Flood Guidance Verification against Impact data

Presenter: Chris Steele

Author: Michael Sharpe

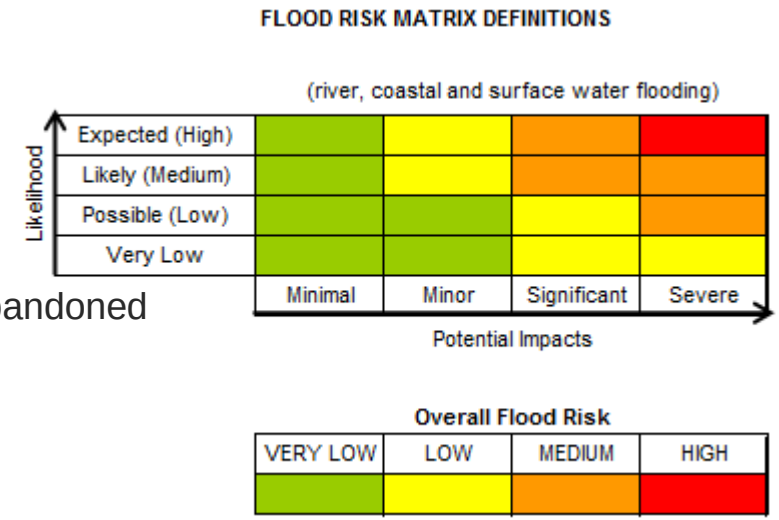
Verification Systems and Operational Products, Weather Science, UKMO

International Verification Methods Workshop

South Africa

May 2024

Risk Matrix Products



- Increasingly, weather-based warning-services are being abandoned and replaced with impact-based alert-services:
  - **not** issued for heavy rain or flooding events,
  - **only** issued if an event:
    - Partially,
    - Significantly, or
    - Severely

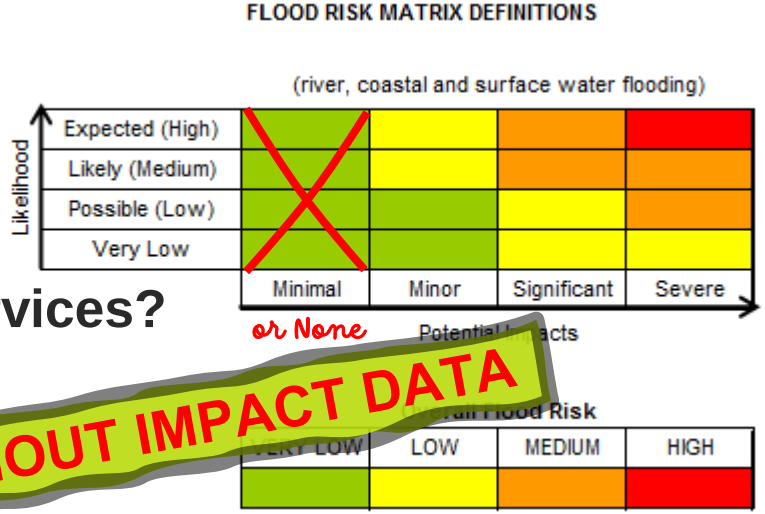
    affects humans or their infrastructure.
- So-called ***Risk Matrices***:
  - express the chance
  - of different severity levels of impact
  - communicate via simple colour-codes

So, ***Risk Matrix*** products are simply **Categorical Probabilistic Forecasting Services**.

Met Office

How do you assess…

Categorical Probabilistic Forecasting Services?



