Expl Agric. (2009), volume 45, pp. 77–91 © *2008 Cambridge University Press*
doi:10.1017/S0014479708007047 *Printed in the United Kingdom*

REML IS AN EFFECTIVE ANALYSIS FOR MIXED MODELLING OF UNBALANCED ON-FARM VARIETAL TRIALS

By D. S. VIRK†*, D. B. PANDIT‡, M. A. SUFIAN‡, F. AHMED§, M. A. B. SIDDIQUE§§, M. A. SAMAD¶, M. M. RAHMAN††, M. M. ISLAM‡, G. ORTIZ-FERRARA‡‡, K. D. JOSHI¶¶ *and* J. R. WITCOMBE†

†*CAZS-Natural Resources, College of Natural Sciences, Bangor University, Wales, UK,* ‡*Wheat Research Centre, Bangladesh Agricultural Research Institute (BARI), Nashipur, Dinajpur, Bangladesh,* §*Regional Agricultural Research Station, BARI, Jamalpur, Bangladesh,* §§*Regional Agricultural Research Station, BARI, Jessore,* ¶*Wheat Testing Station, BARI, Joydebpur, Gazipur,* ††*Regional Wheat Research Centre, BARI, Shyampur, Rajshahi,* ‡‡*International Center for Maize and Wheat Improvement (CIMMYT) South Asia Regional Program, Kathmandu, Nepal and* ¶¶*CAZS-Natural Resources, CIMMYT South Asia Regional Program Office, Kathmandu, Nepal*

(*Accepted 23 July 2008*)

SUMMARY

On-farm participatory varietal selection (PVS) trials are often of two types: mother trials (with all of the entries) and baby trials (each having one, or very few of the entries from the mother trials). We conducted PVS trials on 17 wheat varieties in 12 villages of four districts of Bangladesh over three years but the data were highly unbalanced. Both quantitative and qualitative traits were measured in the on-farm trials. The factors in the trials were both fixed effects (varieties and districts) and random (years and farmers). We used the residual or restricted maximum likelihood (REML) analysis for the mixed model for quantitative traits. For qualitative data on farmers' perceptions, logistic regression procedures were used that are equally applicable to balanced and unbalanced data sets. The REML analysis provided adjusted mean values for quantitative traits for all the varieties, for the mother and baby trials separately, using the data from all years and all locations. It identified varieties BAW 1006 and BAW 1008 that yielded 19–30% more than the control Kanchan and also had a higher 1000-grain weight, were at least as early to flower and had a high overall ranking by farmers in the mother trials. The logistic regression analysis of perception data agreed with the results of the REML analysis as these varieties were most preferred by farmers for grain yield, earlier maturity and better *chapatti* making quality. The less labour-intensive method of recording qualitative perceptions can usefully replace actual yield measurements, particularly when validated by other participatory measures such as intended and actual adoption. In 2005, BAW 1006 was released as BARI Gom 23 or Bijoy and BAW 1008 as BARI Gom 24 or Prodip for the whole of Bangladesh. The validity of the results of the REML analysis was confirmed by the high early adoption trends of the identified varieties. Since REML is an effective analysis for unbalanced PVS trial data using a mixed model, its wider use by researchers would increase the value of the PVS process.

INTRODUCTION

Participatory varietal selection (PVS) is being increasingly used in crop improvement to allow farmers to test and identify varieties (e.g. Sperling and Scheideggers, 1995; Virk

*Corresponding author: d.s.virk@bangor.ac.uk

et al., 2005; 2006; Witcombe, 1999). It increases farmers' awareness of new varieties and enables them to save and exchange the seed of those that they prefer (Grisley and Shamambo, 1993; Sperling and Loevinsohn, 1993; Witcombe *et al.*, 1999). Following the terminology of Snapp (1999), on-farm PVS trials can be of two types: mother trials (that include the full set of varieties) and baby trials (that include one or more varieties alongside the control). Many variations of the mother-baby trial design are used by researchers (Witcombe *et al.*, 2005).

On-farm variety trials have the advantage of testing varieties in the real environment where they will ultimately be grown. However, they are often unbalanced because of the failure of some trials grown under rainfed conditions in farmers' fields and because limited quantities of seed of new varieties dictate in what year and on what scale they enter into the trials. The statistical analysis of such unbalanced mother trials is more complex than for the more balanced data typical of on-station trials. However, analysis of data from paired-plot (new variety versus local control) baby trials only requires a paired *t*-test (Eskridge, 1996; Joshi and Witcombe, 1996). A combined analysis of either the mother or the baby trials that were conducted over several years, locations and farmers involves both fixed and random factors. Fixed effects are those where levels are determined or set by the experimenter, e.g. varieties and districts in our case. Random effects are where the levels of the effects are assumed to be randomly selected from an infinite population of possible levels, e.g. years. While linear fixed effect models are analysable by general linear modelling (GLM), mixed models are more efficiently handled by the residual (or restricted) maximum likelihood (REML) technique (Allan and Rowlands, 2001; Littell *et al.*, 1996; Searle *et al.*, 1992). REML provides efficient estimates of treatment effects that have more than one source of error in a multilevel data structure, e.g. in on-farm trials there is information at the district, village and farm level. The analysis also allows multi-layered data information to be combined from multiple experiments conducted over time and space. Hence, REML is more appropriate than GLM for PVS trials that have complex-unbalanced strata of data in mixed models (Coe, 2002). Despite this, few researchers have used REML analysis for PVS trials (Coe, 2002; Omanya *et al.*, 2007; Virk and Witcombe, 2008).

Through PVS we offered new wheat varieties to farmers in Bangladesh over three years to identify alternatives to the widely grown but old variety, Kanchan (released in 1983). We examine the utility of the REML procedure in analysing these unbalanced trials using mixed models and compare these results with the extent of adoption of the new varieties by farmers.

