

カルマンフィルターを用いた天気予報ガイダンス*

国次雅司**

1. はじめに

天気予報を作成する際に用いるデータには、大気の現在の状態を表す実況（観測値）と、将来の状態を表す予報値がある。予報値の基礎となるデータは、主に数値予報という方法によって作られている。数値予報とは、気温や風などの大気の状態を表す物理量の将来の値を、それらを支配する法則に基づいて現在の値から予想する方法である。大気の運動と状態を支配する方程式を、電子計算機を用いて数値的に解くところからその名前がつけられた。この目的でプログラム化された大気のモデルを数値予報モデルと呼ぶ。

数値予報モデルによって、大気の状態を表す物理量の予報値は、格子点と呼ばれる規則正しく配置された点における値（Grid Point Value : GPV）として求められる。格子点は、気象庁で現在最も解像度の高い数値予報モデル（領域モデル、以下 RSM と呼ぶ）で約 20 km おきに配置されている。

数値予報モデルの出力である GPV を、そのまま予報に用いることができればよいのであるが、実際にはそうはいかない。その理由はいくつかあるが、まず第 1 に、GPV は格子点を囲むまわりの大気の実況値を表しているために、個々の地点の状態にぴったり当てはまるとは限らないことが挙げられる。これに関連して、格子点の標高も実際の地形の高度を平均・平滑化したものを用いているために、個々の地点の標高とは一般に異なっていて、その違いの影響もある。この影響は、具体例で後に示す。

第 2 の理由は数値予報モデルの予報値には必ず誤差

が含まれることである。この誤差の中には、予報のたびにいつも一貫して現れる系統的誤差と呼ばれる数値予報モデルの「くせ」が含まれている。この系統的誤差が分かれば原理的には数値予報の誤差がある程度修正できるはずである。しかし、限られた時間で多数の数値的な修正を人間が行うことは困難である。

また、数値予報モデルが直接予報していない量や、降水確率などの確率的な予報値などは、GPV には含まれていないので、別途計算する必要がある。これが第 3 の理由である。

実際の天気予報作業時には、これらの理由から GPV の他に新たな予報値の作成が必要である。そこで GPV を統計的に処理して、客観的な予報値を得る方法として天気予報ガイダンス（以下ガイダンス）が開発されてきた。なお、このように処理して作成したデータ（予報値）自身もガイダンス（値）と呼ぶ。

平成 8 年 3 月より、気象庁では天気、気温、降水などについて地方天気分布予報や地域時系列予報といったきめの細かい情報を出している。テレビや新聞などのメディアでも、これらの予報形態をよく目にするようになった。ガイダンスはその基礎となるデータであり、重要性をさらに増してきている。

2. ガイダンスの作成

ガイダンス値は GPV やそれらを加工したデータに重みをかけた値の和として求める。この式を予測式と呼ぶ。例えば明日の東京の最高気温を予想する場合を考えてみよう。各種の GPV のうち、東京付近の地上気温の GPV の翌日の日最高値が最も高い相関を持つものと予想されるので、予測式にはまずこれを第一に使うことにする。この他にも関係の深そうな要素がいくつか考えられるので、それらによって補正を行うこと

* Statistical guidance systems using Kalman Filter technique.

** Masashi Kunitsugu, 気象庁数値予報課.

© 1997 日本気象学会

にして、1例として

翌日の東京の最高気温予想値 (ガイダンス値) =
 翌日の東京付近の気温の GPV の日最高値 * 重み 1
 + 翌日の東京付近の風の強さの GPV * 重み 2
 + 翌日の東京付近の雲量の GPV * 重み 3
 + 定数 (重み 4)

という予測式をたててみる。このとき選ばれた要素を説明因子、その重みを係数という。

観測値とガイダンス値の差をガイダンスの予想誤差と呼ぶことにする。ガイダンスの予想誤差が小さい(精度が良い)かどうかは、予測式の説明因子の選び方と係数の値によって決まる。これらの決め方には様々な方法がある。

その方法のうち、観測値と対応する時刻の GPV (を加工したもの) のデータがある期間蓄積し、その統計値に基づいてガイダンスの予想誤差 (の 2 乗の和) が最も小さくなる (1 組の) 説明因子と係数を決定する方法を、MOS (Model Output Statistics) 方式という。MOS 方式によるガイダンス値は、数値予報モデルの系統的誤差を取り除くことができるという利点がある。気象庁はこれまではガイダンスにこの MOS 方式を一貫して用いてきた。

