

## 第2章 メソアンサンブルガイダンス

### 2.1 概説<sup>1</sup>

気象庁は日々の天気予報、防災気象情報および航空気象情報の作成支援のため、メソアンサンブル予報システム (MEPS) (第1章参照) に基づくガイダンス (MEPS ガイダンス) の開発を進めてきた。MEPS ガイダンスは、MEPS の部内試験運用開始後に、風、気温、降水などのガイダンスおよび航空悪天 GPV について、順次、部内試験運用を開始し、2019年10月現在までに表 2.1.1 に示す MEPS ガイダンスを試験運用している (高田 2017, 2018a; 三戸 2018)。

ガイダンスを作成する目的は2つある。1つは、数値予報モデルが直接予測できない要素 (例えば天気や発雷確率など) を予測するためである。もう1つは数値予報モデルの地上気温や降水量などの系統誤差を減らし、予測精度を向上させるためである (高田 2018b)。MEPS ガイダンスにおいては、こうした通常の決定論的なガイダンスの目的に加えて、アンサンブル予報を利用することで、気象予測の不確実性をアンサンブルメンバーの予測のばらつき具合で表現できるため、信頼度情報などの確率的な統計量を予測に活用できる。例えば降水や風の場合、数時間から39時間先の大雨や暴風などの災害をもたらす顕著現象の出現頻度分布などの指標や確率情報が得られ、MEPS ガイダンスの高度な活用が期待できる。

部内試験運用中の MEPS ガイダンスは、MEPS の各アンサンブルメンバーを入力とするガイダンスであり、基本的には MSM ガイダンスの予測式をそのまま適用して作成する (三戸 2018)。この作成方法では、MEPS ガイダンスのコントロールランと MSM ガイダンスが同一となるため、MSM ガイダンスの気象予測の不確実性に基づくガイダンスのアンサンブルメンバーの各予測が得られ、ガイダンスのアンサンブル平均、超過確率、スプレッドなど、MSM ガイダンスの信頼度情報として、様々な統計量が活用できる。このうち、ガイダンスのアンサンブル平均は、これまでの調査から MSM ガイダンスよりも多くの場合で予測精度が向上することを確認しているが、その一方で最大降水量ガイダンスなどのように、現象の最大値や最小値を予測するガイダンスでは、むしろ予測精度が低下することが分かっている (工藤 2018)。MEPS ガイダンスのアンサンブルメンバーの予測の有効な利用方法は、予測要素ごとに異なるため、それぞれの特性を十分に理解して利用していただきたい。

本章では、第 2.2 節～第 2.4 節に気温、降水ガイダンス、航空悪天 GPV について、第 2.5 節に降雪、風、発雷確率ガイダンスについて、その仕様や作成方法、

表 2.1.1 部内試験運用中の MEPS ガイダンス一覧

天気予報ガイダンス	
ガイダンス名	予測要素
降水ガイダンス	最大降水量、平均降水量
降雪ガイダンス	降雪量 (格子)、 降水種別、降雪量 (地点)
風ガイダンス	定時風、最大風速、 最大瞬間風速
気温ガイダンス	時系列気温、最高、最低
発雷確率ガイダンス	発雷確率
航空気象予報ガイダンス	
ガイダンス名	予測要素
降雪ガイダンス	降雪量
風ガイダンス	定時風、最大風速、 ガスト発生確率、 ガスト風速
気温ガイダンス	時系列気温、最高、最低
視程ガイダンス	視程、視程確率
雲ガイダンス	雲量、雲底高度、雲底確率
発雷確率ガイダンス	発雷確率
航空悪天 GPV	乱気流指数、 着水指数、 積乱雲量・雲頂高度、 圏界面気圧、 気温、風、湿度等の一般要素

統計検証、事例検証および利用上の留意点などを解説する。なお、表 2.1.1 の視程と雲ガイダンスについては、現在調査中のため、本章には記載していない。今後、MEPS ガイダンスは、部内試験運用によって、プロダクトの有効性が確認できた予測要素について、アンサンブルの各メンバーのデータ提供を計画している。MEPS ガイダンスの予報作業での利用にあたり、本章の内容を役立てていただければ幸いです。

### 参考文献

- 工藤淳, 2018: 手法の分類. 数値予報課報告・別冊第 64 号, 気象庁予報部, 20–23.
- 三戸洋介, 2018: メソアンサンブルガイダンス. 平成 30 年度数値予報研修テキスト, 気象庁予報部, 11–12.
- 高田伸一, 2017: アプリケーション. 平成 29 年度数値予報研修テキスト, 気象庁予報部, 118–119.
- 高田伸一, 2018a: ガイダンスの今後. 数値予報課報告・別冊第 64 号, 気象庁予報部, 233–236.
- 高田伸一, 2018b: ガイダンス概論. 数値予報課報告・別冊第 64 号, 気象庁予報部, 3–8.

<sup>1</sup> 石川 宜広

## 2.2 気温ガイダンス<sup>1</sup>

### 2.2.1 はじめに

MEPSに基づく気温ガイダンス（MEPS 気温ガイダンス）は、アメダス地点および国内の空港地点の気温を予測するガイダンスである。本節では、MEPS 気温ガイダンスの仕様と作成方法、統計検証結果、利用方法および事例検証について解説する。なお、統計検証は、アメダス地点と国内の空港地点の検証結果に大きな違いが見られなかったため、アメダス地点の結果のみを示す。

### 2.2.2 仕様と作成方法

MEPS 気温ガイダンスは、MEPS の各メンバーを入力とする気温ガイダンスであり、以下の2種類である。

1. 時系列気温ガイダンス
2. 最高・最低気温ガイダンス

各ガイダンスの仕様を、表 2.2.1 および表 2.2.2 に示す。時系列気温ガイダンスは毎正時の気温、最高気温ガイダンスは 09~18JST における最高気温、最低気温ガイダンスは 00~09JST における最低気温をそれぞれ予測するガイダンスである。また、時系列気温ガイダンスと最高・最低気温ガイダンスからは、それぞれのガイダンスの各メンバーを単純平均したアンサンブル平均を作成することができる。アンサンブル平均には、ランダム誤差を軽減する効果があるため、MSM の気温ガイダンスの精度を上回る傾向がみられている（高田 2018）。なお、時系列気温ガイダンスおよび最高・最低気温ガイダンスの詳細については、三戸（2018）を参照願いたい。

MEPS 気温ガイダンスの作成手法は、MSM 気温ガイダンスと同じカルマンフィルタ（工藤 2018）を用いるが、摂動ランの係数は学習せず、同じ初期時刻の MSM 気温ガイダンスの係数を用いる。このため、コントロールランの予測は MSM 気温ガイダンスと一致する。説明変数は、基本的には各メンバーのモデル予測値から算出するが、モデル地上気温の前日との差を表す説明変数「前日との気温差」については、以下の違いがある。前日の地上気温は各メンバーの値ではなく、コントロールランの値から算出する。これは、ある摂動ランと前日の同じメンバーの間には直接的な関連がないことから、メンバー毎に気温差を計算するよりもコントロールランの値を適用した方が妥当と考えたためである。

### 2.2.3 統計検証

本項では、MEPS 時系列気温ガイダンスと、MEPS 最高・最低気温ガイダンスの超過確率予測、誤差の大きさが 3℃ 以上および 5℃ 以上であった事例数（以下、それぞれ 3℃ 外し事例数、5℃ 外し事例数）、地

域特性を検証した結果を示す。

検証期間は、2018 年 6 月 5 日から 2019 年 6 月 4 日までの 1 年間で、00, 06, 12, 18UTC の全ての初期時刻を用いた検証結果を示す。以降、2018 年 6 月 5 日から 2019 年 6 月 4 日を通年、2018 年 6 月 5 日から 2018 年 8 月 31 日を夏期、2018 年 12 月 1 日から 2019 年 2 月 28 日を冬期と表す。本項ではコントロールラン（CNTL）に対するアンサンブル平均（MEAN）の検証スコア（RMSE および 3, 5℃ 外し事例数）の改善率を次式で定義する。

$$\text{改善率} [\%] \equiv \frac{\text{CNTL} - \text{MEAN}}{\text{CNTL}} \times 100 \quad (2.2.1)$$

そのほか、本項で統計的検証に用いる確率予測に関する指標は巻末付録 D.4 を参照いただきたい。

#### (1) MEPS 時系列気温ガイダンスの予報時間別検証

ここでは MEPS 時系列気温ガイダンスの精度を評価するため、ガイダンスの各摂動ランおよび、コントロールラン、アンサンブル平均の二乗平均平方根誤差（RMSE）と平均誤差（ME）を示す。検証対象は全アメダス地点、検証期間は通年で予報時間は FT=1~39 を用いる。

図 2.2.1 (a) に MEPS 時系列気温ガイダンスの RMSE を示す。RMSE で見ると、アンサンブル平均が全ての予報時間で最も精度が高く、次いでコントロールラン、摂動ランの順となっている。予報時間の始めはコントロールランとアンサンブル平均の RMSE の差は小さく、予報前半に比べて後半の方がコントロールランに対するアンサンブル平均の改善率が高い。図 2.2.1 (b) に MEPS 時系列気温ガイダンスおよび MEPS モデル地上気温の ME を示す。ガイダンスはモデルの系統的な誤差を補正しているため、ガイダンスの全てのメンバーの ME は、モデルと比較して 0 にかなり近いことが確認できる。

#### (2) MEPS 最高・最低気温ガイダンスの精度検証

ここでは MEPS 最高・最低気温ガイダンスの精度検証として、夏期の翌日の最高気温がある値を上回る、または冬期の翌朝の最低気温がある値を下回る超過確率予測の検証結果、および 3, 5℃ 外し事例数の改善率を示す。超過確率予測の検証では、最高気温は閾値を 25, 30, 35℃、最低気温は -3, 0, 3℃ とした。エクイタブルスレットスコア（ETS）では、閾値を超過すると予測したメンバー数が過半数であった場合に予測ありとして検証を行う。また、気候値ではなく MSM ガイダンスの超過確率予測（MSM ガイダンスで閾値を超える気温を予測した場合を 1、予測しない場合を 0 とする）に対するブライアスキルスコア（BSS）を  $BSS_{MSM}$  とする。

表 2.2.3 に超過確率予測に対する BSS, ETS を示す。夏期の最高気温、冬期の最低気温共に BSS, ETS は正で、MEPS ガイダンスは気候値予測を改善している。

<sup>1</sup> 山下 浩史、ただし第 2.2.3 項は井上 卓也

表 2.2.1 MEPS 時系列気温ガイダンスの仕様

作成対象地点	アメダス、空港
作成方法	カルマンフィルタ
作成数	コントロールランを含む 21 メンバー
予報対象時間	毎正時
予報期間と間隔	FT=1 から FT=39 まで 1 時間間隔
逐次学習の有無	あり (コントロールラン説明変数で学習)
説明変数	モデル予測値 (地上の東・西・南・北風成分、地上風速、地上気温、中・下層雲量、気温減率、降水量、前日との気温差 (摂動の場合は前日のコントロールランとの差))
目的変数	アメダス：アメダスで観測された毎正時の気温 空港：航空気象定時観測気象報 (METAR)、航空気象自動観測気象報 (METAR AUTO) の毎正時の気温
層別化処理の対象	作成対象地点、初期時刻、予報時間、季節 (暖候期 (4~9 月)、寒候期 (10~3 月)) 暖・寒候期の係数切替前に 1 か月間の事前学習を行う。

表 2.2.2 MEPS 最高・最低気温ガイダンスの仕様

作成対象地点	アメダス、空港																				
作成方法	カルマンフィルタ																				
作成数	コントロールランを含む 21 メンバー																				
予報期間 (対象要素)	<table border="1"> <tr> <th>初期値</th> <th>当日</th> <th>翌日</th> <th>翌々日</th> </tr> <tr> <td>00UTC</td> <td>(最高)</td> <td>最高・最低</td> <td></td> </tr> <tr> <td>06UTC</td> <td></td> <td>最高・最低</td> <td></td> </tr> <tr> <td>12UTC</td> <td></td> <td>最高・最低</td> <td>(最低)</td> </tr> <tr> <td>18UTC</td> <td>最高</td> <td>(最高・) 最低</td> <td></td> </tr> </table>	初期値	当日	翌日	翌々日	00UTC	(最高)	最高・最低		06UTC		最高・最低		12UTC		最高・最低	(最低)	18UTC	最高	(最高・) 最低	
	初期値	当日	翌日	翌々日																	
	00UTC	(最高)	最高・最低																		
	06UTC		最高・最低																		
	12UTC		最高・最低	(最低)																	
18UTC	最高	(最高・) 最低																			
「翌日」などは、日本時間で初期時刻から見た日付を示す。( ) 内はアメダスのみ																					
逐次学習の有無	あり (コントロールラン説明変数で学習)																				
説明変数	モデル予測値 (地上の東・西・南・北風成分、地上風速、地上気温、中・下層雲量、気温減率、降水量、前日との気温差 (摂動ランの場合は前日のコントロールランとの差))																				
目的変数	アメダス：アメダスの気温観測 (1 分値) から算出した最高・最低気温 空港：METAR, METAR AUTO, 航空気象特別観測気象報 (SPECI) から算出した最高・最低気温 当日・翌日の、最高気温は 00~09UTC の最高気温、最低気温は 15~00UTC の最低気温。																				
層別化処理の対象	作成対象地点、初期時刻、季節 (暖候期 (4~9 月)、寒候期 (10~3 月))、予報対象要素 (最高気温、最低気温)、予報時間 (日単位) 暖・寒候期の係数切替前に 1 か月間の事前学習を行う。																				

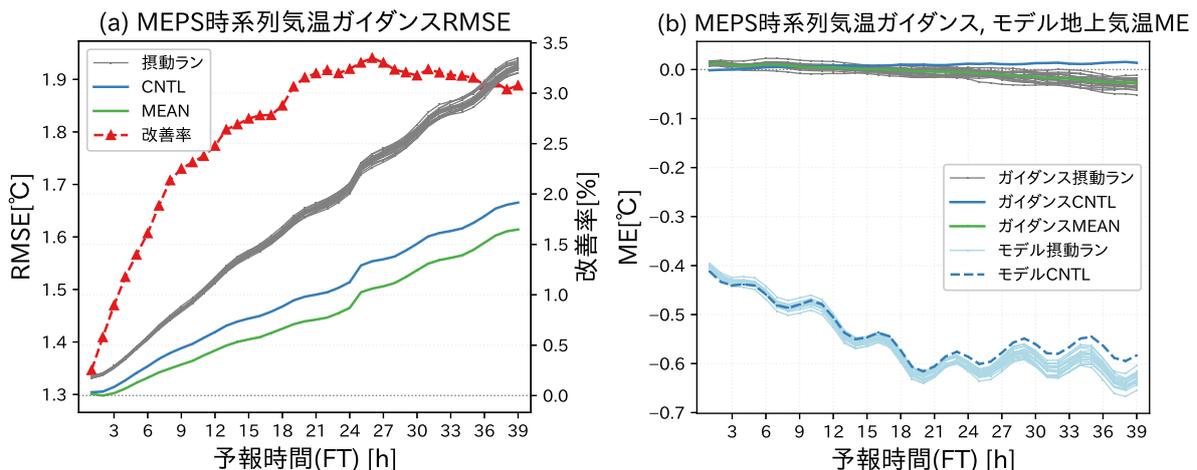


図 2.2.1 予報時間別 (FT) [h] の、(a) MEPS 時系列気温ガイダンスの RMSE [°C] と RMSE の改善率 [%] ((2.2.1) 式)、(b) MEPS 時系列気温ガイダンスと MEPS モデル地上気温の ME [°C]。青実線がコントロールラン、緑実線がアンサンブル平均、灰色実線が各摂動ランを表し、いずれも実線は MEPS 時系列気温ガイダンスを表す。(a) の赤破線はコントロールランに対するアンサンブル平均の RMSE 改善率を表し、(b) の青破線は MEPS モデルのコントロールラン、水色破線は MEPS モデルの各摂動ランの地上気温を表す。検証期間は通年。

表 2.2.3 夏期の最高気温および冬期の最低気温における閾値の超過確率予測に対する BSS, ETS, BSS<sub>MSM</sub>。

(a) 夏期の翌日の最高気温

閾値	25 °C	30 °C	35 °C
BSS	0.72	0.72	0.56
ETS	0.68	0.69	0.52
BSS <sub>MSM</sub>	0.26	0.26	0.22

(b) 冬期の翌朝の最低気温

閾値	-3 °C	0 °C	3 °C
BSS	0.74	0.75	0.75
ETS	0.72	0.72	0.72
BSS <sub>MSM</sub>	0.24	0.24	0.24

特に BSS<sub>MSM</sub> が正であることから、MSM ガイダンスの超過確率予測よりも MEPS ガイダンスの超過確率予測の方が精度が良い。

図 2.2.2 に確率値別出現率図（信頼度曲線）を示す。最高気温の予測と観測出現率の対応は良く、信頼度曲線は skill 領域<sup>2</sup>におおむね位置している。最低気温は最高気温に比べると、予測と観測とで出現率の対応がやや悪く、閾値 -3, 0 °C では信頼度曲線が no-skill line<sup>3</sup>と重なっている部分がある。

表 2.2.4 に、夏期および冬期の翌日の最高気温、翌朝の最低気温について、誤差が 3 °C 以上および 5 °C 以上であった事例数と改善率を示す。夏期、冬期の最高・最低気温はいずれもアンサンブル平均がコントロールランを改善し、3, 5 °C 外し事例数が減少している。

### (3) MEPS 最高・最低気温ガイダンスの地点別検証

MEPS 最高・最低気温ガイダンスの地域特性を検証するため、夏期の翌日の最高気温、冬期の翌朝の最低気温について地点別の RMSE, ME を示す。

図 2.2.3 に RMSE の分布を示す。アンサンブル平均の RMSE の分布は、夏期は北日本の太平洋側沿岸部で誤差が大きく、冬期は北海道で誤差が大きく、コントロールラン（図略、三戸 (2018) を参考）と同様となっている。RMSE の改善率は全国的に正で、夏期の最高気温で約 4%、冬期の最低気温で約 3%となっているが、夏期の最高気温では、岩手県や宮城県の沿岸部で改善となっている。例えば宮城県の RMSE 改善率は県平均で約 0%であるのに対し、太平洋に面していて改善率の高い千葉県は約 7%である。また、アンサンブル平均がコントロールランに対して改善していない地点には、沿岸部や島が多い。MEPS モデルで海面水温には摂動を与えていないため、島など海に面した地点では

<sup>2</sup> BSS に正の寄与を持つ領域のこと。図 2.2.2 では灰色で示されている。

<sup>3</sup> 信頼度と分離度への寄与が等しい大きさを持ち、BSS への寄与は 0 となる直線のこと。

内陸部に比べて地上気温のスプレッドが過小評価されて、アンサンブル平均の改善率が中立となることが考えられる。このように、アンサンブル平均は全国平均で精度の改善が見られるものの、改善率には地域差があることに留意が必要である。

図 2.2.4 に ME の分布を示す。アンサンブル平均の ME は、RMSE の分布もそうであったように、夏期および冬期いずれもコントロールラン（図略）と同様である。ただし、アンサンブル平均からコントロールランを減算した ME の差分は地域によって異なっていることがわかる。例えば夏期の関東平野では、アンサンブル平均の最高気温予測値はコントロールランに比べて低い。また、冬期の ME は全国的に負バイアスだが、北海道の内陸部で正バイアスとなっている。このことについて、冬期の MSM 気温ガイダンスは、北海道内陸部で放射冷却による低温を十分に補正できない特性があり（三戸 2018）、アンサンブル平均も同様の特性を持っている。

### 2.2.4 利用方法

MEPS 気温ガイダンスのアンサンブル平均は、コントロールランより統計的には精度が良いため、そのまま利用することで現在運用中の MSM 気温ガイダンスに比べて、精度の良い予報をすることができ、大外しを軽減することができる。しかしながら、アンサンブル平均だけでは、どの程度の誤差が生じるかという情報を得ることはできない。そこで、全メンバーを概観する手法が有効である。例えば、以下の 3 つの方法がある。

1. スパゲッティ・ダイアグラム（第 2.2.5 項 (1) の図を参照）
2. 箱ひげ図（第 2.2.5 項 (1) の図を参照）
3. 超過確率（第 2.2.5 項 (3) の図を参照）

MEPS 気温ガイダンスのスパゲッティ・ダイアグラムは、特定の地点において、すべてのメンバーのガイダンスを時系列上で重ねた図で、メンバーのばらつき具合を概観できる。箱ひげ図は、中央値、最大値、最小値、第 1 四分位点、第 3 四分位点を表す図で、第 1 四分位点と第 3 四分位点の間を箱とする。箱ひげ図ではメンバーのばらつき具合や分布の偏り具合を定量的に可視化できる。超過確率では、ある閾値の気温を超える、あるいは下回るメンバーの割合を概観できる。これらの図を利用することにより、アンサンブル平均がコントロールランや各摂動ランに対して、どこに位置しているかを手早く確認できる。アンサンブル平均がコントロールランや各摂動ランから外れている場合は、気圧配置の変化や実況の推移、前初期値の MEPS モデルや GSM モデルの予測などをもとに最も実現確率の高い予測を選択することになる。また、第 2.2.3 項 (1) の結果から、MEPS 気温ガイダンスの各摂動ランは、MSM 気温ガイダンスよりも精度が悪いため、単

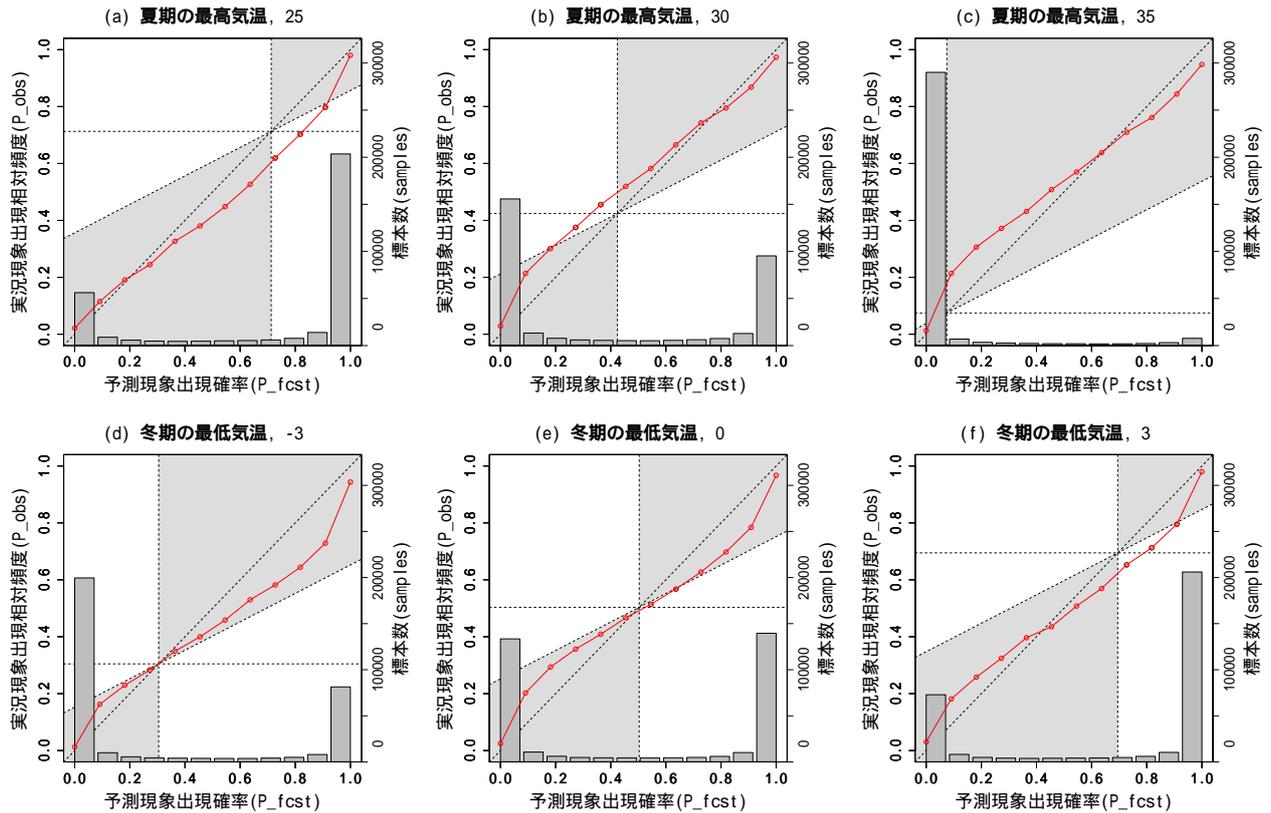


図 2.2.2 最高・最低気温ガイダンスの確率値別出現率図（信頼度曲線）。上段が夏期の翌日の最高気温、下段が冬期の翌朝の最低気温。閾値は (a) が 25 °C, (b) が 30 °C, (c) が 35 °C, (d) が -3 °C, (e) が 0 °C, (f) が 3 °C。縦軸左は実況現象出現相対頻度、縦軸右が標本数、横軸が予測現象出現確率（閾値の超過を予測した事例の出現確率）。赤線が信頼度曲線で、棒グラフが出現標本数を表す。

表 2.2.4 夏期および冬期における翌日の最高気温、翌朝の最低気温の 3, 5 °C 外し事例数。CNTL はコントロールラン、MEAN はアンサンブル平均。改善率は (2.2.1) 式を用いた。

		3 °C 外し事例数			5 °C 外し事例数			全観測事例数
		CNTL	MEAN	改善率 [%]	CNTL	MEAN	改善率 [%]	
夏期	翌日の最高気温	27,690	24,833	10.3	3,765	2,756	26.8	325,979
	翌朝の最低気温	8,207	7,307	11.0	517	410	20.7	326,050
冬期	翌日の最高気温	11,998	11,204	6.6	1,771	1,270	28.3	333,208
	翌朝の最低気温	33,396	30,895	7.49	6,770	5,787	14.5	333,222

独での利用は推奨しない。

## 2.2.5 事例検証

### (1) アンサンブル平均の予測がコントロールランの予測を改善した事例

アンサンブル平均の予測がコントロールランの予測を改善した事例として、2018 年 9 月 12 日 00UTC 初期値の 9 月 13 日を予測対象とした長野県南信濃の事例を示す。図 2.2.5 (a) は、2018 年 9 月 13 日 00UTC の日本域地上天気図である。長野県南信濃では高気圧後面にあたり、11~12JST にかけて日照があったものの、前線の影響を受けて、おおむね曇りであった。図 2.2.5 (b) と (c) に、南信濃の MEPS 気温ガイダンスのスパゲッティ・ダイアグラムと箱ひげ図をそれぞれ

示す。時系列気温ガイダンスの FT=24~33 に着目すると、ばらつきが大きいことが分かる。これは、日本の東からの高気圧の西への張り出しの強弱により、晴れおよび曇や雨のメンバーが見られるなど、天気には差があったためである。その中において、アンサンブル平均がコントロールランに比べて、観測値に近い。このように、本事例ではアンサンブル平均の予測がコントロールランの予測を改善し、第 2.2.3 項 (1) で示した結果が表れている。

### (2) シナリオが 2 つに分かれる事例

MEPS の予測シナリオが 2 つに分かれた事例として、2018 年 12 月 11 日 12UTC 初期値の東京の事例を示す。図 2.2.6 (a) は、2018 年 12 月 12 日 00UTC の日本域

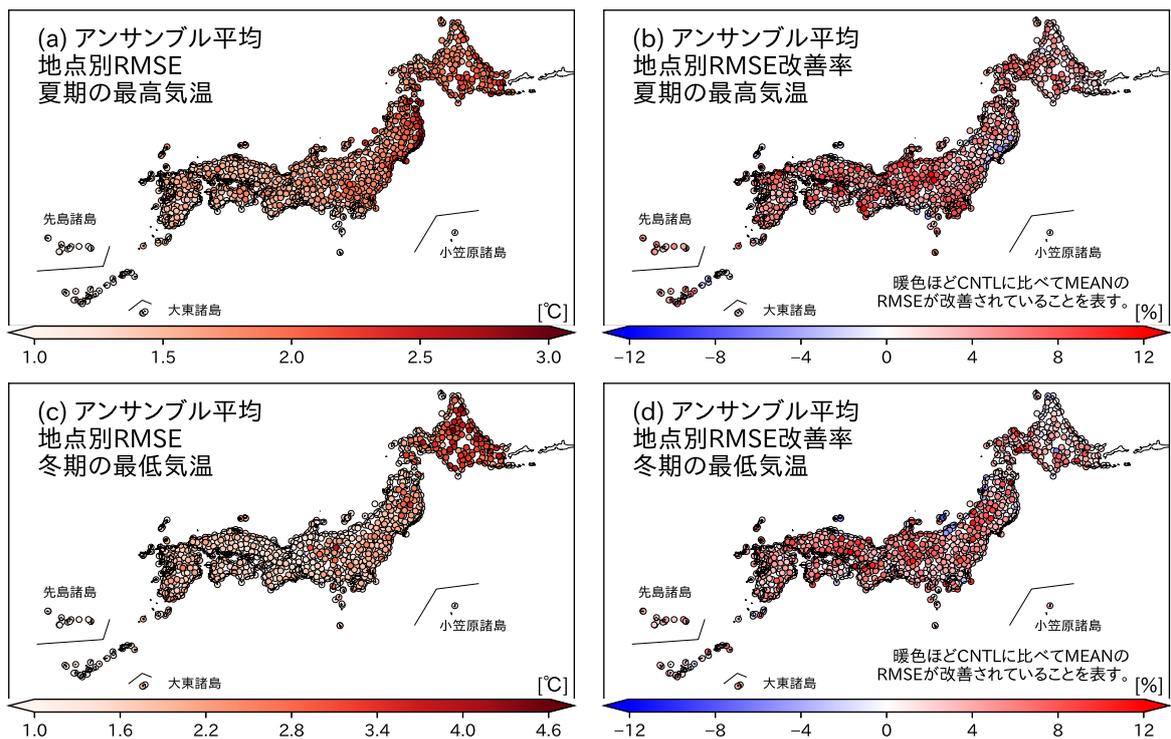


図 2.2.3 夏期の翌日の最高気温および冬期の翌朝の最低気温の地点別 RMSE 分布図。上段は夏期の翌日の最高気温、下段は冬期の翌朝の最低気温で、左列はアンサンブル平均の RMSE[°C]、右列はアンサンブル平均のコントロールランに対する RMSE 改善率 [%] を表す。夏期、冬期共に、コントロールランの RMSE[°C] の分布はそれぞれのアンサンブル平均 (a), (c) と同様であったため、図は省略した。おおむね同様の条件の MSM ガイダンスの分布図として、三戸 (2018) を参考にさせていただきたい。

