Using cost based time series to assess the calibration levels of in-flight major projects

Delft – 13 October 2016

Jamie Walker
Sellafield Limited and
University of Strathclyde Management Science Department

Tim Bedford and Matthew Revie
University of Strathclyde Management Science Department

The Sellafield Site

- Location:
  - Cumbria, North West England.

- Early history:
  - World War II: TNT storage site.
  - 1947: Research site for UK's first atomic bombs.
  - 1956: Windscale fire incident.
  - 1957: World's first commercial nuclear power station.

- Today:
  - Stores legacy waste in ageing structures.
  - Reprocesses spent nuclear fuel.
  - Nearby Moorside site will host a new nuclear power station.

World War II – TNT storage site

University of Strathclyde Business School

1947 - UK Atomic Bomb Programme

Today – Legacy Waste and Reprocessing

University of Strathclyde Business School

The Sellafield Limited Company

- Comprises the largest part (around 75%) of the UK's nuclear estate.
- The estate and Sellafield Limited (SL) is managed by the UK's Nuclear Decommissioning Authority (NDA).
- Annual costs are around £2bn and have steadily increased over the last five years.
- Reprocessing operations are scheduled to finish in 2018.
- From 2018 the main focus will be decommissioning the highly hazardous legacy ponds and silos.
- The long term strategy is to decommission the site and have it become a mostly brownfield area by 2120.

Short-term goals of the organisation

University of Strathclyde Business School

- To create realistic baselines for the durations and costs of all projects, but especially for the major ones (over £100m).
- To drive the durations and costs of projects to be as low as possible to more rapidly increase safety levels and get value for money for UK taxpayers.
- To take intelligently guided risks where interventions might increase short-term risk levels to decrease them in the long term.

Criticisms of Sellafield and the Nuclear Decommissioning Authority (NDA)

- Environmental groups and the Irish and Norwegian governments have repeatedly called for the company to be closed down.

- A 2012 UK National Audit Office (NAO) report highlighted duration and cost concerns – mainly about repeated increases in baseline figures of major projects.

- In April 2016 the privately owned consortium Nuclear Management Partners (NMP) lost the contract to run SL.
  - This followed intense criticism from the UK Government's Public Accounts Committee (PAC) which used the above NAO report as one of its main sources of evidence.

- A recent BBC Panorama exposé expressed safety and cost concerns (mainly from ex-NMP staff).

Why doesn't the organisation appear to be meeting its goals?

- NDA's and Taxpayer's desire for lower baselines could be making them unrealistically optimistic to begin with – this increases the chances of quality problems which then lead to re-work.

- The general understanding of probability is low amongst financial decision makers, so they might be taking bigger risks with duration and cost than they realise.

- The general approach of contractors bidding for large-scale engineering projects is to under-estimate to get the business.
  - But levels of under-estimation are likely to be higher on more uncertain projects.

- Actual spends and performance levels are routinely compared to baseline targets, but they are not routinely compared to model forecasts.

How is the organisation trying to overcome these possible causes?

University of Strathclyde Business School

- To adapt Klein's 'pre-mortem' technique to create a elicitation environments where experts are encouraged to be both optimistic *and* pessimistic, and to reduce the effects of (inner) anchoring bias.

- To try to encourage financial decision makers to openly and officially acknowledge the 'elephant in the room' of contractor under-estimation.

- To develop simple analogies around bias, gambling, and calibration levels to better explain to financial decision makers what a model's probability values are telling them.

- To run comparisons between the probabilistic time series forecasts of models and the officially published accountancy data for those time periods, and then present them as charts.

Comparison Charts

- The following charts will show direct comparisons between model forecasts (solid lines) and official accountancy data (dotted lines).

- The (solid) model lines go from P0 (lowest) to P100 (highest), with P5, P50, and P95 inbetween.

