

XÁC ĐỊNH KÍCH THƯỚC MẪU NGHIÊN CỨU THÍCH HỢP CHO MỘT SỐ CHỈ TIÊU CỦA CÁC THÍ NGHIỆM TRỒNG LÚA

Determining suitable sample size for some characteristics in wet rice experiments

Phạm Tiến Dũng

Khoa Nông học, Trường Đại học Nông nghiệp Hà Nội

TÓM TẮT

Kết quả nghiên cứu được thực hiện trên cơ sở tổng hợp kết quả từ ba thí nghiệm: phân bón, mật độ và so sánh giống được thực hiện tại Hà Tây, Hà Nội trong hai năm 2004 và 2005. Có 4 dòng giống tham gia thí nghiệm là N46, Nếp 87-2, Bắc Thơm số 7 (BTS7) và LT2. Dung lượng mẫu phụ thuộc biến động khác nhau của các chỉ tiêu, điều kiện nghiên cứu nên kết quả được tính trung bình từ ba thí nghiệm trên. Theo dõi 30 cá thể cho mỗi đặc trưng nghiên cứu được tiến hành theo phương pháp hoàn toàn ngẫu nhiên, sau đó tính biến động của chúng để dùng công thức tính ra dung lượng mẫu phù hợp với các loại độ tin cậy và các mức sai số thường chấp nhận khác nhau. Kết quả nghiên cứu đã cho thấy, với độ tin cậy thường áp dụng (95%) và sai số thường chấp nhận (10%) thì kích thước mẫu nên theo dõi là 3 khóm cho đo chiều cao cây, 30 khóm cho theo dõi số dảnh, diện tích lá và khối lượng chất khô, 3 bông cho đo chiều dài cổ bông và lấy 10 bông cho đếm số hạt trên bông.

Từ khoá: Hệ số biến động, kích thước (dung lượng) mẫu, mức ý nghĩa, sai số thực tế, sai số chấp nhận.

SUMMARY

This article was written based on of three field rice experiments conducted at Gialam, Hanoi and Hatay Province in two years, 2004 and 2005 with 4 different rice genotypes, viz. Nếp 87-2, Bac Thom No7 (BTS7), LT2, and N46. Sample size depends on the nature and variability of characteristics under investigation and research conditions. 30 random individual measurements were taken on each characteristic to estimate the coefficient variation that in turn was used to calculate the sample size with different significance levels and acceptable standard errors. The results showed that with the common significance level ($P=0.05$) and standard error of 10%, the sampling size should be 3 hills for measuring of plant height, 30 hills for tiller number, leaf area and dry matter, 3 panicles for measuring panicle neck length and 10 panicles for counting seed numbers per panicle.

Key words: Coefficient of variation, sample size, significance level, standard error.

1. ĐẶT VẤN ĐỀ

Trong nghiên cứu cây trồng nói chung hay nghiên cứu về cây lúa nói riêng, các nhà nghiên cứu thường phải lấy mẫu để đo đếm, theo dõi các chỉ tiêu cần nghiên cứu như các chỉ tiêu chiều cao cây, diện tích lá, số nhánh, cành, các yếu tố cấu thành năng suất, ... Câu hỏi luôn được đặt ra là lấy bao nhiêu cá thể (dung lượng mẫu bằng bao nhiêu) để theo dõi sẽ đủ đại diện cho đám đông cần nghiên cứu, về vấn đề này còn rất ít được quan tâm. Gomez và Gomez (1984) đã đề

xuất với lúa cấy: chiều cao cây nên đo ở 3 khóm, số nhánh đẻ nên đo ở 12 khóm ứng với sai số tiêu chuẩn là 7,1%, số hạt trên bông nên đo ở 12 khóm. Gần đây, Nguyễn Thị Lan (2003, 2005) đã nghiên cứu về dung lượng mẫu cho lúa, đậu tương ở Việt Nam, nhưng kết quả còn kém thuyết phục do chỉ một lần thí nghiệm. Nguyễn Văn Tạo (1998) bằng kinh nghiệm nghiên cứu đã đề xuất một số chỉ tiêu nghiên cứu trên cây chè với dung lượng mẫu khác nhau nhưng chưa có cơ sở khoa học chắc chắn nên độ chính xác chưa được đảm bảo.

Vì vậy, nghiên cứu này được tiến hành với mục tiêu: xây dựng cơ sở cho việc xác định dung lượng mẫu và xác định dung lượng mẫu cụ thể cho một số chỉ tiêu của thí nghiệm trồng lúa nhằm giúp các nhà nghiên cứu có tài liệu tham khảo làm căn cứ cho tiến hành thí nghiệm của mình để đạt độ chính xác cao hơn, tiết kiệm được thời gian, tiền của và nhân lực.