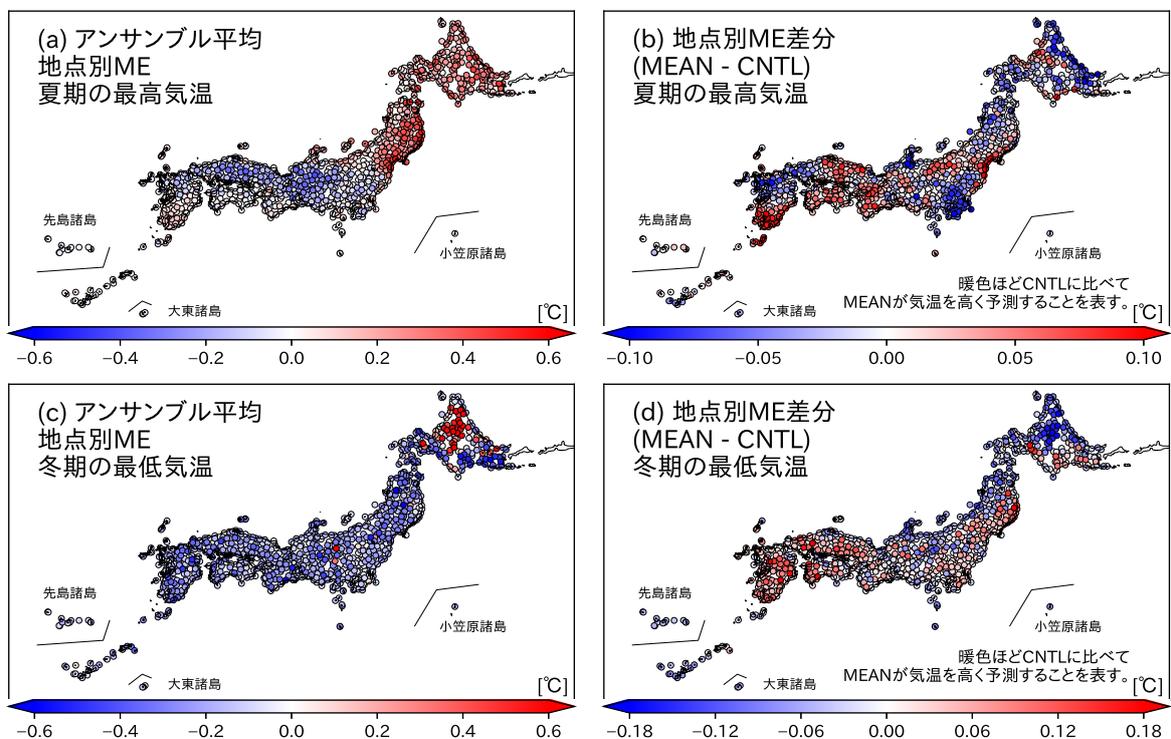


図 2.2.4 夏期の翌日の最高気温および冬期の翌朝の最低気温の地点別 ME 分布図。上段は夏期の翌日の最高気温、下段は冬期の翌朝の最低気温で、左列はアンサンブル平均の ME[°C]、右列は ME の差分 (アンサンブル平均 - コントロールラン) [°C] を表す。図 2.2.3 と同様に、コントロールランの ME の分布はアンサンブル平均と同様であったため、図は省略した。おおむね同様の条件の MSM ガイダンスの分布図として、三戸 (2018) を参考にさせていただきたい。

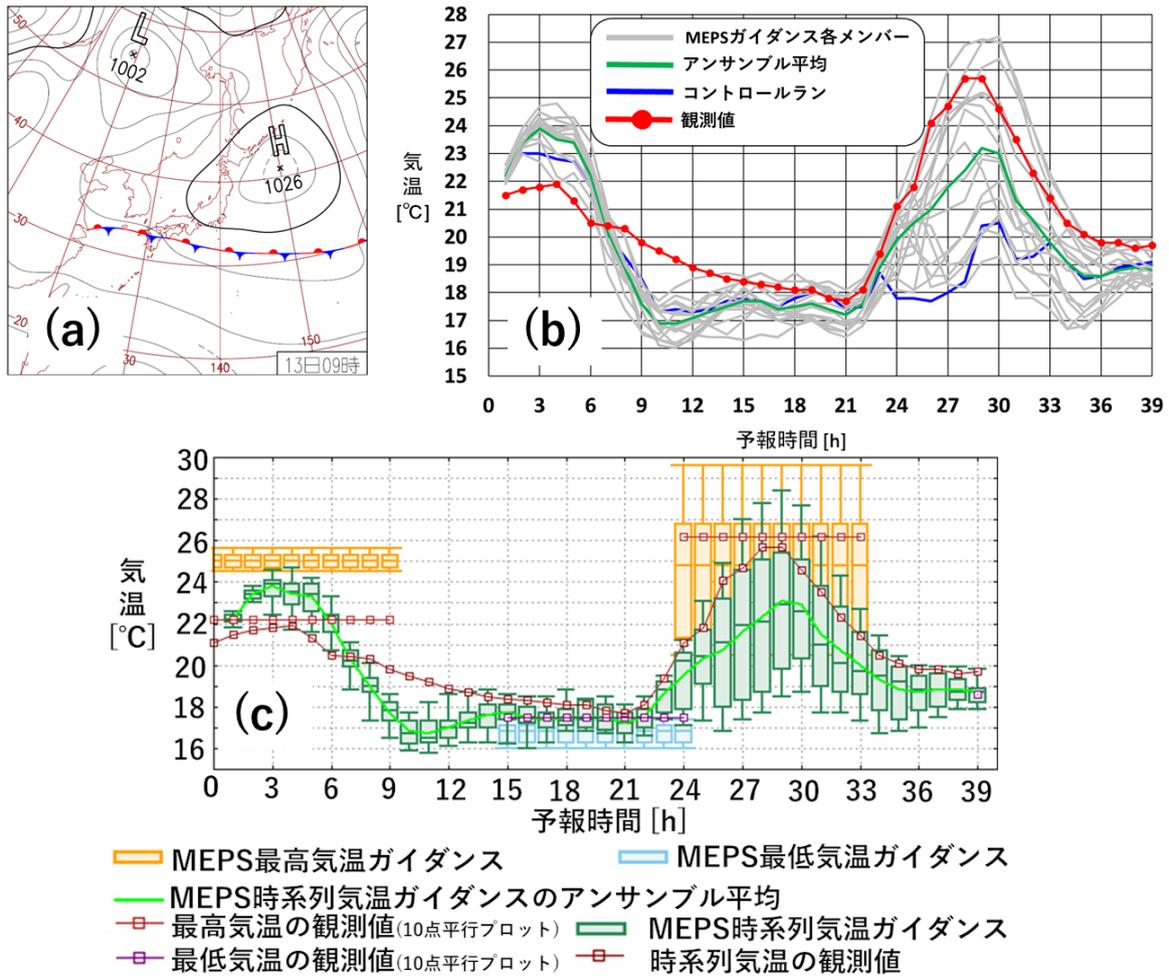


図 2.2.5 (a) 2018 年 9 月 13 日 00UTC の日本域地上天気図、(b) 2018 年 9 月 12 日 00UTC 初期値の長野県南信濃を対象とした時系列気温ガイダンスのスパゲッティ・ダイアグラム、(c) 2018 年 9 月 12 日 00UTC 初期値の長野県南信濃を対象とした時系列気温ガイダンスの箱ひげ図。

地上天気図である。前線を伴った発達中の低気圧が本州の南岸を、また、別の低気圧が日本海を東進している。図 2.2.6 (b) は、東京における MEPS 時系列気温ガイダンスの予測と実況を示す。FT=12~21 に着目すると、コントロールランよりも気温予測が高いシナリオと低いシナリオに分かれていることが分かる。アンサンブル平均は、全メンバーの分布の中央にあり、コントロールランに近い値を予測している。そこで、気温予測が高いアンサンブルメンバー 6 と低いアンサンブルメンバー 4 のモデル予測を確認する。図 2.2.6 (c) と (d) は、MEPS の FT=15 におけるメンバー 6 および 4 の地上気圧、前 3 時間降水量および地上風である。メンバー 6 では、大陸からの高気圧の張り出しにより、北西風が卓越する場となり、天気回復の予測となっている。一方で、メンバー 4 では、上空の気圧の谷が西日本に残り、天気回復が遅れる予測となっている。実況は、図 2.2.6 (a) の天気図の通り、メンバー 4 の予測に近く、アンサンブル平均やコントロールラン予測は

実況から外れる結果となった。このような場合、実況の推移、前初期値の MEPS モデルや GSM モデルの予測などを併用して気温予測を行うことが推奨される。それでも気温予測が難しい場合は、アンサンブル平均を利用することで大外しを軽減できる。

### (3) フェーン現象による高温事例

MEPS のばらつきが不十分だった事例として、2019 年 5 月 25 日 12UTC 初期値の北海道佐呂間の事例を示す。図 2.2.7 (a) は、2019 年 5 月 26 日 00UTC の日本域地上天気図である。2019 年 5 月 26 日は、東北地方から北海道にかけて大陸で形成された暖気が入り込んだことと下層の西風が北見山地を越えたことによるフェーン現象の影響で、北海道東部を中心に高温となった (図 2.2.7 (b))。特に、佐呂間では午後 2 時過ぎに 5 月として全国の史上最高気温となる 39.5℃を観測した。佐呂間における GSM、MSM および MEPS の各最高気温ガイダンスはいずれも 35℃以上の猛暑日を予測していた。ただし、39℃以上を予測できた最高気温

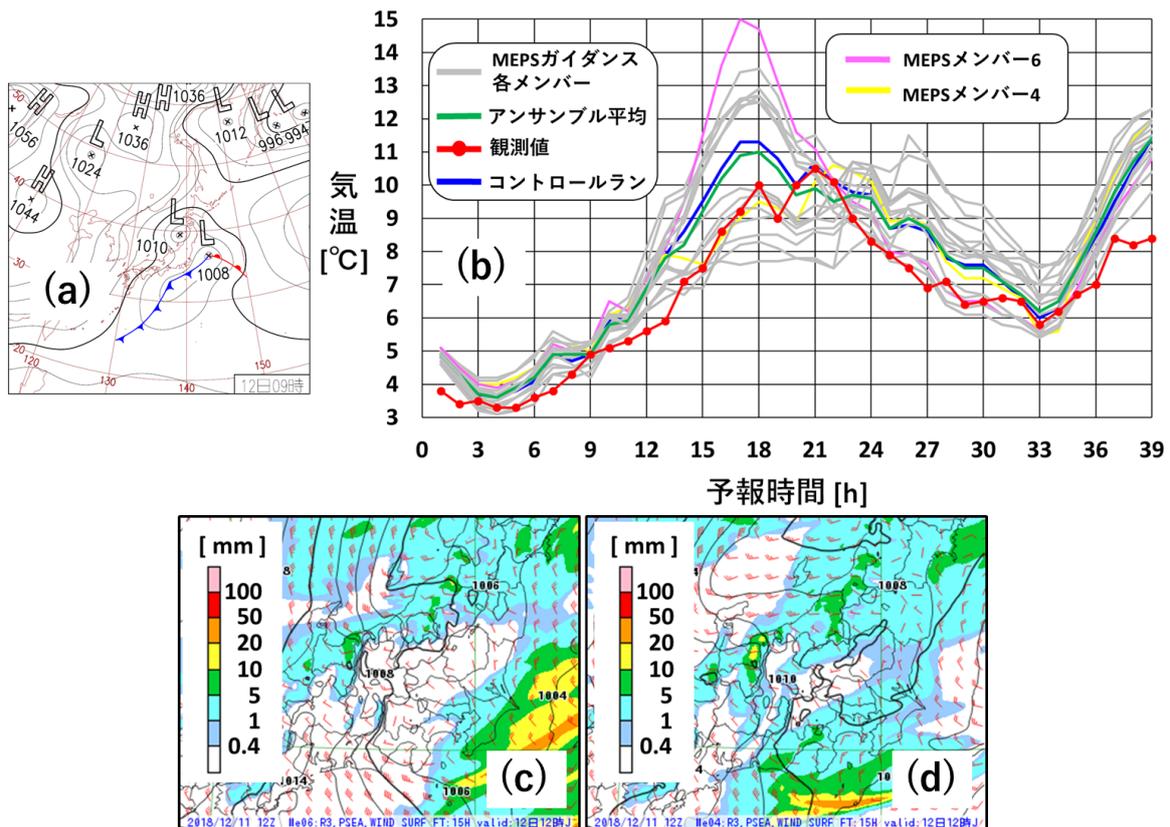


図 2.2.6 (a) 2018 年 12 月 12 日 00UTC の日本域地上天気図、(b) 2018 年 12 月 11 日 12UTC 初期値の東京を対象とした MEPS 時系列気温ガイダンス、(c) と (d) は 2018 年 12 月 11 日 12UTC 初期値の FT=15 における MEPS メンバー 6 とメンバー 4 のモデル予測値の前 3 時間降水量 [mm/3h]、海面更正気圧 [hPa]、地上風 [ノット]。

ガイダンスはなく、MEPS 最高気温ガイダンスの 1 つのメンバーにおける 37.9 °C の予測が最高値であった。気温の急上昇から最高気温が観測された時間帯である FT=8~18 に着目すると、GSM, MSM および MEPS の各時系列気温ガイダンスの予測は、観測値に比べ低かった (図 2.2.7 (c)) が、GSM や MSM のモデル値と比較すると観測値に近づく修正 (例えば 13JST でおおよそ 2~7 °C) を行っていた。MEPS 時系列気温ガイダンスのメンバーの分布に着目すると、気温が上昇する時間帯はばらつきが小さいが、30 °C を超える時間帯からばらつきが大きくなる傾向が見られた。アンサンブル平均は MSM ガイダンスとほぼ同じであった。MEPS 気温ガイダンスの予測特性は MSM 気温ガイダンスと同様であり、MSM ガイダンスにおける極端な高温を予測しづらい傾向があるという利用上の留意点 (三戸 2018) は MEPS 気温ガイダンスにも当てはまる。

続いて、本事例での超過確率の利用事例を紹介する。MEPS の気温ガイダンスから超過確率を利用することで、猛暑日の基準となる 35 °C を超える確率がどの時間帯に出現するか確認できる (図 2.2.7 (d))。図から読み取れることとして、12JST から 15JST までに 35 °C を超える可能性があり、超過確率は最大 40 % と見る

ことができる。

#### (4) オホーツク海高気圧の張り出しによる低温事例

2018 年 6 月 17 日、北日本から関東はオホーツク海に中心を持つ高気圧圏内で気温が上がらず (図 2.2.8 (a), (b))、最高気温は 4 月下旬から 5 月上旬並となり、北海道では平年より 8 °C 以上低い所があった。根室を例に GSM, MSM および MEPS の各ガイダンスの予測結果を確認する (図 2.2.8 (c))。図から、どのガイダンスもモデルよりも高めの気温を予測していたことが確認できる。MEPS ガイダンスに着目すると、実況を予測できていたメンバーはなかった。MEPS の各時系列気温ガイダンスに着目すると、(3) の事例に比べて、ばらつきが小さく、アンサンブル平均が MSM ガイダンスとほぼ同じであった。本事例は、実況では下層雲に覆われて終日気温が低かったのに対し、モデルでは下層雲の予測が不十分で、MSM のお天気マップでは、おおむね晴れを予測していた (図略)。この下層雲は MEPS のメンバーでも予測できておらず (図略)、MEPS 気温ガイダンスの予測が大きく外れる要因となった。なお、GSM のお天気マップでは、おおむね曇りを予測していたため、実況値に近い予測となっていた。モデルの予測が苦手な事例では MEPS 気温ガイダンスの予測

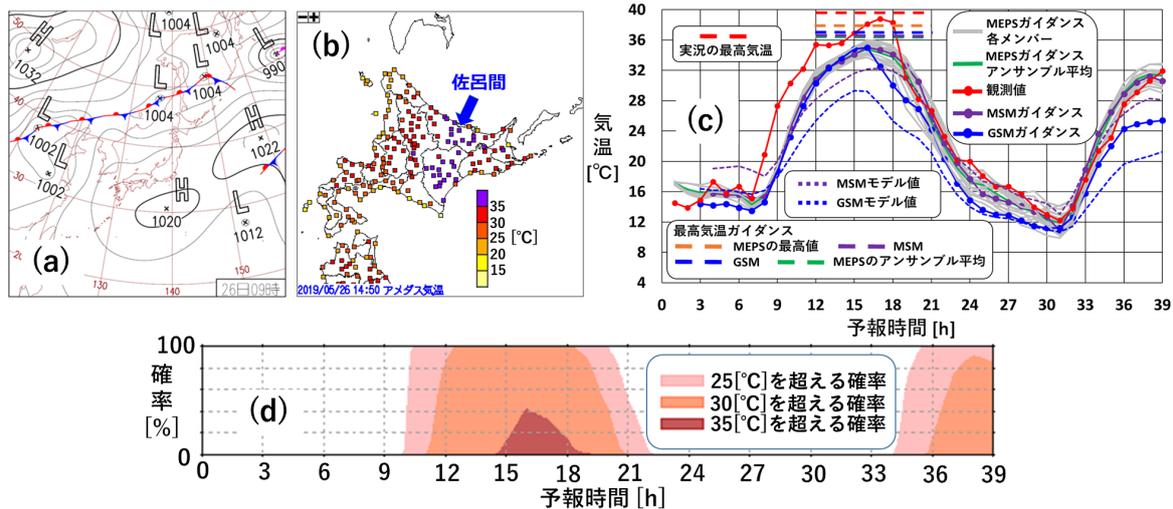


図 2.2.7 (a) 2019 年 5 月 26 日 00UTC の日本域地上天気図、(b) 2019 年 5 月 26 日 0550UTC の北海道のアメダス地上気温分布、(c) 2019 年 5 月 25 日 12UTC 初期値の佐呂間における GSM と MEPS 気温ガイダンスの予測および、5 月 25 日 15UTC 初期値の MSM 気温ガイダンスの予測と実況、(d) 2019 年 5 月 25 日 12UTC 初期値の佐呂間の MEPS 時系列気温ガイダンスの超過確率。

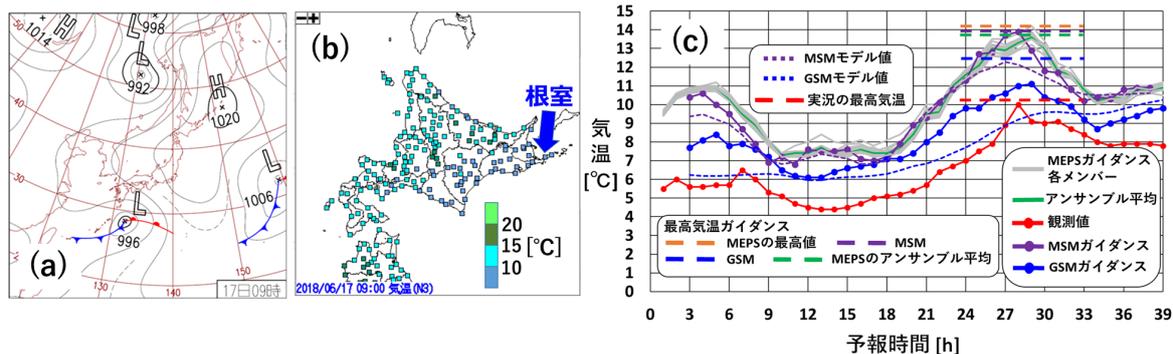


図 2.2.8 (a) 2018 年 6 月 17 日 00UTC の日本域地上天気図、(b) 2018 年 6 月 17 日 00UTC の北海道のアメダス地上気温分布、(c) 2018 年 6 月 16 日 00UTC 初期値の根室における GSM と MEPS 気温ガイダンスの予測および、2018 年 6 月 16 日 03UTC 初期値の MSM 気温ガイダンスの予測と実況。

も不十分になることを示している。

### 2.2.6 まとめ

最後に、MEPS 気温ガイダンスの前項までのまとめを以下に記す。MEPS 気温ガイダンスを利用する際の参考にしていただきたい。

#### 仕様と作成手法

1. 時系列気温ガイダンスおよび最高・最低気温ガイダンスを作成する。
2. 作成手法はカルマンフィルタで MSM ガイダンスと同じだが、摂動ランの係数は学習せず、同じ初期時刻の MSM 気温ガイダンスの係数を用いる。

#### 統計検証と利用方法

1. MEPS 気温ガイダンスのアンサンブル平均は MSM 気温ガイダンスの予測精度を上回り、最高・最低

気温の大外し事例数も減少するため、アンサンブル平均の利用を推奨する。

2. 最高・最低気温ガイダンスのアンサンブル平均は、沿岸部や島を中心に MSM 気温ガイダンスを改善できないアメダス地点があるなどの地域差が存在するため、注意する必要がある。
3. 最高・最低気温ガイダンスのアンサンブル平均の ME は、夏期は東日本で正バイアス傾向、西日本は負バイアス傾向、冬期は全国的に負バイアス傾向（北海道内陸部を除く）であるので、注意する必要がある。
4. MEPS 気温ガイダンスの各摂動ランは、MSM 気温ガイダンスよりも精度が悪いため、単独での利用は推奨しない。

## 事例検証

1. シナリオが2つに分かれる事例の対処法は、実況の推移、前初期値のMEPSモデルやGSMモデルの予測などを併用して気温予測を行うことが推奨される。それでも気温予測が難しい場合は、アンサンブル平均を利用することで大外しを軽減できる。
2. フェーン現象による高温事例は、MEPS気温ガイダンスのアンサンブル平均がMSM気温ガイダンスとほぼ同じであった。MEPS気温ガイダンスの予測特性がMSM気温ガイダンスと同様であることから、極端な高温を予測しづらい傾向があるという利用上の留意点(三戸 2018)は、MEPS気温ガイダンスにも当てはまる。
3. オホーツク海高気圧の張り出しによる低温事例では、モデルの下層雲予測が不十分で、初期の段階から観測値から外れ、MEPS気温ガイダンスの予測も大きく外れる要因となった。モデルの予測が苦手な事例では、MEPS気温ガイダンスの予測も不十分になることを示している。

## 参考文献

- 工藤淳, 2018: カルマンフィルタ. 数値予報課報告・別冊第 64 号, 気象庁予報部, 62-70.
- 三戸洋介, 2018: 気温ガイダンス. 数値予報課報告・別冊第 64 号, 気象庁予報部, 132-143.
- 高田伸一, 2018: 統合型ガイダンス. 数値予報課報告・別冊第 64 号, 気象庁予報部, 234-235.

## 2.3 降水ガイダンス<sup>1</sup>

### 2.3.1 はじめに

降水ガイダンスは、日々の天気予報から警報・注意報といった防災気象情報の発表のための予測資料として、また降雪量ガイダンス等の入力値としても用いられる等、予報作業において幅広く使われている。MEPSの現業予報作業における活用法としては、顕著現象のポテンシャル把握や早期注意情報（警報級の可能性）の判断等への活用が検討されてきており（河野・西本 2018）、MEPS 降水ガイダンスにおいても降水量予測がある閾値を超える確率（超過確率）が有効な資料の1つとして挙げられている（三戸 2018）。本節では、MEPS 降水ガイダンスと、それを入力にして得られる超過確率をはじめとした各種統計量について解説する。まず、第2.3.2項でMEPS 降水ガイダンスの仕様と作成方法を説明し、続いて第2.3.3項と第2.3.4項で統計検証と事例検証の結果を示す。最後に、第2.3.5項と第2.3.6項で本節のまとめと今後の展望について述べる。

### 2.3.2 仕様と作成方法

MEPS 降水ガイダンスの仕様は、コントロールランを含めて21メンバー分のガイダンスを作成する点と、初期時刻が00, 06, 12, 18UTCの6時間毎で予報時間が初期時刻に依らずいずれもFT=39までである点を除けば、MSM 降水ガイダンスと同一である。また、作成方法についても基本的にMSM 降水ガイダンスと同一である。MSM 降水ガイダンスの仕様や作成方法および作成に用いる統計手法の詳細については、白山 (2018) および工藤ほか (2018) を参照いただきたい。MEPS 降水ガイダンスの作成には、各メンバーについてMEPSのモデル予測値およびそれらを変換した値を説明変数に用いている。ただし、カルマンフィルタや頻度バイアス補正等で用いる係数は、いずれのメンバーの予測に対してもMSM 降水ガイダンスで学習・更新して作成したものを利用する。また、ニューラルネットワークや線形重回帰にて用いる予測式についても、MSM 降水ガイダンスの予測式と同一である。したがって、MEPS 降水ガイダンスのコントロールランの予測はMSM 降水ガイダンスの予測と一致する。

### 2.3.3 統計検証

第2.3.2項で述べたように、MEPS 降水ガイダンスは各種統計処理の中で用いる係数や予測式をMSM 降水ガイダンスと共有しており、MEPS 降水ガイダンスのコントロールランの予測はMSM 降水ガイダンスの予測と一致する。MSM 降水ガイダンスは2019年7月4日に改良を行っている（詳細は第5.1節を参照）ため、この変更を反映したMEPS 降水ガイダンスについて、1年間分（2018年6月5日から2019年6月4日

まで）の統計検証の結果を示す。検証には、MEPSのモデル降水量およびMEPS 降水ガイダンスの各予測要素について、いずれの検証でも全初期時刻（00, 06, 12, 18UTC）分のデータを用いる。検証に用いる予報時間は、MEPSの3時間降水量予測（FRR3）やMEPS 降水ガイダンスの3時間平均降水量（MRR3）、1時間最大降水量（RMAX31）および3時間最大降水量（RMAX33）については、FT=3からFT=39まで、24時間最大降水量（RMAX24）については、FT=24からFT=39までの3時間おきである。検証対象格子は、それぞれ全国の陸上および周辺の海上格子である（詳細は白山 (2018) を参照）。MEPS 降水ガイダンスから得られるアンサンブル平均（21メンバーの予測結果を格子ごとに平均）やアンサンブル最大（21メンバーの予測結果の最大値を格子ごとに抽出）、超過確率の検証についてもこれに準じる。

また、本項では先述の検証期間の中で台風事例とそれ以外（以下、非台風事例）で分けてそれぞれ検証を行っている。これは、台風事例と非台風事例とではMEPS 降水ガイダンスの予測特性が大きく異なるためである。表2.3.1に、本節の検証において台風事例として扱った期間を示す。抽出した期間は、いずれも台風が日本に接近または上陸し、広い範囲に強い降水をもたらした時刻周辺の日に絞っており、台風の影響による先行降水の事例等は必ずしも含まれない。

#### (1) MEPS 降水ガイダンス

図2.3.1に、MRR3とFRR3の閾値別バイアスコア（BI）およびエクイタブルスレットスコア（ETS）の比較結果を示す。まず、モデルとガイダンスのコントロールランを比較すると、非台風事例では、モデルがもつ予測頻度の過多傾向をガイダンスが抑制し、特に80 mm/3h以下の降水については概ね適切な予測頻度となっている。また、ETSも概ね80 mm/3h程度までは改善または中立で、予測精度が向上している。一方、台風事例では、80 mm/3h以下の降水についてはガイダンスの予測頻度がモデルよりも大幅に過多傾向となっており、ETSもモデルと比較して劣っている。100 mm/3h以上の降水ではモデルとのETSの差が更に大きくなっており、本検証期間の台風事例ではガイダンスがモデルを適切に修正できていない場合が多かったといえる。この傾向は、2017年の通年を対象とした同様の検証でも、ガイダンスとモデルの精度差は比較的小さかったものの同様であった（図略）。T1821が25年ぶりに非常に強い勢力で上陸した台風であったこと、T1812が近畿地方に上陸後、西日本を東から西に横断した初めての台風であったこと等、2018年は特異な強度や進路の台風が多かったことを踏まえると、特にこのような台風に対してはガイダンスの改悪傾向が顕著になる可能性があるため、留意が必要である。続いて、ガイダンスのコントロールランと摂動ランを比

<sup>1</sup> 森 祐貴

表 2.3.1 台風事例として扱った期間の一覧。各期間の 00, 03, …, 21JST を対象とする予測を検証に用いる。

期間(年/月/日)	関連する台風	備考
2018/07/02–2018/07/04	T1807	沖縄・九州地方に接近
2018/07/29	T1812	強い勢力で近畿地方に上陸、その後西日本を西進
2018/08/15	T1815	九州地方に上陸
2018/08/23–2018/08/24	T1820	強い勢力で四国地方に上陸
2018/09/04	T1821	非常に強い勢力で四国地方に上陸
2018/09/29–2018/10/01	T1824	強い勢力で近畿地方に上陸

