POLITECNICO DI TORINO
Repository ISTITUZIONALE

Lot Acceptance Under Inspection Errors: Benefits of Multiple-Inspector Sampling Versus Single-Inspector 100%

Original
Lot Acceptance Under Inspection Errors: Benefits of Multiple-Inspector Sampling Versus Single-Inspector 100% Inspection / Maisano, D.A.F., Ferrara, L., Franceschini, F.. - In: QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL. - ISSN 0748-8017. - STAMPA. - 42:7(2026), pp. 3736-3749. [10.1002/qre.70291]

Availability:
This version is available at: 11583/3015845 since: 2026-09-23T16:05:16Z

Publisher:
WILEY

Published
DOI:10.1002/qre.70291

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

18 November 2026

Quality and Reliability Engineering International

WILEY

RESEARCH ARTICLE OPEN ACCESS

Lot Acceptance Under Inspection Errors: Benefits of Multiple-Inspector Sampling Versus Single-Inspector 100% Inspection

Domenico A. Maisano | Lucrezia Ferrara | Fiorenzo Franceschini

Politecnico di Torino, Torino, Italy

Correspondence: Fiorenzo Franceschini (fiorenzo.franceschini@polito.it)

Received: 16 February 2026 | **Revised:** 22 April 2026 | **Accepted:** 2 June 2026

Keywords: 100% inspection | conformity assessment | inspection error | multiple inspectors | operating-characteristic curve | single sampling plan

ABSTRACT

In high-value and safety-critical manufacturing contexts (e.g., aerospace, nuclear and defense sectors), reliable and robust conformity assessment of production lots is imperative. While 100% inspection by a single inspector may seem like a solution for this purpose, it can be prone to human error, which can distort lot acceptance decisions. This study investigates whether a multiple-inspector sampling inspection—that is, an inspection limited to a portion of the lot but conducted by a plurality of independent inspectors—can outperform a 100% inspection by a single inspector, in these contexts. Operating-characteristic (OC) curve models were developed for both approaches, taking inspection errors into account and comparing their performance through analytical formulations and realistic examples. To simplify the analysis, this contribution focuses on sampling inspections based on single sampling plans (SSPs). Results demonstrate that incorporating multiple independent inspectors into a SSP can significantly stabilize the OC curve, maintaining lot-acceptance behaviour close to the ideal error-free scenario. Moreover, the multiple-inspector sampling strategy tends to reduce the total number of inspections needed, compared to 100% inspection, especially for moderate numbers of inspectors. In conclusion, these findings suggest that multiple-inspector sampling is a practical and effective way to obtain robust lot-acceptance decisions, protecting them from inspection-error distortions.

1 | Introduction

In several high-value, safety-critical and highly customised manufacturing contexts (e.g., aerospace/defense parts, nuclear, *haute couture* and luxury goods), reliable conformity assessments of production lots are a must. Unfortunately, inspections—especially visual or manual ones—are sometimes subject to human errors by inspectors (i.e., all inspectors evaluate the same sample), thus leading to distorted final *lot-disposition decisions* (LDDs). Multiple-inspector sampling plans have emerged as a practical way to increase the reliability of conformity assessment: A sample of units drawn from a lot is inspected independently by multiple inspectors and then individual assessments are aggregated into a final LDD [1, 2]. Parallel inspections are often required where the potential critical issues and reputational risk associated with producing nonconforming products are high [3, 4]. Moreover, prior studies and industrial practice indicate that redundant human inspections can strengthen the reliability of LDDs in practical contexts where single-inspector accuracy is uncertain [5, 6].

A key motivation for multiple-inspector sampling plans is the intrinsic variability of inspectors' conformity assessments, which is related to factors such as experience, training, attention, subjectivity in judgements and familiarity with product models (which are often subject to change). Redundant, independent assessments on the same sample can not only mitigate subjectivity but also reveal misalignment among inspectors and monitor

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Quality and Reliability Engineering International* published by John Wiley & Sons Ltd.

the internal consistency of the inspection process; these insights enable targeted training and better governance of the whole inspection process [7, 8, 2].

When conformity assessment is not error-free, each inspector may commit two primary error types in inspections [9]: *false conforming* (i.e., misclassifying a defective unit as conforming) and *false defective* (i.e., misclassifying a conforming unit as defective). Consequently, the lot *defectiveness*[1] as perceived by an inspector can appear distorted, with implications for the resulting LDD [10–12]. Inspection errors generally tend to flatten the OC *curve*—that is, the graph of lot acceptance probability (P_a) as a function of lot defectiveness (p)—thereby distorting its discriminatory power [9, 13, 14].

This paper builds on a previous contribution on OC curves for multiple-inspector *single sampling plans* (SSPs) with inspection errors, which provided rigorous Type-A (*hypergeometric*) and Type-B (*binomial*) formulations of the OC curves [2]. That contribution analysed how the *criteria* and *sequences* for aggregating individual inspectors' conformity assessments and the associated inspection errors can shape the OC curves. A key finding was the relative robustness provided by aggregating conformity assessments using the *majority* criterion, which keeps the OC curve close to what it would be in the ideal case of no inspection errors. By contrast, using a *unanimity* aggregation criterion yields OC curves that are excessively severe unless error rates are negligible [8, 15].

Building on that framework, the present study addresses a natural but under-appreciated research question (**RQ0**): How does a multiple-inspector sampling plan—which involves inspecting only a portion of the lot with multiple inspectors—compare with a 100% inspection carried out by a single inspector on the entire lot? Intuition might suggest that 100% inspection of an entire lot by one inspector should dominate any sampling plan. However, in realistic settings where the single inspector is error-prone, a 100% inspection does not guarantee better lot protection and, instead, can produce systematic distortions in LDDs. Conversely, sampling the lot with multiple inspectors can provide (i) A more stable OC curve (with a more controllable protection level), and (ii) Substantial savings in total inspection effort [8, 16, 13].

In an ideal 100% lot inspection by a single inspector (with no inspection errors), the OC curve would be 'step-shaped' [9]: below a certain threshold D^*—corresponding to the maximum tolerated number of defectives, as discussed later—the lot is deterministically accepted (i.e., $P_a = 100\%$), and above that threshold it is deterministically rejected (i.e., $P_a = 0\%$). Translating this into a formula:

$$P_a = \begin{cases} 10 \text{ (acceptance)} & \text{if } \leq D \leq D^* \text{ or } 0 \leq p \leq p^*, \\ 0 \text{ (rejection)} & \text{if } D > D^* \text{ or } p > p^*, \end{cases} \text{ being } p^* = \frac{D^*}{N}, \tag{1}$$

where D is the number of defective units in the entire lot, which consists of N units. Figure 1 illustrates the ideal step-shaped OC curve, which is used throughout the rest of the paper as a reference to quantify the distorting effect of inspection errors

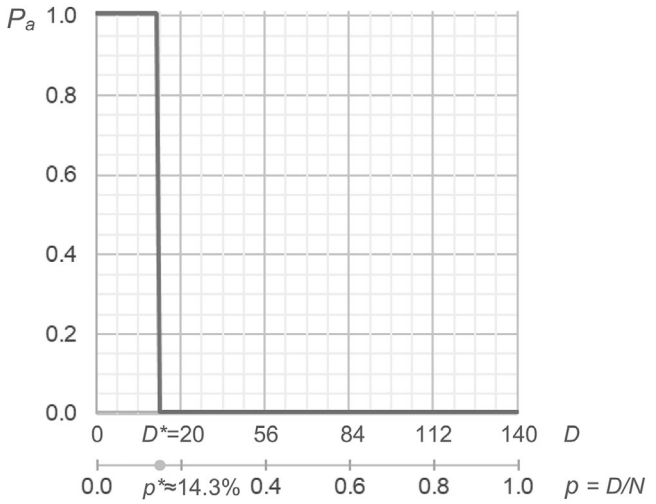


FIGURE 1 | Step-shaped OC curve characterizing the (ideal) case of 100% inspection with no inspection errors. This specific OC curve is constructed for a lot of $N = 140$ units, with an acceptance number of defectives $D^* = 20$; corresponding to a defectiveness of $p^* = D^*/N = 20/140 \approx 14.3\%$ (at the 'step' of the OC curve), which is the maximum tolerated for lot acceptance. To facilitate readability, the horizontal axis includes a dual scale in terms of both D and p.

and to motivate the multiple-inspector approach. Because $D \in \{0, 1, 2, \ldots, N\}$ (being $D \in \mathbb{N}_0$) and consequently $p = D/N$ also takes discrete values, the OC curve is inherently discrete rather than continuous. For graph readability reasons, however, it is plotted as a continuous curve.

Inspection errors can modify this ideal OC curve in an uncontrolled way, as will be shown later. In contrast, this research will evaluate the effectiveness of the multiple-inspector approach in mitigating inspection-error effects. As hinted by the title of this article, the underlying hypothesis is that multiple-inspector sampling can outperform single-inspector 100% lot inspection, in both (i) Control of the protection level and (ii) Inspection effort, especially when inspection errors are non-negligible or uncertain [8, 9, 13, 14]. Based on these preliminary considerations, **RQ0** can be broken down into three more specific research questions that are formulated to guide this study:

RQ1: Between 100% lot inspection by a single inspector and a multiple-inspector sampling plan, which approach ensures a better protection, in the presence of inspection errors?

RQ2: How do the number of inspectors and their individual error rates influence the OC curve in a multiple-inspector sampling plan?

RQ3: From the perspective of inspection effort, which of the two approaches is more advantageous?