MOS 方式では、蓄積データの取得期間が短いと、ある特定の天候しか代表しない係数ができてしまう恐れがある。代表性の良い係数を得るためには、少なくとも 2、3 年のデータを蓄積する必要がある。このため数値予報モデルが変更され、それによって数値予報モデルの予報値の特性が変化してしまうと、せっかく作ったそれまでの予測式が使えなくなるという大きな問題点が存在する。

平成 8 年 3 月の計算機システムの更新とともに気象庁の数値予報モデルが新しくなり、MOS 方式を採用してきたガイダンスはこの問題に直面した。そこで MOS 方式に代わるものとして採用された手法のひとつが、次に述べるカルマンフィルター (KLM) 方式である。

3. カルマンフィルターとは

フィルターとは、ごみや不純物を除去し、必要なものだけを通過させる濾過装置のことである。たとえばある観測値が与えられたとき、それは観測機器から発生するノイズによって乱されていると考えられる。このノイズを除去し、最適な信号を取り出す装置がフィ

ルターと言える。

1960年に R.E.Kalman が発表したフィルタリング理論はカルマンフィルターと呼ばれ、制御工学・通信工学の他に土木工学、社会・経済学、統計学などさまざまな分野に应用されている。カルマンフィルターは、60年代後半におけるアメリカのアポロ計画等の宇宙開発計画において应用されて普及することになった。このときカルマンフィルターは、宇宙船の軌道追跡に利用された。宇宙船の軌道は、基本的には初期の条件を与えて軌道方程式を解いて求められる。しかし実際の軌道はこの計算による基本軌道からはずれる。また宇宙船の位置を観測する際もノイズによって正確な位置を決定できない。そこで刻々と入ってくる実況データをもとにノイズを取り除き、その時刻において最も適していると考えられる宇宙船の位置を割り出すためにカルマンフィルターが用いられた。ここではカルマンフィルターは、時間によって変化する状態 (宇宙船の軌道推定でいえば位置など) のある時刻における最適な値を、観測によって得た情報から推定する (フィルタリングする) システムと言える。

カルマンフィルターを最初にガイダンスに応用したのは A. O. Persson や C. Simonsen で、1991年のことである。気象庁においても基本的には彼らの方法を採用している。

4. カルマンフィルターによるガイダンス

4.1 時間変化する係数

さて、ガイダンスの作成に用いているカルマンフィルターについて説明しよう。ガイダンスでは、2 に述べたような予測式を用いている。その最適な係数を求めるために、MOS 方式に替えてカルマンフィルターを用いる。これを KLM 方式によるガイダンスと呼ぶ。

KLM 方式では、次の予想時間ステップの係数を予想し、それを最適化するという 2 段階の処理を行っている。まず、係数の予想について述べてみよう。係数の予想とは、例えば各項の係数を次の時間ステップに f 倍にするというように、ある法則に従って変化させることを意味する。しかしガイダンスの係数の変化の法則は分からないので、前回と同じ係数を用いる (各項とも $f = 1$)。このことは、係数が時間の経過に関して基本的には変わらないという法則に従うことを仮定することになるが、MOS 方式が一定の成績を示していることからおおむね正しい仮定と考えられる。次の時間ステップの係数の予想値を、第 1 推定値と呼ぼう。

さてこのように予想した係数の第1推定値が、季節を問わずにいつも最適であるとは限らない。そこで、次の時間ステップの係数が、第1推定値からいくらかの大きさの変動範囲内でランダムに変化することを許す。しかし、観測している値は係数ではないので、実際の変化量を知ることはできない。この変化量をシステム雑音と呼ぶ。

