# 第2回統計気候学国際研究集会報告

岡　本　雅　典*

## はじめに

統計気候学に関する最初の国際研究集会は，1979年，東京で開催された(鈴木栄一，1980，「天気」27，461-467)．第2回は1983年9月26日から9月30日，ポルトガルのリスボンにおいて開催された．世界各国から，約100名の参加者があり，日本からの参加者は，組織委員のひとりである鈴木栄一氏(青山学院大)と広尾　純氏（関東学院大)，筆者の3名であった．主要な支援機関は，WMO，U.S. NSF，U.S. ONR であり，その他リスボン科学アカデミー，アメリカ気象学会，アメリカ統計協会，ISI（数理統計，確率のためのベルヌーイ協会)，リスボン大学の統計，OR・計算学部，日本気象学会)，ポルトガル国立科学研究所が協賛した．今回の成功は，組織委員長となった米国オレゴン州立大学大気科学学部マーフィー教授を始め，各組織委員の努力の賜である．

9月末とは言え，日本の真夏の暑さの中で，リスボン郊外のホテルを会場として，17におよぶセッションが1つの会場で次々と行われた．提出され，一応受理された論文は90編，そのうち86編が発表された．あまりに数が多いので，プロシーディングには印刷されるが，時間に余裕が生じた場合にのみ発表が許された保留論文もこの他に多数あった．今回は気候ないし気象に直接関連する論文がほとんどで気象統計・気候統計に関心のある者には非常に有益な研究集会であった．ここでは9項目に分類し主な論文を紹介する．

## 1. 気候・気象の最適観測網の問題

フランスの Megreditchian による「Statistical approach to climatological network optimization」では次の諸点を問題として取り上げた．予算制約等のために情報損失を伴わないで，どの観測地点を残すべきか最大情報を得るための最も効果的地点はどこか，連続変量の特性を失わず離散化するための最適測定単位はいかにとるべきかなどである．観測網での相関による冗長度（相関が高い地点は省略できるとするもの）を見るために，二つの観測網の同等性の尺度として

$$n_{eq}=\frac{[trV]^2}{trV^2}$$

なる量を定義した．ここには$V$相異なる時点，相異なる地点で観測された平均からの偏差の共分散行列で，その要素は $V_{ij}(t,\tau)=E\{\xi_i(t)\xi_j(\tau)\}$ であり，$tr$ はそのトレース(対角和)の意味である．例としてフランス全域（ヨーロッパのみ）で日最高気温と日最低気温について各地点（$i=j$）で，時間のずれ $\tau=t+\Delta t, \Delta t=1,2,\cdots,30$日として計算した結果を示している．この場合は自己相関行列とその行列積のトレースが計算される．手続きは始めのN個の観測点から各1個を除いて $n_{eq}$ が最大となる $i$ 番目の観測点を除く．次に残された N-1 個について同様な手続を繰り返す．その結果全体で48個の中から12番目までの観測点が第1ランクとして示される．地理的にはフランスの周辺海岸，国境周辺の高地が大多数を占め，この方法の有用性が実証されている．空間的にも同様に空間相関々数を距離の関数として仮定するなどの方法により独立な観測点を見出すことも可能であるとしている．　その他この種の問題に対するいくつかのアプローチを与えている．

Gringorten は「Modeling climatology of areal coverage」において，日最大降水量の生起確率が単一観測点から観測区域を拡大するにつれ減少する傾向，あるいはレーダー・エコー内で区域の面積を拡げるにつれてその区域内のセル数が増大する傾向などを，新しく提案した統計的モデルで説明しようとしている．鋸波形を仮定し地点（$u$, $v$）におけるこの波形 $x_i$（$u$, $v$）は（0,1）内で一様分布するとしている（BSW モデル)．さらにここで考えている地点（$u$, $v$）ではこの一様分布する確率変数が重なっている場合を考えて $y(u,v)=\sum_i x_i(u,v)-6$ とすれば正規分布する波形が得られる．各地点の値を $u$, $v$ 平面上に示せば高低を示す分布図が得られる．次に一辺 $s$ の正方形内で，ある気象要素の最大値，次の2番目の最大値が含まれる最小正方形，さらに3番目の最大値が含まれる最小正方形などを順次作り，25,000回のシ

