*Indian J. Genet.*, **49**(2): 203–207 (1989)

# VARIETAL REGRESSION ANALYSES OF INCOMPLETE DATA ON ONION AND TOMATO BY THREE DIFFERENT METHODS OF FITTING CONSTANTS

K. A. BALAKRISHNAN,* J. P. JAIN AND O. P. DUTTA

*Indian Agricultural Statistics Research Institute*
*New Delhi 110012*

**(Received: April 20, 1987; accepted: August 30, 1988)**

## ABSTRACT

**This paper presents analyses of incomplete variety × environment data in respect of two vegetable crops, onion and tomato, by three different methods of fitting constants, viz. Freeman's two-stage least square analysis, Patterson's FITCON analysis, and Digby's modified joint regression analysis. The relative rankings of varieties based on FITCON analysis mean, and modified joint regression analysis are the same for both tomato and onion varieties and the sensitivities of the varieties also do not differ markedly from unity. Hence for varietal analysis in vegetable crops with incomplete data, the FITCON procedure can be used with equal advantage vis-a-vis the operationally more difficult modified joint regression method.**

**Key words:** Varietal regression analyses, incomplete data, fitting constants, modified joint regression analysis.

In regional testing programmes, a sample of promising genotypes is tested each year at different locations. When the results of such trials are combined, it is often seen that data for all genotypes are not available from all locations due to various reasons. This leads to incomplete variety × environment (locations) tables. In such situations, comparison based on means alone favours the varieties which happened to be allocated at better than average locations. Consequently, compensation has to be made in the estimation of varietal means for environments where some of the varieties were missing. The purpose of this paper, besides reporting of results, is to make a critical appraisal of the three different methods of Freeman, Patterson and Digby [1–3], available for analysis of incomplete data, with reference to two important vegetable crops, viz. onion and tomato.

## MATERIALS AND METHODS

### DATA

The data for this study relate to varietal trials in onion and tomato, conducted under the All-India Coordinated Vegetable Improvement Project during 1982-83. The trials at each location were laid in randomized block designs with three replicates.

*Present address: Project Directorate of Vegetable Research, IARI, New Delhi 110012.

The mean yields recorded at different locations are presented in Tables 1 and 2. Out of 10 trials in onion, complete data were received only from six locations while in tomato data were incomplete for 4 out of 11 locations. Of the six onion varieties tested, variety VL-67 is from Almora; Line-102 and Pusa Red from IARI, Delhi; N-53 and N-2-4-1 from Rahuri; and Udaipur 103 from Udaipur in Rajasthan. In tomato, three varieties, i.e. Punjab Chhuhara, Punjab Kesari and S-12 are from Ludhiana; two varieties, Sel-152 and Roma, from Katrain; variety La Bonita from Delhi; and KS-2 from Kalyanpur (Kanpur).

## THE THREE METHODS

One method of dealing with incomplete variety × environment data is the two-stage least squares procedure of Freeman [1]. The first stage involves estimation of parameters $\mu_i$ and $\theta_j$ of varieties and environments in the model:

$$Y_{ij} = \mu_i + \theta_j + e_{ij} \quad (1)$$

which are used to fill the gaps in the incomplete table and the second stage consists of the usual regression analysis of the completed table using the model of Eberhart and Russell [4]:

$$Y_{ij} = \mu_i + \beta_i\theta_j + e_{ij} \quad (2)$$

where $Y_{ij}$— mean yield of i-th variety in j-th environment, $\mu_i$— mean of i-th variety over all the environments in which it is present, $\theta_j$— effect of j-th environment, and $\Sigma\theta_j = 0$, $e_{ij}$— error term, and $\beta_i$— regression coefficient for i-th variety.

The estimates of $\beta_i$ obtained by this method will be biased towards one as this method uses the additive model (1) to estimate the missing values in order to fit the nonadditive model (2). Further, this method makes no adjustment in the variety means for the missing environments.

The other method considered here is the 'FITCON' or 'fitting constants' method of Patterson [2]. In this method also, the estimates of $\mu_i$ and $\theta_j$ are obtained by fitting additive model (1) as in Freeman's method, but the variety means are adjusted for the environments in which they were not tested. The adjustment, however, is not quite appropriate as it does not take into account the differential response or sensitivity of different varieties to different environments. Digby's [3] modified regression analysis makes such an adjustment possible. This is the third method of analysis of incomplete data considered here.

