



9/22/2026

DELIVERED VIA ELECTRONIC MAIL : grain@otc-cta.gc.ca

Secretariat
Canadian Transportation Agency
60 Laval Street, Unit 1-117
Gatineau, Quebec
Canada, J8X 3G9

RE: CN's REPLY - VOLUME-RELATED COMPOSITE PRICE INDEX (VRCPI) FORECASTING MODEL TESTING RESULTS

Dear Sir/Madam,

Canadian National Railway Company ("CN") submits the following comments in response to the Canadian Transportation Agency's (the "**Agency**") consultation on the testing results of the Volume-Related Composite Price Index ("**VRCPI**") forecasting models.

More specifically, this submission addresses the six potential methods identified by the Agency for forecasting the labour, fuel, and material price adjustments incorporated into the VRCPI calculation. Drawing on input from a range of stakeholders, including CN, the Agency has evaluated these methods, and this submission sets out CN's view on the most appropriate approaches¹.

CN appreciates the opportunity to comment on the testing results of the forecasting models and to assist the Agency in selecting the most appropriate approach for use in future VRCPI determinations, beginning next year.

Summary of CN's Position

As detailed in CN's previous submissions, CN's preferred approach remains to eliminate the need for forecasting altogether by using actual price changes from the most recently completed calendar year as the basis for calculating cost inflation for the upcoming crop year. Recognizing, however, that the Agency has now narrowed its options to six forecasting methods for each of the labour, fuel, and material price adjustments, CN considers the multi-year moving average growth methods (Methods 3 or 4) to be an acceptable compromise and a notable improvement over the current models, which are often complex and lack accuracy.

Specifically, **CN recommends that the Agency adopt the three-year moving average growth (Method 3) for all the three cost components – labour, fuel and material**. Adopting a shorter timeframe, such as three years instead of five, would allow railways and shippers to more efficiently recoup discrepancies

¹ Canadian Transportation Agency, VRCPI Forecast Model Consultations - Model Testing Results, <https://otc-cta.gc.ca/eng/extended-consultation-vrcpi-forecast-models/vrcpi-forecasting-test-results>.

between forecasted and actual inflation, enhance responsiveness to changing economic conditions, and better reflect recent cost trends.

Labour and Material Components

The average of CN's and CPKC's railway-specific price indices based on a three-year historical average produced the smallest mean error and cumulative error in the Agency's "back-casting" of labour and material price forecasts. This was determined by comparing the Agency's generated forecasts for the last ten crop years with the corresponding actual price changes.

The mean error measures the degree of systematic over- or under-forecasting, with values closest to zero indicating the least bias and therefore the most accurate forecasting approach. Similarly, cumulative error measures the extent to which forecasted price adjustments align with actual cost changes over time. A cumulative error closest to zero suggests that both railways and shippers are more likely to recover actual costs over the long term. In the Agency's testing, covering the ten crop years from 2015-2016 through 2024-2025, the three-year average of the CN and CPKC indices performed best on both measures for the labour and material components.

Fuel Component

- **Sample Boundary Effects**

While the three- and five-year averaging of historical railway price indices tested relatively well for the labour and material components, the Agency's testing indicated higher error levels for the fuel index. This result is affected by a sample boundary effect at both the beginning and end of the sample period.

At the start of the sample, crude oil prices declined by approximately 70% between mid-2014 and early 2016, largely due to a significant global supply surplus associated with the rapid growth of U.S. shale oil production. This extraordinary market event resulted in a substantial over-forecast of fuel price growth in the 2015-2016 VRCPI with the multi-year moving average method.

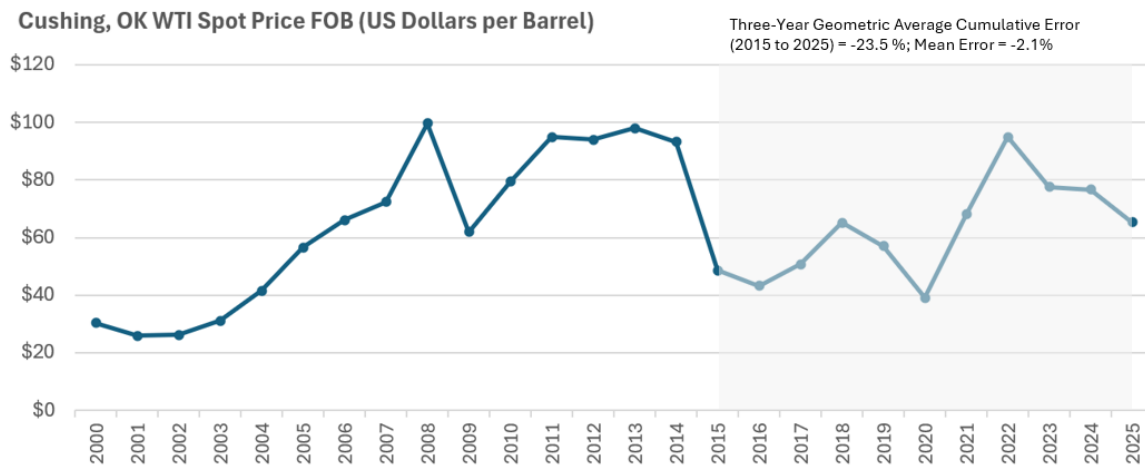
A similar boundary effect occurred at the end of the sample period. By 2025, crude oil prices had declined by approximately 40% from their 2022 peak levels, as the severe and persistent supply shortages anticipated following Russia's invasion of Ukraine did not fully materialize. This sharp correction similarly resulted in a substantial over-forecast in the 2024-2025 VRCPI with the multi-year moving average method.

Because these two exceptional price declines occur at opposite ends of the Agency's ten-year sample period, they disproportionately influence the performance metrics of forecasting approaches based on multi-year moving average growth rates (Methods 3 and 4). As a result, the cumulative and mean errors reported for these methods are artificially inflated and may not be representative of their forecasting performance under more typical market conditions. The Agency should therefore exercise caution in drawing conclusions from a relatively short evaluation period that begins and ends with unusually large fuel price corrections.

• Arithmetic vs. Geometric Averaging

Another key factor that inflated the reported forecast errors for the fuel component under the multi-year moving average methods is the Agency's reliance on arithmetic averaging rather than geometric averaging. Because labour, fuel, and material prices evolve through compounded percentage changes over time, geometric averages will provide a more accurate measure of long-run growth². When growth rates compound through time, the arithmetic average systematically overstates the constant growth factor equivalent to the historical path whenever year-to-year growth varies³. In addition, arithmetic averages can be disproportionately influenced by extreme observations and will overstate both underlying growth trends and forecast errors when large price increases and decreases occur within the evaluation period. This is particularly important for fuel prices, which are subject to significantly greater volatility than the labour and material components.