The remainder of the article is organised into five sections. Section 2 formalizes OC-curve models for a single-inspector 100%

lot inspection, under ideal and error-prone conditions. Section 3 recalls the formulation of the OC curve for a multiple-inspector sampling plan with inspection errors; to simplify the description, the analysis is limited to SSPs. Section 4 presents a quantitative comparison between the two approaches, based on a realistic example that explores the impact of (i) The magnitude of inspection errors (*low* vs. *high*) and (ii) The number of inspectors m. Section 5 discusses implications, limitations and future research directions. Finally, the appendix contains additional material for further study of the analysis carried out.

2 | Modelling the Single-Inspector 100% Lot Inspection

As shown by Figure 1, in the ideal case of 100% inspection by a single inspector who makes no inspection errors, the OC curve is a step-shaped function: the lot is accepted if and only if D does not exceed a conventional threshold D^*, which is generally pre-agreed between supplier and customer (see the LDD in Equation 1). Assuming that the lot includes N total units, this is equivalent to stating that p must not exceed a certain threshold; in the example of Figure 1, with $N = 140$ and $D^* = 20$, this threshold corresponds to:

$$p^* = \frac{D^*}{N} = \frac{20}{140} \approx 14.3\% \tag{2}$$

Let us now consider the more realistic case of a single inspector who makes inspection errors. In this case as well, the inspector must inspect a lot of N units, containing D defective units (a quantity generally unknown to the inspector); the remaining $(N - D)$ units are conforming. After inspecting each unit, the inspector classifies it as conforming or defective. During the inspection, the inspector can make the following two types of errors, with associated probabilities (or error rates):

- a is the probability of misclassifying a conforming unit as defective;
- b is the probability of misclassifying a defective unit as conforming.

Applying the LDD rule of Equation 1, the count of D can be distorted by the errors a and b. Two independent random variables are defined:

- D_1' is the number of 'true' defectives (i.e., truly defective units in the lot) that the inspector correctly classifies as defective. Since the lot contains D defectives and the probability of correctly classifying each of them as defective is $(1 - b)$, D_1' follows a binomial distribution:

$$D_1' \sim Bin\,(D, 1 - b)\,; \tag{3}$$

- D_2' is the number of 'false' defectives among the $(N - D)$ true conforming units (i.e., the number of conforming units misclassified as defective). The probability of misclassifying a conforming unit as defective is a, so:

$$D_2' \sim Bin\,(N - D, a)\,. \tag{4}$$

The total 'apparent' number of defectives in the inspector's view is then:

$$D' = D_1' + D_2'. \tag{5}$$

The adjective 'apparent' emphasises the fact that the inspector's classification of defectives may be distorted by the error rates (a and b). Since D' is obtained by summing two variables with binomial distributions and different probabilities, it will be a discrete variable with a specific probability distribution [15]. Precisely, assuming that D_1' and D_2' are independent, the distribution of D' is the convolution of the two binomial distributions in Equations 3 and 4, which will be discussed later [15]. The lot is accepted if the 'apparent' number of defectives D' is less than or equal to the threshold D^*. In that case, the lot acceptance probability can be expressed as:

$$P_a = \mathrm{P}\,(D' \le D^* | D) = \sum_{D'=0}^{D^*} \mathrm{P}\,(D')\,, \tag{6}$$

Which is the probability of identifying $D' \le D^*$ (apparent) defectives; given that the lot contains D true defectives (see the notation '$|D$' in the second term of Equation 6). The term $\mathrm{P}(D')$, within the summation in the right-hand side of the equation, is the probability that the inspector identifies exactly D' defectives and can be calculated via the following convolution [15]:

$$\mathrm{P}\,(D') = \sum_j \left[\mathrm{P}\left(D_1' = j\right)\right] \cdot \left[\mathrm{P}\left(D_2' = D' - j\right)\right], \tag{7}$$

where:

- j is a dummy discrete variable belonging to the set N, which describes the number of true defectives (among the D truly present in the lot) that are correctly classified as such;
- $(D' - j)$ is the number of units, among the $(N - D)$ truly conforming units in the lot that are incorrectly classified as defective.

In a nutshell, Equation 7 states that, in order to compute the probability of observing D' 'apparent' defectives, all possible combinations of correctly identified (true) defectives (i.e., D_1') and incorrectly identified defectives (i.e., false defectives, D_2') must be considered. The terms D_1' and D_2' are in fact additive components whose sum equals D'. Referring to the above binomial distributions of D_1' and D_2' (cf. Equations 3 and 4), Equation 7 can be expanded as:

$$\mathrm{P}\,(D') = \sum_{j=\max\{0,\ D'-(N-D)\}}^{\min\{D',\ D\}} \left\{ \left[\binom{D}{j} \cdot (1-b)^j \cdot b^{D-j} \right] \cdot \left[\binom{N-D}{D'-j} \cdot a^{(D'-j)} \cdot (1-a)^{(N-D)-(D'-j)} \right] \right\}. \tag{8}$$

Let us focus on the extremes of the summation in Equation 8. The dummy variable j indicates how many of the D truly defectives are correctly classified as such. Since, by definition, j cannot exceed D, it follows that $j \le D$. Furthermore, the quantity $(D' - j)$ counts how many of the $(N - D)$ truly conforming units are misclassified as defective; thus, the additional constraint $(D' - j) \le (N - D)$

holds. Furthermore, neither j nor $(D' - j)$ can be negative. All the above constraints result in: $0 \le j \le D$ and $0 \le (D' - j) \le (N - D)$; in particular, the second condition implies: $j \le D'$ and $j \ge D' - (N - D)$. Combining these constraints, we obtain the expressions $j \ge \max\{0, D' - (N - D)\}$, which corresponds to the lower index, and $j \le \min(D', D)$, which corresponds to the upper index of the summation in Equation 8.

Inserting Equation 8 into Equation 6 leads to the new expression for P_a:

$$P_a = \sum_{D'=0}^{D^*} \sum_{j=\max\{0,\ D'-(N-D)\}}^{\min\{D',\ D\}} \left\{ \left[\binom{D}{j} \cdot (1-b)^j \cdot b^{D-j} \right] \cdot \left[\binom{N-D}{D'-j} \cdot a^{(D'-j)} \cdot (1-a)^{(N-D)-(D'-j)} \right] \right\}. \quad (9)$$

The model in Equation 9 is consistent with the formulation of the step-shaped OC curve (cf. Equation 1 and Figure 1). Specifically, Equation 9 reduces to Equation 1 when $a = b = 0$, which implies $D'_1 = D$ and $D'_2 = 0$, and hence $D' = D'_1 + D'_2 = D$; consequently, $P_a(0 \le D \le D^*) = 1$ and $P_a(D^* < D \le N) = 0$. Conversely, if a and b approach 1, the OC curve becomes the mirror image of the ideal step-shaped function, because conforming units and defective units are systematically swapped. In the latter extreme case, the inspector perceives an 'apparent' defectiveness that is complementary to the true one: That is, $p' = 1 - p$.

Let us now try to generalise the concept of apparent defectiveness. Consistent with Equation 5, the expected value of D' can be defined as the sum of the expected number of true defectives $\mathrm{E}[D'_1] = D \cdot (1 - b)$ and the expected number of false defectives $\mathrm{E}[D'_2] = (N - D) \cdot a$ identified by the inspector. Therefore:

$$\mathrm{E}[D'] = \mathrm{E}\left[D'_1\right] + \mathbb{E}\left[D'_2\right] = D \cdot (1-b) + (N-D) \cdot a. \quad (10)$$

Correspondingly, a general expression of the apparent defectiveness can be obtained as:

$$p' = E[p'] = \frac{\mathbb{E}[D']}{N} = \frac{D}{N} \cdot (1-b) + \left(\frac{N}{N} - \frac{D}{N}\right) \cdot a = p \cdot (1-b) + (1-p) \cdot a. \quad (11)$$

Equation 10 shows that the true defectiveness (p) can be distorted through the 'subjective' lens' of inspection errors (a and b).

Returning to P_a in Equation 9, let us analyse its behaviour by evaluating the effect of error rates a and b. Figure 2 exemplifies six cases, plotting the OC curve (P_a vs. $p = D/N$) for six corresponding (a, b) configurations, and overlaying the ideal step-shaped OC curve (with no errors) as a plotted-line reference (from Figure 1). The cases shown are discussed as follows: (1) with modest error rates ($a = 5\%$, $b = 2\%$) the OC curve already deviates visibly from the ideal one; (2) the predominant effect of $b = 15\%$ raises the right-hand side of the curve, implying more acceptance of truly bad lots; (3) the relatively high value of $a = 17\%$ lowers the left-hand side, implying more rejection of truly good lots; (4) $a = 20\%$ (further increase) continues to lower the left side; (5) $b = 90\%$ (extremely high) greatly raises the right side (many bad lots accepted); and (6) when both error rates are extreme (~100%), the OC curve appears as the mirror image of the ideal step-shaped curve (dotted line).

The above examples highlight the typical effects of a and b. While a depresses the OC curve at relatively low p values (lots are more likely to be rejected because several conforming units are mistakenly classified as defectives), b elevates the OC curve at relatively high p values (lots are more likely to be accepted because some truly defective units are missed). The combination of a and b therefore encodes a tendency toward *severity* at low p values versus *leniency* at high p values. These examples also show that occasional drifts in a and b (due to possible inconsistencies in inspector skill, training, attention, fatigue, familiarity with the product, etc.) can make the OC curve unstable, producing uncontrolled oscillations in the actual severity/protection level relative to the ideal case (with no errors). This aspect can be very detrimental in practice as it can distort LDDs.

