We Don't Know Everything:

Modelling the sugar industry and decision-making in an uncertain world

Jack Fisher
Research & Technology, American Sugar Refining
jack.fisher@asr-group.com

1. Introduction

Almost all companies in the sugar industry nowadays use models to help make decisions. The models might be technical, economic or both, and can range in complexity from a simple cash flow analysis to a full mass & energy balance for a refinery. What the vast majority have in common is that they are deterministic models, where the input data is fixed and static, and therefore so are the outputs.

By using models with fixed inputs, what we are saying is that we know with 100% confidence that our assumptions are correct; for example, that the sugar price will be $500/tonne, or the feed sugar colour will be 1000 IU. In reality, however, there is virtually nothing that we can know or predict with 100% confidence.

This paper will show how we can build this uncertainty into our modelling by using Monte Carlo simulation to convert deterministic into stochastic models. Firstly, a simplified fictional example will be described to demonstrate the methods. Then, three real-world examples will be given: (1) estimating the operating cost for a new refinery; (2) comparing potential refining processes; and (3) identifying areas for optimisation in an existing refinery.

In each example, the paper will show how Monte Carlo simulation, together with holistic modelling, can show the range of potential outcomes and the likelihood of each occurring, the key factors driving the variations, and how this can help with decision making. Finally, the paper will also touch on real-world examples of the expensive consequences of not taking uncertainty into account – of assuming that we know everything.

2. Simplified demonstration of Monte Carlo simulation

Monte Carlo simulation is a computerised mathematical technique that can give the decision-maker a range of possible outcomes and the probabilities that they will occur, for any choice of action. The technique was first used by scientists working on the Manhattan Project, and is used today in areas including finance, project management, energy, manufacturing, engineering, research and development, insurance, oil & gas, transportation, and the environment. The method can be described as comprising the following basic steps:

1. Replace any input parameter which is subject to inherent uncertainty with a range of values, represented by a probability distribution.

2. Recalculate the model over and over again, each time using a different set of random inputs as sampled from the probability distributions.

3. Aggregate the results from each recalculation and generate probability distributions for each output value.

In this paper the technique was carried out using @RISK software[1], which can be used on any Excel spreadsheet.

The method can be illustrated by the following simplified example. Suppose we are contemplating building or investing in a new autonomous sugar refinery. To help make our decision, we've built a very simple economic model, which calculates (a) the annual cash flow based on the profit margin and the production rate; (b) the annualised cost of capital[2], based on the CAPEX, discount rate and economic life; and from the two we get the net annual profit.

There are only six inputs, and we estimate an operating cost of $65/t, a refining margin (difference in price between feed and product sugar) of $110/t (the average world white premium in 2013), a CAPEX of $200m, a production rate of 1,000,000 tonnes per year, a discount rate on capital of 10% and an economic life of 15 years.

Using these assumptions, the results from our simple model are shown in Table 1:

Table 1: Simple model inputs & outputs (single-point estimate)

Inputs	
CAPEX	$200,000,000
Operating costs (/tonne sugar)	$65
White premium (/tonne sugar)	$110
Production rate (tonnes/year)	1,000,000
Discount rate	10%
Economic life (years)	15
Outputs	
Cash flow (annual)	$45,000,000
Cost of capital (annual)	-$26,295,000
Annual Profit	**$18,705,000**

Now it's decision time. Do we proceed beyond the initial stage of the project? The annual profit looks healthy, so we might decide to press ahead. Or, we might acknowledge that there is uncertainty in our assumptions, and decide to test the model using what-if scenarios. One way of doing this is via a three-point estimate. Here we estimate worst case, expected and best case values for the assumptions instead of the previously used single values.

As is common in most models, we were probably conservative with the single-point estimates. For the operating cost, we might actually expect a value of $60/t, with $55 and $80 as the best and worst case. The CAPEX might include 10% contingency and a confidence

[1] http://www.palisade.com/risk/

[2] The annualised cost of capital is calculate by applying a Capital Recovery Factor, which uses an interest rate (i) and project life (n) to determine the rate at which earnings could reasonably be expected if the same funds were invested over a length of time The formula is:
$ (annual) = $ (total) x CRF, where CRF = {i(1 + i)^n} / {[(1 + i)^n]-1}.

level of +-30%. The actual production rate might vary by +5%, -10%, and discount rate on capital between 8 and 12%. For the white premium, we might decide to take the minimum, average and maximum values for the 5-year-average world white premium over the last 25 years: $80/t, $101/t and $111/t.

Using these values, we can run the model and obtain the results for the three cases, as shown in Table 2.

Table 2: Simple model inputs & outputs (single-point estimate)

Inputs	Worst	Expected	Best
CAPEX	$234,000,000	$180,000,000	$126,000,000
Operating costs (/tonne sugar)	$80	$60	$55
White premium (/tonne sugar)	$80	$101	$111
Production rate (tonnes/year)	900,000	1,000,000	1,050,000
Discount rate	12%	10%	8%
Economic life (years)	15	15	15
Outputs	Worst	Expected	Best
Cash flow (annual)	-$70,000	$40,757,000	$58,502,000
Cost of capital (annual)	-$34,357,000	-$23,665,000	-$14,721,000
Annual Profit	**-$34,427,000**	**$17,092,000**	**$43,781,000**

What do the results tell us? We can see that the expected case is similar to our original estimate, that the best case looks very good, but also that the worst case does not look good, and that we could lose a lot of money. The problem here is that while we're aware that there is a significant potential downside, we don't actually know how likely it is. We know the consequences of the risk, but not the likelihood. We're not really any better equipped to take our decision.

This is where Monte Carlo simulation can help. Firstly, we take our three-point estimates from before and convert them into probability distributions. There are various probability distributions we could use (common distributions include normal, lognormal, uniform and triangular) but in this scenario we will use the PERT distribution for each input. The PERT distribution was developed to model expert estimates, and requires three values for its definition: minimum, most likely and maximum. Its shape is similar to a triangular distribution, but with a greater likelihood of values occurring in the region around the most likely value. We will use our three-point estimates from earlier to define the distributions.

