Tree of Knowledge

Tree Risk Assessment
The Good, the Bad and the Ugly

By David Evans
VALID

Correspondence to
david@validtreerisk.com

Following the Likelihood of Failure article in the Autumn 2018 issue of Tree Matters, it's been suggested that what you'd really like to see is an analysis of the strengths and weaknesses in current tree risk assessment systems to help you make up your own mind. Then have a look at how we might do things better.

In order of popularity, the two main players are;

- **TRAQ**, which is a qualitative approach that measures risk using words.
- **QTRA**, which is quantitative approach that measures risk using numbers

This short review looks at some key technical points in how these systems measure risk. It's not about the people who put them together, the trainers, or those of you who use them. Though about the big two, the same points apply to any tree risk assessment system out there.

On the right TRAQ?

One of the attractions of measuring risk with words is that we're all familiar and comfortable with language. Words, like 'high', 'moderate', and 'low' are easy on the eye, and ear. Everyone thinks they know what they mean. The appeal of risk matrices is their 'friendly bingo card' layout and bottom left to top right risk gradient. They're uncomplicated, straightforward to understand, and are widely used. With only four categories at each stage, TRAQ increases the chances that different assessors will get the same risk by simply limiting their choices.

So what's the problem when measuring risk with words and matrices? It turns out, there's quite a few - too many to cover in such a short piece - but in a nutshell, their strengths are also their weaknesses. Before going there though, let's start with an obvious question to ask about any tree risk assessment system. Does it produce reasonable and believable outcomes? Is it credible?

A game of Russian roulette

If you're unfortunate enough to be playing a game of Russian roulette, with 'The Deer Hunter' rules, you might be surprised and relieved to find that the first round is only a Moderate risk according to TRAQ. If you're still alive by round five you'd be astonished to find that the risk is still Moderate. It's only if you make it to round six, when you're going to shoot yourself in the head that the risk increases to become both High AND Extreme at the same time. Why is such an obviously extreme risk being rated as Moderate by TRAQ?

Russian roulette - The Deer Hunter

Entering the Matrix

Despite their popularity, risk matrices are "*often worse than useless*" (Tony Cox, 'What's Wrong with Risk Matrices', Risk Analysis, 2008) because they frequently fail to rank risks sensibly. There's a wealth of research that repeatedly demonstrates this. One of the more obvious issues is that unless they're clearly defined, words like 'high', 'moderate', and 'low' often mean very different things to different people. Whereas, some of the technical reasons, like 'betweenness', can be complicated to understand. Other ingrained problems with matrices are easier to grasp. We'll have a look at two of them - 'range compression' and 'poor resolution'.

Not too wide, not too narrow

Range compression is where a category is too narrow and too accurate to be plausible. One example of range compression in TRAQ is the likelihood of failure category 'Probable'. Probable is described as '*may be expected*'. In other words, more likely than not. In numbers, that's at least more than 50% of the time. Or where there's a greater than a one in two chance of it happening. This is a terrifically narrow range and requires an unrealistic level of accuracy with no room for uncertainty. Particularly, when the likelihood of failure category above it is Imminent; which is '*most likely...in the near future*' in words, a 100% in numbers, or 1/1 as a probability fraction. In Russian roulette, the likelihood goes from Possible to Probable AND Imminent at the same time in round six.

Likelihood of Failure	Likelihood of Impacting the Target			
	Very Low	Low	Medium	High
Imminent	Unlikely	Somewhat likely	Likely	Very likely
Probable	Unlikely	Unlikely	Somewhat likely	Likely
Possible	Unlikely	Unlikely	Unlikely	Somewhat likely
Improbable	Unlikely	Unlikely	Unlikely	Unlikely

Likelihood Matrix - Russian Roulette

On the other hand, 'Possible' suffers from the opposite of range compression, and that's poor resolution. At the top end, Possible must be at least as high as 50%, or 1/2 because it's the next category below Probable; even though Possible is described as '*unlikely*'. And at the bottom end it needs to have stretched a remarkable distance to stand shoulder to shoulder with the lowest likelihood of failure category, 'Improbable'. Possible is far too wide and vague to be useful.

Likelihood of Failure & Impact	Consequences			
	Negligible	Minor	Significant	Severe
Very likely	Low	Moderate	High	Extreme
Likely	Low	Moderate	High	High
Somewhat likely	Low	Low	Moderate	Moderate
Unlikely	Low	Low	Low	Low

Risk Matrix - Russian Roulette

Adding apples and oranges

An ordinal number is a ranking classification. It's like giving 1st place to your favourite meal, you've come 2nd in a three-legged race, and 3rd in a maths test. You can't add ordinal numbers as you would with cardinal numbers, like 1 + 2 + 3 = 6 to come up with a final figure that means anything sensible. It's mathematically wrong. It'd be like suffering five screenings of 'Sex in the City 2' and adding them up to score it the same as one viewing of 'Citizen Kane'.

