II-50　降水観測がもたらす情報量とその配置計画への応用

京都大学工学部　正員　○池淵周一
間組 K.K　正員　岡田知己

1 はしがき

水文資料の観測，収集は元来地味であり，かつ長年月を要するが，得られたデータの豊富さ，あるいは精度のよさは，治水・利水計画を問わず，実際の設計・操作問題においても直接的にその合理性・経済性を大きく左右するものであり，その必要性・重要性はいうまでもない。水問題でもっとも基本となる水文資料は降水であり，その空間的・時間的不確定構造を考えると，降水資料の観測・収集に関しては，最適な観測網体系の確立が最重要課題であろう。本研究では，従来の研究成果をふまえるとともに，観測の行動目的を直接，未知パラメータの推定精度向上におかず，確率的性格の強い日単位以上の降水量の空間的・時間的実現値を知ることにおいて，その情報量を定式化し，予算の許すかぎり観測所をふやす，過去のデータ・非データが内蔵する情報蓄積を最大限利用する，観測所の継続・廃止・新設を同時に考慮したO.R手法を導入する，といった立場を重視して，情報量概念に基づく降水観測網の最適配置基準を設定せんとするものである。したがって，観測がもたらす情報量を，実現値を知る以前にわれわれがもつ不確実さを観測が解消してくれると解釈し，Shannon流情報量で定義するとともに，その展開には事前分布を特徴とするベイズ論的立場をとっている。

2 情報量の一般的表現法

情報という概念は非常に広義なものとして理解されているが，本研究では情報を解析的操作の対象としてとらえていくために，不確実さの測度を減少してくれるものと規定し，ある降水現象を知ることにより得られる情報量として，[情報量]=[事前の不確実さ]−[事後の不確実さ]，で表現する。そして，不確実さの量的表現としては通常エントロピーと呼ばれている，$H(x)=-\sum P(x)\log P(x)$，$H(y|x)=-\sum P(x)\sum P(y|x)\log P(y|x)$ を用いた。したがって，xの実現値を知ることにより得られる情報量$I(x)$は，$I(x)=H(x)=-\sum P(x)\log P(x)$ で，yの実現値はわからないが，xの実現値を知ることによりyに関して得られる情報量$I(y|x)$は，$I(y|x)=H(y)-H(y|x)=-\sum P(y)\log P(y)+\sum P(x)\sum P(y|x)\log P(y|x)$ で与えられる。さて実際に情報量を算定するためには，確率変数x，yの周辺分布あるいは条件付周辺分布を知らなければならない。分布にはそれを特徴づけるパラメータがあり，未知パラメータの推定に最尤法や積率法がよく用いられるが，本研究ではデータ・非データ情報も加味する意味から事前分布を特徴とするベイズ論的立場をとっている。したがって，将来観測にあたっては，過去にデータが豊富であれば，$\xi(\theta|x_1,x_2)=\xi(\theta)f(x_1|\theta)f(x_2|\theta)/\int\xi(\theta)f(x_1|\theta)f(x_2|\theta)d\theta$ を繰りかえすことによって算出される事後確率分布が事前確率分布となり，そのパラメータ分布をもつ確率変数xのいずれかが実現することを観測するわけであり，データがない場合でも何らかの形で事前確率分布を与えれば，同様の考えで確率変数xを観測することになる。

3 観測所の継続・廃止推定および新設により得られる情報量

観測所配置計画にあたっては，各観測所が 1)既設観測所を継続する，2)既設観測所を廃止し，基幹観測所あるいは周辺観測所から推定する，3)観測所を新設する，のいずれかの行動を選択し，その行動にともなって得られる情報量と，その行動に要する費用とのバランス化が指向されよう。そこで以下では，1)，2)，3)の各場合に得られる情報量の一般的定義式を求め，さらに水文情報が正規分布および指数分布にしたがう場合の情報量をそれぞれ算定する。なお，以下では情報源から情報の発生する過程はエルゴード的独立過程としている。