2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Nghiên cứu được thực hiện trên 3 loại thí nghiệm là mật độ, bón phân cho dòng lúa N46 đã thuần (đang được khu vực hoá, được sản xuất chấp nhận rộng rãi, tại khu thí nghiệm khoa Nông học, trường Đại học Nông nghiệp Hà Nội) và thí nghiệm so sánh một số giống lúa chất lượng cao như: Nếp 87-2, Bắc Thơm số 7 (BTS7), LT2 là những dòng giống đã thuần tại Đại Đồng, Thạch Thất, Hà Tây năm 2006.

Các thí nghiệm đều được lặp lại 3 lần, bố trí kiểu khối ngẫu nhiên đầy đủ. Trên mỗi ô thí nghiệm có diện tích 20 m², theo dõi ngẫu nhiên theo sơ đồ 30 cá thể (coi như dung lượng mẫu lớn nhất cho đám đồng thuần nhất) để chứng thể hiện đúng độ biến động của các chỉ tiêu theo dõi như chiều cao cây, số nhánh đê, chất khô tích lũy,... Trên cơ sở của biến động, tính ra dung lượng mẫu cần thiết cho các chỉ tiêu.

Các công thức được sử dụng trong báo cáo:

$$n = \frac{t_{\alpha}^2 \cdot S^2}{(\bar{X})^2 \cdot (\Delta\%)^2} \cdot 10^4$$

Trong đó:

t_{α} - giá trị t lý thuyết của bảng phân bố t với mức xác suất nhỏ α và độ tự do bằng số mẫu trừ đi 1.

S^2 - phương sai mẫu

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}$$

\bar{X} - trung bình mẫu

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

$\Delta\%$ - sai số cho phép được chấp nhận.

Ngoài ra, nghiên cứu còn sử dụng chỉ tiêu hệ số biến động mẫu CV%:

$$CV\% = \frac{S}{\bar{X}} \cdot 100$$

Sau khi tính ra dung lượng mẫu cho mỗi chỉ tiêu của một thí nghiệm, tiếp tục tính trung bình cho mỗi chỉ tiêu đó qua kết quả tính từ mỗi thí nghiệm để đi đến kết luận cuối cùng cho mỗi chỉ tiêu.

3. KẾT QUẢ NGHIÊN CỨU

3.1. Với thí nghiệm bón phân đạm cho dòng lúa N46

3.1.1. Dung lượng mẫu cần thiết cho chỉ tiêu chỉ số diện tích lá

Theo Phạm Chí Thành (1986), dung lượng mẫu không chỉ phụ thuộc vào loại thí nghiệm mà còn phụ thuộc vào từng chỉ tiêu theo dõi. Bởi vì các chỉ tiêu theo dõi khác nhau thường có biến động khác nhau nên dung lượng mẫu theo dõi cũng phải khác nhau.

Từ kết quả theo dõi thí nghiệm phân bón cho dòng N46, chúng tôi đã thu được một số kết quả về dung lượng mẫu cần thiết cho một số chỉ tiêu theo dõi: Chiều cao cây, chỉ số diện tích lá, trọng lượng chất khô, tỷ lệ nhánh hữu hiệu, chiều dài cỏ bông, tổng số hạt/bông, tỷ lệ hạt chắc (Bảng 1).

Qua các công thức bón phân khác nhau, biến động của chỉ tiêu chỉ số diện tích lá không nhiều, vì từ biến động tính ra sai số thực ở độ tin cậy 95% cho thấy tất cả các trường hợp đều đạt sai số ở mức chấp nhận được theo quy định chung.

Từ những biến động thu được, căn cứ mức tin cậy khác nhau, sai số chấp nhận khác nhau để tính ra dung lượng mẫu cần thiết cho kết quả rất rõ là khi yêu cầu độ tin cậy cao hơn với cùng mức sai số chấp nhận thì dung lượng mẫu cần lấy là lớn hơn. Khi sai số chấp nhận nhỏ hơn với cùng độ tin cậy thì dung lượng mẫu cần lấy cũng lớn hơn.

Kết quả nhận được khẳng định cho lý thuyết đã đề cập trên đây là khi mỗi chỉ tiêu biến động khác nhau thì dung lượng mẫu lấy sẽ khác nhau. Chính vì vậy mà phải có các nghiên cứu cụ thể trong điều kiện của Việt Nam để rút ra kết luận về dung lượng mẫu cần lấy sao cho phù hợp.

