Arthur D Little

Data Driven Risk Modelling
A Pragmatic Approach

Marcus Beard and
George Simpson

October 2017

This paper outlines a pragmatic data driven risk modelling approach and considers application to the rail sector

Background

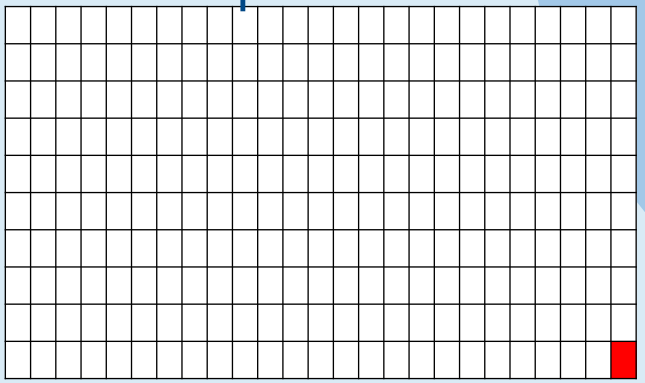
- A utilities sector company needed to prioritise limited safety assurance resource in inspecting large volumes of safety critical installation work
- The work is completed by lone workers over a large geographical area and assurance resources are not sufficient to inspect all work
- These workers can sometimes leave unsafe situations in their work that can lead to low probability, high severity accidents with significant reputational risk

Risk model

- We modelled the risk posed by each individual worker using a composite function of multiple parameters, such as driving behaviour and productivity
- The model outperformed the client's existing risk model in testing and is now being operationalised

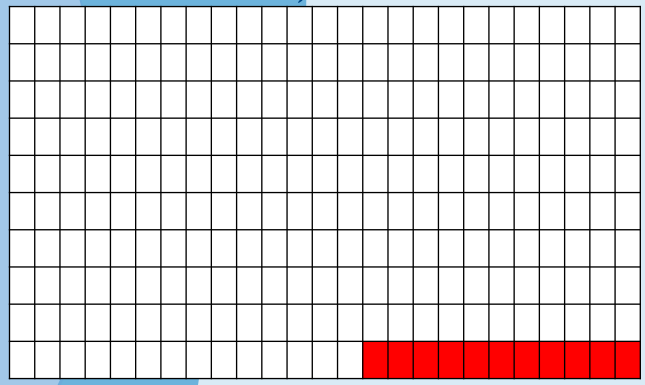
Following a 'deep dive' audit the company found that their actual defect rate was **11 times higher than expected**

