Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/sjar

Spanish Journal of Agricultural Research 2007 5(1), 16-24
ISSN: 1695-971X

Elicitation of subjective crop yield PDF

M. M. Clop-Gallart* and F. Juárez-Rubio

Department of Business Administration and Economic Management of Natural Resources. Universitat de Lleida. Avda. Rovira Roure, 191. 25198 Lleida. Spain

Abstract

The aim of this research is to investigate the coherence and reliability of subjective crop yield probability density functions (PDF) elicited from a series of interviews carried out on a wide group of farmers. Three different elicitation techniques were used: the Two-Step PDF estimation method, the Triangular distribution and the Beta distribution. Subjects who were interviewed gave both estimates for point crop yield (mean, highest possible, most frequent and lowest possible) and for the PDF based on interval estimates. To evaluate persistence, two concepts were used: time persistence and methodological persistence. The results are deemed valuable in order to determine the level of trust in the techniques applied in obtaining data, and in their effectiveness in designing a farm decision support system (DSS).

Additional key words: crop yield, decision maker judgment, estimation reliability, probability density function, subjective judgment.

Resumen

Elicitación de FDP subjetivas de rendimientos de cultivos

En este trabajo se investiga la coherencia y confiabilidad de estimaciones de funciones de densidad de probabilidad (FDP) subjetivas de rendimientos de cultivos realizadas por un amplio grupo de agricultores. Se utilizaron tres técnicas de elicitación diferentes: el método de estimación de FDP en dos pasos, la distribución Triangular y la distribución Beta. Los sujetos entrevistados ofrecieron estimaciones para los valores puntuales de rendimientos de cultivos (medio, máximo posible, más frecuente y mínimo posible) y para las FDP basadas en la estimación de intervalos. Para evaluar la persistencia, se utilizaron los conceptos de persistencia temporal y persistencia metodológica. Los resultados son interesantes para juzgar la adecuación de las técnicas de estimación de probabilidades subjetivas a los sistemas de ayuda en la toma de decisiones en agricultura.

Palabras clave adicionales: coherencia en estimaciones, elicitación subjetiva, función de densidad de probabilidad, juicios a priori, rendimientos de cultivos.

Introduction

The attempts to develop a Decision Support System (DSS) to enhance the outcome of farmers' decision making activities have been complicated by the necessity of assessing Probability Density Functions (PDF) for key variables such as crop yield. In a previous paper, Clop-Gallart and Juárez-Rubio (2006) discussed methodological difficulties that justify additional empirical research about farmers' subjective yield forecasts. In that paper, with the same empirical sample, authors analised persistence of point crop yield estimates (mean, highest possible, lowest possible and most frequent), as a condition of coherence in estimates.

This research is focused on the shape of farmer's PDF estimates. Subjective (personalistic) PDF estimations generally result in skewed probability density functions, and acceptable intervals for the mean value estimation (see Winkler, 1967, for a discussion on the techniques used to obtain data from interviews; Bessler, 1980; Grisley and Kellogg, 1983; Pease, 1992; Smith and Mandac, 1995). Nevertheless, variance is systematically underestimated. Variance underestimation is conjectured in the literature as a consequence of cognitive biases (e.g. Tversky and Kahneman, 1974; Smith and Mandac, 1995).

* Corresponding author: mclop@aegern.udl.es
Received: 24-07-06; Accepted: 29-11-06.

The possible non-normality of the PDF has important implications on the methods which should be used for optimization, as is well known in the decision making theory. In general, the literature has pointed out that crop yield PDF are skewed and non-normal (e.g. Day, 1965; Buccola, 1986; Gallagher, 1986, 1987; Nelson and Preckel, 1989; Nelson, 1990; Taylor, 1990; Kaylen and Koroma, 1991; Moss and Shonkwiler, 1993; Teigen and Thomas, 1995; Kaufmann and Snell, 1997; Ramírez, 1997; Goodwin and Ker, 1998). However, Just and Weninger (1999) have argued in favour of not rejecting normal distributions of crop yield.

A program for the systematic collection of data has been set up in order to elaborate a methodology for the elicitation of crop yield PDF, easily accessible to decision makers; and to accumulate experience on crop yield PDF. Every school year, agronomy students with a background in traditional statistics interviewed farmers to elicit their crop yield PDF.

