Lubrizol

being essential

Understanding and communicating uncertainty around Virtual Engine Tests

Calculating and Communicating Uncertainty Conference
27 January 2015

Andrew Rose, Lubrizol Ltd

Contents

- Introduction to Lubrizol and lubricant performance testing
- Virtual Testing
- Bayesian Variable Assessment approach
- Understanding and Communicating uncertainty

Introduction to Lubrizol and lubricant performance testing

Lubrizol

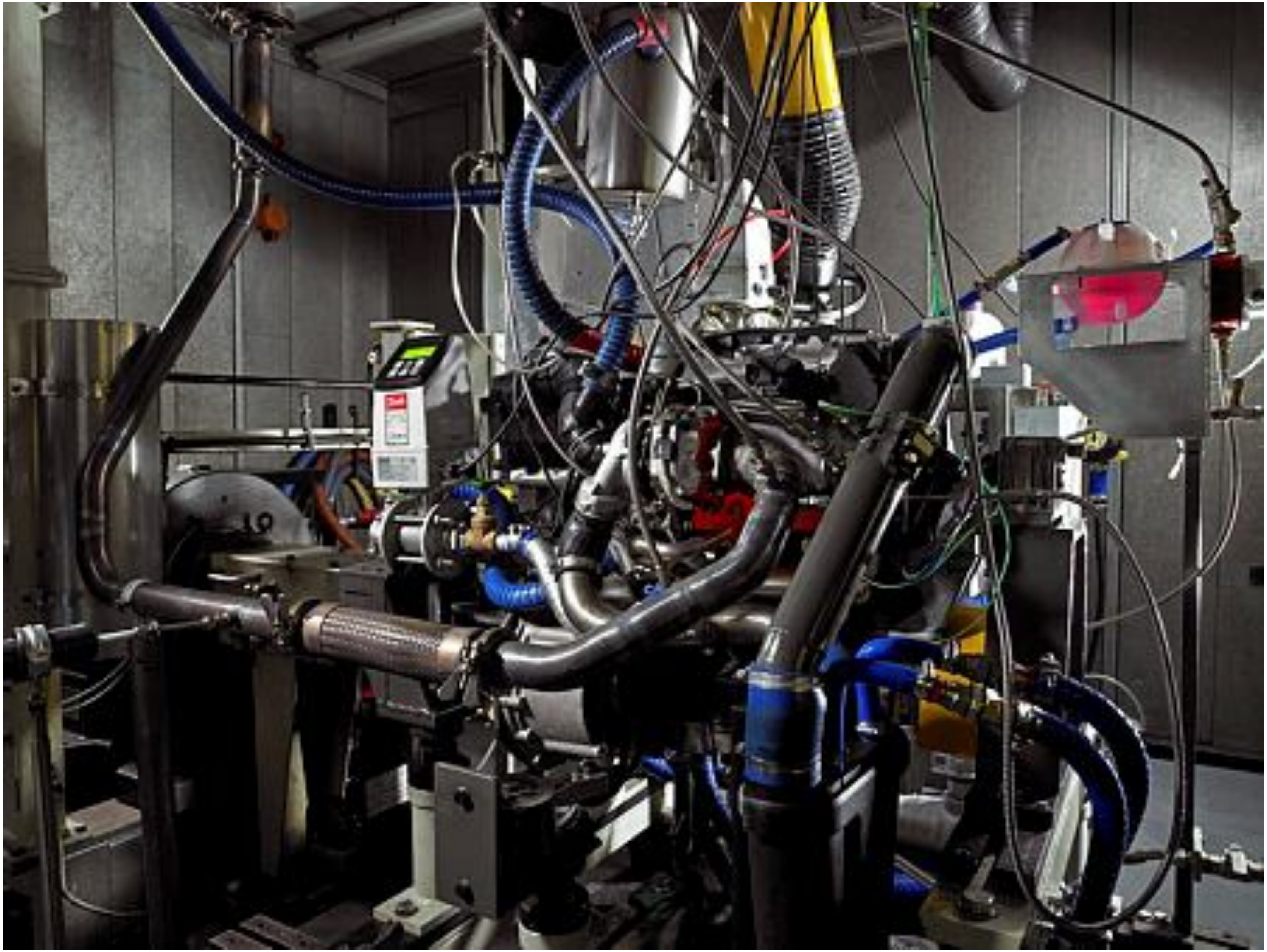
Who are Lubrizol?

- Not a household name… …but over 7500 employees worldwide, $6.4 billion revenues in 2013.
- Speciality chemical manufacturer:
  - Engine oil products
  - Driveline [gearbox, axle, transmission] products
  - Fuel products
  - Industrial Lubricants
  - Engineered polymers (plastics for specialised applications)
  - Personal + Home care (shampoo, hair gels, skin care products)
  - Coatings
  - And many more
- Our products are all around you in everyday life as part of other companies' products. In 1/3 of all cars in the world.
- Our aim: to be an essential ingredient in our customers' success by providing chemistry that solves problems and gives improved performance.