~3 million assets, standard assurance process defect rate 0.4%



Approx. 12,000 high risk defects

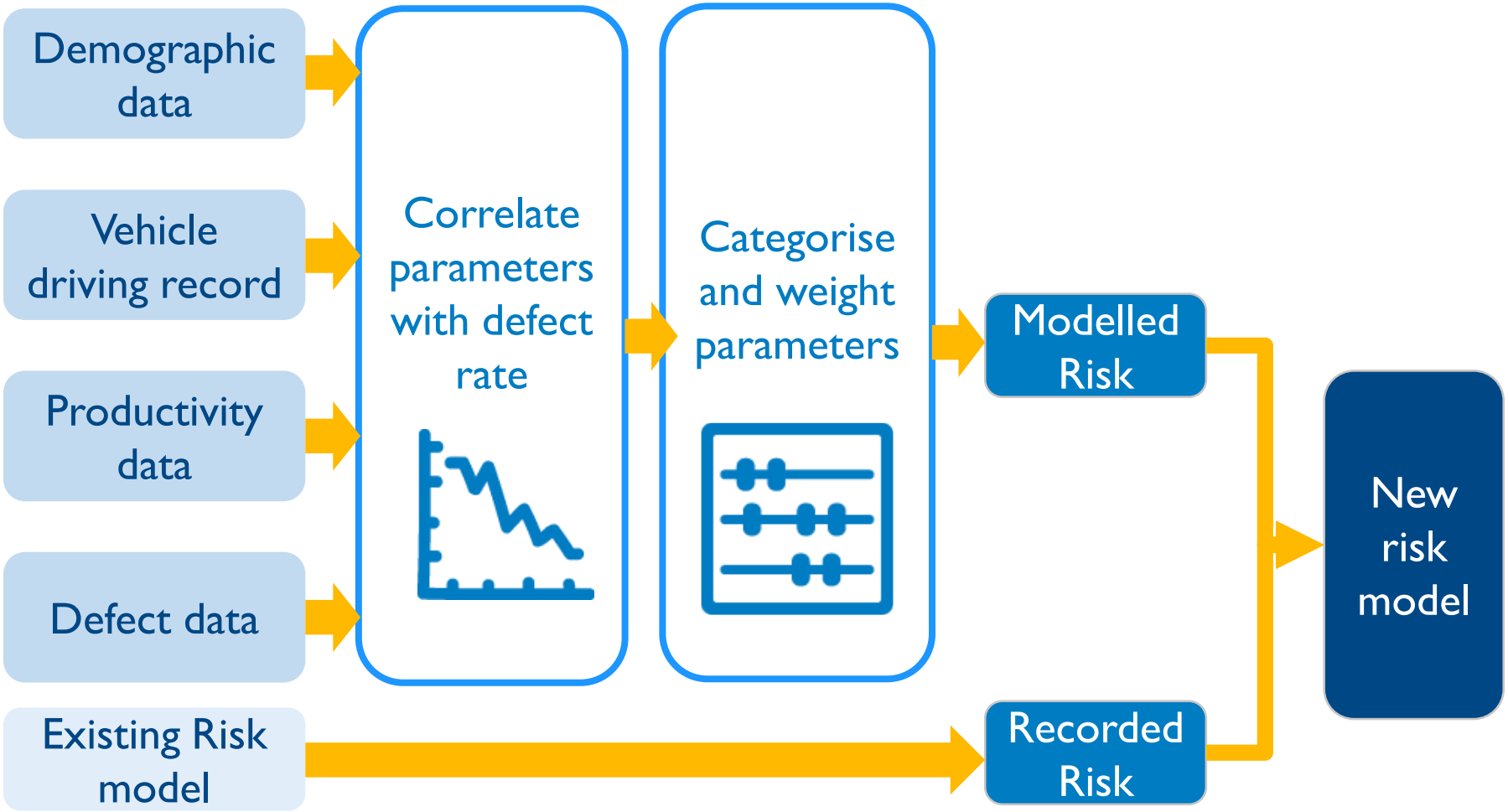
Representative audit sample of ~1650 assets, defect rate 4.4%



