









Sample size matters: optimizing fruit sampling for reliable quality assessment in sweet orange varieties

Maiara Janine Machado Caldas¹, Fabiana Fumi Cerqueira Sasaki², Valter da Silva Rodrigues¹, Lizziane Gomes Leal Santana^{2*}, Carlos Alberto da Silva Ledo², Abelmon da Silva Gesteira², Orlando Sampaio Passos², Walter dos Santos Soares Filho²

¹Universidade Federal do Recôncavo da Bahia/Centro de Ciências Agrárias e Biológicas, Rua Rui Barbosa, 710 – 44380-000 – Cruz das Almas, BA – Brasil.

²Embrapa Mandioca e Fruticultura, Rua Embrapa s/n – 44380-000 – Cruz das Almas, BA – Brasil.

*Corresponding author <lizzianegomes@gmail.com>

Edited by: Sérgio Tonetto de Freitas

Received July 15, 2024

Accepted May 20, 2025

ABSTRACT: This study aimed to determine the optimal sample size of fruits from sweet orange varieties for the evaluation of their physical and chemical quality attributes. Twenty ripe fruits from the following varieties, namely, Diva, Hamlin CNPMF-020, Pera CNPMF-D6, Uruburetama Blood, and Valência Tuxpan CNPMF were used. The variables assessed included fruit mass, longitudinal and transverse diameters, juice mass and yield, peel color and thickness, soluble solids content, titratable acidity, the soluble solids/titratable acidity ratio (SS/TA), and a technological index. The experimental design was completely randomized, with individual analysis of each fruit. Data were analyzed using the Modified Maximum Curvature Method with the aid of the R statistical software. The optimal minimum number of fruits varied according to variety and variable. The smallest required sample sizes were four fruits for Pera CNPMF-D6, six fruits for Diva and Uruburetama Blood, and seven fruits for Valência Tuxpan CNPMF and Hamlin CNPMF-020. Therefore, a standardized sample size of seven fruits is recommended for studies based on the physicochemical quality of sweet orange varieties, to ensure sufficient precision and statistical reliability under the conditions of this study.

Keywords: modified maximum curvature method, titratable acidity, soluble solids

According to data from the Food and Agriculture Organization of the United Nations (FAO, 2021), the citrus industry in Brazil leads the socioeconomic ranking of cultivated and traded fruits, with an emphasis on sweet oranges [*Citrus × sinensis* (L.) Osbeck]. These oranges are of undeniable importance to both domestic and international markets, where they are primarily sold fresh or as processed juice.

For oranges to be well accepted for fresh consumption and industrial processing, they must meet specific physical, chemical, and nutritional quality standards required by domestic and foreign markets (CEAGESP, 2011). Physical attributes include external appearance, size, shape, peel color, diameter (transverse and longitudinal), and juice percentage or yield. Chemical standards include analyses of soluble solids content, titratable acidity, and the ratio of soluble solids to titratable acidity. In experiments involving citrus varieties, it is essential to maximize the information obtained while reducing experimental errors in the analysis of these fruit quality parameters. This goal should be achieved without oversizing the samples analyzed (Lima et al., 2007).

Determining sample size is crucial to the feasibility of research projects, as it has direct implications for both the budget and execution timeline. Calculating the ideal sample size avoids the evaluation of unnecessarily large quantities of material, which would lead to a waste of resources and time. Conversely, samples that are too small may not be representative of the study object and may

lead to imprecise results (Agranonik and Hirakata, 2011; Arellano-Durán et al., 2018; Cargnelutti Filho et al., 2018; Vilvert et al., 2021).

Sample estimation or optimal plot size can be determined using the Modified Maximum Curvature method, which is based on Smith (1938). The data obtained are interpreted through dispersion curves, where the optimal size corresponds to the point of maximum curvature or inflection. This point represents the sample size that reliably characterizes the population. This approach can be applied to determine the appropriate sample size for fruit quality analyses, thereby ensuring efficient use of labor and reliable, accurate results (Faria et al., 2020).

Currently, there is no information in the literature regarding the ideal fruit sample size for determining the physical and chemical quality of citrus cultivars. This study aimed to establish a foundation for future research on sweet orange fruit sampling.