MATERIALS AND METHODS

Experimental sites

The experimental sites were in four diverse districts of Bangladesh: Dinajpur (25°38′N; 88°41′E, 38 m asl) and Rajshahi (24°22′N; 83°36′ E, 17 m asl) in the north west, Jessore (23°11′N; 89°14′E, 7 m asl) in the southwest and Jamalpur (24°56′N; 89°55′E, 19 m asl) in the north of the country (Figure 1). The on-farm experiments

https://doi.org/10.1017/S0014479708007047 Published online by Cambridge University Press

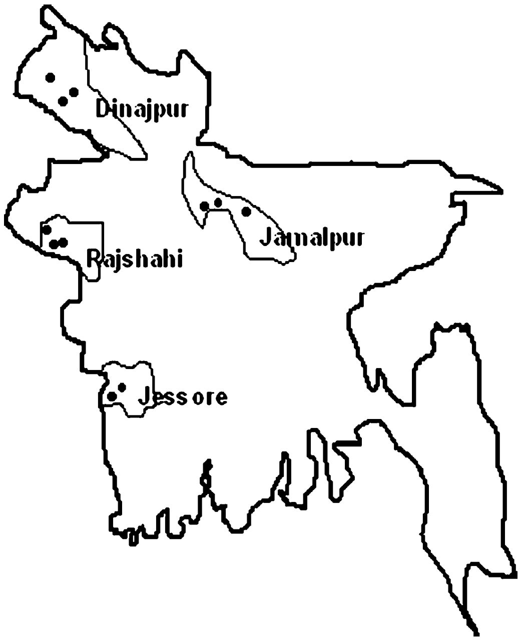

Figure 1. A sketch map of Bangladesh indicating the locations of the project districts and villages.

were conducted in three villages in each district: Daulatpur, Jogdal and Hatiari in Dinajpur district; Duary, Basantapur and Santoshpur in Rajshahi district; Valukghar, Jalalpur and Bandabila in Jessore district; Raier Bakai, Sthall and Norkona in Jamalpur district.

Dinajpur and Jamalpur districts are the major wheat-growing areas of Bangladesh. However, the soils (pH 4.5–5.5) in Dinajpur suffer from micro-nutrient deficiencies that cause spike sterility. High rainfall in Jamalpur causes sprouting of grains in the spike and farmers prefer early maturing varieties. The soils here have pH ranging from 5.0 to 6.5. Rajshahi represents the Barind zone with low rainfall, and alkaline soils of poor fertility. The soils in Jessore are calcareous brown flood plain type with a pH of 7.8.

Varieties for PVS

A total of 17 varieties were tested in the PVS trials. Although there were seven test entries and a control variety (Kanchan) each, these trials were highly unbalanced as varieties were dropped because of poor performance and new varieties were added as seed became available (Tables 1 and 2).

Experimental design and evaluation

A mother and baby design was used in testing the varieties (Snapp, 1999; Witcombe, 2002). Each mother trial always had eight entries in a single replicate of which one was always Kanchan as a control. In the first two years only, Shatabdi was also a

Table 1. The wheat varieties tested in the mother and baby trials in each year in four districts in Bangladesh from 2002/03 to 2004/05.

Variety	2002/03 All	2003/04 All	2004/05 Dinajpur	2004/05 Jessore	2004/05 Rajshahi	2004/05 Jamalpur
Kanchan	Y†	Y	Y	Y	Y	Y
Shatabdi	Y	Y	N	N	N	N
BAW 966	Y	Y	N	N	N	N
BAW 1004	Y	Y	N	N	N	N
BAW 1006	Y	Y	Y	Y	Y	Y
BAW 1008	Y	Y	Y	Y	Y	N
BL 1473	Y	N	N	N	N	N
BL 1887	Y	N	N	N	N	N
BAW 1027	N‡	Y	N	N	Y	Y
BAW 1035	N	Y	Y	Y	Y	Y
BAW 1059	N	N	Y	Y	Y	Y
BAW 1064	N	N	Y	Y	Y	Y
BAW 1065	N	N	Y	N	N	Y
BAW 1066	N	N	N	N	Y	N
BAW 1072	N	N	Y	N	N	Y
BAW 1081	N	N	N	Y	N	N
BAW 1088	N	N	N	Y	N	N

†Y = variety tested.
‡N = not tested.

Table 2. Total number of mother and baby trials conducted for each wheat variety in four districts across three years (2002/03 to 2004/05).

	Dinajpur		Jamalpur		Jessore		Rajshahi	
Variety	Mother	Baby	Mother	Baby	Mother	Baby	Mother	Baby
BAW 1004§	15	19	15	19	15	19	15	19
BAW 1006	24	49	24	49	24	49	24	19
BAW 1008	24	49	15	19	24	49	24	49
BAW 1027	9	15	18	45	9	15	18	45
BAW 1035	18		18		18		18	
BAW 1059	9		9		9		9	
BAW 1064	9		9		9		9	
BAW 1065	9		9					
BAW 1066							9	
BAW 1072	9		9					
BAW 1081					9			
BAW 1088					9			
BAW 966	15	19	15	19	15	19	15	19
BL 1473¶	6	4	6	4	6	4	6	4
BL 1887¶	6	4	6	4	6	4	6	4
Kanchan†	24	133	24	133	24	133	24	133
Shatabdi‡	15	19	15	19	15	19	15	19

†Most grown farmers' control variety released in 1983.
‡Released in 2000.
§Rejected for its lodging susceptibility in 2003/04.
¶Rejected for poor performance in 2002/03.

https://doi.org/10.1017/S0014479708007047 Published online by Cambridge University Press

control. The trials were replicated across farmers and each farmer was given a trial with different randomization. In 2002/03 the mother trial design was four districts × two villages × three farmers (24 replications); in 2003/04 and 2004/05 the mother trials were in four districts × three villages × two farmers (24 replications in each year). The farmers who were given the trials varied across the years.