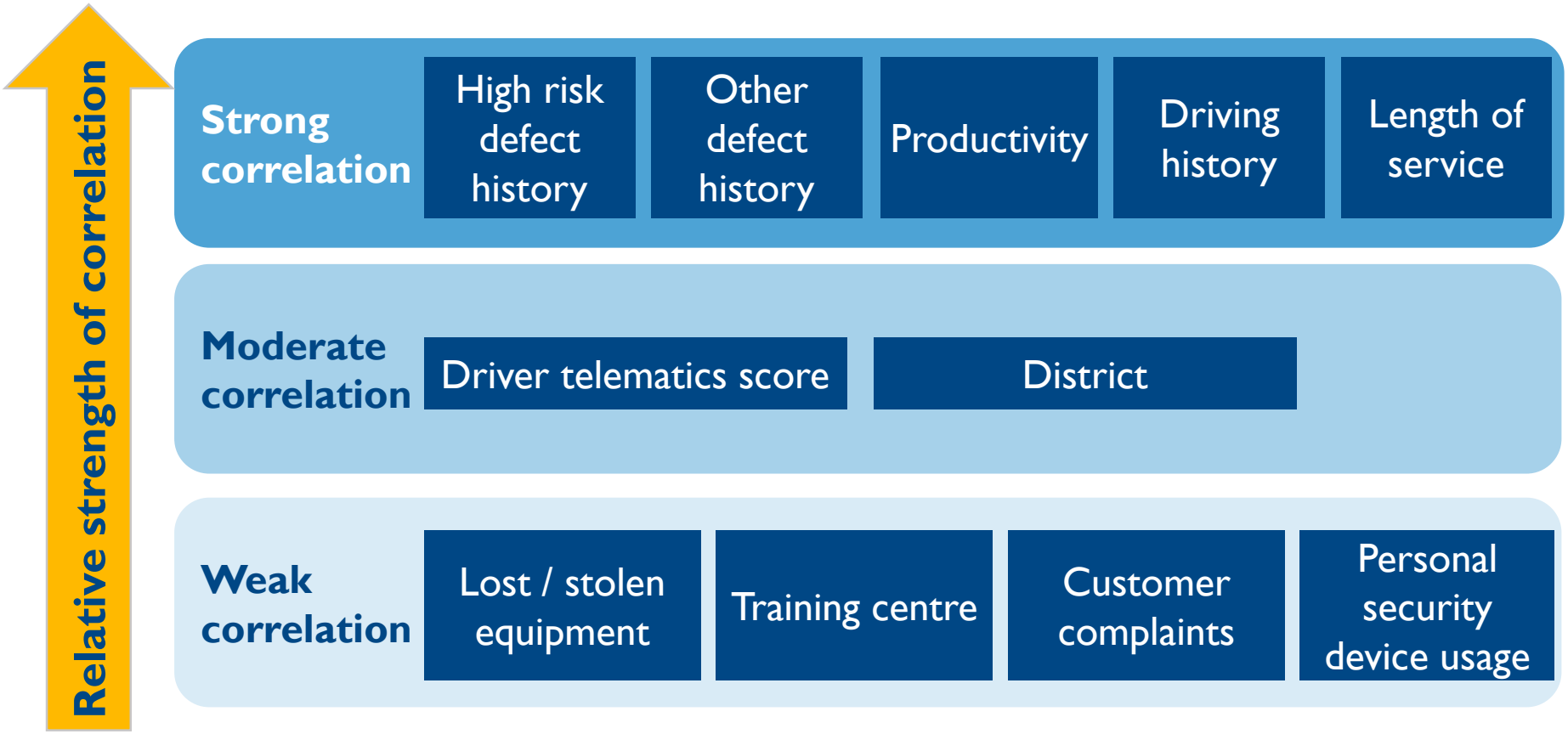
132,000 high risk defects implied

The existing risk score did not differentiate fitters with a propensity to leave high risk defects. So where to focus?

New model: correlating fitters who had previously left high-risk defects with other observable parameters