Figure 1 illustrates the sample boundary effect and presents the cumulative and mean errors calculated using the geometric average of publicly available crude oil prices for two sample periods: (a) 2015-2025 and (b) 2013-2025. When both the starting and ending observations in the sample period (2015 and 2025) coincide with troughs in a highly volatile price cycle, the boundary effect produces a mean error of -2.1%. Extending the sample period to include the price peak in 2013 reduces the mean error to 0.0%, demonstrating the sensitivity of the results to the selection of sample endpoints and the sample periods.

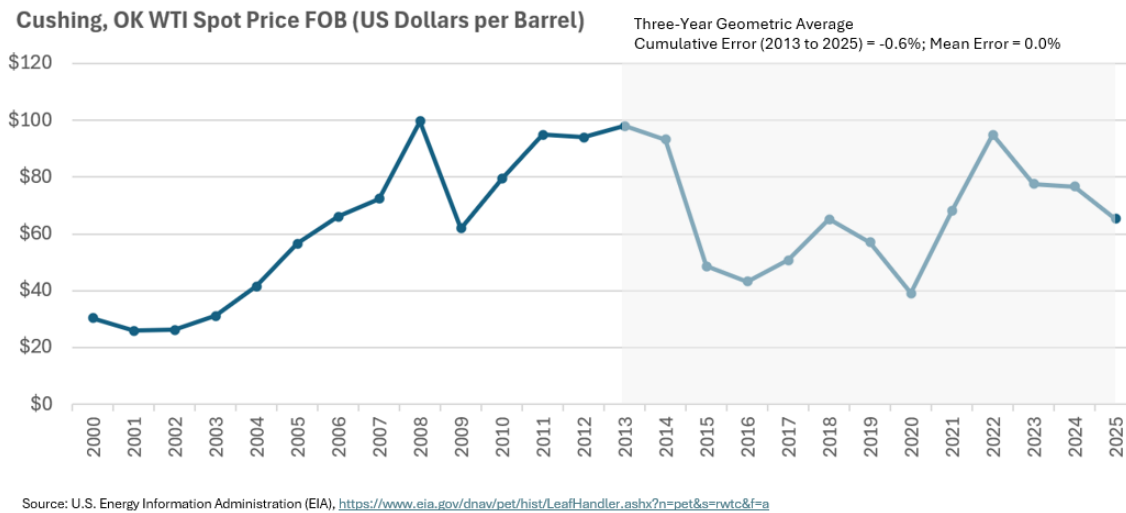


(a) 2015-2025

² Using the arithmetic average will overstate future values when yearly changes vary. For example, a -10% drop followed by a 20% increase turns \$1.00 into $(1 - 10\%) \times (1 + 20\%) = \1.08 . Applying twice the $(-10\% + 20\%)/2 = 5\%$ arithmetic average gives \$1.10, while applying the $(\sqrt{0.9 \times 1.2} = 1.04)$ 4% geometric average correctly gives \$1.08. The two averages are equal only when the yearly changes are the same.

³ This is a well-known theorem called the Arithmetic Mean-Geometric Mean (AM-GM) Inequality.

https://en.wikipedia.org/wiki/AM-GM_inequality



(b) 2013-2025

Figure 1: Historical Crude Oil Prices

- **WTI-Based Regression Methods (Methods 1 and 2)**

In Fuel Methods 1 and 2, the Agency forecasts fuel price changes using a simple linear regression model that estimates the relationship between the railway fuel price index and global crude oil prices, as represented by West Texas Intermediate ("WTI") crude oil at Cushing, Oklahoma. The Agency's regression model linking WTI crude oil prices to CN's purchased fuel prices exhibits a strong statistical fit, with an R-squared value exceeding 90%. However, the model relies on a single explanatory variable, WTI crude oil prices, which are themselves forecasts subject to both inherent forecasting uncertainty and unforeseen external factors, including geopolitical events, supply disruptions, and broader macroeconomic shocks. Regardless of the strength of the regression relationship, the accuracy of the resulting fuel price forecast is ultimately constrained by the accuracy of the underlying WTI price forecast. By definition, the accuracy of third-party WTI forecasts can only be evaluated retrospectively, as forecasted prices may differ materially from actual market outcomes.

Figure 2 compares the WTI price forecasts used by the Agency with the corresponding actual WTI prices. Although the Agency's back-testing results show that Methods 1 and 2 produce relatively small cumulative and mean deviations over the selected ten-year sample period, their performance is necessarily conditional on that historical period. Future application of either method would expose fuel VRCPI forecasts to errors arising from future WTI price forecast inaccuracies, particularly during periods affected by unexpected geopolitical developments or other macroeconomic shocks.

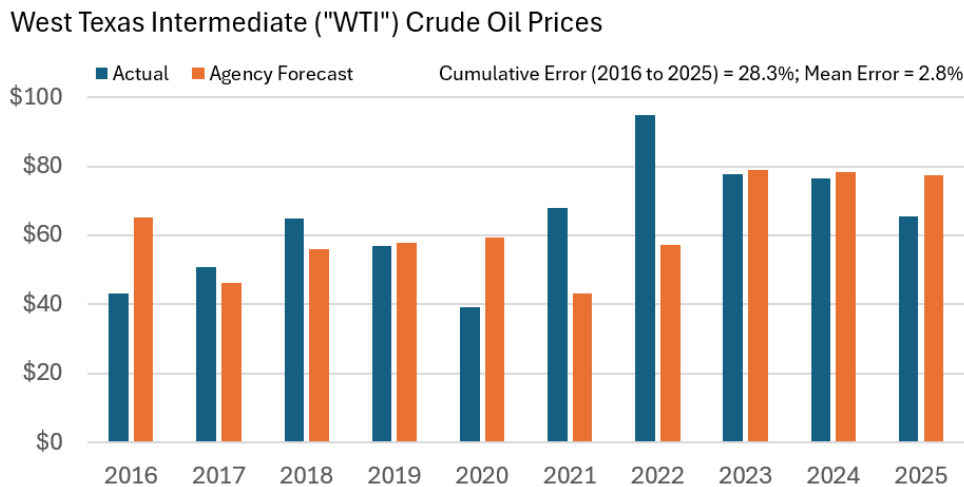


Figure 2: Actual vs Agency Forecast of WTI Crude Oil Prices

- **Advantage of Method 3 for Fuel**

In contrast, Method 3 relies exclusively on observed fuel price changes and does not require forecasting an external commodity price. Because fuel prices evolve multiplicatively over time, the geometric mean is the mathematically appropriate measure of long-run growth. As a result, a three-year moving geometric average of actual fuel price growth will, over a sufficiently long period, recover the full underlying fuel cost increase on average, as periods of over-recovery and under-recovery offset one another. This approach avoids dependence on inherently uncertain third-party commodity price forecasts while providing a robust mechanism for achieving long-run cost recovery to both railways and shippers.