The main objectives of the research were (1) to determine PDF functions reliability or persistence of the answers provided by the farmers and (2) to establish the coherence of the answers through the comparison of: a) the mean values derived from elicited PDF against the point mean values explicitly declared and b) between the mean values and variances derived from the different estimate PDF.

Research method

In 1999, fifty-two farmers were interviewed for the first time. Two students interviewed each farmer. Student number one carried out the interview with what was called the «first day questionnaire». Approximately two weeks later, another student interviewed the same farmer with the «second day questionnaire», which was organized in a different way (questionnaires available from the authors). A total of 104 interviews were carried out with 52 farmers. The second year interviews (2000) followed the same methodology described above. Forty-four different farmers were questioned, providing a total of 88 new interviews. Each farmer indicated the annual crop he would provide information for, depending on his own experience. Out of all the answers obtained, only those with the most number of responses (5 or more) were taken into account[1] (Table 1).

Table 1. Response from selected crops

Annual crop	Cases	
	1999	2000
Non irrigated barley	46	22
Non irrigated wheat	25	12
Irrigated wheat	13	5
Irrigated maize	20	10

Three elicitation techniques were used: Two-Step PDF estimation, Triangular and Beta distributions.

In the Two-Step PDF estimation technique, farmers first assessed the frequency (in percentage) for each of the five crop yield year classes: very poor yield years, poor yield years, normal yield years, good yield years and very good yield years. In this first step, farmers assessed frequencies at five ordinal interval classes. In the second step, farmers indicated the yield interval that they considered appropriate to describe a very poor yield year, a poor yield year, and so. The final product was a crop yields five interval histogram. Mean and variance values of the elicited distribution were then calculated. When x_i is the value given to the midpoint for interval i in the frequency estimation and f_i is the corresponding frequency, the mean (E) and the variance (VE) are given by the expressions:

$$E = \Sigma_i \, x_i f_i \qquad [1]$$

$$VE = \Sigma_i \, (x_i - E)^2 f_i \qquad [2]$$

In the Triangular distribution method, farmers specify the crop's highest possible yield (H), lowest possible yield (L), and most frequent yield (M) (Sonka and Patrick, 1984). The mean (T) and variance (VT) of the Triangular distribution are:

$$T = (1/3) \, (L + M + H) \qquad [3]$$

$$VT = (1/18) \, [(H-L)^2 + (M-L)(M-H)] \qquad [4]$$

In the Beta distribution method, the three previously declared values are fitted to a Beta function with a mean value and variance given by the known PERT approach expressions (Moskowitz and Bullers, 1979).

[1] Farmers come from a wide range of geographical areas with very different environmental and technological conditions from farm to farm. Although Pease (1992) has pointed that «Geographical location plays a larger role than crop in comparison of relative variability of yields», the aim of the present work did not lie in determining a given operational PDF for each of the crops analyzed, but to verify the persistence of the responses.

The mean (B) and variance (VB) of the Beta-PERT distribution are:

$$B = (1/6)\,(L + 4\,M + H) \quad [5]$$

$$VB = (1/36)\,(H - L)^2 \quad [6]$$

One of the requirements for coherence is persistence. Persistence was evaluated by the responses given by the farmers in two different time spans with the use of different subjective probability elicitation methods.

Subjects who were interviewed gave both estimates for point crop yield (mean, H, M and L) and for the PDF based on interval estimates. To evaluate persistence, a concept of persistence (which was called «time persistence») was used, based on measuring the difference between the estimates declared at two different time spans[2]. In the same way, in making a PDF estimate, persistence will be measured by its mean and variance, so that, if the farmer declares a PDF on the first interview and a different one two weeks later, both shall be considered as similar if the values they show are close in their respective means and variances, verifying the (time) persistence criterion.

Estimates of point values time persistence are available for both interview days. Results for the persistence of data, taken from the declared values of the mean (year 2000), lowest, highest and most frequent crop yields, have been analysed in Clop-Gallart and Juárez-Rubio (2006). Persistence criterion seemed to be verified for the mean m (year 2000), H and M declared crop yields, at least in a high percentage of the surveyed population. On the other hand, lowest possible crop yields L showed a great dispersion in relative differences.

