

VẤN ĐỀ ĐÚNG SAI CỦA CÁC THÔNG TIN DỰ BÁO KTTV

TS. Đặng Trần Duy
Vụ Khoa học Kỹ thuật

Bàn về đánh giá một Bản tin dự báo KTTV, trước hết chúng ta lướt qua một số khái niệm liên quan.

Thời tiết bao gồm những hiện tượng tự nhiên xảy ra trong bầu trời và trên mặt đất, những hiện tượng đó biến đổi nhanh chóng và liên tục từ thời điểm này sang thời điểm khác từ địa điểm này sang địa điểm khác, tình hình thủy văn trên các dòng sông cũng vậy. Để ghi lại tình trạng phức tạp và biến đổi không ngừng đó của các hiện tượng KTTV người ta đã thiết lập một mạng lưới những trạm quan trắc cố định và lưu động phân bố trên khắp bề mặt trái đất, nói là quan trắc vì có yếu tố có thể đo đạc bằng máy, có yếu tố chỉ có thể quan sát rồi ghi chép lại. Những quan trắc đồng thời trên toàn mạng lưới cho phép dự báo viên thấy phân bố tình trạng KTTV theo không gian, những quan trắc liên tiếp tại những mốc thời gian khác nhau cho phép dự báo viên nắm được diễn biến tình trạng KTTV theo thời gian. Rõ ràng, tính chính xác của tình hình khí tượng thủy văn và những biến đổi của chúng trong không gian và theo thời gian thông qua những tài liệu quan trắc như vậy là tương đối, bởi lẽ người ta không biết những gì đã xảy ra trong khoảng không gian giữa hai trạm quan trắc và trong khoảng thời gian giữa hai kỳ quan trắc.

Những thông tin về tình hình KTTV nói chung và về dự báo KTTV nói riêng chủ yếu dưới dạng các bản tin mô phỏng bằng lời, có thể kèm theo một vài số liệu cụ thể, song dù viết tin dự báo dài đến mấy cũng không thể diễn đạt sát đúng tình hình KTTV xảy ra ở mọi lúc, mọi nơi, thậm chí nếu sử dụng tất cả những tài liệu quan trắc thì tình trạng như trên đã nói cũng không được cải thiện hơn là bao.

Để có thể mô tả tình trạng KTTV phức tạp được ngắn gọn và sát thực tế, các nhà khí tượng đã đưa ra một số quy định và một số cụm từ với ý nghĩa quy định, người đọc phải nắm được những quy định đó thì mới hiểu được đúng nội dung các thông tin dự báo cũng như tính đúng hoặc sai của nó, sau đây là một số quy định quan trọng nhất:

- Thông thường có nhiều hiện tượng KTTV xảy ra cùng một lúc nhưng trong các thông tin KTTV chỉ nhắc đến 1-2 hiện tượng quan trọng theo thứ tự ưu tiên đã được xếp đặt trước, ví dụ khi trên khu vực đồng thời có nhiều loại mưa cùng xảy ra như mưa dông, mưa rào, mưa thường hoặc mưa phun thì bản tin chỉ nhắc đến mưa dông, loại mưa được khoa học khí tượng xếp vào loại quan trọng nhất, khi xảy ra mưa ở nhiều nơi, chỗ to, chỗ nhỏ thì hiện tượng mưa to sẽ được ưu tiên nhắc đến trong bản tin.
- Để diễn tả mức độ hiện tượng xảy ra trên một khu vực địa lý, các thông tin dự báo khí tượng sử dụng các thuật ngữ "vài nơi", "rải rác" và "nhiều nơi" với những ý nghĩa quy định như sau:
 - "Vài nơi" chỉ hiện tượng có khả năng xảy ra ở $\frac{1}{3}$ số địa điểm có trạm quan trắc trong khu vực.
 - "Rải rác" chỉ hiện tượng có khả năng xảy ra ở $\frac{1}{2}$ số địa điểm có trạm quan trắc trong khu vực.

