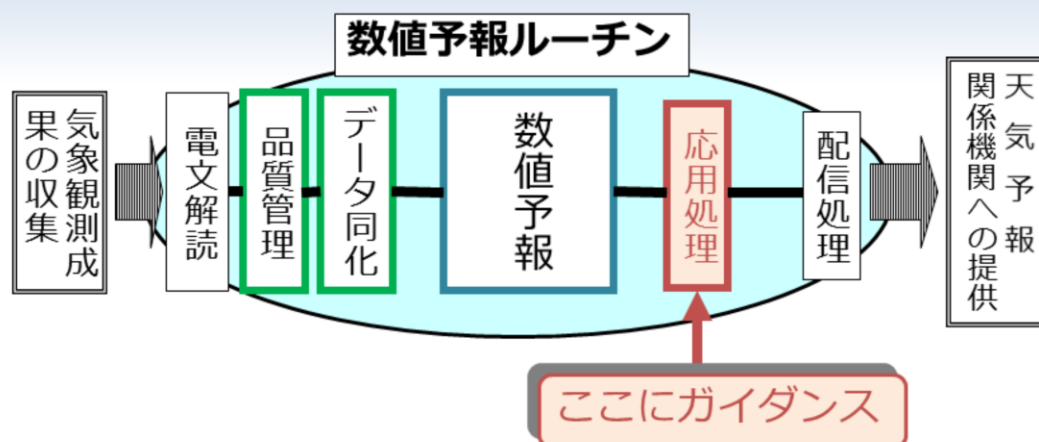




第1章 基礎編

1.6 ガイダンス

数値予報の応用処理



応用処理の例

- 可視化 (Web、FAX図作成)
- データ切り出し (ユーザー向けファイル (GPV) 作成)
- 予報要素への翻訳、統計的な補正 → **ガイダンス**

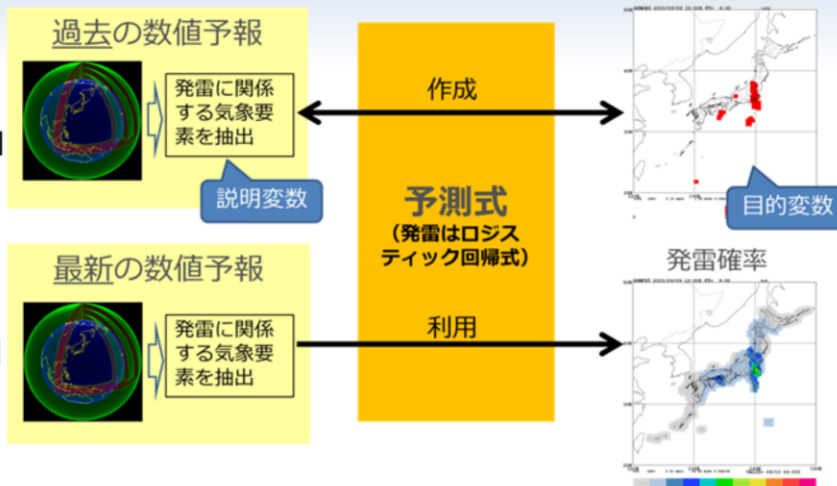
数値予報の後続に位置し、数値予報の結果を利用した各種プロダクトの作成を行う処理を応用処理という。数値予報の結果は膨大な数値データの集まりであり、そのままの状態では利用することは難しい。そこでユーザーの利便性向上のための様々な処理、例えば、数値予報モデルの結果を図に描画する可視化、必要な領域や要素などを抽出するデータ切り出し、予報要素への翻訳や統計的な補正などが応用処理として行われる。これらのうち、最後に挙げた「予報要素への翻訳や統計的な補正」を行う処理およびその結果作成される予測資料をガイダンスと呼ぶ。

限られた時間の中で気温や量などの多くの要素を迅速にモニタリングし、天気予報や注意報・警報・情報を作成するためには、精度よく量的な予測値を直接示す客観的な予測資料が必要であることから、ガイダンスが作成・配信され、利用されている。

