Arboriculture & Urban Forestry 2015. 41(5): 279–285

An Analysis of the Field Precision of the CTLA Trunk Formula Method

James Komen and Donald R. Hodel

Abstract. The CTLA Trunk Formula Method is an industry-standard tool for appraising large trees. The goal of this study was to measure its precision in the field and to look for possible ways to improve the formula or its implementation. Fourteen certified arborists independently appraised the same ten trees, and the results of their appraisals were analyzed. This study focused on the attributes of Trunk Area, Species, Location, and Condition. In the results, the attributes with the highest variance among appraisers were Trunk Area and the Condition Rating. In the past, much of the variation among appraisers has been attributed to personal bias due to lack of experience, and it has been suggested that variance would decrease with experience. These results give evidence to the contrary—the group of appraisers with the highest variance was the group that performed appraisals most frequently. The most valuable information from this study was the identification of four key elements of error involved in the appraisal process: personal value error, personal observation error, measurement error, and systematic error.

Key Words. Cross-Sectional Area; Guide for Plant Appraisal; Measurement Error; Personal Observation Error; Personal Value Error; Rating; Systematic Error; Tree Appraisal; Trunk Formula Method.

Arborists and others in the arboricultural and landscape industries frequently perform monetary tree appraisals. Because such appraisals establish tree values for legal, insurance, monetary transactions, and other purposes, it is critical that they are accurate and precise. Appraisal methods typically follow a formula, and several have been used, compared, and evaluated (Watson 2001; Watson 2002).

One such appraisal method is outlined in the *Guide for Plant Appraisal* (CTLA 2000), which is commonly used in the United States and Canada. It is referred to as the "CTLA Trunk Formula Method" (CTLA TFM) because it is based on measuring the trunk cross-sectional area and multiplying that by a monetary value per square inch. This basic value is then depreciated or reduced by factors for species, condition, and location of the tree. Cullen (2007) provided a comprehensive review of this method.

Precision is desired and expected in any tree appraisal method. An indicator of the precision of a formula is the variance between values calculated by different appraisers on the same tree using the same formula. A small variance suggests that the formula is precise and tends to preclude appraiser bias or inaccuracy. Conversely, a large difference indicates that the formula is imprecise, probably because of appraiser bias and/or inherent problems with the methodology or formula.

The CTLA TFM has been criticized for the high degree of variation sometimes encountered among appraisers, often 100% to 200% or higher (Kielbaso 1979; Rey-Lescure 1985; Abbot and Miller 1991). In previous studies, most of this variation was due to differences in condition and location ratings (Watson 2001), which many practitioners have contended are too subjective (Davis 1983; Tate 1989; Abbot and Miller 1991). Chadwick (1975) felt that large trees have unrealistically high appraisal values, primarily because the error of an area calculation is directly proportional to both the circumference measurement and the error of the circumference measurement. Therefore, the percentage error in measurement size increases with the size of the tree; larger trees will have a larger error in value simply because the error from measurement scales upward.

Watson (2001) suggested a method for improvement: the publication of databases of appraised tree values. These databases would not eliminate the need for a formula but would provide appraisers with established documented values to which they can compare their values. Watson provided tables of comparable appraised values for trees by species and size classes derived from more than 13,000 actual formula method appraisals, which help the appraiser to test the reasonableness of appraised values.

The goal in this study was to compare and evaluate appraised values of several trees by several appraisers using the CTLA TFM. An analysis of the results of this study could help to identify problems of precision and/or methodology of this formula, if any, and suggest ways to improve it. To test the variability between appraisers, the study authors set up a study where individuals could independently appraise the same trees without influencing each other's results.

Although the original intent of the study was to calculate the standard deviation of each of the measured and calculated values, the most important discovery was the identification of the four key elements of error that had not been specifically identified in previous research on tree appraisals: personal value error, personal observation error, measurement error, and systematic error. Although the design of this study did not isolate these variables, the identification of these elements should be the foundation of a future study.

MATERIALS AND METHODS

The authors used the same method Watson (2002) used for tree selection and data collection, but only gathered data for the CTLA method of appraisal and added the measurements of trunk circumference and species rating. It was felt that this method was an effective way of collecting data without placing an undue burden on the study participants.