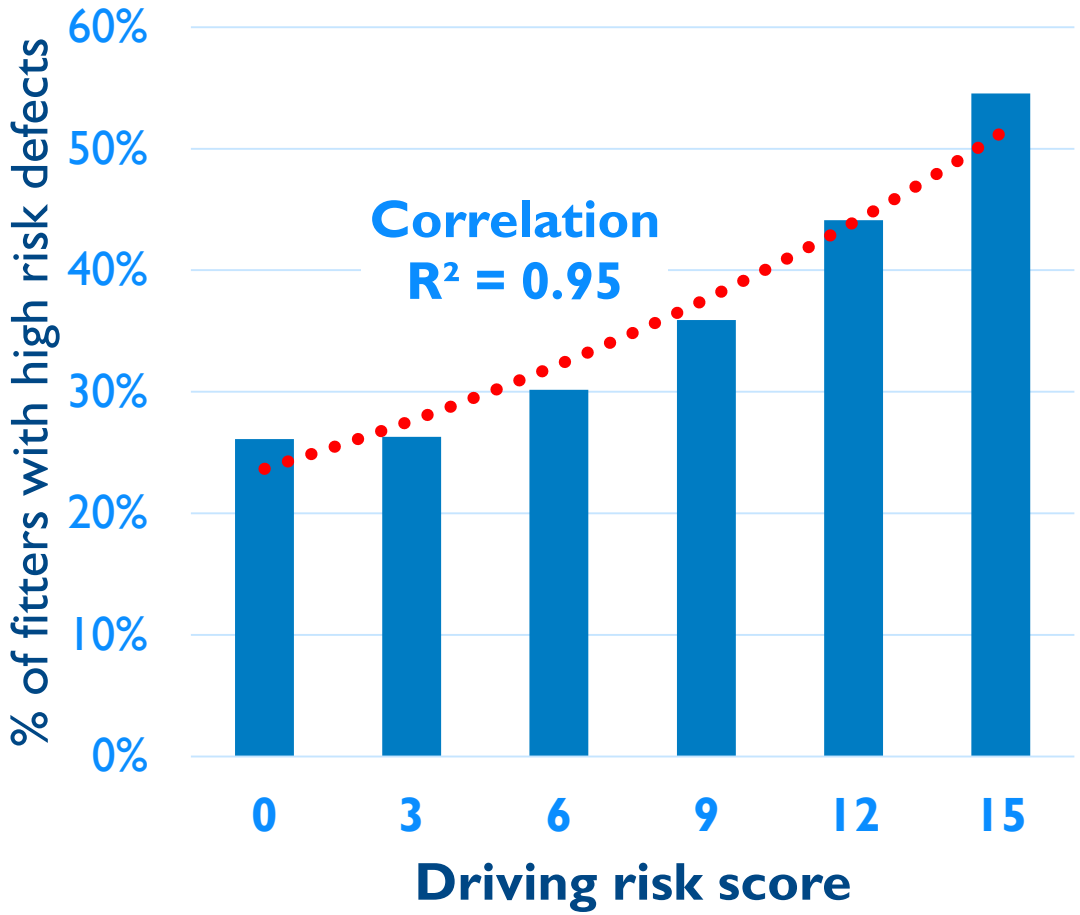


High-risk defects were associated with fitters with high driving risk, high productivity and history of other defects



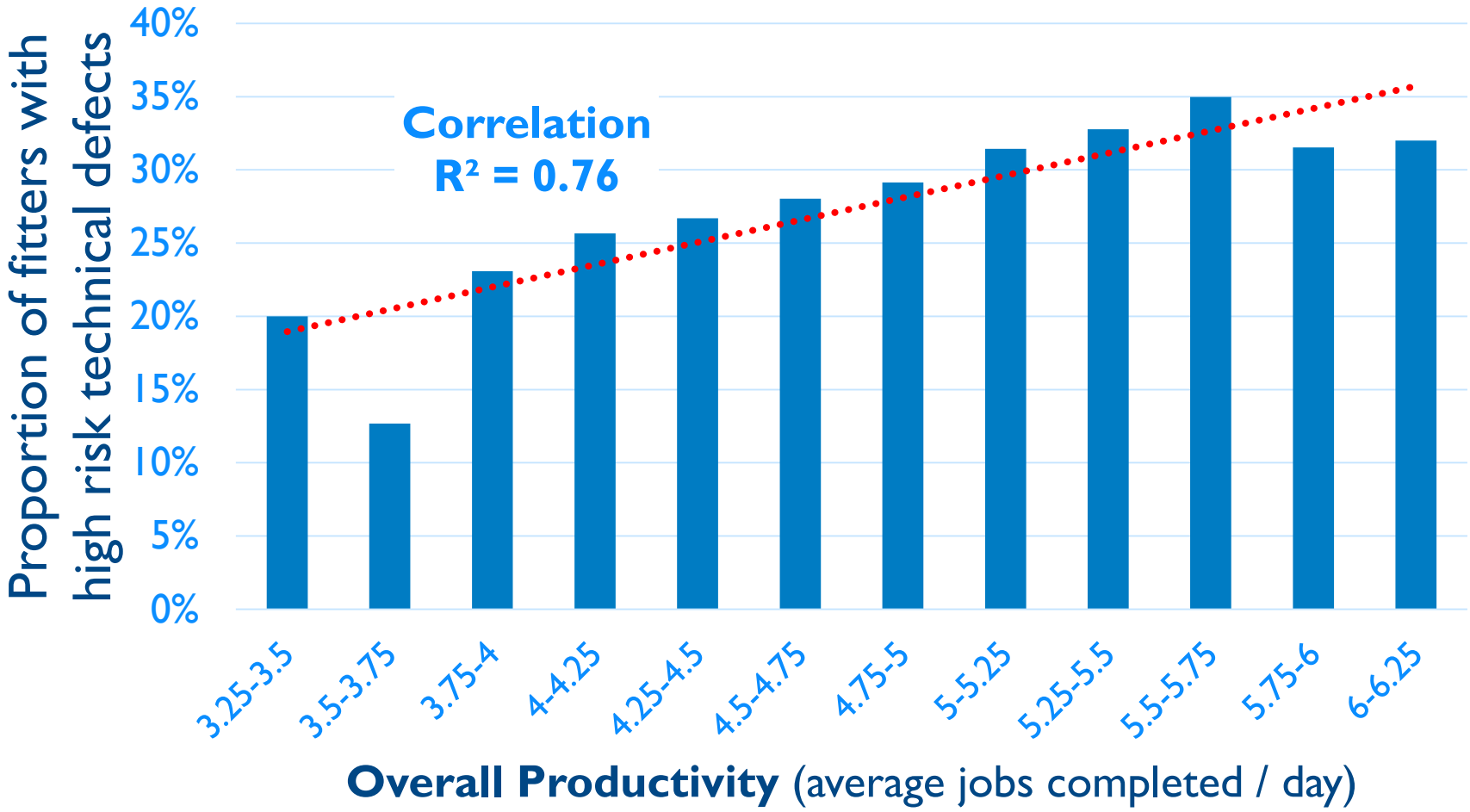
Parameters with a strong or moderate correlation were used to derive the Modelled Risk