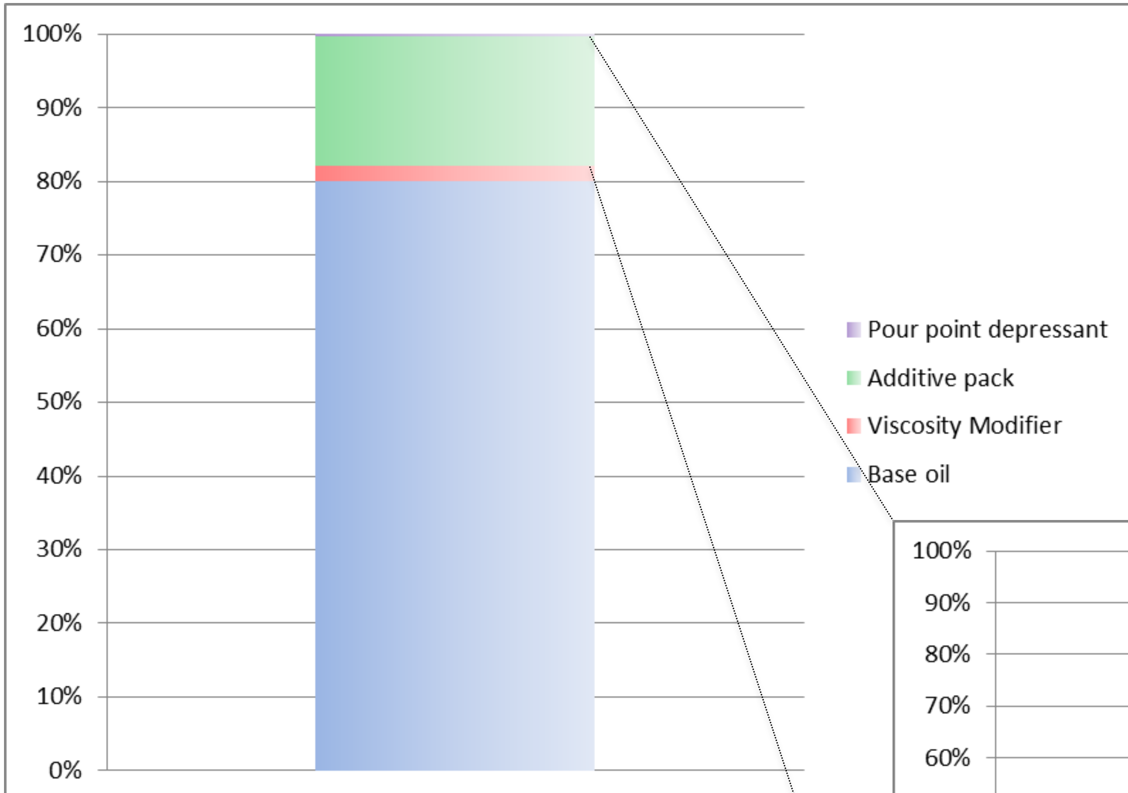
Lubrizol

Engine oil performance testing

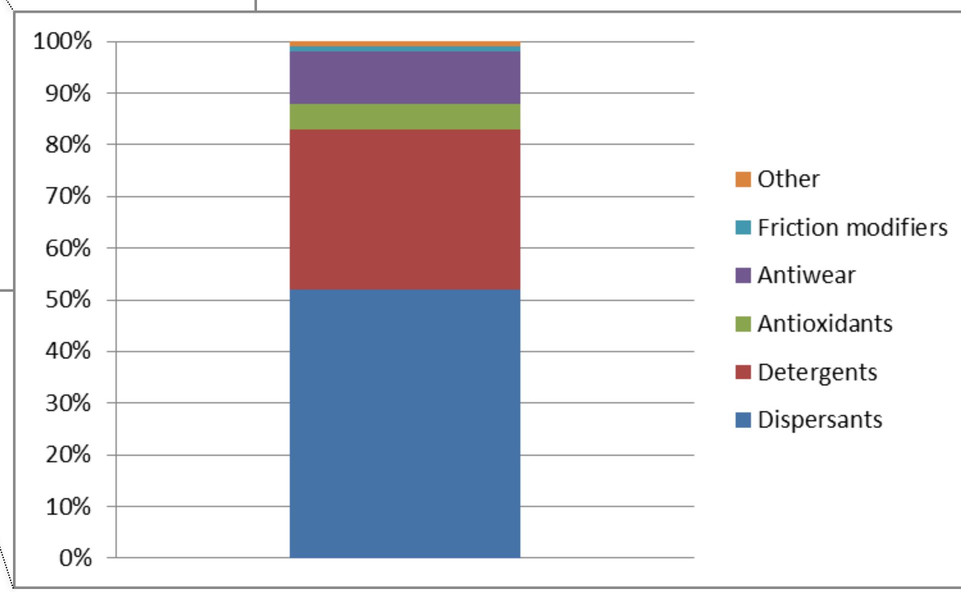
- Specifications set by industry organisation such as API (American Petroleum Institute) and ACEA (European Automobile Manufacturers' Association) set out tests that engine oils must pass in order to be approved for use in vehicles.
- Some manufacturers make additional requirements.
- Some tests are simple laboratory tests, e.g. viscosity, volatility and oxidation.
- Other tests involve running an oil in an actual engine on a test bed for hundreds of hours, and measuring performance properties, e.g. deposits, oil consumption, fuel economy, wear of various engine components, sludge formation.
- These ensure that oils will not cause problems to occur in vehicles out on the road.
- It can cost over £100,000 to run a single engine test.



What's in an engine oil?



- Every component of an engine oil is there to help to pass the performance tests.
- Sometimes a component will help in one test, but be detrimental in another.
- Not easy to balance the requirements of all tests.



We refer to the combination of chemical components within the engine oil as a 'formulation'.