ガイダンス

ガイダンスとは

- 「数値予報を用いた客観的・統計的翻訳」
- 過去の数値予報と実況から統計手法等を用いて予測式を作成し、それを最新の数値予報の結果に適用して予測する。



ガイダンスの役割

- 数値予報が予測していない要素を作成する
- 数値予報の系統誤差を補正する

ガイダンスは、防災気象情報や天気予報等の作成を行う上で重要な基本資料

ガイダンスは「数値予報を用いた客観的・統計的翻訳」(気象庁予報部 1986)とも記述されるように、過去の数値予報と実況から統計手法等を用いて予測式を作成し、それを最新の数値予報の結果に適用することで、客観的・定量的な予測を行うプロダクトである。

ガイダンスには2つの役割がある。一つは、発雷確率や乱気流、視程など、数値予報では直接予測しないが、天気予報、警報・注意報、飛行場予報などの発表に必要な気象要素を作成することである。もう一つは、気温や降水量など、数値予報でも予測するが、その予測値を補正し、より精度を向上させることである。

ガイダンスは防災気象情報や天気予報等の作成を行う上で重要な基礎資料となっている。

現在のガイダンスで用いられている手法

逐次学習

- 入出力関係の変化（モデル変更、季節変化等）に対応可能
- 予測特性に不適切な変化が起きる可能性、開発・運用のコストが大きい

一括学習

- 予測特性が安定、予測根拠がわかりやすい、開発・運用コストが小さい
- モデル変更時に係数再作成が必要になることがある

ニューラルネットワーク

- 1, 3時間最大降水量、最小湿度など

説明変数と目的変数の複雑な関係にも対応できるが、ブラックボックス

カルマンフィルタ

- 気温、風、降水量など

入出力関係が変わるものに対し学習が速い（学習速度はパラメタの設定や現象の頻度に依存）

ロジスティック回帰

- 発雷確率など

確率に利用、開発が容易

線形重回帰

- 24時間最大降水量など

開発が容易

診断手法

- 視程（格子形式）など

過去の研究や目的変数の定義に基き予測式を作成

※ ロジスティック回帰と線形重回帰は逐次学習型としても利用可能、ニューラルネットワークは係数固定型としても利用可能。

現在の気象庁のガイダンスの作成には、ニューラルネットワーク、カルマンフィルタ、ロジスティック回帰、線形重回帰および診断手法が用いられている。

ニューラルネットワークは、神経細胞（ニューロン）の機能の一部をモデル化した機械学習アルゴリズムである。入力値（説明変数）と出力値（目的変数）の関係が非線形である場合にも適用できるという特徴がある。近年のAIに利用されているディープニューラルネットワークは、中間層を多層化したニューラルネットワークである。ニューラルネットワークは説明変数と目的変数の複雑な関係を表現することができるが、その反面、なぜそのような予測になったのか解釈が困難（ブラックボックス）であるという問題がある。ニューラルネットワークは1, 3時間最大降水量、最小湿度などに利用されている。

カルマンフィルタは時系列解析の手法の一つであり、ガイダンスにおいては説明変数と目的変数を結びつける係数を逐次学習する手法として利用されている。カルマンフィルタの予測式は線形形式であり、目的変数と説明変数が線形関係の場合に利用できる。カルマンフィルタは気温、風、降水量などに利用されている。

ロジスティック回帰は、雷の有無などのように、現象を2つのクラスに分類する問題に用いられる統計手法の一つである。ロジスティック回帰により得られる予測値は現象の発生確率として考えることができるため、ロジスティック回帰は確率型のガイダンスである発雷確率などに用いられている。

線形重回帰は、説明変数と目的変数の間に線形関係がある場合に用いられる手法で、予測結果の解釈や開発が容易であるという特徴をもつ。線形重回帰は24時間最大降水量などに利用されている。