- The (dotted) accountancy data are best remember by:
  - **B**lue = **B**aseline
  - **P**ink = **P**erformance (Earned Value)
  - **A**mber = **A**ctual Spends

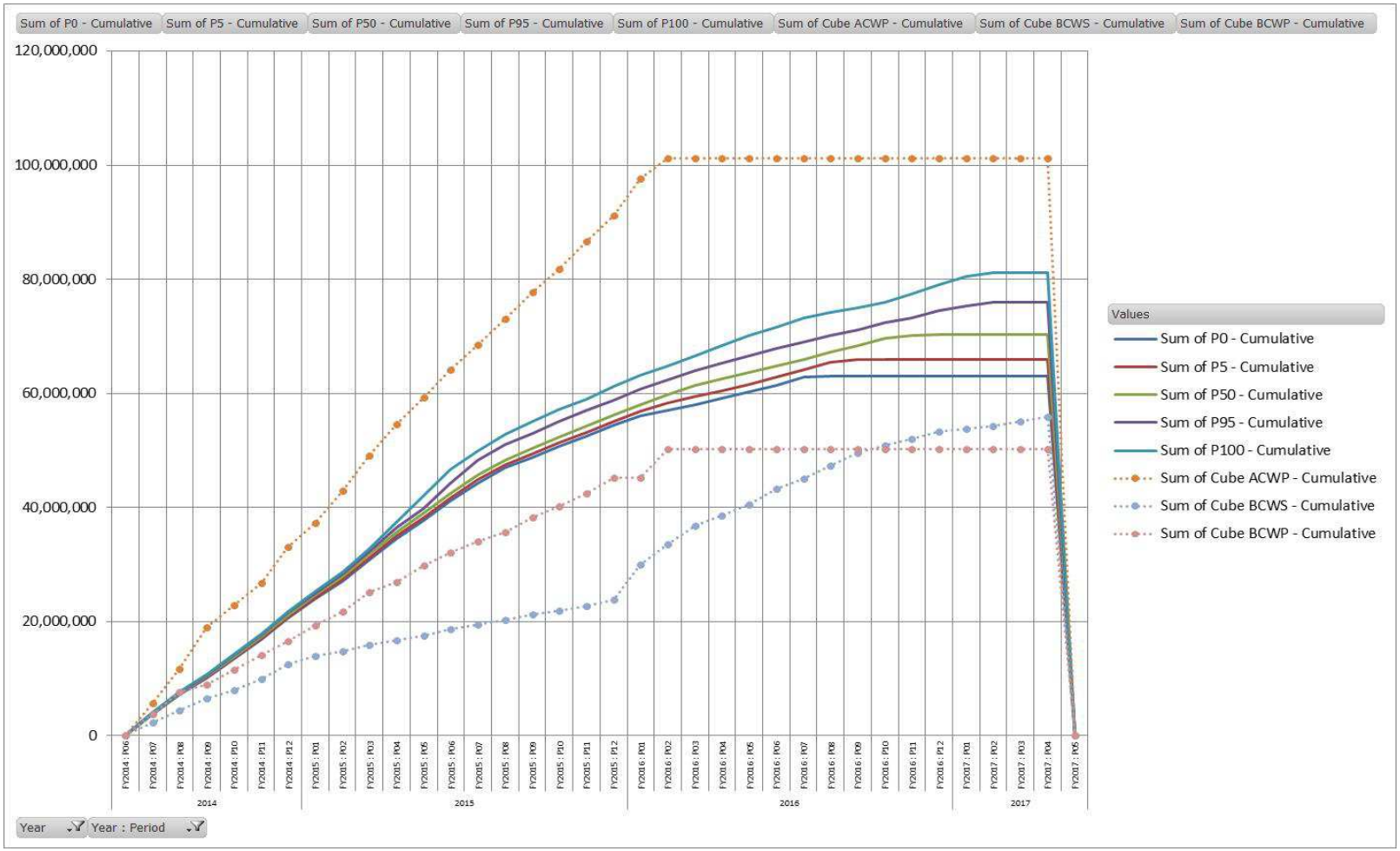
Project 1

- Important note: We generally use project 1's charts as examples of what we **don't** want to get as a later comparison.