- "Nhiều nơi" chỉ hiện tượng có khả năng xảy ra ở 2/3 số địa điểm có trạm quan trắc trong khu vực trở lên. Trong những thông tin dự báo, tính ngữ này thường không được nhắc đến ví như cụm từ "có mưa ở nhiều nơi" được viết gọn hơn thành "có mưa".

Ý nghĩa các thuật ngữ "vài nơi, rải rác, nhiều nơi" được sử dụng chủ yếu trong các thông tin dự báo cho một khu vực, tuy nhiên đối với một địa điểm, có thể hiểu đó là khả năng xảy ra hiện tượng thời tiết ở địa điểm đó tương ứng với các giá trị 30, 50 và 70%.

Qua những điều trình bày trên ta thấy những thông tin KTTV không thể diễn tả đầy đủ, chính xác hiện trạng thời tiết cũng như sự biến đổi liên tục của chúng trong không gian và theo thời gian. Sự nắm bắt, ghi nhận và mô tả tình trạng thời tiết thực tế đã bị giảm ước đi ở nhiều khâu, vì vậy người tiếp nhận thông tin dự báo phải nắm được sự giảm ước đó mới có thể cảm nhận được sự đúng sai của nó. Rất may mắn là đối với những hệ thống thời tiết quy mô lớn, cường độ mạnh, đặc biệt là các hệ thống thời tiết nguy hiểm như bão, gió mùa, mưa lớn diện rộng là những hệ thống tiềm ẩn thiên tai thì số liệu của mạng lưới quan trắc cũng như các thuật ngữ dự báo nói chung đã phản ánh được tương đối chính xác tình hình thực tế, sự sai khác đáng kể chỉ xảy ra trong những hệ thống thời tiết cường độ yếu, không nguy hiểm.

Mỗi hiện tượng thời tiết là kết quả tác động của nhiều nguyên nhân khác nhau, thậm chí có cả nguyên nhân ở đâu đó rất xa trong vũ trụ mà khoa học còn chưa khám phá được, những nguyên nhân đó lại biến đổi không ngừng cả theo không gian và thời gian, vì vậy quan hệ giữa hiện tượng thời tiết và những nguyên nhân hình thành nó không phải là quan hệ nhất-nhất mà là quan hệ thống kê, trong quan hệ này không gian nguyên nhân không có ranh giới rõ rệt giữa hai vùng, vùng có (100%) và vùng không có (0%) khả năng hình thành hiện tượng, mỗi vùng trong không gian nguyên nhân chỉ tương ứng với một giá trị xác suất hình thành hiện tượng nhất định, giá trị đó phân bố trong khoảng từ 0 đến 100%. Lộ trình của mọi phương pháp dự báo KTTV được tiến hành theo hướng ngược lại, xuất phát từ một tập hợp các giá trị ban đầu của các yếu tố được xem là nguyên nhân hình thành hiện tượng, phương pháp sẽ cho biết khả năng hình thành hiện tượng là bao nhiêu, nói chung các phương pháp sau thời gian thử nghiệm cân thiết, có 70% số lần đúng thực tế thì phương pháp đó là tốt và được đưa vào sử dụng trong nghiệp vụ dự báo. Vì vậy *tin dự báo về một hiện tượng KTTV nào đó không phải là thông tin về hiện tượng đó sẽ xảy ra mà là thông tin về khả năng hiện tượng đó sẽ xảy ra, nếu như khả năng đó là 70% có nghĩa là còn 30% khả năng hiện tượng không xảy ra*. Như vậy, quy luật nhân quả trong việc hình thành các hiện tượng KTTV mang bản chất thống kê, các phương pháp dự báo KTTV, dù đó là phương pháp tốt nhất cũng phải mang bản chất đó. Vì vậy trong dự báo là "có" hiện tượng xảy ra đã mang sẵn một phần cái "không" từ trong bản chất của quy luật hình cũng như trong bản chất của phương pháp dự báo chúng. **Dự báo đúng, dự báo sai đều mang tính tất yếu khách quan**, chúng ta chỉ có thể tăng thêm số lần dự báo đúng, giảm dần số lần dự báo sai nhưng không thể làm đúng tất cả những lần dự báo.