3 | Modelling the Multiple-Inspector SSP

The discussion now moves from 100% inspection to sampling inspection via a SSP. In the ideal scenario of no inspection errors, the OC curve of a SSP can be derived from classical acceptance sampling theory [9]. The classical SSP's LDD is that, after drawing a sample of n units from a lot of N units, the number (d) of defective (or nonconforming) units is counted. If $d \le c$, the entire lot is accepted (*pass*); otherwise, it is rejected (*fail*) [13]. SSPs are also codified in international standards, such as ISO 2859-1:2026 [19].

For simplicity, the present work is restricted to the so-called Type-B OC curves, whose construction assumes that the sample is drawn from a large lot (i.e., $n \ll N$ or $n/N \le 0.1$) or, equivalently, from a train of lots randomly selected from a relatively stable process. Under these conditions, the *hypergeometric* distribution of defectives can be approximated by a *binomial* distribution [9]. Thus, for a SSP with parameters n and c, the lot acceptance probability is:

$$P_a = \sum_{d=0}^{c} \binom{n}{d} \cdot p^d \cdot (1-p)^{n-d}. \quad (12)$$

Figure 3 exemplifies the OC curve for a SSP with parameters $N = 140$, $n = 14$ and $c = 1$, constructed using Equation 12. These parameters were chosen such that the threshold defectiveness $p^* = 20/140 \approx 14.3\%$, which corresponds to the 'knee' of the ideal step-shaped OC curve in Figure 1, lies roughly midway between an *acceptable quality level* AQL = 4% and a *lot tolerance percent defective* LTPD = 24.5%[2]. Under this design, the corresponding producer's risk (α) and consumer's risk (β) are approximately 10.6% and 10.8%, respectively. The solid line in Figure 3 represents the OC curve of this single-inspector SSP with no errors and the dashed line shows the ideal 100% inspection OC curve from Figure 1. Both curves correspond to a single-inspector, error-free scenario.

Let us now consider a *multiple-inspector* scenario where m inspectors independently inspect the same n sample units in parallel. The conformity assessments on each unit can be organised in a *conformity-assessment* matrix [8]: each row corresponds to an

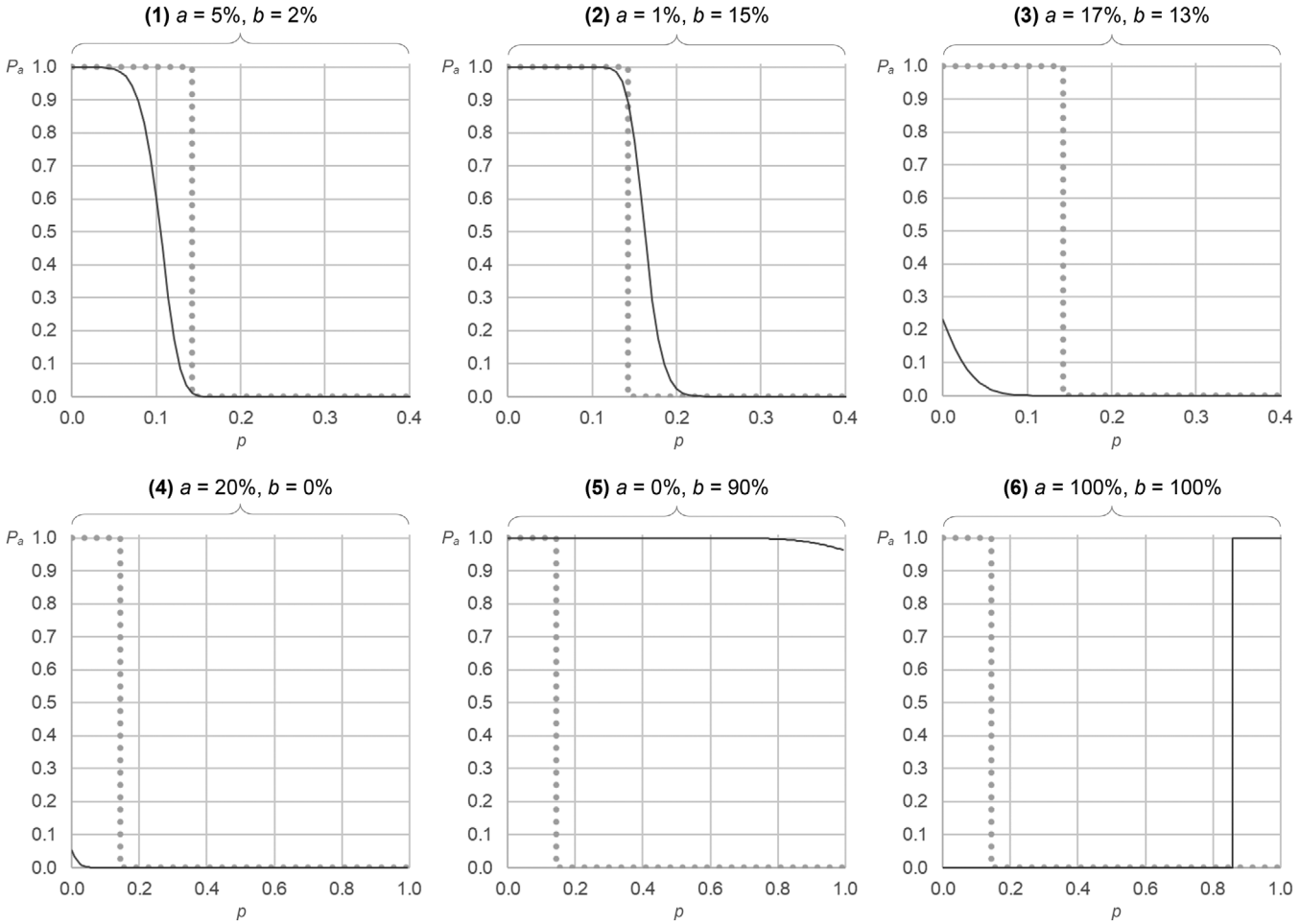


FIGURE 2 | Effect of inspection-error rates on OC curves for a single-inspector 100% inspection, with $N = 140$ and $D^* = 20$. For six combinations of a and b, the OC curves are obtained using the model in Equation 9. The parameters a and b were varied within very wide ranges to make their effects more visible. For greater generality, the horizontal axes of the curves show $p = D/N$; furthermore, the horizontal axis of the top three plots is truncated at $p = 0.4$, for more clarity. In each of the six plots, the ideal step-shaped OC curve (with no errors) is shown as a dotted line for reference (cf. Equation 1 and Figure 1).

inspector (i_1, i_2, …, i_m) and each column to a sample unit (u_1, u_2, …, u_n). An entry in this matrix is marked as '✓', if inspector i classifies unit j as conforming, or '✗', if deemed defective (see the example in Figure 4).

To reach a final LDD from this matrix, a two-stage aggregation is performed [2]. First, the m assessments for each unit (in a specific column of the matrix) are consolidated into a single global classification for that unit (column-wise aggregation). Then, the resulting n global classifications are evaluated against the usual (n, c) SSP's LDD, to decide whether the lot is accepted or rejected (row-wise aggregation). The recent contribution by Maisano et al. [2] also considers an alternative two-stage aggregation, first by rows and then by column, which will not be recalled here, in order to keep the discussion concise.

Returning to the two-stage aggregation initially described (i.e., first by columns and then by row), in this study, a *majority* criterion is assumed for the column-wise aggregation, since it provides a relatively robust OC curve even in the presence of inspector errors [2]. Under the *majority* criterion, a unit is considered to be conforming at a global level (i.e., '✓'), if at least

$$q = \lfloor \frac{m}{2} \rfloor + 1 \tag{13}$$

Of the m inspectors classify it as conforming (*pass*). For example, with $m = 5$ inspectors, a *majority* requires $q = 3$, so a unit yielding the column $[✓, ✓, ✓, ✗, ✗]^T$ from the five inspectors would result in a global conformity assessment of '✓'. By contrast, a *unanimity* rule, requiring all inspectors to agree, is much more stringent [8]. Figure 5 exemplifies an aggregation by columns of the conformity assessments from five inspectors (i_1 to i_5) for three product units (u_1–u_3), using the *majority* criterion.