The mother trials were conducted using farmers' management. The mother trials in farmers' fields were grown in 5 m × 4 m plots with row spacing of 20 cm and thick planting within rows. A seed rate of 120 kg ha^{-1} was used. One set of each mother trial was also conducted on research stations (at Dinajpur, Jamalpur, Rajshahi and Jessore) in a randomized complete block design with three replications. The on-station trials were fertilized with 100N:26P:33K:20S kg ha^{-1} and irrigated 2–3 times. Two-thirds of the N (urea) and all other fertilizers were applied as a basal dose and the remaining N was top dressed after the first irrigation.

Both on-station and on-farm mother trials were researcher designed and completely supervised by the researchers who also recorded all quantitative data. In addition to quantitative data, matrix ranking for several traits was done by farmers at physiological maturity and at the post-harvest stage. Each year, farmers scored the eight varieties in the mother trials on a 1 to 8 scale; 1 = lowest and 8 = highest score for grain yield, maturity, *chapatti* making quality and overall preference. The trials were evaluated by 30 farmers in each village at the pre- and post-harvest stages.

Baby trials corresponded to the two designs described by Witcombe (2002): in Design I, one new variety was grown alongside the control variety, and in Design II, two new varieties were grown side by side with a local control by a farmer under the farmer's management. In 2002/03, Design I trials were replicated across four districts × two villages and two farmers (16 replications) for each of seven new varieties with Kanchan as the control (total 112 trials and 224 individual plots). In 2003/04, Design II was replicated across four districts × three villages × one farmer (12 replications) for each of 15 all possible pairs of six new varieties with Kanchan as the control (total 180 trials and 540 invdividual plots). In 2004/05, Design I was replicated across four districts × three villages × 20 farmers (240 replications) for each of two pairs of varieties selected from the previous year's trials with Kanchan as the control (total 240 trials = 480 individual plots). The farmers who conducted the baby trials varied across years.

Baby trials were evaluated using household level questionnaires (HLQ) following Witcombe (2002). Yield data and plot size were recorded by researchers, and grain yield recorded as kg $plot^{-1}$ was converted to t ha^{-1}.

Assessment of adoption of varieties

Baseline information on varietal diversity was collected through participatory rural appraisal (PRA) tools in four villages in the four districts in August-September 2002. A multidisciplinary team of two agricultural economists, an agricultural engineer, an Agriculture Extension Officer, three agronomists, two breeders, a soil scientist and a plant pathologist, conducted the baseline PRA. Four discussion groups were conducted (one per village) that involved 25 farmers per group and included large, medium, small, marginal and tenant farmers. The PRA covered many topics but only the results for varietal diversity of wheat are presented here. Farmers were asked what were the popular varieties of wheat in their village and the proportion of wheat land that each variety covered.

The varietal adoption surveys were conducted in 2004/05 (three seasons after the start of the PVS) using HLQs to collect information on variety adoption. In each of the 12 villages where the mother trials were grown 30 farmers were interviewed; 10 were participating farmers who grew either a mother or a baby trial; 10 were farmers who did not grow a trial but who had participated in training or demonstrations by researchers; and 10 were new farmers who had had no previous interaction with the researchers. The only female farmers were eight out of the 30 farmers sampled in Dinajpur district

Statistical analysis

All statistical analyses for quantitative and qualitative traits excluded BAW 1004, BL 1473 and BL 1887 since these three varieties had already been rejected by farmers in the first two years.

Quantitative traits. There were a total of 768 farmer plots for analysis in the mother trials and 1244 in the baby trials. Two combined analyses of variance across three years and four districts were computed, one for the mother and one for the baby trials for grain yield, time to heading, 1000-grain weight and overall scores. The analysis of unbalanced data was performed for the mixed model using residual (or restricted) maximum likelihood (REML) analysis of GenStat 8. The REML algorithm estimates the variance components using residual maximum likelihood, and then uses these variance parameter estimates to form the generalised least squares estimates of the treatment effects and the best linear unbiased predictors (BLUPs) of the random effects along with their standard errors (DeLacy *et al.*, 1996; Piepho, 1994).

In our case, varieties were fixed effects because they could be repeated and were summarized by individual means and standard errors. Similarly, districts were treated as fixed effects since they represent zones and the interest was to compare mean values of districts widely placed in Bangladesh. In contrast, the influence of years on the measured traits in a trial is a random effect and the interest is to quantify the variability in a trait due to the year effect rather than comparing individual years. Similarly, farmers within villages represented a random sample of farmers from any village. Years or farmers within villages are considered to be a random sample from a normal distribution, so they are summarized as variance components and as predicted means (BLUPs).

The REML mixed model for all the traits measured in the mother and baby trials had two parts:

(i). Fixed model = Constant + Variety + District + Variety × District
(ii). Random model = Year + Farmer

https://doi.org/10.1017/S0014479708007047 Published online by Cambridge University Press

Analysis of the mixed model with the REML procedure combines both parts (i) and (ii) of the model.

The major interest was to compare the varietal fixed effects and to determine the interaction of varieties with the fixed effects of districts so as to identify district- or zone-specific adaptation. Therefore, year × variety interactions were not estimated.

Average standard errors of the difference among variety means within a district were used to compute the least squares differences (Snedecor and Cochran, 1973).

Logistic regression analysis for qualitative traits. The logistic regression is useful for categorical response variables, unlike the continuous variables used in ordinary regression. Ordinal logistic regression can be used when there are three or more categories with natural ordering of levels. However, when there are only two classes of data that have a binomial distribution a binary logistic regression is used.