Virtual Testing

The Q.LIFE® Engine

- Lubrizol's trademarked data mining and prediction system
- Statistical Sciences mines the data stored in the corporate databases and creates predictive models
- The models and meta data are stored in a model database, which is used to create end-user tools that facilitate the formulating process
- Ongoing effort for many years. Average over 100 users per month, and approx. ~40,000 formulations were predicted via Q.LIFE® in 2014.
- Continuous improvement required (efficiency of the modeling process, and quality of the predictive models and tools).
- Used to guide formulating, optimise products, as a pre-screen before testing and even to demonstrate performance for some customers where actual tests are not required.

The Q.LIFE® Engine

- Predictive models are available to internal users via browser-based and Excel tools.
- If predictions are used with external customers, the Statistical Sciences group create a 'Virtual Test Report' containing predicted results for a given formulation in all relevant tests.
- As part of the Virtual Test process, an extrapolation report is produced and checked, to ensure that the levels of all regressor variables for the new formulation are within (or acceptably close to) the ranges used in the dataset on which the model was built.

Bayesian Variable Assessment and Modelling Strategy

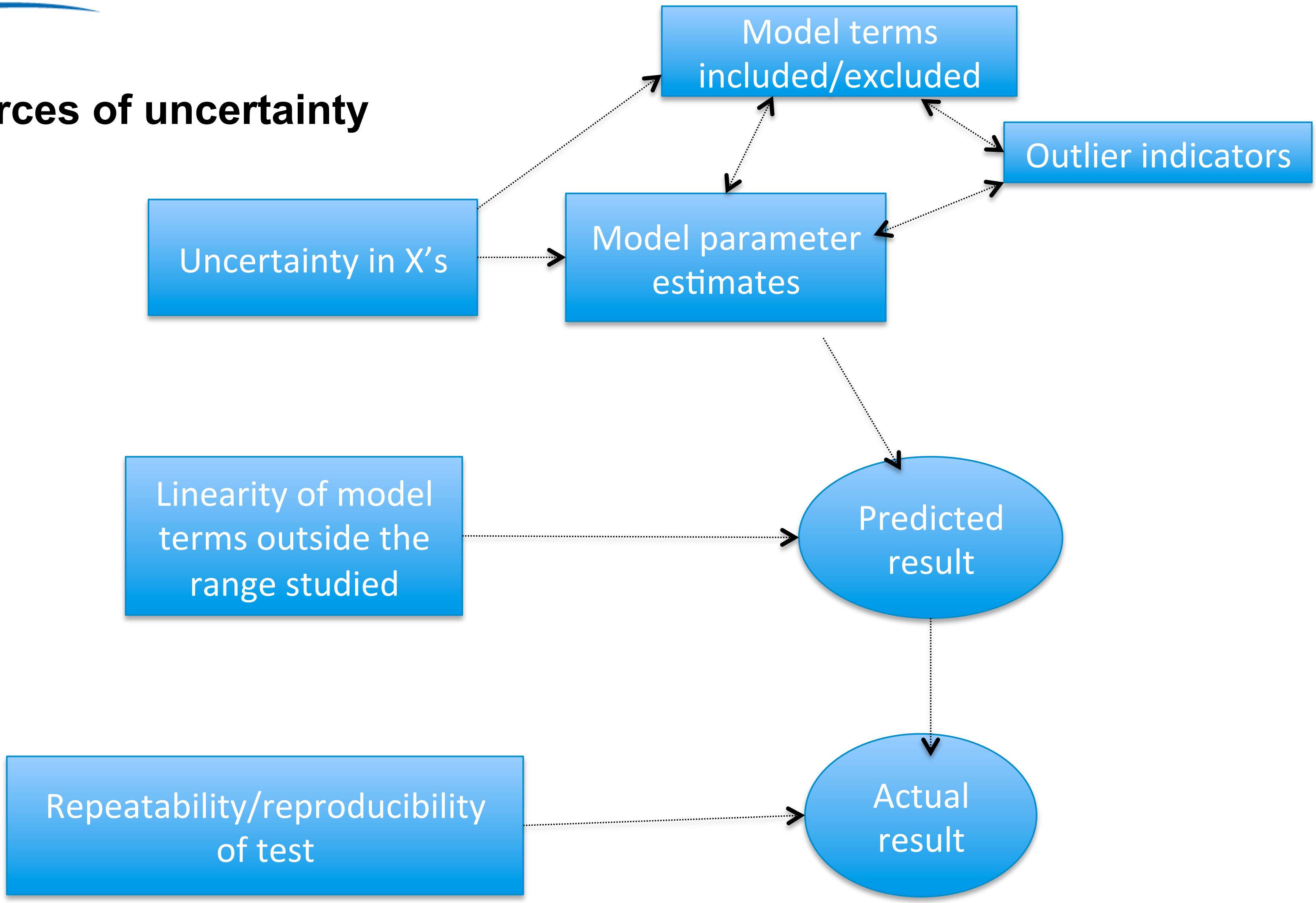
Lubrizol

Bayesian Variable Assessment

- Bayesian model-averaging using mixture priors developed at Lubrizol.
- Assume linear models and effect sparcity. Generally, these are reasonable assumptions in this application.
- Normal-inverse gamma priors with a ‘spike’ at zero for each regressor term.
- All observations have a fixed prior probability of being an outlier.
- Posterior distributions are estimated using a Gibbs sampler.
- Effect ‘activity probability’ is calculated by summing the posterior probability of all models containing that effect.
- Typically, only effects with activity probability >0.35 are included in the final model. We will run several iterations of removing unimportant model terms and re-running.
- Posterior means of model coefficients are used to create predictive model.