(1)既設観測所を継続する場合に得られる情報量　i)一般的定義式；既設観測所において，これまでの観測により得られている情報(観測値)を$X(x_1,x_2,\cdots,x_N)$とすると，情報Xを得た後の未知パラメータθの事後確率分布$\xi(\theta|X)$は

ベイズの定理より，$\xi(\theta|\mathbb{X})=\xi(\theta)f(x_1|\theta)\cdot f(x_2|\theta)\cdots f(x_N|\theta)/\int\xi(\theta)f(x_1|\theta)\cdots f(x_N|\theta)d\theta$ で与えられる。$\xi(\theta|\mathbb{X})$が求められると，$\mathbb{X}$および将来観測$\mathbb{X}^*(x_{N+1},x_{N+2},\cdots,x_{N+n-1})$が与えられての$x_{N+n}$の周辺密度関数$f(x_{N+n}|\mathbb{X},x_{N+1},\cdots,x_{N+n-1})$および$\mathbb{X}$が与えられての$\mathbb{X}^*(x_{N+1},\cdots,x_{N+n})$の結合密度関数$f(x_{N+1},\cdots,x_{N+n}|\mathbb{X})$はそれぞれ次式で与えられる。

$f(x_{N+n}|\mathbb{X},x_{N+1},\cdots,x_{N+n-1})=\int\xi(\theta|\mathbb{X})f(x_{N+1}|\theta)\cdots f(x_{N+n}|\theta)d\theta/\int\xi(\theta|\mathbb{X})f(x_{N+1}|\theta)\cdots f(x_{N+n-1}|\theta)d\theta$ ……(1) $f(x_{N+1},\cdots,x_{N+n}|\mathbb{X})=\int\xi(\theta|\mathbb{X})f(x_{N+1}|\theta)\cdots f(x_{N+n}|\theta)d\theta$ ……(2) ここに，$f(x|\theta)$はθが与えられてのxのしたがう確率法則である。ところで，観測継続の場合は観測により真のxの値を知ることになるから，事後の不確実さは0となり，したがってx_{N+n}を観測することにより得られる情報量は，$I^{(1)}(x_{N+n}|\mathbb{X},x_{N+1},\cdots,x_{N+n-1})=-\int\cdots\int f(x_{N+1},\cdots,x_{N+n-1}|\mathbb{X})f(x_{N+n}|\mathbb{X},x_{N+1},\cdots,x_{N+n-1})\log f(x_{N+n}|\mathbb{X},x_{N+1},\cdots,x_{N+n-1})dx_{N+1}\cdots dx_{N+n}$ ……(3) で与えられ，さらに将来$\mathbb{X}^*(x_{N+1}\cdots x_{N+n})$を観測することにより期待される総情報量は，${}_TI^{(1)}(x_{N+1},\cdots,x_{N+n}|\mathbb{X})=\sum_{n=1}^{n}I^{(1)}(x_{N+n}|\mathbb{X},x_{N+1},\cdots,x_{N+n-1})=-\int\cdots\int f(x_{N+1},\cdots,x_{N+n}|\mathbb{X})\log f(x_{N+1},\cdots,x_{N+n}|\mathbb{X})dx_{N+1}\cdots dx_{N+n}$ ……(4) で与えられる。ii)正規情報の場合；平均値θが未知，分散v_xが既知である正規分布に対してはθの事前確率分布を$N(m_\theta,v_\theta)$で与えると，(3)式，(4)式はそれぞれ次式で与えられる。$I_N^{(1)}(x_{N+n}|\mathbb{X},x_{N+1},\cdots,x_{N+n-1})=\frac{1}{2}(1+\log 2\pi v_x)+\frac{1}{2}\log\{(v_x+n\cdot v_{\theta|x(N)})/(v_x+(n-1)v_{\theta|x(N)})\}$ ……(5)，${}_TI_N^{(1)}(x_{N+1},\cdots,x_{N+n}|\mathbb{X})=\frac{n}{2}(1+\log 2\pi v_x)+\frac{1}{2}\log\left(1+\frac{n\cdot v_{\theta|x(N)}}{v_x}\right)$ ……(6) ここに，$v_{\theta|x(N)}=v_\theta\cdot v_x/(N\cdot v_\theta+v_x)$ である。なお，θの事前分布が，あたかも一様分布とみなせる場合とかNが大きい場合には$v_{\theta|x(N)}=v_x/N$となるので，上式はそれぞれ簡単に，$I_N^{(1)}=\frac{1}{2}(1+\log 2\pi v_x)+\frac{1}{2}\log\frac{N+n}{N+n-1}$ ……(7)，${}_TI_N^{(1)}=\frac{n}{2}(1+\log 2\pi v_x)+\frac{1}{2}\log\frac{N+n}{N}$ ……(8) となる。以上は分散v_xを既知とした場合の展開式であったが，実際には観測数の増加とともに平均値だけでなく，v_xの精度も向上するのであるから，厳密にはθ，v_x両者を未知として解析する必要がある。この場合には事後確率分布$\xi(\theta,v_x|\mathbb{X})$は正規ガンマ分布，$\xi(\theta|\mathbb{X})$は一般化されたステューデント分布で与えられ，これらをまともに使ったのでは以後の計算が非常に煩雑になってしまう。しかし幸いにもNが大きいときには$S_x^2=\sum_{i=1}^{N}(x_i-\bar{x})^2/N-1$を$v_x$の推定値とし$x_1,x_2,\cdots x_N$が既知の分散$v_x=S_x^2$をもって正規分布をしているとみなすことができ，また，いくつかのデータ例について厳密解を求めた結果では，$v_x(=S_x^2)$既知とした場合と大きく異ならないので，一般にも(7)，(8)式のv_xをS_x^2で置きかえれば十分である。iii)指数情報の場合；この場合には日降水量を考えているので，一般に過去の観測数Nは大きく(たとえば，1年であっても季節ごとに分割すればN=90の観測データがあることになる)，事後確率分布$\xi(\theta|\mathbb{X})$は $\xi(\theta|\mathbb{X})=\frac{A^{N+1}}{N!}\theta^N e^{-A\theta}$ で十分近似できる。ここに，$A=\sum_{i=1}^{N}x_i$である。正規情報の場合と同様に展開すれば，$I_E^{(1)}(x_{N+n}|\mathbb{X},x_{N+1},\cdots,x_{N+n-1})=1+\log\frac{x_1+x_2+\cdots+x_N}{N}+\log\frac{N+n}{N+n-1}$ ……(9) ${}_TI_E^{(1)}(x_{N+1},x_{N+2},\cdots,x_{N+n}|\mathbb{X})=n\left(1+\log\frac{x_1+x_2+\cdots+x_N}{N}\right)+\log\frac{N+n}{N}$ ……(10) が得られ，正規情報とよく似た式となる。