A second meaning for the concept of persistence —«methodological persistence»— would be related to the estimation technique. In this way, if a declared value is maintained, not just in time, but also in the techniques it has been declared through, it will be supposed that the criterion of persistence will be verified. «Methodological» persistence is analysed by: (1) The comparison between the direct point estimate of the mean crop yield, made by the farmers, and the estimate calculated from the elicited PDFs; (methodological) persistence is verified if the declared mean crop yield and mean crop yield implicit in PDFs are similar. (2) The comparison between the mean crop yield values and their variances inferred from the different functional estimates, and the relationship between the results and the accessibility of extreme values.

Results

Mode relative position of crop yield PDF

The mode relative position of crop yield PDF in Two-Step PDF elicitation was studied. In general, in both 1999 and 2000 surveys, farmers estimated the frequency histograms corresponding to skewed PDF. For a given crop, farmers described as many modes to the right as to the left of the middle interval. Only in a relatively small percentage of cases, were modes indicated on the middle interval.

The position of the mode in the Triangular and Beta distributions elicited, is described by the $(M\text{-}L) / (H\text{-}L)$ relationship (where M is the mode, L the lowest value and H the highest value). Results (1999 and 2000 surveys) show that farmers generally described skewed crop yield PDF with a relatively small percentage of modes in the centre of the $(H\text{-}L)$ range, where $(M\text{-}L) / (H\text{-}L) = 0.5$.

Time persistence of estimates

PDF estimates for both survey days are available every year. The following analyses the persistence of data, evaluated by the value of the mean crop yield and the variance of the crop yields calculated through Equations [1] to [6].

The T test was carried out in order to compare the calculated mean crop yield answers for each day (mean values calculated from first day answers *vs.* mean values calculated from second day answers). In relation to the data for non irrigated crops in 1999, the calculated means showed correlation coefficients higher than 0.97 on all three PDFs, without any statistically significant differences (p values higher than 0.3 in all cases). In irrigated crops, the correlation was higher than 0.99 for all cases, with p values higher than 0.06, without any statistically significant differences. Similar results were obtained for the T test and correlation coefficients with year 2000 data.

[2] To avoid the biases pointed out by Bland and Altman (1995, 1999) if d_1 and d_2 are the values to be compared, as in, for example, the estimates made by a decision maker on the first and second days, or two values corresponding to different PDFs; relative differences throughout this research will be expressed thus: $(d_1 - d_2) / [(d_1 + d_2) / 2]$.

Table 2. Relative differences in calculated mean crop yields (*E, T* and *B*, day 1 *vs* day 2)[1]

Case	Year	*n*		Mean		SD		Median		Min		Max		K-S[2]	
		NIC[3]	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC
E	1999	68	33	–0.001	–0.014	0.081	0.066	0	–0.004	–0.33	–0.20	0.18	0.15	0.180	0.322
E	2000	33	14	–0.007	–0.021	0.117	0.069	0	0	–0.47	–0.24	0.16	0.07	0.205	0.191
T	1999	68	33	–0.011	–0.018	0.089	0.096	0	0	–0.31	–0.36	0.17	0.22	0.009	0.101
T	2000	33	14	–0.015	–0.029	0.074	0.077	0	–0.030	–0.21	–0.17	0.23	0.16	0.254	0.553
B	1999	68	33	–0.011	–0.023	0.085	0.095	0	0	–0.30	–0.34	0.16	0.16	0.001	0.054
B	2000	33	14	–0.020	–0.040	0.078	0.066	0	–0.037	–0.19	–0.12	0.16	0.12	0.159	0.976

[1] *E:* Two-Step mean values. *T:* Triangular mean values. *B:* Beta-PERT mean values. [2] Kolgomorov-Smirnov test for a normal distribution. [3] NIC: non irrigated crops. IC: irrigated crops.

The following analyses the relative differences between the mean crop yields calculated with all three PDF estimates.