In the study, the authors wanted to mimic the experience of the typical field appraisal, while still retaining control over tree selection, so as to allow comparisons between different appraisers. The local arboretum was chosen as the study location because it possessed a good variety of trees to choose from and good public access. The study was conducted at the Los Angeles County Arboretum & Botanic Garden in Arcadia, California, U.S., between 15 July and 15 August 2014. Ten trees were located and identified, and 14 certified arborists were invited, independently, to appraise these trees according to their availability over the four-week period. Each tree was marked on a map, showing its location and its photograph for identification. The authors selected the trees to include a variety of health, location, and species ratings. Some trees were deliberately selected to include a trunk configuration that is not clearly defined according to the *Guide for Plant Appraisal*. The condition of the trees did not vary significantly over these four weeks. Each of the arborists were provided with a clipboard, a map, a form to fill out, a copy of the *Guide for Plant Appraisal*, and the same tape measure (30.48 m).

Participating arborists were greeted individually upon their arrival at the arboretum and were provided with the study materials. The authors gave a short introduction and overview to the study, explained their participation, duties, and responsibilities, and then observed each conducting a practice appraisal on a tree separate from the 10 subject trees. Explicit instructions on how to conduct the appraisal were not provided; however, the authors did respond to questions pertaining to the proper completion of the data collection form. This was so the authors could avoid adding their own experimenter's biases to the opinions of the participants. The participants were expected to appraise the trees the same way they do in their professional practice.

Participants conducted their data collection independently, without guidance or influence. The authors could have conducted this study all at once, leading the participants to each tree together as a group, but this would have influenced their results because they would have observed each other looking at each subject tree, making them more likely to see defects they would not have otherwise seen.

Participants made their circumference measurements and subjective ratings for the species, location, and condition factors and recorded this data on the forms provided. The study authors provided the standard species ratings published by the local Western Chapter of the International

Society of Arboriculture (ISA); participants only had to adjust these slightly according to their interpretation. The formula restricts participants from adjusting the rating more than 10% from the published value. The standard nursery price and installation cost for each species were also provided. The authors provided some of the input data to save the appraisers' time and to focus on four key variables in this study. If each appraiser were required to collect the additional data, then it would not have been possible for each appraiser to gather the field data for all ten trees in less than four hours, and it would have been difficult to obtain as many voluntary participants. Appraisers did not calculate final appraisal values based on their data; the authors did this for them to avoid mathematical error.

Following the initial data collection period, the participants were polled via email and asked to declare their field of arboriculture, years of experience, and the frequency they use the CTLA TFM. This information was used to segment the participants' data into groups based on their responses. Most notably, the authors tested the effect of frequency of appraisal on precision of results.

Using the participants' data, final appraised values were calculated for each of the 10 trees for each of the 14 arborists. Each of the measured variables (circumference, species, condition, and location) was assumed to be an independent random variable. Means and standard deviations were calculated for the final appraised values and for the four input variables. Finally, the standard deviations were divided by the mean appraised value to arrive at a percentage standard deviation to compare results between trees of different values.

RESULTS

Table 1 provides background information of study participants. They practiced approximately equally between private and municipal operations. One came from nursery production, another from forestry, and one from arboreta management. Experience ranged from two years to 42 years. The frequency of use was evenly distributed between "Frequent," "Occasional," and "Rare/Never."

Mean appraised values ranged from as low as USD $9,000 for *Liriodendron tulipifera* to as much as nearly $80,000 for *Schinus terebinthifolius* (Table 2). Similarly, standard deviations of the mean appraised values ranged from about $2,000 in *L. tulipifera* to over $31,000 in *S. terebinthifolius*. Percentage standard deviation of the mean appraised values of all appraisers ranged from 21% in *Sequoia sempervirens* to 42% in *Jacaranda mimosifolia*.

Of the four variables used to calculate the appraised value, trunk area had the greatest standard deviation (Table 3). The standard deviation percent ranged from 3% in *Liriodendron tulipifera* and

Table 1. Background of Study Participants, Los Angeles County Arboretum and Botanic Garden, Arcadia, California, U.S., 15 July to 15 August 2014.

Field of arboriculture[z]		Experience (years)		Appraisal frequency	
Private	9	<20	4	Frequent	3
Municipal	6	20 < x < 31	4	Occasional	4
Landscape architecture	1	>31	6	Rare/Never	7
Arboretum	1				
Nursery production	1				
Plant health care	1				

[z] Of 14 total participants, several declared multiple fields of arboriculture.

Table 2. Mean appraised value (USD $), standard deviation (USD $), and standard deviation percent (%) of 10 trees, Los Angeles County Arboretum and Botanic Garden, Arcadia, California, U.S., 15 July to 15 August 2014.