* Masanori Okamoto，広島大学経済学部．

ミュレーションにより，$s$を波長の単位で表わして $s=2^Z/340$ の経験的公式を得る．ここに $Z=-1\sim10$，一辺 $s'$km の正方形 $A\,\mathrm{km}^2$ に対しては $s=s'/\Lambda$ とおく，ここに $\Lambda$ はパラメータ波長 (km) である．$Z=ln(340\,\mathrm{s})/ln2$ となるが，実際上の便利さのため，スケール距離 r を次のように定義して $Z$ を表す．

$$r=s'/340\,\mathrm{s\,km}=\Lambda/340\,\mathrm{km}=\sqrt{A}/2^Z\,\mathrm{km}$$

したがって，$Z=ln(\sqrt{A}/r)/ln\,2$.

これらの値を用いて気象要素の最大値 $y$ の生起確率 $P_r(y,f,s^2)$ を面積比率 $f$, $s^2$ の関数として表すことができる．日最大降水量の例では $r=9\sim10$ km，レーダー・エコーの例では $r=1\sim2$ km という．Gringorten 自身も述べているようにこのモデルは全く統計的に求めたもので，物理的構造は全く入っていない．でき得るならば物理的考察からこの種の公式が導かれることが望ましいが，気候学への応用としては気象要素の最大値（最小値）の生起確率を面積とその最大値が含まれる部分領域の割合の関数として表し得たことの意味は大きい．

S.J. Bean と P.N. Somerville は「Modeling visibility for data void regions」において西ドイツ各地の視程予報の問題を扱っている．2パラメーターのワイブル分布 $1-e^{-\alpha x\beta}$ を用いる．$\alpha$，$\beta$をその観測地点の高度とその周囲半径 20 km の平均高度に回帰させ，回帰係数を決めている（非線形回帰）．これらの値は月別と日を3時間毎に分けた8つの時間帯に分けて得られている．この回帰方程式を用い60個の全観測点から54個取り出し，観測値のない地点の予測を行った．この2変量が何故選ばれたかについては説明はないが RMS を調べることによりこの選択はよいとしている．

広尾と伊藤は「Analysis of observed wind directions by statistics of directional data」において，方向性をもつデータの分布である，von Mises 分布，Fisher 分布に基づき大気境界層内での風向とその変動の推定を行った．分布が単峰のときはウインド・ローズを用いるよりはこの分布に基づく平均風向を用いた方がよいことを示した．著者も指摘しているように地上風では単峰型でないことが多く，統計学的には多峰分布の場合の統計量の検定法の開発が望まれる．

K.Y. Amegee と R.H. Cuenca は「Application of regionalized variable analysis to climatic data」においてアメリカ，コロラド州内で熱発散量について観測のない地点での値を推定する問題を扱った．地域的同一性の評価のため，ある観測点からの各距離 $h=k\Delta x$ について準分散 $\gamma(h)=\frac{1}{2N(k)}\sum_{i=1}^{N(k)}(Z_i-Z_{i+k})^2$ を $h$ について示したグラフ (variogram) を用いている．$N(k)$ はラグを $k$ としたときのデータ対の数である．$\zeta(0)=\mathrm{Var}(Z)$，$\zeta(h)=E[\{Z(x)=m\}\{Z(x+h)-m\}]$，$\gamma(h)=\zeta(0)-\zeta(h)$ としたとき，この $\gamma(h)$ がほぼ一定値 ($\zeta(0)$) に接近した距離以上では，一応の空間的独立性を認め，通常の統計的手法が利用できる．その距離以内では推定すべき点の値をまわりの値の線形結合で推定する．その際の重み（係数）を決定しなければならないが，最小2乗残差は推定点から各観測点までの距離の関数となり，通常の最小2乗推定法を用いるわけには行かない．このため各重み係数で残差を微分してこれを0とおいて解く．各重みは地点 $i$，$j$ 間の距離の関係 $h_{ij}$ である $\gamma(h_{ij})$ の連立一次方程式の解として表わされ，観測値から $\gamma(h_{ij})$ を代入して重みは定まる．したがって推定点の予測値が得られる．この方法は統計地理学で開発されて来ており，Kriging 法と呼ばれている．広域のデータを含め気候学上の問題を考えるときもっと利用してよい方法であろう．