In general, mean crop yield values calculated (*E*) from PDFs in Two-Steps for both days were very similar (Table 2). The mean for the relative differences ranged between –0.7% and –0.1% for non irrigated crops and between –2.1% and –1.4% for irrigated crops, with medians being null in non irrigated crops and practically null in irrigated crops. The SDs for that variable were 8.1% and 11.7% in non irrigated crops and 6.6% and 6.9% in irrigated crops. The similarity of both distributions is obvious. There were no differences between the different populations surveyed (Wilcoxon, $\alpha = 0.05$). Similar results were found for the calculated mean crop yield values of the Triangular (*T*) and Beta-PERT (*B*) functions.

There were no significant statistical differences between the mean crop yield results obtained using PDF estimates, showing a persistence in the values declared on each day.

Relative differences between calculated variances were compared using Equations [2], [4] and [6], from the results obtained for first and second days on 1999 and 2000 (Table 3).

In general, Two-Step PDF estimates for non irrigated crops provided very similar values to the variances calculated from the data obtained on both survey days. Nevertheless, there were some cases that showed great differences between the estimates obtained on one day or another, measured using variance. These cases probably showed evidence in some farmers to be biased and to anchor interval values with the mean, having little regard for extreme values.

In the case of relative differences for irrigated crops, unfavourable results in the year 2000 were obtained for the persistence hypothesis; but the general trend was conceptually similar to the one discussed for non irrigated crops.

A result similar to the one discussed for the Two-Step estimation process was noticeable in Triangular and Beta estimations; with very favourable results for the persistence hypothesis —in the case of non irrigated and irrigated crops in 1999— and some problematic answers —in the case of irrigated crops in 2000—, probably magnified by the small size of the sample.

Table 3. Relative differences in calculated crop yield variances (*VE, VT*, *VB*, day 1 *vs* day 2)[1]

Case	Year	*n*		Mean		SD		Median		Min		Max		K-S[2]	
		NIC[3]	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC
VE	1999	68	33	0.001	0.057	0.596	0.466	0	0	–1.93	–0.61	1.41	1.73	0.038	0.011
VE	2000	33	14	0.031	0.267	0.267	0.597	0	0	–0.39	–0.63	0.50	1.62	0.295	0.175
VT	1999	68	33	–0.003	–0.057	0.516	0.369	0	0	–1.20	–0.92	1.78	1.42	0.012	0.026
VT	2000	33	14	0.008	–0.233	0.518	0.819	–0.008	–0.135	–1.15	–1.26	1.30	1.22	0.047	0.816
VB	1999	68	33	–0.008	–0.057	0.515	0.363	0	0	–1.20	–0.92	1.78	1.42	0.009	0.009
VB	2000	33	14	0.009	–0.281	0.514	0.891	0	–0.153	–1.15	–1.89	1.34	1.20	0.006	0.902

[1] *VE:* Two-Step variances. *VT:* Triangular variances. *VB:* Beta-PERT variances. [2] Kolgomorov-Smirnov test for a normal distribution. [3] NIC: nonirrigated crops. IC: irrigated crops.

For all analysed cases, samples for each day appeared to belong to the same population (Wilcoxon, $\alpha = 0.05$).

Methodological persistence

Mean declared crop yield values *m* on the second day of 1999 and the first and second days of 2000 were available. Table 4 shows the declared mean crop yield *m* values compared to the mean values calculated from the Two-Step estimation method —using Equation [1]—, and the mean values calculated with the Triangular and Beta estimates —using Equations [3] and [5], respectively[3].

In the Two-Step method, the mean for relative differences in non irrigated crops tended to be small, around 4% to 6% (with acceptable SDs). The median for the same measure had values ranging between 2% and 6%. This factor indicated a slight underestimation of mean crop yield when expressed through the Two-Step estimation method, in relation to the directly declared mean crop yield value.

For irrigated crops, the mean of relative differences ranged between –2.9% and –0.8%; the median ranged between –1.1% and 0%. These negative values probably led to a slight overestimation of the mean crop yield when expressed through the Two-Step estimation method. SD ranged between 13.2% and 18.9%.

With the Triangular approximation, in the case of non irrigated crop yields, the mean for the relative difference reached values between –2% and –1.7%, the median had values between –2% and 0%. Relative difference followed a similar pattern to the Two-Step estimation for irrigated crops. In this case, relative difference of the declared mean and the mean calculated by the Triangular method lay between –1% and –0.4%, with the median lying between –2.3% and 0%. SD ranged from 7.5% to 11.8%, for non irrigated crops, and 5.8% to 9.9% for irrigated crops.