Tính trung bình cho cả thí nghiệm, dung lượng mẫu cần thiết cho chỉ tiêu chỉ số diện tích lá trong thí nghiệm bón phân cho lúa thơm N4 ở độ tin cậy 95 % và sai số chấp nhận 5 %, 10 % là 109 khóm và 27 khóm. Ở độ tin cậy 99 % sai số chấp nhận 5 %, 10 % là 204 khóm và 51 khóm.

Bảng 1. Dung lượng mẫu cần cho quan sát chỉ số diện tích lá

Công thức	Sai số thực tế %	CV%	Dung lượng mẫu			
			Độ tin cậy 95%		Độ tin cậy 99%	
			Sai số chấp nhận 5%	Sai số chấp nhận 10%	Sai số chấp nhận 5%	Sai số chấp nhận 10%
N1	10,6	28,4	135	34	253	63
N2	9,3	25,0	105	26	196	49
N3	9,3	24,8	103	26	193	48
N4	8,9	23,8	95	24	178	44
TB	9,5	25,5	109	27	204	51

Ghi chú: N1: Lượng đạm bón 60 kg N/ha; N2: Lượng đạm bón 80 kg N/ha; N3: Lượng đạm bón 100 kg N/ha; N4: Lượng đạm bón 60 kg N/ha.

3.1.2. Dung lượng mẫu cần thiết cho các chỉ tiêu còn lại

Bảng cách tính tương tự cho chỉ tiêu diện tích lá, kết quả tính dung lượng mẫu cho các chỉ tiêu còn lại được thể hiện trên bảng 2.

Bảng 2. Dung lượng mẫu cần cho quan sát một số chỉ tiêu khác

Chỉ tiêu	Sai số thực tế %	CV%	Dung lượng mẫu			
			Độ tin cậy 95%		Độ tin cậy 99%	
			Sai số chấp nhận 5%	Sai số chấp nhận 10%	Sai số chấp nhận 5%	Sai số chấp nhận 10%
Tích lũy chất khô	8,9	23,9	96	24	180	45
Chiều cao cuối cùng	2,3	6,2	6	2	12	3
Số nhánh hữu hiệu	10,5	28,1	132	33	248	62
Chiều dài cổ bông	2,5	6,7	8	2	14	4
Tổng số hạt trên bông	7,2	19,2	62	15	115	29
Số hạt chắc/bông	4,2	11,3	21	5	40	10

Trong cùng thí nghiệm, các chỉ tiêu theo dõi khác nhau có mức độ biến động khác nhau. Do có sự biến động khác nhau của các chỉ tiêu theo dõi nên dung lượng mẫu tính ra cần thiết cho mỗi chỉ tiêu cũng khác nhau, khi cùng độ tin cậy và độ chính xác. Trong thí nghiệm này, chỉ tiêu tích lũy chất khô có biến động tới 23,9% nên dung lượng mẫu tại độ tin cậy 95% và sai số chấp nhận 10% là 24 còn chấp nhận sai số 5% là 96. Trong khi đó chỉ tiêu chiều cao cây cuối cùng chỉ có biến động 6,2% nên cùng độ tin cậy 95% và sai số chấp nhận 10% thì số mẫu chỉ cần có 2 cây (Bảng 2). Nhưng thực tế các nhà nghiên cứu hầu hết đo chiều cao cây trên 5 cây và 10 cây.

3.2. Dung lượng mẫu cho các chỉ tiêu của thí nghiệm mật độ cấy

Tiến hành tương tự như thí nghiệm bón phân, trong thí nghiệm mật độ cấy các kết quả tính dung lượng mẫu cho các chỉ tiêu theo dõi được thể hiện trên bảng 3.

Số liệu trong bảng 3 cho thấy, chỉ tiêu chiều cao cây luôn có độ biến động nhỏ nhất nên dung lượng mẫu tính ra trong các trường hợp độ tin cậy và độ chính xác khác nhau đều có dung lượng mẫu nhỏ hơn so với các chỉ tiêu khác. Ví dụ ở độ tin 95% với sai số chấp nhận 10% và 5% hoặc độ tin 99% với sai số chấp nhận 10% và 5% tương ứng dung lượng mẫu là

1, 3, 2 và 6. Trong khi đó cùng độ tin cậy và sai số tương ứng của chỉ tiêu số hạt trên bông do biến động lớn hơn nên dung lượng là 8, 30, 14 và 55. Còn chỉ tiêu tích lũy chất khô có biến động lớn nhất nên dung lượng tương ứng là 26, 103, 47 và 188.

