

FSS(fractions skill score)

共通の座標系に落とした2データについて、辺がLの四角の範囲を近傍域(neighbour)とし、あるしきい値以上の降水の比率を計算する。それが合うかどうかで判断する。0.5以上で空間的一致が満足できるレベルと考える。

原著論文は下記の2本：

Roberts, N., 2008: Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteor. Appl.*, 15, 163-169, doi:10.1002/met.57. ←こっちの方がコンパクトで読みやすいし長期データへの適用例もある。

Roberts, N. and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, 136, 78-97, doi:10.1175/2007MWR2123.1. ←くわしく書いてある。

計算方法

Stage 1：割合fractionの算出

- 予報と観測で同一の変数を同一の格子上に投影する。
- 予報の各格子点について、ある大きさの正方形(近傍域)内にある格子が特定のしきい値(例えば積算降水量1mm)を越える割合を計算する。For every forecast pixel, the fraction of surrounding pixels with in a given square area (neighbourhood) that exceed a specified accumulation threshold (e.g. >1 mm) is computed.
- 観測値についても同様に割合(fraction)を計算する。
- 近傍域の大きさとしきい値を変えてfractionを計算する。つまり、予報値と観測値のfractionの対が、(近傍域の大きさの数)×(しきい値の数)組だけできる。
- Roberts2008では、観測と予報の値の特性の差異の影響を除いて位置精度を評価するため、しきい値は観測値そのものではなくpercentileが同じもの(95th percentileなど)で比較している。

Stage 2：FSSの計算

FSSはBrier Skill Scoreの変種であり、

$$FSS = 1 - \frac{FBS}{FBS_{WORST}} \quad (1)$$

で表される。FBSはFractions Brier Scoreであり、

$$FBS = \frac{1}{N} \sum_{j=1}^N (O_j - M_j)^2 \quad (2)$$

で得られる。ここで、ここで O_j と M_j は各格子における観測と予測の確率値(もしくは割合の値)であり、0から1の間の値を取る。Nは評価する領域の格子数である。通常のBrier Scoreでは観測値 O_j が現象ありで1、なしで0の値を取る。FBSはここを観測値の方も確率値にしたものといえる。FBS_{worst}

$$FBS_{WORST} = \frac{1}{N} \left[\sum_{j=1}^N O_j^2 - \sum_{j=1}^N M_j^2 \right] \quad (3)$$

は O_j と M_j が全く一致しなかったとき(全ての格子でどちらかが0; no collocation of non-zero fractions)の値であり、取り得る最悪の値である。FBSでは最悪=最大であることは(2)を展開すればすぐわかる。FSSは次のような特徴を持つ。

1. 0-1の値を取る。0は完全な不一致、1は完璧な予報。

2. しきい値を超える予報がなく観測はある場合、また予報はあるが観測がない場合、スコアは常に0。
3. 割合を計算する正方形が大きくなるほどスコアは予報と観測の現象頻度の比に漸近する。漸近値が1に近いほど予報バイアスは小さい。percentileを用いることで、近傍域の大きさが解析領域に近づくほど1に漸近するよう設定できる。
4. スコアはまれな値に敏感であり、発生頻度の低い地域では変動しやすい。

多数の予報について計算する場合、FBSとFBS_{WORST}を個々の予報について求め、それらからFSSを求めて良い(式見れば分かる)。

◆データのバイアスについて

99%など、percentileの大きいものは局地的な対流性降水やエコー中のコアなどに対応し、75%はもっと面的に広がった降水に対応する。percentileはデータのバイアスがあまり大きくなない場合、それを除いて比較するのに有効である。この論文で使ったデータは、予報と観測の強雨(>8mm/day)域面積比が0.5から2.0の間であった。バイアスがあまりに大きい場合、そもそもどのスケールでも誤差が大きくなり、場所の一致を議論してもあまり意味はない。

◆a minimum required level of skill for a useful forecast

使える予報の最低(要求)レベルとしては、random skillとperfect skill(=1.0)の真ん中の値(平均)を考えれば良い。これは理想化実験で位置ずれの2倍の近傍域を用いた時の値に近い(Roberts and Lean 2007)。ただ、Vaughan et al. (2014)によると、0.5として構わないようだ。

Roberts 2008にはこのあとスケールや予報時間に対するスキルの値の変化が議論されているが、基本的には、スケールの小さい方が難しい、と理解して、0.5をしきい値として、「この予報は大きい方からどれくらい小さいスケールまで合っていると言えるのか?」という考え方をすれば良いのではないか。