With regard to the Beta approximation on non irrigated crops, the mean for relative differences ranged from –2.6% to –1.5%; on irrigated crops, it ranged from –2.9% to –1.5%. The median for both non irrigated and irrigated crops laid between –2.3% and 0%. SDs varied from 8.6% and 10.3% (for non irrigated crops) and 5.4% and 8.7% (for irrigated crops).

Table 4. Declared mean crop yield (*m*) and PDF calculated mean crop yield (*E, T* and *B*)[1] relative differences

Case	Day and year	*n*		Mean		SD		Median		Min		Max		K-S[2]	
		NIC[3]	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC
m vs E	day 2, 1999	68	33	0.040	–0.008	0.125	0.132	0.042	0	–0.31	–0.24	0.45	0.50	0.600	0.631
m vs E	day 1, 2000	33	14	0.048	–0.029	0.149	0.171	0.022	–0.011	–0.31	–0.45	0.47	0.23	0.940	0.876
m vs E	day 2, 2000	33	14	0.059	–0.020	0.13	0.189	0.059	–0.008	–0.23	–0.50	0.46	0.23	0.994	0.981
m vs T	day 2, 1999	70	33	–0.02	–0.004	0.118	0.099	0	0	–0.33	–0.29	0.18	0.2	0.020	0.261
m vs T	day 1, 2000	33	14	–0.019	–0.01	0.092	0.07	–0.004	0	–0.29	–0.17	0.18	0.08	0.283	0.605
m vs T	day 2, 2000	33	14	–0.017	–0.009	0.075	0.058	–0.02	–0.023	–0.24	–0.09	0.18	0.1	0.388	0.765
m vs B	day 2, 1999	70	33	–0.016	–0.015	0.103	0.087	0	0	–0.32	–0.31	0.4	0.19	0.022	0.042
m vs B	day 1, 2000	33	14	–0.015	–0.019	0.099	0.054	–0.013	0	–0.22	–0.13	0.18	0.05	0.309	0.578
m vs B	day 2, 2000	33	14	–0.026	–0.029	0.086	0.061	–0.023	–0.023	–0.31	–0.12	0.18	0.09	0.156	0.99

[1] *E:* Two-Step mean values. *T:* Triangular mean values. *B:* Beta-PERT mean values. [2] Kolgomorov-Smirnov test for a normal distribution. [3] NIC: non irrigated crops. IC: irrigated crops.

[3] If *m1* is the declared mean crop yield value, and *m2* is the mean value calculated from the Two-Step estimation method or the Triangular and Beta approximations (*m2* = *E* or *T* or *B*), the relative difference will be defined as (*m1-m2*) / [(*m1*+*m2*)/2].

Negative values found for the mean of relative differences —as a result of both the Triangular and Beta approximations— seemed to indicate a slight overestimation of the mean crop yield (regarding the directly declared mean).

For all analysed cases, samples for each day appeared to belong to the same population (Wilcoxon, $\alpha = 0.05$).

Differences between means and variances of crop yields calculated using various PDFs

An interesting practical issue is whether the results obtained by the Triangular and Beta approximations predict values that are very different to those predicted by the Two-Step estimation method, as the first two methods would, initially, appear to be easier ways of interviewing farmers.

Comparing the Triangular and Beta methods holds no empirical interest because differences between the mean and variance values, which are estimated from the same set of values L, M and H (the lowest, most frequent and highest possible crop yields), are mathematically determined by functional forms. In estimating the mean crop yield, the difference given by both functions is the result of using Equation [7] for the mean value, and Equation [8] for the variance.

$$T - B = (1/6)\,(H + L - 2M) \qquad [7]$$

$$VT - VB = (1/36)\,[(H - M)^2 + (M - L)^2] \qquad [8]$$

The difference between the means reaches its lowest value with $M = (1/2)\,(H + L)$ as the variance. This result shows that, when the most frequent value M coincides with the midpoint of L and H, mean and variance values of both approximations also coincide. The difference between variances [8] is always positive, therefore: $VT \geq VB$.