The study was conducted in 2017 at the Active Germplasm Bank of Citrus (AGB Citrus) and the Laboratório de Pós-Colheita of the Embrapa Mandioca e Fruticultura, a research unit of the Embrapa, located in Cruz das Almas, Recôncavo Baiano, in the state of Bahia, Brazil (12°40'39" S, 39°06'23" W, altitude 226 m).

The sweet orange varieties evaluated were Diva, Hamlin CNPMF-020, Pera CNPMF-D6, Uruburetama blood (red pulp), and Valência Tuxpan CNPMF. The experimental design was completely randomized, with 20 replications consisting of one fruit per variety.

Fruits were collected from six-year-old trees at AGB Citrus, planted at a spacing of 5.0 m × 2.0 m and grafted onto 'Indio' and 'Riverside' citrandarin rootstocks [*C. sunki* (Hayata) hort. ex Tanaka × *Poncirus trifoliata* (L.) Raf.]. Harvesting was carried out at the appropriate ripening stage for each variety.

The physical variables analyzed were: a) fruit and juice mass, measured on a commercial scale and expressed in grams; b) longitudinal and transverse diameters and peel thickness, measured with a digital caliper and expressed in millimeters; c) juice yield, calculated as the ratio of juice mass to fruit mass, expressed as a percentage, and d) peel color.

The chemical variables analyzed were: a) technological index (TI), calculated using the equation $TI = [\text{juice yield (\%)} \times \text{total soluble solids (°Brix)} \times \text{box mass} - \text{a standard unit of 40.8 kg}/10,000]$, expressed in kg of total soluble solids (SS) per box; b) titratable acidity (TA), determined by titration according to Moura et al. (2016), expressed as a percentage of citric acid; c) total soluble solids content, measured using a digital benchtop refractometer with temperature correction, expressed in °Brix; d) SS/TA, calculated as the quotient of these two variables.

The data were analyzed using the Modified Maximum Curvature Method proposed by Lessman and Atkins (1963) to determine the optimal plot size. This method relates the coefficient of variation (CV) to plot size, modeled as $CV(x) = a/X^b$, where *a* and *b* are parameters to be estimated, with *a* being a regression constant and *b* a regression coefficient (Meier and Lessman, 1971).

The curvature function derived from this model identifies the abscissa value corresponding to

the point of maximum curvature using the formula: $X_0 = \exp\{[1/(2b+2)]\log[(ab)2(2b+1)/(b+2)]\}$, where X_0 is the estimated optimal sample size (Meier and Lessman, 1971). Statistical analyses were conducted using the R software (R Core Team, 2021).

As regards the physical variables fruit mass and diameter (transverse and longitudinal), the Hamlin CNPMF-020 required the largest sample size, whereas Pera CNPMF-D6 required the smallest (Table 1). Fruit size is influenced by water availability, variety, and environmental temperature (Nawaz et al., 2021; El-Otmani et al., 2020). Water loss reduces fruit mass and alters chemical compound concentrations (Uthman and Garba, 2023) and reported size reductions due to irrigation limitations (El-Otmani et al., 2020). Thus, fruit size directly influences the chemical concentration of compounds, and affects the quality of fruit.

For peel color, peel thickness, mass, juice yield, and the technological index, Hamlin CNPMF-020 again required the largest optimal sample sizes compared to the other varieties (Tables 2 and 3).

In experiments with large plots or many fruit characteristics, large samples demand extensive labor. Therefore, representative but efficiently sized samples are recommended to optimize the use of time, financial, and human resources (Krause et al., 2013).

Optimal minimum sample sizes for juice yield ranged from $X_0 = 1.6$ (Diva) to $X_0 = 5.2$ (Hamlin CNPMF-020). For the Valência Tuxpan CNPMF, Uruburetama blood, and Diva varieties, a minimum of two fruits was sufficient. For Pera CNPMF-D6 and Hamlin CNPMF-020, four and five fruits were optimal, respectively (Table 3).

Table 1 – Estimates of parameters *a* and *b** and sample size (X_0) of the fruit of sweet orange varieties (*Citrus × sinensis*), considering mass and diameter (transverse and longitudinal) of the fruit variables.