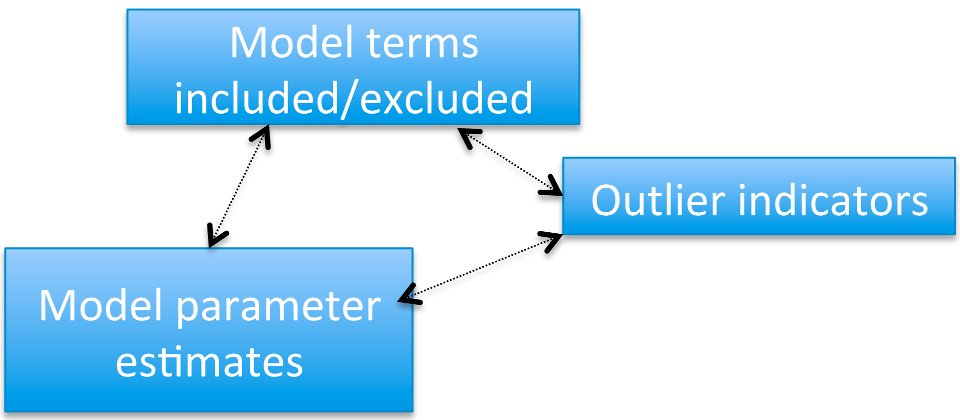
Understanding and communicating uncertainty

Sources of uncertainty



Lubrizol

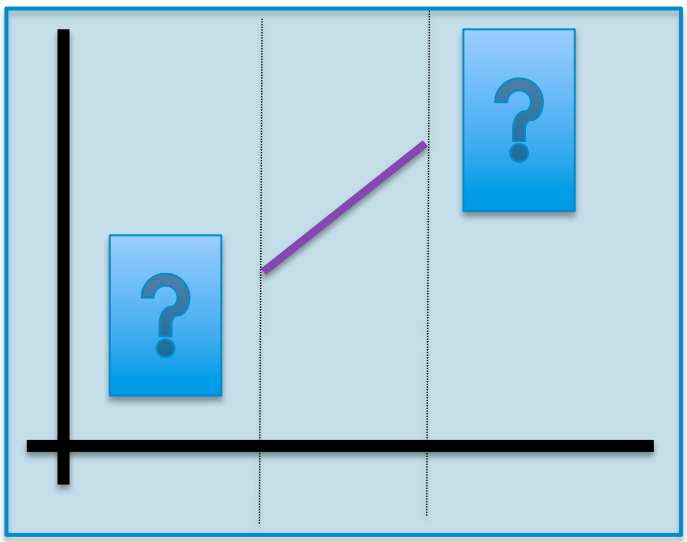
Sources of uncertainty



- The joint posterior distribution of these, as captured in the Gibbs sample chain, summarises the uncertainty from these sources nicely.
- The level of uncertainty varies depending what combination of X's you are predicting for.
- One approach is predicting for all iterations of the chain, and finding the 2.5 and 97.5 percentiles to express uncertainty… …is there a more efficient way?

Sources of uncertainty

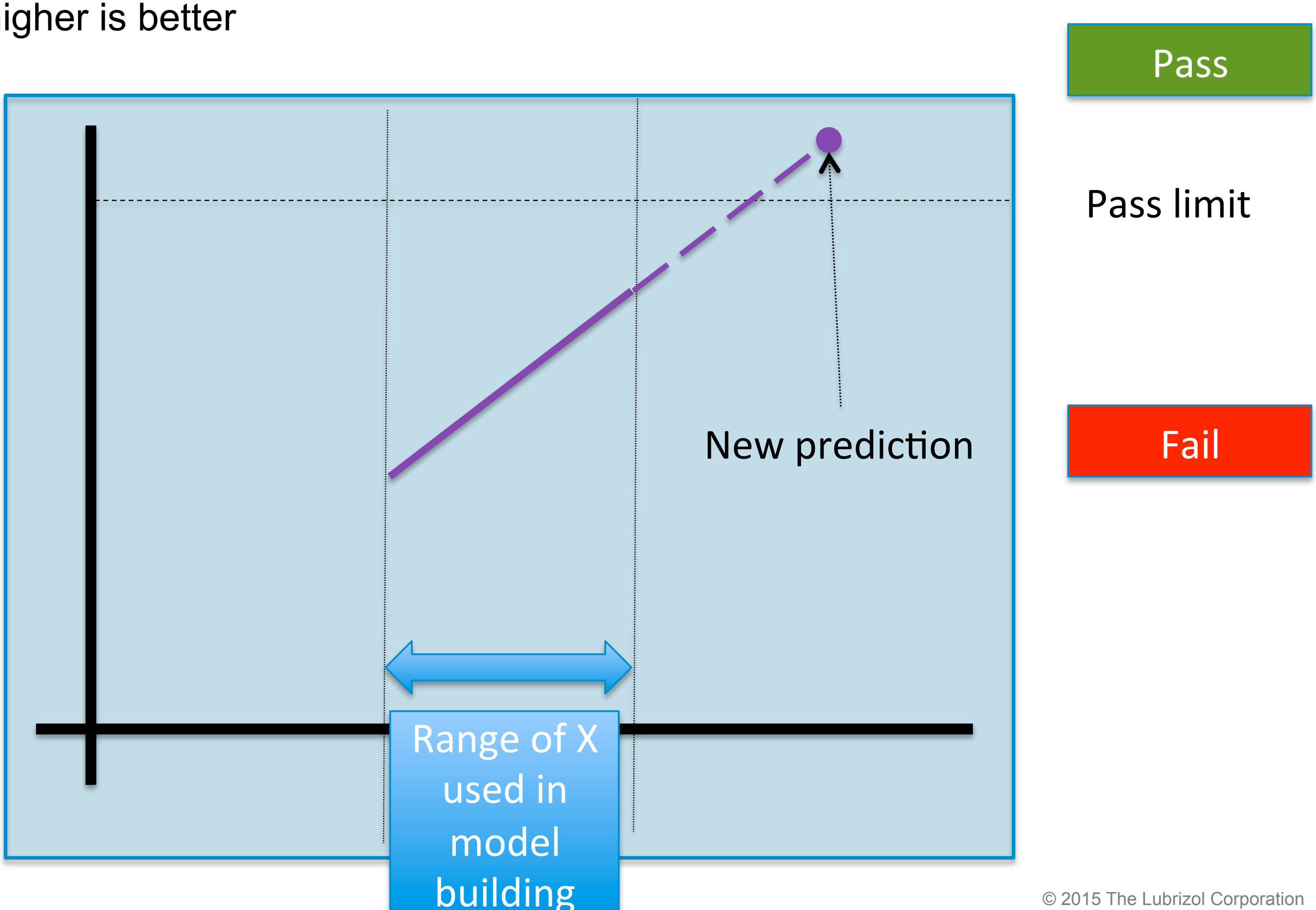
Linearity of model terms outside the range studied