4.2 係数の最適化

第1推定値からの変化量の最適な値を、次の時間ステップに得られたガイダンスの観測値に応じて推定し、その結果最適な係数を得ることができる。これを係数の最適化と呼ぶ。

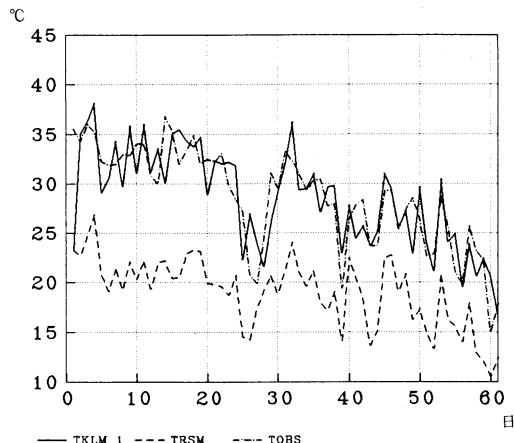
係数の第1推定値を用いてGPVを予測式に代入するとガイダンス値が得られる。そして観測値が得られた時点で、ガイダンスの予想誤差が計算できる。仮に、この誤差全てが時間の経過に伴って予測式と観測値の関係が変わったことによって生じたとするならば、予測式の各項の係数を、その第1推定値から、ガイダンス値に対する寄与に応じて、予想誤差に相当するだけ変化させればよい。

しかし、観測値には予測式に選んだ説明因子では説明できない要素も影響を及ぼす。これには選ばれなかった説明因子や、GPVと相関を持たない小さなスケールの現象の影響などが含まれる。これらはランダムな雑音としてガイダンスの予想誤差に現れる。また説明因子に選ばれたGPVにも、例えば数値予報での高・低気圧や前線の位相のランダムなずれなど系統的ではない誤差が含まれている。この誤差はガイダンス値にランダムな雑音として現れる。ガイダンスの予想誤差に含まれるこれらの雑音を観測雑音と呼ぶ。

そこでガイダンスの予想誤差のうち観測雑音を差し引いた情報を、係数の第1推定値からの変化量に与える。このため、ガイダンスの予想誤差にある重みをかけた量を変化量として第1推定値を修正した値を、係数の最適な推定値として求めることにする。

ガイダンスの予想誤差に含まれる観測雑音を区別することはできないので、係数の真の値を知ることは不可能である。そこで係数の真の値を仮定し、修正した係数からの誤差の大きさ(平均2乗誤差)が最小になるような重みを決定する。重みの値は、係数やシステム雑音、観測雑音の平均値、誤差の大きさなどの統計量を決めれば求めることができる。

重みの値の絶対値は、ガイダンス値との相関が高い項ほど大きく、各項の寄与に応じた情報の分配がな



第1図 甲府の朝9時を初期時刻とするRSMとガイダンスによる翌日の予想最高気温(平成8年8月、9月)。例1: 観測雑音を無とした場合。実線: ガイダンス, 破線: RSM, 一点鎖線: 観測値。

れる。またシステム雑音が小さい項では小さくなり、係数の変化量が制限される。一方観測雑音が大きいほど小さくなり、ガイダンスの予想誤差の情報の取り込みが少なくなる。

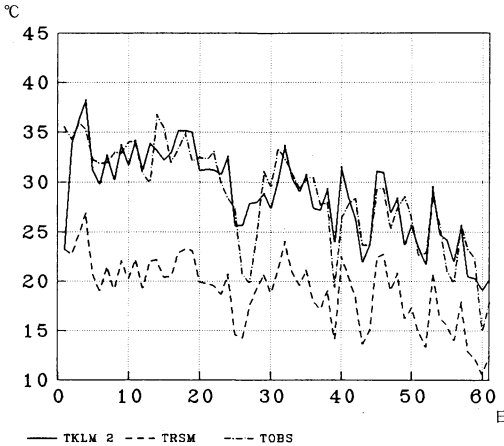
このように、係数を予想し、観測値が得られた時点で雑音を除いた最適な係数を推定(フィルタリング)するアルゴリズムをカルマンフィルターという。

ここで簡単な例についてカルマンフィルターの適用例を見てみよう。取り上げるのは、甲府の日最高気温の予想である。甲府は実際は盆地内にあるが、数値予報モデルは小さなスケールの地形までは解像できないため、数値予報モデル(RSM)の地形上ではかなり高い山の上にある。予測式は、