Monte Carlo Simulation (Appendix A)

Appendix A⁴ provides a more detailed statistical assessment demonstrating that the three-year moving average growth approach (Method 3) is a robust forecasting methodology for the labour, fuel, and material cost components, better than the status quo. Rather than relying solely on the Agency's

⁴ CN has also provided a supporting Excel workpaper to facilitate the Agency's review and replication of the analyses presented in Appendix A. The workpaper contains confidential information and is clearly identified by the word "CONFIDENTIAL" at the beginning of the filename. The workpaper and all its contents should be treated as strictly confidential and must not be disclosed, published, or otherwise made available to the public.

historical back-testing results over a ten-year sample period, CN conducted a Monte Carlo simulation to evaluate the long-run performance of the multi-year moving average forecasting methods under consideration.

A Monte Carlo simulation is a statistical technique that evaluates the impact of uncertainty by repeatedly generating outcomes based on the historical characteristics of the underlying data. Using the average growth rates and historical variability observed for each cost component, CN simulated 1,000 alternative price paths and compared the resulting cumulative and mean forecasting deviations produced by each forecasting method.

The results as summarized in Figure 3 below demonstrate that the three-year moving average growth approach (Method 3) substantially reduced the magnitude of expected bias for the labour, fuel and material indices, is robust to short-term price volatility and provides a reliable estimate of underlying cost trends.

Status Quo Method				Average Growth Rate Method							
				Method 3 (3-yr average)				Method 4 (5-yr average)			
Bias											
Historical	Labour	Fuel	Material		Labour	Fuel	Material		Labour	Fuel	Material
1 record	-0.34%	-4.91%	-7.35%	1 record	0.59%	-1.88%	-2.89%	1 record	0.92%	-4.60%	-3.36%
				1,000 rec.	0.04%	0.02%	0.07%	1,000 rec.	0.01%	-1.29%	-0.01%

Figure 3: Monte Carlo Simulation Results

Adoption of Different Models for Each Prescribed Railway

CN understands that the Agency may also be exploring the possibility of adopting different forecasting models for each prescribed railway company. CN has no objection to such an approach. Given that the operating characteristics, cost structures, and fuel consumption profiles of CN and CPKC may differ, CN recognizes that permitting the use of different models for each railway could allow the Agency to select the method that best reflects each carrier's actual cost dynamics.

Conclusion

For the reasons set out above, CN respectfully recommends that the Agency adopt **the three-year moving geometric average growth (Method 3) for all the three cost components – labour, fuel and material**. Adopting Method 3 would enhance the transparency and credibility of the VRCPI by grounding its forecasts in actual railway costs. It provides mathematical certainty that both railways and shippers will recover actual costs over time, regardless of future price shocks or unforeseen market disruptions. By mitigating the influence of temporary price spikes and abrupt cost swings, Method 3 produces more stable and reliable forecasts while maintaining alignment with underlying cost realities. It would also allow all stakeholders, including shippers, to more effectively anticipate and understand future cost adjustments, as the underlying unit cost trends could be tracked using publicly disclosed financial information reported by the railways.



Financial Planning

935 de La Gauchetière Street West
Montreal, Quebec H3B 2M9

Thank you for your attention to these comments. Again, CN appreciates the opportunity to participate in this consultation and remains available to engage in any follow-up discussions or provide further clarification as needed.

Sincerely,

A handwritten signature in black ink, appearing to read 'Rapik Saat'.

Rapik Saat, Ph.D.
Senior Manager, Regulatory Costing | Financial Planning
Canadian National Railway Company

Appendix A — Technical Assessment of VRCPI Forecasting Methodology

1. Purpose and Overview

This Appendix provides a technical assessment of the data and methodology used to evaluate alternative forecasting approaches for the labour, fuel, and material components of the Volume-Related Composite Price Index (“VRCPI”). Its purpose is to explain why CN’s assessment is based on the latest known version of the relevant historical indices, and why Monte Carlo simulation provides a more reliable basis for comparing forecasting methods than relying on a single historical back-cast sample.

The analysis begins from the Agency’s historical index revisions, then constructs one consistent historical dataset for each index. It then uses those observed historical characteristics to generate alternative price paths and evaluate forecast performance across a broad range of plausible future outcomes. This approach is consistent with the purpose of forecast evaluation: a methodology should be assessed by how well it performs on average across possible outcomes, not merely by how well it happens to fit one realized historical path.

The methodology and results illustrated below are supported by references to the relevant worksheets and cell references in the file “CONFIDENTIAL - VRCPI Calendar Year Price Indices - Monte-Carlo Simulations - CN v4.xlsx”

2. Need for a Single Stable Version of the Historical Indices

The first step in evaluating any forecasting methodology is to establish the actual value against which the forecast should be compared. This is not a trivial issue for the VRCPI indices because the Agency’s determinations are revised after the index is first issued. The evidence in the *Calendar-Year Indices* and *Crop-Year Indices* shows that the Agency’s VRCPI-related index values are not necessarily stable after one or two subsequent determinations. In some cases, the final value continues to change for several years.

For example, the initial labour index for the 2019-2020 crop-year was issued in April 2019. The same index was subsequently updated in April of each year from 2020 to 2024. This example shows that the value against which a forecast is evaluated may continue to change materially after the first and even the second subsequent determination. In that case, the index did not reach a stable value until several years after it was first issued. In the Agency’s evaluation, however, the forecast methods are compared against the two-year-out value, even though the later revisions show that this was not the most accurate final value.

It would therefore be inappropriate to evaluate a forecasting method by comparing its forecast to an interim value that is itself still subject to revisions. Forecast evaluation requires one version of the truth: the forecast must be compared against the latest known stable estimate of the index, rather

than against an earlier vintage simply because that was the value available after a fixed number of years. This issue is well recognized in forecast evaluation where official statistical series are revised over time; using revised or final data can materially affect measured forecast performance, and vintage-consistent data are needed to understand what was known at different release dates and how later revisions changed the target value. Evaluating against interim vintages introduces arbitrary measurement error.

For this Appendix, CN therefore uses the latest known version of each historical index as the benchmark against which forecast methods are assessed. For crop years 2007-2008 to 2013-2014, the benchmark values are taken from the tab *2017-18 R-2017-37 CN Table 1*. For crop years 2014-2015 to 2026-2027, the benchmark values are taken from the tab *2026-27 R-2026-78 CN Table 1*. Together, these two sources provide a single consistent historical series for each of the labour, fuel, and material indices. The analysis excludes the final three index years, 2024-2025 to 2026-2027, because those values remain interim and may be revised in future determinations. It also excludes 2007-2008 to 2012-2013 because Method 4 requires the average of the previous five growth rates. Therefore, the first 6 years from 2007-2008 to 2012-2013 are not available for Method 4 and so were also excluded from the analysis to compare all methods on the same historical record.

The resulting one-version historical dataset is shown in the *One-Version* tab where the table in cells A2:V5 presents 21 calendar years of index levels for labour, fuel, and material that lead to 20 crop-year indices from 2007-2008 to 2026-2027.