The T test was applied to compare the mean crop yields obtained with the Two-Step estimation method against both the Triangular and Beta PDF calculated mean crop yields, for both days of interviews and years 1999 and 2000. For 1999 and non irrigated crops, the resulting correlation coefficients were over 0.93 in all cases, but p values were under 0.05 in all cases; therefore, it may not be statistically established the similarity between both populations. With regards to irrigated crops, correlation coefficients were over 0.99 in all cases and p values were over 0.23 in all cases; this would indicate that there were no statistically significant differences between the calculated mean crop yield using the Two-Step estimation method and the same results calculated from the Triangular and Beta approximations. Nevertheless, this conclusion could be due to the high variances obtained. Results for 2000 were similar to the results for 1999.

Table 5 summarises the differences found between the mean crop yield from the Two-Step estimation method (E) and the Triangular (T) and Beta (B) approximations.

For non irrigated crops, both the Triangular and Beta functions showed a tendency to produce mean crop yield estimates that were slightly higher than the estimates deduced from the Two-Step estimation method, although this difference between them was not signifi-

Table 5. Relative differences for calculated mean crop yield

Year and day	*n*		Mean		SD		Median		Min		Max		K-S[1]	
	NIC[2]	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC
T-E[3]														
1999 day 1	68	33	0.053	–0.008	0.146	0.123	0.043	–0.012	–0.35	–0.25	0.38	0.33	0.612	0.423
1999 day 2	68	33	0.063	–0.003	0.133	0.115	0.035	–0.014	–0.31	–0.18	0.41	0.32	0.453	0.960
2000 day 1	33	14	0.067	–0.018	0.132	0.193	0.040	0	–0.13	–0.46	0.38	0.26	0.665	0.878
2000 day 2	33	14	0.075	–0.010	0.133	0.181	0.077	0.005	–0.13	–0.46	0.51	0.25	0.835	0.963
B-E[3]														
1999 day 1	68	33	0.053	–0.002	0.139	0.131	0.052	–0.019	–0.30	–0.27	0.45	0.34	0.731	0.801
1999 day 2	68	33	0.063	0.007	0.126	0.117	0.048	–0.014	–0.31	–0.21	0.44	0.32	0.744	0.867
2000 day 1	33	14	0.063	–0.011	0.13	0.176	0.041	–0.019	–0.13	–0.40	0.37	0.26	0.673	0.977
2000 day 2	33	14	0.076	0.008	0.148	0.171	0.068	0.019	–0.17	–0.40	0.58	0.26	0.810	0.990

[1] Kolgomorov-Smirnov test for a normal distribution. [2] NIC: non irrigated crops. IC: irrigated crops.[3] *E:* Two-Step mean values. *T:* Triangular mean values. *B:* Beta-PERT mean values.

cant. For irrigated crops, mean crop yield deduced from the histogram was higher than the value obtained using the Triangular estimation process, and there was no clear tendency towards a comparison with the Beta results.

Variances calculated by the Two-Step estimation method and the Triangular and Beta-PERT approximations differed a great deal. In the case of the difference between the crop yield variances deduced from the Two-Step method and the Triangular approximation (*VE-VT*), the relative differences were higher than 34.4% in all cases, with SDs of up to 101%. In the comparison between the Two-Step estimates and the Beta approximation (*VE-VB*), the relative differences were less favourable than in the previous case; the minimum mean for the relative differences being 48.1%, with a SD of up to 100.3%. Results for irrigated crops were slightly more favourable than for non irrigated crops.

Differences between extreme crop yield values

The aforementioned result in the variance may be a consequence of both Triangular and Beta functional specifications, and possible difficulties in the estimation of extreme values, especially *L* (lowest possible crop yield). In order to examine this last issue, the extremes in the density function estimated in Two-Steps were compared to the point estimates that serve to define the Triangular and Beta approximations. In order to achieve this, two relative differences indexes were created (I_m and I_M). If *r* is the lower extreme in the interval defined as «very poor crop yields» and *s* is the upper extreme in the interval defined as «very good crop yields», we have:

$$I_m = (r - L) / [(r + L)/2] \quad [9]$$

$$I_M = (s - H) / [(s + H)/2] \quad [10]$$

Results obtained from the relative differences of indexes [9] and [10] are shown in Table 6.