	Total value		
Tree species	Mean appraised value	SDV value	SDV percent
S. molle	$36,696.32	$12,170.35	33%
Q. agrifolia	$60,415.38	$20,900.34	35%
S. terebinthifolius	$77,586.19	$31,298.00	40%
C. insignis	$29,230.49	$6,740.12	23%
L. tulipifera	$9,044.88	$2,044.13	23%
F. elasticoides	$33,425.21	$11,572.35	35%
J. mimosifolia	$14,312.28	$5,971.45	42%
P. orientalis	$12,552.69	$3,838.26	31%
S. sempervirens	$30,574.91	$6,569.32	21%
E. contortisiliquum	$39,084.30	$14,868.29	38%

Ficus elasticoides to 49% in *Schinus terebinthifolius*. Those trees with single trunks had lower standard deviations than those with multiple trunks, suggesting that the error lies not with the physical measurement component but with the subjective choice of where to place the measuring tape.

Of the three depreciation factors, condition had the greatest standard deviation percent and location and species had the least (Table 4). However, the low standard deviation of the species rating could be due partly to the artificial constraints of the formula that restricts participants from adjusting the rating more than 10% from the published value.

Each tree's calculated percentage deviation was sorted and analyzed by the frequency that the participants declared they use the CTLA TFM. A general trend line was calculated by linear regression and overlaid atop the plotted data (Figure 1). The group that declared they were frequent appraisers tended to have the highest standard deviations while those that declared they were infrequent appraisers tended to have the lowest.

Table 3. Number of trunks and standard deviation percent of 10 trees, Los Angeles County Arboretum and Botanic Garden, Arcadia, California, U.S., 15 July to 15 August 2014.

	Trunk area	
Tree species	Highest number of measured trunks	SDV percent
S. molle	1	6%
Q. agrifolia	1	22%
S. terebinthifolius	5	49%
C. insignis	3	13%
L. tulipifera	1	3%
F. elasticoides	1	3%
J. mimosifolia	2	24%
P. orientalis	1	7%
S. sempervirens	1	6%
E. contortisiliquum	3	17%

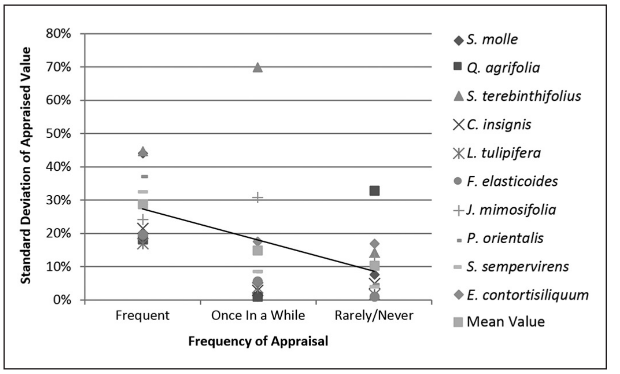

Figure 1. Relationship of frequency of appraisal to calculated standard deviation for each of the 10 trees with overlaid trend line, Los Angeles County Arboretum and Botanic Garden, Arcadia, California, U.S., 15 July to 15 August 2014.

DISCUSSION

While there have been several studies looking at the theoretical mechanics of the CTLA TFM, Watson (2002) was the only previous study involving field comparison of multiple appraisals on the same trees; however, that research specifically stated in its conclusion that its intention was not to judge or rank the appraisal formulas. In contrast, the current study was intended to identify weaknesses and sources of error and to explore options for mitigation of error and improvement of the formula.

The CTLA TFM appraisal method is only as precise as its least precise component. Each of the four variable values in the formula are multiplied together, so any percentage change in one value is reflected by the same percentage change in the final value due to the commutative property of multiplication. Large standard deviations were observed in the condition and size ratings, and these had pronounced effects on the final appraised values, despite the relatively lower species and location rating standard deviations.

Table 4. Means and standard deviations of location, species, and condition ratings of 10 trees, Los Angeles County Arboretum and Botanic Garden, Arcadia, California, U.S., 15 July to 15 August 2014.