As an example, Figure 1 below shows the distribution for operating costs. The vertical axis represents probability, so this distribution indicates that there is a greater probability of values occurring around $60/t. The vertical bars highlight the 90% confidence band: in this case, the distribution implies that we have 90% confidence that the operating costs will be between $56/t and $70/t.

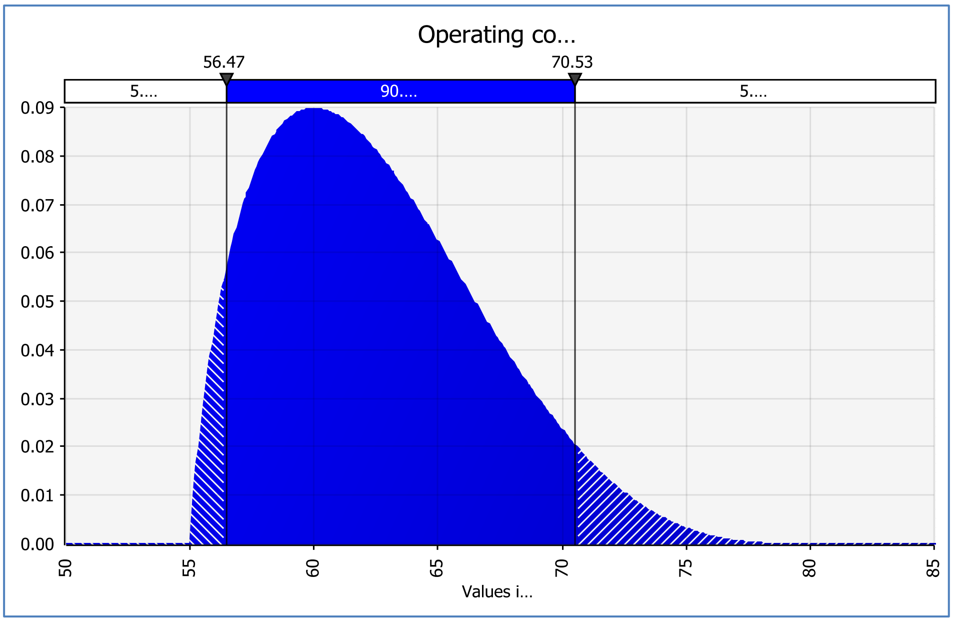

Figure 1: Probability distribution for operating costs ($/tonne)

For the white premium, we can actually use historical data to create the distribution. Figure 2 below shows the five-year-average world white premium for the last 20 years, based on published data.[3]

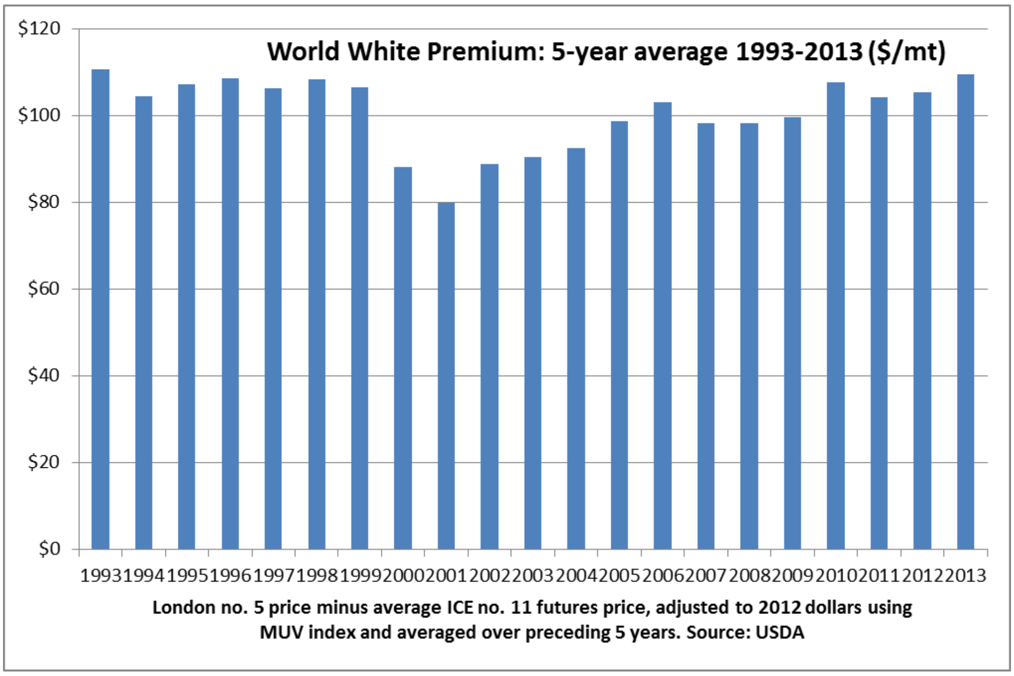

Figure 2: Five-year-average world white premium, 1993-2013.

We can take this same raw data and display it in Figure 3 in the form of a frequency distribution, shown in blue. We can then fit a probability distribution to the data, and the triangular distribution shown in red is the best fit. We can then use this distribution in our model; effectively what we are doing is using past variations as a guide to future variability.

---

[3] http://www.ers.usda.gov/data-products/sugar-and-sweeteners-yearbook-tables.aspx