Data from baby trials were available for three qualitative perceptions (grain yield, maturity and *chapatti* making quality) from the HLQ evaluations. These data were recorded as 'less', 'same' and 'more' in comparison to the control (Kanchan). These classes of data equate to 1, 2 and 3 classes in that order for an ordinal regression analysis. For each variety the total counts in each class were used for the analysis. The means of the counts in each category were used as the reference points in the logistic regression and the analysis was computed for each district separately.

The output from logistic regression analysis provided coefficients of regression, standard errors, t-values and p-levels for testing their significance. The log odds ratio or the estimated coefficient of a factor measures the difference between the given level of a factor (i.e. 'more' or 'less' in our case) compared to the reference level (i.e. 'same' in our case) and is: (log odds of a variety 'A' being high versus low)/(log odds of the control being high versus low). The antilog of the log odds ratio (or coefficient of regression) gives the odds ratio which can be used to compare a new variety with the control directly as the odds ratio (r) of a variety will be $((r-1) \times 100)$ times higher than the control. We computed the odds ratios for each variety for comparison with the control (Kanchan). Estimated regression coefficients of zero or an odds ratio of 1 both imply that the factor has no effect.

Binary logistic regression was used to analyse the two-category data on the intention of farmers to 'grow' (yes) or 'not grow' (no) any variety in the following year. These data were available for five varieties (BAW 1006, BAW 1008, BAW 1027, BAW 966 and Shatabdi) tested in 2002/03 to 2004/05. Counts of 'yes' and 'no' for the varieties in each district separately and pooled over all districts for all three years were used in the analysis.

Variety-on-variety regression for genotype × environment (G × E) interaction. A comparison of the overall mean yield of a new variety with the overall mean of a standard control can be misleading when there are substantial G × E interactions. In the more common regression analysis for G × E interactions individual variety means are regressed onto environmental indexes that use means of all of the varieties in the trials. However, in the present case, the interest centres on a comparison of the new variety with a

Table 3. Significance of the main effects and their interactions tested by chi-squared (Wald statistic/*d.f.*) values in the overall REML analysis for (a) the mother trials for four traits of the wheat varieties and (b) for the baby trials for the grain yield of the wheat varieties, 2002/03 to 2004/05.

		(a) Mother trials				(b) Baby trials
Source	*d.f.*	Grain yield ($t\ ha^{-1}$)	Time to heading (days)	1000-grain weight (g)	Overall score†	Grain yield ($t\ ha^{-1}$)
Variety, V	13 (5)	9.3***	17.90***	83.10***	38.66***	76.9***
District, D	3 (3)	44.7***	128.59***	134.22***	6.41***	6.7***
V × D	26 (15)	3.2***	3.58***	2.79***	4.14***	8.1***

***$p < 0.001$.

†Overall score by matrix scoring of eight varieties on 1 = lowest and 8 = highest scale.

standard control so a regression of the performance of the new variety on that of the standard control is more appropriate (Simmonds, 1979). The standard variety mean values across trials provides an environmental index that is independent of the test variety (the commonly used mean of all of the varieties in the trial is not completely independent). These regressions were computed separately for the mother and the baby trials. The results for the two best varieties (from both the REML analysis and from adoption data) are presented.

RESULTS

Combined REML analysis of variance of mother or baby trials revealed significant variation among varieties and districts for all traits (Table 3). There was also a significant variety × district interaction for all traits in mother trials and for grain yield in baby trials because of the differential response of varieties to the environments in the four districts.

The district mean grain yields of all varieties in mother (baby) trials were: 3.76 (2.52) $t\ ha^{-1}$ for Jamalpur, 3.20 (3.10) $t\ ha^{-1}$ for Rajshahi, 4.20 (3.17) $t\ ha^{-1}$ for Dinajpur and 3.66 (3.31) $t\ ha^{-1}$ for Jessore. Thus the mean grain yields in mother trials compared to the baby trials were 32% higher in Dinajpur, 49% in Jamalpur, 11% in Jessore and 3% in Rajshahi. The on-station mother trials received higher inputs than those on farmers' fields, which could have biased their overall mean yield upwards. However, the lower yields of the baby trials better represented the actual farmers' practices.

The variety mean grain yield ranks across districts were different due to a high level of variety × district interaction (Tables 3, 4). However, the identification of region-specific varieties is difficult because of the highly unbalanced structure of trials with varieties that showed significant interactions having been tested only in one year (BAW 1059 and BAW 1065) or tested only in two years (Shatabdi).

When performance in individual districts was considered, none of the varieties yielded significantly more than Kanchan in Dinajpur in mother trials, but BAW 1006 and BAW 1008 were superior in baby trials (Table 4). In Rajshahi, BAW 1006, BAW 1008 and Shatabdi significantly out-yielded Kanchan in both mother and baby trials,

https://doi.org/10.1017/S0014479708007047 Published online by Cambridge University Press

Table 4. District-wise mean grain yield (t ha^{-1}) of wheat varieties tested in mother (MT) and baby (BT) trials for three years (2002/03 to 2004/05). Blank cells indicate variety was not tested.

	Dinajpur		Jamalpur		Jessore		Rajshahi		Overall	
Variety	MT	BT	MT	BT	MT	BT	MT	BT	MT	BT
BAW 1006	4.1	3.4‡	3.9‡	3.1‡	3.9‡	3.7‡	3.2	3.0	3.8‡	3.3‡
BAW 1008	4.2	3.5‡	3.7‡	2.8‡	4.0‡	4.0‡	3.4‡	3.5‡	3.8‡	3.5‡
BAW 1027	4.0	3.1	3.8‡	2.7‡	4.2‡	3.4‡	3.3†	3.3‡	3.8‡	3.1†
BAW 1035	4.1		4.1‡		3.7†		3.1		3.7‡	
BAW 1059	4.1		4.2‡		3.6		3.0		3.7‡	
BAW 1064	4.3		5.0‡		3.4		2.7		3.8‡	
BAW 1065	4.2		4.5‡							
BAW 1066							2.9			
BAW 1072	4.2		3.9‡							
BAW 1081					3.2					
BAW 1088					3.3					
BAW 966	3.7	2.7	3.4	2.9‡	3.9‡	3.4‡	3.4‡	2.7	3.6‡	2.9‡
Kanchan	3.8	3.0	3.0	2.1	3.2	2.7	2.7	2.9	3.2	2.7
Shatabdi	4.2	3.1	3.6†	3.0‡	4.0‡	3.7‡	3.4‡	3.4‡	3.8‡	3.3‡
s.e.d.	0.24	0.16	0.24	0.16	0.24	0.16	0.24	0.16	0.12	0.08