## 2. 降水データの確率・統計的モデル

H. Alexandersson は「A general stochastic model of the precipitation process with application on integrated precipitation, maximum daily precipitation etc.」において複合ポアソン指数分布を月降水量分布に適合させることを試みている．パラメーターの推定は主として最尤推定法によっている．またもとのパラメーター $m$, $\lambda$ を $m'=mn'/n$，$\lambda'=\lambda$ とおき換える（$n'$ は新しい観測期間，$n$ は従来の観測期間）ことにより，たとえば月降水量から時間降水量を求めることができるとしてシミュレーションも行っている．

P. Allerup と H. Mapsen は「A statistical model for precipitation data —with application to interpolation problems」において月降水量の地域的内挿の問題を取り扱った．$x_{it}$ を対数月降水量として，$x_{it}\sim N(\xi+\alpha_i+\beta_t,\sigma^2)$，$\mathrm{cov}(x_{it_1},x_{it_2})=0$，$(t_1\neq t_2)$，$\mathrm{cov}(x_{it},x_{jt})=\gamma(\beta_t)(D(i,j))^2$ と仮定する．このモデルをもとに推定地点 $q$ と $p$ 地点，$t_0$ 時点の観測値 $x^{(p)}=(x_{i_1t_0},\cdots,x_{ipt_0})$ の同時分布から $x_{qt_0}$ の条件付平均値 $E(x_{qt_0}|x_i,t_0,\cdots,x_{ipt_0})$ を求め，これを内挿値としている．ここで $\gamma(\beta_t)$，$D(i,j)$ は半経験的に決められている．この方法による実測値と内挿値の対比を行っているがすべてがうまく適合してい

るわけではない．それは cov($x_{it_1}, x_{jt_2}$) について何らの考慮もしてないから当然かも知れない．今後改良すべき点であろう．

### 3. 気候・気象データの多変量解析

多変量解析では，空間相関関数を考慮した問題，時間空間両面を同時に考えねばならない問題等統計学の分野でも今後の研究課題となりそうないくつかの研究発表がなされた．K.R. Gabriel は「An approach to analysis and modeling multivariate spatial data in meteorology and climatology」において気象要素の空間分布の解析に空間相関行列とその回転を持ち込んでいる．気象要素としてたとえば降水量をとれば，降水量分布は一般に非対称で，空間相関関係を距離またはその2乗に反比例するという通常の取り扱いは成立しない．風向に沿った方向では距離のスケールを大きく，これに直交する方向では逆に小さくとる必要がある．座標ベクトルを回転し，さらにスケールを変換して得たモデルベクトルを基にして空間相関行列を構成している．相関行列の要素は1から基準化モデルベクトル間の距離の2乗の1/2を引いたものとして与えられている．データをもとにスケール変換行列，回転行列を推測する必要があり，この点に関しては実例もなく今後の研究に俟つ必要がある．

P.P. Jones, T.M.L. Wigley, K.R. Briffa は「Reconstructing surface pressure patterns using principal components regression on temperature and percipitation data」において，地上月平均気圧場で観測値のない格子点の値を気圧，気温，降水量の観測値から回帰方程式により内挿しようという試みである．ただし回帰方程式に入れる変数は主成分分析により選びなおしている．時間的には1780年までの過去，地域的には35～70°N，30°W～40°E の地域を覆い，1,116枚の天気図を作ったという．最終的な推定誤差の理論的検討は何も報告されてない．他の方法により推定された月平均値との相関が示されているが，その結果夏季を除いては高いと言える．