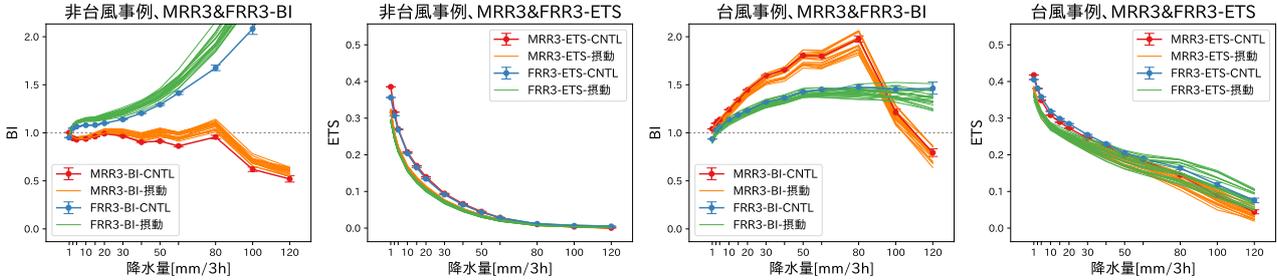


図 2.3.1 MRR3 と FRR3 の閾値別の予測精度比較。図はそれぞれ左から順に、非台風事例の BI、非台風事例の ETS、台風事例の BI、台風事例の ETS を示す。青線はコントロールランの FRR3、緑線は摂動ランの FRR3、赤線はコントロールランの MRR3、橙線は摂動ランの MRR3 について、BI または ETS をそれぞれ表し、エラーバーはブートストラップ法で求めた 95%信頼区間を表している。

較すると、非台風事例・台風事例のいずれも、摂動ランの BI はコントロールランの変化傾向に概ね対応した予測特性となっているが、非台風事例では摂動ランのほうが BI が高く、台風事例では摂動ランのほうが BI が低い傾向がみられた。これについてはモデルにおいても同様の傾向がみられ、ガイダンスはこれに準じる形となっていると考えられる。また、摂動ランの ETS もコントロールランの変化傾向に概ね対応した予測特性となっているが、コントロールランと比較して一段劣ることが分かる。

図 2.3.2 に、RMAX31 と RMAX24 の閾値別 BI および ETS の比較結果を示す。これらの要素についても、摂動ランの BI および ETS はコントロールランの変化傾向に概ね対応した予測特性となっており、摂動ランの ETS はコントロールランと比較して一段劣る傾向がみられるが、RMAX31 は非台風事例でコントロールランよりも BI がやや高く、RMAX24 は非台風事例・台風事例のいずれもコントロールランよりも BI がやや低い傾向がみられた。また、RMAX33 については RMAX31 と同様の傾向がみられた (図略)。

## (2) アンサンブル平均およびアンサンブル最大

図 2.3.3 に、非台風事例における MEPS 降水ガイダンスの MRR3, RMAX33 および RMAX24 のアンサンブル平均とコントロールランの閾値別 BI および ETS の比較結果を示す。いずれの予測要素についても強雨を中心にアンサンブル平均の予測頻度は大幅な過少傾向がみられ、ETS もコントロールランと比較して劣っている。これは、第 1 章でも述べられているとおり、アンサンブル平均によって各メンバーの表現する降水の分布やピークが平滑化されるためであり、降水量のアンサンブル平均は強雨の分布やピークを捉えるには不

向きであることを示している。なお、台風事例の検証 (図略) ではコントロールランの予測頻度の過多傾向をアンサンブル平均による平滑化によって打ち消すことで、BI や ETS が見かけ上改善している場合もみられたが、降水量のアンサンブル平均は強雨の分布やピークを捉えるには不向きであるという点については同様である。図 2.3.4 に、非台風事例および台風事例における MEPS 降水ガイダンスの RMAX31, RMAX33 および RMAX24 のアンサンブル最大とコントロールランの捕捉率および空振り率の比較結果を示す。非台風事例および台風事例のいずれも、アンサンブル最大は全ての閾値で空振り率が高くなっている一方で捕捉率が高く、例えば非台風事例における RMAX31 の 50 mm/h や RMAX24 の 200 mm/24h ではコントロールランの 2 倍以上の捕捉率であることが分かる。降水の予測においては、防災上の観点から見逃しが少ないことは重要であり、アンサンブル最大は顕著現象のポテンシャルを把握する上で有効な資料の 1 つとなると考えられる。ただし、空振り率が高いことを踏まえ、後述の超過確率と併せて現象および目的に応じて使い分けることが重要である。

## (3) 超過確率

超過確率は主に顕著現象のポテンシャル把握や早期注意情報 (警報級の可能性) の判断等への活用を想定しているため、ここでは MEPS の最大降水量ガイダンス (RMAX31, RMAX33, RMAX24) から作成した超過確率の検証結果を示す。

図 2.3.5 に、非台風事例における RMAX31, RMAX33, RMAX24 の超過確率の信頼度曲線とブライアスキルスコア (BSS) を降水量の閾値ごとに示す。いずれについても強い降水ほど信頼度曲線が理想直線

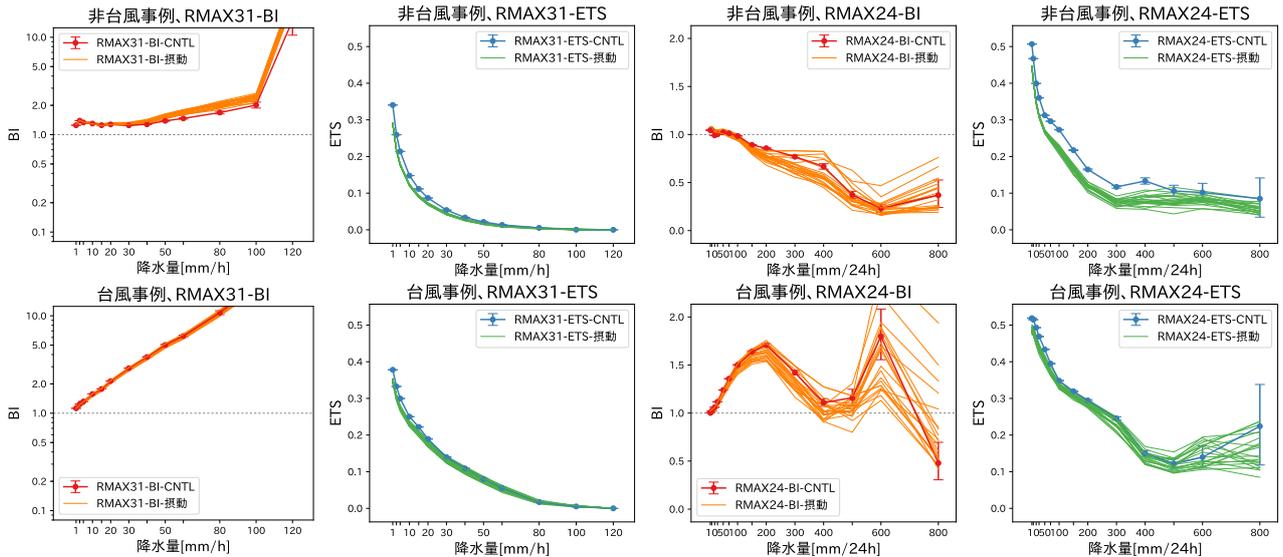


図 2.3.2 非台風事例（上段）および台風事例（下段）における RMAX31 および RMAX24 の閾値別の予測精度比較。図は左から順に RMAX31 の BI、RMAX31 の ETS、RMAX24 の BI、RMAX24 の ETS を示す。赤線はコントロールランの BI、橙線は摂動ランの BI、青線はコントロールランの ETS、緑線は摂動ランの ETS をそれぞれ表す。エラーバーは図 2.3.1 と同じ。なお、RMAX31 の BI については対数軸で描画している。

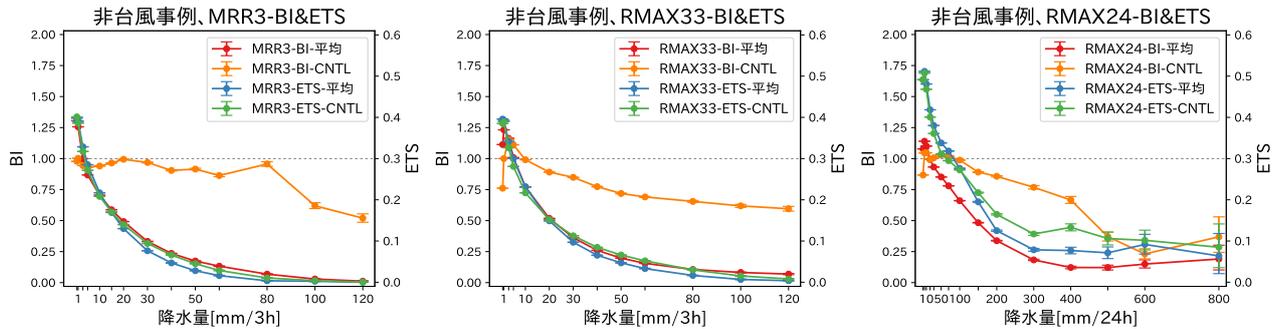


図 2.3.3 非台風事例における、MRR3（左）、RMAX33（中央）、RMAX24（右）のコントロールランとアンサンブル平均の閾値別の BI および ETS の比較。各要素について、赤線はアンサンブル平均の BI、橙線はコントロールランの BI、青線はアンサンブル平均の ETS、緑線はコントロールランの ETS を表し、エラーバーは図 2.3.1 と同じ。

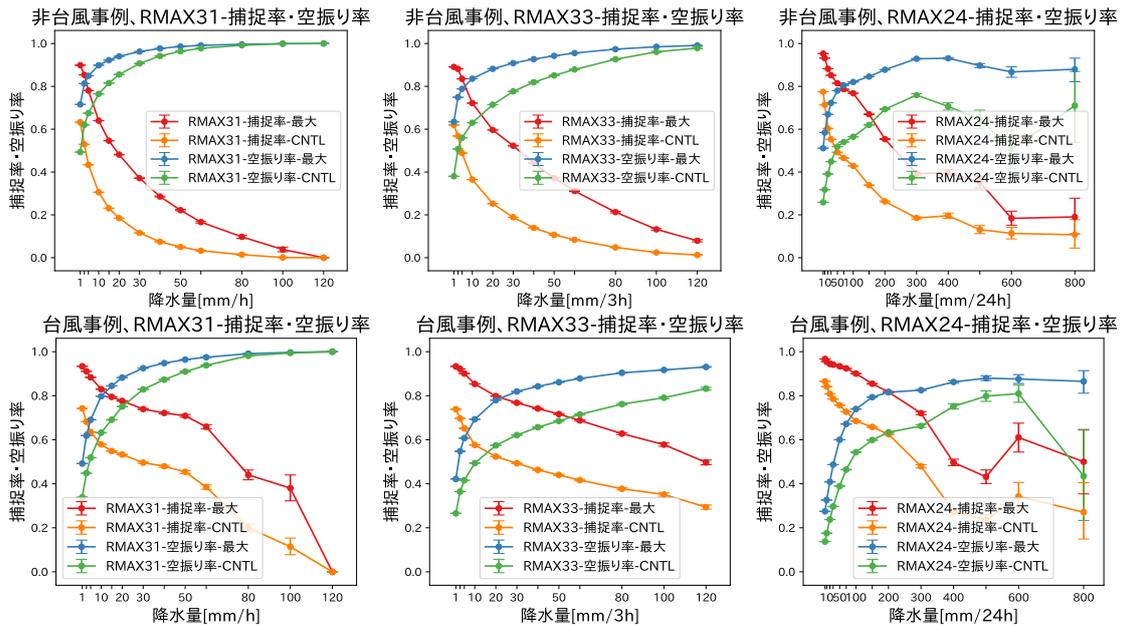


図 2.3.4 非台風事例（上段）および台風事例（下段）における、RMAX31（左）、RMAX33（中央）、RMAX24（右）のコントロールランとアンサンブル最大の閾値別の捕捉率および空振り率の比較。各要素について、赤線はアンサンブル最大の捕捉率、橙線はコントロールランの捕捉率、青線はアンサンブル最大の空振り率、緑線はコントロールランの見逃し率を表し、エラーバーは図 2.3.1 と同じ。

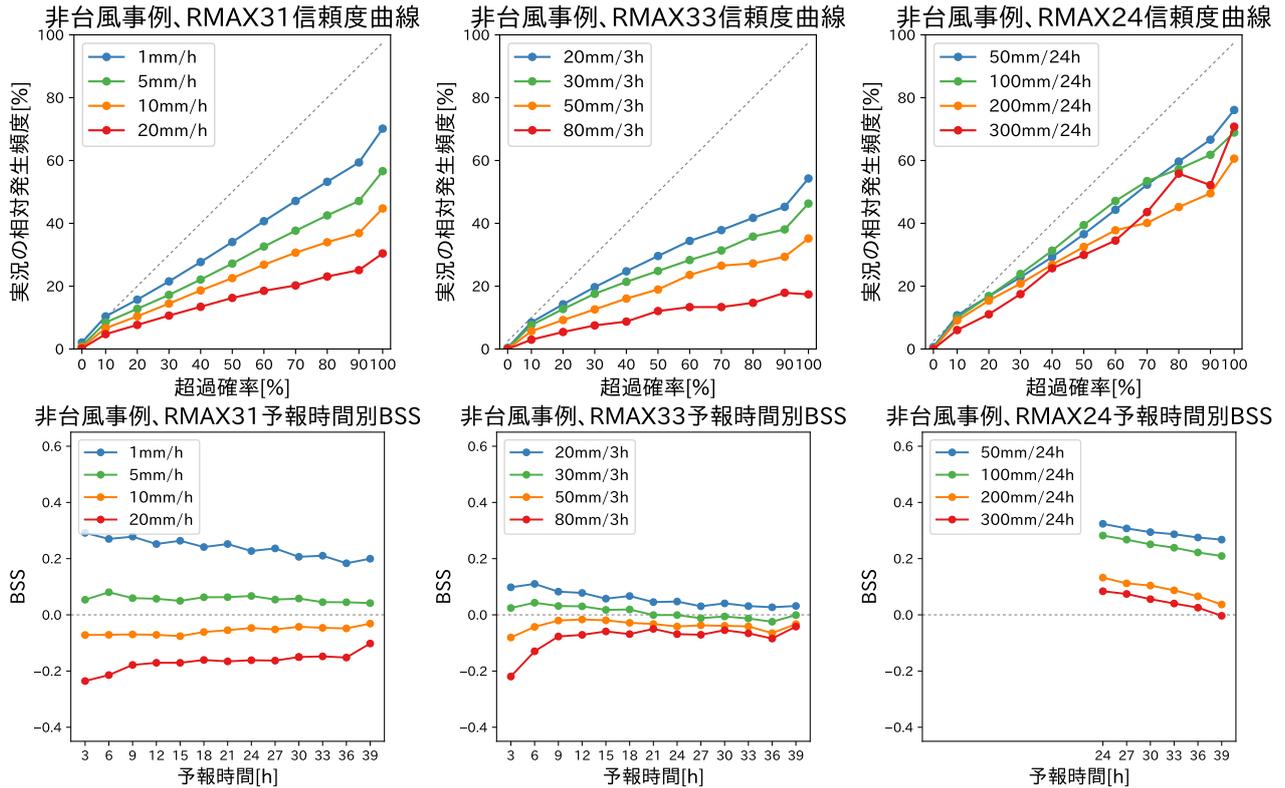


図 2.3.5 非台風事例における RMAX31 (左)、RMAX33 (中央)、RMAX24 (右) の超過確率の信頼度曲線 (上段) および BSS (下段) の検証結果。青線、緑線、橙線、赤線の順に低い閾値の超過確率を表しており、各要素の閾値はそれぞれ低い順に、RMAX31 は 1, 10, 20, 30 mm/h、RMAX33 は 20, 30, 50, 80 mm/3h、RMAX24 は 50, 100, 200, 300 mm/24h である。

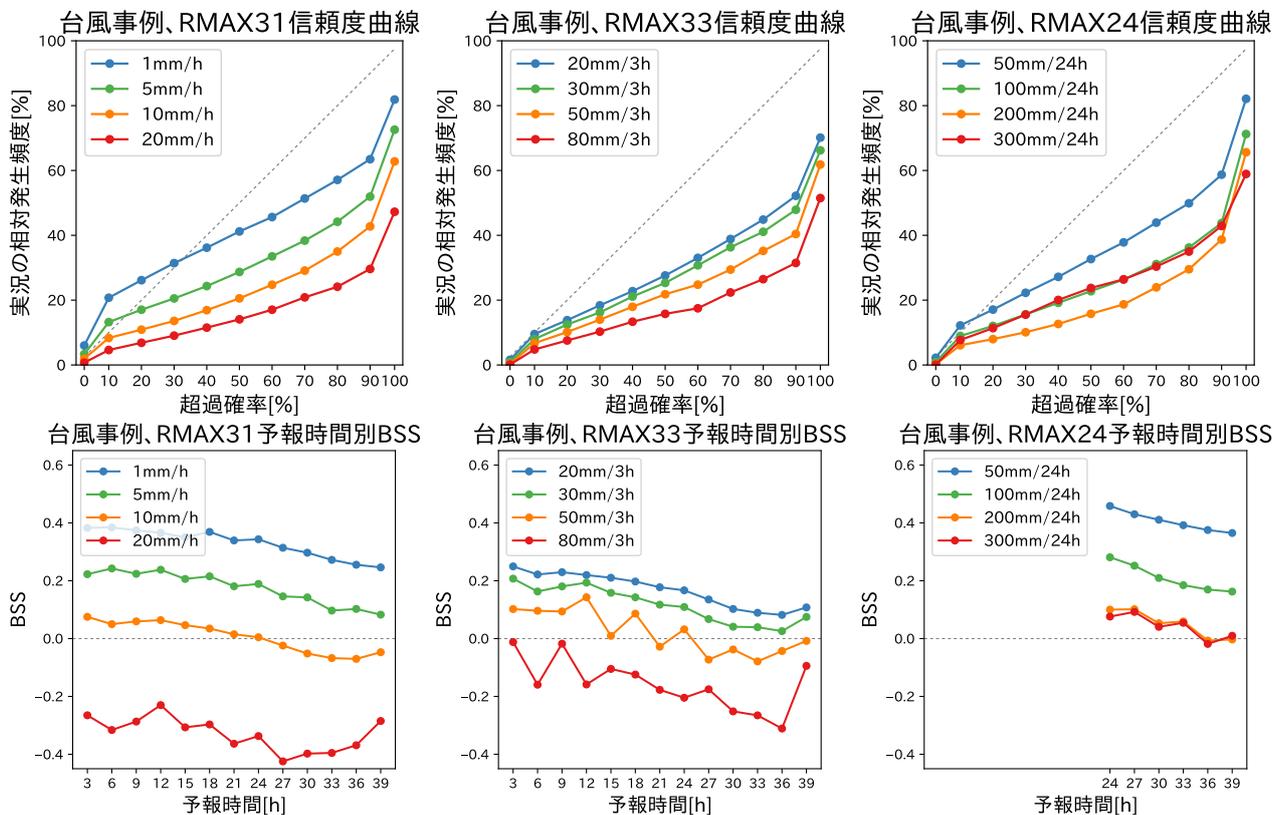


図 2.3.6 図 2.3.5 に同じ。ただし、台風事例における検証。

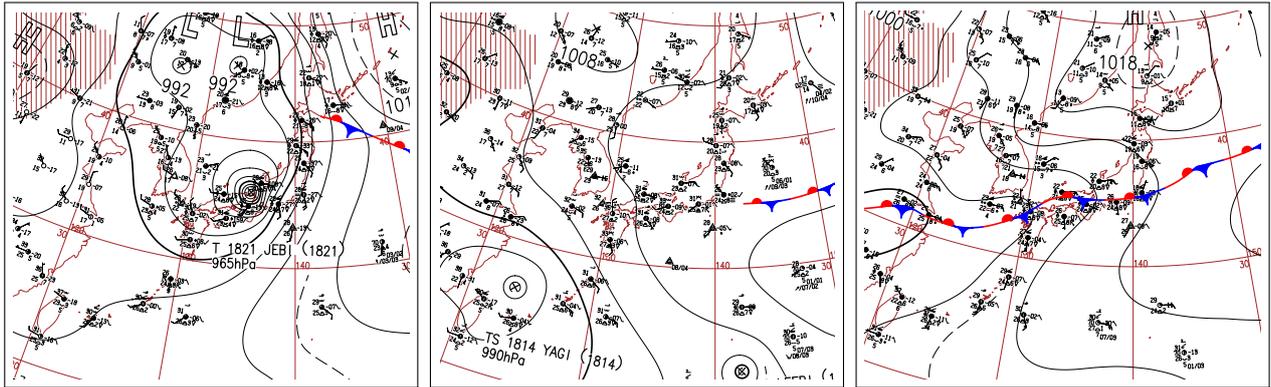


図 2.3.7 2018 年 9 月 4 日 (左)、2018 年 8 月 12 日 (中央)、2018 年 7 月 6 日 (右) のアジア太平洋地上天気図。いずれも日本時間の 15 時の天気図で、日本周辺を拡大して表示している。

から大きく離れて予測の過大傾向が強まり、BSS も全体的に下がる傾向がみられる。超過確率の気候値予測に対する予測スキルをもつ閾値はそれぞれ、RMAX31 は 5 mm/h、RMAX33 は 20–30 mm/3h、RMAX24 は 300 mm/24h (ただし冬季に絞った検証 (図略) では、RMAX31 は 1 mm/h、RMAX33 は 10 mm/3h、RMAX24 は 50 mm/24h) 程度までが目安である。

続いて、図 2.3.6 に、台風事例における RMAX31、RMAX33、RMAX24 の超過確率の信頼度曲線と BSS を降水量の閾値ごとに示す。いずれについても非台風事例の予測特性と概ね同様の傾向がみられるが、高確率の予測を中心に信頼度曲線が理想直線に近づいて予測の過大傾向が弱まっており、BSS は非台風事例と比較して全体的に若干高くなる傾向がみられる。超過確率の気候値予測に対する予測スキルを持つ閾値はそれぞれ、RMAX31 は 5–10 mm/h、RMAX33 は 30–50 mm/3h、RMAX24 は 200–300 mm/24h 程度までが目安である。

RMAX24 の超過確率についてはある程度の大雨事例に対しても統計的には一定の予測スキルを有しており、有効に活用できる場面は多いと考えられるが、一方で RMAX31 や RMAX33 (特に RMAX31) の超過確率については予測スキルが十分でないため、例えば早期注意情報 (警報級の可能性) の判断等への活用という観点では、有効に活用できる場面は限定的であると考えられる。このように、各要素の超過確率の予測スキルは、防災気象情報への利活用の観点でそれぞれ大きく異なる点には留意が必要である。

### 2.3.4 事例検証

ここでは、MEPS 降水ガイダンスおよび超過確率の予測特性や利用上の留意点を表す事例を 3 つ紹介する。まず、本項で示す事例の天気図を図 2.3.7 にまとめて示す。1 つ目は、台風が非常に強い勢力を維持しつつ日本の南から北進して徳島県に上陸し、台風からの湿った空気の影響で、四国から東海地方の太平洋側を中心に大雨となった事例 (左) である。2 つ目は、暖湿気の流入や日中の昇温の影響で全国的に広い範囲で大気が不安定となり、午後には西日本から東日本にかけて雷雨

となった事例 (中央) である。3 つ目は、九州から関東地方にかけて停滞する前線の活動が活発となり、西日本を中心に多くの地点で日降水量の観測史上 1 位の記録を更新した平成 30 年 7 月豪雨の事例 (右) である。

### 台風に伴う降水事例

図 2.3.8 は、2018 年 9 月 4 日 15 時を対象とした MEPS 降水ガイダンスのコントロールランの RMAX33 (左) および RMAX33 の 50 mm/3h を閾値とする超過確率 (中央) と、対応する観測の降水量 (右) である。RMAX33 の超過確率は、実況で閾値を超過した格子が多い紀伊半島南部を中心に高い確率を予測し、淡路島から琵琶湖付近へと北東に伸びる強雨域については 50% 程度の確率を表現している。また、一部の格子で局地的に閾値を超過している四国や静岡県の内陸部等でも実況の降水分布に対応して確率分布を表現している等、実況を確率的に捕捉できていることが分かる。

図 2.3.9 は、図 2.3.8 と同じ時刻を予測の対象とした、24 時間後の初期時刻の MEPS 降水ガイダンスにおけるコントロールランの RMAX33 (左) および RMAX33 の 50 mm/3h を閾値とする超過確率 (中央) と、RMAX33 のアンサンブル最大 (右) である。図 2.3.8 と比較すると、紀伊半島南部ではいずれもコントロールランで強い降水を予測しており、これに対応して超過確率でも引き続き高い確率値を予測している。一方、淡路島から琵琶湖付近へと北東に伸びる強雨域に対応する領域については、超過確率はより高い確率を予測してメリハリの付いた分布を表現している。これは、リードタイムが短くなったことで当該地域におけるメンバー間の降水量予測のばらつきが相対的に小さくなり、降水量が 50 mm/3h を超える雨域の分布の確実性が高まっていることを示唆しており、実際に実況との対応も比較的良いことが分かる。また、超過確率とアンサンブル最大を比較すると、両者の関係から自明ではあるが、アンサンブル最大で 50 mm/3h 以上の降水域は 50 mm/3h を閾値とする超過確率の分布と対応していることが見て取れる。実際に、超過確率とアンサンブル最大のいずれも、コントロールランでは捕捉しきれ

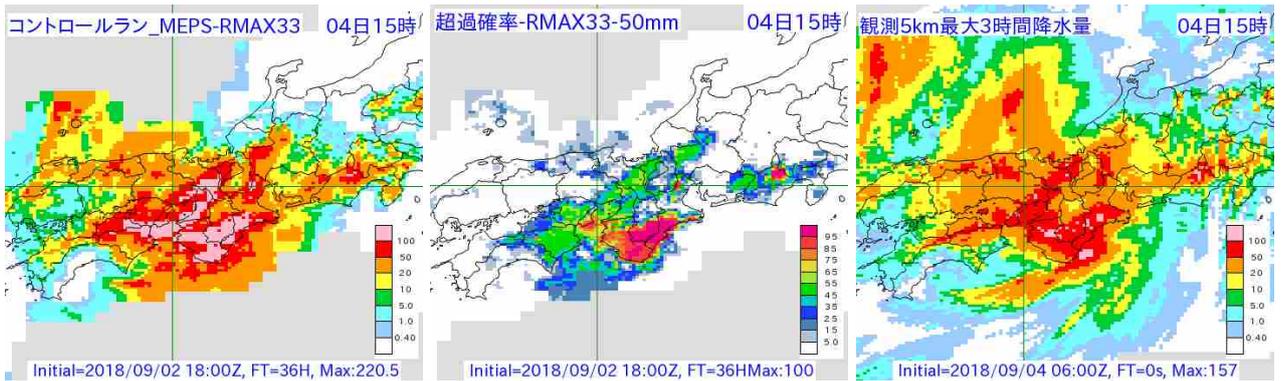


図 2.3.8 2018 年 9 月 4 日 15 時を対象とする、MEPS 降水ガイダンスのコントロールランの RMAX33[mm/3h] (左) および 50 mm/3h を閾値とする超過確率 [%] (中央) と、観測の 5 km 格子最大 3 時間降水量 [mm/3h] (右)。MEPS 降水ガイダンスの初期時刻および予報時間は、2018 年 9 月 2 日 18UTC 初期値の FT=36。

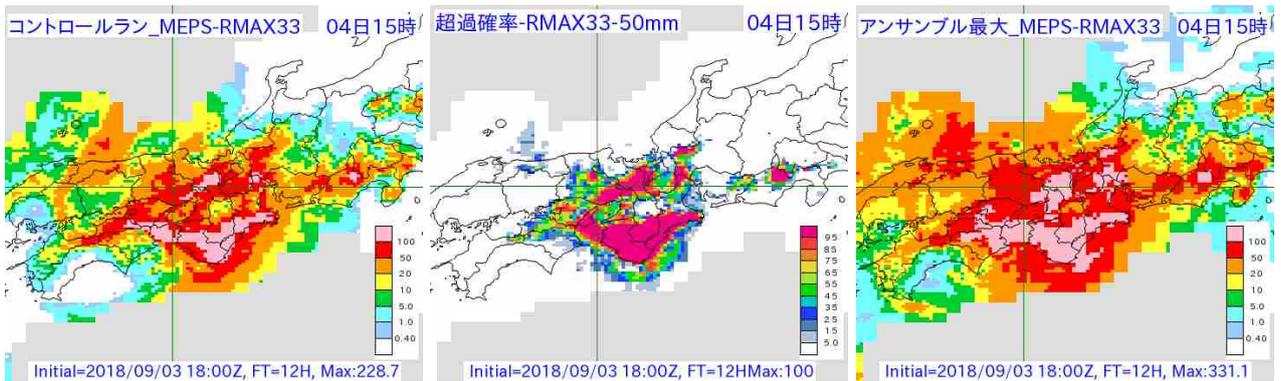


図 2.3.9 2018 年 9 月 4 日 15 時を対象とする、MEPS 降水ガイダンスのコントロールランの RMAX33[mm/3h] (左) および 50 mm/3h を閾値とする超過確率 [%] (中央) と、アンサンブル最大 [mm/3h] (右)。MEPS 降水ガイダンスの初期時刻および予報時間は、2018 年 9 月 3 日 18UTC 初期値の FT=12。

