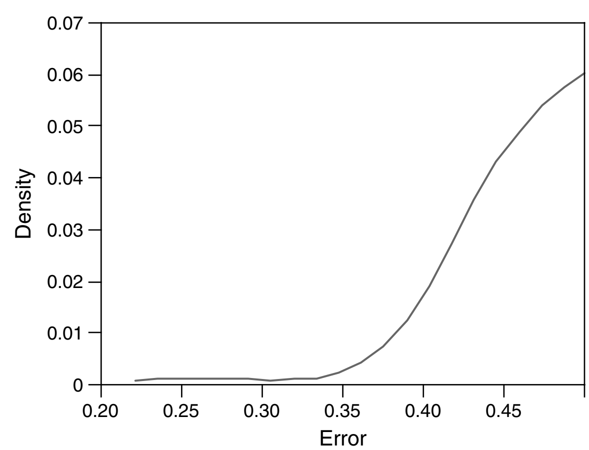

Figure 9: The graph shows product C's underforecast region PDF truncated only to the values representing underforecasts (<0.50 error scores). Under unbiased conditions, the cumulative probability under this curve is 50 percent. For product C, it is ∼ 30 percent, thus indicating a bias to overforecast.

tion of using a median absolute deviation as a more robust sigma estimate when dealing with nonparametric data. However, given the nonparametric nature of Intel's data, and to be conservative, we chose to weight what otherwise would be considered outliers under a median absolute deviation calculation. Equation (5), unlike (3), does not require a maximum criterion because the continuous nature of the kernel estimation will guarantee some positive density to underforecast. Equation (6) will provide an additional weighting that can drive the sigma-modified estimate to zero because an extreme overforecast bias exists.

We have now addressed the heterogeneity at the forward-looking forecast level but still need to account for the bias. We do that by considering the propensity of the bias (either over or under) by weighting Equation (5) with the modeled underforecast likelihood against the expected underforecast likelihood—that is, a ratio of the conditioned, underforecast cumulative distribution function (CDF) up to a value of 0.5 and 50 percent, the CDF for the same region under an unbiased assumption. If the forecast error was unbiased, then this scalar would simply fall out of the equation. If the bias is to overforecast, then the new demand variability is decreased by a relative amount; if the bias is to underforecast, then the new demand variability is increased proportionately. Equation (6) represents the algorithm for determining the weighted point estimate of demand variability conditioned on the forward-looking forecast for use in the SIOP process.

$$\widehat{\sigma}_{\text{Optimized}} = \widehat{\widehat{\sigma}}^{\text{Modified}} * \frac{\int_0^{0.5} P_{\theta\,|\,\text{Forecast}}(\theta)\, d\theta}{0.5}. \qquad (6)$$

It is worth highlighting here that if service levels are deterministic, then we recommend skipping Equations (4)–(6) and using the alternative that Appendix A shows.

Table 1 demonstrates the sigma estimate differences for product C under Equation (6) based on the forecast level. It also displays the more common SDFE approach, Equation (1), to show readers the estimate differences. Figure 10(a) gives the graphical representation of the sigma estimate from Equation (6) at a given forecast level; Figure 10(b) shows product C's error-density field. Note how the Equation (6) sigma estimate peaks at a forecast level of $\sim$1,200 units,

Forecast level Product C	Equation (1) SDFE	Equation (6) $\hat{\sigma}_{\text{Optimized}}$
250	1,141	111
750	1,141	713
1,250	1,141	1,560
1,750	1,141	310
2,250	1,141	72
2,750	1,141	381
3,250	1,141	225

Table 1: The table shows product C's kernel-smoothed, weighted sigma estimate conditioned on a forecast level versus a more conventional root-sum-of-squared-errors approach. Note that the weighted technique does not guarantee a lower sigma estimate for every forward-looking forecast level, even for a product that has an historical overforecast bias.

which corresponds with the historical, increased error density to underforecast at that forecast level. In simpler terms, our modified sigma estimate is negatively correlated with bias.

Validating Kernel-Smoothed, Modified Sigma Estimates

Our approach to validating the kernel-smoothed, sigma-modified approach was threefold: it needed to replicate a theoretical underlying variance estimates for a constant relative variance, a constant unit variance, and a heterogeneous variance.