BUT WE CAN'T DO ANYTHING WITHOUT IMPACT DATA

Si[illegible]

$$RPS = \frac{1}{I-1}\sum_{i=1}^{I}\left[\left(\sum_{k=1}^{i} p_k\right) - \left(\sum_{k=1}^{i} o_k\right)\right]^2$$

where:

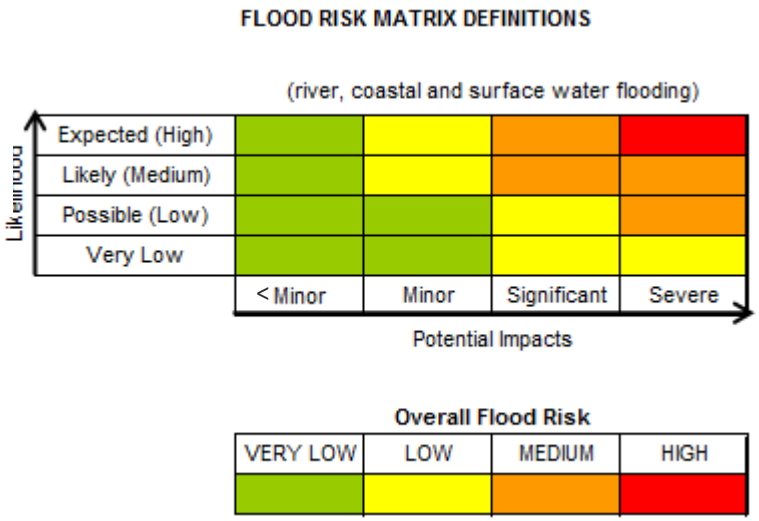
Impact i is:

- ~~Minimal~~ < Minor
- Minor
- Significant
- Severe

p_k = forecast probability of impact category k

…but because, in practice, minimal is too difficult to observe

Is there impact truth data?



Increasingly, YES.

Databases are maintained by

- Environment Agencies
- Local Government authorities
- *Even* the Met Office!
- ?

$$RPS = \frac{1}{I-1}\sum_{i=1}^{I}\left[\left(\sum_{k=1}^{i} p_k\right) - \left(\sum_{k=1}^{i} o_k\right)\right]^2$$

where:

Impact *i* is:

- Minimal
- Minor
- Significant
- Severe

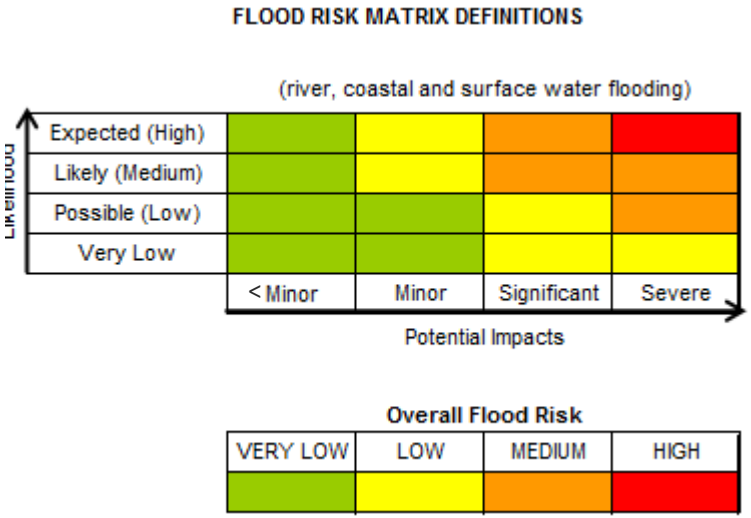
p_k = forecast probability of impact category *k*

What choices are needed?

i.

Likelihood	Probability
Expected	0.8
Likely	0.6
Possible	0.4
Very Low	0.2

but could make another choice (will discuss later)



$$RPS = \frac{1}{I-1}\sum_{i=1}^{I}\left[\left(\sum_{k=1}^{i} p_k\right) - \left(\sum_{k=1}^{i} o_k\right)\right]^2$$