- The effects of some components show diminishing returns over a certain level, or make things much worse if below a certain level.
- We have some intuition in this area, but often limited data.
- Generally, we expect the effect of our model inputs to be monotonic.
- Could average predictions over a range of scenarios if suitable priors could be set up.
- Our current approach is to flag up unacceptable extrapolation according to a series of rules, which guard especially against predicting a pass when a failing result could occur if the extrapolated term was non-linear.

Sources of uncertainty – linearity outside ranges used in model building

- Assume higher is better

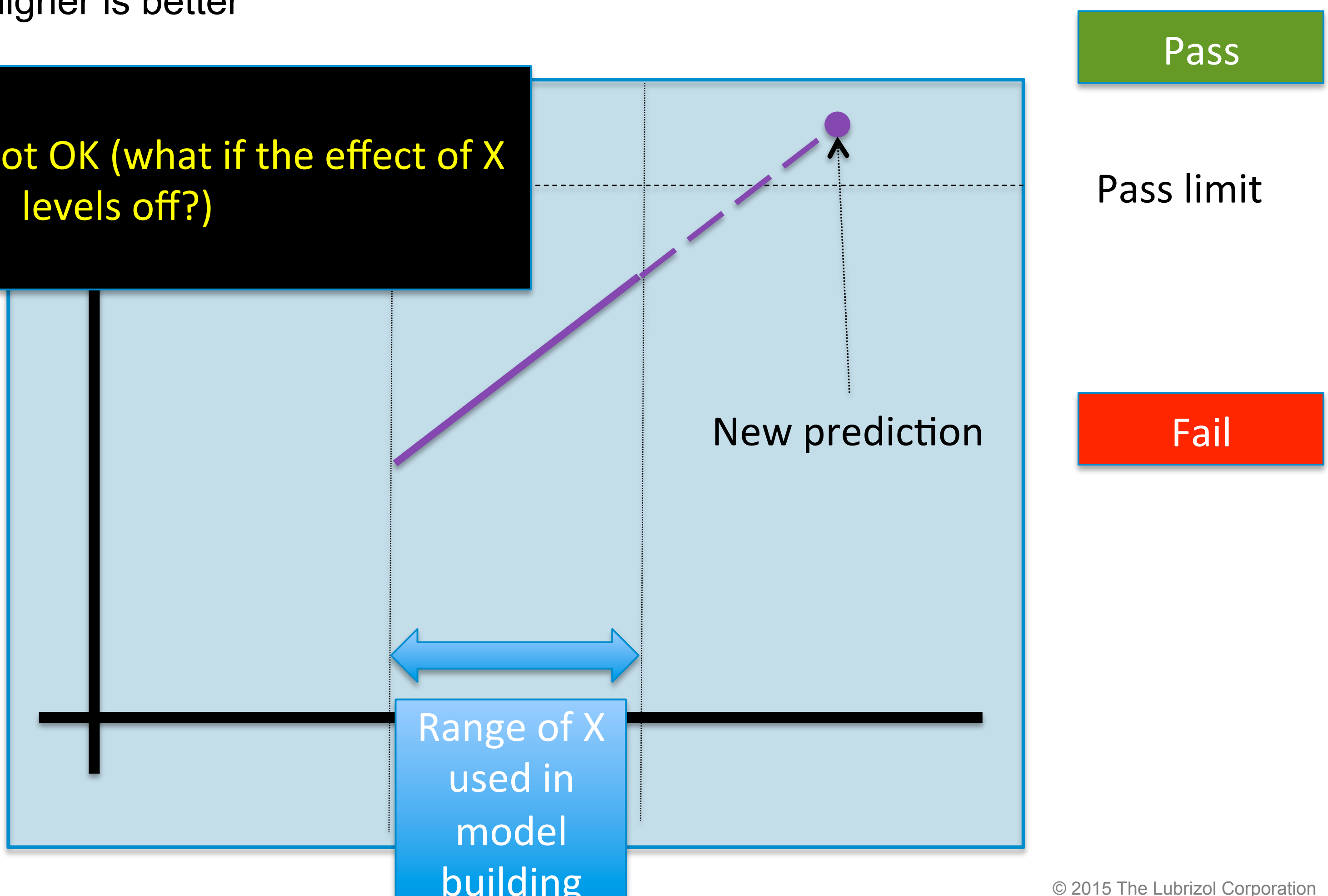


Lubrizol

Sources of uncertainty – linearity outside ranges used in model building

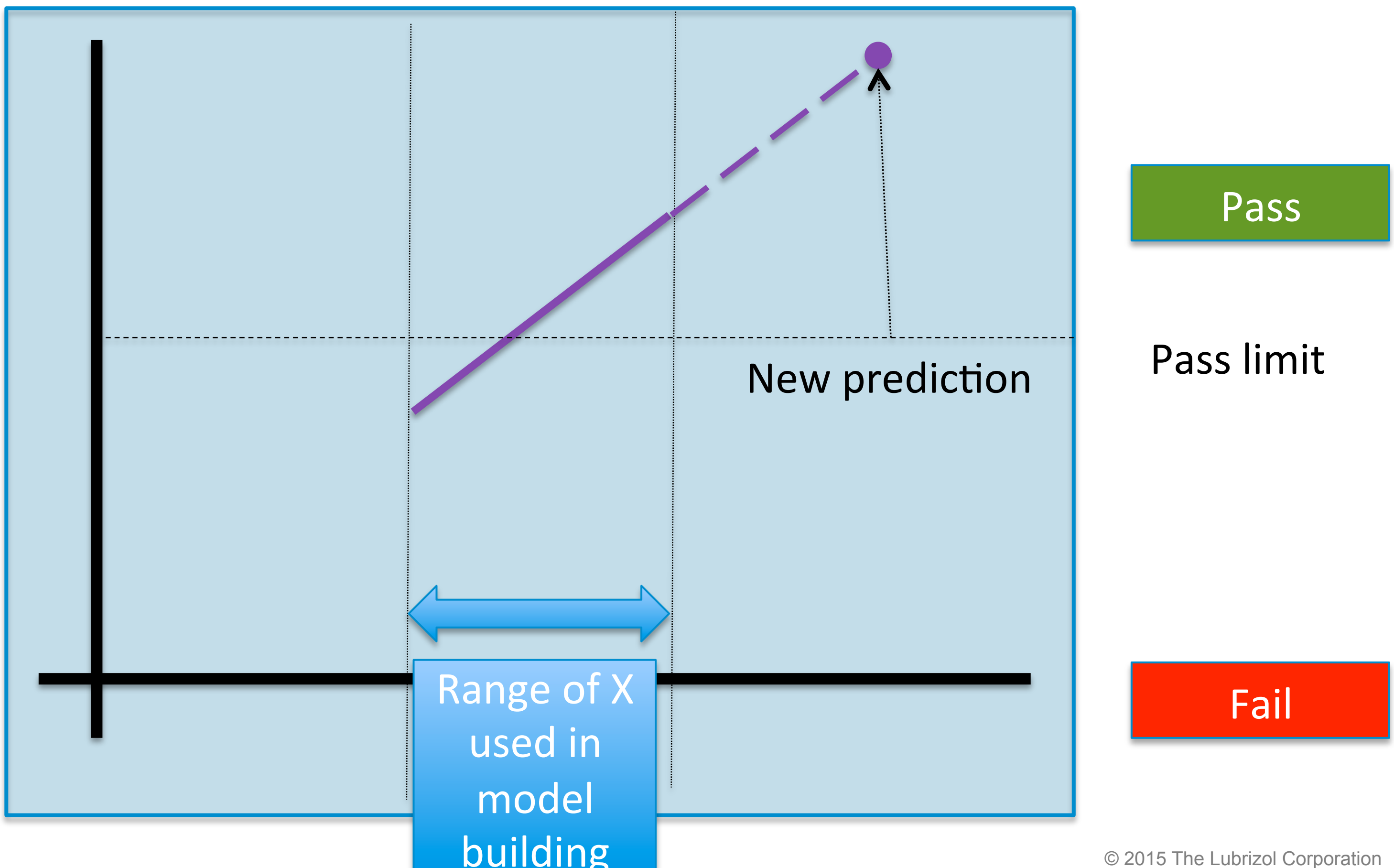
- Assume higher is better

Extrapolation not OK (what if the effect of X levels off?)