†Significantly superior to Kanchan at $p < 0.05$.
‡Significantly superior to Kanchan at $p < 0.01$.

but BAW 966 did so in mother trials only (Table 4). In Jamalpur and Jessore, there were several varieties superior to Kanchan for grain yield in both mother and baby trials, e.g. BAW 1006, BAW 1008 and BAW 1027 (Table 4). These varieties also had bolder grains and higher overall preference scores than Kanchan and flowered at the same time (Table 5).

When performance across all districts was considered, all new varieties and Shatabdi were significantly superior to Kanchan for grain yield. Of the new varieties, BAW 1006 and BAW 1008 were the most promising for all districts; in mother and baby trials BAW 1006 was 19% and 22% higher yielding than Kanchan, and BAW 1008 was 19% and 30% higher yielding (Table 4). They also had higher overall scores and 1000-grain weight than Kanchan in all districts but significant to non-significant earlier flowering (two to three days earlier in Dinajpur and Jamalpur) in different districts (Table 5).

Ordinal logistic regression

For grain yield, four new varieties (BAW 1006, BAW 1008, BAW 1027 and Shatabdi) were perceived by farmers to be superior to Kanchan in each of the four districts (Table 6). Of these all, apart from Shatabdi, were perceived to be earlier to mature. This differed from the quantitative yield data in both the mother and the baby trials where, for example, there were no significant differences between Kanchan and the new varieties in Dinajpur (Table 4). Farmers perceptions on *chapatti* making quality showed that BAW 1004 and BAW 1006 were consistently superior in three of the four districts while BAW 1027 was poor (Table 6).

Table 5. Mean values by district for time to heading (TH in days), thousand-grain weight (TGW in g) and overall score (OS on 1 to 8 scale) of wheat varieties tested in mother trials in four districts over three years (2003–2005).

	Dinajpur			Jamalpur			Jessore			Rajshahi		
Variety	TH	TGW	OS	TH	TGW	OS	TH	TGW	OS	TH	TGW	OS
BAW 1006	68‡	52‡	7.2‡	65‡	45‡	6.6‡	68	43‡	7.2‡	70	49‡	6.4‡
BAW 1008	69	57‡	7.4‡	65‡	51‡	6.9‡	65	47‡	7.3‡	70	55‡	6.9‡
BAW 1027	67‡	48	6.3	65‡	46‡	6.3‡	64	47‡	6.5‡	70	50‡	6.3‡
BAW 1035	70	51‡	5.3	67	44‡	6.4‡	66	42‡	6.0†	71	49‡	5.7
BAW 1059	69	57‡	6.2	64‡	50‡	7.2‡	68	40‡	7.4‡	66‡	52‡	6.7‡
BAW 1064	68†	49†	6.0	65‡	40‡	7.0‡	67	38	7.4‡	68‡	44	6.4‡
BAW 1065	69	47	5.2	65‡	41‡	5.9‡						
BAW 1066										70	44	5.2
BAW 1072	69	43	4.3	66	35	4.6						
BAW 1081							70	37	6.0†			
BAW 1088							70	37	5.7			
BAW 966	66‡	39	5.7	61‡	35	5.4‡	62†	34	5.6	69‡	38	5.1
Kanchan	71	45	5.8	68	35	4.0	66	36	5.2	71	41	5.3
Shatabdi	71	47	7.0‡	68	43‡	6.4‡	70	40‡	6.0†	75	47‡	6.3‡
s.e.d.	1.02	1.50	0.35	1.02	1.50	0.35	1.02	1.50	0.35	1.02	1.50	0.35

†Significantly earlier to flower and with higher grain weight and overall preference than Kanchan at $p < 0.05$.
‡Significantly earlier to flower and with higher grain weight and overall preference than Kanchan at $p < 0.01$.

Table 6. Logistic regression coefficients and odds ratios (OR) for qualitative perceptions recorded in baby trials. Varieties with non-significant parameters are omitted. Blank cells show that estimates were non-significant. All regression coefficients (b) are relative to Kanchan control variety and only significantly positive (increases) estimates for grain yield and *chapatti* quality, and negative (decreases) for maturity are given.