R.F. Cahalan は「EOF spectral estimation in climate analysis」において，EOF（経験的直交関数）の固有値分布を Dyson 分布（1962, Dyson がブラウン運動の物理学的考察から得た）により簡単な解析的型で与えている．累積分布曲線の性質をモンテカルロ・シミュレーションにより確かめ，また降水量の場合に固有値分布に“何らかの信号”が入っているかどうかの判定に利用している．

H.V. Storch は「Statistical aspects of empirical orthogonal functions based on small sample sizes」では標本数mが少ないときの EOF 固有値のバイアスが大きいことをモンテカルロ・シミュレーションにより示している．とくにmが15以下では推定された固有値の分布は正規分布とは言えない．バイアスが大きいと，現象として大きいスケールをもつものは過大に，小さいスケールをもつものは過少に推定される傾向があるという．

M.B. Richman は「specification of complex modes of circulation with T-mode fator analysis」において700 mb 高度偏差値の解析に因子分析を適用している．週間平均値を用い1,000個の格子点（0～150°E，20～60°N），1947年1月から1978年12月までの計1，664個のデータを用いている．このデータによる行列 $Z_{ij}$ の構成は $i=1,2,\cdots,N_{12}$（縦）を時間に，$j=1,2,\cdots,n$(横）を観測地点にとる方法（S-モード）と $i$ を地点，$j$ を時間にとる方法（T-モード）の2つでこれを比較した．この700 mb 高度面の解析では T-モードがよく実態を表わすという．事実，第1因子の因子得点図は Namias の teleconnection 図とよく一致している．

### 4. 時間—空間の両面にわたる確率過程モデルの構成

時間—空間分布の発表では実際例では未だしの感が多かった．V.K. Gupta は「Stochstic modeling of rainfall in space and time」では時間—空間のフレームで降水現象の数学的モデルを構成している．基本的点は点過程とランダム空間場を合成したモデル構成にある．P.L. Lamb と M.B. Richman は「An analysis of the space and time variation of growing season rainfall in the central United States」においてアメリカ中部で夏期対流性降水量についての統計的解析を行ったが，主成分分析の単なる応用と思われる．

### 5. 気候・気象データの時系列解析

時系列解析では岡本が「Time series analysis in climatic system models」において現象のスケールに応じた分析方法を選ぶことの必要性を強調し，短期気候変動の要因を探るために気温，気圧，層厚（850～500 mb），雲量の月平均値を多変量時系列として取上げ，赤池氏のノイズ寄与率によりこれら変量間の影響し合う方向を定めようと試みた．また長期気候変動においては数個の気候時系列から得られた“刈り込み系列”により1つの総

合気候指数を作成した．1200〜1900年にわたるこの指標の全体的傾向は海水面変動と大体一致している．なお岡本と岩瀬は気候変動曲線のトレンドにより，刈り込み系列の基準線（平均値）にずれがあったときの最大誤差を評価する方法を示した（保留論文）．

H.P. Junk は「Dynamic MESA-spectra of climatological time series」において最大エントロピー法(MESA)により推定されたペクトル全体の時間変化を3次元的立体図により示した．観測時点を一定区間だけ順次ずらせ各区間で得られたスペクトルの時間変化が Dynamic MESA-spectra と呼ばれるものである．最大エントロピー法での Burg アルゴリズムによるとフィルター長を増して行くと分解能がゆるくなり，いわゆる spectral shifting が起り，ピークがずれたり見掛け上の小さいピークが現れたりする．これを避けるためいくつかの数値フィルターを用いて予想されるずれたピークを除いて実際の解析を行っている．示されたスペクトルの時間変化図は興味あるものではあるが，非定常スペクトル解析の一般論とは何のつながりも述べられてない．