3. Historical Status-Quo Forecast Results

Before evaluating alternative averaging methods, it is useful to measure how the current status-quo approach performs on the historical record. In this section, the “forecast” is the index value first issued in the Agency’s determination, while the “true value” is the latest known estimate of that same index after subsequent revisions. This comparison measures how well the first determination anticipated the value that was ultimately reflected in the latest available historical series.

The analysis is presented on the *One-Version (Status-Quo)* tab. Because this is the first point in the report where the performance metrics are introduced, each metric is briefly defined here. Bias is the average forecast error percentage and indicates whether the first determination tended, on average, to be above or below the latest estimate.

The historical status-quo results show that the first determination was not neutral across all index components. Labour is close to neutral on average, with a bias of -0.34%. Fuel shows materially larger errors, with a bias of -4.91%. Material also shows a negative average bias, at -7.35%. These results provide the baseline against which the alternative three-year and five-year average growth rate forecasting methods should be assessed.

Table A-1 — Status-Quo Historical Forecast Results			
Historical	Labour	Fuel	Material
1 record	-0.34%	-4.91%	-7.35%

Of particular concern to CN are the sizeable economic losses, especially for fuel and material, where a negative bias can translate into a significant economic impact. Based on CN's VRCPI weights—32% for labour, 16% for fuel, 35% for material, and 17% for other—and a typical current crop-year MRE of approximately \$1.5 billion, the economic loss due to the above observed biases is approximately \$1.6 million per year for labour, \$11.8 million per year for fuel and \$38.6 million per year for material. Although no forecasting method can eliminate loss entirely (neither for the railway in case of under-forecast nor for the shippers in case of over-forecast), the selected method should minimize the expected economic loss as much as reasonably possible.

Of equal concern to CN is the frequency distribution of forecast biases, even though the sample is limited to only eleven crop years, as this is indicative of bias risk. As shown in Table A-2, while 91% of labour forecasts fall within $\pm 5\%$, only 27% of fuel forecasts and 55% of material forecasts fall within this range. This means the risk of a forecast falling outside the $\pm 5\%$ range is 73% for fuel, and 45% for material —levels CN considers unacceptably high.

Table A-2 Status-Quo Method Bias Risk			
Within $\pm$	Labour	Fuel	Material
1%	27%	0%	18%
2%	36%	0%	27%
3%	55%	0%	36%
4%	64%	9%	55%
5%	91%	27%	55%

4. Construction of the Historical Price Paths

In the *One-Version* tab, the table in cells A2:V5 presents 21 years of index levels for labour, fuel, and material from 2007 to 2027. The table in cells A8:V10 converts those levels into year-over-year growth factors. These growth factors are the appropriate basis for the simulation because the indices evolve multiplicatively over time: each year's level is obtained by multiplying the prior year's level by the annual growth factor.

Because each annual growth factor is calculated as the ratio of two positive index levels, it must itself be positive and cannot be negative. Its theoretical lower bound is zero, whereas it has no finite upper bound. A normal distribution is therefore unsuitable because it assigns positive probability to negative values and is symmetric and unbounded in both directions. A log-normal distribution is more consistent with these properties: it is defined only over positive values, is bounded below by zero, has no finite upper bound, and can exhibit a long right-hand tail. Such right-skewed, positive-valued distributions are commonly used to model multiplicative quantities and ratios in finance, engineering, economics, and the construction and analysis of price and cost indices. This does not establish that every growth-factor series is exactly log-normal; rather, it provides a theoretically coherent and widely used modelling assumption that should also be assessed against the observed data and could be useful in modeling the growth factors at hand.

A log-normal distribution being a natural model for positive multiplicative growth factors, the three indices historical growth factors are summarized using the mean and standard deviation of their natural logarithms, as shown in cells X8:Y10. These mean and standard deviation describe the log-normal distribution of the growth factors.

The mean of the natural logarithms of growth factors g_i

$$\mu = \frac{1}{N} \sum_{i=1}^N \ln(g_i) = \frac{1}{N} \ln(g_1 \times g_2 \times \dots \times g_N)$$

$$\mu = \ln((g_1 \times g_2 \times \dots \times g_N)^{\frac{1}{N}})$$

$$\mu = \ln(\text{Geomean}(g_i))$$

The log-normal μ is thus the natural logarithm of the geometric mean of the growth factors - and the geometric mean is $\exp(\mu)$.

The use of logarithms is consistent with compound growth analysis because taking logarithms converts multiplicative changes over time into additive changes, allowing the central tendency and dispersion of growth factors to be measured on the scale on which compounding occurs.¹

The charts beginning at X27, X73, and X119, and shown below in Figures 1, 2, and 3, illustrate the distribution for the labour, fuel, and material growth factors.