It can be observed that subjects tended to give relatively stable estimates for the maximum crop yield and/or the higher extremes of the range. On the other hand, very different estimates for minimum values and/or the lower extremes of the range were given.

Discussion

The similarity between the mean and the variance of estimated functions for each of the days that the survey was conducted on, was used as the criterion to examine the time persistence of PDF estimations given by farmers. It has been found that differences were relatively small, except for the variance on irrigated crops in 2000 (with a smaller number of observations). These results seem to suggest a great stability in the mental image that farmers have of PDFs.

In order to calibrate methodological persistence, the direct estimation of mean crop yields and the calculated from estimated PDFs was compared. The differences were found to be small. In the case of the Two-Step estimation method, PDF in non irrigated crops tended

Table 6. Values of indexes I_m and I_M[1]

Year and day	*n*		Mean		SD		Median		Min		Max	
	NIC[2]	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC	NIC	IC
Index I_m												
1999 day 1	68	33	–0.51	–0.54	0.72	0.71	–0.20	–0.29	–2	–2	0.67	0
1999 day 2	68	33	–0.57	–0.32	0.69	0.49	–0.38	–0.19	–2	–2	0.86	0.18
2000 day 1	33	14	–0.57	–0.66	0.61	0.72	–0.4	–0.46	–2	–2	0	0.4
2000 day 2	33	14	–0.58	–0.49	0.65	0.71	–0.54	–0.25	–2	–2	0.46	0.29
Index I_M												
1999 day 1	68	33	0.052	0.12	0.12	0.11	0.024	0.09	–0.18	0	0.46	0.4
1999 day 2	68	33	0.034	0.09	0.13	0.11	0.038	0.07	–0.48	–0.13	0.31	0.4
2000 day 1	33	14	0.12	0.09	0.18	0.13	0.08	0.07	–0.15	–0.06	0.71	0.45
2000 day 2	33	14	0.09	0.05	0.18	0.08	0	0.03	–0.15	–0.07	0.71	0.24

[1] I_m relative differences between «very poor» crop yield and the lowest possible *L*. I_M relative differences between «very good» crop yield and the highest possible *H*. [2] NIC: non irrigated crops. IC: irrigated crops.

to give a slight underestimation of the mean in relation to the declared; whereas the opposite occurred in irrigated crops. Using the Triangular and Beta approximations, a slight overestimation of the mean crop yield, as opposed to the declared crop yield, was given.

This overestimation of the mean crop yield (shown in the negative figures of the relative differences in Table 4) could be interpreted as the result of a broader range on the frequency histogram (relating to the Two-Step estimation on irrigated crops, and the Triangular and Beta approximations on all crops), contrary to what is cited in the literature regarding variance underestimations. Values were so small, that it is neither possible nor useful to confirm that they may be anything other than zero (meaning that the declared and the calculated crop yields coincide). Nevertheless, it was significant that no general trend of positive relative differences had been detected, which could endorse the tendency to underestimate the variance.

When mean crop yields (calculated by the Two-Step estimation method) were compared to mean crop yields (obtained by the Triangular and Beta approximation methods), similar results were obtained; this showed a great methodological persistence in measuring this variable.

Crop yield variances were very different, possibly due to the functional specifications and problems regarding the estimation of extremes, especially lower ones.

The methods used for the PDF elicitation were clearly acceptable by both the farmers interviewed and the students who carried them out. They are very simple elicitation methods that can be incorporated into DSS with relative ease.

From the data obtained from the interviews, it was possible to conclude that, in general, it cannot be assumed a normal PDF to describe crop yield distributions. The PDF obtained by the farmers were clearly skewed and non normal.

There is a large number of farmers who showed a great accuracy and reliability in their responses for the first and second day interviews. This circumstance was interpreted as an indication of a good knowledge of PDF, although this hypothesis requires additional evidence.

Results show, in general, that simple methods as Two-Step, Beta or Triangular PDF elicitation techniques can be used to estimate farmers' crop yield PDF. The estimations tend to be time and methodologically persistent. The main problem seems to be in the lower possible crop yield estimation, bias that requires future research for its explanation and correction.