A recent study by Maisano et al. [2] modelled the OC curves for the multiple-inspector SSP case, considering different combinations of aggregation sequence (*columns-first* or *rows-first*), column-wise criterion (*majority* or *unanimity*) and statistical model type (Type-A or Type-B). It should be reiterated that, for simplicity, the present study focuses exclusively on the case of *columns-first* aggregation with *majority* criterion and a Type-B model, as this case reflects a relatively common practical scenario

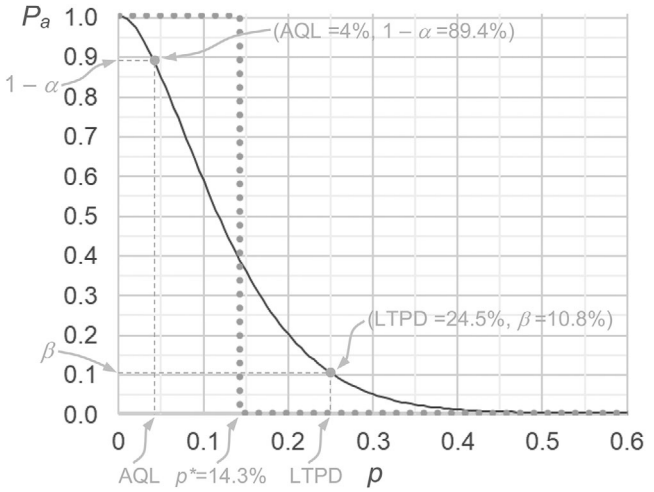


FIGURE 3 | The solid line represents the Type-B OC curve for a SSP with parameters $n = 14$ and $c = 1$. This SSP was designed such that $p^* \approx 14.3\%$ (the 'knee' of the ideal step-shaped OC curve for 100% inspection; dashed line from Figure 1) lies roughly midway between an AQL of 4% and an LTPD of 24.5%, yielding corresponding producer's and consumer's risks of about 10.6% and 10.8%, respectively. In other words, the chosen SSP is characterised by an OC curve that ensures a protection level and/or severity in line with the ideal step-shape OC curve of 100% inspection (cf. Figure 1).

for multiple-inspector SSPs [8, 18, 20]. The OC curve can be modelled by the expression:

$$P_a = \sum_{d=0}^{c} \binom{n}{d} \cdot [p']^d \cdot [1-p']^{n-d}, \tag{14}$$

Which has the same form as that of an ideal SSP (cf. Equation 12), but uses a multiple-inspector apparent defectiveness p' in place of the true defectiveness p. In analogy with the description in Sect. 2 which led to the formulation of Equation 11, the term p' can be expressed as a function of two *aggregated error rates* (a_{agg} and b_{agg}), for a single (fictitious) 'aggregate' inspector:

$$p' = (1 - b_{\text{agg}}) \cdot p + a_{\text{agg}} \cdot (1-p). \tag{15}$$

The parameters a_{agg} and b_{agg} can be interpreted as the characteristic error rates of a hypothetical 'aggregate' inspector, whose conformity assessments result from the aggregation of the individual inspectors' assessments by *majority*. The term p' is thus a distorted version of the true p, as perceived by the hypothetical 'aggregate' inspector; this distortion is driven by the combination of the individual inspectors' error rates into the aggregated parameters a_{agg} and b_{agg} (as described in detail below). Figure 6 highlights this distortion, showing the linear relationship of p' as a function of p, and illustrating the effect of a_{agg} and b_{agg}. Two extreme cases are worth noting:

- If $a_{\text{agg}} = b_{\text{agg}} = 0$ (resulting from zero inspection errors by any inspector, as shown later), then Equation 15 determines a complete correspondence between p' and p, that is, the bisector of the p-p' plane, $p' = p$;
- If $a_{\text{agg}} = b_{\text{agg}} = 1$ (all inspectors make the maximum possible errors, as shown later), then Equation 15 determines $p' = 1 - p$, which corresponds to a line perpendicular to the bisector of the p-p' plane (see the dotted line in Figure 6).

Note that the model of Equation 15 is valid only for the case of aggregation sequence *columns-first* with the column-wise aggregation criterion based on *majority* and Type-B OC curves. For a more general description including other possible cases, see Maisano et al. [2].

Let us now turn attention to the formulation of a_{agg} and b_{agg} as functions of the individual inspectors' error rates:

- a_i denotes the probability that the i-th inspector erroneously classifies a truly conforming unit as *defective* (i.e., false-defective error rate). Equivalently, $(1 - a_i)$ is the probability of correctly classifying a conforming unit.
- b_i is the probability that the i-th inspector misclassifies a truly defective unit as conforming (i.e., false-conforming error rate). Equivalently, $(1 - b_i)$ is the probability of correctly classifying a defective unit.

In practice, the individual error rates a_i and b_i can vary from inspector to inspector, are often hard to estimate and may vary over time, due to factors such as differences in training and experience, fatigue, inherent subjectivity of certain inspections, etc. [8, 2].

Let us return to the aggregated error parameters a_{agg} and b_{agg} introduced in Equation 15. The parameter a_{agg} is formally defined as the probability that a conforming unit is misclassified as defective by the *majority* of the group of m inspectors. Similarly, b_{agg} is the probability that a defective unit is misclassified as conforming by the *majority* of the m inspectors. Mathematically, these probabilities are obtained by summing over all combinations of inspectors that meet the *majority* condition. In particular, the probability that a truly conforming unit will be correctly classified as such by the majority of inspectors is:

$$1 - a_{\text{agg}} = \sum_{s=q}^{m} \left\{ \sum_{\substack{A \subseteq \{1,\ldots,m\} \\ |A| = s}} \left[\prod_{i \in A} (1 - a_i) \right] \cdot \left[\prod_{i \notin A} a_i \right] \right\},$$

from which

$$a_{\text{agg}} = 1 - \sum_{s=q}^{m} \left\{ \sum_{\substack{A \subseteq \{1,\ldots,m\} \\ |A| = s}} \left[\prod_{i \in A} (1 - a_i) \right] \cdot \left[\prod_{i \notin A} a_i \right] \right\}, \tag{16}$$

Where s is the size of a subset A of inspectors who correctly classify the unit ('✓'), while the remaining $(m - s)$ inspectors not in A misclassify it ('✗'). To enumerate all combinations where $q \leq s \leq m$, a contingency table can be constructed. For example, if $m = 5$ and $q = 3$, Table 1 shows all possible combinations in which at least 3 of 5 inspectors (i.e., 3, 4 or 5) fulfil a given condition (i.e., classifying a unit as conforming, in this case).

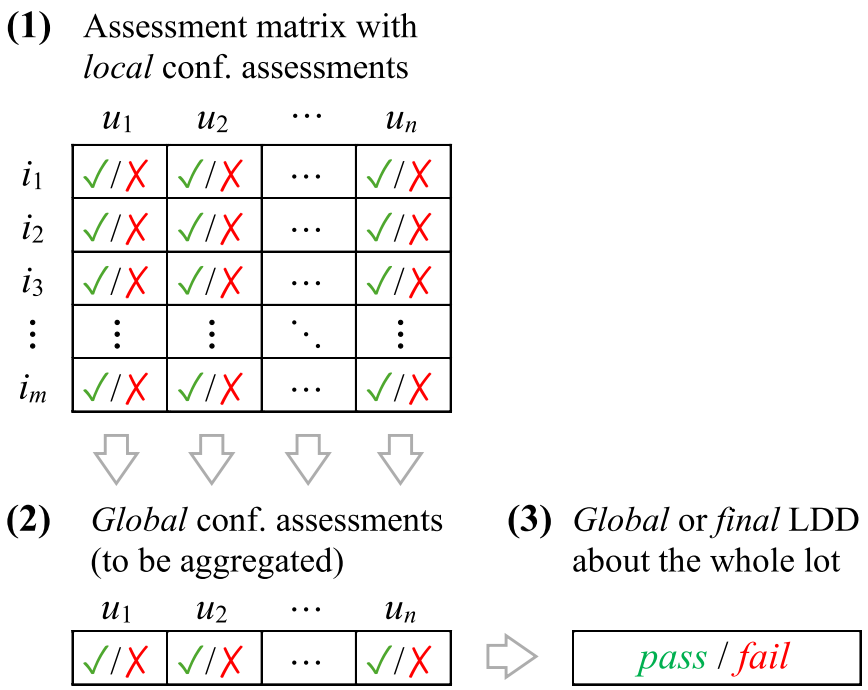


FIGURE 4 | **(1)** Conceptual scheme of a *conformity-assessment matrix* with aggregation sequence *columns-first*, to produce a final LDD for the lot [8]. Specifically, the *local* conformity assessments for each product unit are aggregated by column into a *global* conformity assessment (**2**) Using the *majority* criterion. Then, the global assessments by the inspectors are aggregated by row (**3**) Using the classic SSP's LDD logic, yielding a final LDD.

TABLE 1 | Example of a contingency table for $m = 5$ inspectors, covering all possible combinations in which at least $q = 3$ inspectors fulfill a given condition (e.g., classifying a unit as conforming or defective). 'Y' and 'N' respectively indicate that the condition is fulfilled or not fulfilled for a given inspector.

Inspector 1	Inspector 2	Inspector 3	Inspector 4	Inspector 5	$S = \|A\|$
Y	Y	Y	N	N	3
Y	Y	N	Y	N	3
Y	Y	N	N	Y	3
Y	N	Y	Y	N	3
Y	N	Y	N	Y	3
Y	N	N	Y	Y	3
N	Y	Y	Y	N	3
N	Y	Y	N	Y	3
N	Y	N	Y	Y	3
N	N	Y	Y	Y	3
Y	Y	Y	Y	N	4
Y	Y	Y	N	Y	4
Y	Y	N	Y	Y	4
Y	N	Y	Y	Y	4
N	Y	Y	Y	Y	4
Y	Y	Y	Y	Y	5

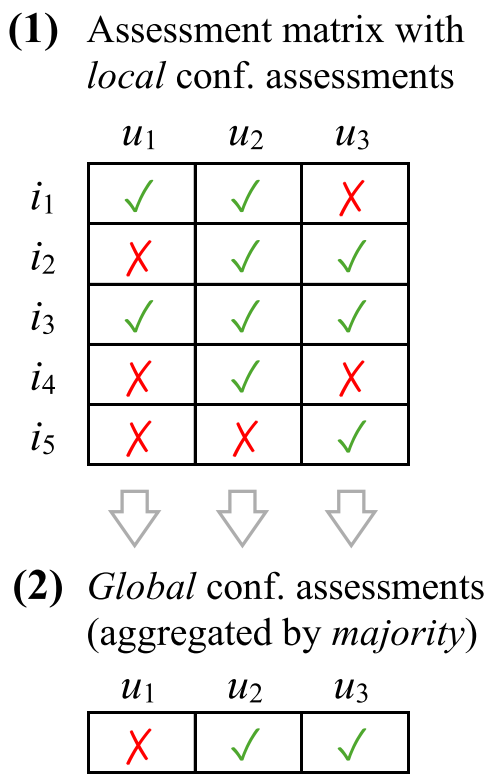


FIGURE 5 | Example of aggregation by column of the conformity assessments from five inspectors (i_1–i_5) for three product units (u_1–u_3), using the *majority* criterion.