Tuy nhiên, việc tăng số lần dự báo đúng trong mấy chục năm qua tỏ ra rất chậm chạp, ngay cả ở những nước phát triển, có nhiều nguyên nhân dẫn đến tình trạng này, hai trong số đó, thứ nhất là, chúng ta còn chưa hiểu biết đầy đủ về những

quá trình xảy ra trong khí quyển và trên các đại dương bởi lẽ còn quá ít số liệu do đặc khảo sát về chúng, nhất là số liệu trên các đại dương mènh mông, ở hai cực trái đất quanh năm băng phủ, trên các tầng cao khí quyển cũng như bên trong các hình thái khí quyển khốc liệt (bão, tố, lốc, mây dông...), thiếu số liệu tức không nắm được bản chất, không phân tích đúng nguyên nhân hình thành và phát triển do đó không thể nói đến dự báo chúng tốt được. Thứ hai là, người ta chưa xây dựng được lý thuyết chuyển động của không khí vùng nhiệt đới đặc biệt là vùng nhiệt đới vĩ độ thấp, do đó chưa có cơ sở để xây dựng lý thuyết và công nghệ dự báo thời tiết hoàn chỉnh cho khu vực này.

Tiêu chuẩn để đánh giá dự báo KTTV cũng mang tính chất tương đối, phụ thuộc vào máy móc thiết bị, trình độ công nghệ dự báo của thế giới nói chung và của mỗi quốc gia nói riêng, khái quát mà nói, tiêu chuẩn đánh giá đúng sai một bản tin dự báo của chuyên môn còn khoảng cách so với quan niệm đúng sai xuất phát từ nhu cầu của cộng đồng.

Trong nhiều năm gần đây, việc đánh giá đúng sai của một tin dự báo chỉ được thực hiện đối với những thông tin định tính dựa trên sự mô tả bằng lời như đã trình bày trên, những giới hạn đúng sai được quy định để đánh giá nói chung đương nhiên phải được mở rộng thêm so với ý nghĩa đã quy định cho các thuật ngữ dự báo, ví dụ dự báo "có mưa một vài nơi" thực tế hoàn toàn không có mưa đến mưa ở dưới 1/2 số trạm có báo cáo quan trắc được xem là đúng, dự báo "có mưa ở nhiều nơi" thực tế chỉ có mưa ở dưới 1/2 số trạm mới bị xem là sai. Những giới hạn quy định này có thể khác nhau giữa các quốc gia tuỳ thuộc trình độ dự báo và tầm quan trọng của yếu tố mà họ quan tâm. Khi chưa có công nghệ dự báo vị trí tâm bão tin cậy thì tầm quan trọng của việc đánh giá yếu tố này phải đứng sau các yếu tố nguy hiểm do bão gây ra như thời gian bắt đầu có gió mạnh và tốc độ gió mạnh nhất, mức độ mưa và mức nước biển dâng đối với các khu vực chịu ảnh hưởng của bão.

Sự quy định những chuẩn mực đánh giá ở đây hoàn toàn không phải vì mục đích nâng cao thành tích dự báo mà xuất phát từ những nguyên nhân và yêu cầu khoa học khách quan, những người làm công tác dự báo luôn luôn phấn đấu để nâng cao mức chính xác và hơn nữa nâng cao hiệu quả phục vụ công đồng, không chấp nhận sự dễ dãi với công việc của mình, một phương pháp dự báo mới phải có cơ sở khoa học, phải qua thử nghiệm thời gian dài đạt mức chính xác từ 70% trở lên và phải có "độ tin cậy" lớn hơn 0,3 mới được chấp nhận đưa vào sử dụng dự báo. Xin nêu ra đây một ví dụ đơn giản, ở những thời kỳ, tần suất xảy ra một hiện tượng khí tượng nào đó tương đối lớn, có thể đạt đến 70% số ngày hoặc hơn thế, vậy trong thời kỳ này ngày nào cũng dự báo hiện tượng đó xảy ra thì các dự báo có thể đúng đến trên 70%, số liệu khí hậu là một cơ sở để dự báo, nhưng những dự báo chỉ dựa vào số liệu khí hậu có "độ tin cậy" bằng 0 vì điều kiện (xin phép không trình bày cụ thể cách tính thông số này) nên phương pháp dự báo kiểu này không được chấp nhận.