	Location		Species		Condition	
Tree species	Mean rating	SDV %	Mean rating	SDV %	Mean rating	SDV %
S. molle	83%	6%	69%	5%	69%	17%
Q. agrifolia	92%	5%	92%	4%	65%	13%
S. terebinthifolius	93%	8%	73%	8%	76%	15%
C. insignis	89%	8%	73%	5%	70%	14%
L. tulipifera	85%	6%	71%	3%	79%	11%
F. elasticoides	79%	13%	53%	6%	80%	12%
J. mimosifolia	80%	12%	89%	3%	70%	15%
P. orientalis	77%	10%	86%	5%	63%	13%
S. sempervirens	89%	7%	74%	6%	83%	11%
E. contortisiliquum	89%	8%	53%	6%	69%	15%

For the *Sequoia sempervirens*, the tree with the smallest level of error, the standard deviation was 21%, meaning that approximately 33% of the possible values deviated more than 21% from the mean value. *Schinus terebinthifolius*, the tree with the highest standard deviation of 40%, had an observed range of values greater than the average value of the tree. These standard deviations illustrate an unacceptable level of subjective error, but there are several ways to improve the method and its application to make it more precise.

Observation of the participants on the practice tree yielded some insight to the field application of the formula. For example, many of the participants expressed dissatisfaction with the limited scale of 1–4 for each of the components of the condition rating. They found themselves frequently indecisive between a 2 (minor defects) and 3 (major defects) rating. In these situations, some biased the higher value and others biased the lower value. This small but biased difference ultimately had a final effect on the condition rating of 3% for each rating action. These small errors can accumulate to a significant effect on the final appraised value. This limited type rating scale provides limited options for the appraiser; it could be mitigated by introducing a larger set of possible values (for example, a scale of 1–9) and one with a single "middle value" as in a scale of 1–5 or 1–9.

On the practice tree, each of the participants observed a unique defect. One saw a large crack in a scaffold branch that none of the others observed; another observed stem galls proliferating in the canopy. Only one arborist correctly identified a white crusting on the top of a section of one branch as bird feces. Furthermore, each participant arrived at his or her conclusion for location and species ratings through different reasoning, even when their final values were the same. These observations illustrate the process by which appraiser bias is incorporated into final appraised values.

From observations in this study, the authors propose delineating the error observed from appraiser bias into two categories: personal *value* error and personal *observation* error. Personal observation error would be the error of an appraiser observing a different set of attributes of a tree than another appraiser. An appraiser who observes a decaying crack in the root crown would assign a lower condition rating than an appraiser who fails to observe the crack. The two arborists may have the same set of personal values, but because of the difference in input (observation) information, they would arrive at different conclusions. Personal value error would occur when an appraiser observed the same set of one tree's attributes but concluded a different value. Two appraisers may observe the same decaying crack, but they may disagree on the severity and its effect on the tree's final appraised value. In this study, these two errors were confounded in the numerical data; a future study should attempt to isolate and measure their relative effect on total appraiser bias. To isolate the personal value error, participants should be given a list of defects on paper and asked to rate the condition with the limited information. To isolate the personal observation error, participants should independently write down their observations on a set of trees, and the lists can be compared for differences.

The set of participants for this study was allowed to include all ISA Certified Arborists® because there is currently no credential that differentiates a qualified plant appraiser and because Certified Arborists are considered Experts in a court of law. Aside from years of arboricultural experience, there is no defining attribute that would make one Certified Arborist's appraisals more valid than another Certified Arborist's.

While Watson (2001) found that variation in the appraised values was mostly due to differences in condition and location, the authors of the current study found the greatest differences in trunk area, followed by condition and location. The standard deviation in trunk area was directly correlated with the highest number of measured trunks. Appraisers calculated trunk area by using different configurations of the measuring tape on the same tree. As a result, the calculated trunk area had a higher standard deviation than on trees with only one observed trunk. Furthermore, the error was greater on larger trees because the error of an area calculation is directly proportional to both the circumference measurement and the error of the circumference measurement. Therefore, the percentage error in measurement size increases with the size of the tree; larger trees will have a larger error in value simply because the error from measurement scales upwards.

The authors found two components of error in the trunk cross-sectional area. The first component is the *measurement* error; two appraisers may place the same tape at the same place on the tree and record different values. This error is minimal, as illustrated by the low observed errors on the large, single-trunk trees. The more critical component of error is the systematic error. In this case, the *systematic* error is the decision of where to place the measuring tape, especially on multi-trunked trees. Although many scenarios are detailed in the *Guide for Plant Appraisal*, the synthesis of these concepts and their practical application have room for improvement. Future training courses in tree appraisal should address a larger set of possible trunk configurations to reduce the discretion left to the appraiser for determining where to measure the tree.

