Spanish Journal of Agricultural Research
14(2), e0906, 14 pages (2016)
eISSN: 2171-9292
http://dx.doi.org/10.5424/sjar/2016142-8175
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)

RESEARCH ARTICLE

OPEN ACCESS

Relations between zero-inflated variables in trials with horticultural crops

Alessandro D. Lúcio[1], Luis F. Nunes[2], Francisco Rego[3] and Maurício P. B. Pasini[4]

[1] *Federal University of Santa Maria, Rural Science Center, Crop Science Department. Santa Maria, Rio Grande do Sul, Brazil.* [2] *Polytechnic Institute of Bragança, School of Agriculture, Mountain Research Centre (CIMO). Bragança, Portugal.* [3] *University of Lisbon, Agronomy Institute, CEABN. Lisbon, Portugal.* [4] *University of Cruz Alta, Cruz Alta, Rio Grande do Sul, Brazil*

Abstract

Certain characteristics of some vegetable crops allow multiple harvests during the production cycle; however, to our knowledge, no study has described the behavior of fruit production with progression of the production cycle in vegetable crops with multiple harvests that present data overdispersion. We aimed to characterize the data overdispersion of zero-inflated variables and identify the behavior of these variables during the production cycle of several vegetable crops with multiple harvests. Data from 11 uniformity trials were used without applying treatments; these comprise the database from the Experimental Plants Group at the Federal University of Santa Maria, Brazil. The trials were conducted using four horticultural species grown during different cultivation seasons, cultivation environments, and experimental structures. Although at each harvest, a larger number of basic units with harvest fruit was observed than units without harvest fruit, the basic unit percentage without fruit was high, generating an overdispersion within each individual harvest. The variability within each harvest was high and increased with the evolution of the production cycle of *Capsicum annuum, Solanum lycopersicum* var. *cerasiforme, Phaseolus vulgaris,* and *Cucurbita pepo* species. However, the correlation coefficient between the mean weight and number of harvest fruits tended to remain constant during the crop production cycle. These behaviors show that harvest management should be done individually, at each harvest, such that data overdispersion is reduced.

Additional key words: multiple harvests; data overdispersion; experimental planning.

Abbreviations used: BU (basic unit); CV (coefficient of variation).

Authors' contributions: Conceived and designed the experiments: ADL, LFN and FR. Performed the experiments: MPBP. Analyzed the data: ADL. Wrote the paper: ADL, LFN, FR and MPBP.

Citation: Lúcio, A. D.; Nunes, L. F.; Rego, F.; Pasini, M. P. B. (2016). Relations between zero-inflated variables in trials with horticultural crops. Spanish Journal of Agricultural Research, Volume 14, Issue 2, e0906. http://dx.doi.org/10.5424/sjar/2016142-8175.

Received: 17 Jun 2015. **Accepted:** 12 May 2016

Funding: The Brazilian Ministry of Education's Graduate Education Support Agency (CAPES) awarded an overseas post doctorate scholarship (Process BEX 1457/14-4).

Competing interests: The authors have declared that no competing interests exist.

Correspondence should be addressed to Alessandro D. Lúcio: adlucio@ufsm.br.

Introduction

In some vegetable species, certain specific characteristics allow multiple harvests during the production cycle. The realization of such multiple harvests is defined in a subjective manner and varies with the season and with each cultivated species. In experiments on species with multiple harvests all over the world, the above variations should be considered together with interference among these variables. Such interferences can inflate any residual variance and induce inadequate estimates in the experimental design because of the lack of adequate information at harvest, favoring overdispersion in the database, with tabulation of a large number of null values.

In several studies, strategies have been developed to identify the most appropriate procedures to minimize data variability in experiments with vegetable crops. Among these are the studies by Lopes *et al.* (1998), Lúcio *et al.* (2006, 2008), Carpes *et al.* (2010), Santos *et al.* (2010), and Haesbaert *et al.* (2011). These studies sought to improve the quality of experiments through the following strategies: determining the plot size and sample (Souza *et al.*, 2002; Mello *et al.*, 2004; Lorentz *et al.*, 2005, Lorentz & Lúcio, 2009; Lúcio *et al.,* 2010; Santos *et al.*, 2010; Haesbaert *et al.*, 2011; Storck *et al.*, 2014), adjusting for the variability of experimental areas and each culture (Lúcio *et al.*, 2006; Carpes *et al.*, 2008, 2010), determining the behavior of variability between plant rows and between harvests (Lúcio *et al.*,

2006, 2008; Benz *et al.*, 2015), studying data transformations (Couto *et al.*, 2009) and using the Papadakis method to minimize the effects of excess zeros and resultant data overdispersion (Lúcio *et al.*, 2016).

Lopes *et al.* (1998), Lorentz *et al.* (2005), Carpes *et al.* (2008) and Lúcio *et al.* (2008), have pointed out significant variability between crop rows and harvests, regardless of the species used, and that such variability significantly alters the estimates of sample sizes, types of sampling, size and form of the parcel, experimental outline, and number of harvests needed to adequately differentiate the study treatments.

The relationship between the observed variables, number and weight of fruits harvested in experiments with vegetable crops, and behavior of these species during the production cycle is important, as it generates information on how multiple harvesting should be planned and carried out. One of the problems associated with repeated measurements is the excess of variables with zero values. An interesting strategy to reduce this problem is to estimate the ideal plot size so that the majority of results have values greater than zero, subsequently reducing the variance. Another strategy is to estimate the ideal plot size that provides the smallest variance between the evaluated plots, because often researchers solve this problem empirically, based on practical sizes for conducting the experiment, available area, or from experience.

In agricultural research, it is common to evaluate the full cycle of a particular species or compare different treatments for crop development. However, to our knowledge, no studies have described the behavior of fruit production with progression of the production cycle in vegetable crops with multiple harvests that present data overdispersion.

This study aimed to characterize the data overdispersion of the zero-inflated variables and identify the behavior of these variables during the production cycle of several vegetable crops with multiple harvests.

Material and methods

Data from 11 uniformity trials were used without applying treatments. These comprise the database from the Experimental Plants Group at the Federal University of Santa Maria, Brazil. The trials were performed on four horticultural species, hibrids, grown in different cultivation seasons, cultivation environments, and experimental structures (Table 1). Each experimental basic unit (BU) was composed of a single plant in each row of plants, except for trials with *Phaseolus vulgaris*, where each BU consisted of two plants because of the indeterminate growth characteristic of the species and the tendency to climb ontoin adjacent plants.

During each harvest, the number and weight (in grams) of fruits harvested from each BU were observed, except for trials with *P. vulgaris,* where only

Table 1. Uniformity trials without treatment application used in the study.

Species and hibrids	Cultivation environment	Growing season	No. cultivation rows	Nº basic unit (BU) per cultivation row	Nº of harvests	Harvests in days after sowing or transplanting
Capsicum annuum - Vidi hybrid	Plastic greenhouse	Summer-Autumn	10	70	5	65, 79, 95, 124, 129
	Plastic greenhouse	Winter-Spring	10	70	4	47, 54, 61, 68
Phaseolus vulgaris - Macarrão hybrid	Plastic greenhouse	Autumn-Winter	6	36 (double BU)	4	61, 74, 88, 112
	Field	Autumn-Winter	3	42 (double BU)	4	61, 74, 88, 112
	Plastic tunnel	Autumn-Winter	3	42 (double BU)	4	61, 74, 88, 112
	Field	Spring-Summer	3	42 (double BU)	3	70, 91, 99
	Plastic tunnel	Spring-Summer	3	42 (double BU)	3	70, 91, 99
Solanum lycopersicum var. *cerasiforme* - Lili hybrid	Plastic greenhouse (250 m^2)	Spring-Summer	8	40	3	66, 82, 101
	Plastic greenhouse (200 m^2)	Spring-Summer	8	30	3	75, 88, 103
Cucurbita pepo - Caserta hybrid	Plastic greenhouse	Autumn-Winter	8	20	12	35, 37, 40, 43, 47, 49, 54, 57, 59, 61, 66, 68
	Plastic greenhouse	Spring-Summer	8	20	30	29, 33, 35, 37, 39, 41, 43, 44, 47, 50, 53, 55, 57, 59, 60, 61, 62, 64, 66, 67, 68, 70, 73, 75, 76, 77, 80, 82, 83, 85

the harvest weight was noted. In the trials, the number of bunches harvested by BU was noted.