G. Padmanabhan and A.R. Rao は「Nonintegers spectral analysis of climatological data」において非整数ハーモニックスを用いたスペクトル解析の応用例を示した．この方法は各観測点での値を正弦波と余弦波に回帰させた多変量回帰モデルに基づく．実際のデータでは独立変数として取られる正弦波と余弦波が常に独立という保証はないし，スペクトル解析の背後にどんな時系列モデルを想定しているかも不明である．その他，各種のスペクトル解析法により有意な周期をいくつか摘出する研究（Schönwiese），同じく摘出された概周期から合成された周期関数を用いる予測に関する研究（Brier, Siddiqui, Haurwitz and Biondini）などの発表があった．

## 6. 気候インパクトの統計的評価

T.M.L. Wigley の極めて包括的ではあるが興味ある報告「The role of statistics in climate impact analysis」があった．社会・経済に対する気候インパクトの統計的解析の困難さ——データ・ベースの非均質性，非正規分布，非定常時系列，トレンドと自己相関との弁別の困難さ，相互に相関をもつ予測子，インパクト変数との見掛け上の相関，見出された回帰式(あるいは統計的関係)の時間による変化，非線形性などを指摘している．しかし従来知られた解析法でも，うまく使えば妥当な統計的結論を得ることを3つの実例で示している．その他水資源とくに水流出・貯蔵量への気候変化の影響（V. Klemeš, J. Něměc），冬季長期運輸業の走行距離と降雪積量との関係（J.P. Palutikof），アフリカにおける降水量変動の研究（J.A.M. Corte-Real など），回帰モデルにより，大麦生産量を平均降水量と気温から予測する研究（F.N. Kogom, 保留論文）があった．

## 7. 気候情報の価値

R. Krzystofowicz が「Oplimal use and economic value of prior data and categorical and probabilitic climate forecasts」において，最小リスクにもとづくベーズ解の立場から三種類の気候予報（単に事前確率に基づく普通のもの，カテゴリカルな予報，確率的予報）を比較している．いくつかの結論のうち興味ある点は次の通りである．(1) リスク最小化法（ベーズ解）を用いるには気候予報だけではなく事前的（歴史的）気候データあるいはそれらの確率モデルをも利用しなければならない．事前的気候データの利得（経済的価値）は予報の技術が減少するにつれて増加する．(2) カテゴリカルな予報の代わりに確率的気候予報を用いたときの利得は，(a) カテゴリカル予報の技術が減少したときかまたは (b) 事前的気候データの利得が少なくなったとき，その分だけ増加する．とくに事前的気候データが存在しないときは予報の計量化が最も重要である．気候予報の経済的価値（利得）について一般にベーズ流の考え方を知る上でよい論文であろう．他にベーズ解法を用いたものとしては流量調節において気候情報の経済的効果の判定(Duckstein et al) がある．ベーズ流の統計解析の利用としては NOAA の観測により北半球の陸上雪被覆面積の推定（E.S. Epstein）がある．そこで事前分布のパラメーターを苦労して半経験的に求めている．

R.A. Madden は「On the prospects for climate prediction」において気候予報可能性に対する1つの客観的評価を与えようとした．たとえば月単位で考えた気候変動をアンサンブル平均内の変動(climate noise と呼ぶ）とアンサンブル平均の変位（potentially predictable signals と呼ぶ）とに分ける．この潜在的予測可能性の測度を信号対雑音の比で表している．線形予測(単回帰)を仮定すれば，この比は相関係数を与えれば計算できる．

Nicolls は「Statistical climate prediction : The state of art」において，回帰予測を無批判に使うべきではなく，変数をふやせば決定係数は大きくなるから，予測が