EXAMPLE. **Driving risk** correlates strongly with fitters who have left high risk technical defects in the past

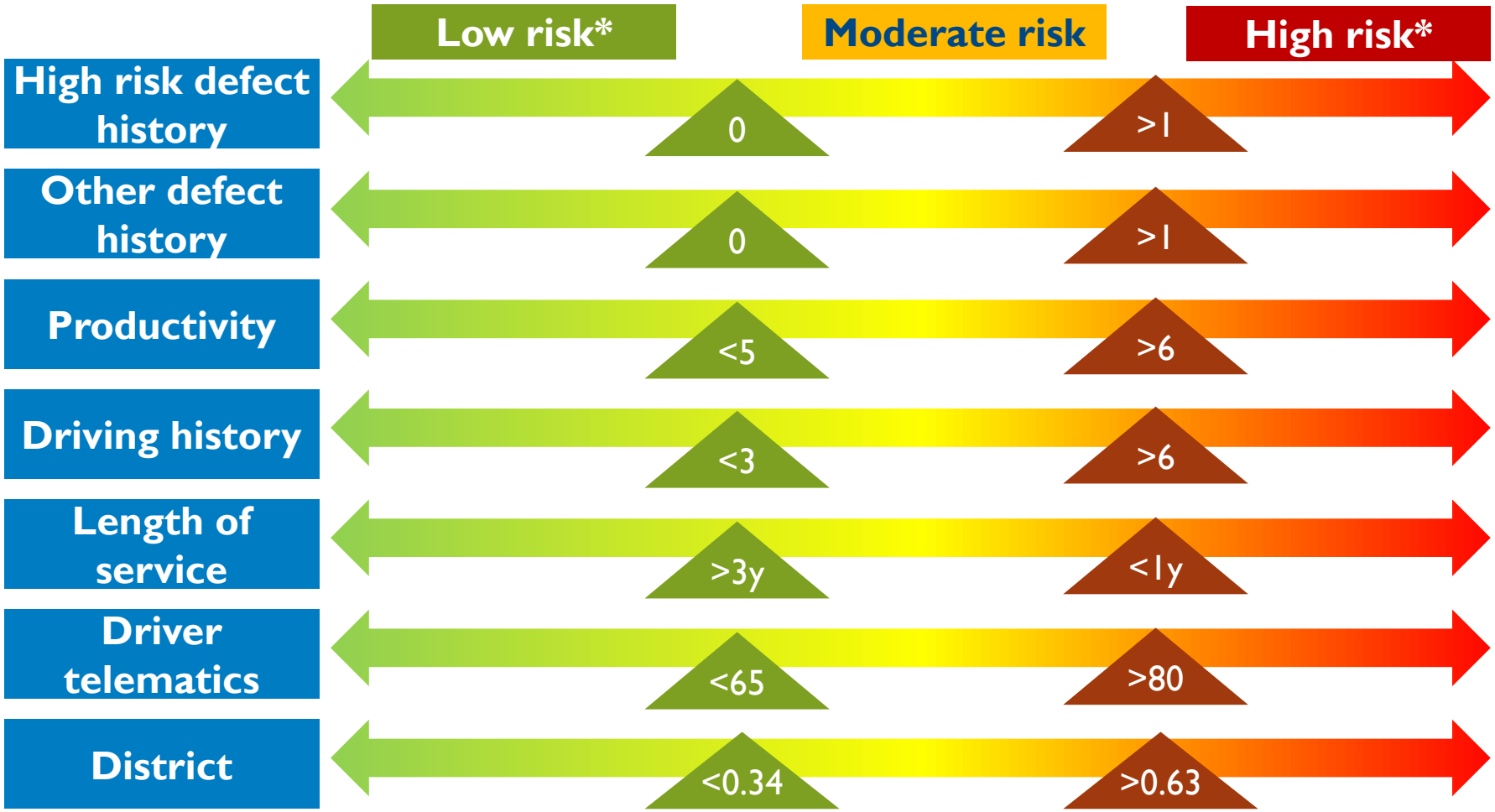


- Driving risk score =
  - Each licence endorsement is 3 points
  - Each vehicle incident recorded (e.g. damaged wing mirror) is 3 points
  - Vehicle write offs are 6 points
- Includes non-blameworthy incidents and remains on an employee's record for 24 months