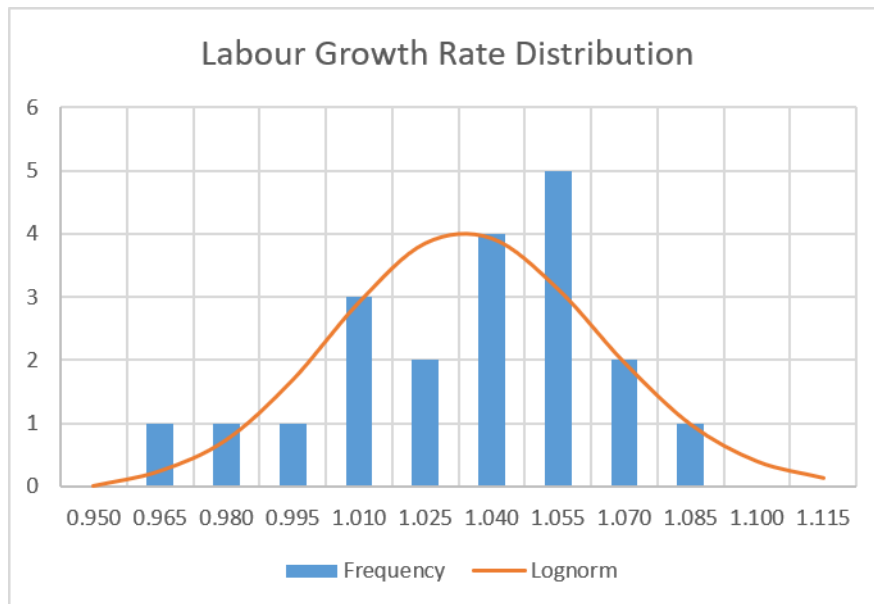


Figure 1 Labour Growth Rate Distribution

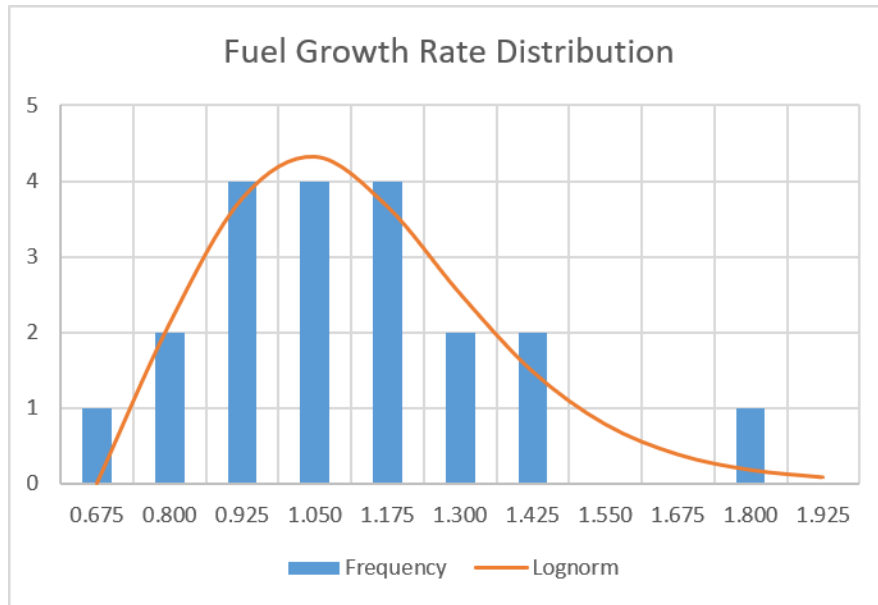


Figure 2 Fuel Growth Rate Distribution

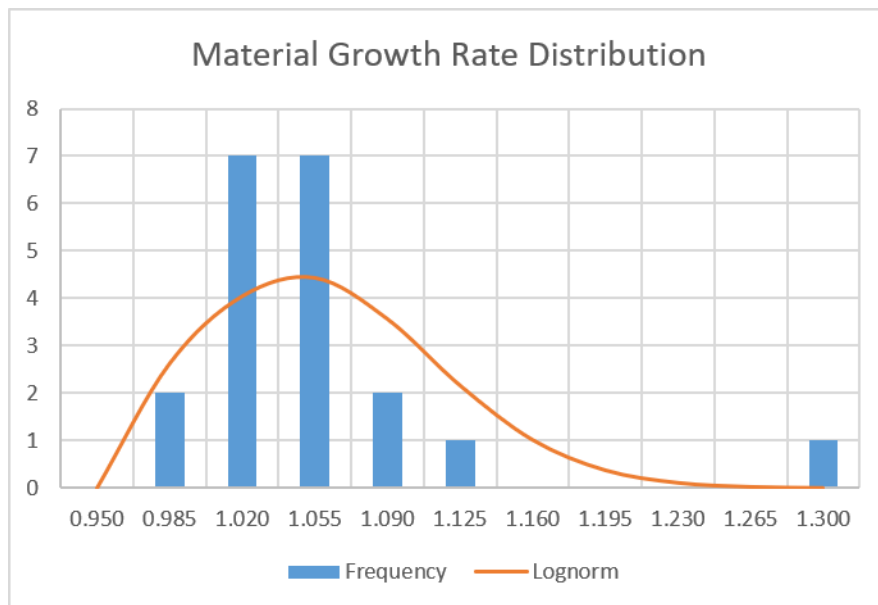


Figure 3 Material Growth Rate Distribution

In each chart, the blue bars represent the empirical distribution of the observed annual growth factors, and the orange curve represents the corresponding log-normal distribution calibrated to the same log-mean and log-standard deviation. These charts are not intended to suggest that the future will precisely replicate the past. Rather, they show that the simulation is grounded in the observed long-run characteristics of each historical index: its average compound growth and its historical dispersion.

5. Why Monte Carlo Simulation Is Used

The last 20 years of VRCPI index movements represent only one realized price path. A forecasting method that performs well on that single path may not only do so because it is genuinely robust, but it may also do so because the realized historical sequence happened to favour that method. Selecting a forecasting methodology solely on the basis of one historical path therefore risks mistaking sample-specific performance for expected performance.

Back-casting tests a forecasting method only against the historical path that actually occurred. There is a very low probability that the next 15 years will reproduce the precise sequence of annual changes observed over the previous 15 years. It is more plausible to assume that the future path may retain broadly similar long-run characteristics - namely, a comparable average compound growth rate together with stochastic year-to-year variation drawn from a probability distribution resembling that observed historically. Monte Carlo simulation is designed to test performance under that broader assumption by generating many alternative sequences consistent with the historical trend and volatility, rather than treating the past sequence itself as the future.

Monte Carlo simulation addresses this problem by repeatedly generating alternative outcomes from a defined probability distribution and then evaluating the model across the resulting set of outcomes. It is widely used for risk and uncertainty analysis because it replaces a single deterministic outcome with a distribution of possible outcomes, allowing the analyst to assess average performance, dispersion, and tail outcomes over many trials.² The result is not “random” in the colloquial sense of being arbitrary. The individual simulated paths are randomly drawn, but the simulation design is disciplined by observed historical growth and volatility, and the conclusions are based on the aggregate performance across many simulated paths.

In this context, the relevant question is not which forecasting method happened to perform best on the one realized historical path, but which method performs best on average across many plausible paths with the same observed long-run characteristics. “All possible price paths” is not computationally necessary for this purpose. A sample of 1,000 simulated paths provides a substantial improvement over judging the methods on only one realized history, while remaining simple to audit and replicate in the supporting Excel worksheets.

6. Simulation Design

Staying on the *One-Version* tab, the tables *LPI_Rand*, *FPI_Rand*, and *MPI_Rand* each contain 1,000 alternative simulated price paths for the labour, fuel, and material indices, respectively. Each path begins from the historical value in 2007 and then progresses forward year by year using a randomly selected annual growth factor drawn from the log-normal distribution calibrated to the historical log-mean and log-standard deviation for that index.

The simulation formula is:

$$P(t) = P(t-1) \times \text{LOGNORM.INV}(\text{RAND()}, \text{log-mean}, \text{log-std.dev})$$

In this expression, $P(t-1)$ is the prior year’s price index, and the log-normal term is the simulated annual growth factor. The use of *RAND()* causes Excel to draw a percentile from zero to one, which is

then converted into a growth factor using the inverse log-normal distribution. Each simulated path therefore represents a plausible sequence of annual index changes with the same compound-growth characteristics and dispersion as the observed historical series.

The first row of each simulation table contains the historical observed index path. That historical path should be understood as one realized outcome among many possible outcomes generated by the same type of price process. If *RAND()* is replaced by 0.5, the simulation selects the median of the log-normal distribution in each year. This produces the smooth median path, equivalent to steady growth at the geometric mean growth rate, which starts from the historical base value, ends at the exact final historical value, reflecting the central compound-growth tendency of the observed data.¹

The Monte Carlo simulation therefore measures the random walk around this median path, with dispersion calibrated to the historical volatility of each index. The charts beginning at AF9, AF55, and AF101 illustrate this concept for labour, fuel, and material, respectively, by showing three illustrative simulated paths, the actual historical index path, and the median path based on the geometric mean growth rate. Users can generate a different random sample of three paths by pressing **F9** or selecting *Formulas > Calculate Now*. One example for each chart is shown below in Figures 4, 5, and 6. The dots represent individual simulated paths, the blue line represents the observed historical values, and the orange line represents the median path. This visual comparison demonstrates that the simulation does not replace the historical data; it uses the historical data as the anchor for generating a broader set of plausible outcomes around it.