		Grain yield		Maturity		*Chapatti* quality	
District	Variety	$b \pm s.e.$†	OR	$b \pm s.e.$	OR	$b \pm s.e.$	OR
Dinajpur	BAW 1006	4.51 ± 0.59	91	-3.38 ± 0.50	0		
	BAW 1008	4.56 ± 0.59	95	-3.65 ± 0.51	0		
	BAW 1027	2.21 ± 0.88	9	-2.43 ± 0.68	0		
	BAW 966	-3.63 ± 0.70	0				
	Shatabdi	5.32 ± 0.79	204	1.54 ± 0.75	5		
Jamalpur	BAW 1006	5.43 ± 0.78	229	-5.09 ± 1.05	0	6.29 ± 1.05	539
	BAW 1008	4.51 ± 0.82	91	-5.59 ± 1.14	0		
	BAW 1027	3.83 ± 0.51	46	-5.04 ± 1.06	0	-2.37 ± 0.44	0
	BAW 966	2.84 ± 0.59	17				
	Shatabdi	4.51 ± 0.80	91	5.83 ± 1.15	341	4.30 ± 0.82	74
Jessore	BAW 1006	6.06 ± 0.82	427	-4.55 ± 0.69	0	8.94 ± 1.52	7620
	BAW 1008	6.06 ± 0.82	427	-5.15 ± 0.71	0		
	BAW 1027	3.28 ± 0.71	27	-2.92 ± 0.84	0	-6.12 ± 1.36	0
	BAW 966	2.54 ± 0.64	13				
	Shatabdi	3.89 ± 0.69	49	4.83 ± 0.81	126		
Rajshahi	BAW 1006	4.20 ± 0.72	66	-6.95 ± 1.27	0	4.83 ± 0.78	125
	BAW 1008	4.91 ± 0.62	136	-4.97 ± 0.79	0	5.88 ± 0.82	358
	BAW 1027	4.57 ± 0.61	96	-3.80 ± 0.79	0	-4.29 ± 0.57	0
	BAW 966	-2.64 ± 0.66	0	-6.95 ± 1.27	0		
	Shatabdi	5.41 ± 0.80	225	4.56 ± 0.88	95	5.31 ± 0.88	203

†Regression coefficients (b) > 1.96 times the standard error are significant at $p < 0.05$.

Table 7. Parameters of binary logistic regression (only given for the significant regressions (b)) and odds ratios for varieties for the 'Yes' or 'No' counts of farmers' replies in baby trials in 2003, 2004, 2005 if they would grow a variety next year.

District	Variety	$b^{\dagger}$	*s.e.*	Odds ratio	Deviance X^2 for 5 *d.f.*[§]
Dinajpur	BAW 1006	4.67[‡]	2.56	106.60	2.21
	BAW 966	−3.35	1.41	0.03	
Jamalpur	Nil				23.81[¶]
Jessore	Nil				0.00
Rajshahi	BAW 1027	−2.74	1.26	0.06	6.23
	BAW 966	−3.08	0.99	0.05	
Overall	BAW 1006	2.39	1.03	10.87	7.25
**	BAW 1008	2.39	1.03	10.87	
	BAW 1027	−2.46	0.82	0.09	
	BAW 966	−2.64	0.74	0.07	

[†]Regression coefficients (b) >1.96 times *s.e.* are significant at $p < 0.05$.
[‡]Probability of significance < 0.10.
[§]Deviance when significant would reject H0 hypothesis of an adequate fit of a model.
[¶]Significant at $p < 0.01$.

Overall, BAW 1006 and BAW 1008 were identified as superior varieties both by the REML analysis for quantitative traits and the logistic regression for qualitative perceptions in all districts.

Binary logistic regression

The binary regression model that tested the significance of whether farmers would grow a variety again varied for its goodness of fit in different districts (Table 7). In the overall analysis, BAW 1006 and BAW 1008 were the most preferred varieties with higher odds ratios, as farmers wanted to grow them again (Table 7). Of these two, only BAW 1006 was significantly preferred in a within-district analysis (Table 7).

Variety-on-variety regression

In both mother and baby trials BAW 1006 and BAW 1008 had significant regression slopes < 1.0 and intercept > 0 in comparison to Kanchan. Hence they performed better in the low-yielding environments than Kanchan, and may have specific adaptation to the lower-yielding fields cultivated by more resource-poor farmers.

On-farm diversity

Baseline PRAs in 2002/03, in the four study districts, revealed that Kanchan (an old variety) was being cultivated on 95–100% of wheat area. The areas under other improved varieties such as Protiva, Gaurab and Sourab were negligible (less than 1% for each). Sonalika (a very old variety) was grown on very small areas in Jessore and Rajshahi (less than 5%). Farmers were thus practicing a near varietal monoculture that gave a high risk of disease epidemics.

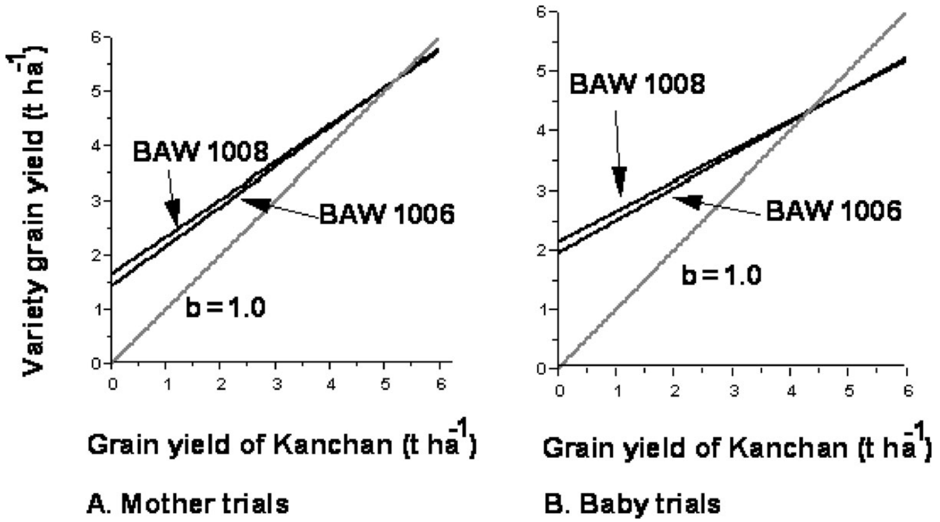

Figure 2. Regression of two varieties on to Kanchan in mother and baby trials conducted in four districts of Bangladesh, 2003 to 2005. The regression equations for the mother trials were: BAW 1006 = $1.4 \pm 0.3 + 0.7x \pm 0.1x$; BAW 1008 = $1.7 \pm 0.3 + 0.7x \pm 0.1$. The regression equations for the baby trials were: BAW 1006 = $2.0 \pm 0.2 + 0.5x \pm 0.1$; BAW 1008 = $2.2 \pm 0.2 + 0.5x \pm 0.1$.