Sources of uncertainty – linearity outside ranges used in model building

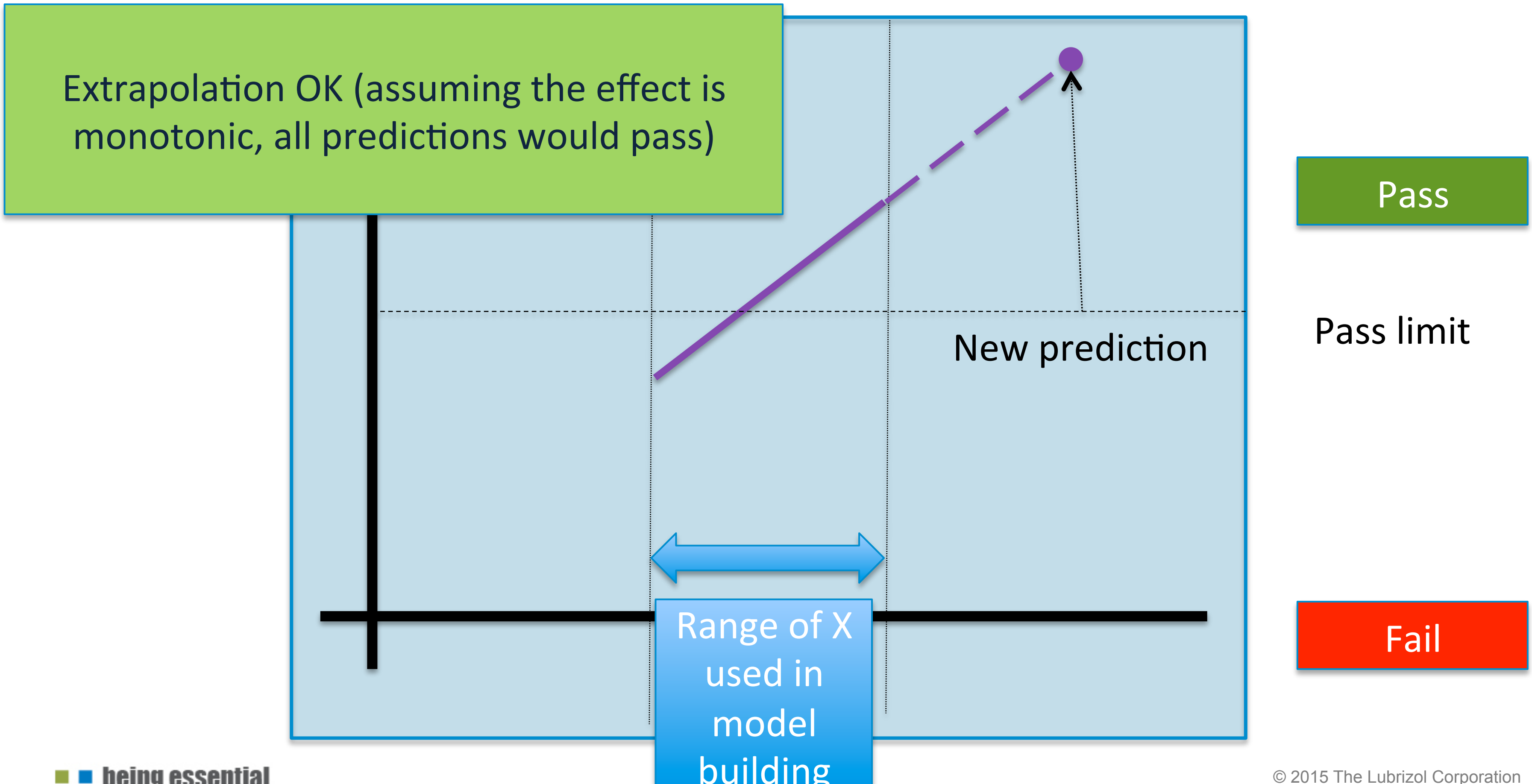
- Assume higher is better



Lubrizol

Sources of uncertainty – linearity outside ranges used in model building

- Assume higher is better



Sources of uncertainty

Uncertainty in X's

- The inputs to our model contain information about the chemistry components, based on either knowledge of its structure/composition (which may be uncertain), or the results of simple analytical tests (which have associated experimental error). We do not currently calculate this source of uncertainty.