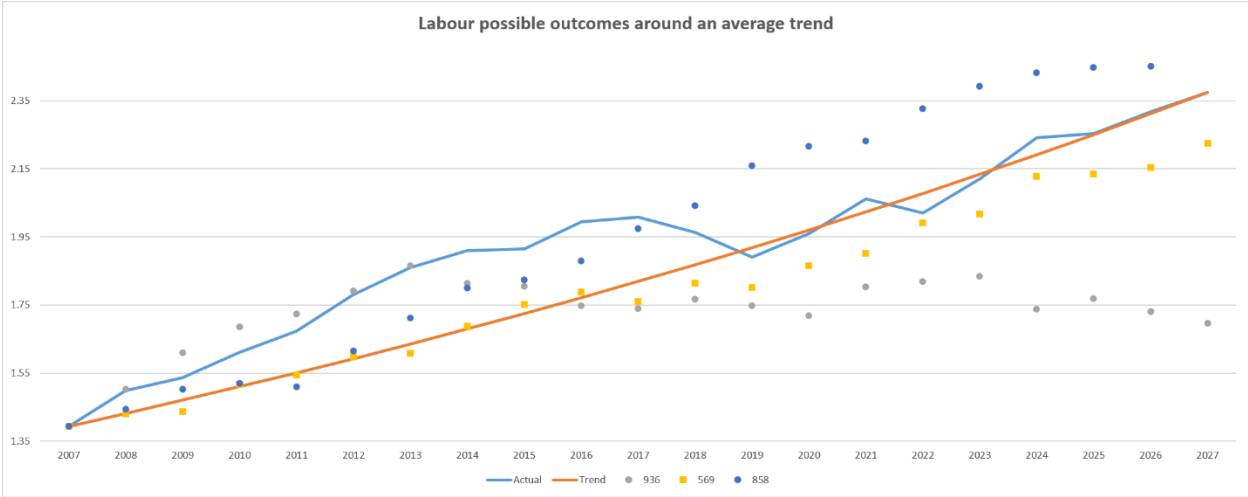


Figure 4 Labour possible outcomes around an average trend

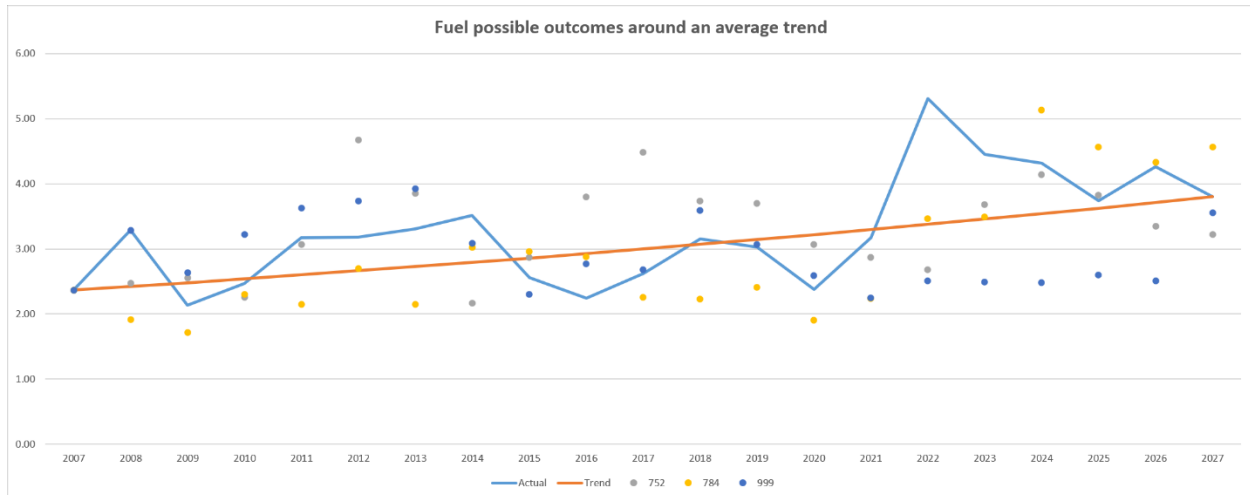


Figure 5 Fuel possible outcomes around an average trend

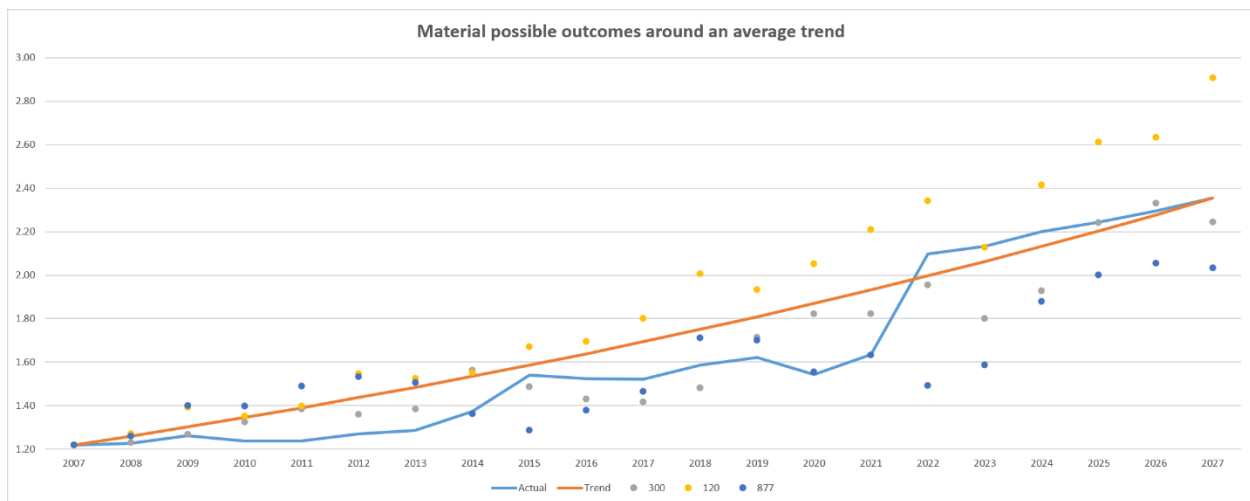


Figure 6 Material possible outcomes around an average trend

7. Forecasting Based on Average Historical Growth Rates

The workbook then evaluates forecasting methods that use recent average growth rates as the basis for projecting the next year's index value. The tab *One-Version (M.3)* performs this analysis using a three-year average growth rate, while the tab *One-Version (M.4)* performs the same analysis using a five-year average growth rate. Both tabs are structured to evaluate the method first against the actual latest version of the historical indices and then across 1,000 Monte Carlo scenarios drawn from the simulated labour, fuel, and material price paths.