Varieties	Fruit mass			Diameter					
	<i>a</i>	<i>b</i>	X_0	Transversal			Longitudinal		
				<i>a</i>	<i>b</i>	X_0	<i>a</i>	<i>b</i>	X_0
Diva	26.25	0.49	5.1	8.67	0.49	2.4	9.94	0.50	2.7
Hamlin CNPMF-020	28.92	0.47	5.4	24.90	0.51	5.0	21.66	0.45	4.4
Pera CNPMF-D6	12.78	0.52	3.2	5.00	0.52	1.7	4.40	0.52	1.6
Uruburetama blood	27.27	0.48	5.3	10.02	0.50	2.7	10.21	0.49	2.7
Valência Tuxpan CNPMF	27.55	0.49	5.3	7.48	0.01	2.2	8.90	0.47	2.4

*Estimates of equation parameters of Smith (1938) by fitting the general equation $CV(x) = a/X^b$.

Table 2 – Estimates of parameters *a* and *b** and sample size (X_0) of the fruit of sweet orange varieties (*Citrus × sinensis*), considering the peel color, fruit peel thickness, and juice mass of variables.

Varieties	Fruit peel color			Fruit peel thickness			Juice mass		
	<i>a</i>	<i>b</i>	X_0	<i>a</i>	<i>b</i>	X_0	<i>a</i>	<i>b</i>	X_0
Hamlin CNPMF-020	25.99	0.54	5.2	28.92	0.52	5.5	33.67	0.53	6.1
Pera CNPMF-D6	12.25	0.49	3.1	16.21	0.52	3.8	10.19	0.48	2.7
Uruburetama blood	13.82	0.51	3.4	16.76	0.53	3.9	16.76	0.53	3.9
Valência Tuxpan CNPMF	12.47	0.50	3.1	26.31	0.50	5.2	26.90	0.51	5.3

*Estimates of Smith equation parameters (1938) by fitting the general equation $CV(x) = a/X^b$.

For the technological index, the required sample size ranged from $X_0 = 2.3$ (Diva) to $X_0 = 4.9$ (Hamlin CNPMF-020) (Table 3). This index, essential to juice-processing cultivars, is a calculation based on juice yield, soluble solids content, and a standard box mass of 40.8 kg (Napoleão et al., 2023). Higher values indicate greater industrial suitability.

Titrateable acidity required sample sizes from $X_0 = 3.2$ (Uruburetama blood) to $X_0 = 6.6$ (Valência Tuxpan CNPMF) (Table 4). For soluble solids, five fruits were needed for Hamlin CNPMF-020, and three fruits for the other varieties. Soluble solids are a key quality parameter in both juice processing and fresh markets on account of their association with sweetness.

The SS/TA ratio required the largest sample size for Valência Tuxpan CNPMF (seven fruits) and the smallest for Uruburetama blood ($X_0 = 3.1$). This ratio is widely used to assess maturity in sweet oranges. Ratios above six are acceptable for fresh consumption, though values vary by variety and growing region (Lado et al., 2014).

This study found that the largest optimal minimum sample sizes, up to seven fruits, were associated with chemical quality variables such as SS/TA ratio and titrateable acidity, particularly for Valência Tuxpan CNPMF (Table 4). The smallest sample sizes were observed in juice yield in Diva ($X_0 = 1.6$), Valência Tuxpan CNPMF ($X_0 = 2.1$), and Uruburetama blood ($X_0 = 1.8$). On average, Pera CNPMF-D6 required the fewest fruits for accurate analysis, with sample sizes ranging from $X_0 = 1.6$ (longitudinal diameter) to $X_0 = 4.1$ (titrateable acidity).

Where multiple variables are of interest, as in this study, the optimal sample size should be calculated for each, and the largest among them adopted as the final sample size (Agranonik and Hirakata, 2011). Accordingly, the optimal sample sizes for determining sweet orange fruit quality are four fruits for Pera CNPMF-D6, six for Diva and Uruburetama blood, and seven for Valência Tuxpan CNPMF and Hamlin CNPMF-020.

Sample size depends on several factors, including the variable of interest, variable type (quantitative or qualitative), whether group comparisons are involved, the number of groups, test power, significance level, and effect size (Agranonik and Hirakata, 2011).

Fruit physical and chemical traits vary with environmental factors such as light, temperature, humidity, as well as endogenous factors like gibberellins, carbohydrate and nitrogen levels, and rootstock used (Nawaz et al., 2021), all of which influence the required sample size. The physicochemical variability in 'Palmer' mangoes differed between traits and growing seasons, necessitating different sample sizes (Vilvert et al., 2021).