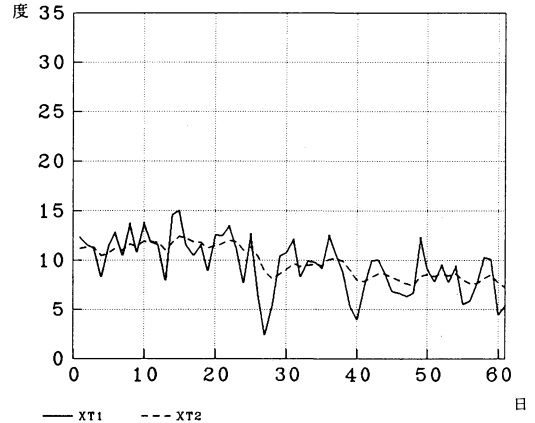
翌日の甲府の最高気温予想値 =

翌日の甲府付近の気温のGPVの日最高値
+ バイアス項

とする。ここでは最適なバイアス項をカルマンフィルターで求める。当日の朝9時を初期時刻とするRSMの予報値(GPV)から、甲府を囲む4格子点の地上気温を甲府の地点に水平内挿する。その値の翌日の日中の最高値が右辺第1項である。第1図、第2図に、平成8年の8月、9月のガイダンス値(左辺)、RSMの予想値(右辺第1項)、観測値の変化を示す。第1図は観測雑音が無いと仮定した例(例1)、第2図は観測雑音を考慮した例(例2)である。バイアス項(右辺第



第2図 第1図と同じ。例2：観測雑音を考慮した場合。実線：ガイダンス、破線：RSM、一点鎖線：観測値。



第3図 例1, 例2におけるバイアス項の変化。実線：例1, 破線：例2。

2項)の初期値はそれぞれ0度と置いた。主に地点の標高の違いによって、RSMの予想値は観測値に比べてかなり低くなっているが、ガイダンス値の方は例1, 例2ともにバイアス項がうまく補正している。バイアス項による補正は第1日から第2日に大きく変化し、その後はほぼ落ちついていく。

例1, 例2を比べてみると、観測雑音を考慮した例2の方が、ガイダンス値の日々の変動がRSMの変動に近い。第3図は、例1, 例2それぞれのバイアス項の変化を示したものである。例1ではガイダンスの予想誤差を全てバイアス項の変化量に取り入れているため、日々大きな変化をしているのに対し、例2では観測雑音を考慮した緩やかな変化をしているのが分かる。

4.3 MOS方式との比較

KLM方式の係数は、時間ステップ毎に観測値から得られる情報を取り入れている。よってMOS方式のように、暖候期、寒候期用の係数を用意して、決まった日時に係数を切り替えたりする必要はない。また、MOS方式における予測式は、予測式の決定に用いたデータ取得期間の最適な状態を表しているため、それとは異なった、例えばまれな猛暑が続くなどの異常な天候には追従できないが、KLM方式では比較的短い期間で最適な係数に近づくことができる。このため、KLM方式においては、数値予報モデルの更新に関する問題を容易に解決することができる。

KLM方式は、予測式の説明因子の選択によって観測雑音の大きさが変わり、これによってガイダンスの

精度が左右されるのに対し、MOS方式は有効であると思われる説明因子の候補を多数用意し、その中からよりよい組み合わせを自動的に選ぶ方法をとっているという点が特徴である。KLM方式の場合には、説明因子の選択は人間に任されており、これがガイダンスの精度に大きな影響を与える。また予測式の作成時に分かりやすい気象要素を説明因子に選べば、その係数の変化を見て説明因子と観測値の関係について解釈することが容易になる。