EXAMPLE. Fitters with **higher overall productivity** are more likely to have left high risk technical defects

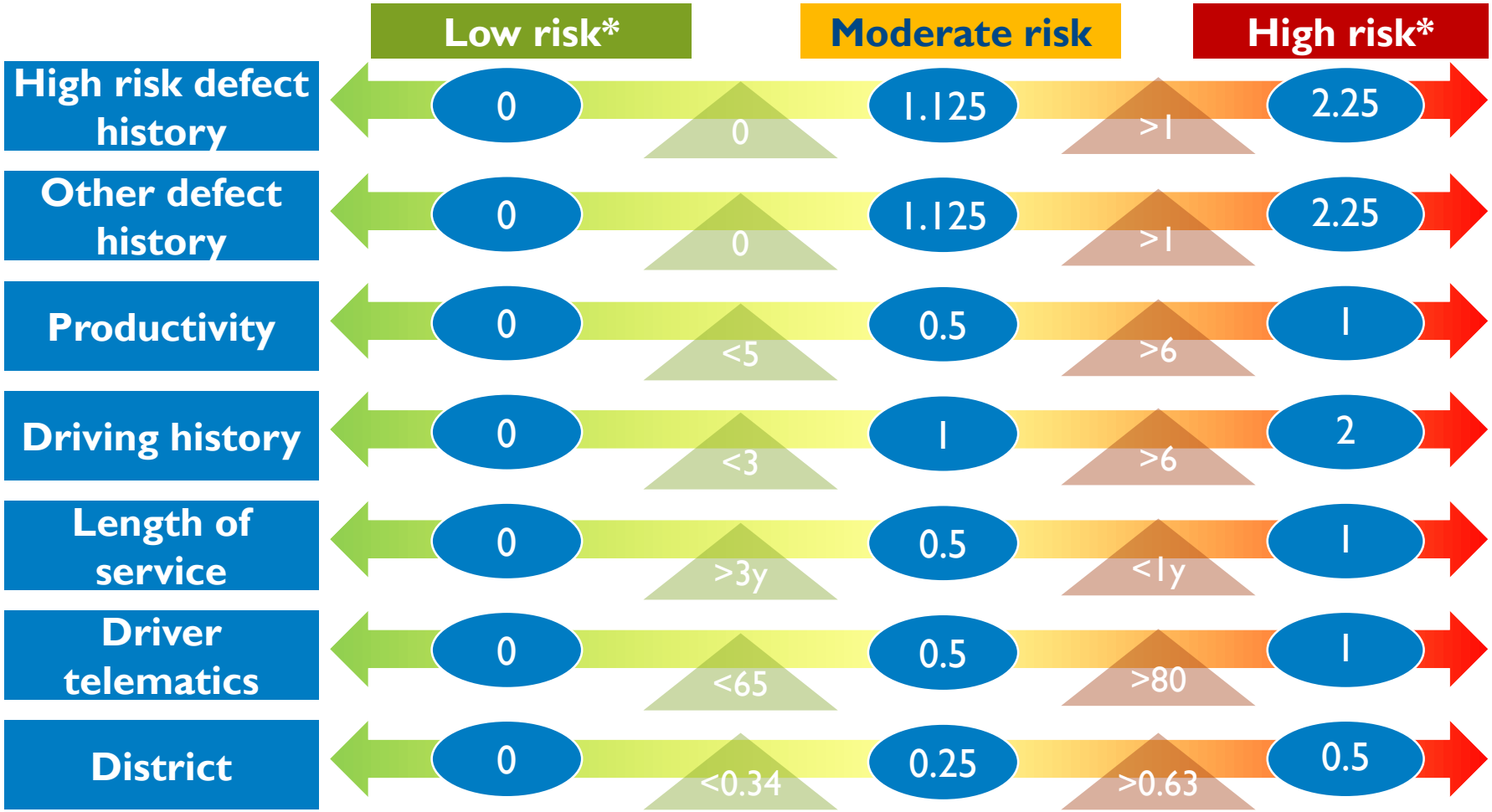


We defined risk boundaries so that approx. 20% of fitters were assigned to each high risk category…