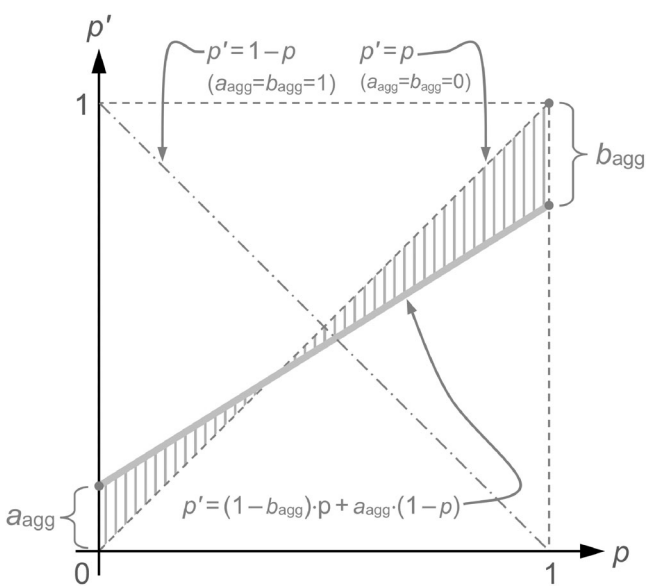


FIGURE 6 | Graphical representation of the relationship in Equation 15, assuming inspections are carried out by m inspectors with error rates a_i and b_i (with $i = 1, 2, \ldots, m$), which are aggregated into corresponding a_{agg} and b_{agg} parameters (cf. Equations 16 and 17). While a_{agg} tends to inflate p' with respect to p (graphically, it sets the starting point of the linear function at $p = 0$), b_{agg} tends to reduce p' with respect to p (graphically, it sets the ending point of the linear function at $p = 1$). The grey shaded area highlights the deviation of p' from the ideal no-error case (where $a_i = b_i = 0$ for all inspectors, yielding $a_{agg} = b_{agg} = 0$ and thus $p' = p$, which corresponds to the bisector of the p-p' plane).

Similarly, the probability that a truly defective unit will be misclassified as conforming by the majority of inspectors is:

$$b_{agg} = \sum_{s=q}^{m} \left\{ \sum_{\substack{A \subseteq \{1,\ldots,m\} \\ |A| = s}} \left[\prod_{i \in A} b_i \right] \cdot \left[\prod_{i \notin A} (1 - b_i) \right] \right\}, \quad (17)$$

s being, in this other case, the numerosity of the subset A of inspectors who misclassify the defective unit as conforming ('✓') and q being an imposed threshold for the previous parameter (cf. Equation 13). Contingency Table 1 can be used again to define all possible combinations in which $q \leq s \leq m$. It is easily verified that if $a_i = 0$ and $b_i = 0\ \forall i$, then from Equations 16 and 17 it is obtained that $a_{agg} = b_{agg} = 0$ and from Equation 15 that $p' = p$. In this ideal case of zero inspection errors, the multiple-inspector model in Equation 14 collapses to the ideal SSP case of Equation 12.

4 | Empirical Comparison

To investigate the effect of the number of inspectors and their error rates (a_i and b_i, for $i = 1, 2, \ldots, m$) on the OC curve of a SSP, a realistic example featuring seven hypothetical inspectors (i_1–i_7) in two different error-rate scenarios is considered (see Table 2). The low-error-rate (L) scenario assumes highly experienced inspectors with a certain familiarity with the inspected units, while the high-error-rate (H) scenario assumes partially experienced inspectors with low familiarity. The a_i and b_i values for L scenario were randomly generated from a normal distribution with $\mu = 4\%$ and $\sigma = 2\%$. This choice was guided by empirical data from scientific literature; for example, the scientific literature reports typical inspector error probabilities on the order of a few percent for a rates and slightly higher percentages, for b rates [17, 21]. On the other hand, the error rates in the H scenario were randomly generated from a normal distribution with $\mu = 18\%$ and $\sigma = 4\%$. These other values are deliberately very large, probably exaggerated, in order to perform a sort of 'stress test' of the multiple-inspector SSP model under extreme conditions. The seven inspectors are further grouped into subsets of sizes $m = 3$, 5 and 7, which will be used to study the effect of parameter m[3].

4.1 | Aggregated Error Rates (a_{agg} and b_{agg})

Using the data in Table 2, the aggregated error rates a_{agg} and b_{agg} were computed for the multiple-inspector SSP as a function of the number of inspectors m. Applying Equations 16 and 17 to the inspectors' a_i and b_i values yields a_{agg} and b_{agg} for each considered value of m. Figure 7 illustrates how these aggregated error rates decrease as m increases, for both the L and H scenarios (the numerical values are also provided in the tables below the respective plots).

The decrease in a_{agg} and b_{agg} with increasing m appears as a convex downward, hinting at an asymptotic trend: as m grows, the aggregated error rates approach zero, converging toward the ideal case of error-free inspections, with p' approaching p (cf.

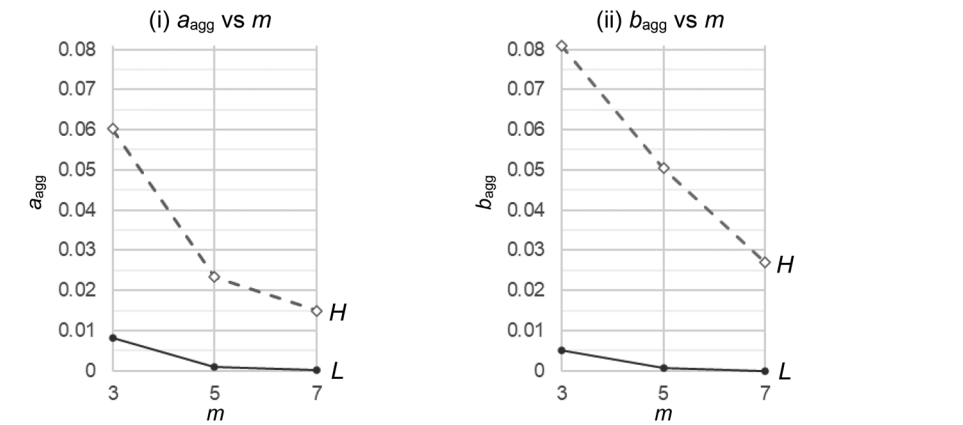
TABLE 2 | Error rates (a_i and b_i) of seven hypothetical inspectors (i_1–i_7) across two error-rate scenarios: *L* scenario, assuming that inspectors have great familiarity with the inspected units, and *H* scenario, assuming that inspectors have only partial familiarity. Inspectors are grouped into three sets of size $m = 3$, 5 and 7.

Inspector	*L* scenario		*H* scenario	
	a_i	b_i	a_i	b_i
i_1	5.3%	2.6%	17.4%	13.2%
i_2	4.3%	4.2%	13.4%	23.0%
i_3	6.5%	6.0%	14.1%	16.8%
i_4	5.5%	3.6%	13.6%	22.9%
i_5	2.9%	4.7%	13.2%	19.7%
i_6	4.1%	2.4%	15.5%	21.3%
i_7	3.1%	3.0%	25.5%	15.6%

group of m=3 inspectors

group of m=5 inspectors

group of m=7 inspectors



m:	3	5	7
H scenario:	6.02%	2.34%	1.49%
L scenario:	0.82%	0.10%	0.01%

m:	3	5	7
H scenario:	8.10%	5.04%	2.69%
L scenario:	0.50%	0.07%	0.01%

FIGURE 7 | Effect of the number of inspectors (m) in terms of reduction in a_{agg} and b_{agg}, considering the two error-rate scenarios (*H* and *L*). The tables below show the numerical values plotted.

Equation 15). In the *L* scenario, the benefit of even a relatively small group of inspectors is evident: with $m = 3$ inspectors, the aggregated error rates are already reduced to under 1%, despite individual inspectors in this scenario having error rates on the order of 4%–6%. Increasing to $m = 5$ and 7 drives a_{agg} and b_{agg} down to essentially negligible levels. In other words, for inspectors who are fairly accurate, having a group of $m = 3$ or more yields an almost error-free overall inspection outcome.

Even in *H* scenario, a substantial reduction in aggregated error rates can be observed as m increases. With $m = 3$, starting from individual error rates in the tens of percent (see Table 2), the aggregated error probabilities drop by about one order of magnitude, reaching only a few percent. With $m = 5$ or $m = 7$, the aggregated errors become even smaller and negligible. This reinforces the hypothesis that the redundancy of inspections has a strong positive effect in reducing the distortions caused by inspection errors.

4.2 | OC Curve Performance

Figure 8 shows the OC curves for the multiple-inspector SSP under varying conditions: two error scenarios (*L* in the top row, *H* in the bottom row) and different numbers of inspectors ($m = 3, 5, 7$ in the columns). In each plot, the solid line is the actual OC curve with inspector errors and the dashed line is the ideal OC curve (with no errors), shown for comparison. These OC curves were constructed using the analytical model derived in Equations 14–17 for a multiple-inspector SSP with inspection errors. The complete equation of P_a as a function of p is given in Appendix A.1.