なおKLM方式では、係数や雑音の大きさの初期値などの決定作業が必要である。

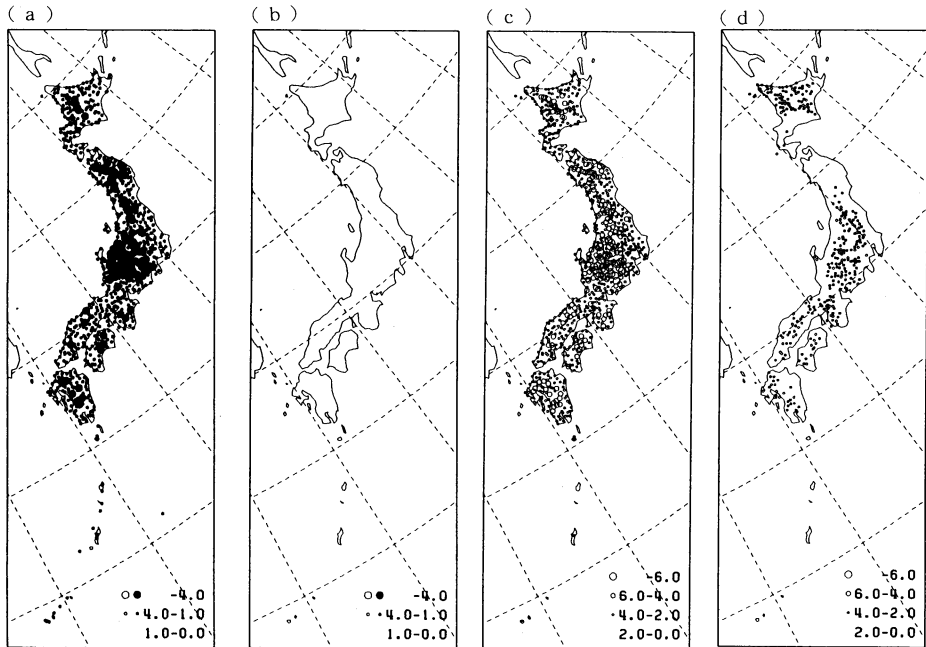
5. 気象庁におけるKLM方式によるガイダンス

気象庁のガイダンスへのKLM方式の導入は、降水量、降水確率、気温ガイダンスを中心に検討され、従来のMOS方式との精度の比較が行われた。その結果、精度はほとんど変わらないことが明らかになった。そして平成6年夏の猛暑における気温ガイダンスでは、KLM方式の方が良い成績を残した。

平成8年3月以降のKLM方式によるガイダンスには次のものがある。

(1) 降水

約20 kmメッシュの各格子点のそれぞれについて、3時間積算降水量(MRR3)と6時間降水確率(PoP6)を予測する。MRR3は、3時間降水量の各格子点での平均値、PoP6は6時間毎の降水確率、つまり各格子点での6時間降水量が1ミリ以上になる確率である。観測値としてはレーダーとアメダスを合成した解析雨量を用いている。予測式には、RSMの降水量や



第4図 朝9時を初期時刻とするRSMとガイダンスによる、アメダス地点の翌日の予想最高気温の平均誤差とRMSE(10月)平均誤差の白丸は正の値、黒丸は負の値。(a)RSMの平均誤差(b)ガイダンスの平均誤差(c)RSMのRMSE(d)ガイダンスのRMSE。

風、水蒸気量、上昇流などから求めた指数を説明因子に使っている。

(2) 気温

全国842のアメダス地点と72の航空気象官署毎に、日中の最高気温、朝の最低気温を求めている。説明因子にはRSMによる各地点の地上気温や中下層雲量、風の東西・南北成分の予想値を用いている。

(3) 風

72の航空気象官署の風速について予想を行っている。説明因子にはRSMによる各地点の地上風速の予想値を用いている。

上記のうち最高気温ガイダンスの成績を示す。ガイダンス値とRSMの予想値(気温GPVを各地点に水平内挿した値の日最高値)を比べるため、第4図に平成8年10月のRSMとガイダンスによる予想結果を示す。(a)は朝9時を初期時刻とするRSMによる、翌日の予想最高気温の誤差(予想値-観測値)を月平均したものである。これをみると全国的に負となっている。(b)にガイダンスによる同様の1か月平均誤差を示すが、ほとんどプラスマイナス0度に近い。次に各

地点で1日毎に誤差を2乗した値を1か月平均し、平方根をとったもの(誤差の2乗の平均値の平方根:RMSE)を見てみる。(c)のRSMでは、負の誤差の大きかった地点で大きくなっているが、(d)のガイダンスの値はこれよりずっと小さい。これらのことからガイダンスはRSMの系統的誤差をよく改善していることがわかる。

6. まとめ

気象庁のガイダンスの作成方式は、数値予報モデルの更新や季節の変化に柔軟に対応できる点が評価され、降水量、降水確率、気温、風速ガイダンスにおいて、MOS方式に代わって現在はKLM方式が用いられている。

一方で、KLM方式においてはガイダンスの予測式や、雑音の大きさなどの決定という作業が必要である。これらの決定には数値予報モデルの精度が影響する。数値予報モデルの精度の向上および特性を確かめながら、よりよいKLM方式のシステムを構築していく必要がある。