## RESULTS AND DISCUSSION

The estimated environmental effects for onion and tomato yield trials as obtained by FITCON and modified joint regression analyses are given in Tables 1 and 2, respectively. Of the 10 environments in onion and 11 in tomato, 4 are rated slightly higher in each case by the modified regression analysis. The estimates for variety means obtained by the three analyses along with their estimates of sensitivity are given in Table 3 (onion) and Table 4 (tomato).

**Table 1. Average yield of onion bulbs and estimated environmental (location) effects**

| Location | Yield of different varieties (tonnes/ha) | | | | | | Estimated environmental effects | |
|---|---|---|---|---|---|---|---|---|
| | VL-67 | Line 102 | Pusa Red | N-53 | N-2-4-1 | Udaipur 103 | FITCON analysis | modified joint regression analysis |
| Hissar | 12.5 | 19.2 | — | 13.7 | 16.0 | 17.0 | −10.51 | −10.55 |
| Pantnagar | 19.2 | 23.9 | 12.8 | 19.0 | 19.0 | 15.5 | −8.11 | −8.36 |
| Sabour | 16.1 | 15.2 | 19.6 | 15.0 | — | 15.7 | −9.63 | −9.64 |
| Almora | 21.0 | 19.8 | 28.7 | 25.3 | 19.8 | 24.7 | −3.14 | −2.99 |
| Solan | 18.2 | — | 16.0 | 15.0 | — | 13.1 | −10.30 | −10.42 |
| Ludhiana | 24.8 | 21.1 | 18.3 | 23.7 | 34.8 | 27.2 | −1.35 | −1.50 |
| Delhi | 26.9 | 36.6 | 39.5 | — | 28.6 | — | 6.74 | 7.27 |
| Rahuri | 21.0 | 33.0 | 39.0 | 37.5 | 43.8 | 31.7 | 7.98 | 8.33 |
| Junagadh | 22.0 | 40.3 | 34.5 | 48.9 | 53.3 | 42.0 | 13.81 | 14.32 |
| Bangalore | 48.0 | 38.6 | 45.8 | 36.7 | 30.8 | 45.6 | 14.53 | 13.55 |

Onion variety VL-67 gave the same average by all the three methods of analysis as it was included in all environments. The adjustment is maximum for variety N-2-4-1, which was not tested in two environments which had high negative effects. No appreciable difference was observed between the estimates of fitting constants

**Table 2. Average yield of marketable fruits of tomato and estimated environmental (location) effects**

| Location | Yield of different varieties (tonnes/ha) | | | | | | | Estimated environmental effects | |
|---|---|---|---|---|---|---|---|---|---|
| | Punjab Chhu-hara | Punjab Kesari | La Bonita | Ks-2 | S-12 | Sel-152 | Roma | FITCON analysis | modified joint regression analysis |
| Bhubaneswar | 3.7 | 2.9 | 2.6 | 2.5 | — | 6.7 | 6.0 | −18.20 | −17.65 |
| Almora (ARU) | 13.9 | 13.5 | 16.1 | 12.4 | 16.9 | 9.7 | 16.3 | −8.65 | −8.58 |
| Sabour | 22.3 | 18.2 | — | 18.4 | 20.8 | — | 16.8 | −3.60 | −3.64 |
| Ludhiana | 25.1 | 21.0 | 21.1 | 12.6 | 26.9 | 12.5 | 13.9 | −3.75 | −4.16 |
| Solan | — | 20.1 | 20.4 | 28.5 | — | 22.3 | 21.0 | 0.58 | 0.70 |
| Hissar | 24.5 | 18.6 | 28.9 | 16.5 | 29.8 | 20.8 | 18.8 | −0.24 | −0.66 |
| Rahuri | 26.6 | 20.9 | 27.7 | 23.0 | 23.2 | 20.3 | 23.8 | 0.87 | 0.73 |
| Junagadh | 29.5 | 18.8 | 29.4 | 25.9 | 28.3 | 20.3 | 27.3 | 2.90 | 2.81 |
| Pantnagar | 22.6 | 20.0 | 30.8 | 30.0 | 30.0 | 39.6 | 27.4 | 5.84 | 5.72 |
| Kalyani | 30.6 | 22.8 | 34.8 | 32.7 | 30.2 | 23.0 | 30.8 | 6.49 | 6.41 |
| Katrain | 40.5 | 41.6 | 31.7 | — | 44.4 | 37.7 | 48.0 | 17.76 | 18.37 |

and modified regression. This is due to lack of difference in the estimates of sensitivity of different varieties. The two best varieties in order of their performance are N-2-4-1 and Pusa Red.