診断手法は、過去の研究や目的変数の定義に基づいて予測式を決定し、ガイダンスの予測値を算出する手法である。ほかの手法と比べて、開発において観測や数値予報モデルの長期間のデータが不要で、観測密度に起因する予測精度の不均一性がなく、メリハリの利いた予測が可能、という特徴がある。診断手法は視程ガイダンス（格子形式）などに用いられている。

過去のデータを用いて係数を一括で学習し、運用においては係数を固定して利用する手法を一括学習と呼び、新しく観測データを入手することに係数を更新する学習方法を逐次学習と呼ぶ。逐次学習型のガイダンスと一括学習型のガイダンスを利用面・運用面から優劣をつけると以下の通りになる（優劣を不等号で表す）。

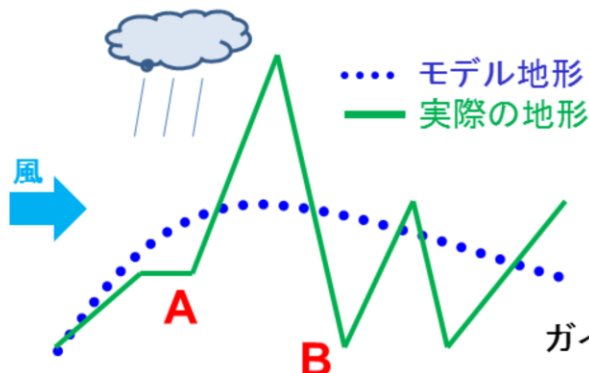
- モデル更新への対応: 逐次 > 一括
- 季節変化への対応: 逐次 > 一括
- 観測所の移設・新設への対応: 逐次 > 一括
- 予測特性の把握: 逐次 < 一括
- 予測の安定性: 逐次 < 一括
- 開発・運用コスト: 逐次 < 一括

カルマンフィルタは基本的には逐次学習のみで利用され、診断手法は一括学習のみで利用されるが、ニューラルネットワークとロジスティック回帰、および線形重回帰では、逐次学習または一括学習のいずれも利用可能である。気象庁のガイダンスでは、ニューラルネットワークには主に逐次学習が、ロジスティック回帰と線形重回帰には一括学習が用いられている。各手法について詳しくは、工藤ほか（2018）を参照。