Table 8. Trends in adoption of new wheat varieties (% of total wheat area) in the PVS villages in 2004/05 compared to 2002/03 in four districts, Bangladesh.

	Dinajpur		Jessore		Rajshahi		Jamalpur		Average	
Variety	2003	2005	2003	2005	2003	2005	2003	2005	2003	2005
Kanchan	100	62.3	95	51.5	98	59.0	100	55.0	98.2	57.0
Sonalika		0.0	4	0.0	2	0.0		0.0	1.5	0.0
Protiva		2.1	1	2.1	0	5.0		5.0	0.3	3.6
Shatabdi		34.3	0	43.8	0	25.0		31.5	0	33.7
Sourav		0.2	0	1.0	0	3.0		3.0	0	1.8
Gourab		0.7	0	1.0	0	5.0		2.0	0	2.2
BAW 1006		0.2	0	0.1	0	1.4		2.0	0	0.9
BAW 1008		0.3	0	0.2	0	1.6		1.5	0	0.9

In a short period of three years (2002/03 to 2004/05) the area of Kanchan decreased significantly in each of the four districts with an overall decline from 98% to 57% of the area (Table 8). Among the new varieties that replaced Kanchan, the released variety Shatabdi was the most adopted because seed was relatively easily available. The new PVS varieties BAW 1006 and BAW 1008 had started to spread and the area of each was comparable to the released varieties Protiva, Gourab and Sourav (Table 8).

DISCUSSION

REML analysis of on-farm trials

In the PVS trials farmers were given a choice of both released and unreleased varieties (Witcombe *et al.*, 1996) that were tested in mother and baby trials. As is

https://doi.org/10.1017/S0014479708007047 Published online by Cambridge University Press

often the case for on-farm trials, unlike on-station trials, they were highly unbalanced (Table 1).

The REML procedure uses the full set of data with missing cells for varieties and computes predicted means for each entry adjusted for the site and year effects. The adjustments allow for the fact that some farmers obtain higher average yields than others. Entries that are tested on 'good' farms will have their means biased upwards compared with those tested on 'poor' farms. The predicted means adjust for these biases and should be used instead of raw means (Coe, 2002). The REML analysis removes the limitation of comparing a control and a new variety grown together and allows a combined analysis of the baby trials across all entries with a variable number of farmers and makes it possible to compare varieties among themselves and with the control variety.

REML analysis and other analyses

The results of REML analysis were confirmed by the logistic regression analysis on the perception data, where varieties BAW 1006 and BAW 1008 were preferred by farmers for grain yield, as well as maturity and *chapatti* making quality. Hence, quantitative yield data could be replaced with more easily recorded farmers' scores or perceptions of varieties. Virk and Witcombe (2008) also found that actual yields and perception scores for yields identified the same varieties. The overall binary regression analysis also confirmed their acceptability as significantly more farmers said they would grow them next year compared with those who would not.

However, when these two analyses (REML for quantitative data and logistic regression for perceptions) were compared by district perception data showed more significant differences. For example, although there were no significant differences between varieties for grain yield in Dinajpur in both the mother and baby trials (Table 4) several varieties differed significantly from Kanchan for yield perceptions and included two of the best varieties overall, BAW 1006 and BAW 1008 (Table 6). This additional sensitivity of perception data is an advantage of participatory trials since over-sensitivity can be ruled out by validation with adoption intentions (binary regression) (Table 7) and actual adoption data (Table 8).

The resource-poor farmers of Bangladesh require varieties with specific adaptation to low fertility soils. The variety-on-variety analysis for genotype × environment interaction showed that the two promising new varieties (BAW 1006 and BAW 1008) were significantly less responsive in the high-yielding environments but were higher yielding in the low-yielding environments than Kanchan, and hence were more adapted to the marginal environments (Figure 2). These two varieties were released in Bangladesh in 2005; BAW 1006 as BARI Gom 23 or Bijoy; and BAW 1008 as BAR Gom 24 or Prodip).

REML analysis and adoption of varieties

Any analysis of on-farm trials is powerful enough if it can identify varieties that are preferred and adopted by farmers. We applied this criterion to the REML analysis and studied the adoption patterns of the identified varieties.

The adoption of REML-identified new varieties increased the replacement rate of the old variety Kanchan as its area fell significantly from 98 to 57% of the total wheat area. Varieties BAW 1006 and BAW 1008 each occupied about 1% of the area of wheat even before their release in 2005. In 2004/05, participating farmers of the four project districts saved 5 and 6 tonnes of seed of the two new varieties BAW 1006 and BAW 1008, and farmer-to-farmer seed spread had already begun (Ortiz-Ferrara *et al.*, 2007).

We have demonstrated that the REML analysis removes the two major problems of analysing the data from on-farm trials, i.e. mixed modelling of factors and non-orthogonality of the data. The use of REML will help to exploit fully the power of PVS trials in which varieties are tested widely over many more locations, though with irregularities, than conventional on-farm testing and where the multilayered data have mixed effect factors. Our results from the REML analysis were confirmed by other analyses and from the rapid adoption of the REML identified varieties by farmers. Using these analyses we found that at least two varieties were most promising and preferred by farmers, and these were later released in Bangladesh. The newly identified varieties have specific adaptation to the lower-yielding environments of the fields of resource-poor farmers. A wider application of REML analysis will remove bottlenecks in handling PVS trial data and encourage the adoption and institutionalization of the PVS process.

Acknowledgements. This paper is an output from a project funded by the UK Department for International Development (DFID) and administered by CIMMYT Nepal and CAZS-Natural Resources for the benefit of developing countries in South Asia. The views expressed are not necessarily those of DFID. The authors gratefully acknowledge Dr D. Harris of CAZS-NR for his suggestions and all those farmers of Bangladesh who participated.