うまく行くと安心してはならないと警告している．やはりデータの不均一性，欠測などを十分吟味してから予測式に用いることの重要さを強調している．EOF による気候予測のヨーロッパの例（M. Déqués），同じくアメリカ東北部の低気圧数の持続性に基づく季節的平均気温の EOF による予測（B.P. Hayden）の例が発表された．

### 8. 大気大循環モデルに基づく気候変動実験結果の統計的検定

大気大循環モデル（GCM）に基づく気候変動実験の実測値による検証は結局平均の差の検定に帰着されるが，多変量の検定統計量を考えねばならない．しかし検定のための標本の大きさは小さく，何等かの方法で検定統計量の次元を引き下げる努力がなされて来ている．検定統計量の標本分布は多変量分布モデルで一般には各成分間に相関がある．当然，1変数での検定結果をあつめても全体の検定とはならない．また逆にグローバルの検定から各成分間の有意性は必ずしも言えない．

R.E. Livezey は「Statistical analysis of general circulation model, climate simulation, sensitivity, and prediction experiments」において，そこに内在する多変量検定問題を考察し，Chevin, Hayashi, Shukla などの研究に対する批判的総合報告を行った．

### 9. 気候・気象統計学上の他の諸問題

ポルトガル科学アカデミーの J.T. de Oliveira が von Mises-Jenkinson type の各種極値分布（グンベル分布，フレシェ分布，ワイブル分布）のパラメーターの最大推定値を求めている．多変量化の一方法，時系列極値分布について言及しているが，これらの大部分はすでに研究された課題であるにすぎない．

K.J.A. Revfeim は「Stochastic process analysis of rainfall totals and extremes」において一雨降水量の総降水量とそこでの最大降水量を一雨降水量という点にある大きさをもった点過程として表現できるとしている．ある時間内で一雨降水という事象系列が与えられると，ポアソン点過程の下では総降水量と最大降水量の確率密度関数は事象の生起確率と事象の大きさの分布の関数として表現される．事象の大きさの分布は指数分布に従うとして漸近展開により近似確率（密度）関数を導いている．そこでのパラメーターは積率法で推定されるが，小標本での推定値の性質はよくない．反復計算法に基づく最尤法により推定する方法も示しているが収束はよくない．また最大風速（gusts）の分布にも適用できる可能性を示しているが，連続的現象ではどこからどこまで“一雨降水”に相当する事象と見做せるかが問題である．

### おわりに

以上多数の発表論文に接して感ずることは気候・気象統計といえども各国の統計学研究の水準の高さと裾野の広ろがりを反映しているということである．15～20年の歴史をもつ多くの統計学部をもつ米国の学者による研究発表は統計学的に一応無難のようにみられる．しかし一方統計後進国と見られる国々からの発表の中には極めて独創的な研究もあったが，そのレベルにかなりのバラツキがある．統計気候学という名前で行われる研究集会での発表内容も多種多様であり，必ずしも“気候”に限定されておらず，多くの従来の気象統計に属する応用実例のみの研究発表もあった点をお知らせしておく．残念なことに今回は全論文そのままが入手可能な形で印刷刊行される予定はない．次の第3回目の研究集会は3年後に Wien で開催予定ということでもあり，日本からはより多くの方々の参加発表があることを希望する．

### 文献

II International Meeting on Statistical Climatology, 1983: Instiuto Nacional de Meteorologia e Geofisica, Lisboa Portugal. (preprint)

## 出版情報

日本海洋学会・沿岸海洋研究部会から下記出版物につき日本気象学会会員への周知方依頼がありましたのでお知らせします．

書名：日本全国沿岸海洋誌

発行期日等：昭和60年2月末，B5版，約 1000頁

価格：1万8千円（ただし海洋および沿岸海洋関係の各学会*会員・*賛助会員・*団体会員は特価1万5千円）

* 日本気象学会も含まれる．

詳細につきましては下記にお問合せ下さい．

〒424　清水市三保 2389
東海大学海洋科学博物館内
沿岸海洋誌事務局