ガイダンスができること、できないこと

できること: 系統誤差の補正

例: モデル地形と実際の地形の違いによる誤差の補正



実際の地形では

A: 急斜面で地形性の降水が多い

B: 標高が低く、気温が高い

モデルの地形では

A: 起伏に乏しく降水が少ない

B: 標高が高く、気温が低い

ガイダンスではこれらの誤差を補正できる

できないこと: (系統誤差ではない)ランダム誤差の補正

例えばモデルが前線の予測位置を外している場合、前線の位置をガイダンスで修正して気温や風を求めることはできない

ガイダンスは数値予報を統計的に補正するが、補正可能な誤差は系統誤差(場所ごとや対象時刻ごとに統計検証した際の平均的な誤差)である。数値予報の系統誤差としては、

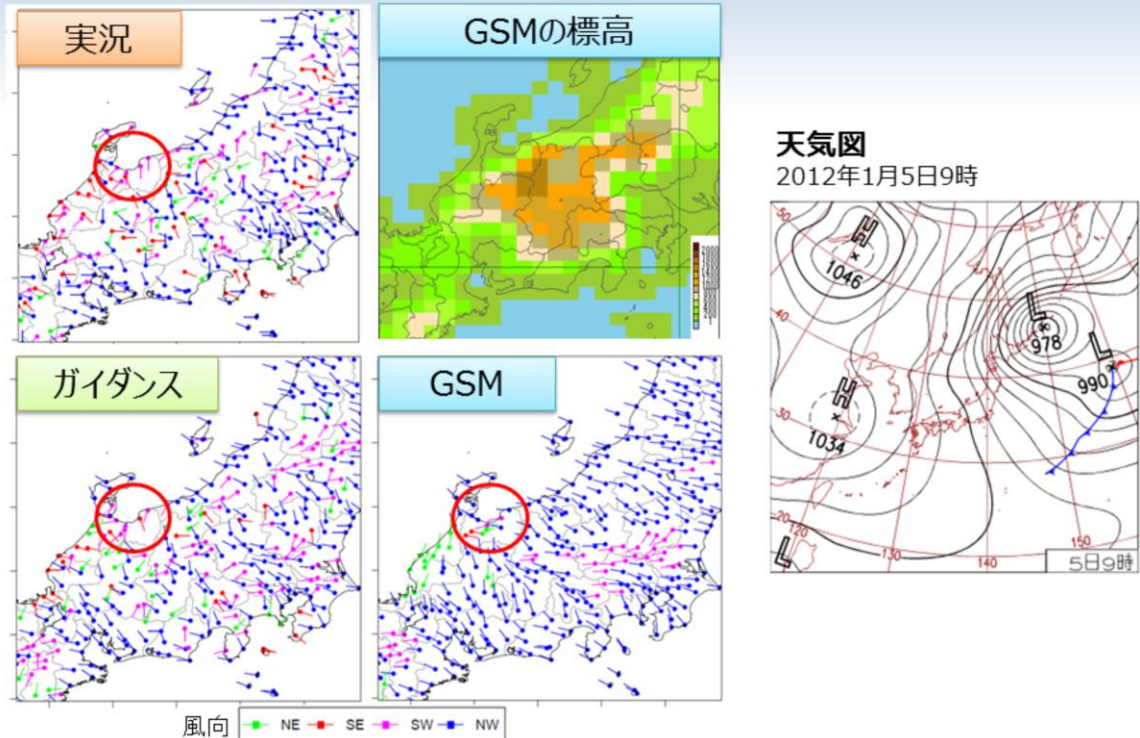
1. 数値予報モデルの地形と実際の地形の違いに起因する誤差
2. 数値予報モデルの海陸分布と実際の海陸分布の違いに起因する誤差
3. 数値予報モデルの不完全性や空間代表性に起因する誤差

などがある。ここでは1の「モデルの地形と実際の地形の違いに起因する誤差」の補正例を示す。

実際の地形とモデルの地形を比較すると、モデルの地形は青点線のように起伏に乏しくなっている。このため、図で左から風が吹いているとすると、A地点では、数値予報の降水は実際の降水よりも弱く予測される傾向がある。またB地点では実際の地形では谷となっているが、数値予報モデルでは谷が表現されていないため、逆転層がない限り、数値予報の地上気温は実況よりも低くなる。

系統誤差ではない誤差(ここではランダム誤差と呼ぶ)はガイダンスで補正することはできない。ランダム誤差の要因は様々であるが、例えば数値予報の前線や低気圧等の擾乱の位置ずれ、擾乱の発生・発達の外れによる誤差がある。モデルが前線の予測位置を外している場合、ガイダンスは前線の位置を修正して気温や風を求めることはできない。