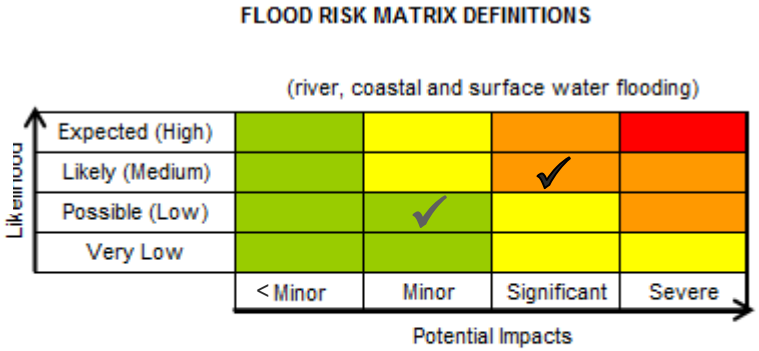
where:

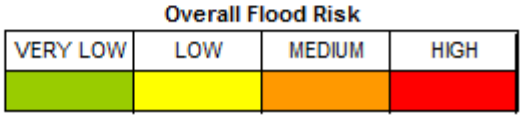
Impact *i* is:

- Minimal
- Minor
- Significant
- Severe

p_k = forecast probability of impact category *k*

What choices are needed? ii.





- We are only given the probability of the **highest** impact level
- We need the rest of the forecast PDF

Example

- Significant/likely (60%)
- What about the remaining 40%?
- Default choice: category $i = 1$, i.e. Minor in this example
- other choices are possible (by arrangement)

$$RPS = \frac{1}{I-1}\sum_{i=1}^{I}\left[\left(\sum_{k=1}^{i} p_k\right) - \left(\sum_{k=1}^{i} o_k\right)\right]^2$$

where:

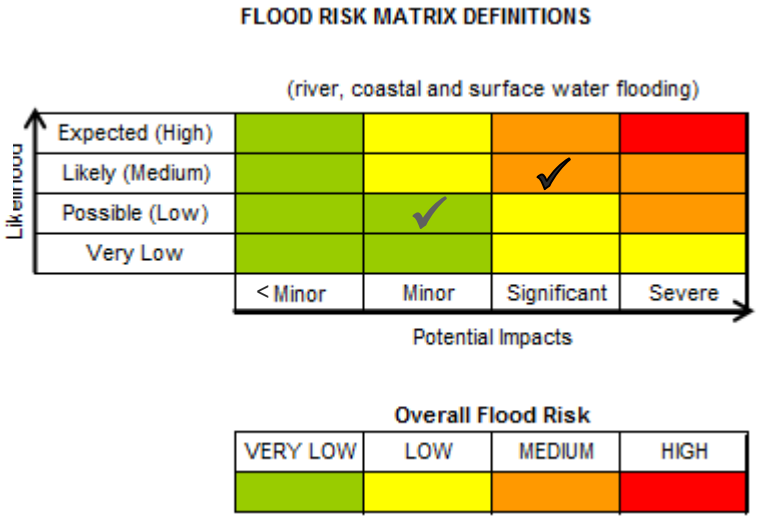
Impact i is:

- Minimal
- Minor
- Significant
- Severe

p_k = forecast probability of impact category k

Score Classifications

- Risk matrix boxes are colour-coded
- So why not performance colour-codes too?



Colour	Performance
	Very Poor
	Poor
	Quite Poor
	Reasonable
	Good

But what should the RPS ranges be for each performance category?

They should be based on the RPS values that correspond to different forecast scenarios

Score Classifications

Likelihood	Probability
Expected	0.8
Likely	0.6
Possible	0.4
Very Low	0.2

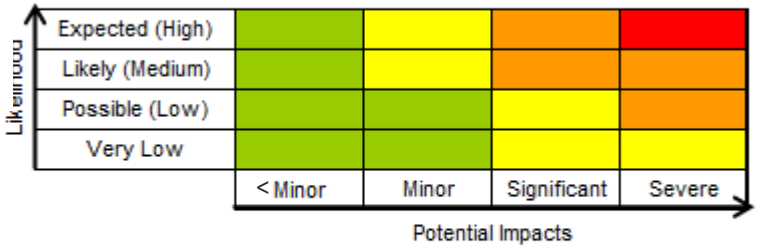
Colour	Performance
	Very Poor
	Poor
	Quite Poor
	Reasonable
	Good