In each harvest, for number and weight of fruits and number of bunches, an initial descriptive statistical analysis was conducted from which we obtained the percentage of estimates of BU with zero values, the minimum and maximum values, medians, means, coefficient of variation (CV, in %), and degrees of asymmetry and kurtosis. Box-plots were constructed for the number and weight of fruits of each harvest, in order to identify the variability and average behavior of these variables with progression of the production cycle of the species evaluated. Further, we compared the proportions of BU with and without harvest fruits, adopting a 50% probability for presence or absence of fruits ready to be harvested.

A linear correlation analysis between the mean weight and the number of fruits per BU was also performed. For *Solanum lycopersicum* var. *cerasiforme* trials, we also estimated the correlation coefficient between the mean weight of fruits and number of bunches per BU for individual species and cultivation season. Next, for each variable, the Shapiro–Wilk test was performed to identify data adherence to a normal distribution and the Levene test to identify variance homogeneity. For all the statistical analyses performed, a probability of error of 5% was adopted, using Action software 2.7 version.

Results

Lack of adherence to a normal distribution was identified within each harvest along with variance heterogeneity among the multiple harvests, independent of species, season, cultivation environment, and observed variable, because of the high variance esti-

Table 2. Descriptive statistics for weight (grams per basic unit) and number of fruits harvested per basic unit uniformity trials for *Capsicum annuum* cultivated in different growing seasons.

	Harvests					Total
	1	2	3	4	5	
SUMMER-AUTUMN						
Zero observations (%)	35.71	27.14	22.71	51.14	38.71	34.23
Weight of fruit						
Maximum	1068.00	1155.00	978.00	810.00	955.00	1155.00
Median	196.50	269.00	225.50	123.50	131.00	192.51
Means	199.09	281.80	251.71	169.28	172.15	214.81
CV(%)	93.13	81.33	81.48	114.51	111.44	96.11
Asymmetry	0.66	0.42	0.72	0.87	1.21	0.78
Kurtosis	0.13	–0.34	0.24	-0.22	1.31	0.14
Number of fruits						
Maximum	6	12	9	6	10	12
Median	1	2	2	1	1	1
Means	0.98	1.51	1.58	1.03	1.47	1.31
CV(%)	93.52	84.29	78.80	116.91	114.03	100.11
Asymmetry	0.80	1.10	0.71	1.03	1.40	1.25
Kurtosis	0.86	5.63	1.20	0.43	2.50	3.33
	Harvests					Total
	1	2	3	4		
WINTER-SPRING						
Zero observations (%)	33.43	36.28	28.86	55.57		38.53
Weight of fruit						
Maximum	1966.00	2554.00	1345.00	1597.00		2554.00
Median	380.50	266.50	357.00	178.00		245.00
Means	411.07	332.25	362.26	144.24		312.46
CV(%)	92.85	101.77	86.19	145.30		106.38
Asymmetry	0.66	1.08	0.47	1.91		1.01
Kurtosis	–0.07	2.23	-0.60	5.23		1.01
Number of fruits						
Maximum	7	9	8	7		9
Median	2	1	2	1		1
Means	1.61	1.49	1.85	0.78		1.43
CV(%)	92.75	98.18	85.12	139.85		103.01
Asymmetry	0.64	0.81	0.47	1.62		0.84
Kurtosis	–0.24	0.51	–0.44	3.10		0.20

Table 3. Descriptive statistics for weight of fruit (grams per basic unit) and number of fruits and bunches harvested per basic unit in uniformity trials of *Solanum lycopersicum* var. *cerasiforme* cultivated in the spring-summer seasons under different environmental conditions.

	Harvests			Total
	1	2	3	
PLASTIC GREENHOUSE 250 M²				
Zero observations (%)	17.12	6.79	10.05	11.32
Weight of fruit				
Maximum	1087.00	2897.00	1375.00	2897.00
Median	253.00	1 006.00	261.00	361.50
Means	266.25	1006.20	288.74	520.40
CV(%)	77.08	51.24	73.69	93.29
Asymmetry	0.60	0.14	0.98	1.29
Kurtosis	0.21	0.54	1.94	1.55
Number of bunches				
Maximum	5	15	13	15
Median	2	5	4	3
Means	1.49	4.94	3.83	3.42
CV(%)	66.44	49.83	58.15	71.98
Asymmetry	0.26	0.25	0.41	0.73
Kurtosis	–0.12	1.05	1.15	0.73
Number of fruits				
Maximum	59	202	132	202
Median	15	71	26	28
Means	16.29	70.64	28.63	38.52
CV(%)	76.40	51.56	70.29	88.78
Asymmetry	0.64	0.17	0.81	1.19
Kurtosis	0.26	0.50	1.55	1.38
PLASTIC GREENHOUSE 200 M²				
Zero observations (%)	3.33	6.25	9.17	6.25
Weight of fruit				
Maximum	1874.00	1591.00	735.00	1874.00
Median	982.25	480.50	99.00	398.50
Means	965.34	489.72	127.68	527.58
CV(%)	42.29	62.19	85.65	86.45
Asymmetry	–0.30	0.57	1.47	0.74
Kurtosis	–0.09	0.50	3.54	–0.41
Number of bunches				
Maximum	10	1i0	10	10
Median	5	4	3	4
Means	4.62	3.84	3.02	3.83
CV(%)	39.56	48.51	61.68	51.16
Asymmetry	-0.29	0.02	0.50	0.05
Kurtosis	0.17	-0.01	0.10	-0.30
Number of fruits				
Maximum	146	143	71	146
Median	66	44	12	38
Means	65.62	44.92	14.75	41.76
CV(%)	43.15	59.78	80.85	75.35
Asymmetry	–0.03	0.49	1.16	0.61
Kurtosis	–0.01	0.27	1.56	–0.35

Table 4. Descriptive statistics for weight of fruit (grams per basic unit) harvested in uniformity trials of *Phaseolus vulgaris* cultivated in different growing seasons and environmental conditions.

	Harvests				Total
	1	2	3	4	
AUTUMN-WINTER					
Plastic greenhouse					
Zero observations (%)	14.35	3.24	2.31	7.41	6.83
Maximum	444.83	670.57	706.30	483.01	706.30
Median	58.73	199.03	140.79	65.26	112.86
Means	83.50	218.41	160.92	90.34	138.29
CV(%)	101.67	63.03	64.44	88.81	85.13
Asymmetry	1.33	0.64	1.17	1.80	1.24
Kurtosis	1.69	0.13	3.05	5.01	1.80
Field					
Zero observations (%)	2.38	0.79	0.79	3.97	1.98
Maximum	293.98	943.84	555.56	221.46	943.84
Median	81.48	269.52	153.77	83.30	130.24
Means	95.91	280.96	177.23	89.50	160.90
CV(%)	69.39	56.73	59.82	55.68	80.64
Asymmetry	0.55	0.70	1.02	0.39	1.55
Kurtosis	–0.32	1.23	1.02	-0.46	3.39
Plastic tunnel					
Zero observations (%)	0.79	0.00	0.79	0.79	0.59
Maximum	548.98	758.50	467.18	538.39	758.50
Median	211.73	345.31	166.70	71.85	187.16
Means	218.64	355.57	181.97	90.17	211.59
CV(%)	50.67	43.29	57.34	81.84	70.28
Asymmetry	0.39	0.17	0.62	2.19	0.87
Kurtosis	–0.18	–0.42	–0.15	9.74	0.34

