The Concept of ‘Deterministic Limit’

Tim Hewson
1 Feb 2007
1. Definitions and Examples

2. Forecast Recalibration

3. Meteorological Complications

- real examples used here relate to strong wind events; concepts apply more generally, to many parameters
1. A New Verification Measure

- Conceptually there should be a ‘deterministic limit’ for predicting a **pre-defined meteorological event** (such as strong winds at a point)
- Simply defined this could be the point in lead time beyond which forecasts **concerning** that event are more likely, on average, to be wrong than right
- This can provide guidance on when to shift emphasis, in forecasts for particular events, from deterministic towards probabilistic
- For rare events at least, correct null forecasts – i.e. the majority - can be ignored as not relevant
The ‘deterministic limit’ for the event in question is then simply the lead time at which, over a suitably large forecast sample, *hits* equals the sum of *misses* and *false alarms* (or CSI = 0.5)

\[
a/(b+c) = 1
\]

\[
\text{misses + false alarms}
\]

\[
\text{hits}
\]

\[
\text{lead time}
\]
Event Examples (numbers are crude estimates)

- Tornado within 2km radius (deterministic limit ~ 5 mins)
- Snow falling at a point (~5 hours)
- Rain falling at a point (~18 hours)
- Gale force gusts at a point (~2 hours)
- Gale force gusts within a UK county (~6 hours)
- Rainfall >15mm in 3 hours somewhere in a UK county (2 hours)
- Cyclonic surface pressure pattern at a point (~120 hours)
- Atmospheric front within 200km of a point (~60 hours)
- Day with maximum above 30C in London (~96 hours)
- ‘Change of synoptic type’ for the UK (~4 days)
Site Specific Example – Mean Winds

2 years data

Deterministic limit for F7 mean winds at Lerwick is ~15 hours

F8 mean winds at Lerwick should be predicted probabilistically at all leads (DL<=0)
Regional Example, \( \geq \) F8 winds

- Similar to Lerwick example, but for all Scottish sites considered collectively.
- Forecast event definition is: ‘gales will occur at a particular site in Scotland, at a particular time’
- Deterministic limit \(<\ 0\)
- Base rate \(<\ 2\%\)

- Forecast event definition is: ‘gales will occur \textit{somewhere} in Scotland, at a particular time’
- Deterministic limit \(~8\) hours
- \textbf{Deterministic forecasts that are geographically less specific are much more valid}
- Time-windowing should increase DL further
- Partly an impact of a higher base rate

\[ \text{Hits} \quad \text{Misses} + \text{False Alarms} \]

With tolerance equal to observation discretization
Some Implications for Windstorm Forecasts

- Deterministic Limit will decrease as areal specificity of the forecast increases
- Hence provides **pointers to forecast and warning content**, and suitable product development, as a function of lead time, eg:
  - day 5: part of continent
  - day 4: country groups
  - day 3: countries
  - day 2: extended regional
  - day 1: regional...

- Partly hypothetical, requires testing!
- In reality, areas may overlap, disappear
- **Aim of using DL is to minimise overlap**
Potential to provide a *meaningful succinct measure* of what to expect from, and therefore what to put into, a forecast. Too many forecast elements are deterministic.

It is something that the public, other customers (and auditors!) could potentially relate to.

Provides a means for inter-comparing the relative merits of ‘operational’ and ‘ensemble’ runs (further work is required on ensemble application).

As always extreme events would be more difficult to represent (though hindcasts from re-analyses are becoming increasingly tractable).

Provides facility also to measure forecast improvements, compare systems, assess forecaster performance.
Further thoughts – “Unbiased Forecasts”

- In the simple case of a fully reliable (unbiased) forecast system, no of false alarms (b) = no of misses (c)

- So the deterministic limit, where \( \frac{a}{b+c} = 1 \), becomes \( a = 2b \)

- Number of events observed, \( O = a+b \)

- Thus \( O = 3b = \frac{3a}{2} \)

- So the ‘deterministic limit’ for an unbiased forecast system is reached when the no of hits (a) drops to two thirds of the no of observed events (O)

- Frequency-preserving recalibration should be used to arrive at an unbiased forecast in most circumstances
3. Recalibration

- Recalibration is a fundamental requirement for model wind speed forecasts, due to biases and local effects. Without this the deterministic limit is less than zero.

- Example below is a contingency table for the windiest site in the UK (N Rona, an island NW of Scotland) for mean winds exceeding 30m/s (58kts), 2004-2006, based on T+0 mesoscale model (12km) data.
Meaning of ‘Reliable Recalibration’

- Example is for London Heathrow T+24 forecasts
- ‘Reliable’ recalibration is so-named because it is frequency preserving. Percentile matches are used to ascertain whether a forecast is above threshold.
- Misses = False Alarms
- Problems can occur (as always!) with model changes
By making a first order assumption of a linear reduction in point density in the two directions shown (s,n), relative box populations can be computed geometrically.

This leads to the result that for the DL to be greater than the lead time to which the plot corresponds (ie hits > misses + false alarms), requires

\[ s > 3n \]

Could be used to instantly assess the validity of making deterministic categorical forecasts for rarer values of a particular parameter at a particular lead time (based on past performance).
- Intersection of $Y=2/3$ with ROC curve corresponds to Deterministic Limit = plot lead time

- Thus a ROC curve (for a particular lead) will tell you the base rate of the event for which the DL equals that lead time

- Event threshold ($T$) then relates directly to base rate

- Assumes recalibration incorporated

$$\text{Base rate} = \frac{3V}{3V+1}$$
Example…

‘Most severe’ windstorm of the 2006/7 European winter period so far (highest gusts anyway!)

Not forecast

Recalibration would not help
‘Arcachon Windstorm’ - 4 hours before hitting French coast

00Z 3rd Oct
Gust 89kts

Gust 82kts

Geostrophic wind reached ~350kts!!!

100km
42-hour lead – feature tracks & feature plumes

Central Pressure

Low level wind max

Upper level wind max
Summary

- New verification concept introduced – the ‘deterministic limit’ (DL)
- Examples, for strong wind exceedances in the UK, illustrate that DL depends on:
  - Base rate
  - Areal (and temporal) specificity
  - Tolerance when judging hits
- Examples also illustrate the requirement for recalibration
- ‘Reliable’ recalibration is needed to maximise forecast information utilisation (maximising DL)
- Scatterplot structure relevant for ascertaining ‘deterministic forecastability’
- DL values denote when to move from deterministic to probabilistic forecasts
- Requirement for expanding sample size, of ‘adverse weather’ events, is clear:
  - Improve model diagnostics (eg multi time-step interrogation for winds)
  - Improve verification data storage (archive all observations of all relevant events)
  - Reserve supercomputer time to perform re-runs of past events with new model versions
- **NEXT**: need to define more user-relevant events, and compute DL for these….
  Leading to a **catalogue of DL values for a wide range of weather events.**
Challenges

- Accounting for observational error
  - e.g. Fewer gales occur than are reported!
- Performing reliable recalibration when observational error has been accounted for
- Performing reliable recalibration in real time
- Data collection needs to be improved
- Meteorological complexity – never forget the physics!
- Extreme events