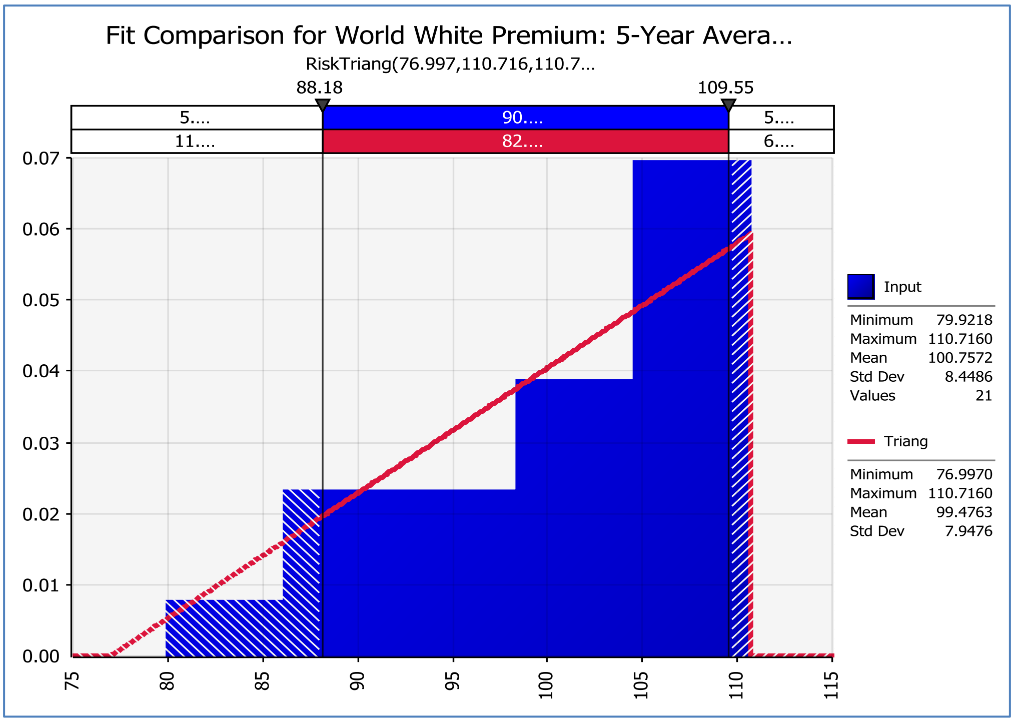

Figure 3: Frequency (blue) and fitted probability (red) distributions for world white premium

Having defined all our model inputs as probability distributions, we can now recalculate our model hundreds or thousands of times, and each time the input data will vary randomly according to the distributions we have defined. We can simulate thousands of what-if scenarios to build up a fuller picture of the likely outcomes. Figure 4 shows the outcomes, in terms of annual profit, of 10,000 model simulations. The result from each simulation is plotted on a frequency distribution, giving a visual representation of the potential variation in results. The average annual profit from the 10,000 simulations is around \$13m/y, significantly lower than the \$18.7m/y and \$17m/y previously estimated using our cruder single-point and three-point estimates. As before we can see the potential for both significant upside and downside, but now we can see the estimated likelihood of this occurring: 5% probability of profit exceeding \$27m, and a 10% probability of making a loss.

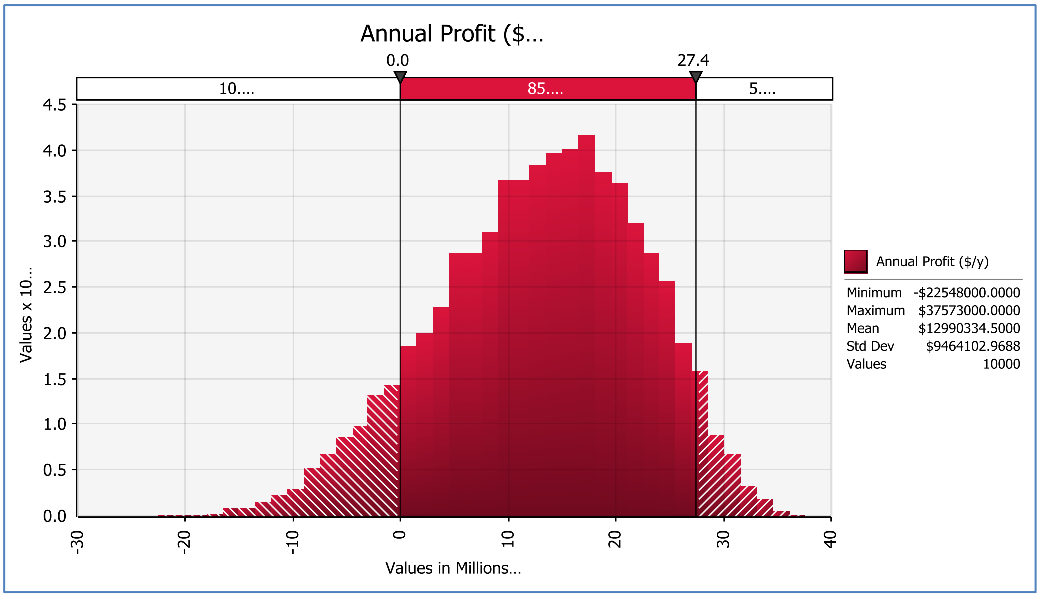

Figure 4: Distribution of Annual Profit output from 10,000 model simulations

How does this help us make our decision? Well, we can see our average expected profit is significantly lower than we thought previously, but more importantly, there is a significant (10%) risk of making a loss, so we might decide to cancel the project on this basis.

Alternatively, we could investigate what is driving this risk and see if we can mitigate it. Figure 5 shows a sensitivity analysis chart for our 10,000 simulations. This ranks each of the variable inputs in terms of the effect of their variability on the annual profit, and we can clearly see that variability in the white premium is our biggest risk. Now we know this, and given that we've seen the project could potentially be lucrative, we might want to explore ways of reducing our exposure to these variations.