**Table 3. Estimated variety parameters for onion (yield in tonnes/ha)**

| Variety | Unadjusted mean | 2-Stage L.S. analysis | | FITCON analysis | | Modified joint regression analysis | | |
|---|---|---|---|---|---|---|---|---|
| | | mean | sensitivity | mean | rank | mean | rank | sensitivity |
| VL – 67 | 23.0 | 23.0 | 0.65 | 23.0 | 6 | 23.0 | 6 | 0.66 |
| Line – 102 | 27.5 | 26.5 | 0.91 | 26.4 | 5 | 26.5 | 5 | 0.91 |
| Pusa Red | 28.2 | 27.0 | 1.10 | 27.1 | 2 | 27.0 | 2 | 1.09 |
| N – 53 | 26.0 | 26.9 | 1.15 | 26.8 | 3 | 27.0 | 3 | 1.16 |
| N-2-4-1 | 30.8 | 28.2 | 1.05 | 28.3 | 1 | 28.1 | 1 | 1.07 |
| Udaipur-103 | 25.8 | 26.7 | 1.14 | 26.6 | 4 | 26.8 | 4 | 1.13 |

In tomato also, two varieties, Punjab Kesari and Roma, which were present in all the 11 environments, do not differ in the estimated means derived from all three analyses. Maximum adjustment was required for variety S-12, which was not tested in the first environment with high negative effect in FITCON as well as modified regression analysis. Further, variety KS-2 shows maximum difference in the mean values obtained by these two methods of analysis due to its high sensitivity to environment. Despite this, there is no difference in the relative ranking of varieties by the FITCON and modified regression analyses. Among the varieties tested, S-12 and La Bonita have the highest score.

From the foregoing empirical results it is obvious that for valid comparison, it is necessary to compensate the varietal means for the missing environments

**Table 4. Estimated variety parameters for tomato (yield in tonnes/ha)**

| Variety | Unadjusted mean | 2-Stage L.S. analysis | | FITCON analysis | | Modified joint regression analysis | | |
|---|---|---|---|---|---|---|---|---|
| | | mean | sensitivity | mean | rank | mean | rank | sensitivity |
| Punjab Chhuhara | 23.9 | 24.0 | 0.97 | 24.0 | 3 | 24.0 | 3 | 0.96 |
| Punjab Kesari | 19.8 | 19.8 | 0.92 | 19.8 | 7 | 19.8 | 7 | 0.92 |
| La Bonita | 24.4 | 24.0 | 0.89 | 24.0 | 2 | 24.0 | 2 | 0.87 |
| KS – 2 | 20.2 | 22.4 | 1.21 | 22.0 | 5 | 22.5 | 5 | 1.25 |
| S – 12 | 27.8 | 26.0 | 0.95 | 25.9 | 1 | 26.1 | 1 | 0.92 |
| Sel – 152 | 21.3 | 20.9 | 0.97 | 20.9 | 6 | 20.9 | 6 | 0.98 |
| Roma | 22.7 | 22.7 | 1.15 | 22.7 | 4 | 22.7 | 4 | 1.15 |

(locations). This is particularly important when these environments have high positive or negative effect. The relative ranking, based on the adjusted mean under the FITCON and modified regression analyses is the same for both tomato and onion. This is because the sensitivity of varieties does not differ markedly from unity except for one onion variety. The FITCON analysis can be used with almost equal advantage in comparison with the operationally more difficult procedure of modified regression analysis unless the varieties differ significantly in their environmental response, or the nonorthogonality measured by the number of missing values is large.

## ACKNOWLEDGEMENT

Thanks are due to Dr. K. N. Agarwal, Senior Scientist, IARI, for useful suggestions given during the investigation.

## REFERENCES

1. G. H. Freeman. 1975. Analysis of interaction in incomplete two-way tables. Appl. Statist., **21**: 46–55.
2. H. D. Patterson. 1978. Routine least square estimation of variety means in incomplete tables. J. Natl. Inst. Agric. Bot., **14**: 401–412.
3. P. G. N. Digby. 1979. Modified joint regression analysis for incomplete variety × environment data. J. Agric. Sci. Camb., **93**: 81–86.
4. S. A. Eberhart and W. A. Russell. 1966. Stability parameters for comparing varieties. Crop Sci., **6**: 36–40.