The location rating did not show a high level of standard deviation for this experiment, which is likely due to the set of subject trees being located exclusively at an arboretum. Each of the trees was either deliberately planted or the landscape was deliberately designed around them, reducing the likelihood that they would receive low placement or contribution ratings. The expert opinions guiding the planning decisions helped to ensure that mistakes were not made in the design of the arboretum, so contribution and placement ratings tended to be high. The arboretum represents the highest value site that can contain a tree, so nearly all the site ratings were between 90% and 100%. It is likely that if this experiment were conducted on a set of trees on private properties, there would be a greater observed standard deviation in the location rating.

The rating with the lowest standard deviation was the species rating. There were two mechanisms that are reduced its variance. The first mechanism is the artificial restriction by the *Guide for Plant Appraisal* that the species rating can be modified by no more than 10% from the value published in regional species rating classification guides. There were several participants who expressed dissatisfaction with their inability to boost or penalize a tree more than 10%. If this artificial restriction were lifted, the species rating would have a higher standard deviation.

The other mechanism buffering the species rating is the use of a "starting value." Rather than asking the participants to assign a value without a standard of reference, the species rating is given as a starting value, and participants were asked if they felt it needed to be adjusted. About half the participants used the same values that were published. This level of correlation would have been unlikely if there were no starting values from which to deviate. The concept of using a set of starting values should be considered for application to the other depreciation ratings as a way to reduce their variability as well.

This study offers evidence against the conventional assumption that valuation variance decreases with level of experience. The precision of the "Frequent" appraisers was the lowest, suggesting that consistency and precision are not necessarily achieved by more appraisal experience; rather, the personal value biases of individual appraisers could be further reinforced through practice. This finding also confirms that the inclusion of participants less experienced in appraisals was still appropriate for this study because their data set did not include outliers that skewed the results. The solution to the problem of error in experienced appraisers is not to encourage appraisers to practice independently more frequently. Rather, it is to offer training courses and literature that identify the four key elements of error and to work toward mitigating them through extensive scenario-based education.

CONCLUSIONS

There are two major categories of improvement for the CTLA TFM: changes to the formula itself and changes to its implementation. If the formula is revised, condition attributes should be rated on a scale of 1–5 to include a "middle value" to reduce the variability from personal bias. Future improvement on the implementation of the formula should include a certification course endorsed by the ISA that reduces error through training. The areas of lowest precision were the trunk area measurements and the condition ratings. These should therefore be the areas that receive the most attention in future training or certification courses.

Future research in tree appraisals should attempt to isolate the components of error and suggest priorities for mitigation of the error. By reducing the error in the formula through improved implementation, the level of the precision will increase, and there will be a higher perceived efficacy of the appraisal methodology and its value to the industry and clientele.

Acknowledgments. We thank Jim Henrich and Frank McDonough of the Los Angeles County Arboretum and Botanic Garden in Arcadia for allowing and facilitating this study. We thank each of the 14 participating arborists: Mike Mahoney, Brandon Watts, Mark Martinez, Delia Juncal, Kathy Copely, Glen Williams, Kevin Holman, Jim Borer, Robert Sartain, Frank Madero, Michael Frilot, Deborah Geisinger, Kerry Norman, and Carl Mellinger. Ken Greby also offered valuable insight into this study.

LITERATURE CITED

Abbot, R.E., and C. Miller. 1991. Utility tree claims. Journal of Arboriculture 17:113–116.

Chadwick, L.C. 1975. ASCA recommendations for modification of the ISTC shade tree evaluation formula. Journal of Arboriculture 1:35–38.

CTLA. 2000. Council of Tree & Landscape Appraisers. Guide for Plant Appraisal, ninth edition. International Society of Arboriculture, Champaign, Illinois, U.S.

Cullen, S. 2007. Putting a value on trees—CTLA guidance and methods. Arboricultural Journal (Great Britain) 30:21–43.

Davis, S.H. 1983. Tree valuation pitfalls. Journal of Arboriculture 9:164–166.

Kielbaso, J.J. 1979.Evaluating trees in urban areas. Journal of Arboriculture 5:70–72.

Rey-Lescure, E. 1985. The location factor when evaluating shade trees and the replacement value in appraising hedges—Two proposals. Journal of Arboriculture 11:113–136.