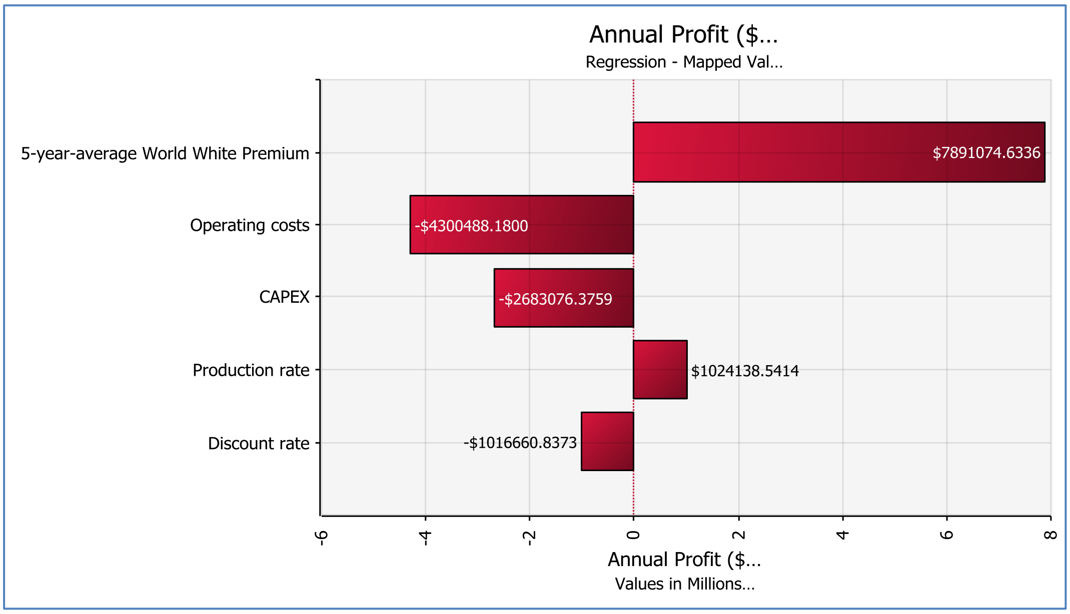

Figure 5: Sensitivity analysis chart showing relative impact of input variability

Let's say that for a fee of $2 per tonne of sugar, which we can add to the operating costs, we can hedge to impose an artificial floor of $95/t to the white premium, so that we are limiting our exposure to the potential downside. We modify our model accordingly, and Figure 6 shows the results from 10,000 simulations.

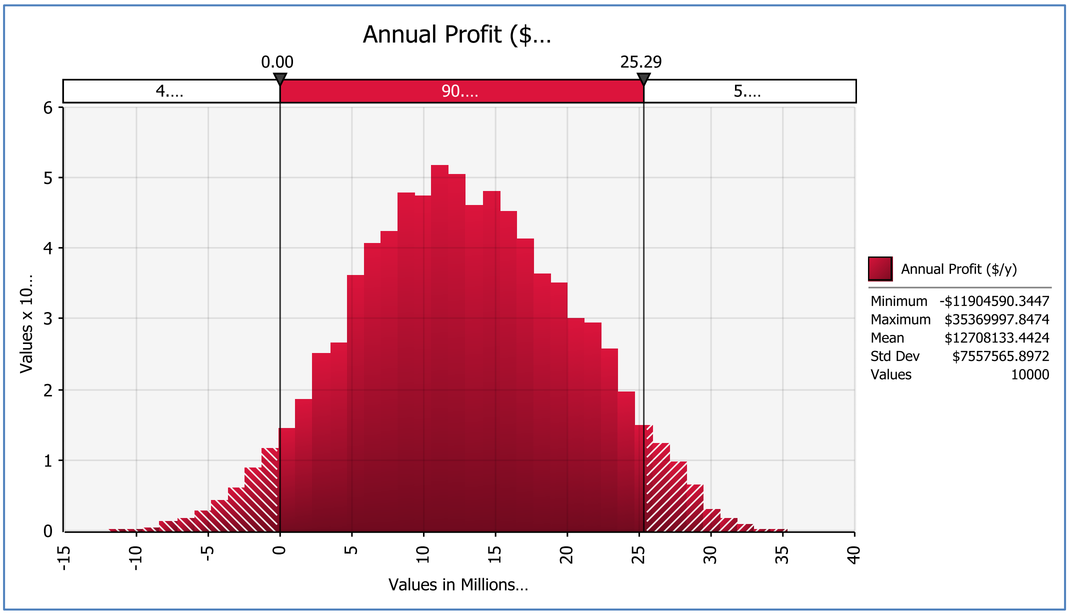

Figure 6: Distribution of Annual Profit output after white premium hedging

We can see that we have slightly reduced our expected average result, and the potential upside, but more importantly, we have reduced the risk of an annual loss to less than 5%. Depending on our appetite for risk, this may be sufficient for us to decide to proceed with the project.

This is a fictional, and very simplified, example of how Monte Carlo simulation can be used to help in decision-making. In reality, nobody is (hopefully!) going to decide on a $200m investment based on such a simple model with only 6 inputs. In the next section, we will explore three real examples of how these methods can be used to help in decision-making.

3. Examples of real applications in the sugar industry

We will now look at three real examples: (i) estimating operating costs for a new refinery; (ii) comparing potential refining processes; and (iii) identifying areas for optimisation in an existing refinery.

3.1. Estimating operating costs for a new refinery

In the model used for the simplified demonstration in section 2 there were only 6 inputs, one of which was the operating costs. In reality, operating cost is a complex variable subject to variability in many other parameters, for example: energy cost; sugar price; feed colour; process efficiencies; chemical prices; waste disposal costs; and so on. In reality, when estimating operating costs for a new refinery, a holistic model should be built including the processes, utilities and associated economics, based on mass & energy balances, with the operating costs estimated from the results. Such a model will have many inputs, and is obviously much more complex than the simple model shown in section 2. Nevertheless, the model can be built in a way that allows the inputs to be changed from static to variable, and defined as probability distributions, just like the simple model.

Figure 7 shows the distribution of operating costs (OPEX) resulting from 5000 simulations of a complex model for a specific refinery project, and Figure 8 shows the sensitivity analysis chart.

