

# PHÂN BỐ THỐNG KÊ CÁC ĐẠI LƯỢNG KHÍ TƯỢNG CAO KHÔNG TRÊN LÃNH THỔ VIỆT NAM VÀ KHU VỰC ĐỊA LÝ LÂN CẬN

TS. Nguyễn Đăng Quế

Trung tâm Tư liệu khí tượng thủy văn

## 1. Một số khái niệm cơ bản

Chuỗi số liệu thống kê là dạng đầu tiên tập hợp các giá trị của đại lượng khí tượng chuẩn bị cho công tác xử lý khí hậu. Khi lượng số liệu của tập hợp thống kê tăng lên đến một mức nào đó người ta phải tìm cách tổ chức chúng về dạng chặt chẽ nhằm biểu diễn rõ các đặc trưng thống kê của chuỗi số liệu cần xử lý. Phân bố thống kê là một trong những dạng được dùng khá phổ biến cho mục đích đó.

Thực chất phân bố thống kê là tập hợp các tần suất xuất hiện của các đại lượng khí tượng trong từng khoảng giá trị từ nhỏ nhất đến lớn nhất có thể có. Có hai dạng phân bố thống kê thường được sử dụng để biểu diễn chuỗi số liệu khí tượng: phân bố vi phân và phân bố tích phân.

### a. Phân bố vi phân

Phân bố vi phân được xây dựng trên cơ sở tần suất tương đối của đại lượng khí tượng cần nghiên cứu. Nếu ký hiệu  $m_i$  là tần suất tuyệt đối của đại lượng một chiều tại khoảng giá trị (i) và n là độ dài của toàn chuỗi, ta có:  $\sum_{i=1}^K m_i = n$ , trong đó K là số lượng các khoảng giá trị. Như vậy, tần suất tương đối (P) tại khoảng giá trị (i) được xác định:  $P_i = \frac{m_i}{n}$  ( $0 \leq P_i < 1$  và  $\sum_{i=1}^K P_i = 1$ ).

Từ các khái niệm tần suất tương đối và tần suất tuyệt đối ta có thêm các khái niệm về mật độ tương đối ( $\alpha$ ) và mật độ tuyệt đối ( $\beta$ ) của phân bố thống kê:

$$\alpha = \frac{P_i}{\Delta x_i} \quad \beta = \frac{m_i}{\Delta x_i}, \text{ trong đó } \Delta x_i \text{ là độ rộng của khoảng giá trị (i) trong thang độ.}$$

### b. Phân bố tích phân

Để xây dựng phân bố tích phân người ta sử dụng tần suất tích lũy, tức là tần suất các giá trị của đại lượng ngẫu nhiên không lớn hơn hoặc không nhỏ hơn một giá trị cho trước nào đó.

$$m_{X \leq x_i} = \sum_{j=1}^i P_j, \quad m_{X \geq x_i} = \sum_{i+1}^k m_i$$
$$P_{X \leq x_i} = \sum_{j=1}^i P_j, \quad P_{X \geq x_i} = \sum_{i+1}^k P_j$$

trong đó  $x_i$  là giá trị biên trên của khoảng giá trị (i) trong thang độ.

### c. Phân bố thống kê nhiều chiều

Ngoài phân bố một chiều trên đây trong công tác xử lý số liệu khí tượng người ta còn dùng các dạng phân bố nhiều chiều. So với phân bố một chiều thì phân bố nhiều chiều rất phức tạp. Trước tiên ta xem xét dạng phân bố hai chiều, tức là tổ hợp của hai đại lượng khí tượng.

Tần suất tuyệt đối của một phân bố hai chiều ( $m_{ij}$ ) là số lần giá trị của đại lượng khí tượng này rơi vào thang bậc (i), đồng thời giá trị tương ứng của đại lượng kia rơi vào thang bậc (j).

Tần suất tương đối của phân bố hai chiều ( $P_{ij}$ ) được xác định:  $P_{ij} = \frac{m_{ij}}{n}$ , trong đó n - số lượng cặp giá trị.

Tương tự như trong trường hợp phân bố một chiều, đối với phân bố hai chiều người ta cũng dùng các đại lượng đặc trưng như tần suất tích lũy tuyệt đối hai chiều ( $m_{x \leq x_i, y \leq y_j}, m_{x \geq x_i, y \geq y_j}$ ) và tần suất tích lũy tương đối hai chiều ( $P_{x \leq x_i, y \leq y_j}, P_{x \geq x_i, y \geq y_j}$ ).