Tate, R. 1989. ISA tree evaluation guide: A critical examination. Journal of Arboriculture 15:145–149.

Watson, G. 2001. A study of CTLA Formula Values. Journal of Arboriculture 27(6):289–297.

Watson, G. 2002. Comparing formula methods of tree appraisal. Journal of Arboriculture 28(1):11–18.

James Komen (corresponding author)
Class One Arboriculture Inc.
1061 Jedburgh St
Glendora, California 91740, U.S.
jameskomen@gmail.com

Donald R. Hodel
University of California
Cooperative Extension
700 W. Main Street
Alhambra, California 91801, U.S.
drhodel@ucanr.edu

Résumé. La méthode de calcul de la surface terrière du tronc que promeut le CTLA (Council of Tree and Landscape Appraisers) est un outil standard de l'industrie pour évaluer les gros arbres. Le but de cette étude était de mesurer son exactitude sur le terrain et de chercher des moyens pour améliorer ladite formule ou sa mise en application. Quatorze arboriculteurs certifiés ont évalué indépendamment les mêmes dix arbres et les résultats de leurs évaluations ont été analysés par la suite. Cette étude a porté sur les éléments de la surface terrière, de l'essence, de la localisation et de la condition. Suite à l'analyse des résultats, les caractéristiques qui comportaient la plus grande variation entre les évaluateurs étaient la surface terrière et la cote de la condition. Par le passé, une grande partie des disparités entre les évaluateurs était attribuée à des opinions préconçues du fait d'un manque d'expérience et il a été suggéré que la variance diminuerait avec l'accumulation des années d'expérience. Les résultats actuels témoignent du contraire, le groupe d'experts ayant le plus grand écart était le groupe qui effectuait le plus fréquemment des évaluations. L'information la plus pertinente découlant de cette étude fut l'identification de quatre causes d'erreurs impliquées dans le processus d'évaluation: erreur d'expertise personnelle, erreur d'observation personnelle, erreur de mesure et erreur systématique.

Zusammenfassung. Die CTLA Trunk Formula Methode ist eine verbreitete Standartmethode zur Abschätzung von großen Bäumen. Das Ziel dieser Studie war, die Präzision vorort zu überprüfen und nach möglichen Verbesserungen der Formel oder ihrer Anwendung zu suchen. Vierzehn zertifizierte Arboristen untersuchten unabhängig die gleichen zehn Bäume und die Resultate der Untersuchungen wurden analysiert. Diese Studie fokussiert auf den Attributen der Stammanteil, Baumart, Standort und Kondition. In den Resultaten waren die Attribute mit der höchsten Varianz der Stammanteil und Vitalitätsbeurteilung. In der Vergangenheit wurden viele Variationen bei der Beurteilung dem Personal zugeschrieben, wegen mangelnder Erfahrung, und es wurde vermutet, dass die Varianz sich mit zunehmender Erfahrung reduziert. Diese Ergebnisse liefern Beweis für das Gegenteil – die Gruppe der Kontrolleure mit der höchsten Varianz war die Gruppe, die die meisten Untersuchungen durchführten. Die wertvollste Information aus dieser Studie war die Identifikation von vier Schlüsselelementen von Fehlern in dem Begutachtungsprozeß: persönlicher Bewertungsfehler, persönlicher Beobachtungsfehler, Meßfehler und systematischer Fehler.

Resumen. El método de la Fórmula del Tronco, CTLA, es una herramienta estándar de la industria para evaluar los árboles grandes. El objetivo de este estudio fue medir su precisión en el campo y buscar posibles formas de mejorar la fórmula o su implementación. Catorce arboristas certificados evaluaron de forma independiente los mismos diez árboles y se analizaron los resultados de sus evaluaciones. Este estudio se centró en los atributos de Área del Tronco, Especies, Ubicación y Condición. En los resultados, los atributos con mayor varianza entre los evaluadores fueron Área del Tronco y Clasificación de Condición. En el pasado, gran parte de la variación entre los tasadores se ha atribuido a prejuicios personales, debido a la falta de experiencia, y se ha sugerido que la varianza disminuiría con la experiencia. Estos resultados dan evidencia de lo contrario - el grupo de tasadores con mayor varianza fue el grupo que lleva a cabo evaluaciones con más frecuencia. La información más valiosa de este estudio fue la identificación de cuatro elementos clave de error que intervienen en el proceso de evaluación: error personal de valor, error de observación personal, error de medición y error sistemático.