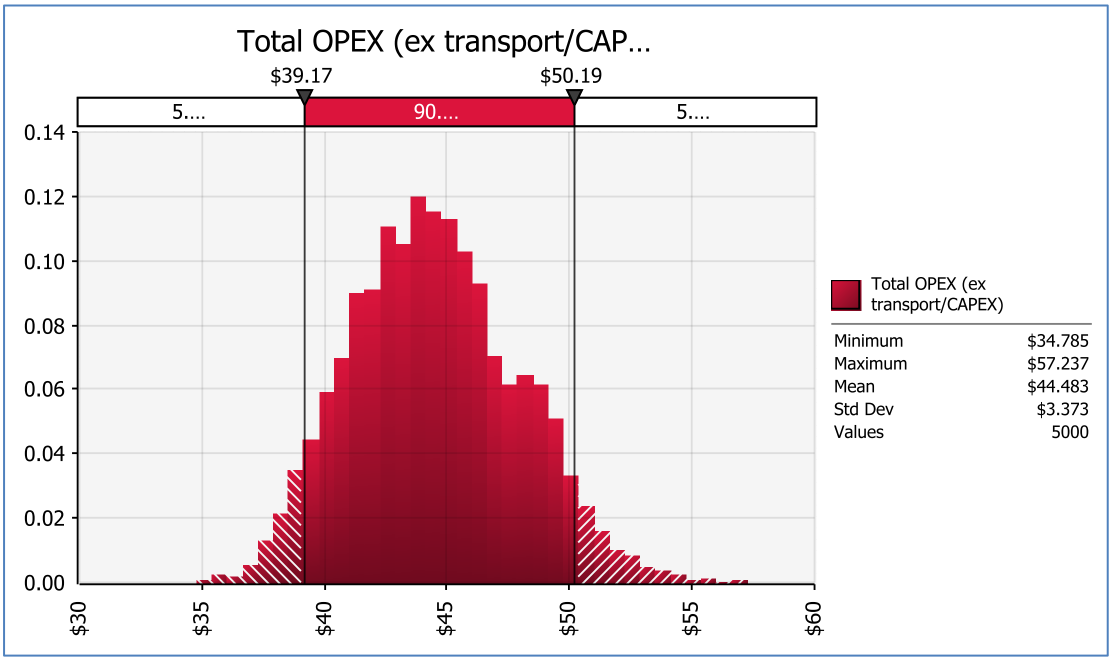

Figure 7: Distribution of operating costs from 5,000 simulations of new refinery model

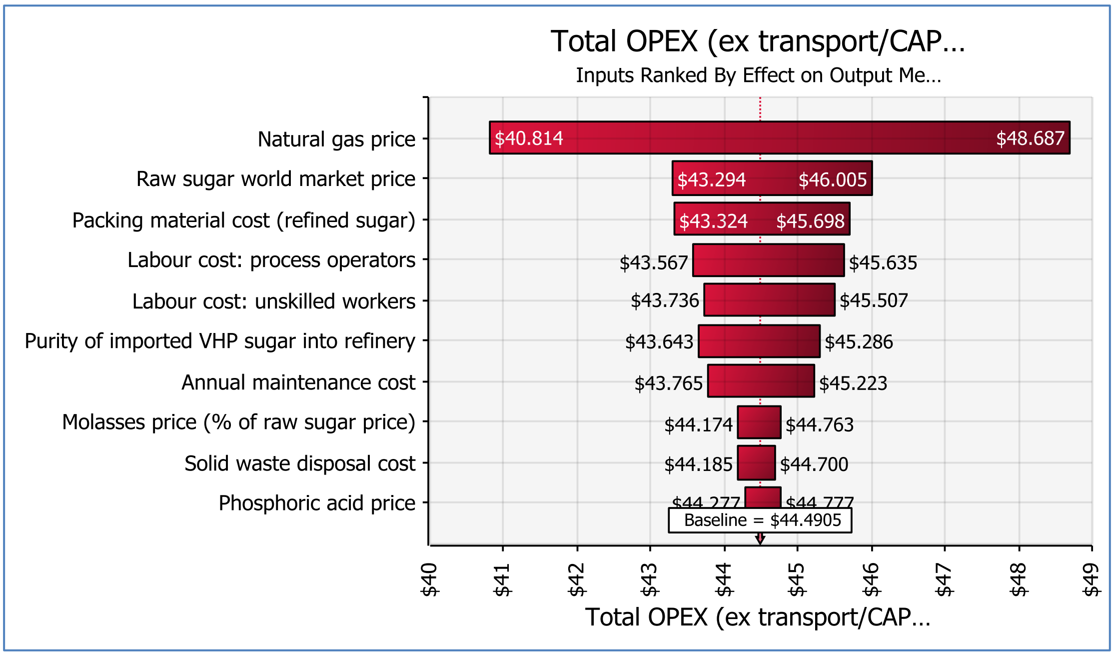

Figure 8: Sensitivity analysis chart for operating costs for a new refinery

What the simulation shows is that instead of a single fixed operating cost, there is a range of potential values between $35 to $57/t. It shows that the average, or expected, cost is $44.5/t, and it also shows that we can be around 90% confident that the costs will be between $39 and

$50/t. It also shows clearly that the biggest risk to the operating costs is the gas price. Overall, this means that we can (a) give more confidence in the economic viability of the refinery, and (b) highlight the importance of securing a supply of natural gas at a reasonable price.

3.2. Comparing potential refining processes

For new (and some existing) refineries, the choice of refining processes is not fixed and automatic. Probably the most common and significant decisions lie in choosing which processes to use for clarification (or primary decolourisation) and for secondary decolourisation. For clarification, the choice is usually between phosphatation and carbonatation, and for secondary decolourisation, the choice is usually between ion exchange and granular (or in some cases powdered) activated carbon, or a combination of each.

Although the choice of process is not solely a financial one, for a sensible and informed decision to be made it is important to estimate the operating costs for each option. Models of each process can be built to calculate the operating costs. The inputs to the model can vary according to the scenario; for example, the steam or electricity cost, or the waste disposal options and cost, can vary from refinery to refinery. These potential variations can be explored via Monte Carlo simulation, to build a picture of the likely benefits of one process over another, and the key drivers behind those benefits.

For example, Figure 9 shows the distribution of results from a comparison model of phosphatation and carbonatation. The key result from the model is the operating cost differential: the operating cost for carbonatation minus the cost for phosphatation. This shows, interestingly, that in approximately half of the 10,000 simulations the results favoured carbonatation and in the other half, phosphatation, i.e. the average operating cost differential is around zero. It should be noted here that this does not necessarily mean that half of all real-life scenarios favour carbonatation over phosphatation, and vice versa. The variability of inputs to the model was set to explore the drivers behind the choice of process, not to accurately reflect the actual variability in scenarios worldwide. However, it does indicate that purely on the basis of operating costs, the choice between carbonatation and phosphatation is not a straightforward one.