Đối với các nước phát triển, nói chung công nghệ dự báo của họ đều đã được số hoá kể cả phương pháp synoptic truyền thống vẫn được xem là phương pháp định tính, nội dung các thông tin dự báo của họ nói chung là những thông tin định lượng, vì vậy thay vì đánh giá dự báo theo kiểu đúng-sai cổ điển, họ tiến hành đánh giá sai số các thông tin dự báo đã phát ra mà đơn giản là độ chênh lệch giữa trị số dự báo và trị số thực tế quan sát được, hướng phấn đấu để nâng cao chất lượng,

hiệu quả của dự báo là phần đầu để giảm dần độ chênh lệch này, việc đánh giá theo kiểu đúng-sai ở những nước này hoặc không còn sử dụng hoặc chỉ sử dụng đối với những yếu tố dự báo không đo đạc được. Bảng 1 trình bày sai số quân phương trung bình (m/s) tốc độ gió dự báo ở độ cao 1500 mét vùng nhiệt đới của Trung tâm dự báo thời tiết hạn vừa Châu Âu năm 1997 (ECMWF).

Bảng 1

Tháng	1	2	3	4	5	6	7	8	9	10	11	12	TB
Dự báo 24 giờ	2,8	2,9	3,0	3,0	2,6	2,4	2,4	2,5	2,5	2,5	2,5	2,5	2,6
Dự báo 48 giờ	3,7	3,7	3,8	3,7	3,3	3,2	3,2	3,4	3,4	3,4	3,2	3,2	3,4
Dự báo 72 giờ	4,2	4,2	4,2	4,0	3,8	3,6	3,7	4,0	3,9	3,8	3,6	3,6	3,9
Dự báo 96 giờ	4,6	4,5	4,6	4,3	4,1	4,0	4,0	4,4	4,3	4,2	3,9	4,0	4,2
Dự báo 96 giờ (quán tính)	6,0	5,9	6,3	5,8	5,6	5,5	5,5	6,1	5,7	5,7	5,4	5,5	5,8

Ô cuối cùng của hàng và cột cuối cùng trong bảng 1 biểu thị sai số quân phương trung bình năm 1997 về dự báo số trị tốc độ gió ở mực 1500 mét (với những tổn kém rất lớn) của ECMWF cho vùng nhiệt đới là 4,2 m/s, so với sai số của phương pháp quán tính (đơn giản, ít tổn kém) là 5,8 m/s, như vậy sai số dự báo số trị mới giảm được 27% sai số so với phương pháp quán tính.

Trong các tin dự báo bão của Trung tâm bão Tokyo (RSMC), ở phần dự báo vị trí trung tâm bão với các thời hạn khác nhau thường có hai thông số, một là vị trí (theo kinh vĩ độ địa lý) trung tâm bão hoặc chính xác hơn là vị trí tâm vòng tròn có 60% khả năng tâm bão dự báo nằm phía trong nó, thứ hai là bán kính (tính bằng km) của vòng tròn khả năng này (từ sau 1 tháng 7 năm 1997 mốc vòng tròn khả năng quy định tăng thêm 10% là 70%), bán kính vòng tròn khả năng thường được tính lại sau một số lần dự báo quy định, kết quả tính cho giai đoạn trước được dùng đưa vào tin dự báo những cơn bão tiếp theo. Bán kính vòng tròn khả năng 60% sau 113 lần dự báo vị trí tâm bão 24 giờ trong mùa bão 1997 trước ngày 30 tháng 6 là 158 km, của 470 lần dự báo sau mồng 1 tháng 7 là 180 km, bán kính những vòng tròn tương ứng của dự báo vị trí tâm bão trước 48 giờ năm 1997 là 316 và 362 km, bạn thử đặt những vòng tròn này lên vùng ven biển nước ta sẽ thấy nó bao phủ một vùng rộng lớn như thế nào, thế mà cũng chỉ có 60-70% khả năng tâm bão dự báo nằm trong nó, còn 30-40% khả năng là nằm ngoài, đây là những kết quả đánh giá dự báo bão của người Nhật.