ていない兵庫県北部の 50 mm/3h 以上の降水について概ね捕捉できていることが分かる。ただし、アンサンブル最大は実況と比較して予測の空振りが非常に目立ち、また、確率的な情報を持たない決定論的な予測資料であるため、防災気象情報の発表の判断等において直接活用するには適していない。このため、基本的には超過確率を参考に現象発生の可能性や範囲を見積もりつつ、予測を組み立てていくことが推奨される。

このように、超過確率は MEPS 降水ガイダンスの予測のばらつきに対応した確率分布を表現しており、実況を広く捕捉しつつ予測の不確実性を確率値として把握することができる。また、初期時刻毎の確率値の変化を追うことで大雨の範囲の絞り込みに利用できる等、リードタイムに応じた防災気象情報の発表の判断等において有効に活用できると考えられる。

### 夏季の不安定降水の事例

図 2.3.10 に、2018 年 8 月 12 日 18 時を対象とした MEPS 降水ガイダンスのコントロールランの RMAX31 (左) および摂動メンバーの 1 つ (メンバー 12) (中央) と、対応する観測の降水量 (右) を示す。中部地方を中心とした強雨については、ガイダンスのコントロールランは多少の位置ずれや空振り等はみられるものの、実況に近い強い降水は予測しており、強雨のシグナルをある程度表現できていると考えられる。しかしその

一方で、九州・四国・中国地方の降水についてはほとんど表現できていない。これに対してメンバー 12 では、中部地方を中心とした強雨については全く予測できていない一方で、九州・四国・中国地方の降水についてはある程度表現できていることが分かる。

続いて、図 2.3.11 に、図 2.3.10 に対応する RMAX31 の 20 mm/h (左) および 1 mm/h (中央) を閾値とした超過確率と、MEPS 降水ガイダンスのアンサンブル最大を示す。図 2.3.10 で示したとおり、本事例では MEPS 降水ガイダンスの予測は各メンバーでばらつきが非常に大きく、例えば 20 mm/h を閾値とする超過確率は岐阜県内を中心に数 10%程度と表現されるのみである。しかし、例えば 1 mm/h を閾値とした超過確率をみると、実況で強い降水が集中している中部地方を中心に確率値を広く表現しており、不安定降水のポテンシャルについては確率的に表現できていることが分かる。ただし、この場合の超過確率の確率値については強雨のポテンシャルを直接的に表したのではなく、定性的に捉える必要がある点には留意が必要である。アンサンブル最大についても、降水を広く予測して不安定降水のポテンシャルを表現している点については同様であるが、確率的な情報を持たないという点で超過確率にはやや劣ると言える。その一方で、アンサンブル最大は予測される強雨の程度を降水量分布として表現

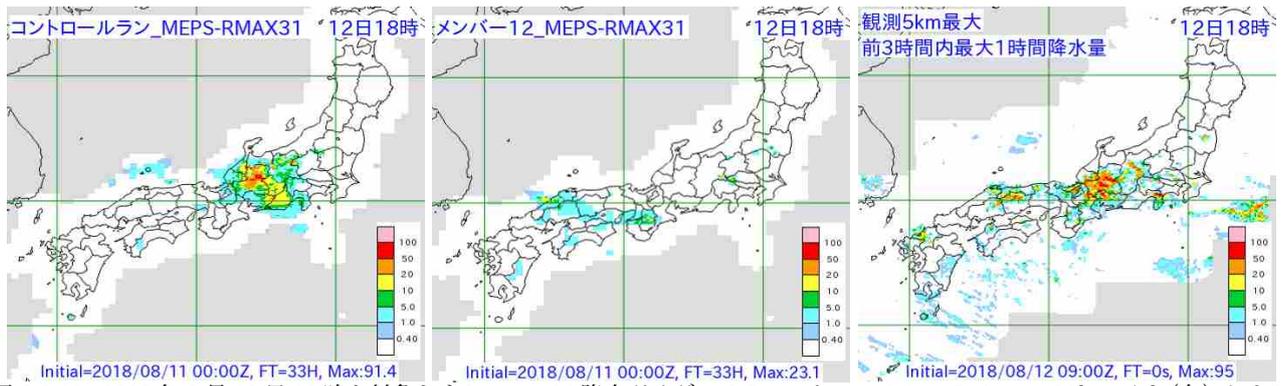


図 2.3.10 2018 年 8 月 12 日 18 時を対象とする、MEPS 降水ガイダンスのコントロールランの RMAX31[mm/h] (左) および 摂動メンバーの 1 つ (メンバー 12) [mm/h] (中央) と、観測の 5 km 格子最大前 3 時間内最大 1 時間降水量 [mm/h] (右)。MEPS 降水ガイダンスの初期時刻および予報時間は、2018 年 8 月 11 日 00UTC 初期値の FT=33。

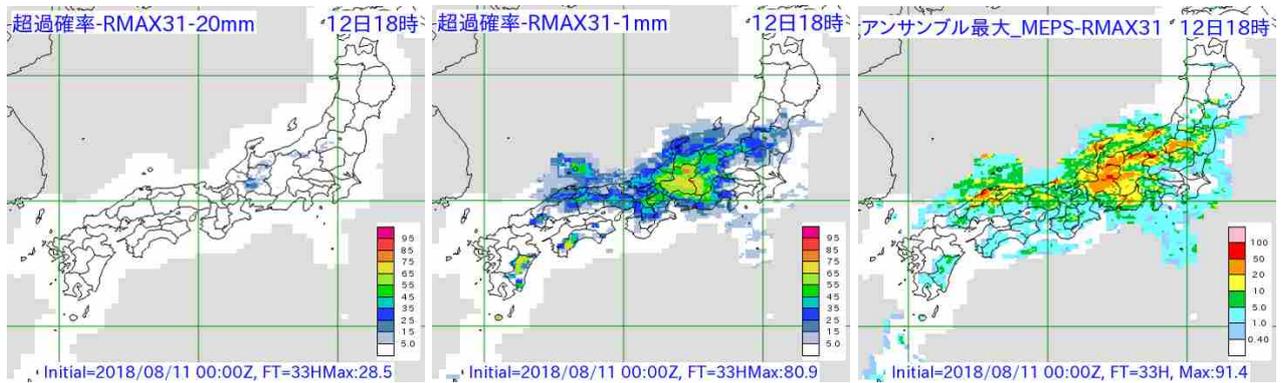


図 2.3.11 2018 年 8 月 12 日 18 時を対象とする、MEPS 降水ガイダンスの RMAX31 の 20 mm/h (左)、1 mm/h (中央) を 閾値とする超過確率 [%] と、アンサンブル最大 [mm/h] (右)。MEPS 降水ガイダンスの初期時刻および予報時間は、2018 年 8 月 11 日 00UTC 初期値の FT=33。

することで、強雨のポテンシャルを表現しており、実際に中部・中国地方の強雨との対応も良い。この点では、アンサンブル最大は超過確率にない特長をもっていると言える。

このように、本不安定降水事例のような局地的・短時間に発生する降水現象の予測においては、MEPS 降水ガイダンスのアンサンブルメンバーの個々の予測は他の事例に比べてばらつきが非常に大きく、超過確率については特に高い閾値において確率値を適切に表現できない場合が多い。不安定降水のポテンシャルの確率的な把握にあたっては低い閾値の超過確率の分布を参照することが有効であるが、強雨のポテンシャルについてはアンサンブル最大の分布を参照することが有効であると考えられ、目的に応じて両者を使い分けることが重要である。

#### 前線に伴う降水事例 (平成 30 年 7 月豪雨)

図 2.3.12 に、2018 年 7 月 6 日 15 時を対象とした MEPS 降水ガイダンスのコントロールランの RMAX24 (左) および 200 mm/24h を閾値とする超過確率 (中央) と、対応する観測の降水量 (右) を示す。高知県東部や近畿地方、岐阜県周辺等、コントロールランの予測値や超過確率の分布は概ね実況と良く対応していることが分かる。一方で、九州北部の 200 mm/24h を超える降水については、超過確率では確率値がほとん

ど表現されておらず、実況を十分に捕捉できているとは言いがたい。

続いて、図 2.3.13 に、図 2.3.12 に対応した MEPS 降水ガイダンスの RMAX24 の 150 mm/24h (左) および 100 mm/24h (中央) を閾値とする超過確率と、アンサンブル最大 (右) を示す。九州北部ではアンサンブル最大で 100 mm/24h や 150 mm/24h を超える予測の格子はある程度面的な広がりをもっており、これに対応して超過確率の分布もそれぞれ表現されていることが分かる。一方、九州北部ではアンサンブル最大で 200 mm/24h を超える予測の格子はほとんどなく、超過確率の分布もこれに対応している。一般に、超過確率はアンサンブルメンバーの予測のばらつきの幅を超えて確率値を表現することはできず、また、アンサンブル最大付近を閾値とした場合にも、超過確率の面的な分布を適切に表現することは難しい。このため、本事例でも九州北部においてアンサンブル最大に迫る 200 mm/24h を閾値とした超過確率の分布を参照するのは適切ではないと考えられる。このような強雨のポテンシャルについては、アンサンブル最大の分布や 100 mm/24h あるいは 150 mm/24h といった少し低い閾値の超過確率を参考に、定性的な見積もりを行うのが適当であると考えられる。

このように、本事例のような空間的・時間的に連続

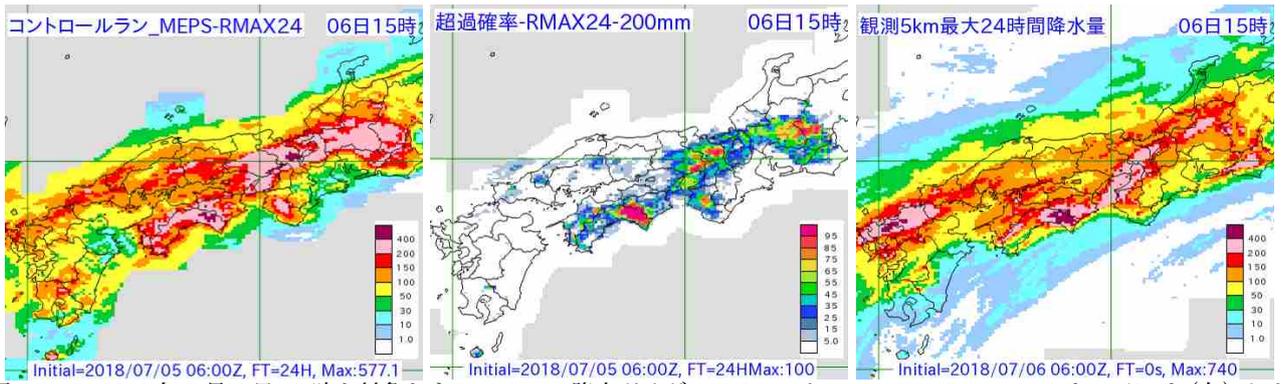


図 2.3.12 2018 年 7 月 6 日 15 時を対象とする、MEPS 降水ガイダンスのコントロールランの RMAX24[mm/24h] (左) および 200 mm/24h を閾値とする超過確率 [%] (中央) と、観測の 5 km 格子最大 24 時間降水量 [mm/24h] (右)。MEPS 降水ガイダンスの初期時刻および予報時間は、2018 年 7 月 5 日 06UTC 初期値の FT=24。

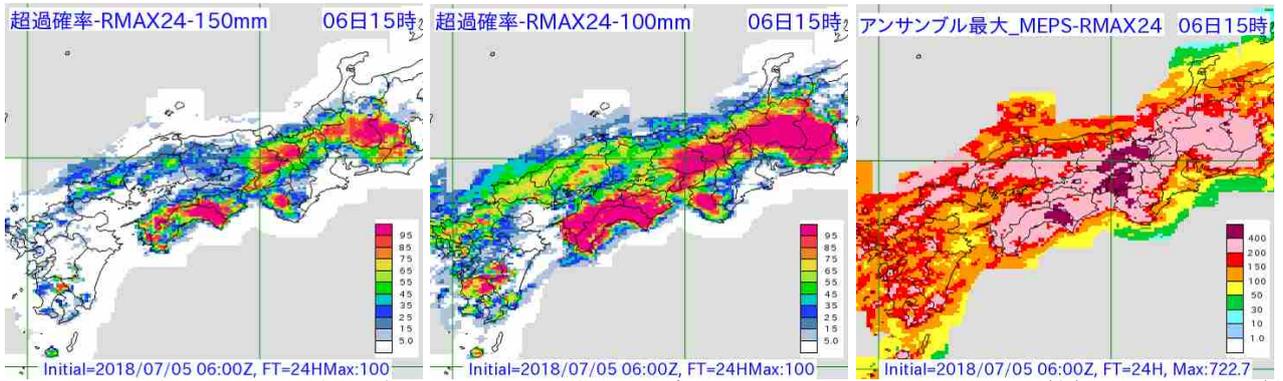


図 2.3.13 2018 年 7 月 6 日 15 時を対象とする、MEPS 降水ガイダンスの RMAX24 の 150 mm/24h (左)、100 mm/24h (中央) を閾値とする超過確率 [%] と、アンサンブル最大 [mm/h] (右)。MEPS 降水ガイダンスの初期時刻および予報時間は、2018 年 7 月 5 日 06UTC 初期値の FT=24。

的な降水現象の予測においても、超過確率はガイダンスのアンサンブルメンバーの予測のばらつきを踏まえて閾値を適切に選択する必要がある点には留意が必要である。強雨のポテンシャルの定性的な見積りにあたっては、まずアンサンブル最大の分布を確認し、ある程度面的な分布が表現されている少し低めの閾値の超過確率を参考にする、といった工夫が必要であると考えられる。

### 2.3.5 まとめ

MEPS 降水ガイダンスは、MEPS の各メンバーの予測値を入力とするガイダンスであり、MSM 降水ガイダンスの係数や予測式を用いて作成する。コントロールランは MSM 降水ガイダンスと同一であり、各メンバーの予測のばらつきを利用することで、MSM 降水ガイダンスに関する確率情報等の作成が可能となる。その 1 つとして、アンサンブルメンバーの降水量予測がある閾値を超える確率（超過確率）が挙げられる。本節では、MEPS 降水ガイダンスと、超過確率をはじめとした各種統計量について精度検証および事例検証を行った。これらの結果について、利用上の留意点と併せて以下にまとめる。

### MEPS 降水ガイダンスについて

非台風事例の検証では、ガイダンスはモデルを適切に修正し、予測頻度・予測精度ともに改善がみられる。一方で、台風事例の検証ではガイダンスはモデルよりも更に予測頻度過多かつ予測精度もやや劣る結果となり、ガイダンスがモデルを適切に修正できていない。摂動ランの予測特性は、非台風事例・台風事例のいずれもコントロールランの変化傾向に概ね対応する形となっており、予測精度はコントロールランと比較して一段劣る傾向がみられる。また、予測要素によってコントロールランよりも摂動ランのほうが BI がやや高いあるいはやや低い傾向がある場合もみられる。

### アンサンブル平均およびアンサンブル最大について

アンサンブル平均については平均処理によって各メンバーの表現する降水の分布やピークが平滑化されるため、強雨の分布やピークを捉えるのには適さない。アンサンブル最大についてはコントロールランと比較して捕捉率が高く、顕著現象のポテンシャルを把握する上で有効な資料の 1 つとなると考えられる。ただし、空振り率が高い点には十分に留意が必要であり、例えば不安定降水事例における強雨のポテンシャルの見積もりへの利用等、現象や目的に応じた超過確率との使い分けが重要である。

## 超過確率について

超過確率の検証では、いずれの要素についても強い降水ほど予測の過大傾向が強まり、予測精度も低下する傾向がみられた。非台風事例と台風事例とでは、全体としては後者の方が予測精度が高くなる傾向が見られ、統計的な目安としては、RMAX31は5–10 mm/h、RMAX33は20–50 mm/3h、RMAX24は200–300 mm/24h（冬季では、RMAX31は1 mm/h、RMAX33は10 mm/3h、RMAX24は50 mm/24h）程度までは気候値予測に対してスキルがあると考えられ、各要素の超過確率の予測スキルは、防災気象情報への利活用の観点でそれぞれ大きく異なる点には留意が必要である。超過確率はMEPS降水ガイダンスの予測のばらつきに対応した確率分布を表現でき、実況を広く捕捉しつつ予測の不確実性を確率値として把握することができる。また、初期時刻毎の確率値の変化に着目することで、リードタイムに応じた防災気象情報の発表の判断等において有効に活用することが可能である。ただし、不安定降水のように各メンバーの予測のばらつきが非常に大きい事例では、超過確率が適切に表現されない場合もみられる。このような場合、不安定降水のポテンシャルの確率的な把握にあたっては低い閾値の超過確率の分布を参照することが有効であるが、強雨のポテンシャルについてはアンサンブル最大の分布を参照することが有効である。また、空間的・時間的に連続的な降水事例においても、アンサンブルメンバーの予測のばらつきを踏まえて超過確率の閾値を適切に選択する必要がある点には留意が必要である。強雨のポテンシャルの定性的な見積りにあたっては、アンサンブル最大の分布を踏まえつつ、少し低めの閾値の超過確率を参考にすると、といった工夫が必要である。

今回のMEPS降水ガイダンスの検証では、台風事例においてガイダンスがモデルを適切に修正できていない場合が多いことが示唆された。ただし、これはMEPS降水ガイダンスの課題ではなく、MSM降水ガイダンスとしての課題である。このため、今後もMSM降水ガイダンスの改良に継続的に取り組んでいくことが重要である。

### 2.3.6 今後の展望

上述の通り、超過確率は確率情報としてある程度の予測スキルを有しているものの、RMAX31やRMAX33の超過確率等については、例えば早期注意情報（警報級の可能性）の判断等への活用という観点では十分でなく、より高い精度をもつ確率情報の作成が重要である。そこで以下では、現在開発を行っている、予測の時空間的なずれや地形等を考慮した確率情報（以下、補正超過確率）について、作成方法や事例検証の結果を簡単に紹介する。

### (1) 補正超過確率の作成方法

まず、補正超過確率の作成において精度向上の寄与が最も大きい手法であるNeighbourhood Processingについて紹介する。Neighbourhood Processingは、英国気象局等で降水確率等の確率情報を作成する際に用いられている手法の1つで、ある閾値を超える格子のカウントにおいて対象格子の周辺の格子も参照する手法である(Clark et al. 2016; Schwartz et al. 2010)。図2.3.14に、Neighbourhood Processingのイメージ図を示す。Neighbourhood Processingでは、各メンバーの各格子について、その格子を含む周辺の11×11格子のうち何格子が閾値を超えるかをカウントし、それらを全メンバー分足して、周辺格子を含む全参照格子数に対する割合を求める。アンサンブルのメンバー数を $M$ 、ある閾値を超える格子の数の合計を $N$ とした場合、計算される確率値は $100 \times N / (11 \times 11 \times M) \%$ となる。この値は、超過確率を11×11個の周辺格子で平均した値と同じである。

次に、補正超過確率の作成方法について簡単に解説する。まず、上述のNeighbourhood Processingにおいて、周辺格子について対象格子との水平距離や標高差が大きいほど重みが小さくなるように重み付けをして計算を行う。続いて、各FT・各格子で計算されたこの値について、前後のFTの同一格子とで重み付き平均（予報時間後半ほど前後のFTの重みを徐々に大きくする）を行う。これにより、予測の時空間的なずれおよび地形を考慮することができる。

### (2) 補正超過確率の事例検証

図2.3.15に、図2.3.8の12時間後を予測対象とした200 mm/24hを閾値とするRMAX24の超過確率（左）および補正超過確率（中央）と、対応する観測の降水量（右）を示す。超過確率および補正超過確率は高確率格子と実況との対応が良く、四国東部・紀伊半島・岐阜県北西部・東海地方内陸部のそれぞれについて、実況で200 mm/24hを超えた格子の周辺にピークを持つ確率分布が表現されている。一方で、静岡県と長野県の県境付近（図中の四角枠内）に着目すると、実況では長野県の南部まで200 mm/24hを超えている格子が広がっているが、超過確率は長野県内では概ね0%となっている。これに対し、補正超過確率は長野県南部まで40%前後の確率を表現しており、実況を確率的に捕捉できている。また、滋賀県と三重県の県境付近（図中の破線枠内中央付近）では補正超過確率の分布がややシャープに表現されており、地形に沿って確率分布を広げていることが分かる。本事例の場合、当該地域において実況で200 mm/24hを超えた格子は1格子のみではあるものの、降水量はその周辺と比較すると相対的に多くなっており、補正超過確率は地形による降水特性の違いを反映した分布を表現していると言える。

このように、補正超過確率は予測の位置ずれ等を考

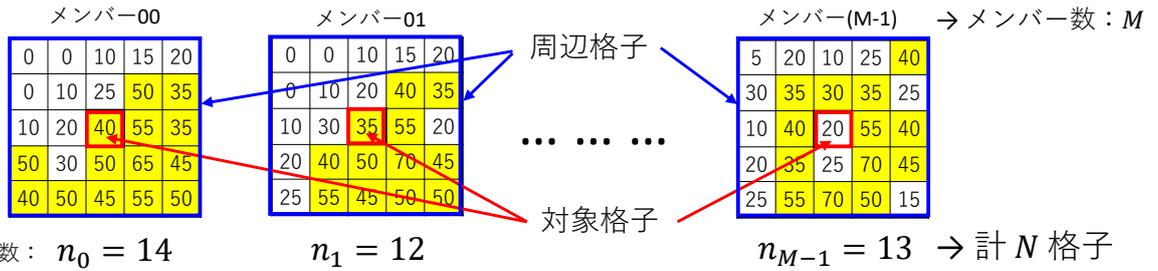


図 2.3.14 Neighbourhood Processing のイメージ図。アンサンブルメンバー数を  $M$ 、閾値を 30 mm とした場合の例を示す。閾値を超えた格子を黄色で塗りつぶして示しており、それらの格子数の全メンバー分の合計を  $N$  としている。なお、ここでは簡単のため、対象格子が参照する周辺格子を  $5 \times 5$  (本来は  $11 \times 11$ ) の格子で図示している。

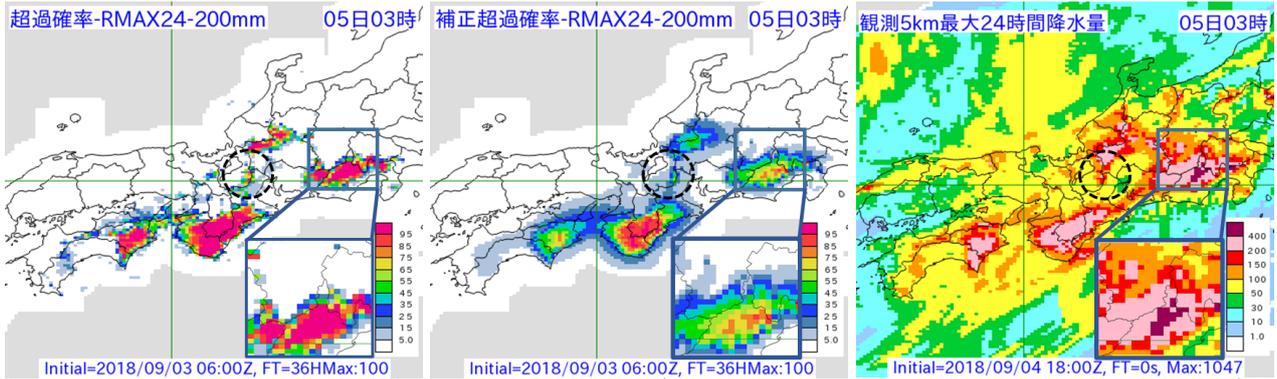


図 2.3.15 2018 年 9 月 5 日 3 時を対象とする、MEPS 降水ガイダンスの RMAX24 の 200 mm/24h を閾値とする超過確率 [%] (左) および補正超過確率 [%] (中央) と、観測の 5 km 格子最大 24 時間降水量 [mm/24h] (右)。四角枠線内は拡大図も示している。MEPS 降水ガイダンスの初期時刻および予報時間は、2018 年 9 月 3 日 06UTC 初期値の FT=36。

慮して超過確率の分布を広げることで、実況を確率的に捕捉しやすくなっている一方で、地形が急峻な領域では確率分布の広がりを抑え、ある程度地形に沿った確率分布の表現が可能であることが分かる。

今後、超過確率の代わりとなるより精度の高い確率情報の作成を目指し、補正超過確率のさらなる改良に取り組んでいく。

### 参考文献

河野耕平, 西本秀祐, 2018: MEPS の利用方法の検討. 平成 30 年度数値予報研修テキスト, 気象庁予報部, 10-11.

工藤淳, 後藤尚親, 高田伸一, 2018: ガイダンスの作成技術. 数値予報課報告・別冊第 64 号, 気象庁予報部, 19-85.

三戸洋介, 2018: メソアンサンブルガイダンス. 平成 30 年度数値予報研修テキスト, 気象庁予報部, 11-12.

白山洋平, 2018: 降水ガイダンス. 数値予報課報告・別冊第 64 号, 気象庁予報部, 95-118.

Clark, Peter, Nigel Roberts, Humphrey Lean, Susan P. Ballard, and Cristina Charlton-Perez, 2016: Convection-permitting models: a step-change in rainfall forecasting. *Meteorol. Appl.*, **23**, 165-181.

Schwartz, Craig S., John S. Kain, Steven J. Weiss, Ming Xue, David R. Bright, Fanyou Kong, Kevin W. Thomas, Jason J. Levit, Michael C. Coniglio,

and Matthew S. Wandishin, 2010: Toward Improved Convection-Allowing Ensembles: Model Physics Sensitivities and Optimizing Probabilistic Guidance with Small Ensemble Membership. *Wea. Forecasting*, **25**, 263-280.