	Harvests			Total
	1	2	3	
SPRING-SUMMER				
Field				
Zero observations (%)	12.70	0.79	2.38	5.29
Maximum	462.00	1468.89	845.94	1468.89
Median	118.72	484.43	196.84	232.93
Means	152.16	525.58	263.22	313.65
CV(%)	81.00	60.76	87.76	90.81
Asymmetry	0.63	0.45	0.75	1.1
Kurtosis	–0.40	–0.47	–0.58	0.87
Plastic tunnel				
Zero observations (%)	1.59	1.59	9.52	4.23
Maximum	1310.34	1290.46	169.66	1310.34
Median	574.90	402.10	37.47	288.16
Means	552.70	433.97	43.99	343.56
CV(%)	52.07	55.31	90.35	89.45
Asymmetry	0.08	0.74	0.77	0.70
Kurtosis	–0.37	0.95	–0.09	–0.34

Table 5. Data descriptive statistics for weight (grams per basic unit) and number of fruits harvested by basic unit in the uniformity trials of *Cucurbita pepo* grown in the autumn-winter season.

Weight of fruits	Harvests						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Zero observations (%)	90.62	44.37	39.35	66.87	88.75	66.87	48.75
Maximum	784.00	407.00	352.00	677.00	576.00	442.00	828.00
Median	0.00	95.50	104.00	0.00	0.00	0.00	106.50
Means	22.58	96.59	114.77	71.57	27.58	57.00	93.14
CV(%)	415.95	107.35	95.81	158.89	316.56	162.86	120.87
Asymmetry	6.02	0.78	0.40	1.70	3.67	1.68	2.09
Kurtosis	41.69	–0.20	–1.09	3.98	14.73	2.57	10.20
	8	**9**	**10**	**11**	**12**	**Total**	
Zero observations (%)	69.37	40.62	36.25	72.50	78.12	61.87	
Maximum	322.00	797.00	1032.00	1096.00	588.00	1096.00	
Median	0.00	172.00	209.50	0.00	0.00	0.00	
Means	64.61	183.86	235.42	100.01	37.95	92.09	
CV(%)	159.06	102.69	99.72	222.01	242.37	164.44	
Asymmetry	1.16	0.76	0.83	2.87	3.51	2.33	
Kurtosis	–0.27	–0.10	0.22	8.35	15.33	7.39	
Number of fruits	**Harvests**						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Maximum	2	2	3	2	2	3	4
Median	0	1	1	0	0	0	1
Means	0.12	0.72	0.78	0.36	0.13	0.40	0.61
CV(%)	322.75	101.11	95.08	149.59	297.92	160.08	112.39
Asymmetry	3.45	0.48	0.55	1.14	3.05	1.44	1.14
Kurtosis	11.65	–1.01	–0.41	0.27	9.14	1.46	2.46
	8	**9**	**10**	**11**	**12**	**Total**	
Maximum	2	4	7	3	3	7	
Median	0	1	1	0	0	0	
Means	0.32	0.90	1.39	0.35	0.26	0.53	
Variance	0.24	0.86	2.01	0.39	0.29	0.65	
CV(%)	153.09	103.04	101.99	178.43	207.12	152.12	
Asymmetry	1.08	0.86	1.01	1.72	2.18	1.96	
Kurtosis	–0.13	0.31	1.00	2.32	4.94	5.91	

mates, and consequently, the CV (Tables 2 to 6). When plotting the weight and number of fruits (bunches in one case) variability in each of the multiple harvests, independent of the above conditions, we could not identify similar behavior of variability with progression of the production cycle of the species (Figs. 1 to 4).

When comparing the proportion of BU with and without harvest fruits, within each of the multiple harvests, in 13.3% of the harvests (10 of 75 harvests under all study conditions), the proportions did not differ; that is, they had statistically the same number of BU with and without harvest fruits in the specific season. In 15 (23.1%) of the 65 harvests the BU proportion without harvest fruit was significant greater than that with harvest fruit (Figs. 5 to 8). This result is interesting and indicates that in 56 of the 75 study harvests (74.7%), a significant difference was noted with greater number of BUs with harvest fruits than those without harvest fruits.

Within each harvest, significant correlations coefficients were noted between the mean weight and number of fruits and/or bunches harvested per BU, with estimates of around 0.6 for *C. annuum* and *S. lycopersicum* var. *cerasiforme* species (Figs. 5 and 6). As for *C. pepo*, the estimates varied as the production cycle progressed, were significant and presented maximum values around 0.8 (Fig. 8). As previously described, *C. pepo* presented different characteristics during fruit maturation, which

Table 6. Data descriptive statistics for weight (grams per basic unit) and number of fruits harvested by the basic unit in the uniformity trials of *Cucurbita pepo* grown in the spring-summer season.

Weight of fruits	Harvests					
	1	**2**	**3**	**4**	**5**	**6**
Zero observations (%)	80.00	80.00	41.87	32.50	7.50	24.37
Maximum	424.00	629.00	648.00	554.00	869.00	884.00
Median	0.00	0.00	154.50	191.50	408.00	188.00
Means	43.51	50.23	160.48	183.06	390.23	186.02
CV(%)	219.17	239.77	101.80	85.65	49.74	76.39
Asymmetry	2.17	2.73	0.62	0.33	–0.17	0.81
Kurtosis	3.87	7.44	–0.59	–0.86	–0.31	2.54
	7	**8**	**9**	**10**	**11**	**12**
Zero observations	31.25	63.75	76.87	26.25	8.75	15.62
Maximum	691.00	901.00	658.00	1062.00	1530.00	1062.00
Median	200.00	0.00	0.00	247.00	804.00	386.00
Means	180.31	112.30	88.67	303.21	749.85	380.54
CV(%)	82.98	160.78	201.06	87.10	50.36	62.23
Asymmetry	0.39	1.88	1.86	0.67	–0.33	0.01
Kurtosis	–0.31	3.84	2.13	–0.33	–0.38	–0.49
	13	**14**	**15**	**16**	**17**	**18**
Zero observations	26.25	33.12	52.50	84.37	68.12	51.87
Maximum	867.00	1150.00	601.00	368.00	537.00	543.00
Median	293.50	239.50	0.00	0.00	0.00	0.00
Means	285.37	252.20	107.54	30.80	61.18	118.99
CV(%)	78.49	94.36	122.20	247.86	162.34	121.11
Asymmetry	0.22	0.82	1.07	2.54	1.66	0.91
Kurtosis	–0.95	0.43	0.74	6.15	3.13	–0.20
	19	**20**	**21**	**22**	**23**	**24**
Zero observations	54.35	41.87	30.62	54.37	20.00	38.13
Maximum	517.00	1045.00	537.00	1441.00	855.00	604.00
Median	0.00	156.50	192.00	0.00	297.00	139.50
Means	109.51	165.03	188.89	110.88	294.72	154.37
CV(%)	122.17	106.57	82.20	151.18	68.65	100.86
Asymmetry	0.87	1.17	0.27	3.62	0.09	0.79
Kurtosis	–0.18	2.68	–0.84	24.23	–0.70	–0.12
	25	**26**	**27**	**28**	**29**	**30**	**Total**
Zero observations	31.87	48.13	26.88	53.12	58.75	43.13	43.54
Maximum	855.00	341.00	1 001.00	513.00	304.00	756.00	1 530.00
Median	219.50	98.00	272.00	0.00	0.00	109.00	146.00
Means	227.51	91.29	306.82	109.11	63.89	139.06	188.19
CV(%)	84.53	107.89	85.89	122.43	130.92	114.57	121.88
Asymmetry	0.39	0.57	0.53	0.93	0.89	1.24	1.63
Kurtosis	–0.46	–0.84	–0.53	–0.03	–0.54	1.70	3.60

(Cont.)