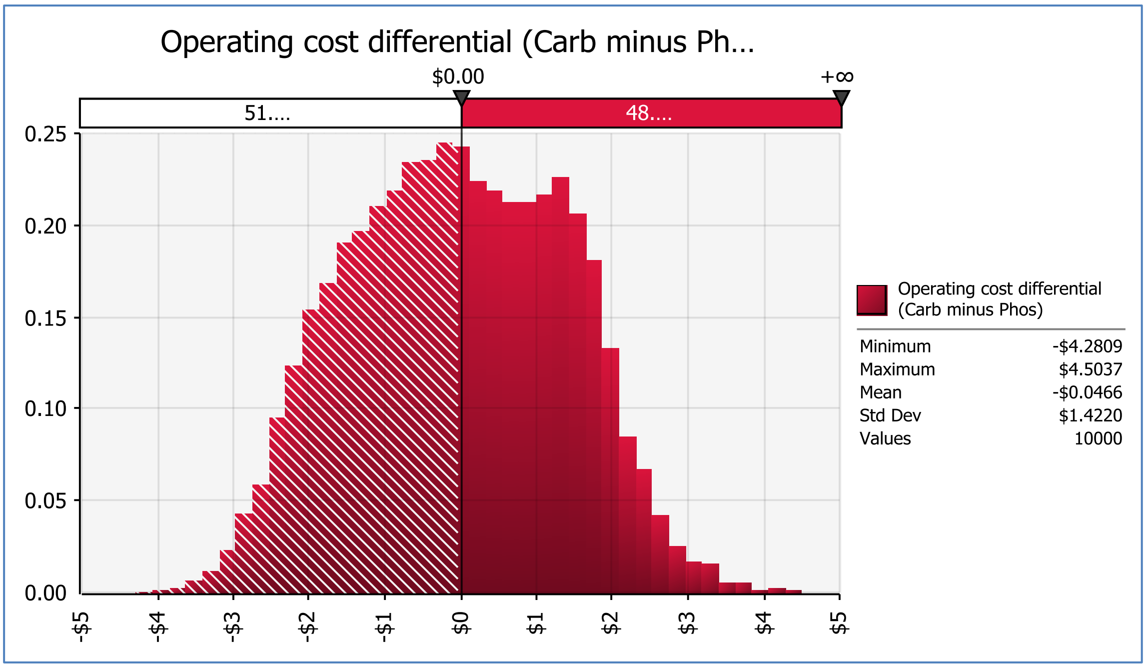

Figure 9: Distribution of operating cost differential from 10,000 simulations of a phosphatation vs carbonatation model

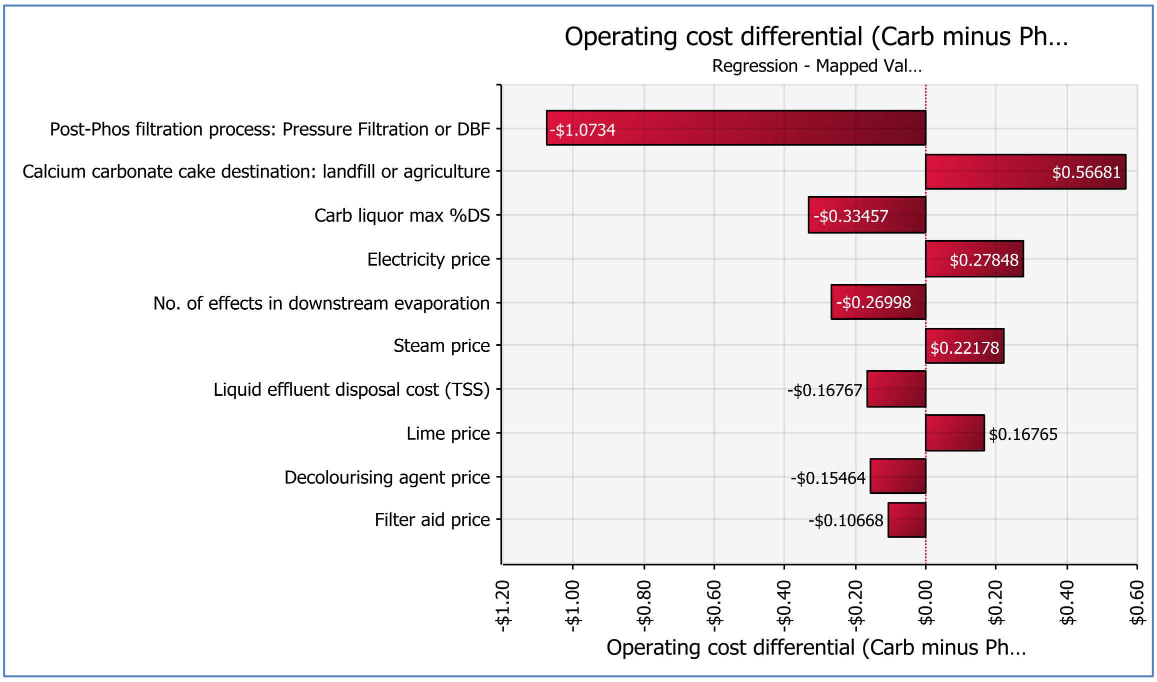

Figure **10** shows the sensitivity analysis chart for the 10,000 simulations. Bars on the right-hand-side of the axis indicate that an increase in the parameter favours phosphatation, and vice versa. This shows that the biggest drivers between the two processes are: (1) the type of filtration process required after phosphatation; (2) the available destination for calcium carbonate cake; (3) the maximum liquor Brix to the post-carbonatation filters; (4) the electricity price; and (5) the number of effects in downstream evaporation.