## 2.4 航空悪天 GPV<sup>1</sup>

### 2.4.1 はじめに

MEPSに基づく航空悪天 GPV（以下、MEPS 航空悪天 GPV）は、MSMに基づく航空悪天 GPV（以下、国内航空悪天 GPV）と同じ手法で作成される空域予報用 GPV である。各アンサンブルメンバーの予測値や、アンサンブル予報から算出される平均値、超過確率、スプレッドなどの統計量は、空域予報のほか、将来の航空交通システムでも利用される予定である。MEPS 航空悪天 GPV は 2015 年 4 月から 1 日 1 回 11 メンバーで部内試験運用されており、2018 年 6 月には 1 日 4 回 21 メンバーに増強されている。

本節では MEPS 航空悪天 GPV の主な要素である乱気流指数、着氷指数、および積乱雲量のアンサンブル平均や超過確率などの検証結果と予測事例を示し、利用上の留意点を述べる。

### 2.4.2 仕様

MEPS 航空悪天 GPV は国内航空悪天 GPV と同じ領域・座標系・要素で作成しており、MEPS と同様に 1 日 4 回 21 メンバーでの予測を行う。MEPS 航空悪天 GPV の各要素は国内航空悪天 GPV と同じ手法・係数・パラメータで作成されており、コントロールランの予測は同じ初期時刻の国内航空悪天 GPV と同一である。国内航空悪天 GPV の詳細な仕様は付録 A.2.10 を、各要素の作成方法は井藤 (2018a,b) および工藤 (2018a,b) を参照していただきたい。

### 2.4.3 統計検証

以下では MEPS 航空悪天 GPV の乱気流指数、着氷指数、積乱雲量について、アンサンブル平均や超過確率とコントロールランを比較した検証結果を述べる。検証期間は 1 日 4 回 21 メンバーでの部内試験運用を開始した 2018 年 6 月 5 日 00 UTC 初期値から 2019 年 6 月 4 日 18 UTC 初期値までの 1 年間で、00, 06, 12, 18 UTC の全ての初期時刻を用いた検証結果を示す。

#### (1) 乱気流指数の検証

ここでは MEPS 航空悪天 GPV から算出した乱気流指数のアンサンブル平均と超過確率をコントロールランと比較した検証結果を示す。また比較のために、鉛直シアーのアンサンブル平均とコントロールランの検証結果も示す。

各統計量のうち、乱気流指数 (TBindex) の超過確率 ( $p(\text{TBindex} > 3)$ ) は並 (MOD) 以上の乱気流を予測する目安である TBindex が 3.0 を超えたメンバーの割合に 100 を掛けた値、鉛直シアー (VWS) のアンサンブル平均は各メンバーの単純平均とする。TBindex については、アンサンブル平均は各メンバーの単純平均ではなく、ロジスティック回帰で求めた確率値のアンサ

ンブル平均から算出した TBindex とする。この理由は次のとおりである。TBindex はロジスティック回帰で確率値を求めた後に以下の変換を行うことで算出している (工藤 2018a)。

$$\text{TBindex} = \frac{\ln p - \mu}{\sigma} + 3.0 \quad (2.4.1)$$

ここで  $\mu$  と  $\sigma$  は調整パラメータ、 $p$  はロジスティック回帰で求めた確率値である。この TBindex を単純にアンサンブル平均すると、 $p$  の相乗平均に対して (2.4.1) 式を適用することになる。一方で、アンサンブル平均の方法としては  $p$  の相加平均に対して (2.4.1) 式を適用することも考えられる。事前調査の結果、相加平均を用いた方が予測精度が高かったため、ここでは  $p$  の相加平均から算出した TBindex を TBindex のアンサンブル平均 (TBindex(MEAN)) とし、次式で定義する。

$$\text{TBindex(MEAN)} \equiv \frac{\ln\left(\frac{1}{M} \sum_m p_m\right) - \mu}{\sigma} + 3.0 \quad (2.4.2)$$

ここで  $m$  は各アンサンブルメンバーを、 $M$  はアンサンブルメンバー数を表す。

検証では、C-PIREP<sup>2</sup>による乱気流観測を実況データとして利用する。実況と比較する予測値は、実況があった時刻・地点に最も近い格子の値を用いる。なお、乱気流の遭遇地点については、「地点 A で揺れた」と通報される場合と、「地点 A から地点 B まで揺れた」と通報される場合の 2 つのパターンがある。後者の場合は 2 地点間の中点を実況があった場所とするが、2 地点間の距離が水平 240 km 以上または鉛直 6,000 ft 以上離れていた場合は中点に代表性がないと判断し、検証には使用しない。また、天候状態 (晴天、雲中など) による区別はしない。

図 2.4.1 に MOD 以上の乱気流に対する MEPS 航空悪天 GPV の高度別・閾値別のスキルスコアを示す。図の FL はフライトレベル<sup>3</sup>を表す。スキルスコアの最大値で比べると、TBindex(MEAN) と  $p(\text{TBindex} > 3)$  は、高度と FT によらずコントロールラン (TBindex(CNTL)) を上回っており、予測精度が高いといえる。スキルスコアが最大となる  $p(\text{TBindex} > 3)$  の閾値は 20-30% 付近であり、これを目安とすることで TBindex(CNTL) よりも精度が高く最適な予測を行うことができる。

<sup>2</sup> PIREP は Pilot Report (操縦士報告) の略。C-PIREP は Common-PIREP の略で、従来は各航空会社内でのみ利用されていた PIREP (カンパニー PIREP) を国土交通省航空局のシステムで集約して形式を統一した PIREP。乱気流、着氷、天候状態などが報じられる。乱気流の強さは、静穏 (SMTH)、弱<sup>-</sup> (LGTM)、弱 (LGT)、弱<sup>+</sup> (LGTP)、並 (MOD)、強 (SEV) の 6 段階、着氷の強さは、弱 (LGT)、並 (MOD)、強 (SEV) の 3 段階で報じられる。天候状態は、晴天 (CLR)、雲中 (INC)、雲頂 (TOP)、雲底 (BSE) などが報じられる。

<sup>3</sup> 標準大気気圧高度 [ft] を 100 で割った高度。「FL100」など “FL” に続けて 3 桁の数値で表される。FL100 = 10,000 ft、1 ft  $\approx$  0.3048 m。

<sup>1</sup> 工藤 淳

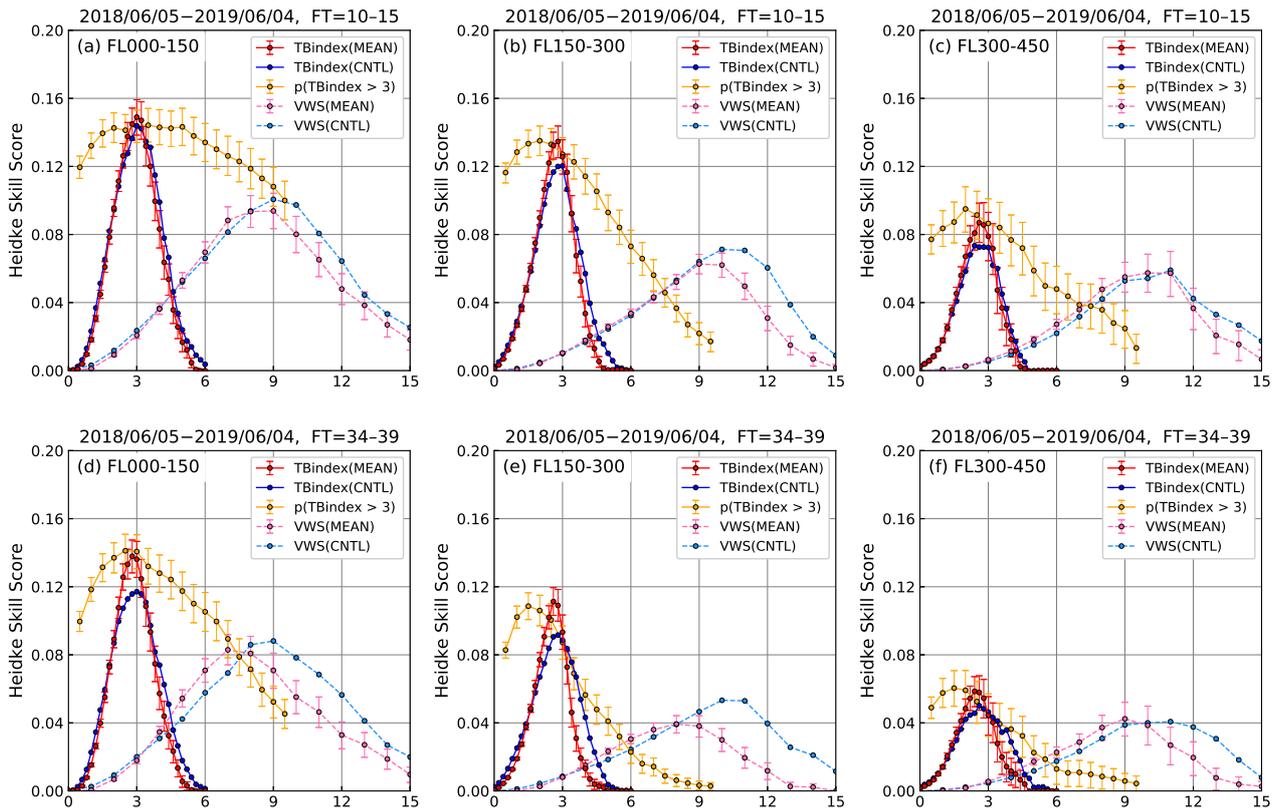


図 2.4.1 MOD 以上の乱気流に対する MEPS 航空悪天 GPV の高度別・閾値別のスキルスコア。(a)~(c) はそれぞれ FT=10-15 の (a) FL000-FL150, (b) FL150-FL300, (c) FL300-FL450 の検証結果、(d)~(f) はそれぞれ FT=34-39 の (d) FL000-FL150, (e) FL150-FL300, (f) FL300-FL450 の検証結果。TBindex(MEAN) は確率のアンサンブル平均から求めた乱気流指数、TBindex(CNTL) はコントロールランの乱気流指数、 $p(\text{TBindex} > 3)$  は TBindex が 3.0 を超える超過確率、VWS(MEAN) は鉛直シアアのアンサンブル平均、VWS(CNTL) はコントロールランの鉛直シアアの検証結果。横軸は TBindex の値または  $p(\text{TBindex} > 3)$  [%] または VWS [kt/1000ft] を表すが、同一の図に表示するために、 $p(\text{TBindex} > 3)$  は確率値 [%] を 10 で割った値で示している。TBindex は 0 から 0.2 毎に、VWS は 0 から 1 kt/1000ft 毎に検証した結果を、 $p(\text{TBindex} > 3)$  は 5% から 95% まで 5% 毎に検証した結果を示している。TBindex(MEAN),  $p(\text{TBindex} > 3)$ , VWS(MEAN) にはブートストラップ法で求めた 95% 信頼区間を示すエラーバーを付加している。

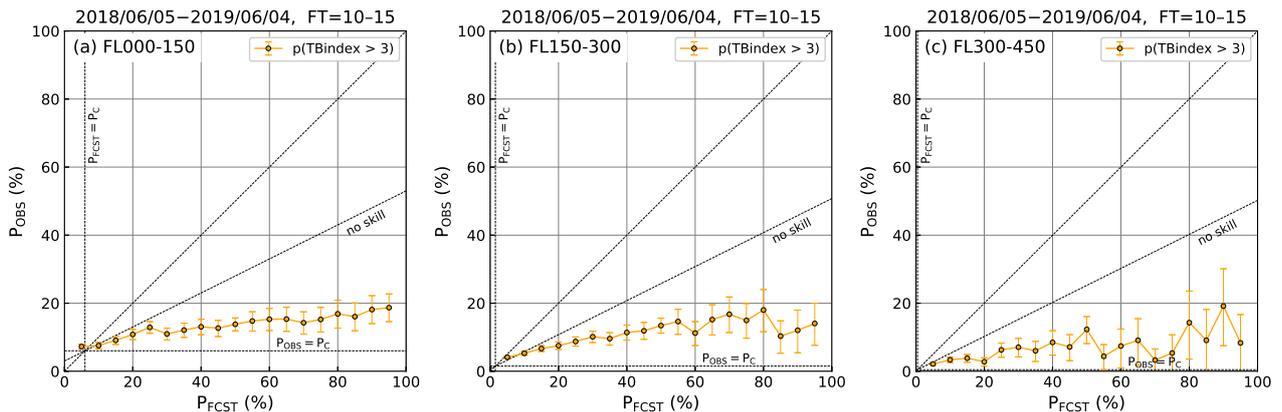


図 2.4.2 MOD 以上の乱気流に対する  $p(\text{TBindex} > 3)$  の信頼度曲線。FT=10-15 の (a) FL000-FL150, (b) FL150-FL300, (c) FL300-FL450 の検証結果。図の見方は付録 D.4.4 を参照。エラーバーはブートストラップ法で求めた 95% 信頼区間。

FT=10-15 では、TBindex(MEAN) でスキルスコアが最大となる閾値は 3.0 付近であり、TBindex(CNTL) と比べて利用上も大きな違いはない。一方 FT=34-39 では、TBindex(MEAN) でスキルスコアが最大となる閾値は 3.0 よりも 0.2 ~ 0.6 小さくなっている。これは予報時間の後半でアンサンブルのばらつきが大きくなり、

アンサンブル平均することで予測のピーク値が小さくなるためである。このため、予報時間後半でアンサンブル平均の TBindex を用いる場合には閾値を下げて利用する必要がある。

鉛直シアアについて、アンサンブル平均とコントロールランを比較すると (図 2.4.1 の VWS)、FL300 以下

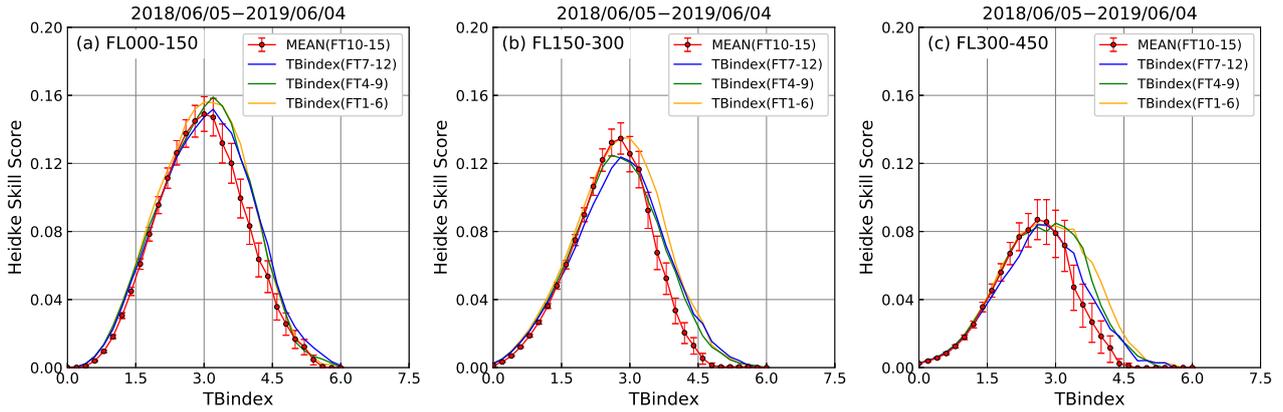


図 2.4.3 MEPS 航空悪天 GPV のアンサンブル平均の乱気流指数と 3, 6, 9 時間後の初期値の国内航空悪天 GPV の乱気流指数の比較。(a) FL000–FL150, (b) FL150–FL300, (c) FL300–FL450 の検証結果。MEAN は MEPS 航空悪天 GPV のアンサンブル平均の乱気流指数を、TBindex は国内航空悪天 GPV の乱気流指数の検証結果を示す。アンサンブル平均の乱気流指数は 00, 06, 12, 18 UTC 初期値の FT=10–15 の検証結果、国内航空悪天 GPV の乱気流指数は、青が 03, 09, 15, 21 UTC 初期値の FT=7–12 の検証結果、緑が 06, 12, 18, 翌日 00 UTC 初期値の FT=4–9 の検証結果、橙色が 09, 15, 21, 翌日 03 UTC 初期値の FT=1–6 の検証結果。MEAN にはブートストラップ法で求めた 95%信頼区間を示すエラーバーを付加している。

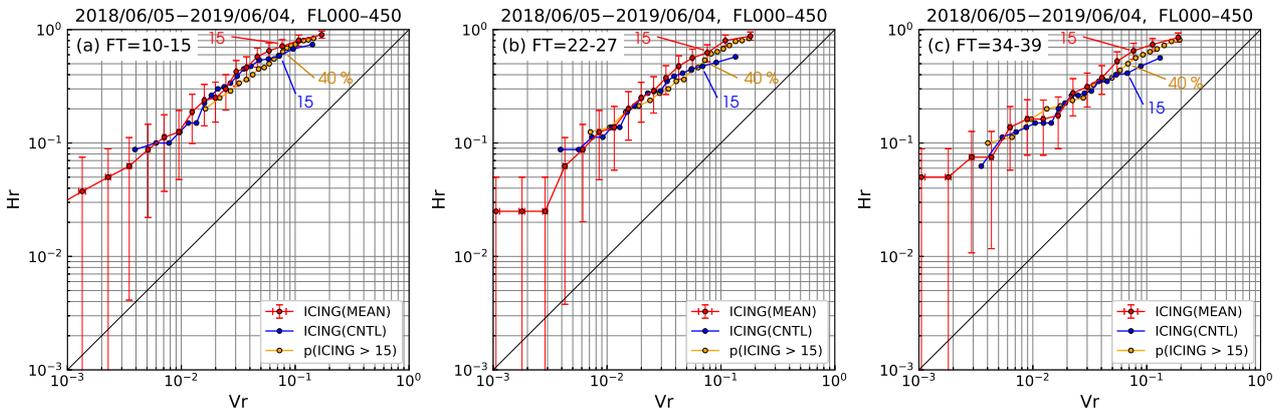


図 2.4.4 MOD 以上の着氷に対する MEPS 航空悪天 GPV の捕捉率 (Hr) – 体積率 (Vr) の関係。(a) FT=10–15, (b) FT=22–27, (c) FT=34–39 の検証結果。ICING (MEAN) は着氷指数のアンサンブル平均、ICING (CNTL) はコントロールランの着氷指数、 $p(\text{ICING} > 15)$  は ICING が 15 を超える超過確率 [%] についての検証結果。ICING (MEAN) にはブートストラップ法で求めた 95%信頼区間を示すエラーバーを付加している。

ではアンサンブル平均の方がスキルスコアのピーク値がやや小さく、予測精度はやや悪いといえる。またアンサンブル平均では、スキルスコアが最大となる閾値はコントロールランよりも低めにシフトしており、コントロールランと同じ閾値で予測することは適切ではない。

図 2.4.2 に、 $p(\text{TBindex} > 3)$  の信頼度曲線を高度別に示す。いずれの高度でも、信頼度曲線が理想直線 (図中の 45 度の直線) よりも下にあって大きく離れており、予測確率が実況と比べて過大であることを示している。 $p(\text{TBindex} > 3)$  はあくまで超過確率であり、値そのものを現象の発生確率と見ることはできない。一方、予測の確率が高くなるほど実況の確率も高くなる傾向があり、超過確率が高いほど乱気流の可能性が高いとみて利用することはできる。

ここまでで示したように、TBindex (MEAN) は TBindex (CNTL) よりも予測精度が高い、すなわち、TBindex (MEAN) は同一初期時刻の国内航空悪天 GPV

の TBindex よりも予測精度が高いといえる。しかしながら国内航空悪天 GPV は 1 日 8 回作成されることから、MEPS 航空悪天 GPV よりも新しい初期値の情報を利用できるというメリットがある。そこで新しい初期値の国内航空悪天 GPV と比較した場合にそれ以前の初期値の TBindex (MEAN) がどの程度有効か確認する。

図 2.4.3 に、00, 06, 12, 18 UTC 初期値の FT=10–15 の TBindex (MEAN) に対する、3 時間後の初期値の FT=7–12、6 時間後の初期値の FT=4–9、および、9 時間後の初期値の FT=1–6 の国内航空悪天 GPV の TBindex のスキルスコアを示す。図から、3 時間後の初期値の予測 (青線) と比べて TBindex (MEAN) は同等または予測精度が高く、特に FL150–300 では、9 時間後の初期値と同程度の精度を持っていることがわかる。図には示さないが、同様の検証をほかの FT に対して行ったところ、TBindex (MEAN) が優位である期間は予報時間後半ほど長くなる傾向が見られた。

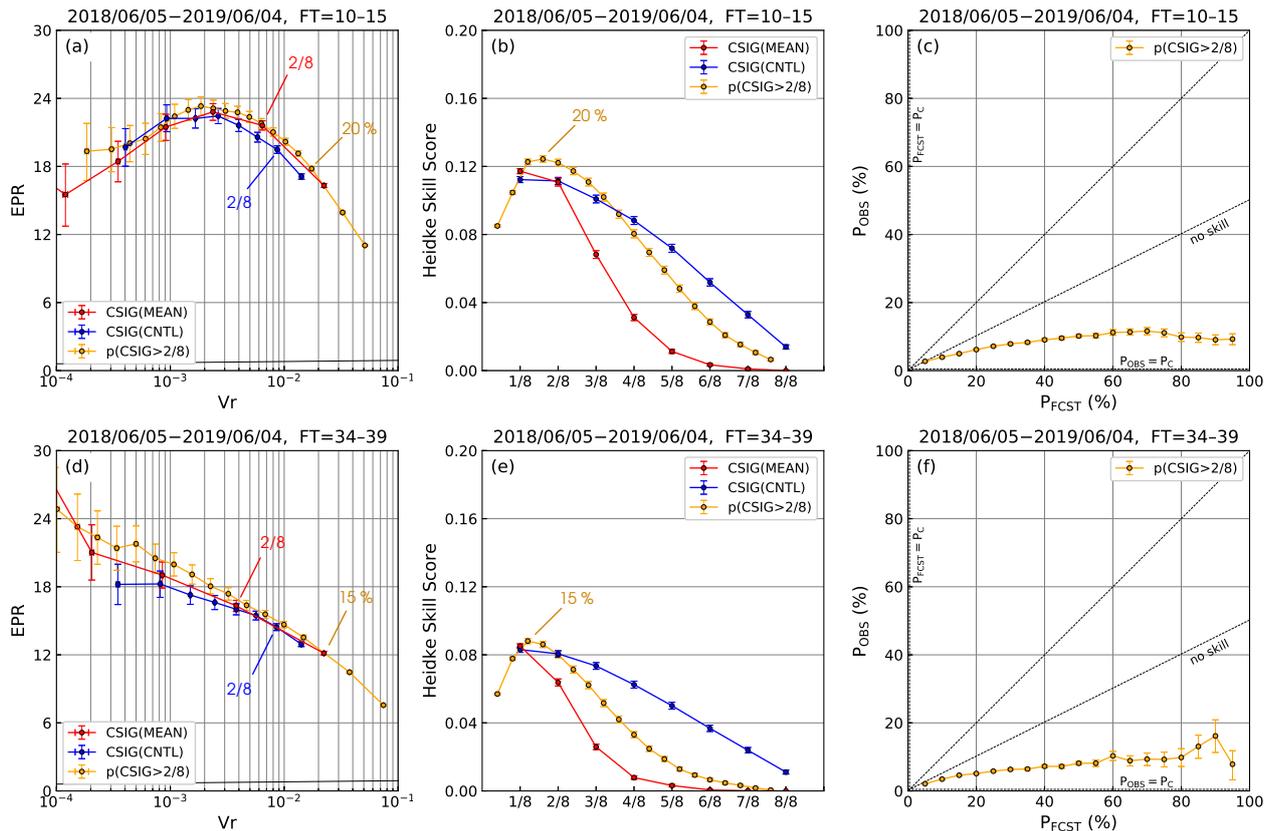


図 2.4.5 MEPS 航空悪天 GPV の積乱雲量の検証結果。(a)~(c) はそれぞれ FT=10-15 の (a) 体積率-遭遇率比、(b) スキルスコア、(c) 信頼度曲線。(d)~(f) はそれぞれ FT=34-39 の (d) 体積率-遭遇率比、(e) スキルスコア、(f) 信頼度曲線。CSIG(MEAN) は積乱雲量のアンサンブル平均、CSIG(CNTL) はコントロールランの積乱雲量、 $p(\text{CSIG} > 2/8)$  は CSIG が  $2/8$  を超える超過確率 [%]。積乱雲量は  $1/8$  から  $8/8$  まで  $1/8$  毎に検証した結果を、 $p(\text{CSIG} > 2/8)$  は 5% から 95% まで 5% 毎に検証した結果を示している。(b), (e) の横軸は積乱雲量  $[0-1]$  または  $p(\text{CSIG} > 2/8)$  [%] を表すが、同一の図に表示するために  $p(\text{CSIG} > 2/8)$  は確率値 [%] を 100 で割った値で示している。信頼度曲線の見方は付録 D.4.4 を参照。エラーバーはブートストラップ法で求めた 95% 信頼区間。

このように、TBindex(MEAN) は新しい初期値の国内航空悪天 GPV と比べても同等以上の予測精度を持つことから、基本的には最新の MEPS 航空悪天 GPV を利用することで精度の高い予測ができる。特に予測場のランダム誤差が大きくなる予報時間の後半では、アンサンブル平均による予測精度の改善幅が大きくなる。

## (2) 着氷指数の検証

ここでは MEPS 航空悪天 GPV から算出した着氷指数のアンサンブル平均と超過確率、コントロールランの検証結果を示す。着氷指数 (ICING) のアンサンブル平均 (ICING(MEAN)) は各メンバーの単純平均で、超過確率 ( $p(\text{ICING} > 15)$ ) は MOD 以上の着氷の目安としている ICING が 15 を超えるメンバーの割合に 100 を掛けた値とする。実況データは C-PIREP による着氷観測を利用し、前項の乱気流実況と同様に扱う。

図 2.4.4 に MOD 以上の着氷に対する予報時間別の捕捉率 (Hr) と体積率 ( $V_r$ ) の関係を示す。着氷は実況の通報数が非常に少ないため、高度別には検証せず、FL000-450 でまとめた検証結果を示している。体積率が等しい予測を比べた場合には捕捉率が大きいほど実況を多く捉えた良い予測であり、捕捉率が等しい予測を

比べた場合には体積率が小さいほど予測領域を限定した良い予測といえるため、図では線が左上にあるほど予測精度が高いといえる。ICING=15 付近の閾値で比べると、ICING(MEAN) は予報時間に関わらずコントロールランの着氷指数 (ICING(CNTL)) の予測精度を上回っている。また、ICING(MEAN) は ICING(CNTL) よりも体積率が若干大きいことから、アンサンブル平均の方がコントロールランよりも広めに予測しているといえるが、それ以上に捕捉率が大きくなり予測精度としては向上している。 $p(\text{ICING} > 15)$  の体積率・捕捉率が ICING=15 と同程度となる超過確率 40% の閾値と比較すると、 $p(\text{ICING} > 15)$  は ICING(CNTL) と同等以上の予測精度を持つが、ICING(MEAN) と比べると精度が低い。

## (3) 積乱雲量の検証

ここでは MEPS 航空悪天 GPV から算出した積乱雲量のアンサンブル平均、超過確率とコントロールランの検証結果を示す。積乱雲量 (CSIG) のアンサンブル平均 (CSIG(MEAN)) は各メンバーの単純平均で、超過確率 ( $p(\text{CSIG} > 2/8)$ ) は国内悪天 12 時間予想図 (FBJP112) で積乱雲域を予測する閾値である CSIG

が2/8を超えるメンバーの割合に100を掛けた値とする。検証では工藤(2007)と同様にLIDENによる雷観測を用い、予報対象時刻の前後30分以内にMEPS航空悪天GPVの格子内で対地雷が1回以上観測された場合に実況ありとする。

図2.4.5に積乱雲量と超過確率の閾値別検証結果を示す。図の上段(a)~(c)はFT=10-15の検証結果を、下段(d)~(f)はFT=34-39の検証結果を示している。図2.4.5(a), (d)は遭遇率比(EPR)と体積率(Vr)の関係である。遭遇率比は付録D.3.1の分割表を用いて、

$$EPR = \frac{FO/(FO + FX)}{(FO + XO)/(FO + FX + XO + XX)} \quad (2.4.3)$$

であり、全領域での積乱雲の発生率(気候値的発生率)に対する積乱雲を予測した領域での実況の発生率の比を表す。同じ体積率の予測を比べた場合、遭遇率比が大きいほど予測精度が高いと見ることができる。図(a), (d)を見ると、いずれの予報時間でも $p(CSIG > 2/8)$ はコントロールラン(CSIG(CNTL))を上回っていることがわかる。閾値2/8付近で比較すれば、CSIG(MEAN)の精度もCSIG(CNTL)よりも高いといえる。しかし、CSIG(CNTL)と比べて体積率が小さくなっており、2/8を閾値とすると予測領域が狭くなってしまふ。

図2.4.5(b), (e)は閾値別のスキルスコアである。スキルスコアの最大値で比べると、いずれのFTでも $p(CSIG > 2/8)$ が最も高く、予測精度が高いことがわかる。スキルスコアが最大となる $p(CSIG > 2/8)$ の閾値は15-20%付近であり、これを目安とすることでCSIG(CNTL)よりも精度が高く最適な予測を行うことができる。CSIG(MEAN)は閾値1/8ではCSIG(CNTL)を多少上回るものの、2/8以上の閾値では同等か精度が悪い。 $p(CSIG > 2/8)$ が15-20%であるときの体積率(図2.4.5(a), (d))を見ると、CSIG(CNTL) = 2/8の体積率よりも大きくなっている。体積率が大きくなって予測精度が高くなる理由は第2.4.4項(3)の事例で示す。

図2.4.5(c), (f)は $p(CSIG > 2/8)$ の信頼度曲線である。いずれの予報時間でも、信頼度曲線は理想直線の下にあって大きく離れており、予測確率が実況と比べて過大であることを示している。一方、確率60%程度までは予測の確率が高くなるほど実況の確率も高くなる傾向を示しており、超過確率が高いほど積乱雲発生の可能性が高くなる傾向があるといえる。 $p(CSIG > 2/8)$ はあくまで閾値を超えたメンバーの割合であり、値そのものを現象の発生確率と見ることはできないが、発生のしやすさを表す指標としては利用できる。

#### 2.4.4 事例検証

以下では統計検証期間中に発生した、乱気流、着氷、積乱雲の事例を示し、それぞれについて、コントロールランとアンサンブル平均、超過確率の予測結果を比較する。

#### (1) 乱気流指数の予測事例

2018年11月22日は低気圧と前線に伴う雲域の雲頂付近でMODの乱気流が多数発生した。図2.4.6は11月22日00UTCの高頻度衛星雲観測プロダクトの雲頂高度情報である。22日00UTCの地上天気図(図略)によると、関東と四国の南海上に前線を伴った低気圧があり、これらの低気圧と前線に伴う雲域が西日本から東日本の広い範囲に広がっている。西日本から関東付近の雲の雲頂高度はFL200-270(20,000-27,000ft)付近で、紀伊半島周辺など一部ではトランスバースバンド状の雲となり雲頂高度がFL300を超えているところもある。

図2.4.7は11月22日00UTCの前後1時間以内にC-PIREPで報じられたMOD以上の乱気流実況である。22日は西日本から関東の広い範囲でMODの乱気流が報じられており、その発生高度は主にFL200-300であった。

図2.4.8に11月21日12UTC初期値のFT=12のMEPS航空悪天GPVによるFL270の乱気流予測を示す。乱気流指数の色はそれぞれ、LGTM, LGT, LGTP, MOD, SEVの目安となる閾値で塗り分けており、黄色がMOD以上の乱気流の目安となるTBindex > 3.0の領域を示している。コントロールランの乱気流指数(左図)では、西日本や関東の東海上での乱気流は捕捉できているものの実況よりも予測領域が狭い。一方アンサンブル平均の乱気流指数(中図)では、東海地方や関東南部など、より広い領域でMOD以上の乱気流が予測されている。また、 $p(TBindex > 3)$ (右図)でも、20-30%を閾値とするとアンサンブル平均と同様に乱気流を捕捉できている。中図や右図では、個別のメンバーの予測が平滑化されることでランダム誤差が軽減され、より実況に近い予測となっている。

図2.4.9に、乱気流発生時刻の直前の初期値である、11月21日21UTC初期値の国内航空悪天GPVによるFT=3, FL270の乱気流指数の予測を示す。最新初期値のMSMに基づく乱気流予測では、12UTC初期値のコントロールラン(図2.4.8左)と比べて予測領域に大きな差はなく、依然として不十分であった。この例のように雲頂付近での鉛直シアの強まりなどに伴う乱気流では、アンサンブル平均や超過確率を利用することで位置ずれが緩和され、最新初期値のMSMの予測よりも予測精度が高くなる場合がある。

図2.4.10は図2.4.8と同じ初期時刻と対象時刻の鉛直シアの予測である。コントロールランの鉛直シアでは、近畿・中国・四国地方の一部で、空域予報で晴天乱気流を予測する目安である12kt/1000ftを超える値が予測されている。一方アンサンブル平均の鉛直シアでは、平均化によりピークが弱まっていることがわかる。この事例のように、鉛直シアのアンサンブル平均で乱気流を予測する場合にはコントロールランよりも低い閾値を用いる必要があるが、統計検証(図