Table 6 (cont.). Data descriptive statistics for weight (grams per basic unit) and number of fruits harvested by the basic unit in the uniformity trials of *Cucurbita pepo* grown in the spring–summer season.

Number of fruits	Harvests					
	1	**2**	**3**	**4**	**5**	**6**
Maximum	2	2	2	2	3	5
Median	0	0	1	1	2	1
Means	0.22	0.22	0.72	0.86	1.68	0.98
CV(%)	213.20	208.30	96.22	82.22	45.33	74.98
Asymmetry	1.98	1.92	0.43	0.30	–0.41	0.96
Kurtosis	3.15	2.87	–0.90	–1.00	–0.10	3.93
	7	**8**	**9**	**10**	**11**	**12**
Maximum	3	3	2	3	5	3
Median	1	0	0	1	2	2
Means	0.87	0.40	0.26	1.06	2.09	1.46
CV(%)	82.08	152.07	196.12	76.64	46.88	60.49
Asymmetry	0.39	1.56	1.76	0.32	–0.37	–0.10
Kurtosis	–0.34	2.87	2.23	–0.56	0.16	–0.76
	13	**14**	**15**	**16**	**17**	**18**
Maximum	3	3	3	2	2	2
Median	1	1	0	0	0	0
Means	1.05	0.92	0.57	0.17	0.33	0.58
CV(%)	74.99	86.27	118.99	249.57	151.51	115.66
Asymmetry	0.21	0.44	1.01	2.38	1.01	0.71
Kurtosis	–0.72	–0.52	0.74	5.13	–0.30	–0.61
	19	**20**	**21**	**22**	**23**	**24**
Maximum	2	2	3	2	3	3
Median	0	1	1	0	1	1
Means	0.50	0.69	0.83	0.56	1.21	0.82
CV(%)	116.62	96.13	79.92	121.11	65.07	92.07
Asymmetry	0.66	0.42	0.32	0.79	–0.08	–0.38
Kurtosis	–0.56	–0.79	–0.26	–0.55	–0.85	–0.91

	25	**26**	**27**	**28**	**29**	**30**	**Total**
Maximum	3	2	4	2	2	4	5
Median	1	1	1	0	0	1	1
Means	0.92	0.59	1.11	0.54	0.44	0.74	0.78
CV(%)	82.06	104.48	79.05	115.65	124.48	103.80	104.94
Asymmetry	0.21	0.54	0.40	0.71	0.71	0.88	0.80
Kurtosis	–0.99	–0.64	–0.33	–0.50	–0.64	0.84	0.15

generated results different from those obtained with the other studied species that showed similar correlation coefficient estimates in the multiple harvests.

Discussion

The high variability and overdispersion identified in the data is a direct consequence of the high number of BUs with observed values equal to zero. This fact changes the entire behavior of the descriptive statistics estimates, such as the asymmetry and degree of kurtosis (Tables 2 to 6). This situation means that in most cases, the data show a positive asymmetrical distribution and high degree of kurtosis with a platykurtic distribution.

The appearance of fruits on the plants on different days, causing variation in growth among them; early

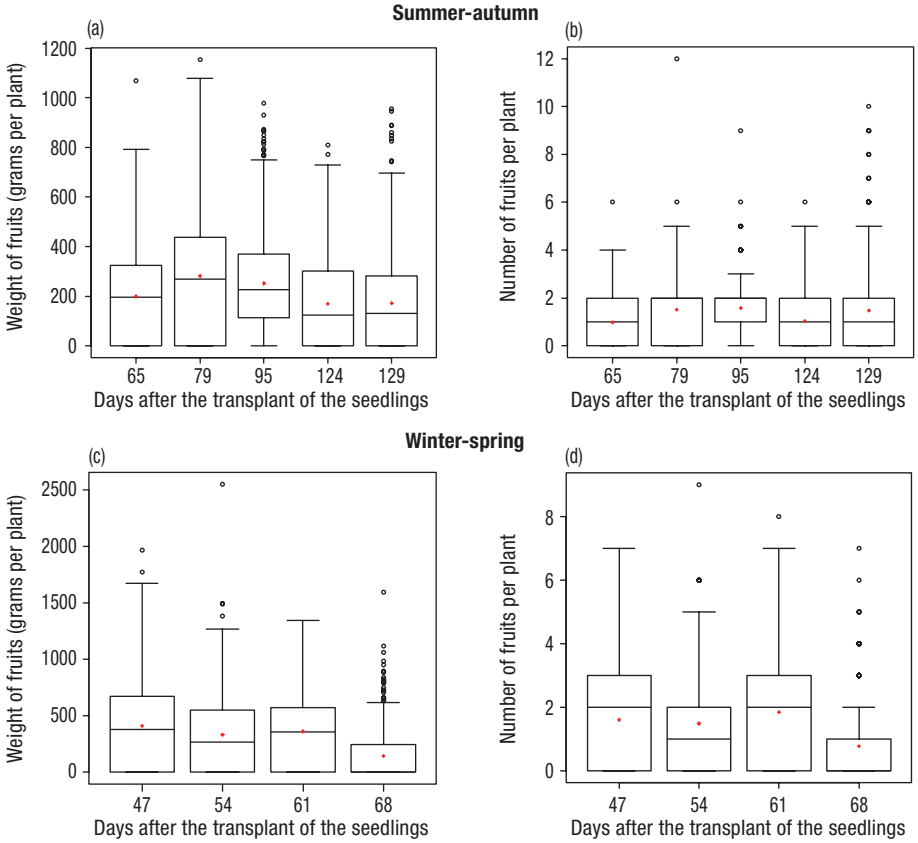

Figure 1. Box-plot for weight (grams per basic unit) (a,c) and number of fruits (b,d) per harvest by basic unit in *Capsicum annuum* uniformity trials in a plastic greenhouse in different growing seasons.