As expected, the multiple-inspector OC curves converge toward the ideal OC curve as m increases or as error rates reduce. In fact, under the *L* scenario, the OC curve with $m = 3$ is practically indistinguishable from the ideal curve; for $m = 5$ and $m = 7$ (*L* scenario), the OC curves are essentially identical to the ideal case. This means that even with some small inspection errors

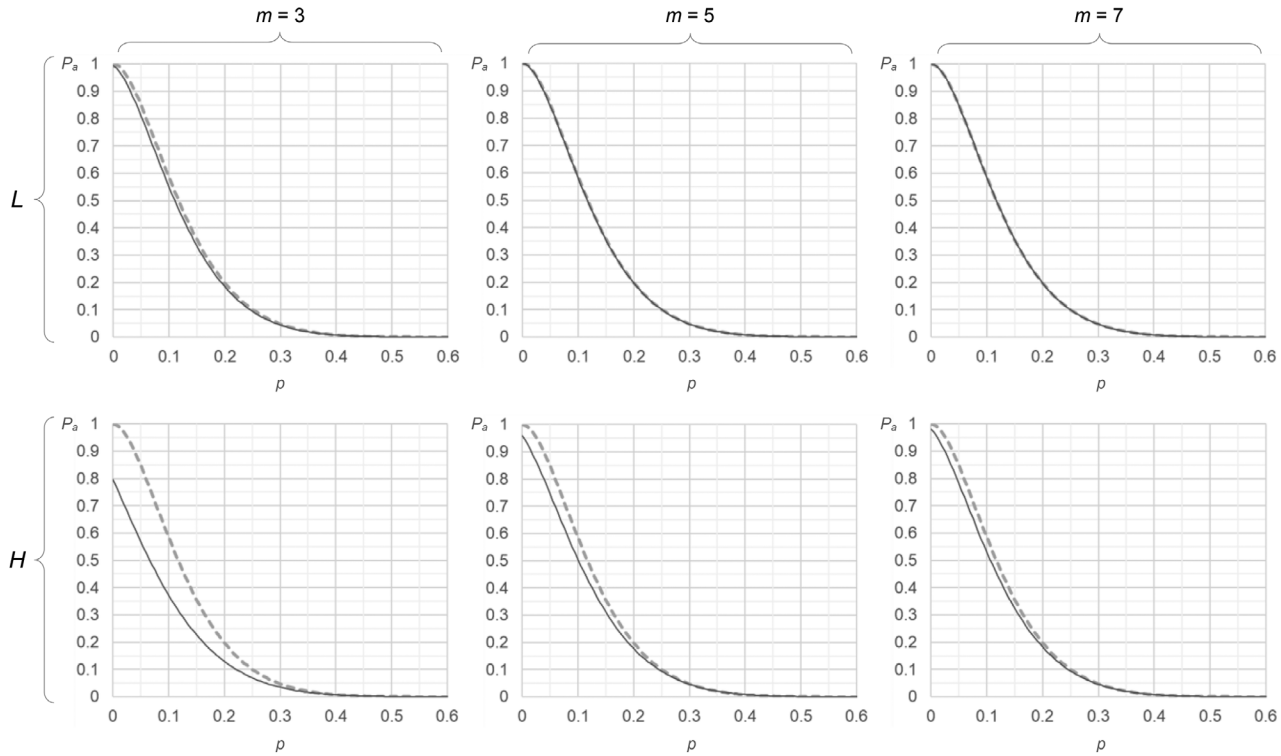


FIGURE 8 | OC curves for the multiple-inspector SSP under varying conditions: two error scenarios (*L* in top row, *H* in bottom row) and different number of inspectors ($m = 3$, 5 and 7, in the columns). In each plot, the solid line represents the actual OC curve with inspectors' errors and the dashed line represents the ideal OC curve (with no inspection errors). In the *L* scenario, even a group of $m = 3$ inspectors produces an OC curve virtually identical to the ideal one. In the *H* scenario, larger groups ($m = 5$ or 7) are needed to closely approximate the ideal OC curve, although $m = 3$ still shows some deviation. Overall, increasing the number of inspectors drives the OC curve of the sampling plan to converge toward the error-free ideal one, illustrating the compensatory effect of multiple inspectors.

by individual inspectors, the group's *majority* decision yields lot acceptance behaviour nearly the same as if there were no inspection errors at all.

Under the *H* scenario, the trend is similar, but a larger m is needed to approach the ideal OC curve. With $m = 3$, there remains a noticeable deviation from the ideal OC curve: in the example shown, the multiple-inspector SSP appears slightly more severe, meaning that it tends to reject lots more often than the ideal case. This behaviour is probably attributable to the effect of a_{agg}, which arises from the relatively high a_i error rates: even a few percent of false rejections (false defectives) can shift the OC curve downward (making acceptance harder). However, by increasing to $m = 5$, the OC curve moves much closer to the ideal one, and with $m = 7$ it aligns even more closely. Thanks to this error-compensation effect, the multiple-inspector SSP can achieve OC curves that are nearly as discriminating as the ideal one, even when individual inspectors have non-negligible error rates.

From a practical standpoint, this result is highly encouraging: implementing a multiple-inspector strategy provides a form of protection against uncontrolled variations in the OC curve. To complete the analysis carried out so far, Sect. A.2 (in the Appendix) shows how typical error rates, which do not significantly 'disturb' the OC curves in the multiple-inspector case, would have a significant effect if a SSP with just a single inspector was considered.

4.3 | Inspection Effort

Another crucial aspect is the inspection effort associated with multiple-inspector sampling versus the 100% lot-inspection approach. Intuitively, using multiple inspectors increases the number of inspections per sampled unit (and the related cost), but this must be counterbalanced by the fact that a SSP inspects far fewer units than a 100% inspection of the entire lot. In other words, a key question is whether a multiple-inspector SSP can still offer efficiency benefits in terms of total inspections needed compared to inspecting every unit by a single inspector (cf. **RQ3**).

Assuming a Type-B OC curve, the total number of inspections ($n{\cdot}m$) may remain significantly less than N. For example, returning to the earlier case with $N = 140$ and $n = 14$, if a single inspector performs 100% inspection of the lot, that requires 140 inspections. By contrast, a SSP with $m = 3$ inspectors examining 14 sampled units involves $n{\cdot}m = 14{\cdot}3 = 42$ total inspections; with $m = 5$ inspectors, $14{\cdot}5 = 70$ inspections; and even with $m = 7$ inspectors, $14{\cdot}7 = 98$ total inspections. All of these values are lower than 140, meaning that the multiple-inspector sampling strategy can still require fewer inspections than checking the entire lot unit-by-unit. Figure 9 illustrates this comparison, showing the total number of inspections per lot for each of the above approaches. Even in the 'worst case' here (i.e., $m = 7$ inspectors), the sampling plan uses about 30% fewer inspections than full inspection, while in the other cases ($m = 5$ and 3) the savings are on the order of 50%–70%.

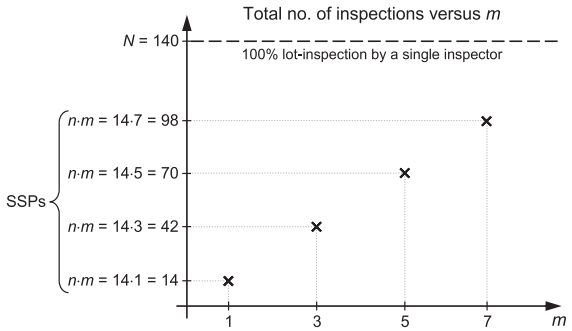


FIGURE 9 | Comparison of the total number of inspections required per lot, under different inspection strategies: full 100% inspection of an entire lot by a single inspector, versus SSPs (with sample size n) executed by either a single inspector ($m = 1$) or by a group of inspectors ($m = 3, 5, 7$). The example assumes $N = 140$ and $n = 14$. Even with multiple inspectors, the SSPs use substantially fewer inspections than 100% lot inspection (dashed line).

5 | Conclusions

The analysis shows that multiple-inspector SSPs can achieve OC curves very close to the ideal error-free scenario. In practice, even if individual inspectors occasionally make inspection errors, combining their independent inspections yields an overall LDD that is relatively robust. In fact, multiple-inspector SSPs perform nearly as if each unit were inspected perfectly, despite the errors at the level of individual inspector (cf. **RQ1**).

Another significant finding is that the number of inspectors and their individual error rates influence the shape and stability of the OC curve: increasing the number of inspectors and/or decreasing individual error rates contributes to improving performance (cf. **RQ2**). Aggregating independent conformity assessments—typically using the *majority* criterion—can be beneficial, even when inspector error rates are uncertain or difficult to estimate. In fact, the error-compensation effect allows reliable outcomes without the need for precise and often impractical estimation of each inspector's error rates. For instance, three inspectors with error rates of around a dozen percentage points can jointly yield the performance of a single 'aggregate' inspector with error rates of just a few percentage points. Therefore, redundancy in inspection introduces a natural form of protection, even in the presence of uncertain or heterogeneous inspector performance.

Another key finding is that a multiple-inspector SSP tends to require fewer total inspections than a single-inspector 100% lot inspection. While each sampled unit is inspected multiple times, the total number of inspections—measured as $n{\cdot}m$—is typically much lower than inspecting the entire lot unit-by-unit (cf. **RQ3**). This is especially evident when the number of inspectors is moderate, striking a favourable balance between redundancy and workload. From a practical standpoint, this means that a multiple-inspector sampling strategy can save time and cost while ensuring a good level of protection against distorted LDDs.