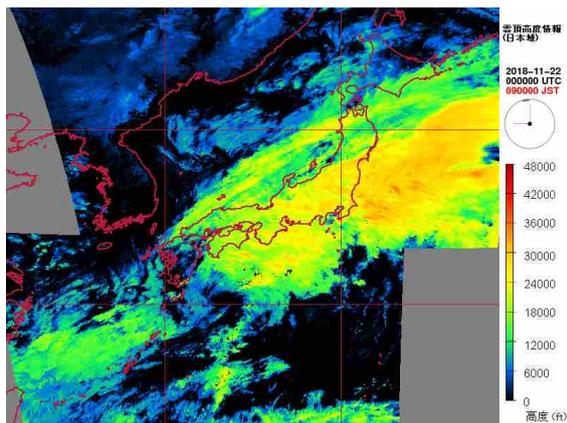


図 2.4.6 2018 年 11 月 22 日 00 UTC の高頻度衛星雲観測プロダクトの雲頂高度情報。

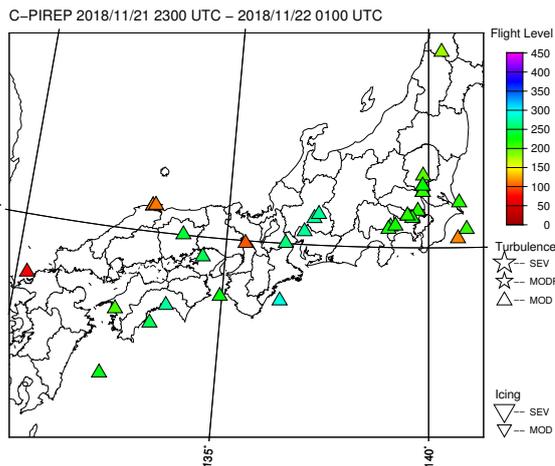


図 2.4.7 2018 年 11 月 22 日 00 UTC の前後 1 時間以内に C-PIREP で通報された MOD 以上の乱気流の実況。△は MOD の乱気流を、色は高度 (FL) を示す。

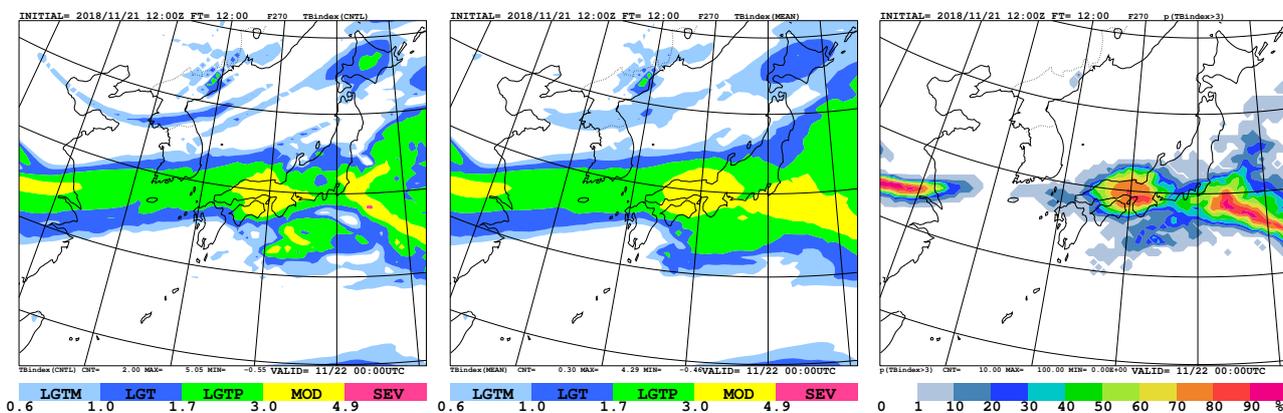


図 2.4.8 2018 年 11 月 21 日 12 UTC 初期値の MEPS 航空悪天 GPV の FT=12, FL270 の乱気流予測。(左) TBindex(CNTL)、(中) TBindex(MEAN)、(右)  $p(\text{TBindex} > 3)$  [%]。

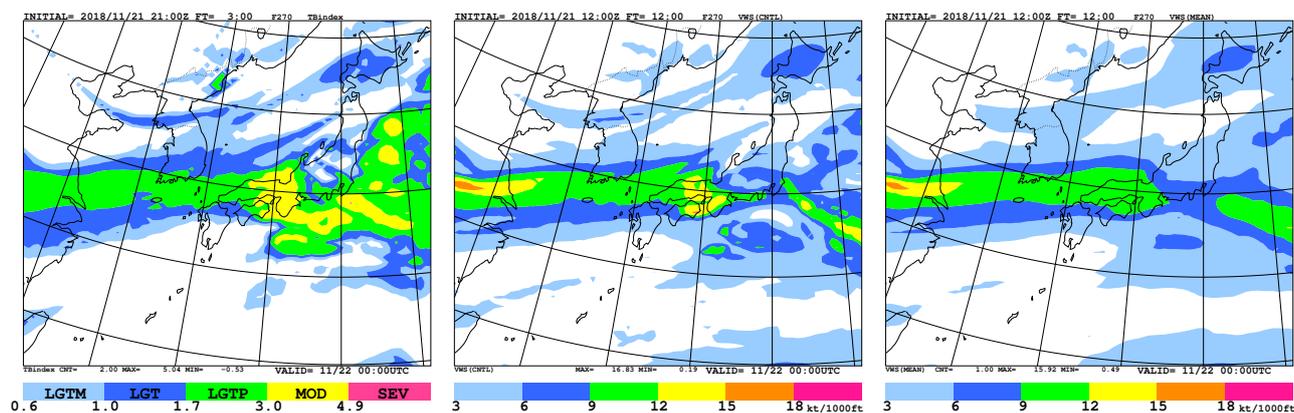


図 2.4.9 2018 年 11 月 21 日 21 UTC 初期値の国内航空悪天 GPV の FT=3, FL270 の TBindex の予測。

図 2.4.10 2018 年 11 月 21 日 12 UTC 初期値の MEPS 航空悪天 GPV の FT=12, FL270 の予測。(左) VWS(CNTL) [kt/1000ft]、(右) VWS(MEAN) [kt/1000ft]。

2.4.1) で示したように、その予測精度はコントロールと比べても低い。

## (2) 着氷指数の予測事例

2018 年 12 月 4 日 07 UTC 頃、和歌山県上空で MOD の着氷が観測された。図 2.4.11 は 12 月 4 日 07 UTC の高頻度衛星雲観測プロダクトの雲頂高度情報である。

4 日 06 UTC の地上天気図 (図略) によると、北海道にある低気圧からのびる寒冷前線が日本海から東シナ海にあり、前線と低気圧に伴う雲域が近畿地方を含む日本の広い範囲に広がっている。

図 2.4.12 は 12 月 4 日 07 UTC の前後 1 時間以内に C-PIREP で報じられた MOD 以上の着氷実況である。

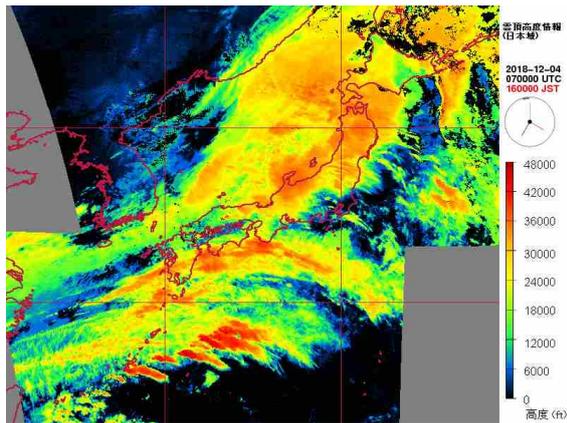


図 2.4.11 2018 年 12 月 4 日 07 UTC の高頻度衛星雲観測プロダクトの雲頂高度情報。

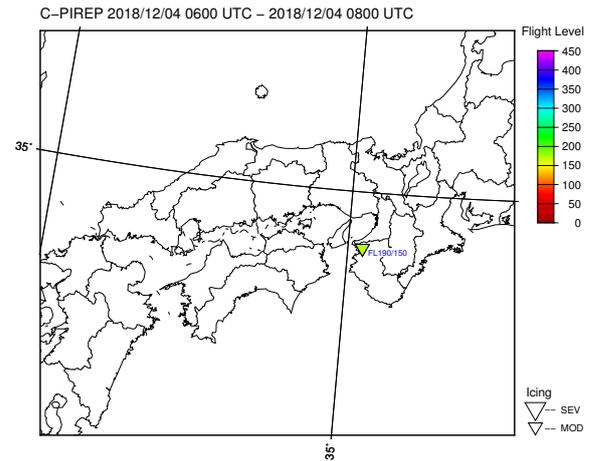


図 2.4.12 2018 年 12 月 4 日 07 UTC の前後 1 時間以内に C-PIREP で通報された MOD 以上の着氷の実況。▽ は MOD の着氷を、色とシンボル横の数値は高度 (FL) を示す。

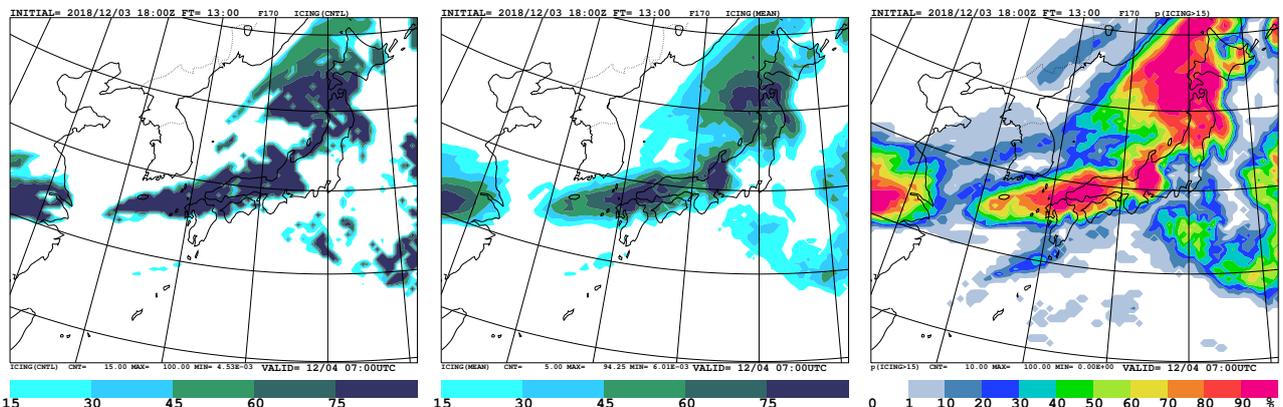


図 2.4.13 2018 年 12 月 3 日 18 UTC 初期値の MEPS 航空悪天 GPV の FT=13, FL170 の着氷予測。(左) ICING(CNTL)、(中) ICING(MEAN)、(右)  $p(\text{ICING} > 15)$  [%]。

和歌山県上空の FL190-150 で MOD の着氷が報じられている。

図 2.4.13 に 12 月 3 日 18 UTC 初期値の FT=13 の MEPS 航空悪天 GPV による FL170 の着氷予測を示す。コントロールランの予測 (左図) では、着氷指数 15 を閾値とした場合、和歌山上空での着氷を捕捉できていない。一方アンサンブル平均の着氷指数 (中図) では、紀伊半島でも 15 以上の予測となっており、着氷を予測できている。また、 $p(\text{ICING} > 15)$  (右図) でも、30-40%を閾値とするとアンサンブル平均と同様に着氷を捕捉できている。

### (3) 積乱雲量の予測事例

2019 年 3 月 12 日は日本海などで対流雲に伴う雷が観測された。図 2.4.14 は 2019 年 3 月 12 日 12 UTC の地上天気図である。日本海に低気圧があり、寒冷前線が対馬海峡にのびている。

図 2.4.15 は 2019 年 3 月 12 日 12 UTC の前後 30 分以内に LIDEN で観測された対地雷の実況である。低気圧中心と寒冷前線付近およびその前面の所々で雷が観測されている。

図 2.4.16 は 2019 年 3 月 12 日 00 UTC 初期値の FT=12 の MEPS 航空悪天 GPV による積乱雲量の予測である。コントロールランの積乱雲量 (左図) は、低気圧と寒冷前線近傍で積乱雲が予測されているが、積乱雲量 2/8 以上の予測領域は LIDEN の実況と比べて狭い。アンサンブル平均の積乱雲量 (中図) でも、低気圧と寒冷前線近傍で積乱雲が予測されているが、アンサンブル平均することで雲量の値が全般に小さくなっており、コントロールランと同じ雲量 2/8 を閾値とすると予測領域はかなり狭くなってしまふ。一方、 $p(\text{CSIG} > 2/8)$  (右図) では、コントロールランと比べて広がりを持った予測となっており、20%程度を閾値とすることで低気圧と前線付近の雷との対応が良い。また、東北の日本海側や北陸、近畿および太平洋上の散発的な雷については、コントロールランやアンサンブル平均ではほとんど捕捉できていないが、右図ではある程度捕捉できている。

現在の積乱雲量の予測では対象時刻の瞬間値における気温や水蒸気の前報を用いており、時間的な広がりにはほぼ考慮していない。このためコントロールランのように時間的な広がりのない予測となってしまうのだ

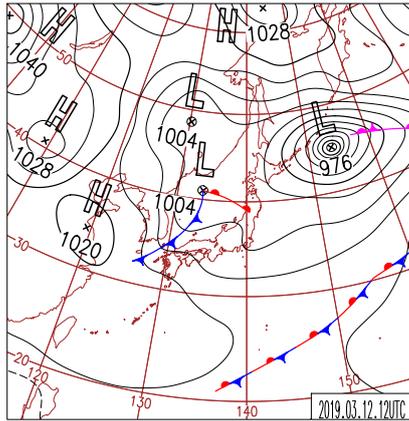


図 2.4.14 2019 年 3 月 12 日 12 UTC の日本域地上天気図。

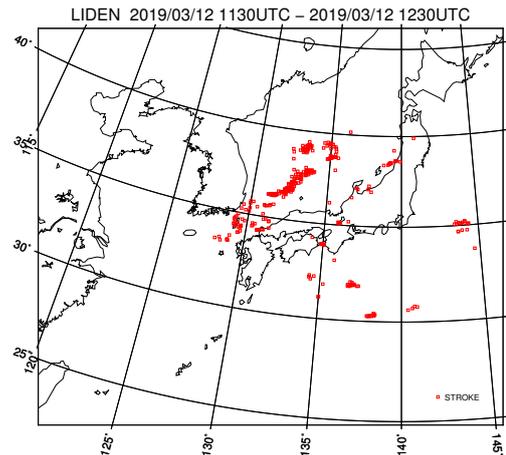


図 2.4.15 2019 年 3 月 12 日 12 UTC の前後 30 分以内に LIDEN で観測された対地雷の実況。

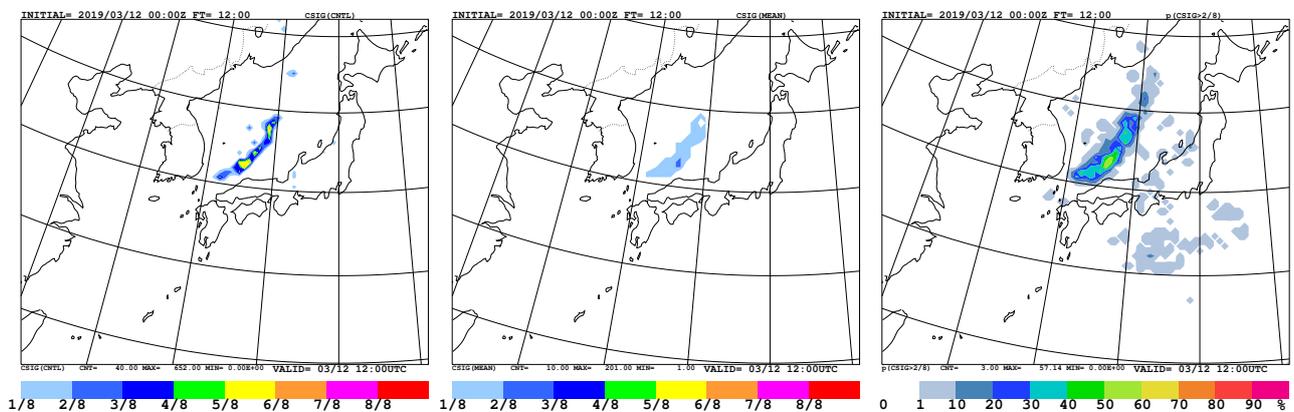


図 2.4.16 2019 年 3 月 12 日 00 UTC 初期値の MEPS 航空悪天 GPV の FT=12 の積乱雲予測。(左) CSIG(CNTL)、(中) CSIG(MEAN)、(右)  $p(\text{CSIG} > 2/8)$  [%]。

が、超過確率ではアンサンブルメンバーのばらつきにより空間的な広がりを持った予測となり、予測精度が高くなる。

### 2.4.5 利用上の留意点

統計検証の結果から、アンサンブル平均の乱気流指数と乱気流指数が 3.0 を超える超過確率は、コントロールランの乱気流指数と比べて予測精度が高く、国内航空悪天 GPV の乱気流指数をこれらに置き換えて利用することが可能である。アンサンブル平均の乱気流指数を用いる場合、空域予報で主に利用する FT=15 程度までは 3.0 を閾値とすることで適切な予測を行うことができるが、それ以降の予報時間では閾値を下げて利用する必要がある。また、乱気流指数が 3.0 を超える超過確率を用いる場合には、20-30%を閾値とすることで適切な予測を行うことができるが、信頼度曲線は理想直線を大きく下回っており、超過確率を現象の発生確率として利用することはできない。

アンサンブル平均の鉛直シアはコントロールランの鉛直シアと同等か精度が悪く、またスキルスコアが最大となる閾値もアンサンブル平均とコントロールランとで異なる。このため鉛直シアを用いて乱気流

を予測する場合には、アンサンブル平均ではなく、コントロールランを利用することを推奨する。

着氷指数のアンサンブル平均は、コントロールランの着氷指数と比べて予測精度が高く、国内航空悪天 GPV の着氷指数を単純に置き換えて利用することが可能である。コントロールランと比べてアンサンブル平均では全般に値が小さくなり、閾値 15 以上の領域は多少広がるが、それ以上に捕捉率が高くなるため予測精度が向上している。

積乱雲量が 2/8 を超える超過確率は、閾値 15-20%でコントロールランの積乱雲量よりも予測精度が高い。ただし信頼度曲線は理想直線から大きく下回っており、超過確率を現象の発生確率として利用することはできない。積乱雲量のアンサンブル平均はコントロールランと比べても予測精度が低く、利用には適さない。

### 2.4.6 まとめ

本節では MEPS 航空悪天 GPV の予測精度と予測事例を示した。MEPS 航空悪天 GPV から算出する乱気流指数のアンサンブル平均 ((2.4.2) 式) と超過確率、着氷指数のアンサンブル平均、積乱雲量の超過確率は、コントロールランすなわち国内航空悪天 GPV を上回

る予測精度を持つ。予報作業においては、国内航空悪天 GPV の予測を MEPS 航空悪天 GPV に置き換えて利用することも可能であり、空域予報の予測精度向上に寄与するものと思われる。本節では示さなかったが、積乱雲の雲頂高度が FL300 を超える超過確率や、LGT 以上の乱気流が予測される超過確率、風の予測誤差が大きい領域の予測などにも MEPS 航空悪天 GPV が利用できると思われる。前項で述べた点に留意し、空域予報などに有効に活用していただきたい。

#### 参考文献

- 井藤智史, 2018a: 着氷指数. 数値予報課報告・別冊第 64 号, 気象庁予報部, 219–225.
- 井藤智史, 2018b: 積乱雲量・雲頂高度. 数値予報課報告・別冊第 64 号, 気象庁予報部, 226–232.
- 工藤淳, 2007: 国内航空用悪天 GPV の積乱雲予測手法の開発. 航空気象ノート第 66 号, 気象庁航空気象管理官, 11–18.
- 工藤淳, 2018a: 乱気流指数. 数値予報課報告・別冊第 64 号, 気象庁予報部, 204–218.
- 工藤淳, 2018b: 国内航空悪天 GPV および LFM 航空悪天 GPV の改良. 平成 30 年度数値予報研修テキスト, 気象庁予報部, 32–41.

## 2.5 その他のガイダンス（降雪、風、発雷確率）<sup>1</sup>

### 2.5.1 はじめに

前節までに MEPS を入力としたガイダンスなどの応用プロダクトである、気温ガイダンス・降水ガイダンス・航空悪天 GPV について解説してきた。本節では、第 2.1 節の MEPS ガイダンスのうち、降雪ガイダンス・風ガイダンス・発雷確率ガイダンスについて紹介する。本節で紹介するガイダンスは、統計検証や事例検証についてこれから新たな知見が蓄積されていくプロダクトではあるが、これまでの検証から得られた結果や利用方法について紹介する。

次項から降雪ガイダンス、風ガイダンス、発雷確率ガイダンスの順に仕様と統計検証を述べる。仕様では各 MEPS ガイダンスで作成する要素と作成方法を、統計検証では MSM ガイダンスとの予測特性の違いを中心に述べる。降雪ガイダンスと発雷確率ガイダンスについては、統計検証の後で事例を紹介する。最後に、第 2.5.5 項でガイダンス毎の利用上の留意点をまとめる。

### 2.5.2 降雪ガイダンス

#### (1) 仕様

MEPS 降雪ガイダンスでは、MSM 降雪ガイダンスと同様に降水種別、降雪量（格子）<sup>2</sup>、降雪量地点の 3 つのガイダンスを作成する。各ガイダンスは対応する MSM ガイダンスと同じ手法を MEPS の各メンバーに適用することにより作成する。ただし、降雪量地点ガイダンスではニューラルネットワークおよび頻度バイアス補正の係数に MSM ガイダンスの係数を利用し、MEPS 降雪ガイダンスでは係数更新を行わない。予測手法の詳細については土田（2018a）を参照頂きたい。

#### (2) 統計検証

MEPS 降雪ガイダンスの統計検証の結果を示す。はじめに降水種別ガイダンスの予報要素および超過確率の検証結果、次に降雪量（格子）、降雪量地点の順に両ガイダンスの予報要素の検証結果、最後に降雪量（格子）および降雪量地点ガイダンスの超過確率の検証結果を示す。検証期間は 2018 年 12 月～2019 年 2 月である。

#### 降水種別ガイダンス

全国の地上気象官署（特別地域気象観測所を含む）における天気観測を降水種別とした降水種別ガイダンスの検証結果を示す。まず、降水種別のカテゴリ検証の結果を示す。地上気温の観測値が  $-3 \sim 5^{\circ}\text{C}$  かつ

<sup>1</sup> 小林 大輝（第 2.5.1 項、第 2.5.5 項）、土田 尚侑（第 2.5.2 項、第 2.5.4 項）、井藤 智史（第 2.5.3 項）

<sup>2</sup> これまで「最大降雪量ガイダンス」と呼称していたが、本テキスト以降は「降雪量ガイダンス」と呼称する。ただし、本節のように降雪量地点ガイダンスと区別する必要がある場合は「降雪量（格子）ガイダンス」と格子形式のガイダンスであることを注記する。

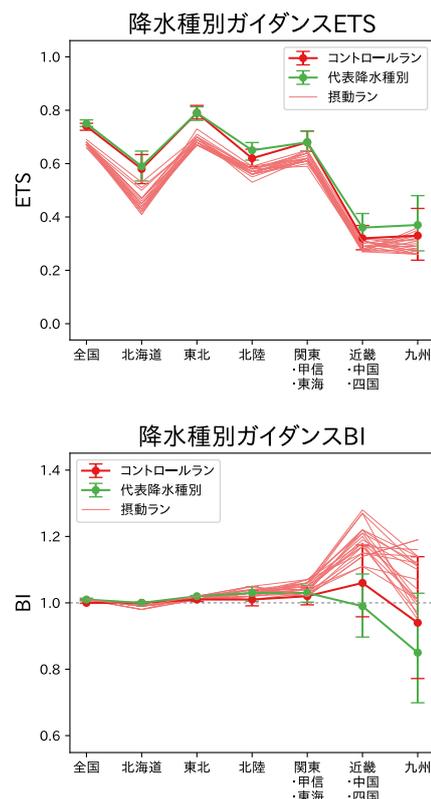


図 2.5.1 降水種別ガイダンスの雪予測の地域別 ETS（上）および BI（下）の比較。赤線は各メンバー（太線がコントロールラン、細線が摂動ラン）、緑線は代表降水種別で、それぞれ左から全国、北海道、東北、北陸、関東・甲信・東海、近畿・中国・四国、九州の順に地域ごとの検証結果を示す。エラーバーはブートストラップ法による 95%信頼区間を示す。

降水（みぞれ・あられなどは除く）が観測されている場合（天気が霧雨、地雨、しゅう雨、地雪、しゅう雪の場合）のみを検証対象とし、降水種別を「雨」と「雪」の 2 カテゴリに分類して実施した。実況の降水種別は、天気が霧雨、地雨、しゅう雨を「雨」、地雪、しゅう雪を「雪」とし、予報の降水種別は、「雨か雪」を「雨」、「雪か雨」を「雪」として扱う。なお、アンサンブルメンバーを代表する降水種別（以下、代表降水種別）は、「雨か雪」または「雨」のメンバーが半数以上を占めれば「雨」、「雪か雨」または「雪」のメンバーが半数以上を占めれば「雪」とした検証結果を示す。検証地点の予報値には最近接格子の値を用いた。また、検証は 00, 06, 12, 18UTC 初期値の FT=15 ~ 36 に対して実施した。図 2.5.1 に「雪」予測を現象あり、「雨」予測を現象なしとした地域別のエクイタブルスコア (ETS) およびバイアススコア (BI) を示す。

代表降水種別について、ETS は北陸と西日本でコントロールランを上回っており、雪予測を改善する。そのほかの地域ではコントロールランと同等である。BI は北陸でコントロールランより高く、雪予測を増やしている一方、西日本ではコントロールランよりも低く、雪

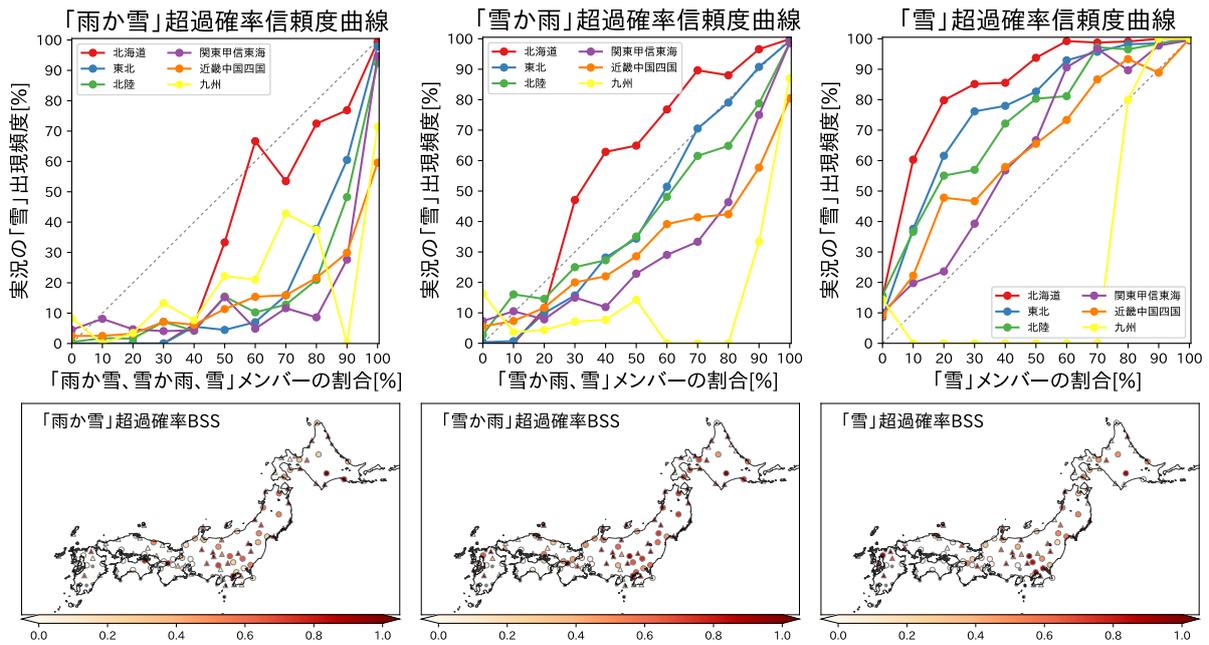


図 2.5.2 降水種別ガイダンスの雪予測に対する超過確率の信頼度曲線（上段）と BSS のスコアマップ（下段）。左から「雨か雪」超過確率、「雪か雨」超過確率、「雪」超過確率の順に示す。信頼度曲線は赤線が北海道、青線が東北、緑線が北陸、紫線が関東・甲信・東海、橙線が近畿・中国・四国、黄線が九州の各地域毎の検証結果を示す。BSS のスコアマップで丸印は有人官署、三角印は特別地域気象観測所を表し、灰色の点は検証期間中に対象となるカテゴリを超える予測と雪の観測がともになかった地点を表す。

予測を減らしている。そのほかの地域ではコントロールランと同等である。コントロールランでは BI が 1 を超えていた近畿・中国・四国では 1 に近づき適切な頻度での予測となったが、1 を下回っていた九州では予測頻度がより過小な傾向となっている。なお、西日本で雪予測を減らしているにもかかわらず ETS がコントロールランを上回っているのは、捕捉数を維持しつつ空振りを減らしているためである（図略）。

