SEJ Flyover

Cooke

RFF

Jan 25 2018
Quantifying Uncertainty: Structured Expert Judgment

Procedures guide for structured expert judgment
Anno 2016 over 200 professional applications

<table>
<thead>
<tr>
<th>Area</th>
<th>Organizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear</td>
<td>EU, USNRC</td>
</tr>
<tr>
<td>Aerospace</td>
<td>ESTEC, NASA</td>
</tr>
<tr>
<td>Chemical Process</td>
<td>VROM, SHELL</td>
</tr>
<tr>
<td>Dose Response</td>
<td>VROM</td>
</tr>
<tr>
<td>Environmental Transport</td>
<td>EU, USNRC, VROM</td>
</tr>
<tr>
<td>Banking / Investment</td>
<td>SHELL, AMS Optie</td>
</tr>
<tr>
<td>Volcanoes</td>
<td>UK, EU</td>
</tr>
<tr>
<td>Aeronautics</td>
<td>VROM, AIRBUS, BA</td>
</tr>
<tr>
<td>Project mngt</td>
<td>Robert Woods Johnson</td>
</tr>
<tr>
<td>Public Health</td>
<td>Health Canada</td>
</tr>
<tr>
<td>Civil Infrastructure</td>
<td>UK, NL, EPA</td>
</tr>
<tr>
<td>Invasive Species</td>
<td>NOAA</td>
</tr>
<tr>
<td>Ice Sheets</td>
<td>RL Foundation, UK</td>
</tr>
<tr>
<td>Global Burden of Disease</td>
<td>WHO, CDC</td>
</tr>
</tbody>
</table>
Prob FALSE rejection:
Expert 1 = 0.4
Expert 7 = 0.33
Expert 4 = 0.000006
Statistical Accuracy, 323 post 2006 experts

Expert Nr
WHO Global burden of disease
72 experts, 135 panels, Remote elicitation by novices

Global Burden of Disease: 72 WHO experts
First Miracle of SEJ
EW tends to give good statistical performance (at the expense of informativeness)

Global Burden of Disease: 72 experts, Equal weight
Second Miracle of SEJ
PW preserves statistical accuracy and recovers informativeness

Global Burden of Disease: 72 experts, Equal weight, Performance weight
Peer Rankings DON’T Predict Performance
Out of Sample Cross-Validation: of Classical Model

62 studies, per study: geomeans of comparisons of PW/EW combined score ratios. Eggstaff, Mazzuchi, Sarkani (2013 RESS);
Questions?
Average over all studies per % training set size of the average $PWS_a$ and average $EWS_a$ (post 2006)
Average over all studies per % training set size of the average \textit{PWInf} and average \textit{EWInf}\bigskip

![Bar graph showing average PWInf and EWInf across different % training set sizes.](image-url)

- \text{PWInf}
- \text{EWInf}

% calibration variables in training set

- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
Average over all studies per % training set size of the average PWComb and average EWComb

% calibration variables in training set