The above findings align with industry practices since in several high-stakes domains, the use of multiple inspectors— whether human experts or automated systems—is standard practice to mitigate the risk of unreliable LDDs, due to inspection errors. For example, critical inspections in sectors such as manufacturing and medical diagnostics are often independently double-checked [18]. Similarly, hybrid inspection systems exploit the error compensation effect by combining automated vision technologies with human oversight, in ambiguous or complex cases [22]. Any group of independent inspectors (whether human or machine) tends to counterbalance individual misclassifications, producing a final LDD closer to the ideal of perfect inspection.

Despite these advantages, the proposed study has some limitations. First, it assumes statistical independence between inspectors' errors—a condition that may not always hold in real settings, as correlated errors can arise when inspectors share the same 'blind spots' or are influenced by similar contextual factors. This assumption is crucial for enabling the error-compensation effect highlighted in this study: when inspectors operate independently, their individual misclassifications tend to offset each other through the aggregation process, thereby enhancing the robustness of the final LDD. Conversely, the presence of dependence among inspectors reduces their ability to effectively 'cross-verify' each other. In the extreme case of complete dependence—where all inspectors systematically produce identical conformity assessments (i.e., perfect correlation)—the aggregation process becomes ineffective, as the aggregated error rates would coincide with those of the individual inspectors, thus nullifying the benefits of redundancy [2]. Although this study does not explicitly model statistical dependence among inspectors, this limiting case provides an intuitive understanding of its potential impact. One way to mitigate this limitation is to assemble groups of inspectors with complementary skills, where each inspector focuses on different aspects of the product, thereby reducing the risk of overlapping errors. A second limitation lies in the specific sampling model adopted: this study focused on multiple-inspector SSPs with a *columns-first* aggregation sequence and column-wise aggregation based on *majority*. As such, the results may not directly generalize to more complex schemes (e.g., double or sequential sampling) or to different aggregation logics.

To conclude, this study demonstrated that a multiple-inspector sampling strategy can significantly enhance the robustness and efficiency of conformity assessments in the presence of inspection errors. Addressing the research questions, it was shown that involving multiple independent inspectors stabilizes the OC curve, mitigates the impact of misclassifications and may reduce the total inspection effort compared to 100% inspection by a single inspector. Extending the results obtained, greater reliability of the final LDD would certainly be guaranteed by a 100% inspection by multiple inspectors, but this would be excessively costly in terms of inspections ($N{\cdot}m$ total inspections required). Employing additional inspectors entails proportionally higher costs related to training, salaries and inspection time, among other factors. Hence, the logic of the multiple-inspector SSPs seems a good compromise.

Future research will aim to overcome the present study's limitations by exploring alternative aggregation rules, relaxing the independence assumption and extending the analysis to more complex sampling frameworks. In particular, further efforts will be devoted to developing practical methods to assess and monitor the degree of independence among inspectors, thereby

supporting the design of suitably diversified inspection teams. A deeper understanding of how redundancy and sampling strategies interact—especially in the presence of correlated inspection errors—will enable the design of inspection systems that are both highly reliable and practically viable.

List of acronyms and variables

Acronyms

AQL	Acceptance quality level;
H	High error-rate scenario;
ISO	International Organization for Standardization.
L	Low error-rate scenario;
LDD	Lot-disposition decision;
LTPD	Lot tolerance percent defective;
OC curve	Operating-characteristic curve;
RQ	Research question;
SSP	Single sampling plan.

Variables

a_i	Probability that the i-th inspector misclassifies a conforming unit as defective (false-defective error rate);
a_{agg}	Aggregated false-defective error rate resulting from the *majority* aggregation of individual inspectors' assessments (as perceived by an equivalent "aggregate" inspector);
α	Producer's risk, defined as probability of rejecting a lot with relatively low defectiveness, at the AQL;
A	Subset of inspectors who classify (correctly or incorrectly) a given unit as conforming;
b_i	Probability that the i-th inspector misclassifies a defective unit as conforming (false-conforming error rate);
b_{agg}	Aggregated false-conforming error rate, from the *majority* aggregation of individual inspectors' assessments (as perceived by an equivalent "aggregate" inspector);
β	Consumer's risk, defined as probability of accepting a lot with relatively large defectiveness, at the LTPD;
c	Acceptance number in a SSP (maximum defective units allowed in the sample for acceptance);
d	(True) number of defective units in the sample;
D	(True) number of defective units in the entire lot;
D^*	Acceptance number for a 100% inspection of the lot (i.e., agreed maximum count of defectives in the lot for acceptance);
$D' = D'_1 + D'_2$	Total number of defective units identified by inspection (i.e., the apparent number of defectives, to be compared with D^*);
D'_1	Number of true defectives correctly identified by inspection in the lot;
D'_2	Number of false defectives (i.e., conforming units misclassified as defective ones) identified by inspection in the lot;
$i_1, i_2, i_3, \ldots$	Generic inspectors involved in the (sample) inspection;
m	Number of inspectors performing the sample inspection independently;
n	Sample size in a SSP;
N	Lot size;
p	True *defectiveness* (or *fraction nonconforming* or *fraction defective*) of a lot (i.e., D/N);
p^*	Threshold defectiveness corresponding to D^* (i.e., D^*/N);
p'	"Apparent" defectiveness of a lot from the combined perspective of the m inspectors (as perceived by an equivalent "aggregate" inspector);
p'_3	"Apparent" defectiveness of a lot, from the perspective of the inspector i_3;
P_a	Probability of lot acceptance (which can be calculated both in the case of 100% inspection and SSPs);
$\mathbb{P}(D')$	Probability that the inspector will identify D' defective units during a 100% inspection of the lot;
q	*Majority* threshold of inspectors (e.g., at least q out of m inspectors);
s	Numerosity of the subset of inspectors (A);
$u_1, u_2, \ldots$	Generic product unit of the inspected sample;
✓	Conformity (*pass*) result relating to the assessment of a single product unit;
✗	Nonconformity (*fail*) result relating to the assessment of a single product unit.

Author Contributions

The authors have contributed equally to the drafting of the paper.

Acknowledgements

This research was carried out under the MICS (Made in Italy, Circular and Sustainable) Extended Partnership and partially funded by the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza–Missione 4, Componente 2, Investimento 1.3, D.D. 1551.11-10-2022, PE00000004). This manuscript reflects only the authors' views and opinions; neither the European Union nor the European Commission can be considered responsible for them.

Open access publishing facilitated by Politecnico di Torino, as part of the Wiley - CRUI-CARE agreement.

Ethics Statement

The authors respect the Ethical Guidelines of the Journal.

Consent to Participate

Not applicable.

Consent to Publish

Not applicable.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Endnotes

[1] Also known as *fraction nonconforming*, it represents one of the major parameters in inspections for *attributes* and *defectives* [9].

[2] The AQL is defined as a relatively low value of p, reflecting the typical defectiveness of the producer's process; this value is generally considered fully acceptable by the customer and should therefore be associated with a relatively high P_a and a low risk of rejection, also known as "producer's risk": $\alpha = 1 - P_a(p = \text{AQL})$. On the other hand, the LTPD is defined as a relatively high value of p, considered to be at the limit of acceptability for the customer; this other value generally corresponds to a relatively low P_a, also known as "customer risk": $\beta = P_a(p = \text{LTPD})$ [9]. Additionally, this Type-B OC curve with parameters $n = 14$ and $c = 1$ has values of α and β that are very close to each other (see Figure 3).

[3] Only odd values of m have been chosen as these are more practical for applying the column-wise aggregation criterion based on *majority*, given that they exclude the possibility of ties.

References

1. F. Franceschini, D. Maisano, and L. Mastrogiacomo, *Rankings and Decisions in Engineering: Conceptual and Practical Insights*, International Series in Operations Research & Management Science Series (Springer International Publishing, 2022).

2. D. A. Maisano, L. Ferrara, and F. Franceschini, "Characteristic Curves for Multiple-Inspector Sampling Plans With Inspection Errors," *Quality and Reliability Engineering International* 41 (2025): 3639–3656, https://doi.org/10.1002/qre.70058.

3. Y. H. Chun, "Improved Method of Estimating the Product Quality After Multiple Inspections," *International Journal of Production Research* 54, no. 19 (2016): 5686–5696, https://doi.org/10.1080/00207543.2015.1128128.

4. C. Mazza and J. L. Alvarez, *The Aesthetic Turn in Management* (Routledge, 2017), 157–178, https://doi.org/10.4324/9781351147965.

5. J. Bokrantz, A. Skoogh, C. Berlin, T. Wuest, and J. Stahre, "Smart Maintenance: A Research Agenda for Industrial Maintenance Management," *International Journal of Production Economics* 224 (2020): 107547, https://doi.org/10.1016/j.ijpe.2019.107547.

6. Y. H. Chun, "Improving Product Quality by Multiple Inspections: Prior and Posterior Planning of Serial Inspection Procedures," *IIE Transactions* 41, no. 9 (2009): 831–842, https://doi.org/10.1080/07408170802389324.

7. O. Battaïa, A. Dolgui, and N. Guschinsky, "An Exact Method for Machining Lines Design With Equipment Selection and Line Balancing," *International Journal of Production Research* 62, no. 1-2 (2024): 71–91, https://doi.org/10.1080/00207543.2023.2289182.

8. F. Franceschini, D. Maisano, and L. Mastrogiacomo, "Acceptance Sampling With Multiple Inspectors: Critical Aspects in Aggregating Individual Assessments," *Quality Engineering* 37, no. 2 (2025): 318–329, https://doi.org/10.1080/08982112.2025.2485175.

