

アンサンブルデータの統計解析

上野 玄太 モデリング研究系 教授

【問題意識】

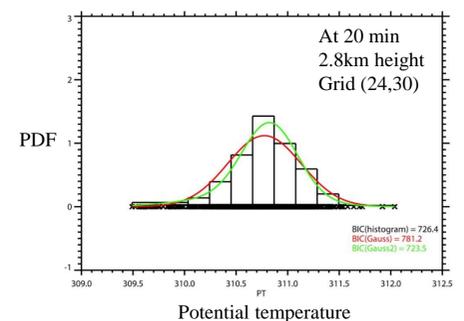
- 粒子フィルタで、確率分布を正確に再現できているはず
- アンサンブル計算データをどのように扱うか？
- 従来の、平均・分散・スプレッド(最大値-最小値)以上の情報を抽出すべき
- とりあえず「非ガウス」

【(試行)高次モーメントの限界】

- ガウス分布 ⇒ 3次モーメントがゼロ
- 3次モーメントがゼロでない ⇒ ガウス分布でない
 - ✓ 有限個のサンプルから計算するので、ガウス分布からのサンプルであったとしても、厳密にゼロになることはまずない
 - ✓ 逆に、ゼロが得られた場合でも、本当にゼロ？
 - ✓ では、ゼロ条件を緩和して、3次モーメントが小さければガウス、大きければ非ガウス、・・・でいいの？ では敷居値はいくつ？
 - ✓ ふたこぶの分布だと、3次モーメントはゼロ

【ガウス分布かどうか？ という主張】

- 「ガウス分布ではない」という主張は価値がない
- ガウス分布の補集合が大きすぎる(非ガウス分布が多種類ありすぎる)ため、情報の価値が低く、後ろ向きであり、無責任である



【アンサンブルデータのモデル選択】

- 正規分布モデル

$$L(\{\mathbf{x}(m)\} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \prod_{m=1}^M \phi(\mathbf{x}(m); \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

- 正規混合分布モデル

$$L(\{\mathbf{x}(m)\} | \pi_1, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \pi_2, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) = \prod_{m=1}^M \{\pi_1 \phi(\mathbf{x}(m); \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) + \pi_2 \phi(\mathbf{x}(m); \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)\}$$

- ヒストグラムモデル

高さ: $n(i) / \{n \Delta a\}$

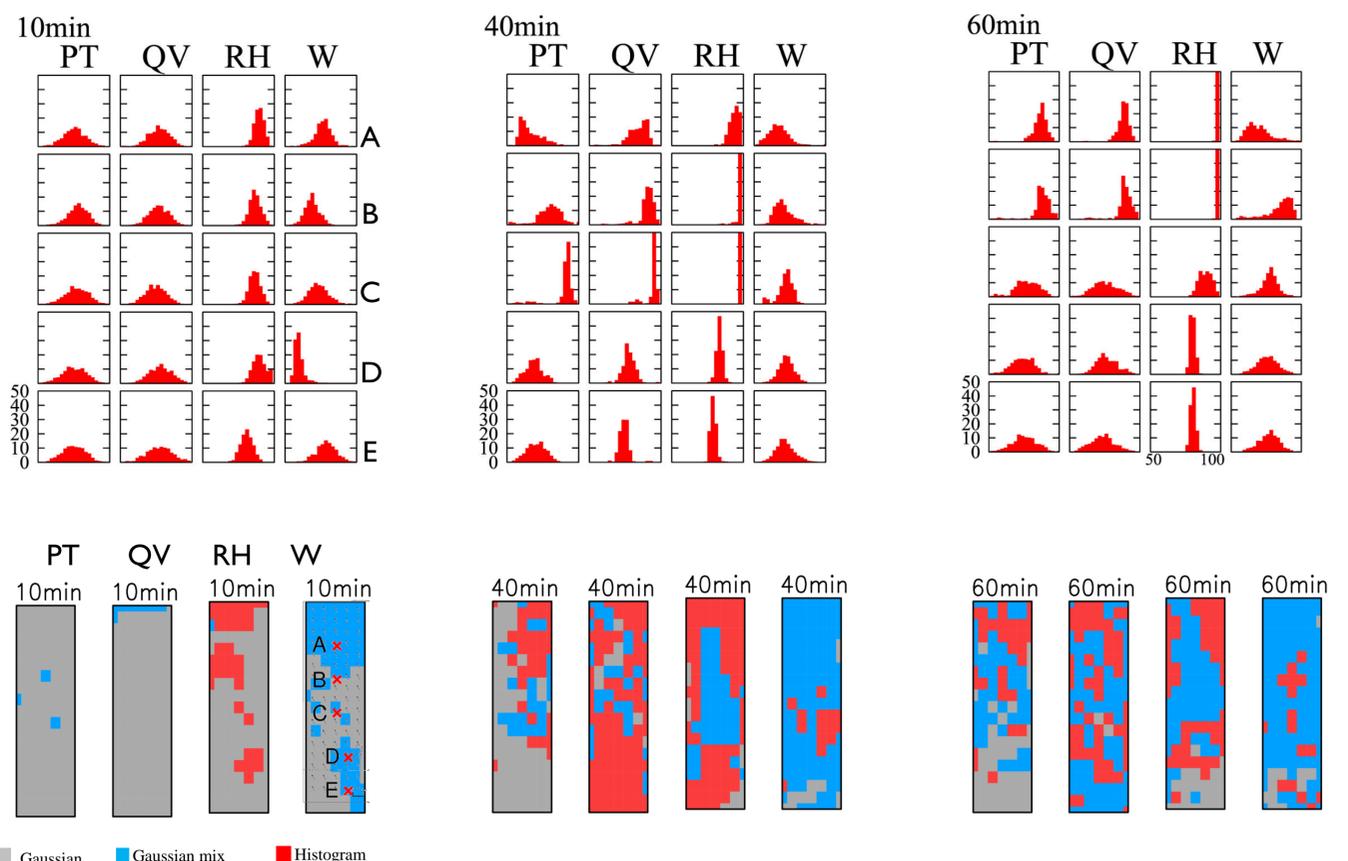
底辺: Δa

$$L(\{\mathbf{x}(m)\} | \{p(i)\}) = \prod_{m=1}^M I(\mathbf{x}(m) \in [a_{i-1}, a_i]) \left(\frac{p(i)}{\Delta a} \right)$$

【同化実験】

気象庁数値予報モデルと粒子フィルタにより推定された、積乱雲の発達過程で見られる各種物理量の時間変化(温度PT、水蒸気量QV、湿度RH、鉛直風W)。A-Eは解析領域を北から南に向かって選んだ地点。

積乱雲発達時の各種物理量の確率分布



情報量規準により選択された、各確率分布の統計モデル(正規分布、正規混合分布、ヒストグラムから選択)。

粒子フィルタにおいては、粒子数は1000、退化を防ぐために適応的観測誤差推定を導入している。上向きの鉛直風(上昇気流)の確率分布に非ガウスの形状がまず見られ、続いて水蒸気量・温度の確率分布の形状の変化が現れることが示されている。

【参考文献】

Kawabata, T., and G. Ueno, 2020: Non-Gaussian Probability Densities of Convection Initiation and Development Investigated Using a Particle Filter with a Storm-Scale Numerical Weather Prediction Model. *Mon. Wea. Rev.*, **148**, 3–20, <https://doi.org/10.1175/MWR-D-18-0367.1>.