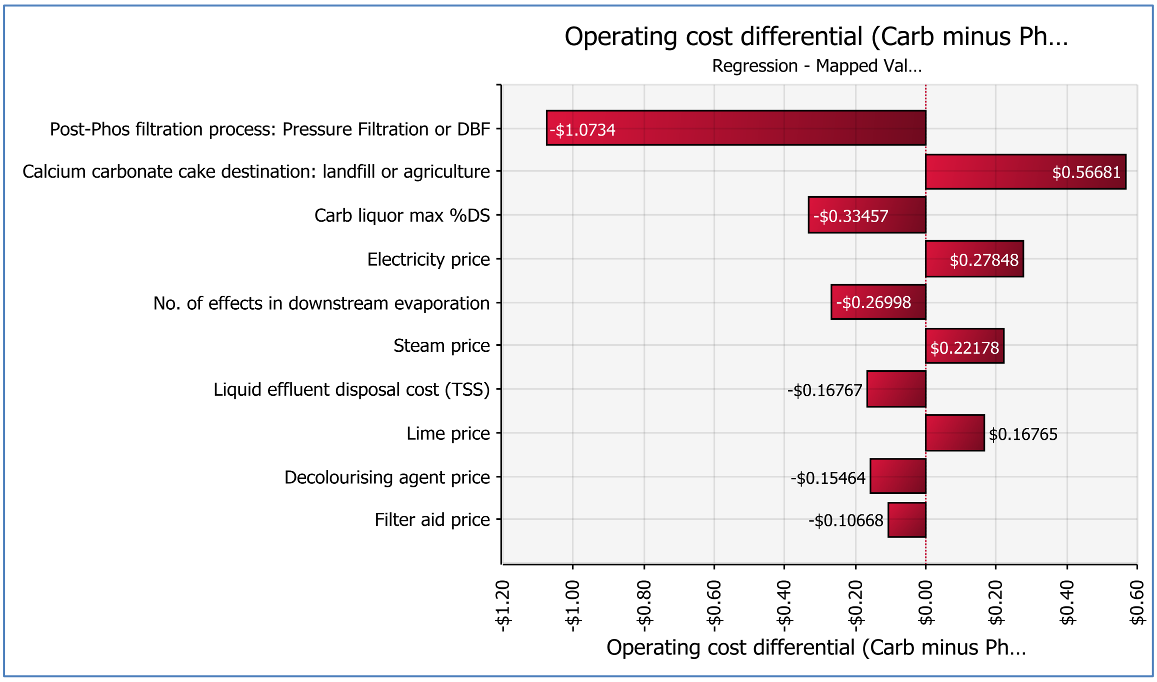

Figure 10: Sensitivity analysis chart for operating cost differential between phosphatation and carbonatation

In this example, Monte Carlo simulation has shown us the potential difference in operating costs between the two processes. It has also shown the scenario- and location-specific factors that are most likely to impact the difference in operating costs. This helps us to understand the decision between the processes better, and which areas to focus on to make each process more cost-effective.

3.3. Identifying areas for optimisation in an existing refinery

When analysing an existing refinery, just as for the new refinery example in section 3.1, we can build a holistic model of the refinery, based on mass and energy balances. We can then use that model to explore various what-if scenarios and identify areas with scope for optimisation. Here, again, Monte Carlo simulation can be usefully applied.

For example, we can explore how energy usage can be reduced by changing day-to-day operating parameters rather than by capital projects. Model inputs related to these parameters are set to variable, with distributions defined according to practically feasible ranges. Figure 11 below shows the results of 1,000 simulations of the model, in terms of overall energy usage (expressed in mBtu per cwt raw sugar). This gives an indication of the potential saving achievable by varying operating parameters. The sensitivity analysis chart in Figure 12 then highlights which parameters have the biggest effect on energy usage. Bars on the right-hand-side mean that an increase in the parameter will cause an increase in energy usage, and vice versa.

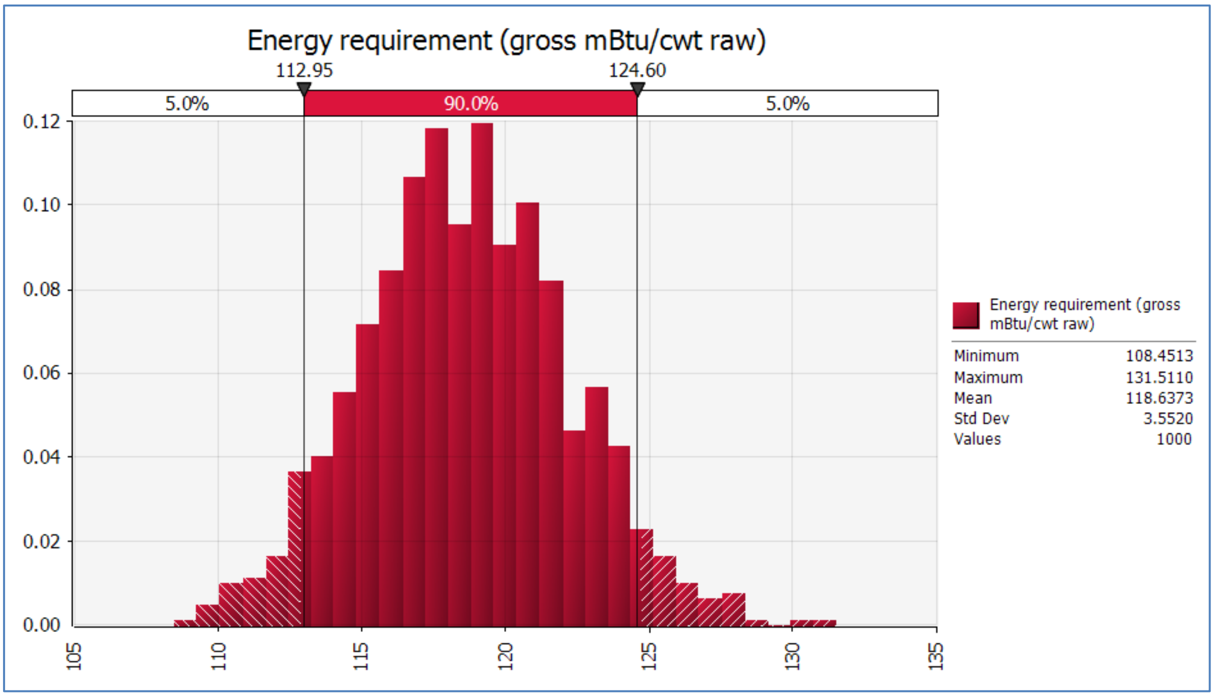

Figure 11: Distribution of energy usage (mBtu/cwt feed sugar) for an existing refinery model.