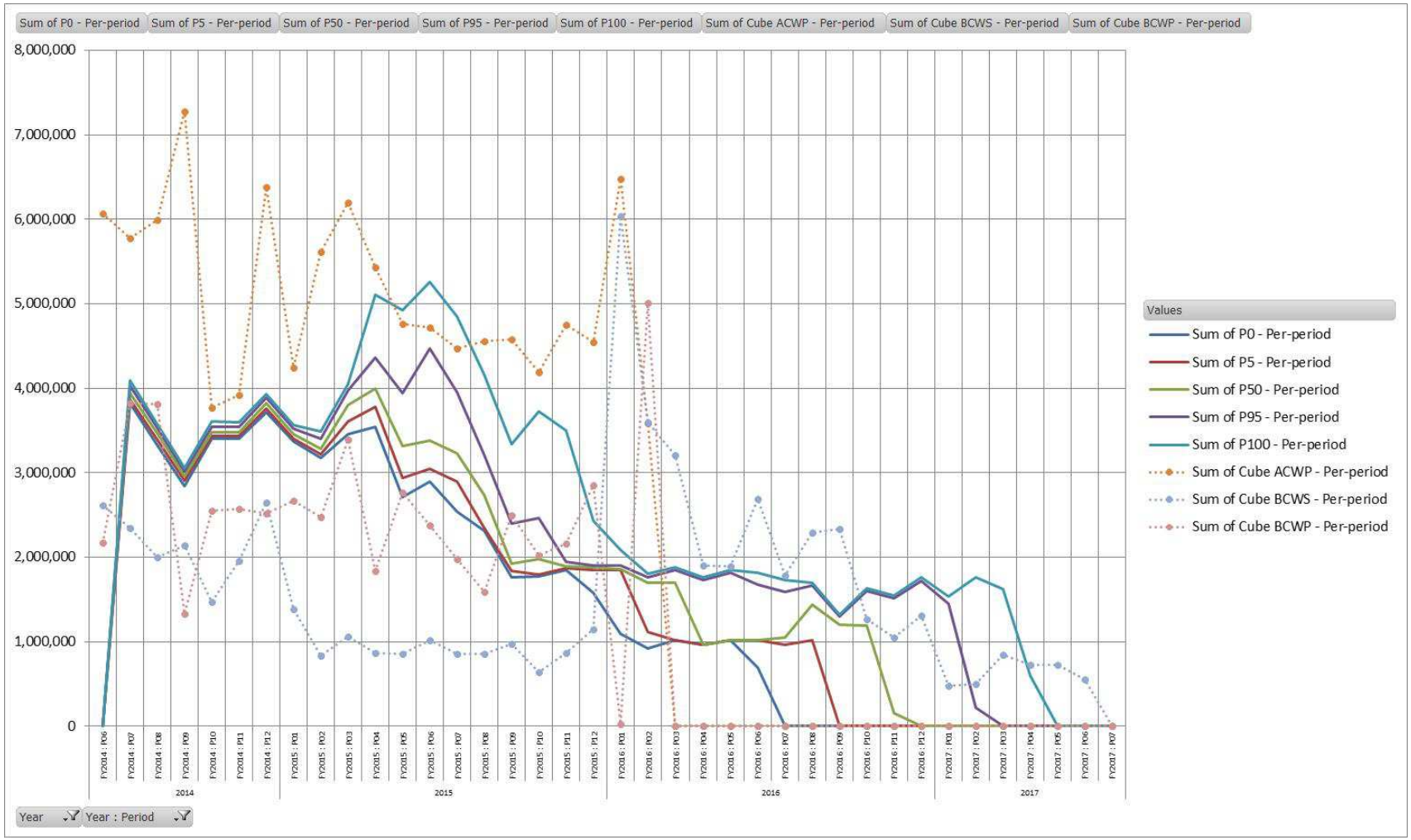
- Almost every aspect of the comparisons is 'wrong' in the sense that the model's ranges have turned out to be extremely unrealistic.

- But the comparisons are still useful as they can be used as a useful educational tool to help people to understand how probability relates to calibration.