Similarly, you can't add ordinal numbers in risk assessments. It's a problem with what's known as 'Matheny & Clark', where ordinal number categories are added to give 'hazard rating' scores from 3 low to 12 high. For example, a high rating like 9 could be an extremely low risk because it's a big tree that's just about to fall, but it's in the middle of nowhere. On the other hand, a 100mm diameter cracked branch over a park bench would only have a rating of 6 even though it's a much higher risk than the 9.

Likelihood of Failure	Likelihood of Impacting the Target			
	Very Low 1	Low 2	Medium 3	High 4
Imminent 4	Unlikely 4+1=5	Somewhat likely 4+2=6	Likely 4+3=7	Very likely 4+4=8
Probable 3	Unlikely 3+1=4	Unlikely 2+3=5	Somewhat likely 3+3=6	Likely 3+4=7
Possible 2	Unlikely 2+1=3	Unlikely 2+2=4	Unlikely 2+3=5	Somewhat likely 2+4=6
Improbable 1	Unlikely 1+1=2	Unlikely 1+2=3	Unlikely 1+3=4	Unlikely 1+4=5

Behind the curtains of the Likelihood Matrix

2+2 = 5

This is why the International Society of Arboriculture's 'Best Management Practices, Tree Risk Assessment' says, "*Risk professionals caution that addition or multiplication of ordinal numbers is mathematically incorrect* ." What's odd about this sound advice from risk professionals is that it's ignored. When you draw back the curtains on the Likelihood Matrix, you can clearly see that it's been built by adding ordinal ranking numbers. As has most of the Risk Matrix.

Trying to make sense of added up ordinal numbers converted into words to measure risk can be a bit of a headache. Are cardinal numbers that you can do maths with the answer?

The numbers game

Measuring risk with numbers makes a lot of sense because it solves many of the problems of measuring risk with words. Numbers are not ambiguous or open to interpretation. They can be compared to tolerable or acceptable levels of risk that we know could be reasonably imposed because of the benefits that trees provide. We can also work out the likelihood of occupation and damage to property by using measurable values rather than words, and then letting you try to work out what those words mean.

That's Numberwang!

What are some of the problems when we measure with numbers, like QTRA does? Firstly, numbers can all too easily baffle and many of us are not very comfortable with maths. Let's explore an example of how numbers look like they can help us, but then cause problems. TRAQ describes a 'minor consequence' in words as '*moderate monetary damage to a vehicle*'. This use of words to measure consequences is not very helpful. Not least because minor damage to a new car could cost a lot more than moderate damage to a second-hand car, or writing off a very old car.

Say we clear up the ambiguity of words and use numbers to agree the average value of a vehicle. You've now got to perform some mental gymnastics to work out how long a parking bay might be occupied for. That means accounting for the hours, days, nights, weekends, holidays, numbers of cars parked, and then coming up with an average occupancy. Not only are you likely to be doing this calculation with incomplete knowledge, but it's a tax on your thinking time, and it's very easy for you to get the maths wrong.

1 Day ÷ 24 ÷ 60 ÷ 60 × y = ?

Similarly, with the number of vehicles per day. Even if you can get your hands on the traffic data, a useful figure to start with can easily become punctured by uncertainty when it's converted into what you might actually see during an assessment. It's complicated to divide day rates into manageable numbers, then make some kind of unknown adjustment because far more vehicles are likely to be there at 5pm than 5am, and on Monday to Friday than the weekend. Each step in your calculation is an opportunity for you to mess it up.

1 in what?

Numbers can not only be tricky when assessing risk, but also challenging for the tree owner or manager. They often struggle to make sense of risk outputs as probability fractions, such as 1/3 000, 1/40 000, or 1/500 000. More importantly, there's too much uncertainty in tree risk assessment to reasonably claim a level of accuracy to one significant figure, like 1/20 000. Or to be able to tell the difference between a 1/20 000 risk and a 1/50 000 risk with enough certainty to justify the difference between them.

Many of us can struggle with maths and numbers

Not dead by 8.6

We've already talked about range compression and poor resolution. QTRA used to suffer from it with its Targets before the release of v5. However, range compression and poor resolution is still very evident in the Size Ranges. The boundaries from the top are 1/1 – 1/2 – 1/8.6. Where 1/1 is a death and 1/2 is half a death.

Similar to TRAQ's Probable, a range of 1/1 – 1/2 at a factor of ×2 is too narrow and accurate for a consequence that has such a high level of uncertainty. The next range is also very slender at just over ×4 from 1/2 to 1/8.6. And then we encounter another credibility issue. Can you really assess the extent of an injury to someone from being hit by a tree part to one decimal place like 8.6? Especially when a 600mm diameter is the 1/1 value, and there's little basis or evidence to mark this diameter as equalling a death. It's simply chosen as 1/1 because it's the largest diameter in the allometric data set used. To compound the problem, 600mm is the weakest part of the data, yet it's the most important measurement because all the Size Ranges are worked from it.

At the other end, we have poor resolution, with 1/82 – 1/2500 being a factor of ×30. And another issue with accuracy. An injury to a person that's 2500 times less painful than death is a questionable level of accuracy to confidently claim or measure. After all, the medical profession doesn't try to measure an injury 1000 times less than a fatality with its Abbreviated Injury Scale, so how can we?