(2)既設観測所を廃止し，基幹観測所から推定することにより得られる情報量　i)一般的定義式；推定観測所におけるこれまでの観測値を$\mathbb{Y}(y_1,\cdots,y_M)$，それに対応する基幹観測所の観測値を$\mathbb{X}(x_1,x_2,\cdots,x_M)$とすれば(一般には$N\geq M$)，$\beta_1$，$\beta_2$を回帰係数とし，$\varepsilon$を攪乱項とする線形回帰モデルは，$y_i=\beta_1+\beta_2x_i+\varepsilon_i$ で与えられ，$\mathbb{X}$，$\mathbb{Y}$が観測された後の$\xi(\beta_1,\beta_2|\mathbb{X},\mathbb{Y})$は，$\xi(\beta_1,\beta_2|\mathbb{X},\mathbb{Y})=\xi(\beta_1,\beta_2)f(y_1|\beta_1,\beta_2,x_1)\cdots f(y_M|\beta_1,\beta_2,x_M)/\iint\xi(\beta_1,\beta_2)f(y_1|\beta_1,\beta_2,x_1)\cdots f(y_M|\beta_1,\beta_2,x_M)d\beta_1d\beta_2$ で与えられる。$\xi(\beta_1,\beta_2|\mathbb{X},\mathbb{Y})$が求められると，$\mathbb{X}$，$\mathbb{Y}$および$x_{M+n}$が与えられての$y_{M+n}$の周辺密度関数$f(y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+n})$は，$f(y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+n})=\iint\xi(\beta_1,\beta_2|\mathbb{X},\mathbb{Y})f(y_{M+n}|\beta_1,\beta_2,x_{M+n})d\beta_1d\beta_2$で表わされ$y_{M+n}$の事後の不確実さが求まるので，$\boxed{2}$の議論から観測所を廃止し推定した場合に得られる情報量$I^{(2)}(y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+n})$および総情報量${}_TI^{(2)}(y_{M+1},\cdots,y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+1},\cdots,x_{M+n})$は，それぞれ次式で与えられる。$I^{(2)}(y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+n})=-\int f(y_{M+n}|\mathbb{Y})\log f(y_{M+n}|\mathbb{Y})dy_{M+n}+\iint f(x_{M+n}|\mathbb{X})f(y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+n})\log f(y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+n})dx_{M+n}dy_{M+n}$ ……(11) ${}_TI^{(2)}(y_{M+1},\cdots,y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+1},\cdots,x_{M+n})=\sum_{n=1}^{n}I^{(2)}(y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+n})$ ……(12) ここに $f(y_{M+n}|\mathbb{Y})=\int\xi(\theta_y|\mathbb{Y})f(y_{M+n}|\theta_y)d\theta_y$ である。