With these *Likelihood* probability choices, when a **missed Impact** occurs:

	Type of missed event		
	Minor	Significant	Severe
RPS	0.33	0.67	1.00
	Reasonable*	Poor*	Very poor*

FLOOD RISK MATRIX DEFINITIONS

(river, coastal and surface water flooding)



Overall Flood Risk			
VERY LOW	LOW	MEDIUM	HIGH

*Sharpe's personal assessment

Met Office

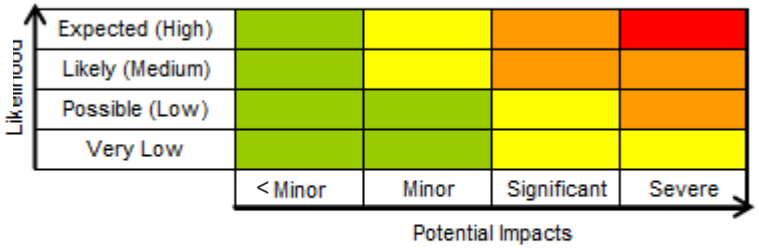
Score Classifications

Likelihood	Probability
Expected	0.8
Likely	0.6
Possible	0.4
Very Low	0.2

Colour	Performance
	Very Poor
	Poor
	Quite Poor
	Reasonable
	Good

FLOOD RISK MATRIX DEFINITIONS

(river, coastal and surface water flooding)



With these *Likelihood* probability choices, when a **Severe Impact** occurs:

Overall Flood Risk

VERY LOW	LOW	MEDIUM	HIGH

Severe Occurred	Forecast Impact		
Forecast Likelihood:	Minor	Significant	Severe
Expected	0.68	0.35	0.01
Likely	0.72	0.39	0.05
Possible	0.79	0.45	0.12
Very Low	0.88	0.55	0.21
	*Poor – Very Poor	*Reasonable – Poor	*Good – Reasonable

*Sharpe's personal assessment

Score Classifications

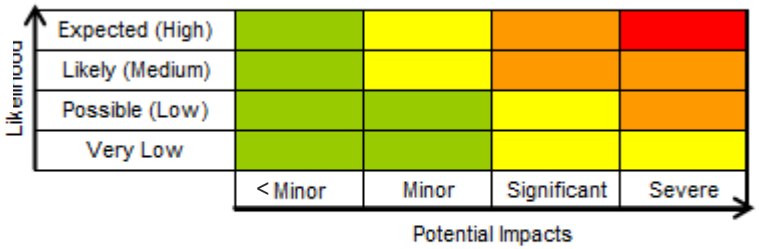
Likelihood	Probability
Expected	0.8
Likely	0.6
Possible	0.4
Very Low	0.2

Colour	Performance
	Very Poor
	Poor
	Quite Poor
	Reasonable
	Good

After categorising each forecast vs observed impact scenario…

FLOOD RISK MATRIX DEFINITIONS

(river, coastal and surface water flooding)



Overall Flood Risk

VERY LOW	LOW	MEDIUM	HIGH

an RPS range for each performance category is obtained…

So

- RPS values can be hidden,
- colour-coded categories used instead, and
- definitions supplied for each category

Colour Definitions Table

Represents:

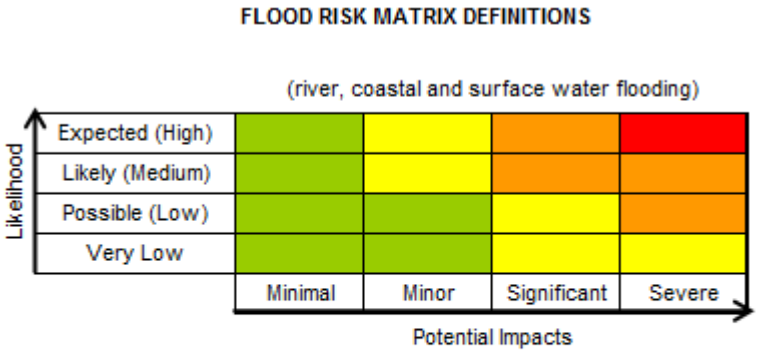
- Forecasting that a:
  - *Severe* Impact
  - is *Likely*
- What occurs:
  - nothing