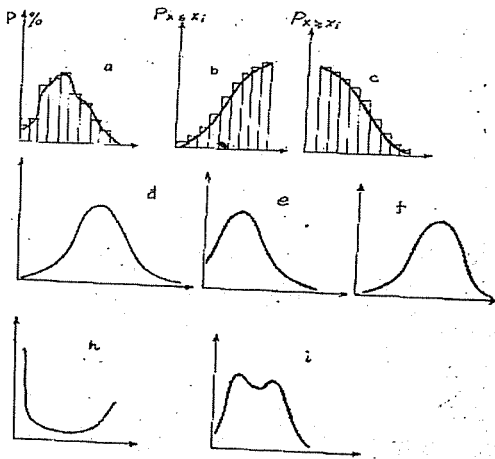
Phân bố ba chiều có dạng phức tạp hơn và trên thực tế ít khi được sử dụng.

#### d. Dạng đồ thị của phân bố thống kê

Hàm phân bố thường được biểu thị dưới dạng đồ thị. Trên hình 1 minh họa thí dụ về đồ thị hàm phân bố vi phân (hình 1a) và các đồ thị hàm phân bố tích phân (các hình 1b, 1c). Tuy vậy, trên thực tế, các đồ thị hàm phân bố vi phân được sử dụng phổ biến hơn. Đồ thị hàm phân bố vi phân thường có các dạng: đường cong một đỉnh (các hình 1d, 1e, 1f), dạng hình chữ U (hình 1h) và dạng nhiều đỉnh (hình 1i), trong đó dạng một đỉnh là phổ biến nhất.

Dạng đường cong phân bố một đỉnh được phân thành nhiều loại: đối xứng (hình 1d), bất đối xứng dương (hình 1e), bất đối xứng âm (hình 1f). Ở các đại lượng vật lý có giá trị khó đạt đến giá trị ngưỡng giới hạn, đường cong phân bố thường đối xứng hoặc bất đối xứng nhẹ. Sự bất đối xứng mạnh thường xảy ra ở đồ thị hàm phân bố của đại lượng có giá trị bị hạn chế bởi các giới hạn "cứng".

Dạng đường cong hình chữ "U" thường quan sát thấy ở các đại lượng có giới hạn "cứng" từ hai phía.



Hình 1. Các dạng đường cong phân bố thống kê: phân bố vi phân (1a), phân bố tích phân (1b, 1c). Các dạng đồ thị phân bố vi phân thường gặp: đối xứng một đỉnh (1d), bất đối xứng dương (1e), bất đối xứng âm (1f), dạng hình chữ "U" (1h), dạng nhiều đỉnh (1i).

## 2. Phân bố thống kê số liệu cao không mùa hè trên lãnh thổ Việt Nam và khu vực địa lý lân cận

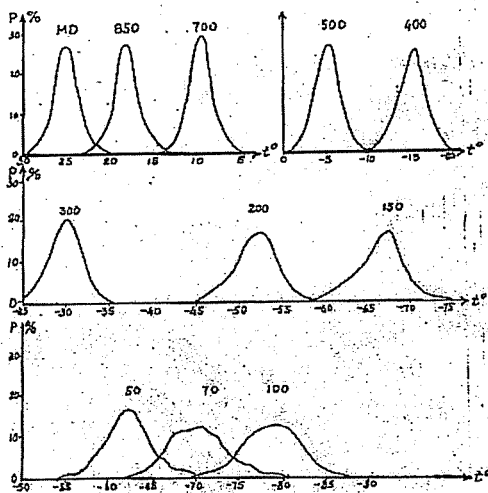
Để có được nhận thức cơ bản làm định hướng cho công tác xử lý số liệu khí tượng cao không, trong công trình này đã xây dựng một số các đường cong phân bố dựa trên chuỗi số liệu của các trạm thám không vô tuyến của Việt Nam và khu vực địa lý lân cận. Vì điều kiện hạn chế trong việc thu thập số liệu ban đầu, trong bài này

chỉ lập được các phân bố thống kê cho từng tháng và cả mùa hè. Hàm phân bố được xây dựng cho các đại lượng: độ cao địa thế vị, nhiệt độ không khí, độ ẩm, hướng và tốc độ gió trên các mực đẳng áp chuẩn trong phạm vi tầng đối lưu.