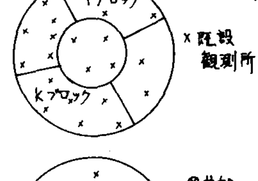
ii)正規情報の場合；$\mathbb{X}$，$\mathbb{Y}$が観測されてのβ_1，β_2の事後確率分布$\xi(\beta_1,\beta_2|\mathbb{X},\mathbb{Y})$は事前確率分布$\xi(\beta_1,\beta_2)$が正規分布で，その平均が$m=(m_1,m_2)'$，分散共分散行列が$\mathbb{V}=\|v_{ij}\|$，$i,j=1,2$である場合，やはり正規分布となり，$x_{M+n}$が与えられての$y_{M+n}$の周辺密度関数は $f(y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+n})=\frac{1}{\sqrt{2\pi v_{xy}}}e^{-\frac{1}{2v_{xy}}(y_{M+n}-m_1'-m_2'x_{M+n})^2}$，ここに $v_{xy}=v_{11}'+2v_{12}'x_{M+n}+v_{22}'x_{M+n}^2+v_\varepsilon$ であり，$\begin{bmatrix}v_{11}' & v_{12}'\\ v_{21}' & v_{22}'\end{bmatrix}=\begin{bmatrix}h_{11}+N/v_\varepsilon & h_{12}+\sum x_i/v_\varepsilon\\ h_{21}+\sum x_i/v_\varepsilon & h_{22}+\sum x_i^2/v_\varepsilon\end{bmatrix}$，$\mathbb{V}^{-1}=\|h_{ij}\|$，$v_\varepsilon$＝攪乱項$\varepsilon$の分散である。[1)] したがって，この場合の情報量は(11)式より，$I_N^{(2)}(y_{M+n}|\mathbb{X},\mathbb{Y},x_{M+n})=\frac{1}{2}\{\log 2\pi v_y-\log 2\pi(v_{11}'+v_\varepsilon-\frac{v_{12}'^2}{v_{22}'})\}+\frac{1}{2}\log\left(\frac{v_y+n\cdot v_{\theta|y(M)}}{v_y+(n-1)v_{\theta|y(M)}}\right)-\frac{1}{2}\left\{\frac{1}{\sigma^2 2\delta}-\frac{1\cdot3}{2\sigma^4(2\delta)^2}+\frac{1\cdot3\cdot5}{3\sigma^6(2\delta)^4}-\cdots\right\}$ ……(13) で与えられる。なお上式の誘導においては$\log\{(x_{M+n}-\frac{v_{12}'}{v_{22}'})^2+(v_{11}'v_{22}'+v_{22}'v_\varepsilon-v_{12}'^2)/v_{22}'^2\}$をマクローリン展開している。ここに，$\sigma^2=(v_{11}'v_{22}'+v_{22}'v_\varepsilon+v_{12}'^2)/v_{22}'^2$，$2\delta=(v_x+(n-1)v_{\theta|x(N)})/(v_x+nv_{\theta|x(N)})$ である。