- Forecasting that a:
  - *Minor* Impact
  - with *Very Low* probability
- What occurs:
  - a *Severe* Impact

Colour	Type of Missed event	Type of False Alarm		Type of Hit		
		Forecast Impact	Forecast Likelihood	Observed Impact	Forecast	
					Impact	Likelihood
	Severe	Severe	Expected	Severe	Minor	Very Low
	Significant	Severe	Likely, Possible, Very Low	Minor	Severe	Expected
		Significant	Expected	Significant	Minor	Very Low
				Severe	Minor	Possible, Likely, Expected
					Significant	Very Low
	-	Significant	Likely, Possible	Minor	Severe	Likely, Possible
				Significant	Minor	Likely, Possible
					Severe	Likely, Expected
				Severe	Significant	Likely, Possible
	Minor	Minor	Likely, Expected	Minor	Minor	Very Low, Possible
		Significant	Very Low		Significant	Likely, Expected
					Severe	Very Low
				Significant	Minor	Expected
					Significant	Very Low, Possible
					Severe	Very Low, Possible
				Severe	Significant	Expected
					Severe	Very Low, Possible
	-	Minor	Very Low, Possible	Minor	Minor	Likely, Expected
					Significant	Very Low, Possible
				Significant	Significant	Likely, Expected
					Severe	Very Low, Possible
				Severe	Severe	Likely, Expected

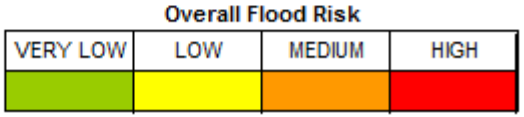
Met Office

Real results example…



- Calendar View of Performance

  What was the performance of all alerts issued for (or all events to occur on) a particular calendar day?



- Lead-time View of Performance

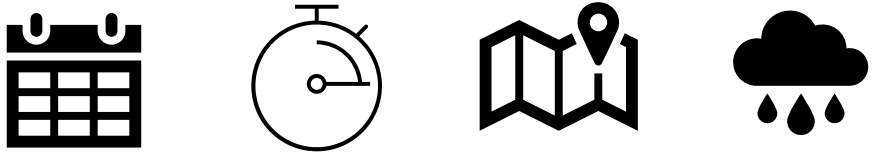
  What was the performance of all forecasts issued at a particular lead-time? (e.g. all day-3 forecasts)

- Regional View of Performance

  What was the performance of the service within each alert area?

- Forecast View of Performance

  What as the overall performance of each 5-day forecast?



Met Office

Real results example…

Calendar View of Performance

What was the performance of all alerts issued for (or all events to occur on) a particular calendar day

FLOOD RISK MATRIX DEFINITIONS

(river, coastal and surface water flooding)

Likelihood	Minimal	Minor	Significant	Severe
Expected (High)				
Likely (Medium)				
Possible (Low)				
Very Low				

Potential Impacts

December 2022						
1	2	3	4	5	6	7
8	9	10	11	**12** 1	13	14
15 4	**16** 5	**17** 5	**18** 72	**19** 72	**20** 1	21
22 24	**23** 8	**24** 8	**25** 7	**26** 16	**27** 28	**28** 30
29 32	**30** 40	**31** 20				

Header-colour indicates mean monthly performance

Overall Flood Risk

VERY LOW	LOW	MEDIUM	HIGH

5-day flood-impact forecast issued daily

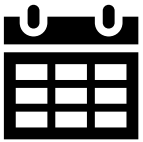
19 forecast areas

Box colour denotes performance of all forecasts valid on that day

White means no impact was forecast or observed on that day

The number of alert areas on this day where:

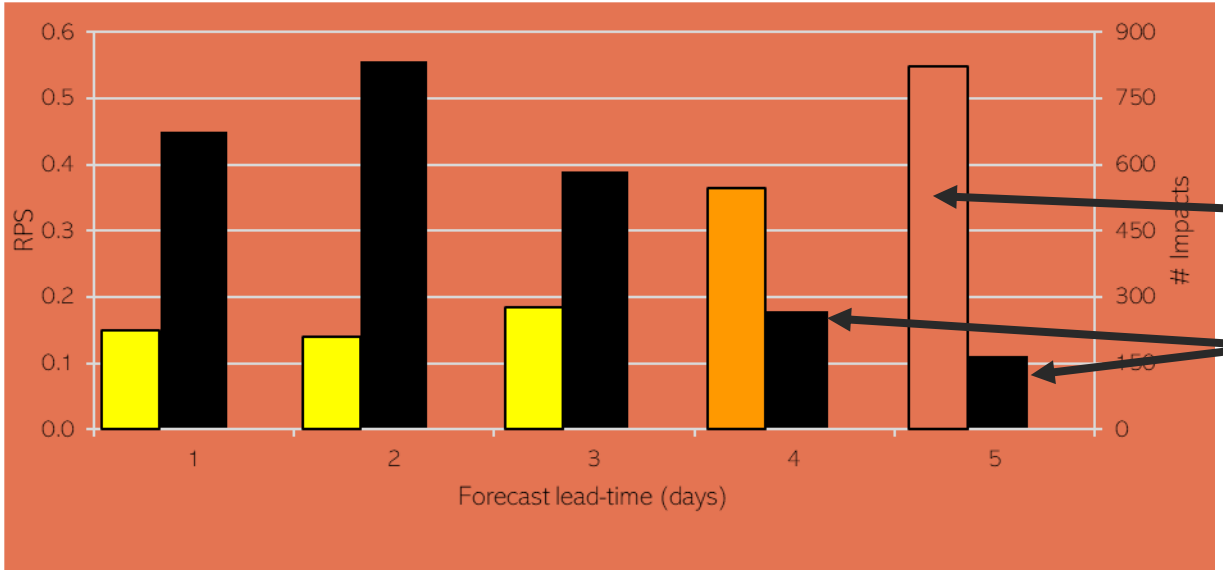
- an impacts occurred, or
- an impact was forecasts



Real results example…

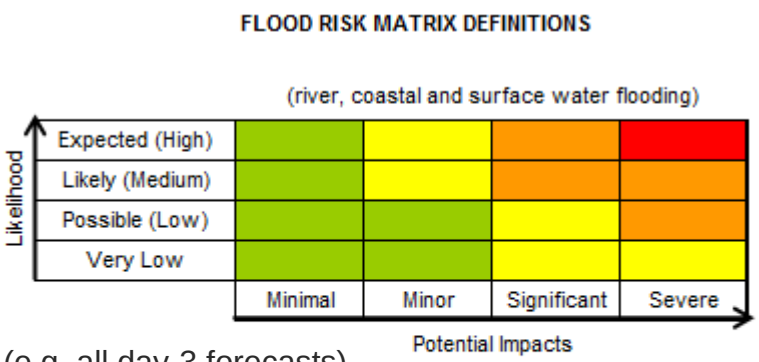
Lead-time View of Performance

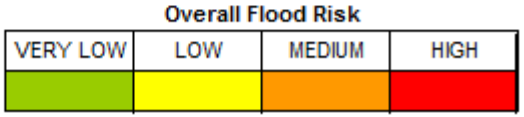
What was the performance of all forecasts issued at a particular lead-time? (e.g. all day-3 forecasts)

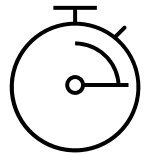


Bar colour gives performance category

#forecast + observed impacts
(so this includes false alarms)