### a. Nhiệt độ không khí

Hình 2 là đồ thị phân bố thống kê nhiệt độ cho cả mùa hè trên các mực đẳng áp khác nhau. Xem xét các đồ thị trên hình 2 nhận thấy hàm phân lớp của nhiệt độ không khí trên tất cả các mực đều có một cực đại. Tần suất lặp lại của các giá trị nhiệt độ đều khá tập trung xung quanh giá trị trung bình nhiều năm. Độ rộng của dải biến thiên nhiệt độ trên các mực trong lớp từ mặt đất đến 300hPa không lớn (xấp xỉ  $10^{\circ}\text{C}$ ), nhưng từ độ cao mực 200hPa trở lên dải biến thiên bắt đầu mở rộng ra. Độ tập trung giá trị xung quanh giá trị trung bình nhiều năm càng lên cao càng giảm (độ dốc của đường cong phân bố giảm dần).

Nhìn chung, dạng đường cong của phân bố nhiệt độ gần với đường cong phân bố chuẩn. Sự tăng nhẹ hiện tượng bất đối xứng trên các mực trong lớp trên, theo chúng tôi, liên quan đến hoạt động của lớp đối lưu hạn và ảnh hưởng của đối gió đông nhiệt đới (đồ thị hàm phân bố trên các mực này có xu hướng bất đối xứng dương - độ dốc đồ thị phía nhiệt độ cao lớn hơn phía nhiệt độ thấp).



Hình 2. Đồ thị hàm phân bố thống kê của nhiệt độ không khí trên các mực: mặt đất, 850, 700, 500, 400, 300, 200, 150, 100, 70 và 50hPa

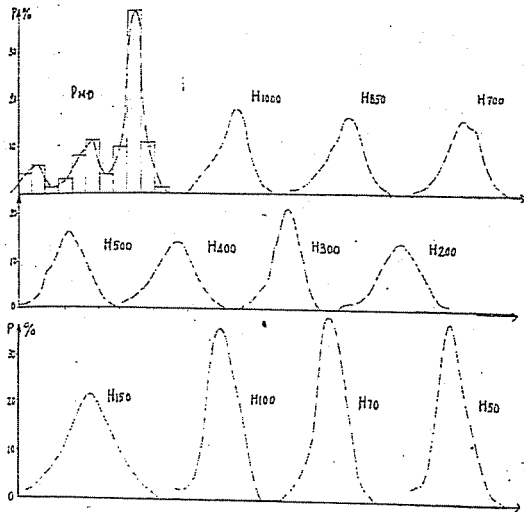
Như vậy, xử lý số liệu nhiệt độ không khí có thể áp dụng phương pháp tính toán đối với dạng hàm phân bố chuẩn nhưng cần lưu ý khi xử lý số liệu các mực từ 200hPa trở lên.

### b. Độ cao địa thế vị

Trên hình 3 là đồ thị hàm phân bố áp suất mặt đất và độ cao địa thế vị trên các mực đẳng áp chuẩn khác nhau. Qua đồ thị có thể thấy trên các mực tầng thấp có hiện tượng bất đối xứng âm. Càng lên cao các đồ thị chuyển dần sang bất đối xứng dương.

Trên đồ thị hàm phân bố của áp suất thể hiện khá rõ tính bất đồng nhất của chuỗi số liệu. Ngoài một cực đại chính nằm trong khoảng từ 1005-1010hPa, trên đồ thị còn có 2 cực đại phụ nằm ở phía có áp suất thấp hơn 1000hPa. Tính bất đồng nhất này, theo chúng tôi, liên quan đến bất đồng nhất về mặt đệm và sự tranh chấp ảnh hưởng tại vùng giáp ranh giữa các vùng lân cận có điều kiện khí hậu khác nhau. Thực tế cho thấy số liệu áp suất thấp hơn 1000hPa chủ yếu xuất hiện tại các trạm thuộc vùng núi phía bắc và Tây Nguyên. Đây là các vùng có chế độ mặt đệm đa dạng và

phức tạp nhất. Phương án tối ưu nhất trong trường hợp này là tách số liệu khí áp của các trạm vùng núi phía bắc và Tây Nguyên để xử lý riêng. Vấn đề này cần được nghiên cứu kỹ khi khảo sát hàm phân bố thống kê trên các chuỗi số liệu khí tượng bề mặt.



Hình 3. Đồ thị hàm phân bố thống kê của áp suất không khí  $P_{MD}$  và độ cao địa thế vị các mực: 1000, 850, 700, 500, 400, 300, 200, 150, 100, 70 và 50hPa.

Các đồ thị hàm phân bố thống kê của độ cao địa thế vị trên tất cả các mực đẳng áp đều chỉ có một đỉnh. Đặc biệt, tại các mực trong nửa dưới của tầng đối lưu, các đồ thị hàm phân bố có dạng gần phân bố chuẩn. Càng lên cao tính bất đối xứng dương càng tăng dần lên.