ところで，上式は事前分布 $\xi(\theta_x)$, $\xi(\theta_y)$ の分散要素，$\xi(\beta_1,\beta_2)$ の分散－共分散行列要素および過去の両観測所のデータ構造要素をすべて含んでおり，要素間の大小関係により推定情報量は変化する。しかし，各要素が同じように情報量増減に寄与するわけでなく，事前分布の分散が大きいほど情報量は大きくなる。いま，$\xi(\theta_y)$, $\xi(\beta_1,\beta_2)$ の変動が観測データの個数に比して十分小さく，あたかも一様分布であるかのように計算を進めることが許される場合を考えると，上式はより簡単になり，さらにM，nが大きいと近似的に，$I_N^{(2)}=\frac{1}{2}(\log 2\pi v_y-2\pi v_\varepsilon)$ ……(14)
総情報量も $\frac{n}{2}(\log 2\pi v_y-\log 2\pi v_\varepsilon)$ で近似される。厳密には，(13)式を計算するわけであるが，(14)式と比較すると，他の項はかなり小さくなり，実用的には(14)式で十分近似できる。(ii)指数情報の場合；この場合は，y_{M+n} の条件付周辺密度関数が，0次の変形ベッセル関数を含むので，情報量算定は解析的にも数値計算的にも困難である。そこで本研究では少なくとも数値計算が可能となるよう，この場合に限り伝統的立場にたった，情報量算定式を挙げておく。$I_E^{(2)}(y_{M+n}|X,Y,x_{M+n})=1+\log\frac{y_1+y_2+\cdots+y_M}{M}+\int_0^\infty\int_0^\infty \hat{\theta}_x e^{-\hat{\theta}_x x_{M+n}}\cdot\frac{1}{\hat{\beta}_1}e^{-\frac{1}{\hat{\beta}_1}(y_{M+n}+\hat{\beta}_2 x_{M+n})}\times I_0\left(2\sqrt{\frac{\hat{\beta}_2}{\hat{\beta}_1}x_{M+n}y_{M+n}}\right)\times\log\left\{\frac{1}{\hat{\beta}_1}e^{-\frac{1}{\hat{\beta}_1}(y_{M+n}+\hat{\beta}_2 x_{M+n})}I_0\left(2\sqrt{\frac{\hat{\beta}_1}{\hat{\beta}_2}x_{M+n}y_{M+n}}\right)\right\}dx_{M+n}dy_{M+n}$，${}_TI_E^{(2)}(y_{M+1},\cdots,y_{M+n}|X,Y,x_{M+1},\cdots,x_{M+n})=nI_E^{(2)}(y_{M+n}|X,Y,x_{M+n})$ ここに，$\hat{\theta}_x=M/\sum x_i$，$\hat{\theta}_y=M/\sum y_i$，$\hat{\beta}_1=\sum y_i/M-\hat{\beta}_2\sum x_i/M$，$\hat{\beta}_2=\hat{\rho}\cdot\sum y_i/\sum x_i$，$\hat{\rho}=M\cdot\sum x_i y_i/\sum x_i\sum y_i-1$ ただし積率解）である。