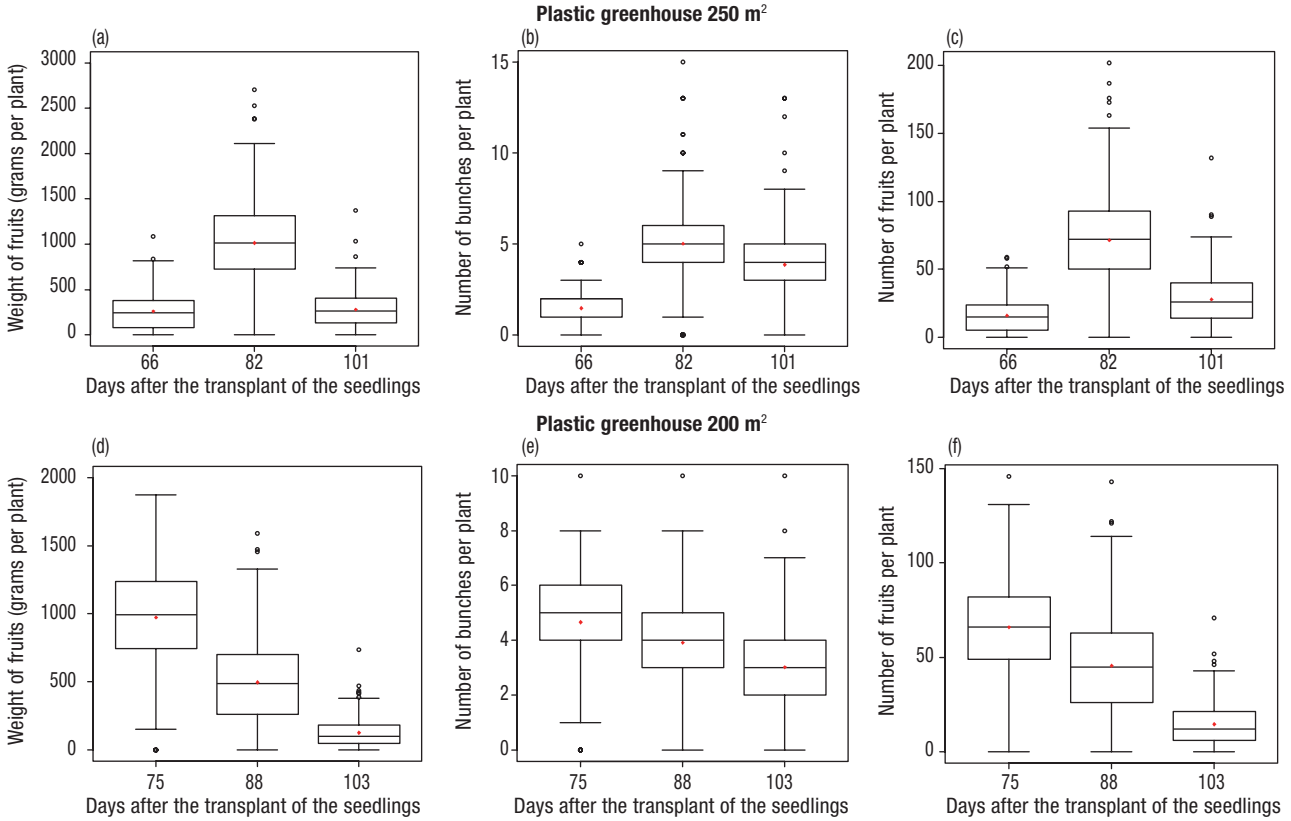

Figure 2. Box-plot for weight of fruits (grams per basic unit) (a,d), number of bunches (b,e) and number of fruit (c,f) per harvest by basic unit in *Solanum lycopersicum* var. *cerasiforme* uniformity trials in the spring-summer season under different environmental conditions.

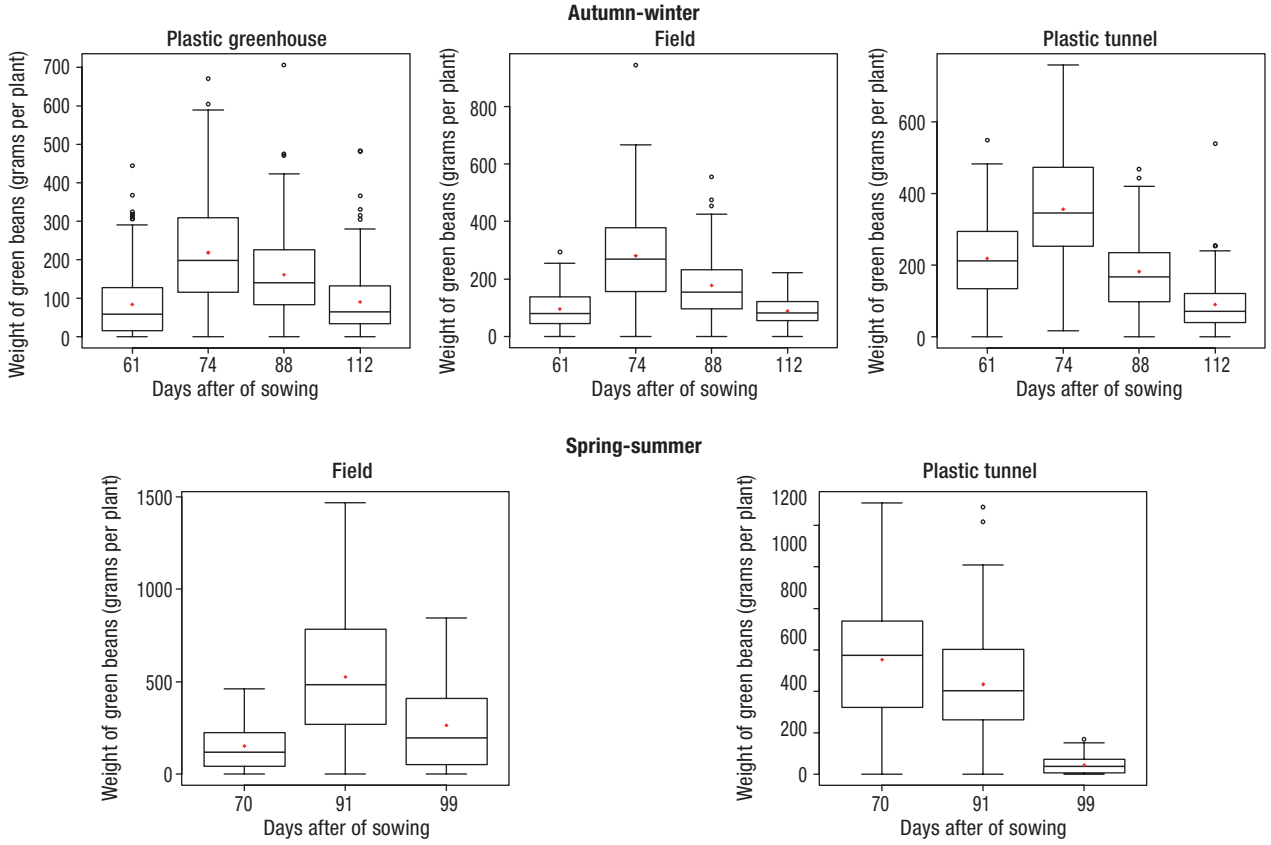

Figure 3. Box-plot for weight of fruit (grams per basic unit) per harvest in *Phaseolus vulgaris* uniformity trials in the autumn-winter and spring-summer seasons under different environmental conditions.

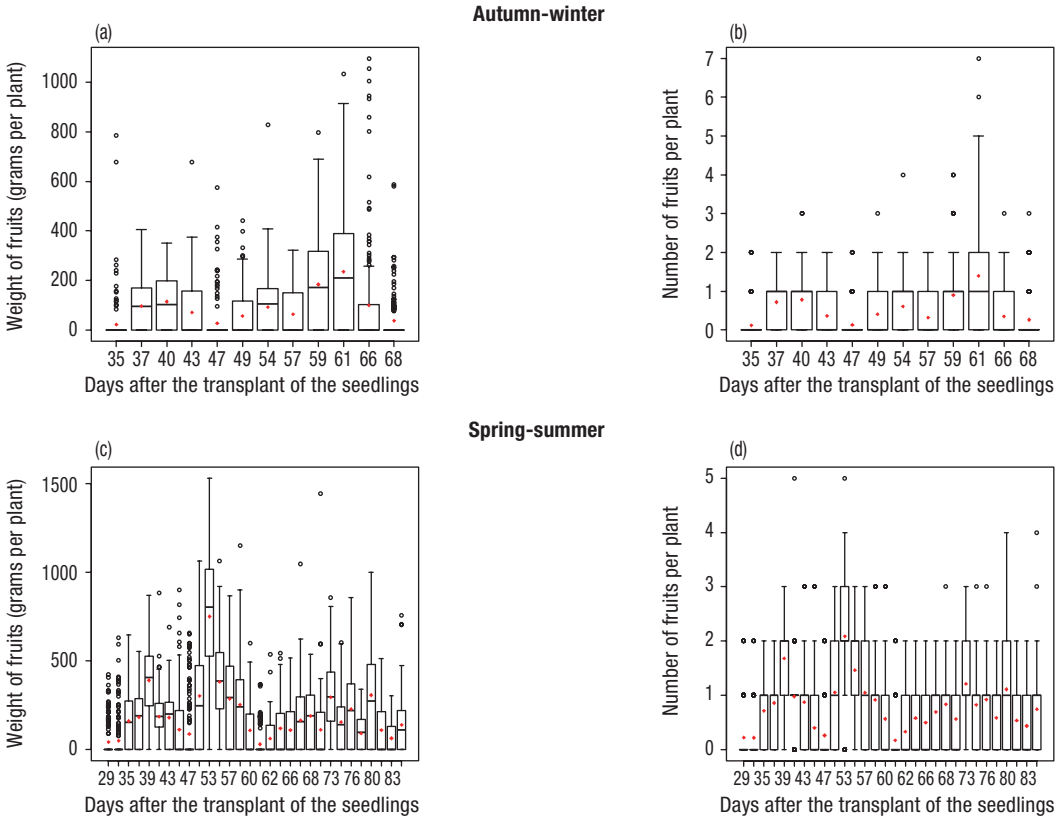

Figure 4. Box-plot for weight (grams per basic unit) (a,c) and number of fruits (b,d) per harvest by basic unit in *Cucurbita pepo* uniformity trials in autumn-winter and spring-summer seasons.