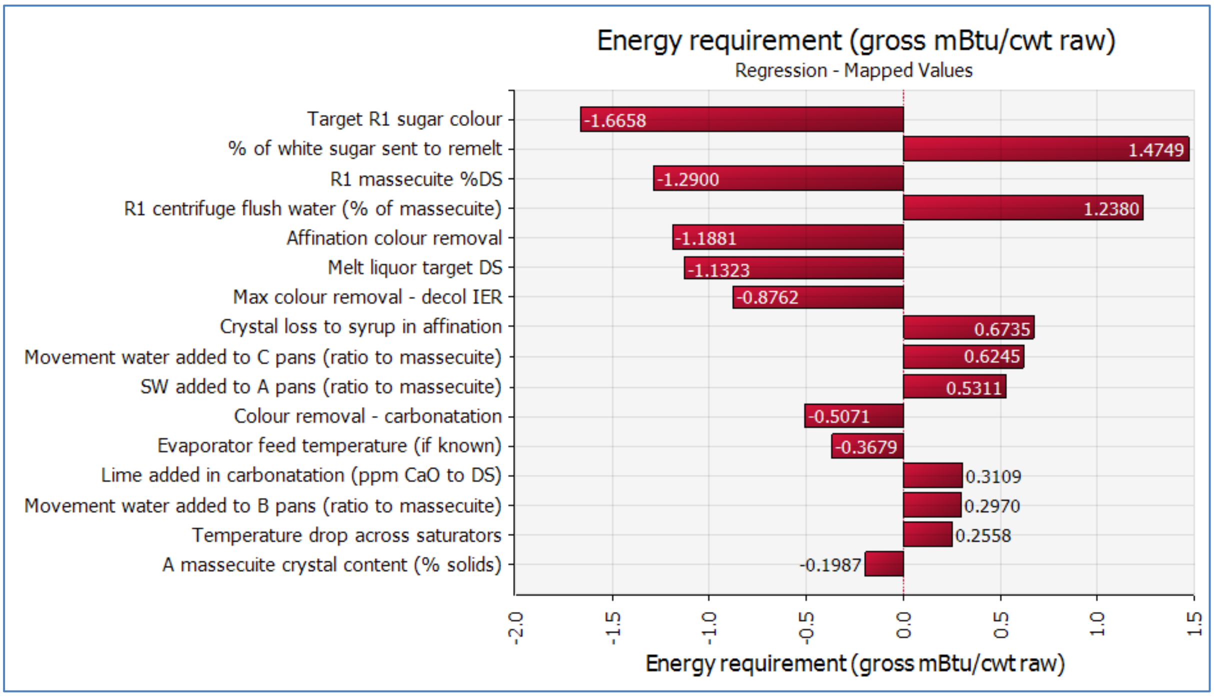

Figure 12: Sensitivity analysis chart for energy usage in an existing refinery.

Here, the simulation has helped us to understand the relative impact of altering process parameters and highlight which specific areas of operation to focus on to achieve our goal of reducing energy consumption.

4. Where Monte Carlo simulation should have been applied?

In the sugar industry, like in any other walk of life, not all decisions turn out to be successful. Often this is because the decision was based on a set of assumptions which didn't apply in reality. Here are some examples of real-life situations were Monte Carlo simulation might have highlighted the impact of those assumptions and helped to avert or mitigate the failure or problem, or at least highlighted the potential range of outcomes.

1. Feed quality. More than one sugar refinery has suffered due to a worse than expected feed colour, ash, or purity. In one particular refinery lacking crystallisation, this resulted in higher than expected operating costs, and higher effluents than the treatment plant was designed for, leading to problems maintaining production. The financial cost was significant. Monte Carlo simulation would have highlighted the sensitivity to feed quality and perhaps led to a stricter feed specification, averting the problem altogether.
2. Market demand. Overconfidence in market demand for products is not uncommon, and has led to problems with process plant turndown, lower than expected revenues, and in some cases has led to whole (or parts of) refineries being shut down temporarily or permanently. Again, Monte Carlo simulation would have highlighted the sensitivity to this assumption and shown a range of potential financial outcomes.
3. Energy price. Energy prices are notoriously unstable, but even so, investments are sometimes made on the basis of a fixed energy price assumption. There are examples including ethanol plants being mothballed for years due to lower than predicted oil prices, and alternative-fuel boilers being financially unviable due to an unexpected spike in fuel prices, all involving investments of millions of dollars.
4. Chemical price. Chemical prices too can be highly volatile. One plant was forced to shut down for several months due to a spike in the price of the humble chemical caustic soda. Again, Monte Carlo simulation would have highlighted this perhaps-unexpected sensitivity.
5. Process performance. There are numerous examples of refineries suffering worse than expected financial returns due to overconfidence in the performance of a process or processes, such as decolourisation, yields, Brix levels, reliability, throughput or thermal efficiency. Monte Carlo simulation could have been used to show the expected range of outcomes based on a range of potential process performance parameters, rather than just the expected values.

Of course, saying that Monte Carlo simulation would have mitigated all these problems is speculation – but hopefully this paper has shown that it is not unreasonable speculation.

5. Summary and conclusions

The purpose of this paper was to give a description of the Monte Carlo simulation technique and a flavour of how it can be usefully applied in the sugar industry. The technique is not perfect and it has to be applied intelligently, just like any modelling or simulation. But when applied properly it can shine more light onto the complexities behind decision-making. Nowadays we have access to more and more data, with greater levels of detail, but we still cannot predict the future – and we should beware the illusion of accuracy.

As human beings, we are modelling all the time, reducing the complexity of the world to a simplified and comprehensible version of it in order to make decisions efficiently. For example: in planning my journey to airport for this conference, my brain built a simplified model of the London traffic and transport network, based on its experience. The inputs to the model included the flight time, time of day, weather, probability of accidents, delayed trains, etc. The output was my departure time. In this model, my brain took into account the inherent uncertainty in the assumptions, considered the range of potential outcomes, the consequences of missing the plane (high!), and made a balanced decision on that basis.

The bottom line is that reality is uncertain, and if human nature tries to account for this uncertainty, so models should too. It is still people that make the decisions, but having a better understanding of the probable outcomes should help those people to make better decisions.