We first validated the kernel-smoothed, modified sigma technique for the sigma of a well-behaved (i.e., homogeneous, unbiased) forecast. We generated a series of forecasts and randomly determined the actuals around the forecast using a percent error represented by an $N(0, 5\%^2)$ distribution. Equation (7), a coefficient of variation (COV) that is the ratio of sigma at a given forecast and the forecast, measures the relative error for our case:

$$\text{COV} = \frac{\sigma_{|\text{Forecast}}}{\text{Forecast}}. \qquad (7)$$

With this approach, we were looking to see if the kernel-smoothing technique could consistently replicate the same 5 percent standard deviation across any forecast level. Figures 11(a) and 11(b) display the forecast error distribution that was randomly generated for the purposes of validating and the COV at each forecast level generated by Intel's kernel-smoothing technique.

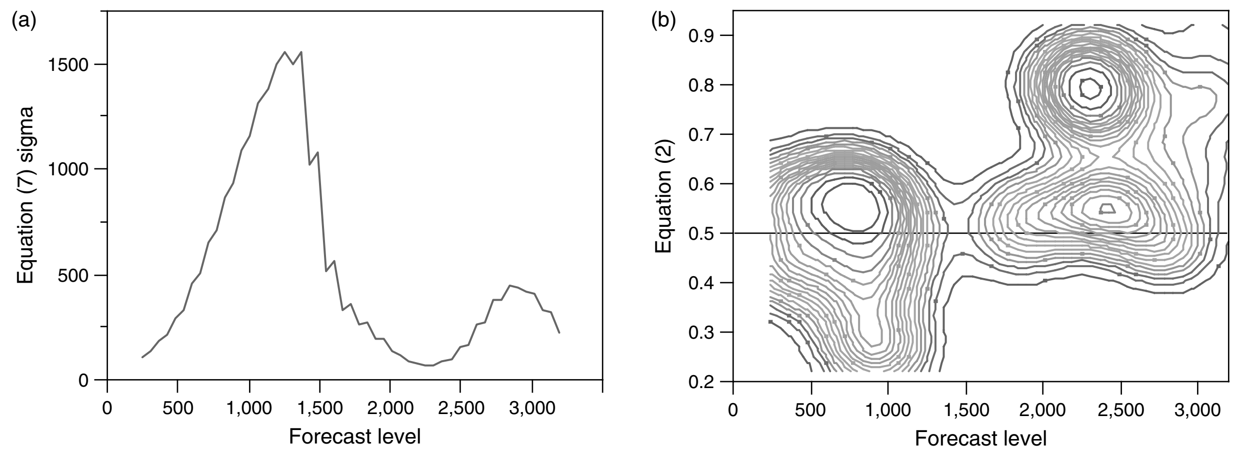

Figure 10: The graphs illustrate product C's new kernel-smoothed, weighted sigma estimate conditioned on a forecast level and the forecast-density error field. As expected, the new procedure estimates its largest value for sigma where (global max in panel a) the product has historically been underforecast most severely (i.e., density around 1,200 in panel b).

After relative homogeneity, we tested unit homogeneity. In a similar fashion, we considered a series of actuals that were normally distributed around a forecast by a constant unit variance, regardless of forecast value. Using the same arbitrarily chosen forecast range from the test for relative homogeneity, we generated the actuals around the forecasts with an $N(0, 25^2)$ error pattern. Figures 12(a) and 12(b) demonstrate both the resulting randomly generated forecast error pattern and sigma estimate from the kernel-smoothing technique. As desired, the kernel-smoothing approach replicated the underlying variance pattern.

With constant relative error and constant unit error validated, the final validation check was to see if the smoothing technique could replicate a heterogeneous error pattern, a particularly important aspect given its intent to address SKUs that fail the homogeneity test. Figure 13(a) represents a split-forecast error pattern in which forecasts from 100 to 350 have a relative error structure of $N(0, 5\%^2)$ and the relative error