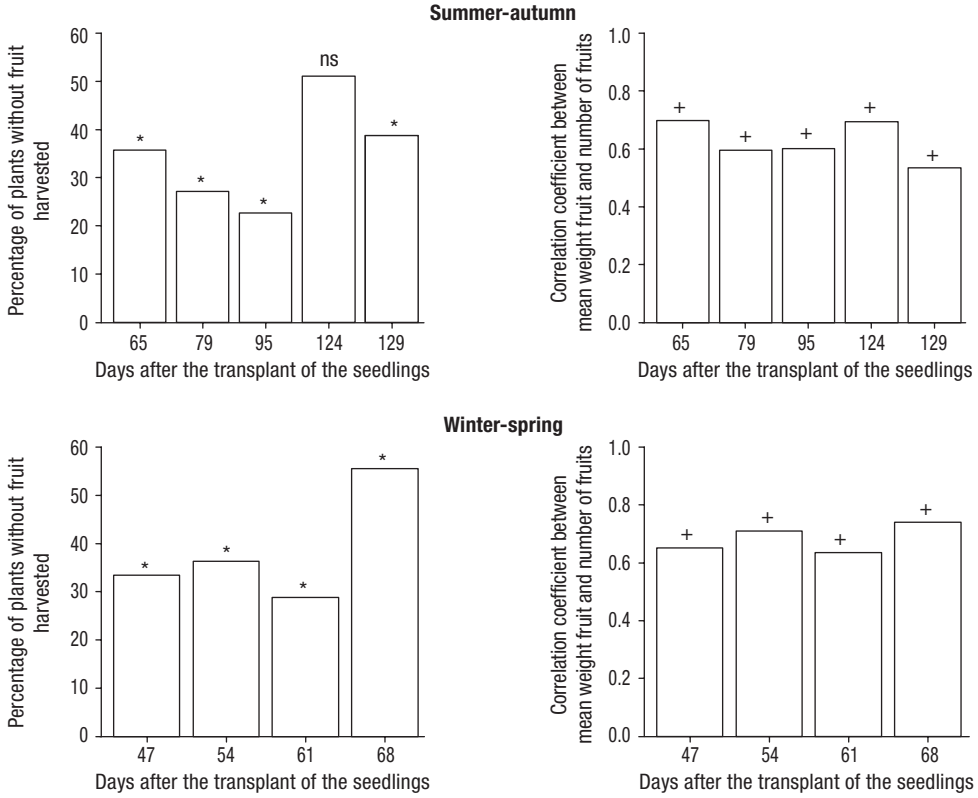

Figure 5. Percentage of basic units without fruits harvested and correlation coefficient of the mean weight of fruits per basic unit and the number of fruits per basic unit, in *Capsicum annuum* cultivated in a plastic greenhouse in different seasonal stations.*: Significant difference between the proportions of fruit present or not fit to be harvest, at 5% probability of error; +: Significant correlation coefficients at 5% probability of error. ns: not significant.

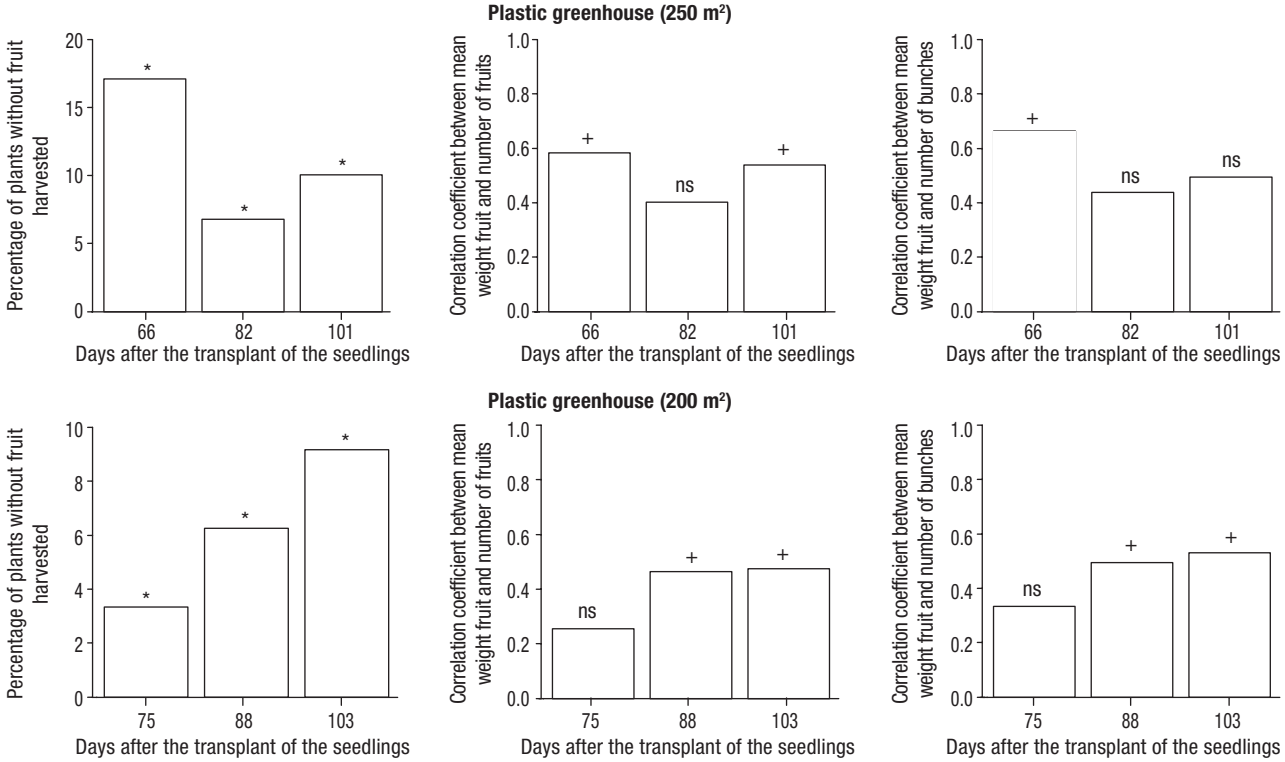

Figure 6. Percentage of basic units without fruits harvested and correlation coefficients of the mean weight and number of fruits per basic unit and mean weight and number of bunches per basic unit, in *Solanum lycopersicum* var. *cerasiforme* cultivated in the spring-summer season in a plastic greenhouse 250 × 200 m².*: Significant difference between the proportions of fruit present or not fit to be harvest, at 5% probability of error; +: Significant correlation coefficients at 5% probability of error. ns: not significant.