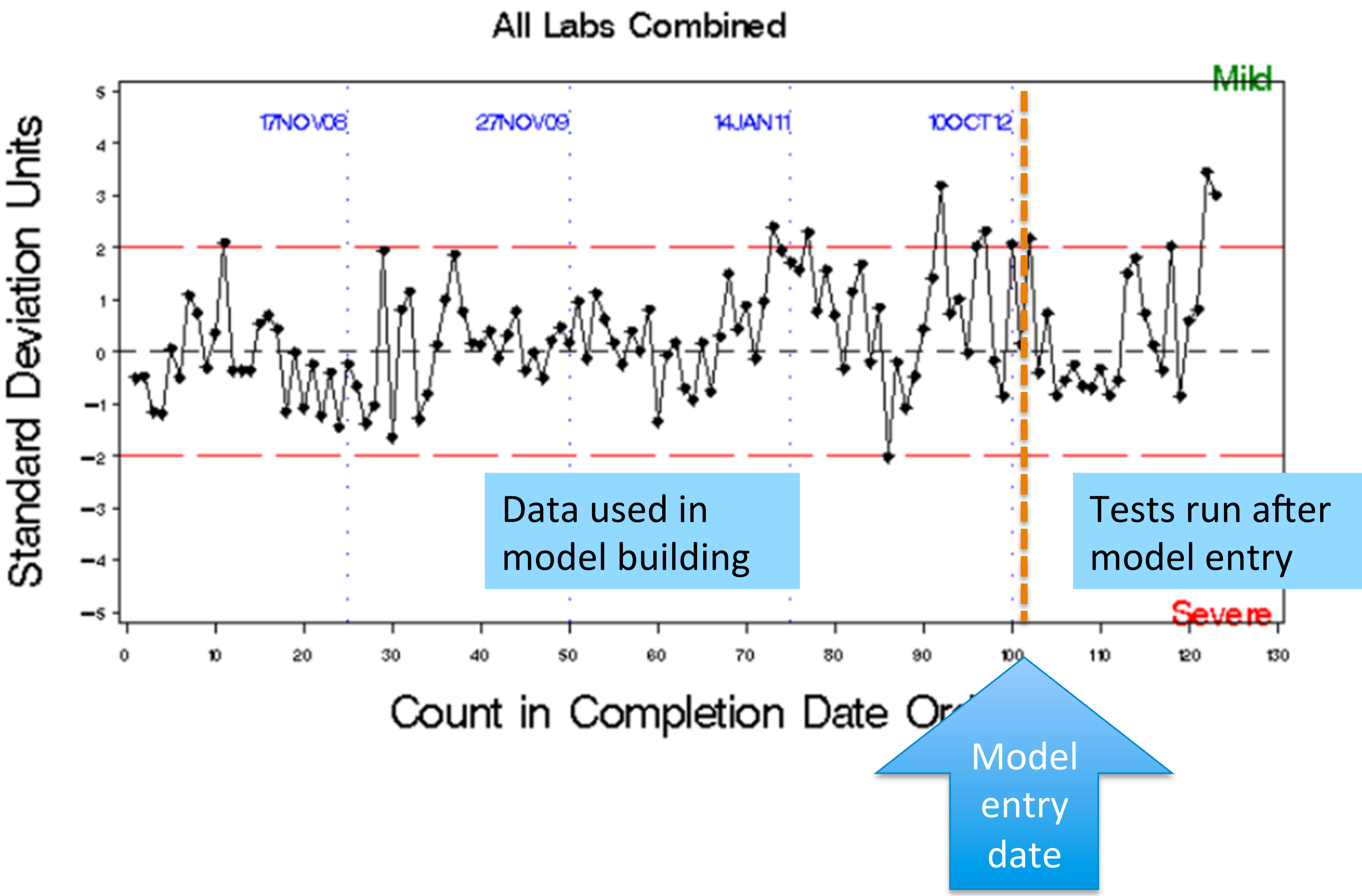
Repeatability/reproducibility of test

- Estimates of this often exist for industry standard tests, calculated from inter-laboratory studies. Otherwise, we are able to estimate this from reference data.
- We use this in evaluating the difference in effective pass limit when multiple attempts are allowed at passing a single test. Virtual test results can be adjusted accordingly.

All in together – labmon (laboratory monitor)

- There are many sources of uncertainty surrounding our models, and it is difficult to quantify them all. Despite this, our models are still useful, and can often be more reliable at predicting mean performance than a single real engine test result.
- The final proof of performance, and quantification of uncertainty, lies in predictions made by the model for new tests, on new formulations, carried out after the model building period.
- This is visualised via our ‘labmon’ tool.
  - Shows the standardised residuals (actual result – predicted result)/model SD
  - Can be used to track severity shifts at individual labs, and to assess model performance.

Lubrizol


All Labs Combined
Mild
Severe
17NOV08
27NOV09
14JAN11
10OCT12
Standard Deviation Units
Count in Completion Date Order
Data used in model building
Tests run after model entry
Model entry date

Challenges

- How can we best refine the extrapolation rules to accurately reflect uncertainty?
- We state that model predictions can sometimes be a better reflection of an oil's true performance than a single engine test result. Can we work out when this is the case?
- How do we communicate the model performance, from labmon, to the end user, in order to give the right level of confidence in the result?

Summary

- Virtual testing can be useful as tool to complement running expensive engine tests, in product development, optimisation and demonstration of performance.
- Bayesian model averaging is used to generate predictive models. The posterior distribution from this has the potential to help quantify uncertainty.
- Extrapolation rules are used to guard against the most damaging effects of model uncertainty.
- End users are typically only presented with point estimate predictions. An expression of uncertainty might be helpful to some users.

References

- 1. Meyer, R.D. and Wilkinson, R.G., “Bayesian Variable Assessment,” Communications in Statistics – Theory and Methods, 27(11), 2675-2705 (1998).

- 2. Scinto, P.R., “The Virtual Engine Test,” SAE Paper No. 2001-01-1905, May 2001.

Lubrizol

Bayesian Variable Assessment

- Assume that there are k predictor variables, all predictor variables were scaled to have std. 1 and mean 0.
- Let M_i be the label for the i^{th} subset of predictor variables, k_i be the number of predictor variables in this subset. Assume that every predictor variable has some probability π that it is active, i.e., that it is contained in the correct model.
  - For typical problems, we often choose π to be 0.25.
  - Predictor variables are active independently
- We assume the response variable $Y \sim \mathrm{Normal}(\sum X_i \beta_i , \sigma^2 I)$.
- The prior for the intercept and error standard deviation is noninformative, i.e., $p(\beta_0 ,\sigma) \sim 1/\sigma$.
- $\beta_i \sim \mathrm{Normal}(0,\gamma^2 \sigma^2 / k_i)$ where $\gamma^2 \sim$ inverse gamma $(\lambda/2,\nu/2)$, i.e., $p_{\gamma^2 \lambda,\nu} \propto 1/\gamma^{\lambda+2} \exp(-\nu/2\gamma^2)$, we chose the default values of λ *and* ν *to be 2*.
- Every observation has a prior probability of being an outlier (default is 0.05). The outliers are down weighted when updating the regression coefficients and standard deviation.