Như vậy, để xử lý số liệu địa thế vị, cũng có thể áp dụng phương pháp phân tích tính toán theo dạng phân bố gần chuẩn, tuy cần lưu ý hiện tượng bất đối xứng dương trên các mực phía trên của tầng đối lưu. Riêng đối với chuỗi số liệu áp suất, cần áp dụng biện pháp xử lý đặc biệt nhằm giải quyết hiện tượng bất đồng nhất.

### c. Độ ẩm không khí

Như trong phần mở đầu đã nêu, độ ẩm là đại lượng có khoảng giá trị nằm giữa hai giá trị biên 'cứng' nên đồ thị có dạng hai đỉnh. Qua các đồ thị trong hình 4, nhận thấy trên mặt đất và mực 1000hPa, đường cong hàm phân bố có dạng hai đỉnh rõ rệt. Tần suất cực đại rơi vào các khoảng giá trị độ ẩm từ 15 đến 20% và từ 90 đến 100%. Nhưng càng lên cao hiện tượng hai đỉnh dần dần mất đi và thay vào đó là dạng đường cong một đỉnh với cực đại rơi vào khoảng giá trị độ ẩm từ 50 đến 60%. Điều này phù hợp với thực tế trên các mực này độ ẩm không khí không đạt tới giá trị biên cứng thứ hai (100%).

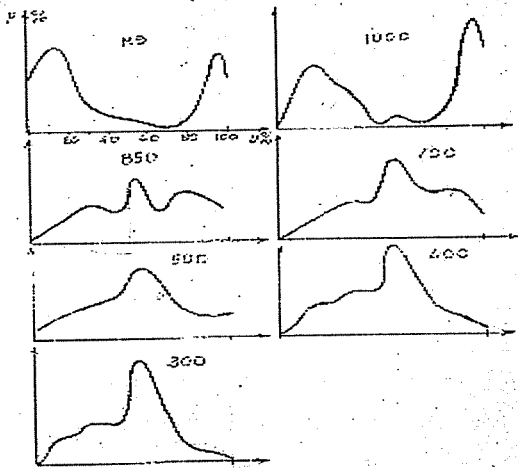
Cũng có thể nhìn thấy hiện tượng bất đối xứng từ trái sang phải khi đi từ các mực thấp lên các mực cao. Cụ thể ở tầng thấp, tần suất của không khí có độ ẩm cao hơn tần suất của không khí có độ ẩm thấp. Càng lên cao quan hệ này càng có xu hướng chuyển dần sang chiều ngược lại.

Như vậy, khi xử lý số liệu độ ẩm không khí cần tách các mực sát mặt đất ra để xử lý theo phương pháp riêng (tách chuỗi thành các chuỗi nhỏ tương ứng với số đỉnh). Đối với các mực trên cao, việc xử lý số liệu được tiến hành theo dạng một đỉnh, song cần lưu ý đến tính đặc thù của các đường cong phân bố độ ẩm.

Như vậy, chuỗi số liệu ẩm là chuỗi bất đồng nhất. Tính bất đồng nhất đó thể hiện khá rõ sự tranh chấp ảnh hưởng của các đới gió hoạt động trong các lớp khác nhau thuộc tầng đối lưu.

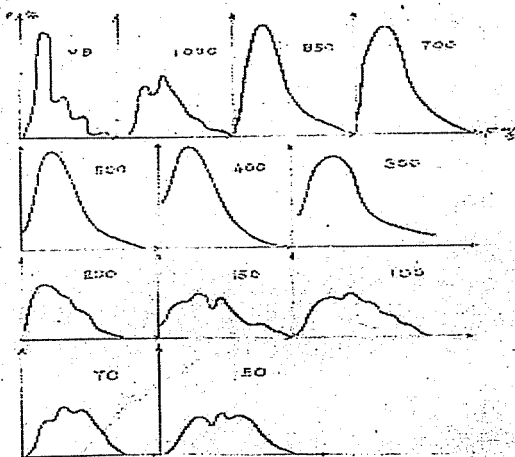
#### d. Hướng và tốc độ gió

Biết rằng trên khu vực địa lý, về mùa hè chủ yếu có sự tranh chấp của các đới gió tây-nam và gió đông nhiệt đới. Vì vậy, tần suất các đới gió này chiếm ưu thế trên biểu đồ tần suất hướng gió. Vấn đề này đã được xem xét kỹ và đã được nêu trong [1]. Trong công trình này chỉ đề cập đến hàm phân bố tốc độ gió trên các mực đẳng áp khác nhau.