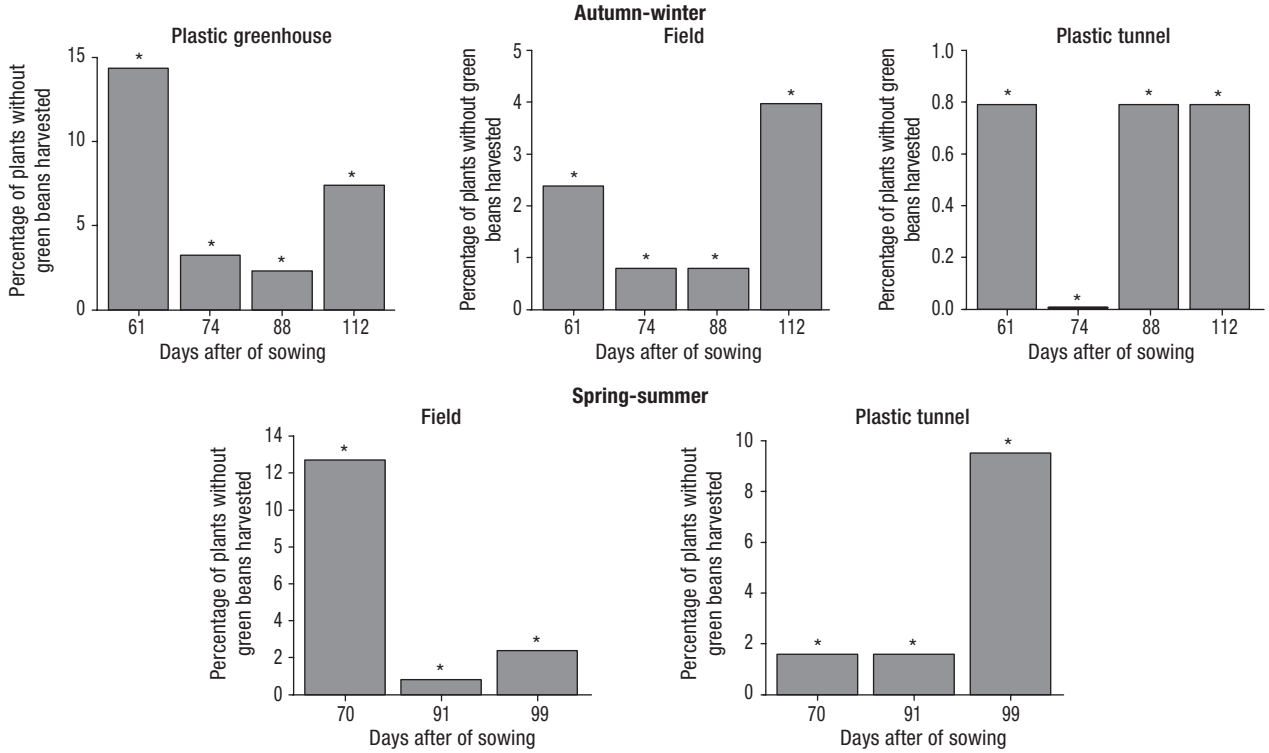

Figure 7. Percentage of basic units without fruits harvested in *Phaseolus vulgaris* cultivated in different seasonal stations and cultivation environments.*: Significant difference between the proportions of fruit present or not fit to be harvest, at 5% probability of error.

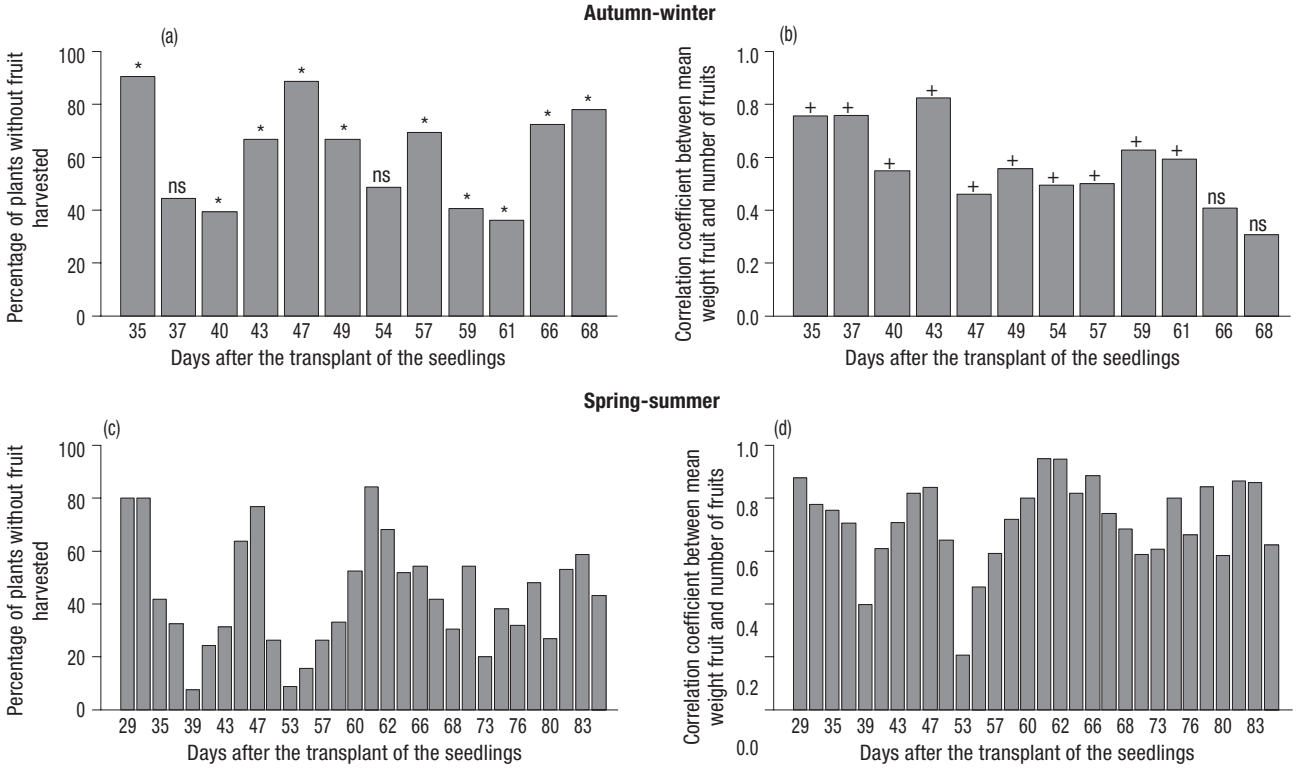

Figure 8. Percentage of basic units (BU) without fruits harvested and the correlation coefficient between the mean weight of fruits/BU and the number of fruits/BU in *Cucurbita pepo* cultivated in a plastic greenhouse in different seasonal stations and cultivation environments. a,b) *: Significant difference between the proportions of fruit present or not fit to be harvest, at 5% probability of error; +: Significant correlation coefficients at 5% probability of error. ns: not significant. c): There was no significant difference between the proportions of fruit present or not fit to be harvest at 60, 64, 66, 70, 77, 82 and 85 days after the transplant of the seedlings. In the other crops there was no significant difference between the proportions. d): All correlation coefficients were significant at 5% probability of error, except those obtained 39 and 53 days after the transplant of the seedlings.

or late maturation of some fruits; and lack of uniformity in size at harvest, beyond their lack of uniformity in size fruit at harvest, resulting in the inability to define the ideal harvest point, are factors increasing variability in the fruit number and weight, causing data overdispersion and consequent variations in the statistical analysis. Cargnelutti Filho *et al.* (2004) obtained higher CV% values at the beginning and end of the tomato harvest, because the beginning and end of the fruit production were not uniform among the plants. In another study on tomatoes, Lúcio *et al.* (2010) found that the largest production of fruits occurred mid-way and at the end of the production cycle, and that the variability increased due to physiological aspects of the plants, because they go into a state of senescence.