Bảng 3. Dung lượng mẫu cần thiết cho theo dõi một số chỉ tiêu của thí nghiệm mật độ cấy

Chỉ tiêu	CV%	Độ tin cậy P = 95%		Độ tin cậy P = 99%	
		Sai số 10%	Sai số 5%	Sai số 10%	Sai số 5%
Cao cây cuối cùng	4,35	1	3	2	6
S.nhánh/khóm thời kỳ trổ	22,40	22	86	39	156
Số bông/khóm	21,03	19	76	34	138
Chỉ số diện tích lá	22,07	21	84	38	152
Tích lũy chất khô	24,35	26	103	47	188
Số hạt/bông	13,42	8	30	14	55
Số hạt chắc /bông	18,10	14	55	25	100

3.3. Dung lượng mẫu cần thiết cho một số chỉ tiêu trong thí nghiệm so sánh giống

Trong thí nghiệm so sánh giống, để thấy rõ hơn khi theo dõi trên các giống khác nhau, mức độ biến động của mỗi chỉ tiêu theo dõi có khác nhau nhiều không, kết quả được trình bày riêng cho từng giống qua các lần nhắc lại (Bảng 4).

Cùng chỉ tiêu theo dõi, trên các giống khác nhau có biến động khác nhau không nhiều nên

khi tính ra dung lượng mẫu cần thiết có chênh lệch không đáng kể.

Chỉ tiêu chiều cao cây giữa 3 giống có biến động chỉ dao động từ 3,90 đến 5,24%, do vậy dung lượng mẫu tính ra là 1 (độ tin cậy 95% và sai số 10%) hoặc là từ 1 đến 2 (độ tin cậy 99% và sai số 10%). Nếu chấp nhận sai số nhỏ hơn (5%) thì cũng chỉ là 3 đến 5 hoặc từ 5 đến 9 (tương ứng).

Bảng 4. Dung lượng mẫu theo dõi cần thiết cho các chỉ tiêu

Giống	Chỉ tiêu theo dõi	CV%	Dung lượng mẫu			
			Độ tin 95%		Độ tin 99%	
			Sai số 10%	Sai số 5%	Sai số 10%	Sai số 5%
Nếp 87-2	Cao cây	5,24	1	5	2	9
	Số nhánh	28,99	37	148	67	268
	Tích lũy chất khô	31,35	41	166	75	301
	Đếm số lá	28,99	37	148	67	268
LT2	Cao cây	4,24	1	3	1	5
	Số nhánh	27,82	32	131	59	236
	Tích lũy chất khô	25,07	26	106	48	192
	Đếm số lá	27,82	32	131	59	236
BTS7	Cao cây	3,90	1	3	1	5
	Số nhánh	29,87	38	152	69	276
	Tích lũy chất khô	25,78	28	113	51	204
	Đếm số lá	26,22	29	116	53	210

3.4. Tổng hợp kết quả nghiên cứu của các thí nghiệm

Chỉ tiêu chiều cao cây có độ tin cậy thông thường trong nghiên cứu nông nghiệp là 95% với sai số thường được chấp nhận cho thí nghiệm ngoài đồng là 10%, từ biến động thực tế của chúng, dung lượng mẫu được tính có biến động từ 1 - 2 cây cho các thí nghiệm và tính trung bình là 1 (Bảng 5). Điều này có nghĩa là khi theo dõi chỉ tiêu chiều cao cây chỉ cần theo dõi trên mỗi ô 1 cây đã đủ sức đại diện cho cả ô. Tuy nhiên, không nên theo dõi như vậy mà ít nhất cũng nên theo dõi 3 cây để tính trung bình sẽ tốt hơn.

Tương tự cách nhìn nhận và phân tích trên đây, dung lượng mẫu cho các chỉ tiêu khác lần lượt tính được: 23 - 27 và trung bình là 33 cá thể cho theo dõi số nhánh đẽ; 21 - 27 và trung bình là 24 cá thể cho theo dõi diện tích lá; 29 - 37 và trung bình là 33 cá thể cho đếm số lá; 24 - 41 và trung bình là 29 cá thể cho theo dõi trọng lượng chất khô; 2 - 2 và trung bình cùng là 2, nên theo dõi 3 cá thể cho chỉ tiêu đo dài cỏ bông; 8 - 15 và trung bình là 11 cho chỉ tiêu đếm số hạt trên bông; và 9 - 14 và tính trung bình là 9 cá thể cho đếm số hạt chắc trên bông (Bảng 5).