9. D. C. Montgomery, *Introduction to Statistical Quality Control*, 8th Edition (Wiley, 2019).

10. S. O. Duffuaa and A. El-Ga'aly, "Impact of Inspection Errors on the Formulation of a Multi-objective Optimization Process Targeting Model Under Inspection Sampling Plan," *Computers & Industrial Engineering* 80 (2015): 254–260, https://doi.org/10.1016/j.cie.2014.07.025.

11. F. Franceschini, D. Maisano, and L. Mastrogiacomo, "Empirical Analysis and Classification of Database Errors in Scopus and Web of Science," *Journal of Informetrics* 10, no. 4 (2016): 933–953, https://doi.org/10.1016/j.joi.2016.07.003.

12. A. Kolus and S. Duffuaa, "Determining Optimal Process Means in a Multi-stage Production System With Inspection Errors in 100% Inspection," *Quality Technology & Quantitative Management* 22, no. 1 (2025): 105–130, https://doi.org/10.1080/16843703.2024.2312016.

13. E. G. Schilling and D. V. Neubauer, *Acceptance Sampling in Quality Control*, 2nd Edition (CRC Press, 2009), https://doi.org/10.1201/9781584889533.

14. A. Tzimerman and Y. T. Herer, "Off-line Inspections Under Inspection Errors," *IIE Transactions* 41, no. 7 (2009): 626–641, https://doi.org/10.1080/07408170802331250.

15. S. M. Ross, *Introductory Statistics* (Academic Press Elsevier, 2017).

16. R. S. Kenett, S. Zacks, and P. Gedeck, "Sampling Plans for Batch and Sequential Inspection," in *Industrial Statistics: A Computer-Based Approach With Python* (Springer International Publishing, 2023), 397–442.

17. M. Colledani, T. Tolio, A. Fischer, et al., "Design and Management of Manufacturing Systems for Production Quality," *CIRP Annals* 63, no. 2 (2014): 773–796, https://doi.org/10.1016/j.cirp.2014.05.002.

18. D. H. Cymek, "Effects of Blinded and Nonblinded Sequential Human Redundancy on Inspection Effort and Inspection Outcome in Low Prevalence Visual Search," *Scientific Reports* 14 (2024): 23003, https://doi.org/10.1038/s41598-024-72210-8.

19. ISO 2859-1:2026 *Sampling Procedures for Inspection by Attributes – Part 1: Sampling Schemes Indexed by Acceptance Quality Limit (AQL) for Lot-by-lot Inspection* (ISO, 2026).

20. Y. H. Lin, J. M. Chen, and Y. C. Chen, "The Impact of Inspection Errors, Imperfect Maintenance and Minimal Repairs on an Imperfect Production System," *Mathematical and Computer Modelling* 53, no. 9-10 (2011): 1680–1691, https://doi.org/10.1016/j.mcm.2010.12.040.

21. C. G. Drury and P. G. Dempsey, *Handbook of human Factors and Ergonomics* (John Wiley & Sons, Inc, 2021), 853–879, https://doi.org/10.1002/9781119636113.

22. W. N. van Wieringen and J. de Mast, "Measurement System Analysis for Binary Data," *Technometrics* 50, no. 4 (2008): 468–478, https://doi.org/10.1198/004017008000000415.

Appendix A

A.1 | Complete formula of the multiple-inspector SSP's OC curve

By combining Equations 14–17 for a multiple-inspector plan with inspection errors, one can derive a complete closed-form expression for the lot acceptance probability (with aggregation sequence *columns-first* and column-wise aggregation based on *majority*):

$$P_a = \sum_{d=0}^{c} \binom{n}{d} \cdot \left[\left(1 - \sum_{s=q}^{m} \left\{ \sum_{\substack{A \subseteq \{1,\ldots,m\} \\ |A|=s}} \left[\prod_{i \in A} b_i \right] \cdot \left[\prod_{i \notin A} (1-b_i) \right] \right\} \right) \cdot p + 1 - \sum_{s=q}^{m} \left\{ \sum_{\substack{A \subseteq \{1,\ldots,m\} \\ |A|=s}} \left[\prod_{i \in A} (1-a_i) \right] \cdot \left[\prod_{i \notin A} a_i \right] \right\} \cdot (1-p) \right]^{d}$$
$$\cdot \left[1 - \left(1 - \sum_{s=q}^{m} \left\{ \sum_{\substack{A \subseteq \{1,\ldots,m\} \\ |A|=s}} \left[\prod_{i \in A} b_i \right] \cdot \left[\prod_{i \notin A} (1-b_i) \right] \right\} \right) \cdot p + \left(1 - \sum_{s=q}^{m} \left\{ \sum_{\substack{A \subseteq \{1,\ldots,m\} \\ |A|=s}} \left[\prod_{i \in A} (1-a_i) \right] \cdot \left[\prod_{i \notin A} a_i \right] \right\} \cdot (1-p) \right]^{n-d}. \tag{A1}$$

Despite its apparent complexity, Equation A1 results from substituting the expressions of a_{agg} and b_{agg} (Equations 16 and 17) into the relationship defining p' as a function of p, a_{agg} and b_{agg} in Equation 15, and subsequently substituting p' into Equation 14. Further details on the step-by-step derivation of Equation A1 are provided in Maisano et al. [2].

A.2 | Comparison with a single-inspector SSP

This section highlights the benefit of the multiple-inspector approach compared to the single-inspector one, for a generic SSP. Returning to the previous example and to the data in Table 2, let us consider inspector i_3, whose individual error rates are $a_3 = 6.5\%$, $b_3 = 6.0\%$ in the L scenario, and $a_3 = 14.1\%$, $b_3 = 16.8\%$ in the H scenario. The OC curves for a corresponding single-inspector SSP, with inspection errors can be obtained by adapting the binomial model in Equation 12, simply replacing the parameter p with the 'apparent' defectiveness perceived by i_3, due to his/her specific inspection errors. Specifically, i_3 will actually perceive an 'apparent' defectiveness:

$$p'_3 = (1 - b_3) \cdot p + a_3 \cdot (1 - p), \tag{A2}$$

Where $(1 - b_3) \cdot p$ represents the fraction of truly defective units that i_3 correctly identifies and $a_3 \cdot (1 - p)$ represents the fraction of truly good units that are incorrectly marked as defective. Equation A2 can be seen as a variant of Equation 11, adapted from the specific perspective of inspector i_3. The term p'_3 can then be plugged into Equation 12 in place of the (true) p, in order to obtain the equation governing the single-inspector OC curve with inspection errors:

$$P_a = \sum_{d=0}^{c} \binom{n}{d} \cdot [(1 - b_3) \cdot p + a_3 \cdot (1 - p)]^d \cdot \{1 - [(1 - b_3) \cdot p + a_3 \cdot (1 - p)]\}^{n-d}. \tag{A3}$$

Figure A1 presents the resulting OC curves for i_3 under L and H scenarios, with the ideal OC curve (with no errors) also shown for reference. The difference from the multiple-inspector case (cf. Figure 8) is striking. Even in the L scenario, the OC curve with a single inspector i_3 deviates substantially from the ideal one. For instance, at moderate p levels, P_a is significantly lower than it would be ideally, reflecting the i_3's false rejections (due to a_3) that erroneously inflate the perceived/apparent defectiveness. Conversely, at higher levels of defectiveness, the acceptance probability is higher than ideal, due to the missing detection of a portion of the defective units (due to b_3), making the lot seem better than it is. These distortions become extreme in the H scenario, where the single-inspector OC curve tends to become unreasonably selective.

Comparing Figure A1 with the multiple-inspector OC curves in Figure 8, it is clear that the same individual errors, which cause large distortions for one inspector, result in only minor deviations in the multiple-inspector OC curves (and those deviations shrink further as m increases). Therefore, the multiple-inspector approach transforms large, unpredictable OC-curve distortions at the individual level into small, manageable differences at the aggregated level.

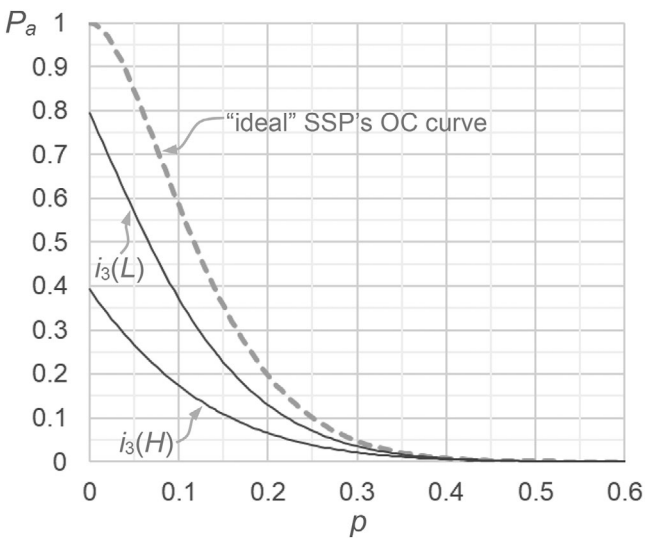


FIGURE A1 | OC curves for a single inspector (i_3), in the L scenario (with $a_3 = 6.5\%$ and $b_3 = 6.0\%$) and in the H scenario (with $a_3 = 14.1\%$ and $b_3 = 16.8\%$). The OC curves were constructed using the model in Equation A3. For ease of comparison, the ideal OC curve with no errors, constructed using the model in Equation 12, is also shown (as a dashed line).