With larger number of harvests, an increase in variability is noted. Souza *et al.* (2002), Oliveira *et al.* (2005), and Lúcio *et al.* (2006) recommend that homogeneous variances be maintained during the production cycle of the crop. Further, they suggest that researchers clearly define the ideal number of harvests to be performed, which then must be planned and executed by considering each row as a block, thereby allowing experimental repetition. Variabilities were also noted in the studies by Lúcio *et al.* (2004), Mello *et al.* (2004), and Lorentz & Lúcio (2009) on *C. annuum*; Carpes *et al.* (2008) and Lúcio *et al.* (2008) on *C. pepo*; and Storck *et al.* (2014) on *Passiflora edulis*. According to these authors, an increase in the number of replicates is recommended, along with possibly increasing the plot size.

The non-adherence to a normal distribution of the data in the *C. pepo* trials can be explained by the number of harvests, which was larger than those of *C. annuum* and *S. lycopersicum* var. *cerasiforme*. With the increased number of harvests in these trials, within each individual harvest, lesser number of fruits per BU was observed, consequently, showing a tendency of non-adherence to the normal distribution with smaller data overdispersion than for species with fewer harvests with greater number of harvested fruits in the individual harvests (Tables 2 to 6).

These variance behaviors show that harvest management should be done individually, at each harvest, such that data overdispersion is reduced. Appropriate definition of each fruit harvest time can be a practical alternative, as well as defining time intervals for each harvest rather than identifying a specific day. Thus, the BU number with fruit ready for harvest can be increased, with reduction in the data amplitude within each multiple harvests.

The main characteristic of the studied species in plants without fruits ready to be harvested throughout their production cycle was a lack of reduction of variability during the course of crop production cycles. Even without fruit harvest in a BU, the variability of the data remained high and kept increasing, because in this particular case, the value of the crop n remained identical to the value obtained at harvest $n-1$, while in the BU with harvested fruits, the value increased; thus, variability in the values in each harvest tended to increase (Figs. 1 and 4).

One way to reduce data variability, and thus, overdispersion, is to increase the number of BUs with harvested fruits within each harvest, since this will also increase the number of harvested fruit and total weight of fruit within each BU. As previously mentioned, a practical and viable manner to promote this situation is to clearly define the harvest point and identify time intervals between each harvest.

In summary, within each harvest, there were more basic units (BU) with than without harvest fruit. However, the BU percentage without fruits was high, generating data overdispersion within each harvest. The variability within each harvest is high and increases as the production cycle progresses in *C. annuum, S. lycopersicum* var. *cerasiforme, P. vulgaris,* and *C. pepo*. The correlation coefficient values between the average fruit weight and number of harvested fruits tended to remain constant during the crop production cycle. These behaviors show that harvest management should be done individually, at each harvest, such that data overdispersion is reduced.

References

Benz V, Lúcio AD, Lopes SJ, 2015. The spatial and temporal independence of Italian zucchini production. Acta Scientiarum 37: 257-263. http://dx.doi.org/10.4025/actasciagron.v37i2.19398.

Cargnelutti Filho A, Radin B, Matzenauer R, Storck L, 2004. Número de colheitas e comparação de genótipos de tomateiro cultivados em estufa de plástico. Pesquisa Agropecuária Brasileira 39: 953-959. http://dx.doi.org/10.1590/S0100-204X2004001000002.

Carpes RH, Lúcio AD, Storck L, Lopes SJ, Zanardo B, Paludo AL, 2008. Ausência de frutos colhidos e suas interferências na variabilidade da fitomassa de frutos de abobrinha italiana cultivada em diferentes sistemas de irrigação. Ceres 55: 590-595.

Carpes RH, Lúcio AD, Lopes SJ, Benz V, Haesbaert FM, Santos D, 2010. Variabilidade produtiva e agrupamentos de colheitas de abobrinha italiana cultivada em ambiente protegido. Ciência Rural 40: 294-301. http://dx.doi.org/10.1590/S0103-84782010005000007.

Couto MRM, Lúcio AD, Lopes SJ, Carpes RH, 2009. Transformação de dados em experimentos com abobrinha italiana em ambiente protegido. Ciência Rural 39: 1701-1707. http://dx.doi.org/10.1590/S0103-84782009005000110.

Haesbaert FM, Santos D, Lúcio AD, Benz V, Antonello BI, 2011. Tamanho de amostra para experimentos com feijão-de-vagem em diferentes ambientes. Ciência Rural 41: 38-44. http://dx.doi.org/10.1590/S0103-84782011000100007.

Lopes SJ, Storck L, Heldwein AB, Feijó S, Ros CA, 1998. Técnicas experimentais para tomateiro tipo salada sob estufas plásticas. Ciência Rural 28: 193-197. http://dx.doi.org/10.1590/S0103-84781998000200002.

Lorentz LH, Lúcio AD, 2009. Tamanho e forma de parcela para pimentão em estufa plástica. Ciência Rural 39: 2380-2387. http://dx.doi.org/10.1590/S0103-84782009005000202.

Lorentz LH, Lúcio AD, Boligon AA, Lopes SJ, Storck L, 2005. Variabilidade da produção de frutos de pimentão em estufa plástica. Ciência Rural 35: 316-323. http://dx.doi.org/10.1590/S0103-84782005000200011.

Lúcio AD, Mello RM, Storck L, Carpes R., Boligon, AA, Zanardo B, 2004. Estimativa de parâmetros para planejamento de experimentos com a cultura do pimentão em área restrita. Horticultura Brasileira 22: 766-770. http://dx.doi.org/10.1590/S0102-05362004000400020.

Lúcio AD, Lorentz LH, Boligon AA, Lopes SJ, Storck L, Carpes RH, 2006. Variação temporal da produção de pimentão influenciada pela posição e características morfológicas das plantas em ambiente protegido. Horticultura Brasileira 24: 31-35. http://dx.doi.org/10.1590/S0102-05362006000100007.

Lúcio AD, Carpes RH, Storck L, Lopes SJ, Lorentz LH, Paludo AL, 2008. Variância e média da massa de frutos de abobrinha-italiana em múltiplas colheitas. Horticultura Brasileira 26: 333-339. http://dx.doi.org/10.1590/S0102-05362008000300009.

Lúcio AD, Carpes RH, Storck L, Zanardo B, Toebe M, Puhl OJ, Santos JRA, 2010. Agrupamento de colheitas de tomate e estimativas do tamanho de parcela em cultivo protegido. Horticultura Brasileira 28: 190-196. http://dx.doi.org/10.1590/S0102-05362010000200009.

Lúcio AD, Santos D, Cargnelutti Filho A, Schabarum DE, 2016. Método de Papadakis e tamanho de parcela em experimentos com a cultura da alface. Horticultura Brasileira 34: 66-73. http://dx.doi.org/10.1590/S0102-053620160000100010.

Mello RM, Lúcio AD, Storck L, Lorentz LH, Carpes RH, Boligon AA, 2004. Size and form of plots for the culture of the Italian pumpkin in plastic greenhouse. Scientia Agricola 61: 457-461. http://dx.doi.org/10.1590/S0103-90162004000400017.

Oliveira SJR, Storck L., Lopes SJ, Lúcio AD, Feijó S, Damo HP, 2005. Plot size and experimental unit relationship in explanatory experiments. Scientia Agricola 62: 585-589. http://dx.doi.org/10.1590/S0103-90162005000600012.

Santos D, Haesbaert FM, Puhl OJ, Santos JRA, Lúcio AD, 2010. Suficiência amostral para alface cultivada em diferentes ambientes. Ciência Rural 40: 800-805. http://dx.doi.org/10.1590/S0103-84782010000400009.

Souza MF, Lúcio AD, Storck L, Carpes RH, Santos PM, Siqueira LFF, 2002. Tamanho da amostra para peso de massa de frutos, na cultura da abóbora italiana em estufa plástica. Revista Brasileira de Agrociência 8: 123-128.

Storck L, Lúcio AD, Krause W, Araújo DV, Silva CA, 2014. Scaling the number of plants per plot and number of plots per genotype of yellow passion fruit plants. Acta Scientiarum 36: 73-78. http://dx.doi.org/10.4025/actasciagron.v36i1.17697.