Trong đánh giá dự báo vị trí tâm bão, người Nhật thường sử dụng 3 thông số, thứ nhất là sai số về khoảng cách (D) được đo bằng khoảng cách giữa hai vị trí tâm bão thực tế và dự báo, thứ hai là giá trị quân phương trung bình (SD) của những khoảng cách sai số đó và thứ ba là tỷ lệ phần trăm giữa sai số khoảng cách của các mô hình dự báo nghiệp vụ (EO) và của dự báo theo phương pháp quán tính (OP) ký hiệu là EO/EP%. Bảng 2 trình bày những thông số đó theo số liệu tổng kết chính thức năm 1997 của Trung tâm bão Tokyo.

Bảng 2

Bão	Dự báo 24 giờ				Dự báo 48 giờ			
	Số DB	D km	SD km	EO/EP %	Số DB	D km	SD km	EO/EP %
9701 ISA	37	124	86	67	35	254	118	54
9702 JIMY	5	338	209	92	1	212	0	
9703 KELLY	5	233	86	67	1	365	0	
9704 LEVI	2	265	98		0			
9705 MARIE	11	221	58	99	7	615	169	128
9706 NESTOR	22	124	75	59	18	216	97	44
9707 OPAL	16	120	57	45	12	332	127	58
9708 PETER	15	193	79	54	11	505	201	69
9709 ROSIE	28	123	42	63	24	154	56	38
9710 SCOTT	18	307	141	125	13	682	299	136
9711 TINA	33	123	61	79	29	213	96	63
9712 VICTOR	6	105	52	69	2	318	50	
9713 WINNIE	38	104	51	82	34	182	89	69
9714 YULE	11	237	70	70	7	537	183	80
9715 ZITA	5	269	79	225	1	306	0	
9716 AMBER	28	103	53	109	24	182	94	80
9717 CASS	1	512	0		0			
9718 BING	21	102	52	29	17	185	90	25
9719 OLIWA	48	129	93	68	44	259	157	57
9720 DAVID	24	167	84	60	20	244	124	42
9721 FRITZ	6	126	39	134	2	442	59	
9722 GINGER	20	194	80	50	16	462	215	60
9723 IVAN	38	178	80	104	31	354	136	79
9724 JOAN	35	104	57	50	30	262	125	45
9725 KEITH	35	80	57	44	31	96	64	19
9726 LINDA	5	236	25	80	1	517	0	91
9727 MORT	15	107	74	51	11	148	78	36
9728 PAKA	55	146	65	105	51	268	131	80
Trung bình năm		144km	92km	70%		267km	181km	58%

Qua bảng 2 ta thấy sai số về khoảng cách dự báo vị trí tâm bão trung bình năm của Nhật là 144 km cho thời hạn dự báo 24 giờ, là 267 km cho thời hạn 48 giờ, tương ứng rút ngắn được 30 và 42% so với sai số của phương pháp dự báo quán tính, sai số về khoảng cách khác nhau rất lớn giữa các cơn bão, lớn hơn 200 km cho dự báo 24 giờ chiếm đến 1/3 số cơn, lớn nhất đến 307 km (không kể giá trị 512 vì chỉ phát có 1 bản tin), nhỏ nhất là 80 km, còn đến 8 cơn bão kết quả dự báo nghiệp vụ 24 giờ tương hoặc kém thua kết quả dự báo theo phương pháp quán tính ($EO/EP \approx$ hoặc $> 100\%$).

Kết quả đánh giá dự báo của hai Trung tâm dự báo hiện đại nhất thế giới hiện nay trình bày trên để minh họa cụ thể việc đánh giá của các trung tâm dự báo

hiện đại, đồng thời cho thấy chất lượng dự báo của họ còn khoảng cách lớn so với yêu cầu phục vụ, đã tốt hơn nhưng không phải đã tốt hơn nhiều so với phương pháp dự báo quán tính và so với kết quả dự báo ở những trung tâm dự báo còn trong tình trạng lạc hậu.