3)観測所新設により得られる情報量　新設の場合には，観測資料がないために将来の観測情報に対する事前確率分布はデータ・非データ情報から判断される $\xi(\theta)$ そのものとなり，継続情報量の展開式から明らかなように，正規情報に関する新設情報量および総情報量は，$I_N^{(3)}(x_n)=\frac{1}{2}(1+\log 2\pi v)+\frac{1}{2}\log\frac{v+nv_0}{v+(n-1)v_0}$ ……(15)，${}_TI_N^{(3)}(x_1,\cdots,x_n)=\frac{n}{2}(1+\log 2\pi v)+\frac{1}{2}\log\left(1+\frac{n}{v}v_0\right)$ ……(16)で与えられ，また指数情報に関してもいままでの展開から近似的に，$I_E^{(3)}(x_n)=1+\log\bar{x}$ ……(17)，${}_TI_E^{(3)}(x_1,x_2,\cdots,x_n)=n(1+\log\bar{x})$ ……(18)　で与えられよう。ところで問題は新設地点の v_0，v，$\bar{x}$ の与え方であり，とくに量的効果としては v，$\bar{x}$ の推定が重要となる。既存データがある場合には，1)近接地点の値で代表する，2)既設観測所の値から描いた等値線から内・外挿する，3)地形，気象因子との回帰構造から推定する，などの方法が考えられ，データが豊富なほど後の方法に依存するであろう。なお，対象区域にほとんど既存データがない場合には，4)類似の気候帯・地形帯での降水情報および従来の経験・判断情報をふまえて推定せざるをえないであろう。

4 降水観測網配置計画への応用

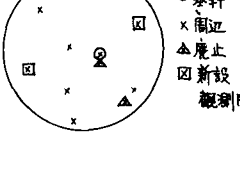
不確実な降水現象を的確に把握するためには，それ相当に密な観測網を配置すればよいのであろうが，現実には経済的にも技術的にも得策とはいえない。ここでは前述の情報量概念を基礎に，最適な観測網配置計画の手順を概述する。手順1；まず，観測網配置計画における対象区域を図のようにK個のブロックに分割する。分割には各ブロック内に存在する既設観測所の継続情報量が類似している区域を1つのブロックとし，情報量が多く得られる区域では観測所を密に，またそのブロックの面積が広い場合にはより多くの観測所が設置できるように全投資額をそのブロックの平均情報量と面積の積に比例するよう分割する。手順2；各ブロック内の基幹観測所を選ぶとともに，観測所が一様に分布するよう図のごとく新設したり廃止したりする。手順3；一様に分布された観測所に必要な費用が，そのブロックへの投資額より小であれば問題はないが，その逆の場合には，$I_{k,\ell}^{(1)}-I_{k,\ell}^{(2)}/I_{k,\ell}^{(1)}<\varepsilon$ なる基準のもとに観測所を廃止する。手順4；$\sum_{\ell=1}^{m_k}\{C_{k,\ell}^{(1)}\delta_{k,\ell}+C_{k,\ell}^{(2)}(1-\delta_{k,\ell})\}+\sum_{\ell=m_k+1}^{m_k+m_k'}C_{k,\ell}^{(3)}\delta_{k,\ell}\leq C_k$，$\delta_{k,\ell}=0$ or 1 の制約下に，$\sum_{\ell=1}^{m_k}\{I_{k,\ell}^{(1)}\delta_{k,\ell}+I_{k,\ell}^{(2)}(1-\delta_{k,\ell})\}+\sum_{\ell=m_k+1}I_{k,\ell}^{(3)}\delta_{k,\ell}\rightarrow \max$



ここに，m_k，m_k' はkブロック内の既設観測所および新設観測所数，$I_{k,\ell}^{(1)}$, $I_{k,\ell}^{(2)}$, $I_{k,\ell}^{(3)}$ は継続・廃止推定・新設情報量，$C_{k,\ell}^{(1)}$, $C_{k,\ell}^{(2)}$, $C_{k,\ell}^{(3)}$ はそれぞれの費用，$\delta_{k,\ell}$ は決定変数で，継続新設なら1，廃止推定なら0，C_k はkブロックへの投資額である。

5 あとがき

本論文では情報量の展開式に終始したが，こうした理論的アプローチを琵琶湖流域にも適用しているので，時間が許せば講演時に述べたい。



参考文献；1)宮沢光一：情報・決定理論序説，岩波書店，pp.248~250，1971，11.