Project 1 – Cumulative



Project 1 – Per period



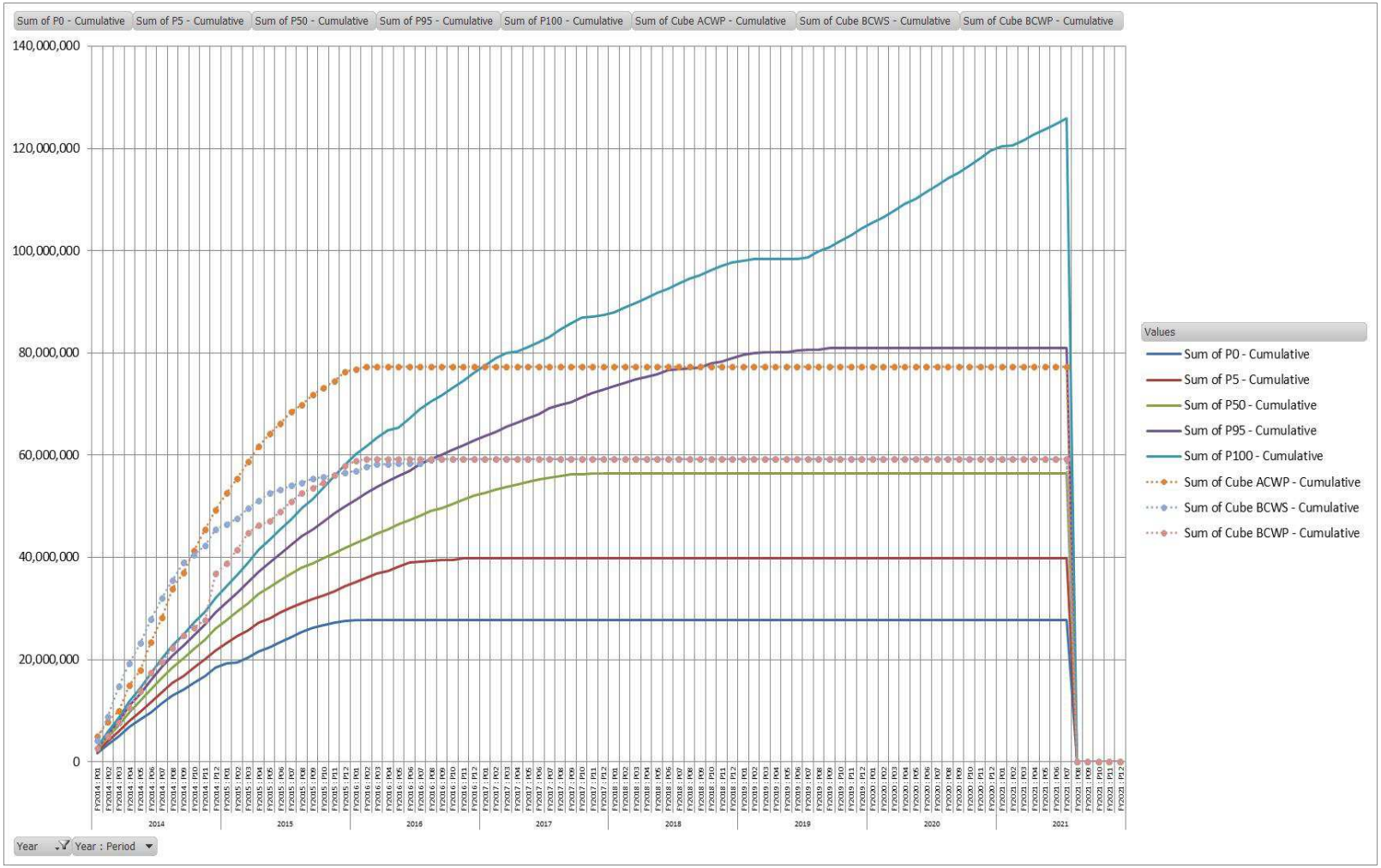
Project 2

- The next slides shows an example of a model that turned out to be:
  - Over-optimistic (too low) and over-confident (too narrow) with regard to cost.
  - Over-pessimistic (too high) and under-confident (too wide) with regard to duration.

- This wasn't surprising, since we used 'inner anchoring' methods for cost elicitation, and 'outer anchoring' methods for duration elicitation.

- The model and its related target P50 baseline values were rejected by the NDA at the time for being too pessimistic for *both* duration and cost – they were half right at least!