摂動ランについて、各メンバーの ETS はコントロールランと比較して全国的に低くなっており、統計的な予測精度は低いことがわかる。各メンバーの BI は北日本ではコントロールランと同等であるが、東日本～西日本ではコントロールランよりも大きく予測頻度が過大になっている<sup>3</sup>。

次に「雪」の観測に対して各カテゴリを超える超過確率の検証結果を示す。検証対象の初期値や観測値による対象事例の抽出方法はカテゴリ検証と同じである。図 2.5.2 に、実況の「雪」出現頻度に対する「雨か雪」・「雪か雨」・「雪」を予測したメンバーの割合（以下、「雨か雪」超過確率）、「雪か雨」・「雪」を予測したメンバーの割合（以下、「雪か雨」超過確率）、「雪」を予測したメンバーの割合（以下、「雪」超過確率）の地域別の信頼度曲線および BSS のスコアマップを示す。

「雨か雪」超過確率の信頼度曲線は北海道の 60% 以上では傾き 45 度の理想直線に近く比較的予測頻度が適切であるが、北海道の 50% 以下とそのほかの地域では理想直線の下にあり、予測頻度が過大である。「雪か雨」超過確率の信頼度曲線は東北の 70% 以上では理想直線上にあり予測頻度が適切である。北海道では 30% 以上で理想直線より上にあり予測頻度が過小、東北の 60% 以下とそのほかの地域では理想直線の下にあり予測頻度が過大である。「雪」超過確率の信頼度曲線は実況の事例がなかった九州の 70% 以下を除き理想直線の上であり、予測頻度が過小である。

ここで各カテゴリを超える超過確率の BSS のスコアマップ<sup>4</sup>を比較すると、北海道から近畿・中国・四国までは「雪か雨」超過確率のスコアが高く、九州では「雪」超過確率のスコアが高い地点が多い。地域ごとにまとめて検証した結果では、北海道、東北、北陸、関東・甲信・東海は「雪か雨」超過確率、近畿・中国・四国は「雪か雨」超過確率および「雪」超過確率、九州では「雪」超過確率のスコアがそれぞれ最も高くなっている（図略）。この結果は各地域の信頼度曲線が最も理想直線に近くなるカテゴリとも対応が良い。各地域で BSS が最も高くなるカテゴリの超過確率を、代表降水

<sup>3</sup> 北日本ではコントロールランと比べて空振り事例と見逃し事例がどちらも多い一方で、東日本～西日本ではコントロールランと比べて空振り事例のみが多く、見逃し事例は同等のためである（図略）。

<sup>4</sup> 自動観測地点では地上気温などに応じた雨雪判別を行うが、雨雪予測が難しい 0℃ 付近では「みぞれ」が観測される事が多い。しかし「みぞれ」は検証対象に含まれないことから、自動観測地点では目視観測地点よりも BSS が高めの傾向があることに留意していただきたい。

種別による雪予測の補助資料としていただきたい。

以上より降水種別ガイダンスでは、代表降水種別を用いてカテゴリ予測を行った場合に統計的な予測精度が最も高くなる。また、北海道、東北、北陸、関東・甲信・東海では「雪か雨」超過確率、近畿・中国・四国では「雪か雨」または「雪」超過確率、九州では「雪」超過確率を目安とすることで雪予測の補助資料として利用できる。

### 降雪量（格子）ガイダンス

全国のアメダス地点で観測された1時間積雪深差から求めた降雪量を観測値とした6, 12, 24時間の降雪量（格子）ガイダンスの検証結果を示す。まずアンサンブル平均および摂動ランの検証結果について、図2.5.3および図2.5.4に降雪量閾値別のETSおよびBIを示す。検証地点の予測値は、検証地点の周囲の4格子点の値を線形内挿することにより求めた。検証対象は00, 06, 12, 18UTC初期値で、予報時間は6時間降雪量がFT=9～30、12時間降雪量がFT=15～36、24時間降雪量がFT=27～36を対象とした。

アンサンブル平均について、ETSは6, 12, 24時間降雪量のそれぞれで3 cm/6h, 5 cm/12h, 10 cm/24h程度までの降雪量を閾値とした場合はコントロールランと同等であるが、それ以上の降雪量を閾値とした場合にはコントロールランを下回っている。BIは6, 12, 24時間降雪量のそれぞれで3 cm/6h, 5 cm/12h, 10 cm/24h程度まではコントロールランより高く、それよりも大きな降雪量閾値では低くなっている。アンサンブル平均では各メンバー予測値が平滑化されるため、強い降雪が予測されにくい。

摂動ランについて、ETSはほぼすべての降雪量閾値に対してコントロールランを下回っており、統計的な予測精度は低い。BIは概ねコントロールランの周りでばらついており、6, 24時間降雪量のそれぞれで30 cm/6h, 40 cm/24h以上および12時間降雪量の30～40 cm/12hを閾値とした場合にはコントロールランよりも低くなる傾向が見える。

次にアンサンブル最大の検証結果について、図2.5.5に降雪量閾値別の捕捉率、図2.5.6に降雪量閾値別のETSとBIを示す。検証設定はアンサンブル平均および摂動ランと同様だが、各初期値、予報時間に対し都道府県毎<sup>5</sup>の最大の予測値と観測値を用いた検証を実施した（地域最大の予測精度検証）。

アンサンブル最大の捕捉率は全閾値でコントロールランを上回っており、コントロールランで予測できていない大雪を捕捉可能なことがわかる。また、ETSは6時間降雪量の7 cm/6h以上でコントロールランを上回っており、12, 24時間降雪量でも15 cm/12h, 30 cm/24h以上でコントロールランと同等か上回っていることか

<sup>5</sup> 北海道は地方毎。関東および東海の太平洋側と四国、九州は検証対象外。

ら、都道府県単位を対象とした大雪可能性の検討資料として利用できると考えられる。また、6時間降雪量はコントロールランのBIが大雪予測に対して1を大きく下回っているため、アンサンブル最大でBIが1に近づき適切な予測頻度となっている。そのため、6時間降雪量は都道府県単位の短時間の大雪に対する定量的な資料としても利用可能である。

以上より、降雪量（格子）ガイダンスはアンサンブル最大を用いた大雪可能性の検討資料として利用できる。特に6時間降雪量では定量的な大雪予測資料としても利用できる。ただし地点単位規模では予測頻度がかなり過大となるため（図略）、今回検証図を示した都道府県単位（北海道は地方単位）規模以上での利用を推奨する。

### 降雪量地点ガイダンス

全国のアメダス地点で観測された1時間積雪深差から求めた降雪量による6, 12, 24時間の降雪量地点ガイダンスの検証結果を示す。まずアンサンブル平均および摂動ランの検証結果について、図2.5.7および図2.5.8に降雪量閾値別のETSおよびBIを示す。検証対象の初期値、予報時間は降雪量（格子）ガイダンスと同様である。

アンサンブル平均について、ETSは6時間降雪量では7 cm/6h程度まではコントロールランを上回っているか同等であるが、それ以上の降雪量を閾値とした場合はコントロールランを下回る。一方12, 24時間降雪量は全ての閾値でコントロールランを上回っているか同等である。BIは6, 12, 24時間降雪量のそれぞれで3 cm/6h, 5 cm/12h, 10 cm/24h程度まではコントロールランよりも高く、それよりも大きな降雪量閾値では低くなっており、アンサンブル平均することで強い降雪が予測されにくくなることを示している。なお、第1.4節でも示されている通りアンサンブル平均では予測値が平滑化されるため、予測の不確実性が大きい現象は表現できない。降雪量地点ガイダンスでも予測のばらつきが大きな事例では大雪が表現できないため、アンサンブル平均の利用は推奨しない。

摂動ランについて、ETSはほぼすべての降雪量閾値に対してコントロールランを下回っている。BIは概ねコントロールランの周りでばらついており、6, 12, 24時間降雪量のそれぞれで30 cm/6h, 30 cm/12h, 40 cm/24h以上を閾値とした場合にはコントロールランよりも低くなる傾向が見える。

次にアンサンブル最大の検証結果について、図2.5.9に降雪量閾値別の捕捉率、図2.5.10に降雪量閾値別のETSとBIを示す。検証設定は降雪量（格子）ガイダンスと同様である。

アンサンブル最大の捕捉率は全閾値でコントロールランを上回っており、コントロールランで予測できていない大雪を捕捉可能なことがわかる。しかしながら

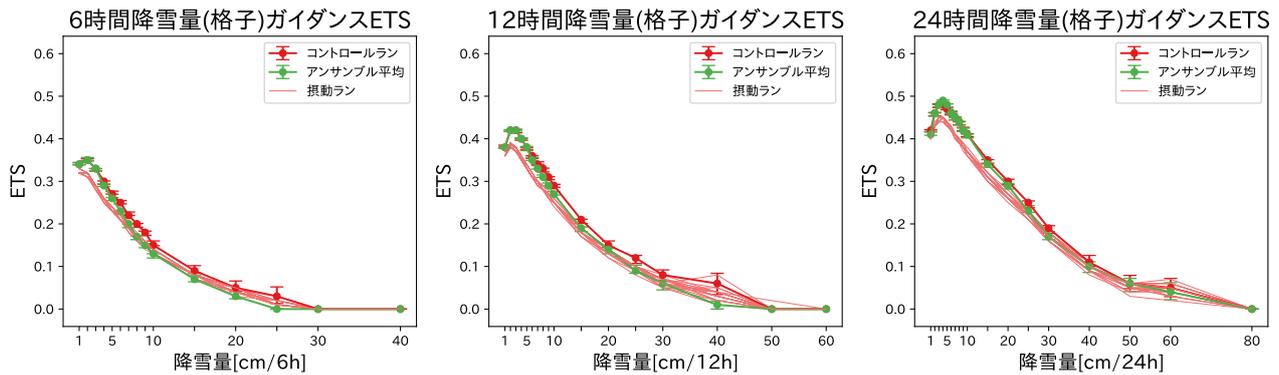


図 2.5.3 降雪量（格子）ガイダンスの降雪量閾値別の ETS。左図が 6 時間降雪量、中図が 12 時間降雪量、右図が 24 時間降雪量を対象とした検証で、線種は図 2.5.1 と同じである。

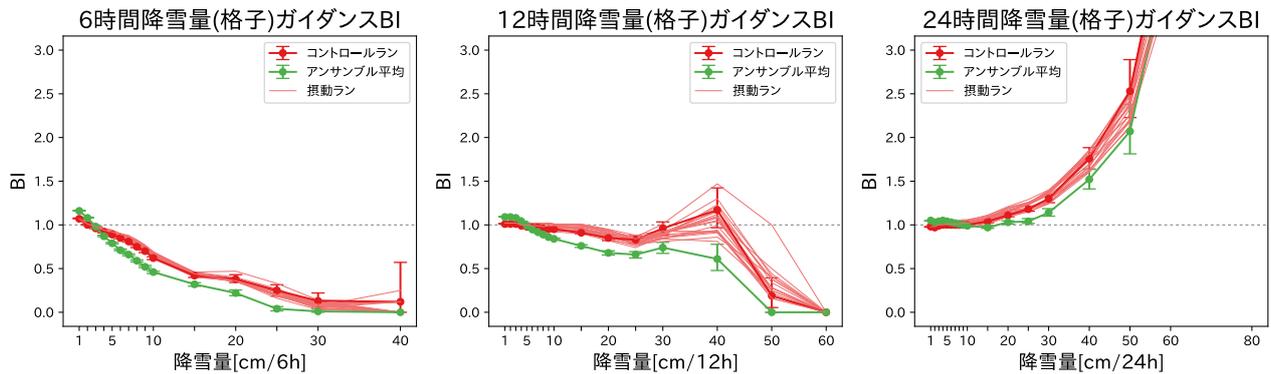


図 2.5.4 降雪量（格子）ガイダンスの降雪量閾値別の BI。図の見方は図 2.5.3 と同じ。

ETS は全閾値でコントロールランを下回っており、定量的な資料としてはコントロールランを改善しない。BI がコントロールランよりもかなり大きく予測過大傾向を示しており、捕捉の改善よりも空振りの増加の影響が大きいと考えられる。コントロールランの時点で捕捉率が高いことから、アンサンブル最大では予測過大となる事例が多いと考えられる。低気圧の接近に伴う降雪事例など、予測のばらつきが大きくコントロールランでの捕捉率が比較的低い事例において、大雪可能性の検討資料として利用いただきたい。

以上より、降雪量地点ガイダンスでは予測のばらつきが大きい事例において、アンサンブル最大がコントロールランの補助資料となり、都道府県単位規模以上での大雪可能性の検討に利用できる。降雪量（格子）ガイダンスのアンサンブル最大よりも捕捉率が高いことから、降雪量（格子）ガイダンスのアンサンブル最大が示す大雪範囲を広げる場合の検討資料などに利用いただきたい。

#### 降雪量超過確率

降雪量（格子）および降雪量地点ガイダンスの超過確率の検証結果を示す。検証対象の初期値は 00, 06, 12, 18UTC で、予報時間は信頼度曲線が各ガイダンスの検証と同じ FT を、予報時間別の BSS は 6 時間降雪量が

FT=9 ~ 39、12 時間降雪量が FT=15 ~ 39、24 時間降雪量が FT=27 ~ 39 のそれぞれ 3 時間毎を対象とした。図 2.5.11 に 6, 12, 24 時間降雪量の超過確率の信頼度曲線と予報時間別 BSS を示す。それぞれ破線が降雪量（格子）ガイダンス、実線が降雪量地点ガイダンスである。

信頼度曲線を見ると、両ガイダンスとも図に示した全閾値について傾き 45 度の理想直線より下にあり、予測頻度が過大であることがわかる。また、閾値が大きいほど理想直線から下に離れており、予測頻度がより過大になる。ただし両ガイダンスを同じ閾値同士で比較すると、降雪量地点ガイダンスの方が理想直線に近く、超過確率と実況の出現頻度との対応が良い。ここで予報時間別の BSS を見ると、両ガイダンスを同じ閾値同士で比較した時に降雪量地点ガイダンスの方が予測精度が高く、また降雪量地点ガイダンスの方がより大きな閾値まで BSS>0 であり気候値予測に対して予測スキルを持っていることがわかる。面的な確率分布を降雪量（格子）ガイダンスで把握し、超過確率の値は降雪量地点ガイダンスを用いて大雪可能性を検討するといった利用方法が有効であると考えられる。ただし、予測のばらつきが大きい事例では、アンサンブル平均によって現象が平滑化されるように、超過確率は各地点において低い値を示すことしかできず、有効に

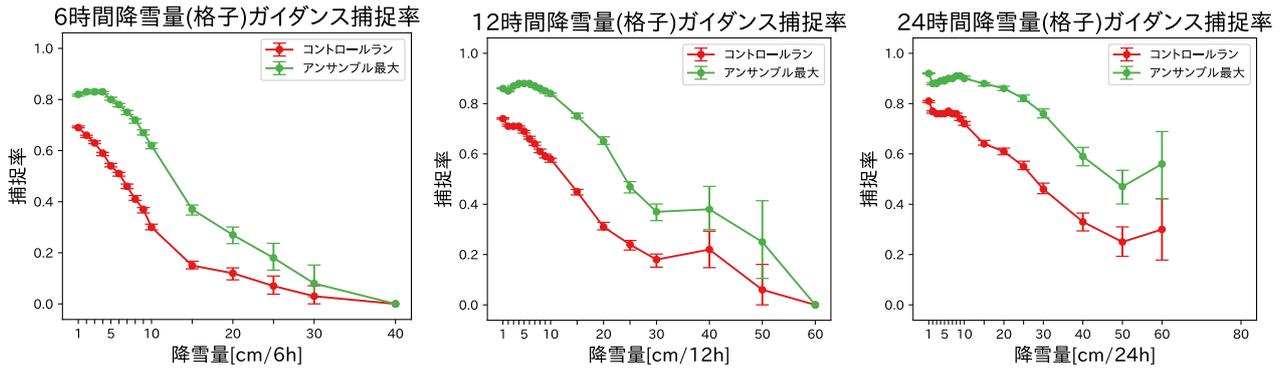


図 2.5.5 降雪量（格子）ガイダンスのアンサンブル最大の降雪量閾値別の捕捉率。左図が6時間降雪量、中図が12時間降雪量、右図が24時間降雪量を対象とした検証で、赤線がコントロールラン、緑線がアンサンブル最大である。エラーバーはブートストラップ法による95%信頼区間を示す。

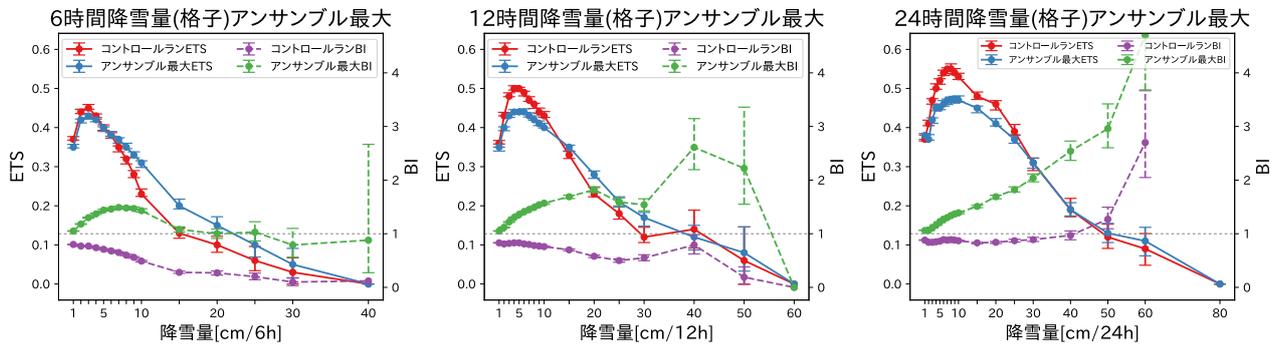


図 2.5.6 降雪量（格子）ガイダンスのアンサンブル最大の降雪量閾値別の ETS と BI。左図が6時間降雪量、中図が12時間降雪量、右図が24時間降雪量を対象とした検証で、赤線がコントロールランの ETS、青線がアンサンブル最大の ETS で、紫線がコントロールランの BI、緑線がアンサンブル最大の BI である。エラーバーはブートストラップ法による95%信頼区間を示す。

利用することはできない。そのような事例では降雪量（格子）および降雪量地点ガイダンスのアンサンブル最大を利用して大雪可能性を検討いただきたい。なおここで示した検証結果は全国を対象としたものであるが、地域別で見ると降雪量地点ガイダンスが気候値予測に対して予測スキルを持つ閾値は北海道で10 cm/12h、東北で30 cm/12h、北陸、関東・甲信・東海、近畿・中国・四国で20~30 cm/12h程度までである（図略）。

### (3) 事例

最後に、降雪量（格子）および降雪量地点ガイダンスでアンサンブル最大が大雪可能性を示すことができた事例を紹介する。図 2.5.12 は2018年2月5日から7日にかけて北陸地方の平野部を中心に大雪となった事例について、石川県から福井県にかけて降雪が強まった時間帯の予測事例である。6日03UTCまでの12時間で石川県から福井県にかけての平野部で40 cmを超える大雪を観測したが、MSM ガイダンスは平野部では40 cmを超える降雪は予測できなかった。また、特に降雪が強まる位置の予測は降雪量（格子）ガイダンスでは南側にずれていた。これに対し MEPS ガイダンスのアンサンブル最大は降雪量（格子）、地点ガイダ

ンスともに40 cmを超える値を予測しており、コントロールランの予測値よりも大雪となる可能性を示していた。また、降雪量（格子）ガイダンスでは実際に大雪となった石川県～福井県の平野部で降雪が特に強まる可能性を示すことができている。

## 2.5.3 風ガイダンス

### (1) 仕様

MEPS 風ガイダンスでは、国内のアメダス観測地点および空港地点の地上風を予測している。ガイダンスの種類として、正時の風速と風向を予測する定時風ガイダンス、前3時間または前1時間内の最大風速とその風向を予測する最大風速ガイダンス、前3時間内にガストが発生する確率を予測するガスト発生確率ガイダンス、前1時間内のガスト風速とその風向を予測するガスト風速ガイダンス、前1時間内の最大瞬間風速とその風向を予測する最大瞬間風速ガイダンスがある。

MEPS 風ガイダンスの作成手法は、MSM 風ガイダンスと同じ手法を用いているため詳しい内容は松本(2018)を参照して頂きたい。なお、MEPS 風ガイダンスの係数はMSM 風ガイダンスで学習した係数を利用し、各アンサンブルメンバーに同じ係数を使用して

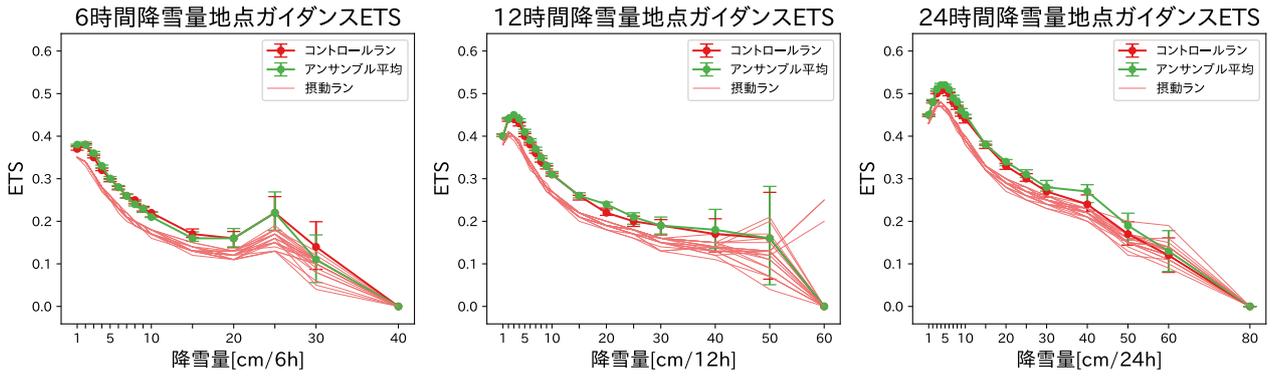


図 2.5.7 降雪量地点ガイダンスの降雪量閾値別の ETS。図の見方は図 2.5.3 と同じ。

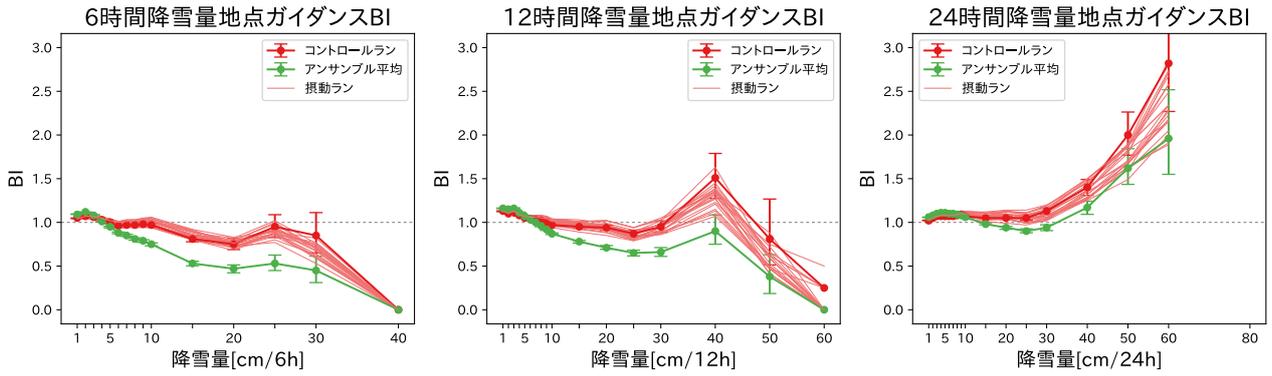


図 2.5.8 降雪量地点ガイダンスの降雪量閾値別の BI。図の見方は図 2.5.3 と同じ。

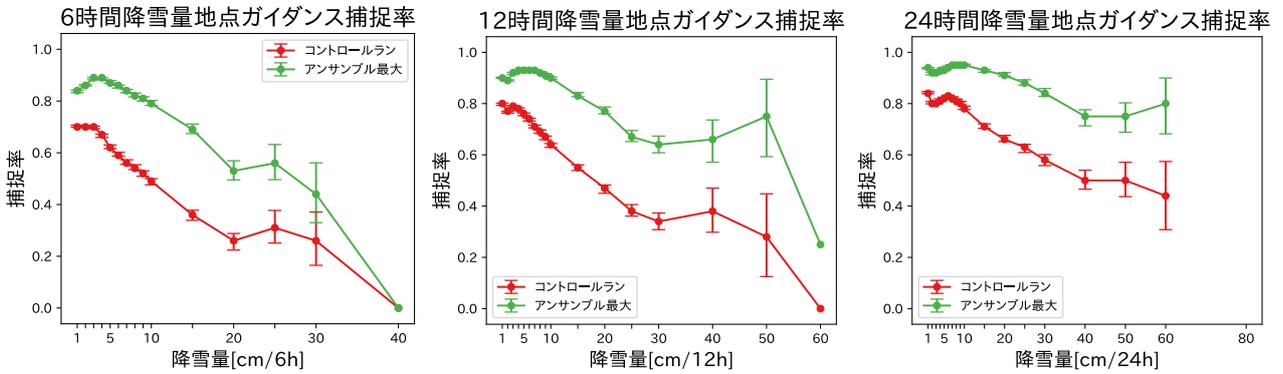


図 2.5.9 降雪量地点ガイダンスのアンサンブル最大の降雪量閾値別の捕捉率。図の見方は図 2.5.5 と同じ。

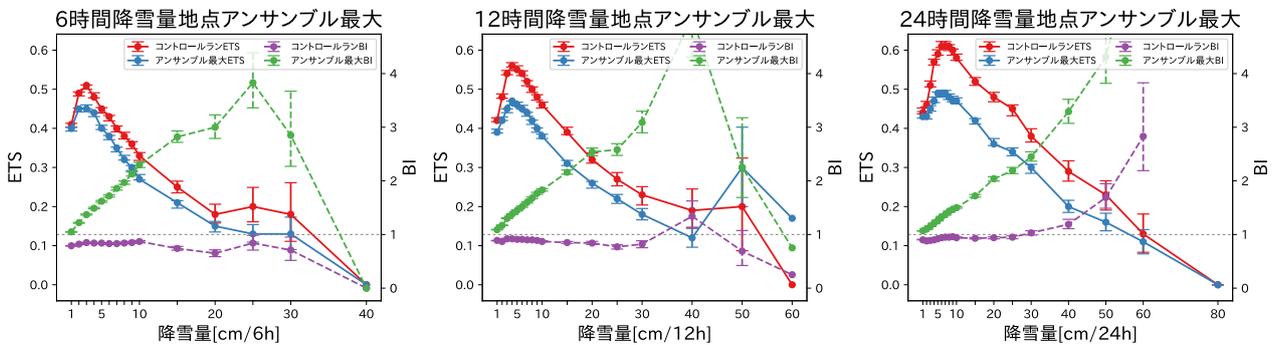


図 2.5.10 降雪量地点ガイダンスのアンサンブル最大の降雪量閾値別の ETS と BI。図の見方は図 2.5.6 と同じ。

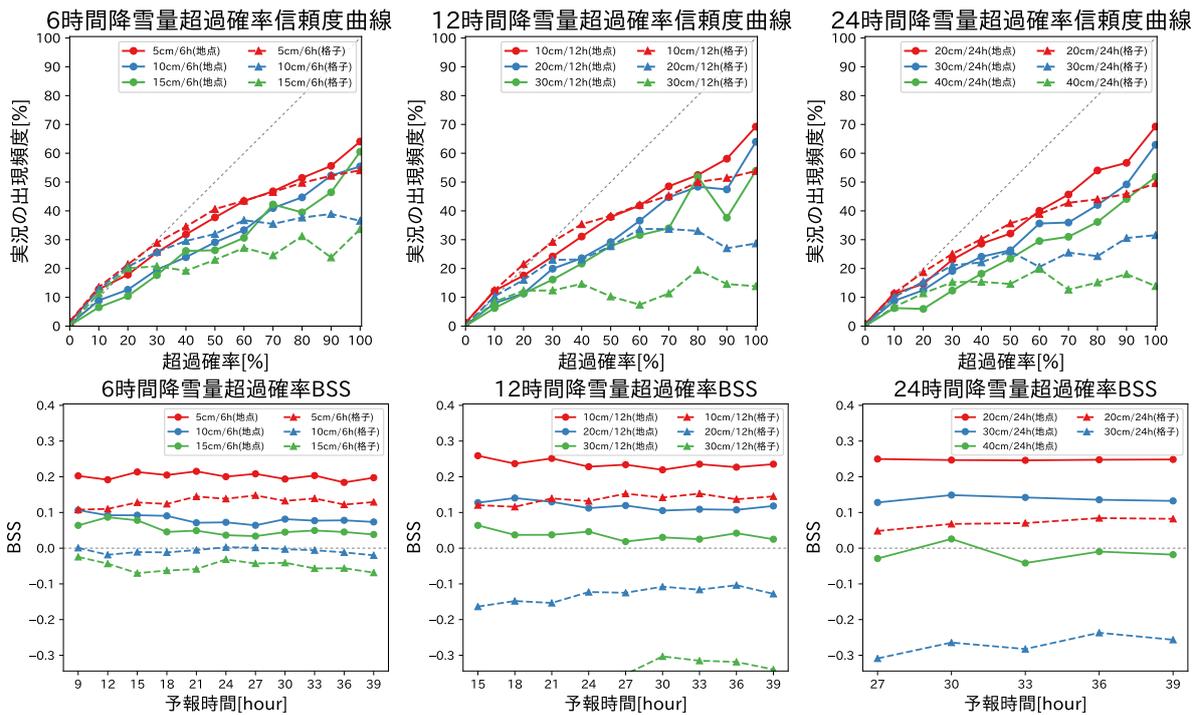


図 2.5.11 降雪量（格子）ガイダンスおよび降雪量地点ガイダンスの超過確率の信頼度曲線（上段）と予報時間別の BSS（下段）。左図が 6 時間降雪量、中図が 12 時間降雪量、右図が 24 時間降雪量を対象とした検証を示す。破線が降雪量（格子）ガイダンス、実線が降雪量地点ガイダンスを示し、それぞれ 6 時間降雪量は赤線が 5 cm/6h、青線が 10 cm/6h、緑線が 15 cm/6h、12 時間降雪量は赤線が 10 cm/12h、青線が 20 cm/12h、緑線が 30 cm/12h、24 時間降雪量は赤線が 20 cm/24h、青線が 30 cm/24h、緑線が 40 cm/24h を閾値とした超過確率の検証結果を示す。（降雪量（格子）ガイダンスの 40 cm/24h の検証結果は負の方向に大きくスケールアウトしているため非表示とした。）