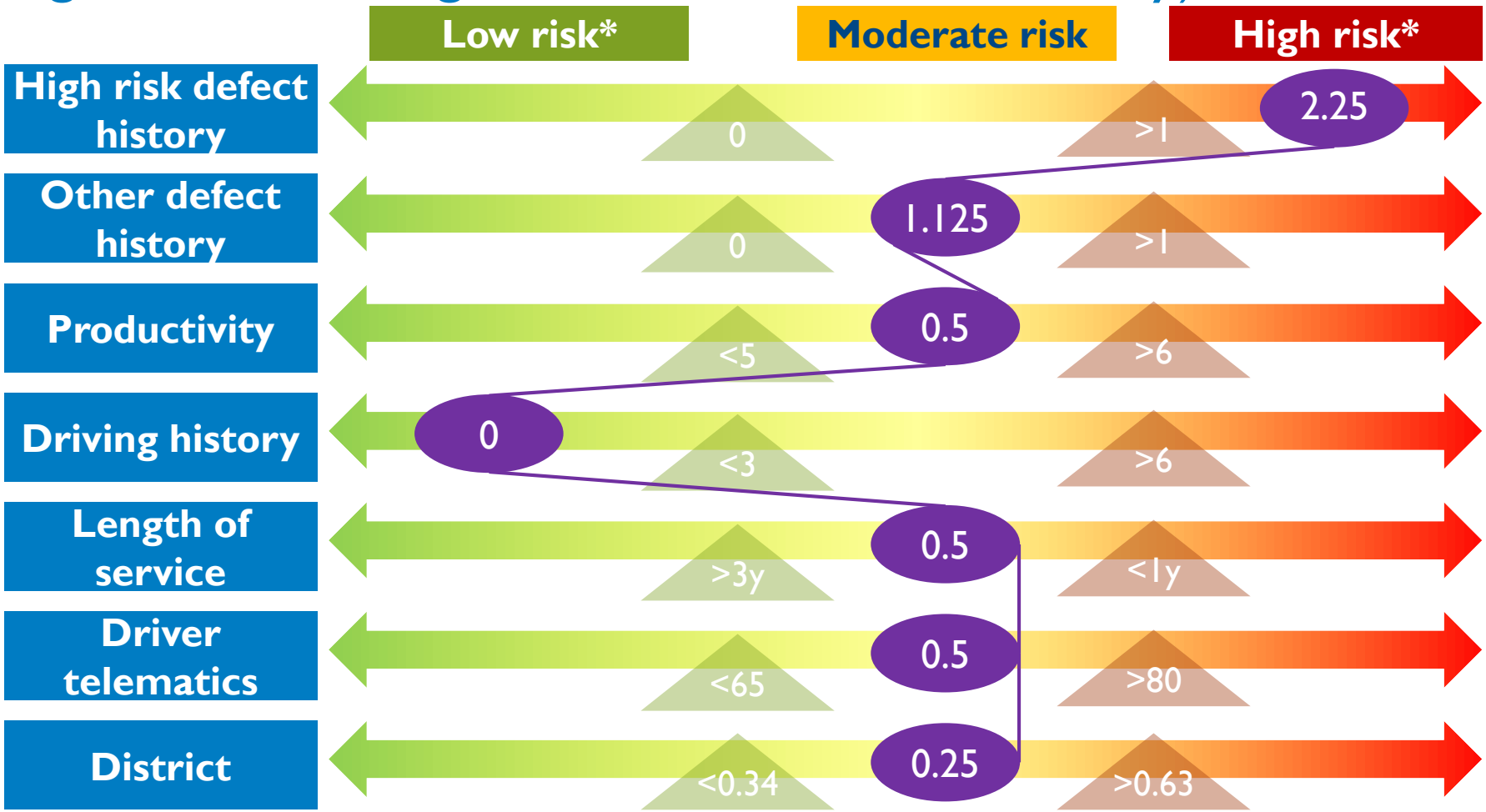
*All values shown are illustrative

…we then assigned weightings to each parameter, the sum of which gives the fitter's modelled risk