Hình 4. Đồ thị hàm phân bố thống kê của độ ẩm không khí trên các mực: mặt đất, 1000, 850, 700, 500, 400 và 300hPa.

Xem xét đồ thị hàm phân bố thống kê tốc độ gió trên các mực đẳng áp khác nhau (hình 5), nhận thấy trên mặt đất và mực 1000hPa cũng như trên các mực từ mực 200hPa trở lên, đường cong hàm phân bố thể hiện khá rõ nét tính bất đồng nhất trong chuỗi số liệu về tốc độ gió. Riêng trên các mực trong lớp từ 850hPa đến 300hPa đồ thị hàm phân bố của tốc độ gió có dạng một đỉnh với tính bất đối xứng dương khá mạnh.



Hình 5. Đồ thị hàm phân bố thống kê của tốc độ gió trên các mực: mặt đất, 850, 700, 500, 400, 300, 200, 150, 100, 70 và 50hPa.

Trên mặt đất, hiện tượng bất đối xứng trên đây liên quan chặt chẽ với điều kiện bất đồng nhất của mặt đệm. Sự bất đối xứng trên các tầng cao, theo chúng tôi, một

phân không nhỏ chịu sự ảnh hưởng từ hoạt động của các dòng chảy xiết trong đới gió đông nhiệt đới về mùa hè tại khu vực địa lý đang nghiên cứu.

### 3. Kết luận và kiến nghị

Từ kết quả tính toán và phân tích trên đây cho thấy hàm phân bố thống kê của các đại lượng khí tượng cao không trên lãnh thổ Việt Nam và khu vực địa lý lân cận rất khác nhau. Hàm phân bố của nhiệt độ không khí từ tầng thấp lên tầng cao đều khá gần với hàm phân bố chuẩn, trong khi đó, hàm phân bố của các yếu tố khác đều ít nhiều bị ảnh hưởng bởi tính bất đồng nhất. Trên mặt đất và các mực gần mặt đất trong các chuỗi số liệu đều tồn tại hiện tượng bất đồng nhất, đặc biệt các chuỗi số liệu về áp suất, độ ẩm không khí và tốc độ gió. Càng lên cao tính bất đồng nhất càng giảm dần. Trên các mực phía trên của tầng đối lưu xuất hiện trở lại hiện tượng bất đồng nhất nhẹ.

Kết quả trên phản ánh khá đầy đủ các đặc điểm nổi trội của điều kiện địa lý khí hậu khu vực nhiệt đới gió mùa đang nghiên cứu, đồng thời cũng phù hợp với quy luật vật lý cũng như các điều kiện địa lý khí hậu quy mô lớn.

Theo kết quả nghiên cứu này, khi xử lý số liệu cao không bằng các phương pháp thống kê cần quan tâm đến tính đặc thù khá nổi trội của chuỗi số liệu từng loại yếu tố khí tượng trên từng độ cao khác nhau. Trên các mực gần sát mặt đất, đối với hầu hết các đại lượng khí tượng đều cần phải quan tâm đến tính bất đồng nhất của chuỗi. Trong số đó, cần đặc biệt chú trọng đến chuỗi số liệu áp suất không khí, độ ẩm không khí và tốc độ gió.

Phương pháp cần sử dụng trong trường hợp có bất đồng nhất là phân chia chuỗi số liệu lớn thành các chuỗi nhỏ, trong đó tính đồng nhất được thoả mãn một cách tương đối, để xử lý theo phương thức riêng.

Trong phạm vi tầng đối lưu, càng lên cao các chuỗi số liệu càng có dạng gần với dạng phân bố chuẩn. Các chuỗi số liệu nhiệt độ không khí từ tầng thấp lên tầng cao đều gần với dạng phân bố chuẩn. Đối với các chuỗi số liệu này có thể mạnh dạn áp dụng các phương pháp xử lý phổ biến trong lý thuyết toán xác suất thống kê.

Đối với tất cả các đại lượng khí tượng, trong khi xử lý số liệu trên các mực phía trên của tầng đối lưu đều cần phải điều chỉnh giá trị các tham số cho phù hợp với đặc tính biến thiên của chúng.

### Tài liệu tham khảo

1. Nguyễn Đăng Quế, Hoàng Phương Hồng. Đặc điểm cấu trúc thống kê trường gió trên cao trên lãnh thổ Việt Nam và khu vực địa lý lân cận.- *Tuyển tập công trình Hội thảo khoa học "Cơ học thủy khí với thiên niên kỷ mới"*, Hà Nội, 2001.