Project 2 – Cumulative



Project 2 – Per period

In-flight analyses

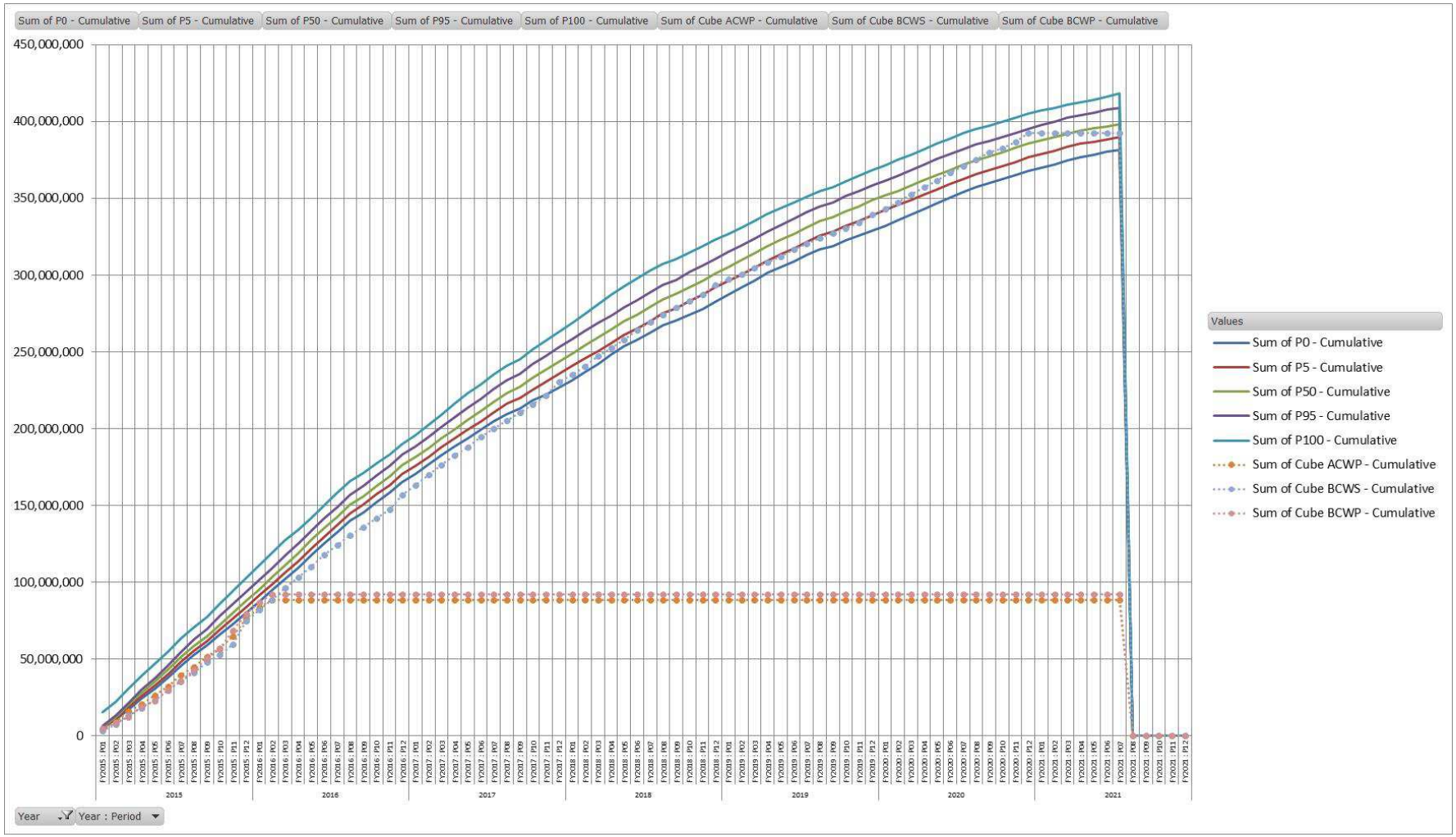
- The previous project's probabilistic time series was compared to its actual cost and performance values *after* it had been completed, but we want to make use of such comparisons for in-flight projects as well.

- This would give us early warning indicators of a model's possibly poor calibration levels within a few months.

- The project on the next slides is still in-flight.

- The model appears to be over pessimistic (too high) and over confident (too narrow) with regard to both cost and duration, but its general performance levels have improved between 2015 and 2016 (so far at least…).

Project 3 – Cumulative



Project 3 – Per period

What's next?

- We are developing elicitation and modelling techniques that will cover both the past and the future.

- This would allow us to run comparison analyses for historical data that would help us to calibrate a model's inputs –
  - If the model is well calibrated over its recent historical data then we should have more confidence in its forecasts.

- A Bayesian approach is probably the way forward, but our schedules can tend to be extremely complex so this might not be viable in the short term.

- Thank you for your time – questions please!

University of Strathclyde Glasgow