Một ý tưởng mới để đánh giá đúng-sai đối với các thông tin dự báo KTTV là phương pháp thăm dò ý kiến đánh giá của người tiếp nhận và sử dụng thông tin, đã có những kết quả thử nghiệm thú vị theo hướng này, một công ty dịch vụ thời tiết ở Mỹ đã cung cấp dự báo thời tiết suốt 3 tháng mùa cưới trong 2 năm 1998, 1999, những cặp nam nữ đã hứa hôn ở 10 thành phố thuộc nước Mỹ căn cứ vào đó chọn một thời kỳ được dự báo có thời tiết tốt để tổ chức hôn lễ của mình, năm 1998 có 77% và năm 1999 có 83% các cặp uyên ương cho biết hoàn toàn vừa ý với sự lựa chọn ngày hôn lễ của mình theo dự báo thời tiết được cung cấp trước, công ty dịch vụ thời tiết này đã xem kết quả thăm dò đó là kết quả dự báo của họ.

Như vậy, đã có ba phương pháp để đánh giá các thông tin dự báo thời tiết, thứ nhất là phương pháp đánh giá cho kết quả đúng-sai trên cơ sở những quy định ước lệ, chủ yếu để đánh giá những thông tin dự báo nặng về mô phỏng, thứ hai là phương pháp đánh giá bằng sai số tuyệt đối của những yếu tố dự báo, chủ yếu để đánh giá kết quả dự báo số trị, thứ ba là phương pháp đánh giá thông qua thăm dò ý kiến của những người tiếp nhận và sử dụng thông tin dự báo, mỗi phương pháp đều có cơ sở khoa học, có điểm mạnh, điểm yếu riêng, tùy điều kiện mà áp dụng song đều là những thông tin quý giá để những người công tác trong ngành KTTV noi chung và những dự báo viên nói riêng làm căn cứ phấn đấu nâng cao chất lượng và hiệu quả phục vụ của mình.

(tiếp theo trang 4)

Sau năm 1990, gắn với các mục tiêu phát triển kinh tế - xã hội do Đại hội VII và Đại hội VIII của Đảng đề ra, hầu hết các đề tài đều tập trung vào những vấn đề cấp bách do thực tiễn của sản xuất và đời sống đặt ra. Cùng với việc thực hiện 12 đề tài thuộc 5 chương trình trọng điểm cấp Nhà nước trong các lĩnh vực: Đánh giá hệ quả biến đổi khí hậu; Điều kiện KTTV biển; Bảo vệ và sử dụng tài nguyên nước; Đánh giá hiện trạng và biện pháp bảo vệ môi trường không khí và nước; Các hiểm họa thiên tai và biện pháp phòng tránh,... Tổng cục đã tiến hành 96 đề tài cấp ngành thuộc 6 chương trình: Phòng chống thiên tai; Biến đổi khí hậu; Thủy văn và tài nguyên nước; KTTV biển; Môi trường và chương trình KTTV phục vụ kinh tế - xã hội miền núi. Kết quả nghiên cứu toàn bộ hoặc từng phần của nhiều đề tài nói trên đã và đang được đưa vào phục vụ sản xuất và đời sống. Hàng chục báo cáo khoa học được đúc rút từ các đề tài này đã trình bày và được đánh giá cao tại các Hội thảo khoa học quốc tế và quốc gia về: Biến đổi khí hậu; Các chất phá hủy tầng ô-dôn; Khai thác và quản lý hồ chứa; Quản lý tổng hợp tài nguyên nước, cấp nước và kiểm soát nhiễm bẩn nước,....

Chuẩn bị bước vào thiên niên kỷ mới, tuy còn nhiều khó khăn thử thách, song cũng có nhiều thuận lợi rất cơ bản Đảng và Nhà nước ngày càng quan tâm hơn đối với các hoạt động KTTV, đội ngũ cán bộ CNVC trong Ngành có bản lĩnh chính trị, có kiến thức và khả năng tiếp thu, ứng dụng và phát triển các phương pháp và công nghệ hiện đại, đó là cơ sở chắc chắn để Ngành KTTV tiếp tục vươn lên đóng góp ngày càng có hiệu quả hơn cho sự nghiệp công nghiệp hoá, hiện đại hoá đất nước.