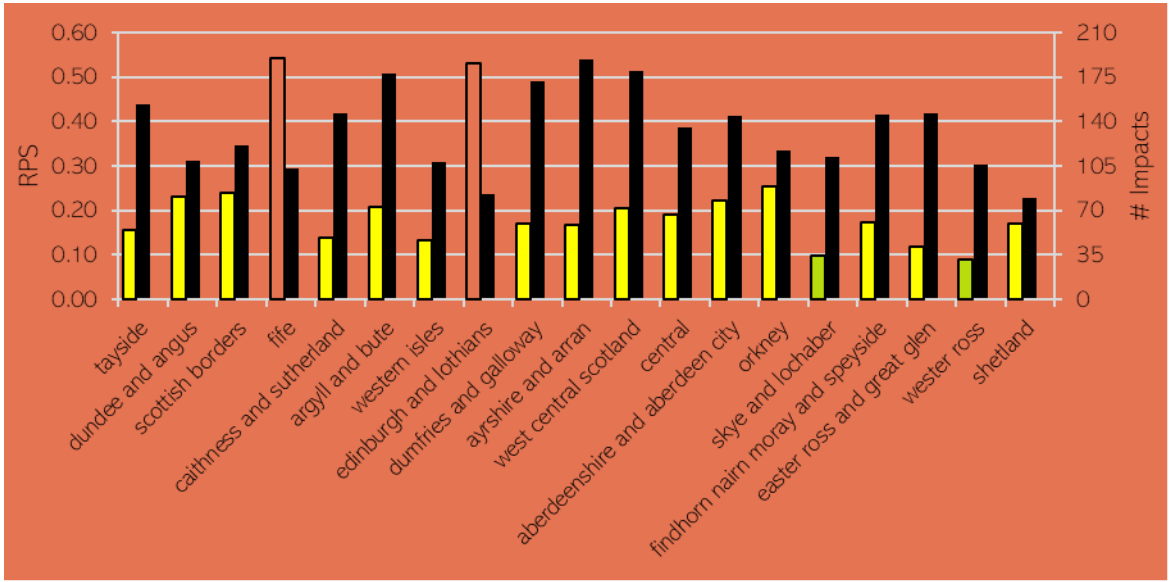


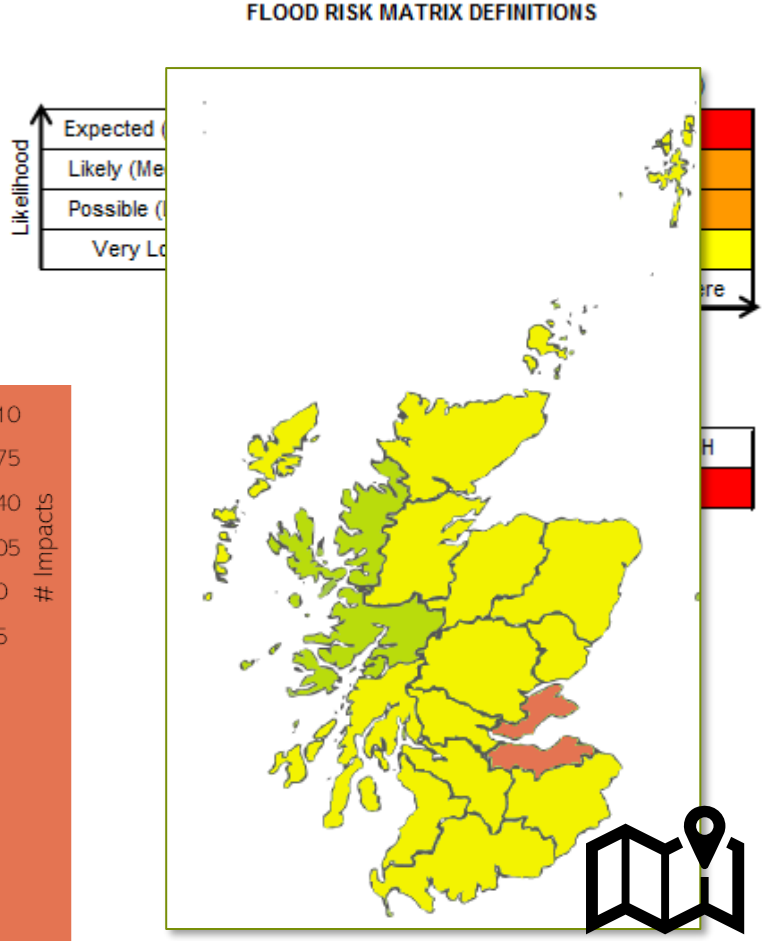


Real results example…

Regional View of Performance

What was the performance of the service within each alert area

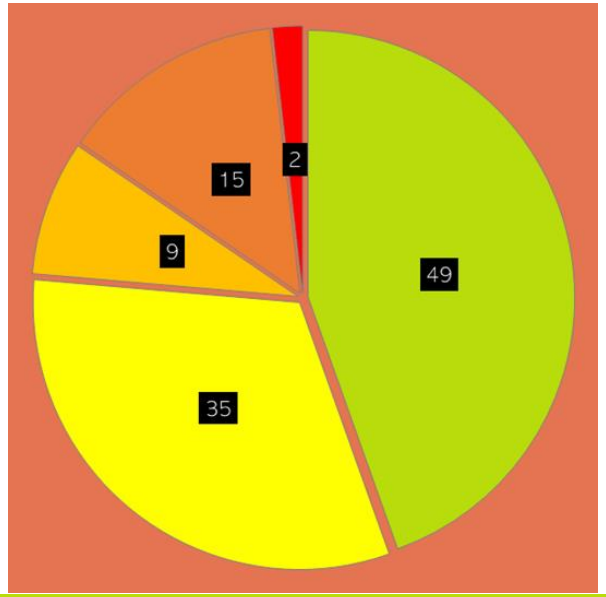




Real results example…

Forecast View of Performance

What overall performance of each 5-day forecast?



FLOOD RISK MATRIX DEFINITIONS

(river, coastal and surface water flooding)

Likelihood	Minimal	Minor	Significant	Severe
Expected (High)				
Likely (Medium)				
Possible (Low)				
Very Low				

Potential Impacts

Overall Flood Risk

VERY LOW	LOW	MEDIUM	HIGH

How many daily 5-day forecasts showed:

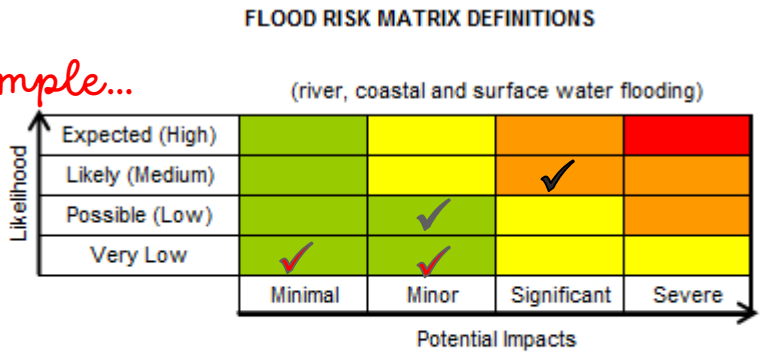
- Good
- Reasonable
- Quite Poor
- Poor
- Very Poor

performance?

Met Office

But isn't this all arbitrary?

For example...



BUT…

Overall Flood Risk			
VERY LOW	LOW	MEDIUM	HIGH

Likelihood	Probability
Expected	0.8
Likely	0.6
Possible	0.4
Very Low	0.2

arbitrary?

- Forecast Significant/likely (60%)
- Remaining 40%: category $i - 1$, i.e. Minor

arbitrary?

Could configure alternatives within a forecasting tool

Different choices => different RPS values **BUT** the same performance colours

Thank you, any questions?

Presenter: Chris Steele

Author: Dr Michael Sharpe, BSc., DipStat(Open), FRMetS

Verification Systems and Operational Products, Weather Science, UKMO

International Verification Methods Workshop

South Africa

May 2024