Identifying skilful spatial scales using the Fractions Skill Score

Marion Mittermaier and Nigel Roberts
Outline

1. Brief overview of the method

2. Synthetic cases
   • Geometric cases
   • Perturbed cases

3. Domain size influences

4. Conclusions
The method
Verification approach

We want to know:

3. How forecast skill varies with neighbourhood size.
4. The smallest neighbourhood size that can be used to give sufficiently accurate forecasts.
5. Does higher resolution provide more accurate forecasts on scales of interest (e.g. river catchments)

Compare forecast fractions with fractions from radar over different sized neighbourhoods (squares for convenience)

Use rainfall accumulations to apply temporal smoothing
Schematic comparison of fractions

Threshold exceeded where squares are blue

observed

forecast

\[
\text{Fraction} = \frac{6}{25} = 0.24
\]

\[
\text{Fraction} = \frac{6}{25} = 0.24
\]
Idealised example
In summary

This verification method provides a way of answering some important questions about forecasts from ‘storm-resolving’ NWP models.

• How does forecast skill vary with spatial scale?
• At what scales are higher resolution forecasts more skilful (if any)?
• At what scales are forecasts sufficiently accurate?............

(There are other questions that need different approaches)
How we are using it

6hr Precip Accumulation (10.0%): Analysis
Southern UK area (scale rainfall)

10% threshold
12.5 mm

FSS
0.5

L(FSS>0.5)
20 km

Freq. Bias

See poster by Mittermaier and Thompson
Synthetic cases
THE WINNER

200 km

> 600 km

> 600 km
Summary:
\[ L(FSS > 0.5) \]

<table>
<thead>
<tr>
<th>Case</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>200 km</td>
</tr>
<tr>
<td>002</td>
<td>&gt; 600 km</td>
</tr>
<tr>
<td>003</td>
<td>&gt; 600 km</td>
</tr>
<tr>
<td>004</td>
<td>550 km</td>
</tr>
<tr>
<td>005</td>
<td>?</td>
</tr>
</tbody>
</table>
Conclusions: geometric cases

- Case 005 has the largest FSS at the grid scale but in terms of $L(\text{FSS} > 0.5)$ never reaches a skillful scale.

- Case 001 is the most skillful in terms of scale but has no skill at all at the grid scale, based on the FSS.
Perturbed forecasts
**Summary : L(FSS > 0.5)**

<table>
<thead>
<tr>
<th>Case</th>
<th>&gt; 0.5 mm</th>
<th>&gt; 32 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>4 km</td>
<td>20 km</td>
</tr>
<tr>
<td>002</td>
<td>4 km</td>
<td>40 km</td>
</tr>
<tr>
<td>003</td>
<td>10 km</td>
<td>80 km</td>
</tr>
<tr>
<td>004</td>
<td>50 km</td>
<td>190 km</td>
</tr>
<tr>
<td>005</td>
<td>160 km</td>
<td>&gt; 200 km</td>
</tr>
<tr>
<td>006</td>
<td>10 km</td>
<td>?</td>
</tr>
<tr>
<td>007</td>
<td>35 km</td>
<td>90 km</td>
</tr>
</tbody>
</table>

© Crown copyright   Met Office
Conclusions: perturbed cases

• Small shifts/timing errors are initially unharful to the FSS and scale. Gross timing errors result in large degradations in skillful scales and the overall magnitude of the FSS.

• Changes in the actual magnitudes (006 and 007) show greater impact. Subtle differences (due to near-threshold level misses) seem to be potentially more serious in terms of FSS magnitude.
Domain size influences

t+24h 1-hr forecast accumulation
• With such a large domain the **wet-area ratio is quite small**, even for extensive precipitation areas.

• *How sensitive are the results to domain size?*

• Perform tests where around ~30% of the domain is used
Impact of domain size 2

- Domain size affects the magnitude of the FSS and the spatial scales.
- Smaller domains are faster to compute.
- Sub-regions take into account that neither precipitation nor skill is uniform over a large domain.
General conclusions

- The absolute value of FSS is less potentially useful than the scale where an acceptable level of skill is reached.

- The skillful scales in the fake cases ties in well with the 1-d idealized example results.

- Frequency thresholds are potentially useful but not when the domains are very large. [The zeros dominate the cdf.]

- Consider a “roving” (fixed size) verification domain that focuses on an area of interest.

- Other optimizing ideas are being explored.
Questions?