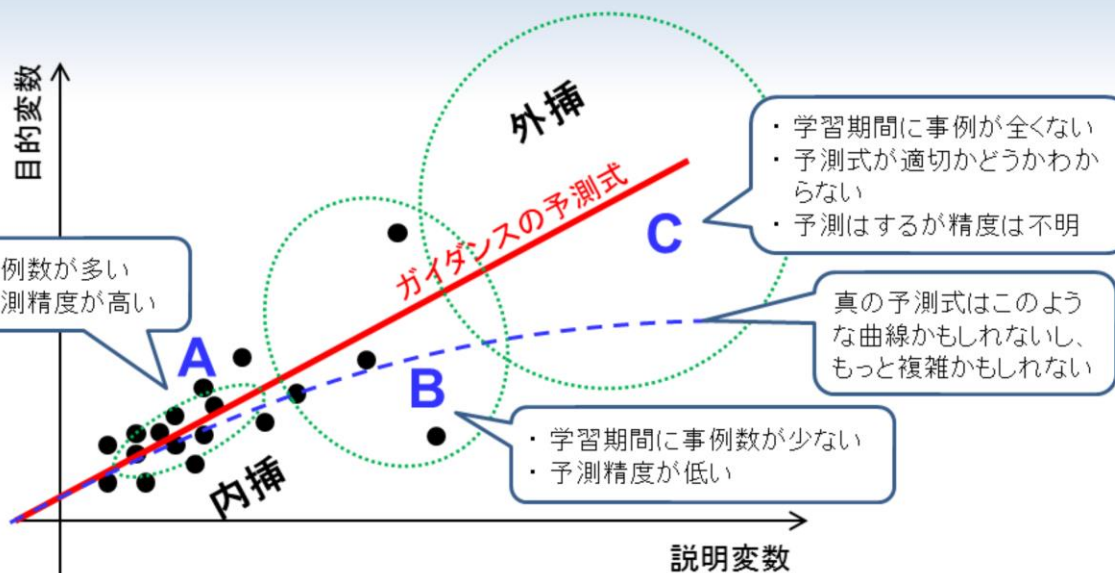
粗いモデル地形に伴う系統誤差の補正例



ここでは地形による系統誤差をガイダンスが適切に補正した例としてGSM定時風ガイダンスの予測事例(2012年1月4日12UTC初期値のFT=24の事例)を示す。

GSMの地形(2012年当時)は中央上図に示したように粗いため、GSMは一様に近い滑らかな風を予測している(中央下図)。しかし実況では実際の地形を反映した複雑な風となっている(左上図)。ガイダンス(左下図)では数値予報を適切に補正し、実況の風に近づけていることが分かる。例えば富山県(図中の赤丸)では、冬型の気圧配置の際に地形の影響で南寄りの風になることが多いが、GSMでは滑らかな地形のために西から北西の風を予測する系統誤差がある。ガイダンスはこれを補正して南寄りの風を予測していることがわかる。

統計手法でできるのは過去事例の内挿



- ・主にAで学習した予測式をBやCにも適用している
- ・学習期間中にほとんど起きたことがない現象(大雨、強風など)は予測精度が低い
- ・学習期間中に一度も起きたことがない現象に対しては適切な予測はできない

統計手法を用いたガイダンスでは、過去の予測(説明変数)と実況(目的変数)から作成した関係式を用いて未来の予測を行うが、実況は偏った頻度分布をしている場合が多い。例えば降水量や風速では、降水量が少ない(または無い)事例や弱風の事例が非常に多く、大雨や強風の事例は少ない。このように実況の頻度分布に偏りがある場合、ガイダンスは頻度が多い事例で学習した予測式を頻度が少ない(または全くない)領域まで適用することで予測値を求めていることに注意が必要である。

図の例では、Aの領域では事例数が多いが、Bの領域では事例数が少なく、Cの領域では事例が全くない状況になっている。Aの領域に対する予測は過去事例の内挿であり、統計手法による予測精度は一般的に高い。これに対してBの領域では学習期間に事例数が少ないため予測精度は低くなる。またCの領域では学習期間中に事例が一つも存在せず、ガイダンスの予測(赤実線)はAやBのデータからの外挿になっており、過去データから予測精度を検証することもできない。真の予測式は図の青点線で示したような曲線かもしれないし、もっと複雑な曲線かもしれないが、過去データからはそれを判別することはできない。

この例のように、学習期間中にほとんど起きたことがない現象に対しては、統計手法を用いたガイダンスは一般に予測精度が低く、また、一度も起きたことがない現象に対しては適切な予測を行うことはできない。

参考文献

- 気象庁予報部, 1986: 短期間予報用ガイダンスの解説, 予報技術資料第35号.
- 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.