Bảng 5. Dung lượng mẫu cần theo dõi cho các thí nghiệm trồng lúa tại các độ tin và sai số chấp nhận khác nhau (trung bình và khoảng biến động)*

Chỉ tiêu nghiên cứu	Độ tin 95%		Độ tin 99%	
	Sai số 10%	Sai số 5%	Sai số 10%	Sai số 5%
Chiều cao cây	1 (1 - 2)	4 (3 - 6)	2 (1 - 3)	7 (5 - 12)
Số nhánh đẽ	33 (23 - 37)	131 (92 - 132)	60 (42 - 69)	228 (116 - 276)
Chỉ số diện tích lá	24 (21 - 27)	96 (84 - 109)	44 (38 - 51)	178 (152 - 204)
Đếm số lá	33 (29 - 37)	131 (116 - 148)	60 (53 - 67)	238 (210 - 268)
Trọng lượng chất khô	29 (24 - 41)	117 (96 - 166)	53 (45 - 75)	213 (192 - 301)
Chiều dài cỏ bông	2 (2 - 2)	8 (7 - 8)	4 (3 - 4)	14 (13 - 15)
Số hạt/bông	11 (8 - 15)	46 (30 - 62)	21 (14 - 29)	85 (55 - 115)
Số hạt chắc/bông	9 (5 - 14)	38 (21 - 55)	17 (10 - 25)	70 (40 - 100)

* Ghi chú: Số trong ngoặc đơn là biến động của dung lượng mẫu, số đậm phía trên là trung bình

Tại các độ tin cậy và sai số khác nhau cho các trường hợp khác, các dung lượng mẫu cần thiết cũng đã xác định, tham khảo cho mỗi thí nghiệm (Bảng 5). Kết quả thu được có sai khác đôi chút so với kết quả nghiên cứu của Nguyễn Thị Lan (2005). Đây là tổng hợp từ nhiều thí nghiệm và kết quả được tính ra trên cơ sở các biến động đủ đại diện cho mỗi ô thí nghiệm nên độ tin cậy cao hơn.

4. KẾT LUẬN VÀ ĐỀ NGHỊ

4.1. Kết luận

Trong mỗi thí nghiệm, trên mỗi ô, biến động của mỗi chỉ tiêu nghiên cứu là khác nhau.

Giữa các chỉ tiêu nghiên cứu cũng có mức độ biến động khác nhau, chỉ tiêu chiều cao cây có biến động ít nhất. Biến động nhiều nhất là số nhánh đẽ và tiếp theo là số lá trên khóm.

Giữa các loại thí nghiệm khác nhau, biến động của các chỉ tiêu cũng khác nhau do đất đai không giống nhau, mặt khác yếu tố tác động cũng khác nhau như giống, phân bón,...nhưng không thể có quy định dung lượng mẫu cho từng loại thí nghiệm được nên các số trung bình và khoảng biến động trên bảng 5 là có ý nghĩa cho ứng dụng.

Dung lượng mẫu cho một số chỉ tiêu cần theo dõi tại độ tin cậy thông thường 95% và sai số thường chấp nhận 10% đã được xác định: chỉ

tiêu chiều cao cây cần theo dõi: 3 khóm; Các chỉ tiêu số nhánh, diện tích lá, số lá, trọng lượng chất khô theo dõi 30 khóm; Chỉ tiêu chiều dài cổ bông theo dõi 3 bông; Chỉ tiêu số hạt trên bông và hạt chắc trên bông nên theo dõi 10 bông.

4.2. Đề nghị

Cần có các nghiên cứu bổ sung để xác định được nhiều hơn loại chỉ tiêu, trên nhiều điều kiện khác nhau để trung bình hoá được đại diện hơn.

5. TÀI LIỆU THAM KHẢO

Kwanchai A. Gomez & Arturo A. Gomez (1984). *Statistical Procedures For Agricultural research*. Copyright 1984 by John Wiley & sons, Inc. Printed in Singapore.

Nguyễn Thị Lan (2003). *Xác định dung lượng mẫu cho một số chỉ tiêu nghiên cứu với cây đậu tương*, Tạp chí Khoa học kỹ thuật nông nghiệp ĐHNHI tập I số 4/2003. tr. 96-101.

Nguyễn Thị Lan (2005). *Xác định dung lượng mẫu cho một số chỉ tiêu với cây lúa*. Tạp chí Khoa học kỹ thuật nông nghiệp ĐHNHI tập III số 4/2005. tr.278-284.

Nguyễn Văn Tao (1998). *Các phương pháp quan trắc thí nghiệm đồng ruộng chè*. Tuyển tập các công trình nghiên cứu về chè (1988 - 1997). NXB Nông nghiệp, Hà Nội.

Phạm Chí Thành (1986). *Giáo trình phương pháp thí nghiệm đồng ruộng*. NXB Nông nghiệp, Hà Nội.