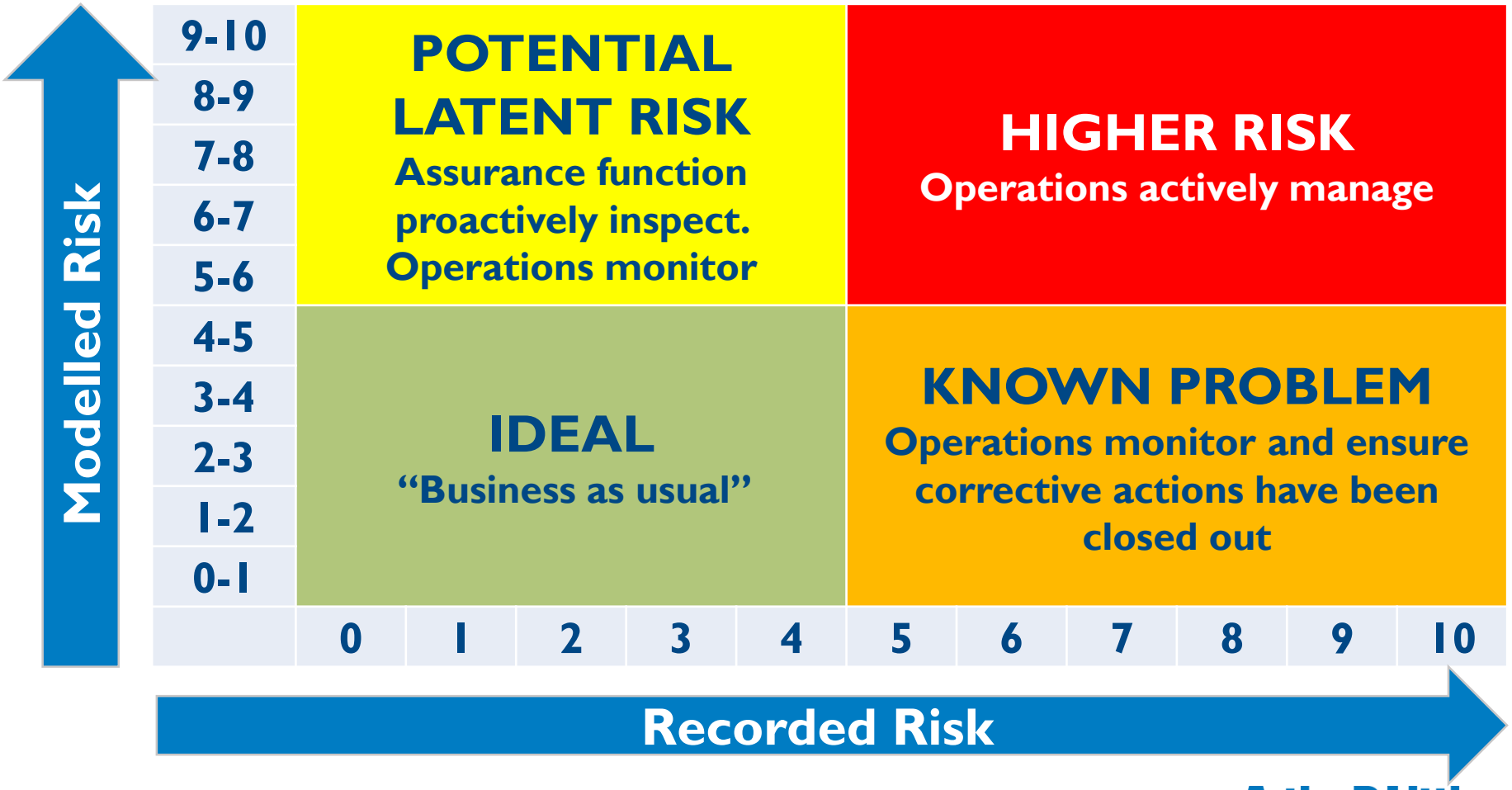
*All values shown are illustrative

EXAMPLE. This fitter's modelled risk is 5.125 (which is higher than average due to their defect history)



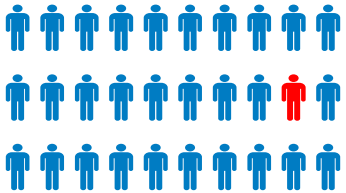
*All values shown are illustrative

Roll out: Model results are updated monthly and guide operational and assurance managers with focused intervention



There is broad application potential to the rail sector: bridges, lineside boxes, cabling, p-way?

Workforce



Dispersed large workforce

Limited supervision or lone working

Assets

Large number and dispersed

Long asset lives

Low probability high consequence failures

Inspections

Frequent inspections impractical

Inspections cannot reveal all defects

Productivity

Productivity = $

But….

too much focus can increase error rate

+ Massive quantities of data (opportunity and challenge)