REFERENCES

Allan, E. and Rowlands, J. (2001). *Mixed Models and Multilevel Data Structures in Agriculture*. Statistical Service Centre, The University of Reading, UK.

Coe, R. (2002). Analyzing data from participatory on-farm trials. In *Quantitative Analysis of Data from Participatory Methods in Plant Breeding*, 18–34 (Eds R. Mauricio, and J. Reeves). Mexico, DF: CIMMYT.

DeLacy, I. H., Basford, K. E., Cooper, M., Bull, J. K. and McLaren, C. G. (1996). Analysis of multi-environment trials – a historical perspective. In *Plant Adaptation and Crop Improvement*, 39–124 (Eds M. Cooper and G. L. Hammer). CAB International, Wallingford, Oxon, UK.

Eskridge, K. M. (1996). Analysis of multiple environmental trials using the probability of outperforming a check. In *Genotype-by-environment Interaction*, 283–308 (Eds M. S. Kang and H. G. Gauch). CRC Press, Boca Raton, FL, USA.

Ortiz-Ferrara, G., Chand, R., Bhatta, M. R., Mudwari, A., Thapa, D. B., Sufian, M. A., Saikia, T. P., Chatrath, R., Witcombe, J. R., Virk, D. S. and Sharma, R. C. (2007). Partnering with farmers to accelerate adoption of new technologies in South Asia to improve wheat productivity. *Euphytica* 157:399–407.

Grisley, W. and Shamambo, M. (1993). An analysis of adoption and diffusion of Carioca beans in Zambia resulting from an experimental distribution of seed. *Experimental Agriculture* 29:379–386.

Joshi, A. and Witcombe, J. R. (1996). Farmer participatory crop improvement. II. Participatory varietal selection, a case study in India. *Experimental Agriculture* 32:461–477.

Littell, R. C., Nilliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996). *SAS System for Mixed Models*. SAS Institute Inc., Cary, NC, USA.

Omanya, G. O., Weltzein-Rattunde, E., Sogodogo, D., Sanogo, M., Hanssens, N., Guero, Y. and Zangre, R. (2007). Participatory varietal selection with improved pearl millet in west Africa. *Experimental Agriculture* 43:5–19.

Piepho, H. P. (1994). Best linear unbiased prediction (BLUP) for regional yield trials: A comparison to additive main effects and multiplicative interactions (AMMI) analysis. *Theoretical Applied Genetics* 89:647–654.

Searle, S. R., Casella, G. and McCulloch, C. E. (1992). *Variance Components*. Wiley, New York.

Simmonds, N. W. (1979). *Principles of Crop Improvement*. London: English Language Book Society and Longman Group Ltd.

Snapp, S. (1999). Mother and baby trials: a novel trial design being tried out in Malawi. In *TARGET. The Newsletter of the Soil Fertility Research Network for Maize-Based Cropping Systems in Malawi and Zimbabwe*. Jan. 1999 issue. CIMMYT, Zimbabwe.

Snedecor, G. W. and Cochran, W. G. (1973). *Statistical Methods*, 6th edn. Iowa: Iowa State University Press.

Sperling, L. and Sceidegger, U. (1995). Participatory selection of beans in Rwanda: Results, Methods Institutional issues. *International Institute for Environment and Development Gatekeeper Series* No. 51. International Institute for Environmental and Development, London.

Sperling, L. and Loevinsohn, M. E. (1993). The dynamics of adoption: distribution and morality of bean varieties among small farmers in Rwanda. *Agricultural Systems* 41:441–453.

Virk, D. S., Chakraborty, M., Ghosh, J. and Harris, D. (2006). Participatory evaluation of horsegram (*Macrotyloma uniflorum*) varieties and their on-station responses to on-farm seed priming in eastern India. *Experimental Agriculture* 42:411–425.

Virk, D. S., Chakraborty, M., Ghosh, J., Prasad, S. C. and Witcombe, J. R. (2005). Increasing the client-orientation of maize breeding using farmer participation in eastern India. *Experimental Agriculture* 41:413–426.

Virk, D. S. and Witcombe, J. R. (2008). Evaluating cultivars in unbalanced on-farm participatory trials. *Field Crops Research* 106:105–115.

Witcombe, J. R. (1999). Do farmer participatory methods apply more to high potential areas than to marginal ones? *Outlook on Agriculture* 28:43–49.

Witcombe, J. R. (2002). A mother and baby trial system. In *Breeding Rainfed Rice for Drought-prone Environments: Integrating Conventional and Participatory Plant Breeding in South and South-east Asia. Proceedings of a DFID Plant Science Research Programme/IRRI Conference 12–15 March 2002, IRRI, Los Baños, Philippines*, 79–89 (Eds J. R. Witcombe, L. B. Parr and G. N. Atlin). Department for International Development (DFID) Plant Science Research Programme (PSP), Centre for Arid Zone Studies (CAZS), Bangor and International Rice Research Institute (IRRI), Manila.

Witcombe, J. R., Joshi, A., Joshi, K. D. and Sthapit, B. R. (1996). Farmers participatory crop improvement. I. Varietals selection and breeding methods and their impacts on biodiversity. *Experimental Agriculture* 32:445–460.

Witcombe, J. R., Joshi, K. D., Gyawali, S., Musa, A. A., Johansen, C., Virk, D. S. and Sthapit, B. R. 2005. Participatory Plant Breeding is better described as highly client-oriented plant breeding. I. Four indicators of client orientation in plant breeding. *Experimental Agriculture* 41:1–21.

Witcombe, J. R., Petre, R., Jones, S. and Joshi, A. (1999). Farmer participatory crop improvement. IV. The spread and impact of a rice variety identified by participatory varietal selection. *Experimental Agriculture* 35:471–487.