Genetic improvement programs often involve evaluating numerous genotypes and traits in both plants and animals. However, limitations in labor, time, and resources make sample size optimization crucial (Faria et al., 2020). An appropriate sample size should accurately represent the variety studied and ensure reliable assessments of fruit traits while minimizing resource use.

The Modified Maximum Curvature Method demonstrated that the ideal minimum sample size for sweet orange fruits varies according to variety and variable. Based on physical and chemical quality analyses, the recommended minimum sample sizes are four fruits for Pera CNPMF-D6, six for Diva and Uruburetama blood, and seven for Valência Tuxpan CNPMF and Hamlin CNPMF-020. Thus, using seven fruits can be considered a standardized minimum sample size to ensure accuracy and statistical reliability across different varieties and quality traits under the conditions studied.

Table 3 – Estimates of parameters a and b^* and sample size (X_0) of the fruit of sweet orange varieties (*Citrus × sinensis*), considering the juice yield and technological index of the variables.

Varieties	Juice yield			Technological index		
	a	b	X_0	a	b	X_0
Diva	4.45	0.50	1.6	7.81	0.49	2.3
Hamlin CNPMF-020	26.21	0.54	5.2	24.17	0.50	4.9
Pera CNPMF-D6	17.63	0.52	4.0	14.99	0.52	3.6
Uruburetama blood	54.02	0.51	1.8	11.64	0.51	3.0
Valência Tuxpan CNPMF	6.76	0.50	2.1	11.56	0.52	3.0

*Estimates of equation parameters of Smith (1938) by fitting the general equation $CV(x) = a/x^b$.

Table 4 – Estimates of parameters a and b^* and sample size (X_0) of the fruit of sweet orange varieties (*Citrus × sinensis*), considering the titrateable acidity (TA), soluble solids (SS), and the ratio of total soluble solids/titrateable acidity (SS/TA Ratio) of the variables.

Varieties	TA			SS			SS/TA Ratio		
	a	b	X_0	a	b	X_0	a	b	X_0
Diva	28.60	0.48	5.4	7.93	0.51	2.3	24.67	0.50	5.0
Hamlin CNPMF-020	26.52	0.54	5.3	24.07	0.50	4.9	22.42	0.47	4.6
Pera CNPMF-D6	18.27	0.53	4.1	7.79	0.51	2.3	13.99	0.49	3.4
Uruburetama blood	12.52	0.50	3.2	8.63	0.50	2.5	11.91	0.50	3.1
Valência Tuxpan CNPMF	37.60	0.52	6.6	8.02	0.50	2.4	41.19	0.50	7.0

*Estimates of equation parameters of Smith (1938) by fitting the general equation $CV(x) = a/x^b$.

Authors' Contributions

Conceptualization: Caldas MJM, Sasaki FFC, Soares Filho WS. **Data curation:** Caldas MJM. **Formal analysis:** Caldas MJM. **Investigation:** Caldas MJM, Sasaki FFC, Rodrigues VS. **Methodology:** Ledo CAS, Soares Filho WS. **Resource:** Caldas MJM, Sasaki FFC, Rodrigues VS, Gesteira AS, Passos OS. **Software:** Ledo CAS. **Supervision:** Sasaki FFC, Soares Filho WS. **Visualization:** Santana LGL, Gesteira AS, Passos OS. **Writing-original draft:** Caldas MJM, Sasaki FFC, Soares Filho WS. **Writing-review & editing:** Caldas MJM, Sasaki FFC, Ledo CAS, Gesteira AS, Passos OS, Soares Filho WS.

Conflict of interest

The authors declare no conflict of interest.

Data availability statement

Not applicable.

Declaration of use of AI Technologies

No technology was used.