7.1 Results from the Three-Year Method (Method 3)

In *One-Version (M.3)*, the table in cells A2:V40 performs the analysis using the actual latest version of the indices. The table in cells Y2:AS40 performs the same analysis for one selected row drawn from the possible outcomes in the *LPI_Rand*, *FPI_Rand*, and *MPI_Rand* simulation tables described above. Finally, Excel's *Data > What-If Analysis > Data Table* functionality, shown in cells AQ50:AS1049, runs

the same forecasting analysis across all 1,000 simulated rows and gathers the performance metrics needed to assess the robustness of the forecasting method across a wide range of possible price paths, rather than only the historical path that happened to occur.

The dynamic table in Y2:AS40 is useful for explaining the mechanics because it shows how the method performs for any selected simulated path. The selected rows from the random draw tables are controlled by cells X3:X5, which depend on the row index in cell AP48. Cell AP48 is in turn varied by the *What-If Analysis > Data Table*. When the index is set to 1, cells Y3:AS5 contain the historical observations. When it is set to another value, N, those same cells contain the values from the Nth row of the random draw tables. The static table in A2:V40 works in the same manner, except that it remains fixed to the historical version of the indices.

Within the dynamic table, cells Z8:AP10 calculate the year-over-year growth rates. Cells AC13:AQ15 then calculate the average three-year growth rate as the geometric mean of the prior three annual growth rates. Cells AC18:AQ20 calculate the forecast for the first year of the VRCPI by multiplying the prior year's index value by that three-year average growth rate. Cells AD23:AR25 then calculate the forecast for the second year of the VRCPI by multiplying the first year forecast by the same three-year average growth rate. Cells AC28:AQ30 combine these two forecasts into the crop-year VRCPI forecast, and cells AC33:AO35 combine the calendar-year actual index values to obtain the crop-year actual VRCPI. Finally, cells AE38:AO40 calculate the forecast error ratio by comparing the forecasted value with the actual value for the same crop-year.

A word here is necessary to explain the calibration factor used in the VRCPI forecasting. As explained above, historical growth rates exhibit a right-skewed (lognormal) distribution, where high outcomes (from μ to ∞) weigh more on the arithmetic mean even when they do not occur more frequently than low outcomes (bound between 0 and μ). This statistical asymmetry creates a small upward bias in measured mean forecast errors, even when forecasts are unbiased (equal frequency and probability of highs and lows). The calibration factor removes this mean bias using the historical variability of the series, ensuring that forecast mean results are centered around actual outcomes over time.

Mathematically, the arithmetic mean of a lognormal distribution is not $\exp(\mu)$ like its median but rather $\exp\left(\mu + \frac{\sigma^2}{2}\right)$.³ Calculating the forecast error percentage involves both the forecast (F) and the actual growth factor (G). Thus, the expected value of the error (R) is

$$E[R] = E\left[\frac{F}{G}\right] = E[F] \times E\left[\frac{1}{G}\right]$$

Assuming that both have the same distribution

$$E[F] = e^{\mu + \frac{\sigma^2}{2}} \quad \text{and} \quad E\left[\frac{1}{G}\right] = e^{-\mu + \frac{\sigma^2}{2}}$$

$$E[R] = e^{\sigma^2}$$

Because of the lognormal distribution properties, the forecast error is therefore overstated by a factor of e^{σ^2} . The calibration factors in cells AS28:AS30 are used to reduce the forecasts by these values derived from the historical observations ('One-Version'!Y8:Y10). The calibration factors are

rather small for labour and material (0.083% and 0.357%, respectively), but non-negligible for the more volatile fuel (5.33%).

The one-run performance metrics are summarized in cells AQ38:AQ40 and these results are used to populate the 1,000 rows of the *What-If* table below it. Finally, the means of these 1,000 runs are in cells AS38:AS40. Both results are presented in Table A-3.

Table A-3 — Results of 3-year Average Growth Rate			
Bias	Labour	Fuel	Material
1 record	0.59%	-1.88%	-2.89%
1,000 rec.	0.04%	0.02%	0.07%

Bias is the average forecast error percentage and indicates whether, over the simulated runs, the method tends to over-forecast, under-forecast, or remain close to neutral. These results indicate that the Monte-Carlo forecasts of Method 3 are neutral. Moreover, Method 3 results over the historical record have improved over the status-quo method. Comparing Table A-3 with Table A-1 for the historical record, the three-year average growth-rate method substantially reduced the magnitude of expected bias for the labour, fuel and material indices.

More importantly, Method 3 substantially reduces the bias risk, especially for fuel, as presented in Table A-4 that shows the bias frequency distribution of the 1,000 runs.

Table A-4 Method 3 Bias Risk			
Within $\pm$	Labour	Fuel	Material
1%	80%	12%	46%
2%	99%	23%	79%
3%	100%	34%	93%
4%	100%	44%	99%
5%	100%	53%	100%

Table A-4 summarizes the risk profile by showing the share of simulated bias outcomes that fall within selected error bands. For the three-year method, 100% of labour and material bias outcomes fall within the $\pm 5\%$ range. Fuel is materially more volatile: only 12% of fuel bias outcomes fall within $\pm 1\%$, 34% fall within $\pm 3\%$ and 53% fall within $\pm 5\%$. Put differently, there remains a 47% chance of a fuel bias error greater than 5% in absolute value, and approximately a 66% chance of an error greater than 3%. This result is significant because it shows that, for a highly volatile index such as fuel, positive and negative forecast errors do not necessarily always cancel each other out, even over a 10-year evaluation horizon.

Comparing Table A-4 with Table A-2, 100% of labour and material forecasts are expected to fall within $\pm 5\%$ for Method 3, compared to the status quo 91% and 55%, respectively. For fuel, 53% is now expected to fall within $\pm 5\%$ compared to only 27% for the status-quo that additionally had no results within $\pm 3\%$. The risk of a fuel forecast falling outside the $\pm 5\%$ range has dropped significantly

from the status-quo value of 73% to the Method 3 value of 47%, a result that is more acceptable to CN.

Charts of the bias distributions summarized in Table A-4 provide an additional way to assess risk. The labour, fuel, and material bias distributions are shown respectively in the charts beginning at AA47, AA62, and AI47. These charts show not only the average result but also how widely the method's performance can vary across plausible price paths. This distinction is important because a method that is unbiased on average may still create unacceptable risk if its error distribution is too wide or if its tail outcomes are too severe.

This result is based on the 1,000 simulated price paths to evaluate the expected performance of the three-year method across a broader range of plausible outcomes, rather than relying only on the single historical path that occurred.

Tables A-3 and A-4, when compared to Tables A-1 and A-2, indicate that the three-year method is expected to perform materially better than the status-quo for all three indices.