References

BESSLER D.A., 1980. Aggregated personalistic beliefs on yields of selected crops estimated using ARIMA processes. Am J Agr Econ 62(4), 666-674.

BLAND J.M., ALTMAN D.G., 1995. Comparing methods of measurement: why plotting difference against standard method is misleading. Lancet 346, 1085-1087.

BLAND J.M, ALTMAN D.G., 1999. Measuring agreement in method comparison studies. Stat Methods Med Res 8, 135-160.

BUCCOLA S.T., 1986. Testing for nonnormality in farm net returns. Am J Agr Econ 68, 334-343.

CLOP-GALLART M.M., JUÁREZ-RUBIO F., 2006. Persistence of point crop yield subjective estimates. Span J Agric Res 4(3), 202-206.

DAY R.H., 1965. Probability distributions of field crop yields. J Farm Econ 47(3), 713-741.

GALLAGHER P., 1986. U.S. corn yield capacity and probability: estimation and forecasting with nonsymmetric disturbances. N Cent J Agr Econ 8(1), 109-122.

GALLAGHER P., 1987. U.S. soybean yields: estimation and forecasting with nonsymmetric distributions. Am J Agr Econ 69(4), 796-803.

GOODWIN B.K., KER A.P., 1998. Nonparametric estimation of crop yield distributions: implications for rating group-risk crop insurance contracts. Am J Agr Econ 80(1), 139-153.

GRISLEY W., KELLOGG E.D., 1983. Farmers' subjective probabilities in Northern Thailand: An elicitation analysis. Am J Agr Econ 65(1), 74-82.

JUST R.E., WENINGER Q., 1999. Are crop yields normally distributed? Am J Agr Econ 81(2), 287-304.

KAUFMANN R.K., SNELL S.E., 1997. A biophysical model of corn yield: integrating climatic and social determinants. Am J Agr Econ 79(1), 178-190.

KAYLEN M.S., KOROMA S.S., 1991. Trend, weather variables, and the distribution of U.S. corn yields. Rev Agr Econ 13(2), 249-258.

MOSKOWITZ H., BULLERS W.I., 1979. Modified PERT versus fractile assessment of subjective probability distributions. Organ Behav Hum Perform 24(2), 167-194.

MOSS C.B., SHONKWILER J.S., 1993. Estimating yield distributions with a stochastic trend and nonnormal errors. Am J Agr Econ 75(4), 1056-1062.

NELSON C.H., 1990. The influence of distributional assumptions of the calculation of crop insurance premia. N Cent J Agr Econ 12, 71-78.

NELSON C.H., PRECKEL P.V., 1989. The conditional beta distribution as a stochastic production function. Am J Agr Econ 71(2), 370-378.

PEASE J.W., 1992. A comparison of subjective and historical crop yield probability distributions. S J Agr Econ 24(2), 23-32.

RAMÍREZ O.A., 1997. Estimation and use of a multivariate parametric model for simulating heteroskedastic, correlated, nonnormal random variables: the case of corn

belt corn, soybean, and wheat yields. Am J Agr Econ 79(1), 191-205.

SMITH J., MANDAC A.M., 1995. Subjective versus objective yield distributions as measures of production risk. Am J Agr Econ 77(1), 152-161.

SONKA S.T., PATRICK G.F., 1984. Risk management and decision making in agricultural firms. In: Risk management in agriculture (Barry P.J., ed). Iowa State University Press, Ames, Iowa. pp. 95-115.

TAYLOR C.R., 1990. Two practical procedures for estimating multivariate nonnormal probability density functions. Am J Agr Econ 72(1), 210-218.

TEIGEN L.H., THOMAS M.Jr., 1995. Weather and yield, 1950-94: Relationships, distributions and data. USDA Commercial Agriculture Division, Economic Research Service, Washington, DC.

TVERSKY A., KAHNEMAN D., 1974. Judgment under uncertainty: heuristics and biases. In: Judgment under uncertainty: heuristics and biases (Kahneman D., Slovic P., Tversky A., eds). Cambridge University Press, Cambridge. pp. 3-20.

WINKLER R.L., 1967. The assessment of prior distributions in bayesian analysis. J Am Stat Ass 62, 776-800.