Exposing yourself in public

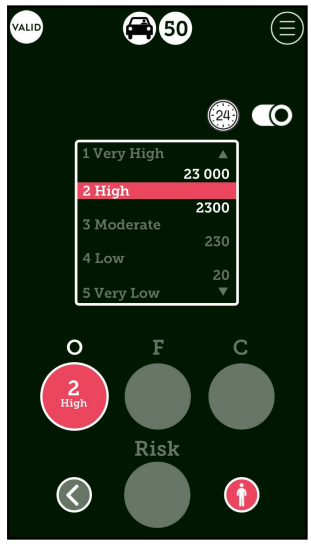
Something that's great about measuring risk with numbers is its transparency. But that can also be a point of great anxiety because if you get the maths wrong it's there for everyone to see, leaving you naked and your reputation exposed. It's noteworthy that the highest profile cases involving QTRA are when there's been a mistake measuring the risk with numbers.

Do we really need another one?
The answer is yes, if it sorts out those problems we've looked at in this article, and more that there wasn't the space to cover. Whether you're assessing or managing tree risk, here's how VALID gives you a helping hand. VALID takes a 'best of both worlds' approach that works on the strengths and ditches the weaknesses when measuring risk with words or numbers. That means collaborating with an independent maths professor who's an expert in measuring risk to model what you're trying to measure, so that it's credible and realistic. Define categories that are just right and neither too wide, nor too narrow. Where necessary, embrace imprecision. Describe the likelihood of occupancy and consequence categories so that you don't need maths, and they can easily be understood and recognised when you're out in the field. Have risk outputs that are believable, that are not too vague or misleadingly accurate. Then test the model for uncertainty and user error to check its resilience.

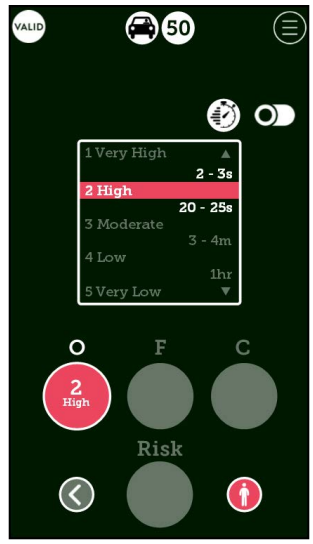
Occupancy greater than 'constant' in words or '1/1' in numbers is common

Where it matters most
One of the more eye-opening revelations when putting VALID together was that the likelihood of occupation on the busiest roads and in city centres is often greater than constant when measured in words. Or 1/1 when measured in numbers. In other words - and numbers - if a tree part falls it's likely more than one person or vehicle will be hit. What that means is the most important assessments, where the 'targets' have the highest value, are being systemically undervalued by every tree risk assessment system out there.

The VALID answer
How VALID deals with occupancy when it's higher than constant in words, or 1/1 in numbers is a window onto how we can make things better. The complex stuff is dealt with in the engine room of the App, and a combination of words and numbers help you easily recognise it when you see it. It goes like this. You're assessing a tree next to a main road in a city with a 50kph speed limit. There's five likelihood of occupation categories available, and from this description alone it won't be three of them. It'll be 1 Very High, or 2 High. Switch the App from '24hr' to 'stopwatch'; which is calibrated to show figures for 7am - 7pm, Monday - Friday. If you're typically seeing a vehicle pass every 2 - 3 seconds or more, then the occupancy is 1 Very High. If it's less than that, then the occupancy is 2 High.



24 hour clock vehicles per day



Switch to stopwatch for user-friendly per vehicle

Loitering without intent
Working out the likelihood of occupation where people loiter or mill around is especially difficult to measure. So let's describe some of them to make it easy for you. A signalised pedestrian crossing in a city centre will most likely be a 1 Very High occupancy. Similarly, a market, an entrance to a mainline train station, or an event like a Royal A&P Show are all 1 Very High. The entrance to a school or college is 2 High. A car park is 3 Moderate. And so on.

Anyone can do (most) of it
Much hard work has gone into VALID's design to make the likelihood of occupation and consequence decisions really straightforward for you when you're on site. That way you don't to have to interrogate your client at length, use a thesaurus, or take out your calculator. It's done through a combination of easy-to-understand descriptive words supported by simple numbers, as we saw earlier with the 50kph road.

The Full Monty
With previous tree risk assessment systems, what's often been overlooked is the tree owner or manager. After all, they're the ones that hold a duty of care and the liability. They're also the ones paying for the risk assessments and any tree works that are necessary to reduce unacceptable risks. We should be looking to help them out as best we can. That's why VALID provides the complete package of tree risk-benefit policy, plan, and assessment.

Back to the likelihood of failure future
By taking care of the occupancy and consequences part of the risk-benefit assessment. Then making them super user-friendly so that you can confidently identify them with a bit of training and a bit of thought. VALID frees arborists up to focus your efforts on what you know best. The tree. And what's the likelihood of failure? (Autumn 2018 issue Tree Matters).