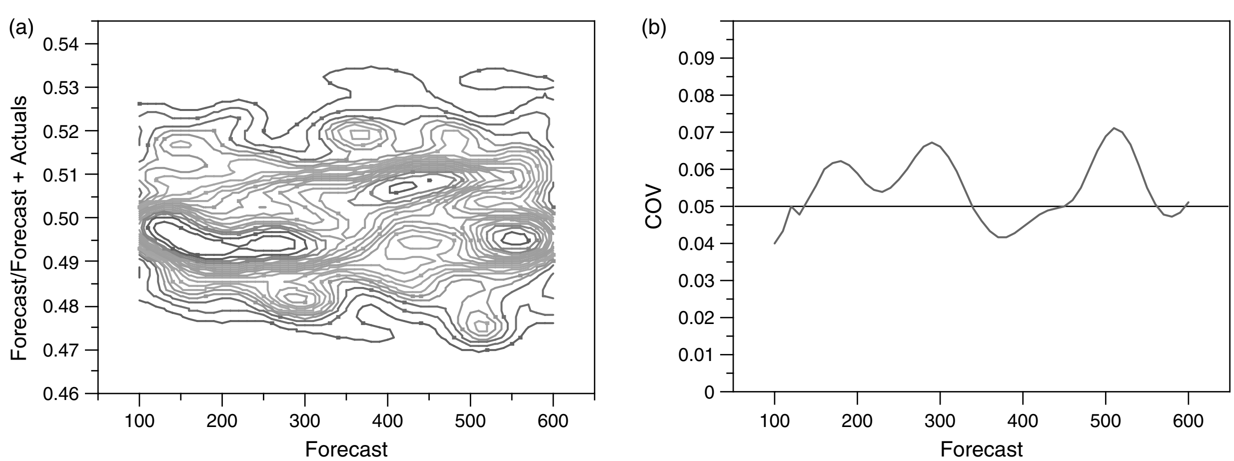

Figure 11: Panel a depicts the Equation (2) forecast error that was randomly generated from an $N(0, 5\%^2)$ distribution. Panel b demonstrates the kernel-smoothed sigma appropriately estimated the underlying 5 percent sigma across all forecast levels.

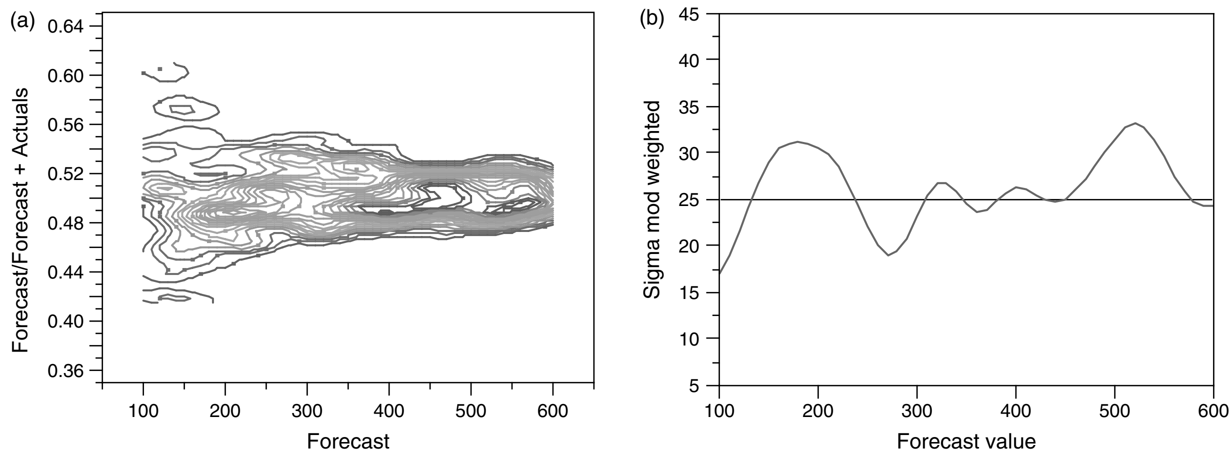

Figure 12: Panel a depicts the Equation (2) forecast error that was randomly generated from an $N(0, 25^2)$ distribution. Panel b demonstrates the kernel-smoothed sigma estimate, which appropriately estimated the underlying sigma (25) across all forecast levels.

for forecasts above 350 are structured as $N(0, 15\%^2)$. Figure 13(b) demonstrates how the kernel-smoothing technique tracks to the underlying error pattern.

In validating, we learned that a larger kernel standard deviation estimate than Bowman and Foster's (1993) recommendation can sometimes result in faster convergence to the true underlying sigma when bias is present, but is otherwise homogeneous. However, a smaller standard deviation estimate helps minimize the transition lag of sigma-modified estimates when variance is heterogeneous (e.g., in the range of 300 to 400, as in Figure 13(b)), but the smaller kernel-smoothing standard deviation creates higher variability in the sigma-modified estimate at any given forecast level. Optimal standard deviation estimate selection for the kernel-smoothing technique provides an area of further research for Intel.