7.2 Results from the Five-Year Method

The tab *One-Version (M.4)* replicates the same evaluation framework using a five-year average growth rate rather than a three-year average growth rate. In other words, the worksheet uses the same one-version historical benchmark, the same Monte Carlo price paths, the same dynamic row-selection mechanism, the same *Data Table* approach, and the same performance metrics. The only substantive difference is the look-back period used to calculate the average growth rate that feeds the forecast. This structure allows the three-year and five-year methods to be compared on a like-for-like basis, with any performance difference attributable to the forecasting horizon rather than to a difference in benchmark data, simulation design, or metric definition.

The five-year method produces slightly lower average bias than the three-year method for labour and material, but a higher negative bias for fuel, as shown in Table A-5. For the less volatile labour and material indices, the longer averaging period reduces already small biases from 0.04% to 0.01% and from 0.07% to -0.01%, respectively, which is not a materially important improvement. For the more volatile fuel index, however, the longer window appears to delay the forecast's adjustment to price shocks. Before the effect of one shock is fully reflected, another may occur, causing the forecast to lag persistently and increasing its negative bias from 0.02% to -1.29%. The three-year method responds more quickly to changing fuel prices and therefore produces a lower bias.

Table A-5 — Results of 5-year Average Growth Rate			
Bias	Labour	Fuel	Material
1 record	0.92%	-4.60%	-3.36%
1,000 rec.	0.01%	-1.29%	-0.01%

The more important effect of the five-year method is to widen the distribution of possible bias outcomes and increase bias risk as shown in Table A-6.

Table A-6 Method 5 Bias Risk			
Within $\pm$	Labour	Fuel	Material
1%	72%	10%	43%
2%	97%	20%	72%
3%	100%	30%	87%
4%	100%	39%	96%
5%	100%	47%	99%

This wider dispersion is material across all three index components. For example, the $\pm 1\%$ band for labour was reduced from 80% to 72%, for fuel it was reduced from 12% to 10%, and for material the reduction was from 46% to 43%. At the $\pm 5\%$ level, labour and material remained at ~100%, but fuel dropped from 53% to 47%.

The practical implication is that the five-year method increases forecast risk without providing a significant improvement in average accuracy. Wider bias distributions increase the probability of larger forecast errors, particularly for the more volatile fuel index.

7.3 Summary Table Supporting the Method Comparison

Results Summary

Status Quo Method

Bias

Historical	Labour	Fuel	Material
1 record	-0.34%	-4.91%	-7.35%

Bias Risk

Within ±	Labour	Fuel	Material
1%	27%	0%	18%
2%	36%	0%	27%
3%	55%	0%	36%
4%	64%	9%	55%
5%	91%	27%	55%

Average Growth Rate Method

Method 3 (3-yr average)

	Labour	Fuel	Material
1 record	0.59%	-1.88%	-2.89%
1,000 rec.	0.04%	0.02%	0.07%

Within ±	Labour	Fuel	Material
1%	80%	12%	46%
2%	99%	23%	79%
3%	100%	34%	93%
4%	100%	44%	99%
5%	100%	53%	100%

Method 4 (5-yr average)

	Labour	Fuel	Material
1 record	0.92%	-4.60%	-3.36%
1,000 rec.	0.01%	-1.29%	-0.01%

Within ±	Labour	Fuel	Material
1%	72%	10%	43%
2%	97%	20%	72%
3%	100%	30%	87%
4%	100%	39%	96%
5%	100%	47%	99%

8. Conclusion: The Three-Year Method Provides the Better Forecasting Balance

From a policy perspective, the choice between a three-year and a five-year averaging method is not a simple mechanical question. Each option reflects a different balance between responsiveness, accuracy, dispersion, and risk. A five-year method may appear attractive because it smooths short-term movements and may seem less exposed to year-to-year volatility. That smoothing, however, comes at a cost: it can make the forecast less responsive when cost conditions are changing – either up or down – and, as the simulation results show, it does not necessarily improve average accuracy.

The three-year method, by contrast, places greater weight on more recent index movements and is therefore more responsive to current cost conditions. That responsiveness could be viewed as a potential weakness if it merely amplified temporary noise. The evidence here does not support that concern. On the contrary, the slow responsiveness of the five-year method produces a large negative bias for fuel. On a like-for-like Monte Carlo comparison, the three-year method achieves average bias results that are broadly comparable to the five-year method, and better for fuel, while producing materially lower dispersion and a lower probability of large errors, especially for fuel. For that reason, the policy choice is not between a “stable” five-year method and a “volatile” three-year method; it is between a longer-window method that increases forecast-risk dispersion and a shorter-window method that preserves responsiveness without sacrificing average accuracy and reduces the probability of large positive and negative errors.

The comparison between the three-year and five-year methods is methodologically clean because both approaches are evaluated using the same benchmark data, the same simulated price paths, and the same performance metrics. The only substantive difference is the look-back period used to calculate the forecast growth rate. On that like-for-like basis, the Monte Carlo results support the three-year method.

The five-year method does not materially improve average accuracy. Its average bias is broadly similar to the three-year method, meaning that the longer look-back period does not produce a more accurate central forecast. Instead, its main effect is to increase dispersion materially across all three index components. This is not a minor trade-off: it means the five-year method produces a wider range of possible forecast errors without a corresponding gain in average accuracy.

That wider distribution is especially important for fuel. Under the three-year method, only 53% of simulated fuel bias outcomes fall within $\pm 5\%$. Under the five-year method, that share falls to 47%. For the most volatile component of the VRCPI, the five-year method therefore increases the probability of larger forecast errors while offering no material improvement in average bias.

The stronger conclusion is therefore that the three-year method provides the better forecasting balance. It remains responsive to recent changes in the underlying indices, while avoiding the materially wider bias distributions produced by the five-year method. The five-year method may appear more stable in concept because it averages over a longer period, but the simulation results demonstrate that, in practice, it increases forecast risk rather than reducing it, and for the volatile fuel its low responsiveness produces a large negative bias.

Accordingly, the three-year method should be preferred. It achieves comparable or better average bias with materially lower dispersion and a lower probability of large forecast errors, particularly for fuel. From a regulatory and forecasting perspective, that is the more defensible result: a method should not be adopted merely because it uses a longer historical window if that longer window does not improve accuracy and instead increases the risk of larger errors.

9. References

1. "[Geometric Mean: Formula & When to Use It.](#)" CASRAI, last updated 23 August 2026. The source states that the geometric mean is the appropriate average for growth rates, ratios, and log-normally distributed data. This supports the use of the geometric mean for multiplicative index changes and compound growth calculations. <https://casrai.org/guides/geometric-mean>
2. Marsden, Eric. "[Monte Carlo Methods for Risk Analysis: Stochastic Simulation and Numerical Experiments.](#)" *Risk Engineering*, 2018. The source describes Monte Carlo methods as numerical experiments used to observe average outcomes over many runs of a stochastic model, involving repeated random sampling from input probability distributions, model execution using those stochastic inputs, and aggregation of results over many executions. The accompanying slide deck presents the Monte Carlo workflow. <https://risk-engineering.org/monte-carlo-methods/>
3. https://en.wikipedia.org/wiki/Log-normal_distribution