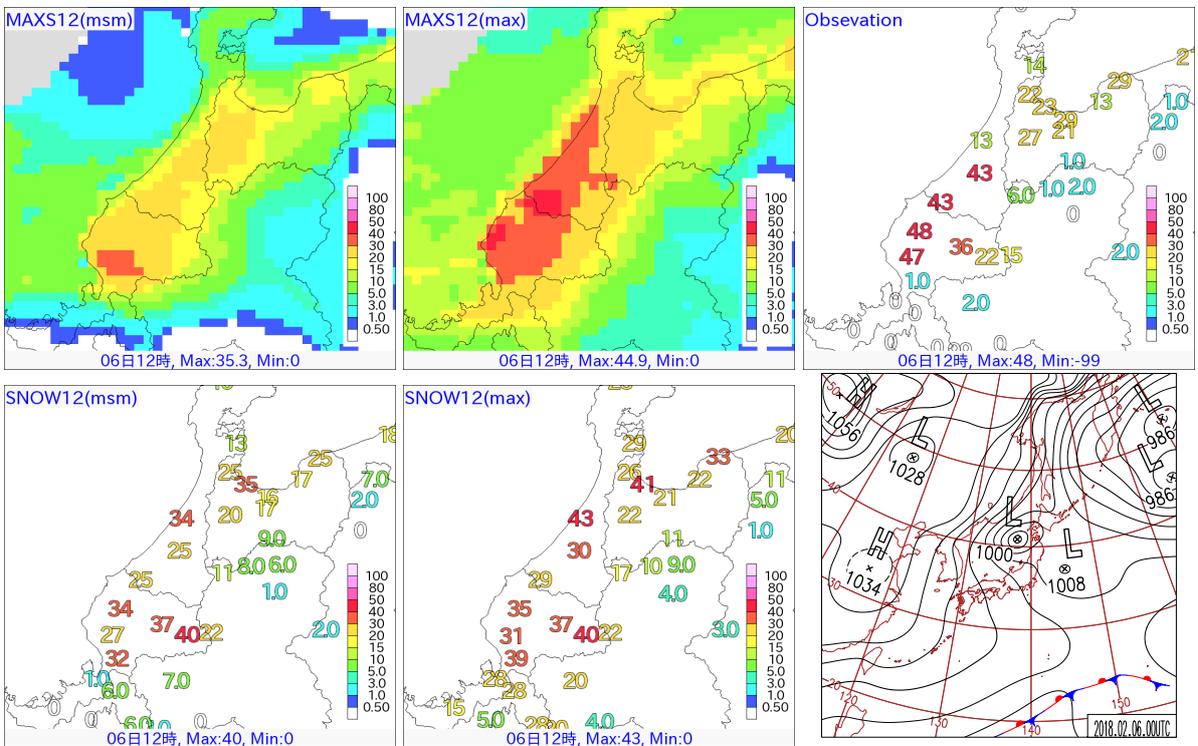


図 2.5.12 2018 年 2 月 6 日 03UTC を対象とした 12 時間降雪量の予測事例。予測の初期時刻は 4 日 18UTC。上段左から MSM 降雪量（格子）ガイダンス、MEPS 降雪量（格子）ガイダンスのアンサンブル最大、アメダスによる前 1 時間積雪深差の 12 時間積算値、下段左から MSM 降雪量地点ガイダンス、MEPS 降雪量地点ガイダンスのアンサンブル最大、6 日 00UTC の日本域地上天気図。降雪量の単位は [cm/12h]。

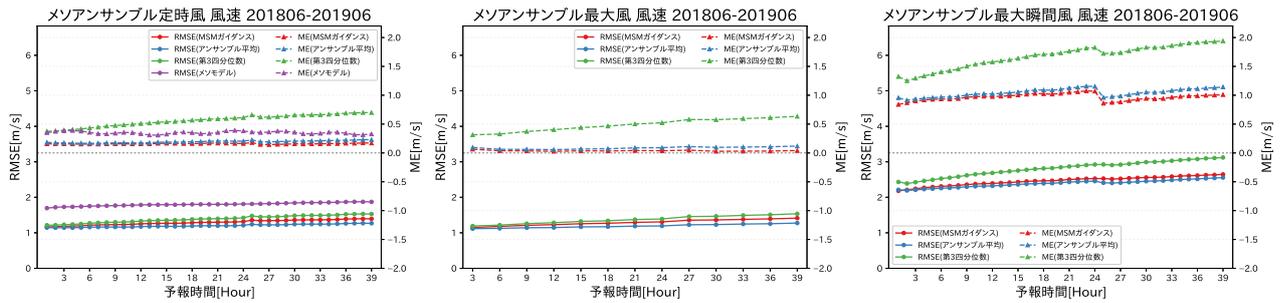


図 2.5.13 予報時間毎の風速の RMSE (実線) と ME (破線)。(左) 定時風、(中) 最大風、(右) 最大瞬間風。赤線が MSM ガイダンス、青線がアンサンブル平均、緑線が第 3 四分位数、紫線がメソモデルのスコアを示す。

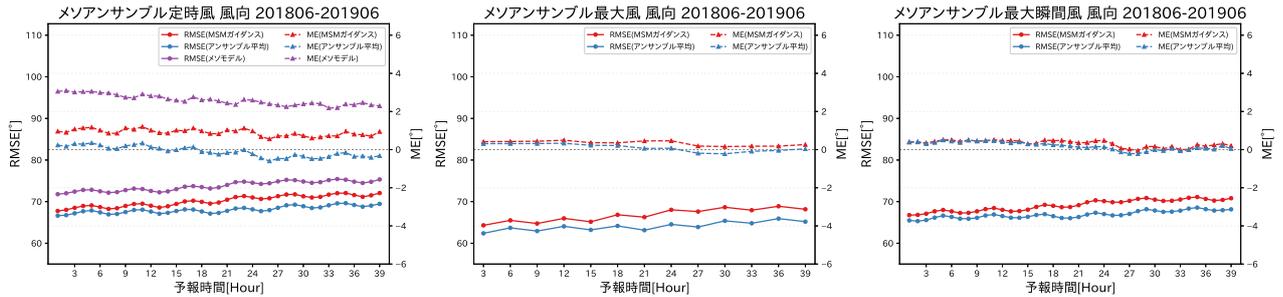


図 2.5.14 図 2.5.13 と同じ。ただし、予報時間毎の風向のスコアを示す。

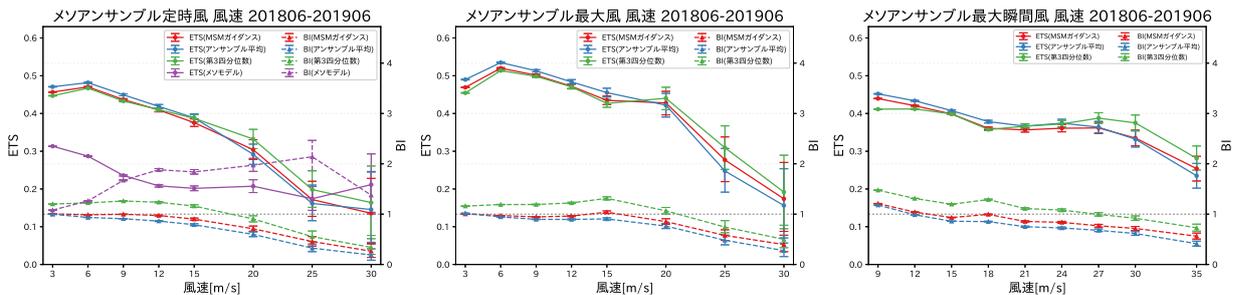


図 2.5.15 図 2.5.13 と同じ。ただし、風速の閾値毎の ETS (実線) と BI (破線) を示し、エラーバーは 95%信頼区間である。

いる。

## (2) 統計検証

MEPS 風ガイダンスの統計検証の結果を示す。検証期間は 2018 年 6 月 5 日から 2019 年 6 月 4 日までの 1 年間で、アメダス地点での検証である。検証には 00, 06, 12, 18UTC 初期値のガイダンスを用いた。空港地点の MEPS 風ガイダンスは作成開始が 2019 年 2 月であり、検証期間が短いことから本項では省略する。

図 2.5.13 に、予報時間毎の風速の二乗平均平方根誤差 (RMSE) と平均誤差 (ME) を示す。左から定時風、最大風、最大瞬間風のガイダンスの検証結果を示している。定時風ガイダンスの結果を見ると、ME は MSM ガイダンス (赤線) が最も 0 に近く、次いでアンサンブル平均<sup>6</sup> (青線) が 0 に近づいている。第 3 四分位数<sup>7</sup> (緑線) は最も正バイアスが大きくなっている。

<sup>6</sup> 風速のアンサンブル平均は、メンバー毎の風速をスカラー平均して求めた。ベクトル平均するよりも精度が良かったため、スカラー平均としている。

<sup>7</sup> 第 3 四分位数は、アンサンブルメンバー毎の風速を小さい順に並べた時の 75% に位置する値として求めた。

RMSE は、アンサンブル平均が最も小さく、MSM ガイダンス、第 3 四分位数、メソモデルの順に大きくなる。また、予報時間初期は RMSE にあまり差はないが、予報時間が先になるほどそれぞれのスコアの差が大きくなっており、アンサンブル平均の優位性が高くなるのが分かる。このことは、最大風、最大瞬間風のガイダンスの結果でも同様である。

図 2.5.14 に、予報時間毎の風向の RMSE と ME を示す。こちらは全てのガイダンスで ME, RMSE とともに、アンサンブル平均<sup>8</sup>が最もスコアが良い。

以上から、基本的には MEPS 風ガイダンスのアンサンブル平均を利用すれば、統計的には最も予測精度が高くなると言える。ただし、風速の閾値毎の ETS と BI を示した図 2.5.15 を見ると、風速 20 m/s 未満ではアンサンブル平均の ETS が最も高いが、20 m/s 以上では第 3 四分位数の方が ETS は高くなっている。これは、風ガイダンスでは風速が 20 m/s 以上では BI が 1

<sup>8</sup> 風向のアンサンブル平均は、メンバー毎の風の単位ベクトルを平均して求めた。ベクトル平均するよりも精度が良かったため、単位ベクトル平均としている。

より小さくなっており、アンサンブル平均を取ること  
で、さらにBIが小さくなるためである。第3四分位数  
は捕捉率が高く(図略)、アンサンブル平均やMSMガイ  
ダンスよりBIが1に近いのでETSが高い。このた  
め、常にアンサンブル平均を利用するのではなく、風  
速20 m/s以上の強風が予測されるような場合では、第  
3四分位数の利用も検討して欲しい。

## 2.5.4 発雷確率ガイダンス

### (1) 仕様

MEPS発雷確率ガイダンス(以下、MEPS-PoT)は、  
MEPSを入力としMSM発雷確率ガイダンス(以下、  
MSM-PoT)と同様の手法で作成する。ただしMEPS-  
PoT独自で係数は作成せず、MSM-PoTの係数を各メ  
ンバーの予測に適用する。また、LAF(Lagged Average  
Forecast)法による重み付け平均はコントロールランの  
み行い、摂動メンバーでは実施しない。MEPS-PoTの  
LAFでは過去2初期値のそれぞれFT=39までを利用  
し、最新初期値と過去初期値を2対1の割合で重み付け  
平均を行う<sup>9</sup>。そのほか予測手法の詳細は土田(2018b)  
を参照頂きたい。

### (2) 統計検証

MEPS-PoTの統計検証の結果を示す。なお、MEPS-  
PoTのコントロールランとMSM-PoTのLAFを適用  
した予測値はわずかに異なるが、統計的な精度は同一  
のため検証図ではMEPS-PoTのみ示す。図2.5.16は  
MEPS-PoTのコントロールラン、アンサンブル平均、  
摂動ランの予測精度比較で、閾値別のETS、予報時間別  
のBSS、信頼度曲線を示している。検証期間は2018年  
6月5日~2019年6月4日である。検証対象は00, 06,  
12, 18UTC初期値で、ETSと信頼度曲線はFT=15~  
36を、予報時間別のBSSはFT=6~39の3時間毎を  
それぞれ対象とした。まず図2.5.16(a)より、MEPS-  
PoTのコントロールラン、アンサンブル平均、摂動ラン  
はいずれも発雷確率20%を閾値とした検証でETSが  
ピーク値をとり、20%を閾値として発雷の有無を判断  
した場合にそれぞれの予測精度が最も良いことがわか  
る。したがって、MEPS-PoTとMSM-PoTの間で判  
断の閾値を変更することなく利用できる。

続いてアンサンブル平均の精度を確認する。図2.5.16  
(b)より、信頼度曲線が理想直線よりも上にあり、実況  
の出現頻度よりも低めの確率となる傾向があることが  
わかる。また、その傾向はコントロールランよりも強  
い。図2.5.16(c)からは、BSSはコントロールランと  
同等か、予報時間の後半ではコントロールランをやや  
上回っていることがわかる。ここで図2.5.16(a)を見  
ると、閾値10~20%まではコントロールランのETS

を上回っているが、25%以上では逆にコントロールラン  
を下回っていることがわかる。信頼度曲線からもわか  
る通りアンサンブル平均することにより高い確率値  
が予測されにくくなるため、高確率を閾値とした場合  
には見逃しが増えてしまう。その一方で空振りを大き  
く減らすことができ、10~20%を閾値とすればコント  
ロールランよりも多くの発雷を捕捉することができる。  
以上のようにアンサンブル平均の発雷予測精度は統計  
的にはコントロールランと同等か、予報後半でコント  
ロールランを上回っている。一般に予報時間が長いほ  
ど予測のばらつきは大きいことから、高確率を閾値と  
した場合の空振りを減らし、低確率を閾値とした場合  
に多くの発雷を捕捉することによる精度向上効果が大き  
いと考えられる。特に予報後半については不確実性  
が取り込まれた情報として参考にできる可能性がある。

最後に摂動ランについて、信頼度曲線は理想直線よ  
りも下にあり、実況の出現頻度よりも高めに確率を予  
測する傾向があることがわかる。また、BSSはコント  
ロールランを下回っていることから、ある固定のメン  
バーを常に利用しつづけた場合にはコントロールラン  
を改善しない。その一方で、閾値50%以上では摂動ラン  
のETSはコントロールランと同等であることがわか  
る。これは閾値50%以上を発雷の判断基準とした場合  
には、摂動ランは空振りが増えるものの、コントロール  
ランよりも多くの発雷を捕捉できていることを意味す  
る。ここで、発雷の有無の判断閾値である20%を境と  
して、GSMガイダンス(以下、GSM-PoT)が20%以  
上、MSM-PoTが20%未満を予測しているような事例  
を考える。土田(2018b)に示されている通り高確率を  
閾値とした場合はGSM-PoTのETSがMSM-PoTを  
上回る。このような事例でGSM-PoTのBSSはFT後  
半についてはMSM-PoTを上回るものの、FT前半で  
はMSM-PoTの方が高いためGSM-PoTを採用する  
ことが難しい場合が多い。このとき、MEPS-PoTの  
アンサンブル最大を判断基準とすることで、FT前半  
も含めてGSM-PoTを採用することが可能である。図  
2.5.17はGSM-PoTが20%以上、MSM-PoTが20%未  
満を予測している事例を対象とした、GSM-PoTおよ  
びMSM-PoTの予報時間別のBSSである。図2.5.17  
(b), (c)に示したとおり、アンサンブル最大の予測値  
20%以上を閾値としてGSM-PoTを採用すればMSM-  
PoTを改善可能であり、予測値50%を閾値とした場合  
にMSM-PoTに対するGSM-PoTの改善率が最も大  
きくなる。

### (3) 事例

MSM-PoTとGSM-PoTで予測が異なった場合に、  
MEPS-PoTメンバーの予測が参考となった事例を紹介  
する。図2.5.18は2018年6月24日15JSTを対象  
とした事例で、北海道では寒冷前線の通過に伴い発雷  
が観測された。この北海道での発雷予測に着目すると、

<sup>9</sup> ガイダンス実行の初期値間隔がMSM-PoTとは異なるた  
め、LAF法の重み付けの仕様はMEPS-PoTのコントロ  
ールランとMSM-PoTで統一していない。このため両者の予  
測結果にはわずかながら差が生じる。

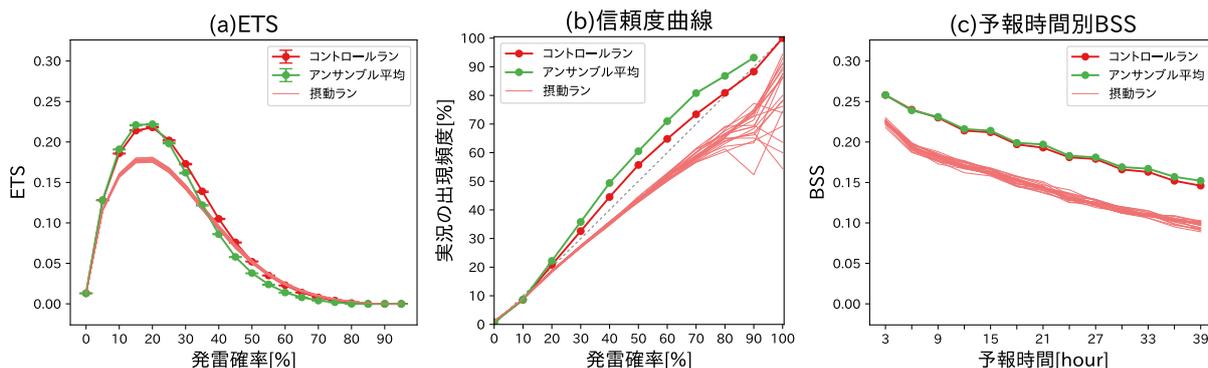


図 2.5.16 検証期間 2018 年 6 月 5 日～2019 年 6 月 4 日を対象とした MEPS-PoT の精度比較。(a) は閾値別の ETS、(b) 信頼度曲線、(c) は予報時間別の BSS で、赤線が各メンバー（太線がコントロールラン、細線が摂動ラン）、緑線がアンサンブル平均を示す。(a) のエラーバーはブートストラップ法による 95%信頼区間を表す。

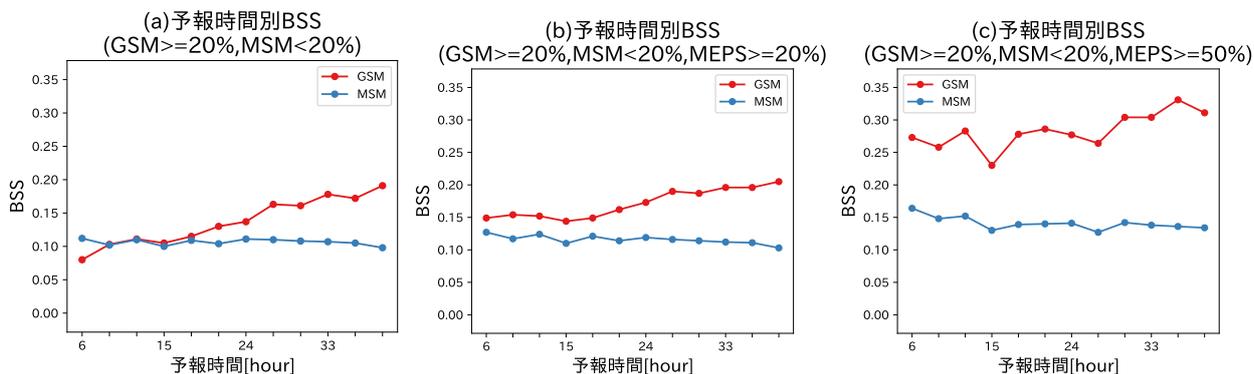


図 2.5.17 GSM-PoT が 20%以上かつ MSM-PoT が 20%未満である事例を対象とした予報時間別の BSS。検証期間は 2018 年 6 月 5 日～2019 年 6 月 4 日で、検証対象は GSM, MSM ともに 00, 06, 12, 18UTC 初期値。(a) は MEPS-PoT を参照せずに各ガイダンスを利用した場合、(b) は MEPS-PoT のアンサンブル最大が 20%以上の条件で各ガイダンスを利用した場合、(c) は MEPS-PoT のアンサンブル最大が 50%以上の条件で各ガイダンスを利用した場合の検証結果を示す。

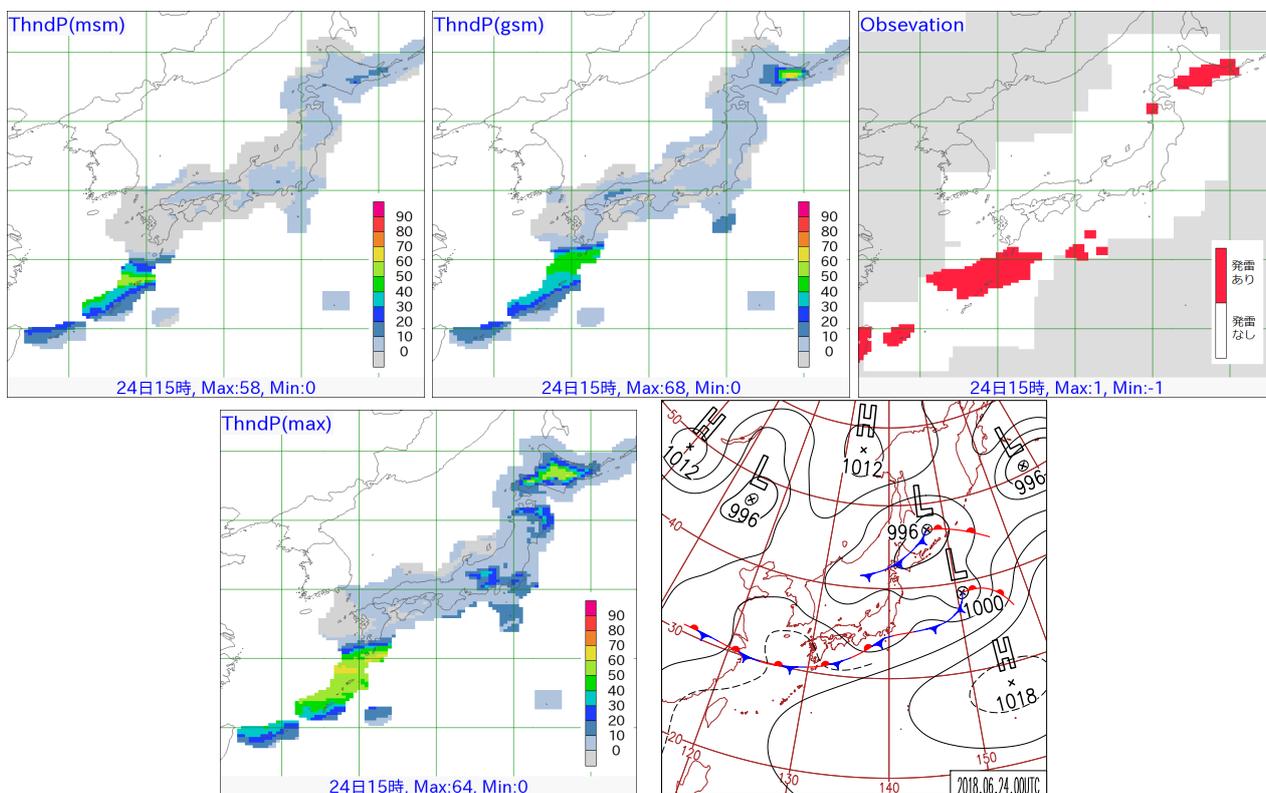


図 2.5.18 2018 年 6 月 24 日 06UTC の前 3 時間を対象とした予測事例。予測の初期時刻は 23 日 00UTC。上段左から MSM-PoT[%]、GSM-PoT[%]、発雷実況、下段左から MEPS-PoT[%] のアンサンブル最大、2018 年 6 月 24 日 00UTC の日本域地上天気図。発雷実況は赤い格子が発雷ありを示す。

GSM-PoT は範囲は狭いが高い確率値を予測している一方で、MSM-PoT は確率値が低く、予測範囲も狭い。このとき MEPS-PoT のアンサンブル最大を確認すると、北海道に高い確率値が予測されていた。このような場合、北海道地方の発雷確率については GSM-PoT の予測を採用することで、MSM-PoT の予測を改善することが可能である。また、前線に伴う発雷が観測された鹿児島県の種子島や屋久島付近についても同じ理由で GSM-PoT の予測を採用することで MSM-PoT の予測を改善できる。

### 2.5.5 まとめ

各ガイドンスについて、利用上の留意点を中心にまとめを記す。

#### 降雪ガイドンス

降水種別ガイドンスでは、代表降水種別を用いてカテゴリ予測を行った場合に、統計的な予測精度で MSM ガイドンスを改善する。また、雨雪判別の確率的予測資料として、北海道、東北、北陸、関東・甲信・東海では「雪か雨」超過確率、近畿・中国・四国では「雪か雨」または「雪」超過確率、九州では「雪」超過確率が利用できる。

降雪量（格子）ガイドンスおよび降雪量地点ガイドンスではアンサンブル最大を大雪可能性の検討資料として利用できる。まず降雪量（格子）ガイドンスで面的な大雪範囲を検討し、より捕捉率の高い降雪量地点ガイドンスによりその範囲を広げる検討を行うといった利用方法を推奨する。なお、6時間降雪量（格子）ガイドンスのアンサンブル最大は7 cm/6h 以上の降雪について MSM ガイドンスの予測精度を上回るため、短時間の大雪に対する定量的な予測資料としても利用できる。利用上の留意点として、アンサンブル最大は地点単位規模では予測頻度がかなり過大なため、都道府県単位での最大値を見るなど、ある程度の広さの中で利用していただきたい。また、降雪量地点ガイドンスについては特に予測頻度が過大となることから、主に予測のばらつきが大きい事例での大雪捕捉に利用していただきたい。

降雪量超過確率は降雪量（格子）ガイドンスよりも降雪量地点ガイドンスの方が予測精度が高いことから、面的な確率分布を降雪量（格子）ガイドンスの超過確率で把握し、超過確率の値には降雪量地点ガイドンスを用いて大雪可能性を検討するといった利用方法を推奨する。利用上の留意点として、予測のばらつきが大きい事例では超過確率は低い値しか示すことができないため、そのような事例ではアンサンブル最大を利用して大雪可能性を検討していただきたい。

#### 風ガイドンス

MEPS 風ガイドンスでは、基本的にアンサンブル平

均の方が MSM 風ガイドンスより予測精度が高いが、風速 20 m/s 以上の強風時には第 3 四分位数の方が予測精度が高くなる。このため、常にアンサンブル平均を利用するのではなく、強風が予測される場合は予測地点の箱ひげ図を見るなど、第 3 四分位数の予測も活用していただきたい。

#### 発雷確率ガイドンス

アンサンブル平均は MSM ガイドンスと比べると確率値と実況出現率との対応が悪くなるが、発雷の予測精度は MSM ガイドンスを改善する。10~20%を閾値に MSM ガイドンスの補足情報として利用いただきたい。摂動ランは 50%以上を閾値とすることで統計的には MSM ガイドンスと同等の予測精度で利用できる。ただし、各メンバーの予測値をその都度検討することは難しいことから、アンサンブル最大を MSM ガイドンスを改善するための補助資料として利用いただきたい。例えば、GSM ガイドンスで 20%以上の予測があるが MSM ガイドンスでは 20%未満の場合に、MEPS ガイドンスのアンサンブル最大が 20%以上であれば、当該地域では MSM ガイドンスではなく GSM ガイドンスを利用した方が予測精度が高い。また、MEPS ガイドンスのアンサンブル最大の閾値を 50%以上とすれば、実況との対応がさらに良くなる。

以上のように、必ずしもアンサンブル平均の精度が良いというわけではなく、比較的強い現象を予測する場合は風ガイドンスのように第 3 四分位数の予測を利用するなど、利用方法を検討する必要がある、今後の開発課題となっている。本節で紹介した MEPS ガイドンスについて、現象の強さを予測するために有効な情報を提供できるよう開発を続けている。

#### 参考文献

- 土田尚侑, 2018a: 降雪ガイドンス. 数値予報課報告・別冊第 64 号, 気象庁予報部, 119-131.
- 土田尚侑, 2018b: 発雷確率ガイドンス. 数値予報課報告・別冊第 64 号, 気象庁予報部, 165-176.
- 松本智文, 2018: 風ガイドンス. 数値予報課報告・別冊第 64 号, 気象庁予報部, 144-153.