Optimized Targets Based on New Modified SDFE

Since it implemented the modified sigma approach in early 2008, Intel has seen safety stock inventory-

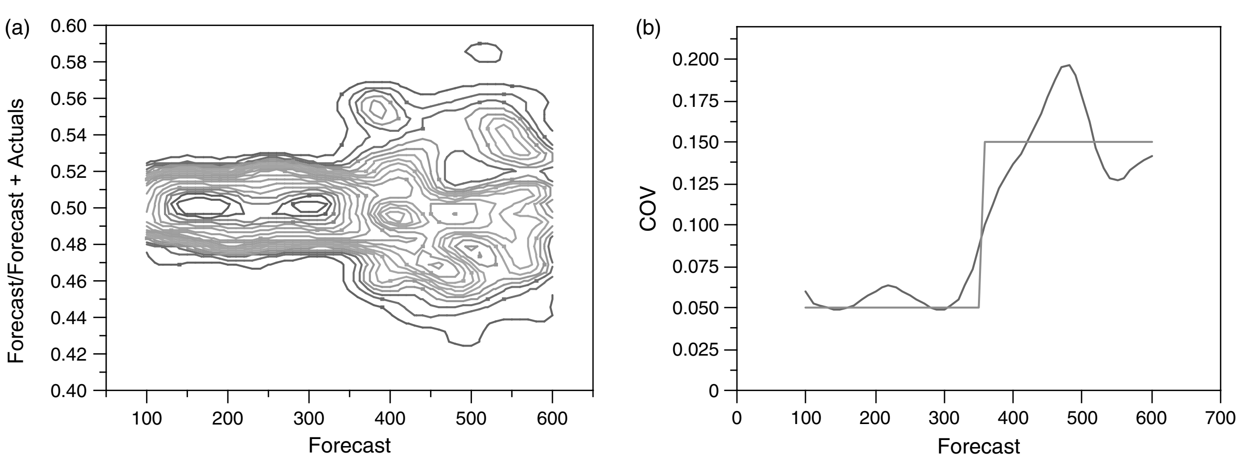

Figure 13: Panel a depicts the Equation (2) randomly generated heterogeneous forecast error. Panel b demonstrates the kernel-smoothed COV estimate, which appropriately estimated the underlying theoretical sigma (step function) across all forecast levels.

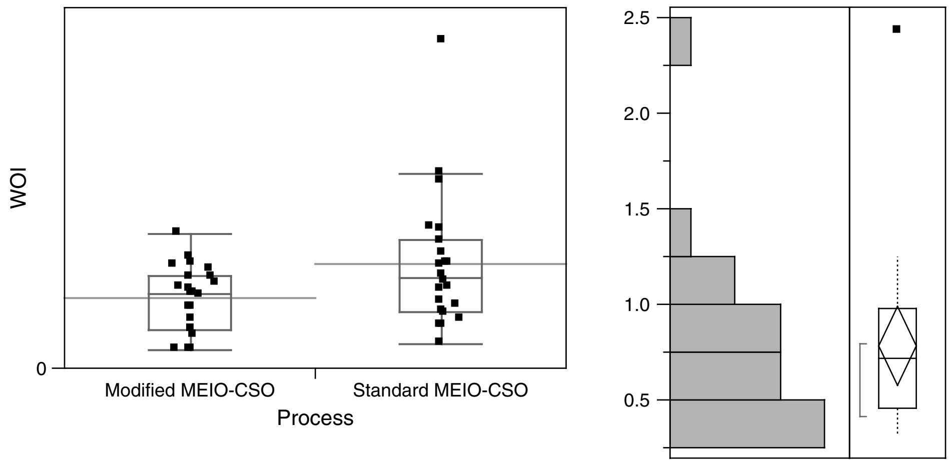

Figure 14: The graphs show representative snapshots of boxed CPU inventory targets comparing the modified sigma approach with a squared error standard deviation estimate. Panel a shows the overall reduction in both variance and average weeks of inventory (WOI) targets; Panel b displays the inventory target under sigma modified relative to the squared error standard deviation.

level reductions of approximately 15 percent without any reduction in customer service levels. Figure 14(a) demonstrates a snapshot of its inventory targets in the first half of 2008. Both average targets and targeting variability have been reduced. Figure 14(b) provides insight into inventory targeting from a paired perspective in which values are measured as the inventory target under the modified sigma approach standardized to the previous method of a squared error standard error. Approximately 70 percent of the SKUs had a significantly lower safety stock target than they would had the sigma not been modified; approximately 10 percent saw a significant increase in safety stock targets. Because of the propensity to overforecast, as we expected, the realized safety stock reductions (~15 percent) were less than the targets called for (~50 percent) because the forecast bias inherently carries with it extra "safety stock" that is realized in finished goods when the demand does not materialize.