References

- Agranonik M, Hirakata VN. 2011. Sample size calculation: proportions. *Revista HCPA* 31: 382-388 (in Portuguese, with abstract in English).
- Arellano-Durán N, Villegas-Monter Á, Muñoz-Orozco A. 2018. Optimum sample size in quantitative characteristics of seeds of polyembryonic mango. *Revista Brasileira de Fruticultura* 40: e519. <https://doi.org/10.1590/0100-29452018519>
- Cargnelutti Filho A, Alves BM, Santos GO, Wartha CA, Kleinpaul JA, Silveira DL. 2018. Sample size to estimate the mean and median of traits in flax. *Revista Brasileira de Ciências Agrárias* 13: e5492. <https://doi.org/10.5039/agraria.v13i1a5492>
- Companhia de Entrepósitos e Armazéns Gerais de São Paulo [CEAGESP]. 2011. Normas de Classificação de Citros de Mesa. CEAGESP, São Paulo, SP, Brazil (in Portuguese).
- El-Otmani M, Chouaibi A, Azrof C, Bouchaou L, Choukr-Allah R. 2020. Response of Clementine Mandarin to water-saving strategies under water scarcity conditions. *Water* 12: 2439. <https://doi.org/10.3390/w12092439>
- Faria GA, Lopes BG, Peixoto APB, Ferreira AFA, Maltoni KL, Pigari LB. 2020. Experimental plot size of passion fruit. *Revista Brasileira de Fruticultura* 42: e-125. <https://doi.org/10.1590/0100-29452020125>
- Food and Agriculture Organization of the United Nations [FAO]. 2021. Citrus Fruit Fresh and Processed Statistical Bulletin 2020. Available at: <https://openknowledge.fao.org/server/api/core/bitstreams/4760a5b5-f3b2-41c7-8713-ccdb1a5f8c08/content> [Accessed Apr 21, 2025]
- Krause W, Storck L, Dal'Col Lúcio A, Nied AH, Gonçalves RQ. 2013. Optimum sample size for fruits characters of pineapple under fertilizations experiments using large plots. *Revista Brasileira de Fruticultura* 35: 183-190 (in Portuguese, with abstract in English). <https://doi.org/10.1590/S0100-29452013000100021>
- Lado J, Rodrigo JM, Zacarias L. 2014. Maturity indicators and citrus fruit quality. *Stewart Postharvest Review* 10: 1-6.
- Lessman KJ, Atkins RE. 1963. Comparisons of planning arrangements and estimates of optimum hill plot size for grain sorghum yield tests. *Crop Science* 3: 477-481.
- Lima JF, Peixoto CP, Ledo CAS, Faria GA. 2007. Optimal size of plots for experiments with papaya plants under greenhouse conditions. *Ciências Agrotecnologia* 31: 1411-1415 (in Portuguese, with abstract in English). <https://doi.org/10.1590/S1413-70542007000500021>
- Meier VD, Lessman KJ. 1971. Estimation of optimum field plot shape and size for testing yield in *Crambe abyssinica* Hochst. *Crop Science* 11: 648-650. <https://doi.org/10.2135/cropsci1971.0011183X001100050013x>
- Moura LO, Carlos LA, Oliveira KG, Martins LM, Silva EC. 2016. Physicochemical characteristics of purple lettuce harvested at different ages. *Revista Caatinga* 29: 489-495. <https://doi.org/10.1590/1983-21252016v29n227rc>
- Napoleão GM, Leonel S, Souza JMA, Leonel M, Martins RC, Cardoso CP, et al. 2023. Germplasm diversification in citrus orchards in a mesothermal climate in Brazil. *Agriculture* 13: 1551. <https://doi.org/10.3390/agriculture13081551>
- Nawaz R, Khan MA, Hafiz IA, Khan MF, Khalid A. 2021. Climate variables effect on fruiting pattern of Kinnow mandarin (*Citrus nobilis* Lour x *C. Deliciosa* Tenora) grown at different agro-climatic regions. *Scientific Reports* 11: 18177. <https://doi.org/10.1038/s41598-021-97653-1>
- Smith HF. 1938. An empirical law describing heterogeneity in yields of agricultural crops. *Journal Agricultural Science* 28: 1-23.
- Uthman A, Garba Y. 2023. Citrus Mineral Nutrition and Health Benefits: A Review. *Citrus Research-Horticultural and Human Health Aspects*. Intechopen, London, UK. <https://doi.org/10.5772/intechopen.107495>
- Vilvert JC, Freitas ST, Ferreira MAR, Costa EBS, Aroucha EMM. 2021. Sample size for postharvest quality traits of 'Palmer' mangoes. *Revista Brasileira de Fruticultura* 43: e-014. <https://doi.org/10.1590/0100-29452021014>