Although initially developed only to address products failing the homogeneity test, we found the kernel-smoothing technique robust enough to replace the need for the initial Bartlett test (Bartlett and Kendall 1946) and RMSE estimate of sigma modified. Intel fully implemented the kernel-smoothing technique (with retirement of the homogeneity testing) in fall 2008.

Intel's boxed processor inventory targeting has employed MEIO optimization techniques for several years. Intel's experience has shown that appropriate demand characterization is a critical driver of the model's efficacy. Demand characterization also tends to be the variable that departs most frequently and severely from the conventional assumptions upon which most optimization models and software are built. The adaptations that Intel developed to address ill-behaving demand characterization have focused on correcting bias and heterogeneity impacts to optimized inventory target levels. In cases in which imperfections, such as bias or heterogeneity, cannot be removed from the raw forecast data, these approaches have proven effective in allowing a formal SIOP process to achieve desired inventory and service-level targets.

Appendix A

If the service level is predetermined, then Equations (4)–(6) can be replaced with a simple cumulative density problem and the original sigma-modified calculation in Manary and Willems (2008) in which the $\bar{\theta}$ solved for in Equation (A1),

$$\int_{\bar{\theta}}^{1} P_{\theta \mid \text{Forecast}}(\theta)\, d\theta = \text{Service Level Selection}, \qquad \text{(A1)}$$

is then used in Equation (3); the result is

$$\hat{\hat{\sigma}}^{\text{Modified}} = \text{Max}\left[\frac{((1-\bar{\theta})/\bar{\theta})-1}{t_{\beta,\,df}} * \mu,\ 0\right], \quad \text{(A2)}$$

where the probability density function is defined by the kernel-smoothed density field localized at the current forecast.

Appendix B

To test for homogeneity, we calculated an adaptation of Bartlett and Kendall's (1946) statistic using the sigma-modified calculation. In the adapted Bartlett's test, $(n-1)$ represents the historical observations to draw against, k reflects the number of service levels (SL) to test against, and sigma modified is calculated from Equation (3) for k service levels, as Equation (B1) shows:

$$T = \frac{(n-1)(k*\ln(\sum_{\text{SL}}(\hat{\sigma}^2_{\text{Modified}}/k)) - \sum_{\text{SL}}\ln(\hat{\sigma}^2_{\text{Modified}}))}{1+(\sum_{\text{SL}}(1/(n-1)) - 1/(k*(n-1)))/(3*(k-1))}. \quad \text{(B1)}$$

Acknowledgments

We thank the additional team members at Intel: Brian Wieland and Major Govindaraju for their role in managing decision support, Paul Bloomquist for his technical expertise in the programming environment, and Pat Mastrantonio for his advocacy at the management level. The success of this project would not have been possible without their contributions. M. P. Manary's current affiliation is Duke University's Fuqua School of Business, Durham, NC 27708.

References

Bartlett, M. S., D. G. Kendall. 1946. The statistical analysis of variance-heterogeneity and the logarithmic transformation. *J. Royal Statist. Soc. Supplement* **8**(1) 128–138.

Bowman, A., A. Azzalini. 1997. *Applied Smoothing Techniques for Data Analysis*. Oxford University Press, Oxford, UK.

Bowman, A., P. Foster. 1993. Density based exploration of bivariate data. *Statist. Comput.* **3**(4) 171–177.

Brown, M. B., A. B. Forsythe. 1974. Robust tests for the equality of variances. *J. Amer. Statist. Assoc.* **69**(346) 364–367.

Crum, C., G. E. Palmatier. 2003. *Demand Management Best Practices: Process, Principles and Collaboration*. J. Ross Publishing, Boca Raton, FL.

Franses, P. H. 1998. *Time Series Models for Business and Economic Forecasting*. Cambridge University Press, Cambridge, UK.

Levene, H. 1960. Robust tests for the equality of variances. I. Olkin, ed. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press, Palo Alto, CA, 278–292.

Manary, M. P., S. P. Willems. 2008. Setting safety-stock targets at Intel in the presence of forecast bias. *Interfaces* **38**(2) 112–122.

Montgomery, D. C., L. A. Johnson, J. S. Gardiner. 1990. *Forecasting and Time Series Analysis*, 2nd ed. McGraw-Hill Companies, New York.

O'Brien, R. G. 1979. A general ANOVA method for robust tests of additive models for variances. *J. Amer. Statist. Assoc.* **74**(368) 877–880.