QUALITY ASSESSMENT IN THE PRESENCE OF ADDITIONAL DATA IN PHOTOVOLTAICS

SABINE MEISEN[1], ANDREY PEPELYSHEV[2] AND ANSGAR STELAND[3]
[1] INSTITUTE FOR MEDICAL STATISTICS, RWTH AACHEN UNIVERSITY
[3] INSTITUTE OF STATISTICS, RWTH AACHEN UNIVERSITY
[3] INSTITUTE OF STATISTICS, RWTH AACHEN UNIVERSITY

ABSTRACT. Acceptance sampling represents an important tool for quality control. The practical methods of choice for non-normal variables are attribute sampling and variables sampling assuming normality applied to averages instead of single observations. Both methods usually lead to very large sample sizes and are therefore infeasible in practice if observations are expensive. We discuss and extend recent results developed for the photovoltaic industry and actively used there. Here - and presumably in other industries as well - additional data are available which can be used to construct valid and asymptotically optimal sampling plans for non-normal measurements. Consistency and asymptotic optimality of the sampling plans, which are random in our setup, as well as asymptotic normality of the required sample size are established under weak assumptions. We also provide sensitivity studies dealing with the effects of a systematic bias (shift) between the additional data and the lot (shipment), which may matter in practice. The new plans are investigated by simulations to some extent.

1. INTRODUCTION

In industry, quality control of lots or shipments of produced items is an important practical problem, particularly when high-quality leadership is a strategic goal. Delivering shipments of bad quality to customers may result in expensive law suits. In the photovoltaic industry, the expectations of customers in terms of quality of delivered photovoltaic modules (PV modules) are very high, presumably due to the fact that state of the art semi-conductor technologies are at the core of the business, thus being associated with digital precision. Acceptance sampling, which deals with the problem to determine the minimal sample size which controls the producer's as well as the consumer's risk, is therefore an important practical approach to the problem, cf. the recent monograph [11]. The customer is interested in the quality of his or her shipment of modules and not in the average outgoing quality. Thus, manufacturers interested in customer satisfaction should control production on the basis of outgoing shipments.

In this article, we study the classic acceptance sampling problem under the general assumption that the measurements (the power output in photovoltaics) may follow an arbitrary continuous distribution function (d.f.). To handle this case, we use additional data in the form of a historic data set, a situation which is typically present in photovoltaics. We derive sampling plans assuming a less general distributional model as in our previous work [13], but the results of the present article are valid under less restrictive assumptions. Further, the model of the present article nicely allows us to investigate sensitivity and robustness issues, respectively. We study the effect on the sampling plans, when the distributions of the shipment and the lab samples differ, which is of substantial interest for applications, where a systematic bias is of primary concern.

Our setup is as follows: We consider a shipment $X_1,\dots,X_N \sim F$ with a common distribution function F which is assumed to be continuous and strictly increasing with a finite fourth moment. Here and in what follows, $\sim$ always indicates that the random variables are independent and identically distributed. However, the stochastic relationship *between* samples may be arbitrary, i.e. they are not required to be independent. Let $\mu = E(X_1)$ and $\sigma^2 = E(X_1-\mu)^2 \in (0,\infty)$. In the photovoltaic problem motivating our work, X_i represents the true but random power output of the ith module, which is classified as non-conforming, i.e. being of low quality, if $X_i \le \tau$ for some constant τ. In practice, one puts $\tau = \mu(1-\varepsilon)$ where ε is the tolerance. The additional data are provided in the form of a historic sample $Y_1,\dots,Y_m \sim F$ of size m. Conditions on m will be given in Section 3. Clearly, the expected fraction of non-conforming modules in the shipment is given by

$$p = P(X_1 \le \tau) = F(\tau).$$

Suppose we have fixed two numbers $0 < AQL < RQL < 1$, namely the *acceptable qualilty level* (AQL) and the *rejectable quality level* (RQL), such that the lot should be accepted if $p \le AQL$, whereas it should be rejected when $p \ge RQL$. Since p is unknown to us and checking all modules is infeasible, our aim is to decide on the basis of a control sample $X'_1,\dots,X'_n$ of n measurements, n as small as possible, whether or not the shipment has to be accepted with controlled error probabilities of false decisions. Suppose that the decision is based on a statistic $T'_n = T'_n(X'_1,\dots,X'_n)$ using the decision rule to accept the shipment if and only if $T'_n > c$. A natural choice to use a standardized sum statistic, i.e. to accept if

$$T'_n = \sqrt{n}\frac{\overline{X}'_n - \tau}{\sigma} > c.$$

Then a solution (n,c) to the above problem, defined rigorously in Section 3, is called *sampling plan*.

In the case of normally distributed items, the optimal solution is well known and indeed based on the statistic T_n'. The resulting procedure is called *variables sampling*. However, in photovoltaics the power output measurements of PV modules are usually non-normal, such that variables sampling yields invalid sampling plans and is therefore not applicable. Indeed, all kinds of distributional shapes appear in practice. We discuss some of the factors leading to that unpleasant empirical fact in the next section. It is well known, and we shall provide the relevant arguments in Section 3 when deriving the plans, that the optimal sampling plan for an arbitrary d.f. F depends on that unknown F.

In photovoltaics, our key application for the methodology discussed in the present article, additional data from the production line, the so-called *flasher report tables*, are available and can be used in the construction of sampling plans. This key idea has been used in [1] to develop a photovoltaic-specific two-stage decision procedure using those flash data in order to construct valid sampling plans for normal as well as non-normal data. The procedure has also been implemented in a software tool which is used in the photovoltaic industry, cf. [3]. In [13], we elaborated on the new procedure addressing the non-normal case by establishing its asymptotic optimality as well as its asymptotic distribution assuming a general location scale model for the additional data. The mathematical proofs required advanced tools from probability theory such as empirical process theory and the functional delta method in metric spaces. Unfortunately, the formula for the asymptotic variance turned out to be rather complex and first simulations indicated that the variance of the estimated sample size is rather high even when hundreds of additional measurements are available. Thus, a natural question is whether it is possible to construct similar sampling plans leading to simpler formulas, simpler derivations and better accuracy in practice. For another recent proposal we refer to [4]. We show that indeed concise proofs can be given for the approximation theorem behind the construction of the sampling plans as well as for the central limit theorem for the estimated sample size. Here the classic Bahadur-Kiefer representation of sample quantiles plays a key role in the derivations. Further, we provide a new result on the strong consistency of the estimated sample size, thus strengthening the weak consistency implied by the central limit theorem, and the weak consistency of the estimated critical value under weak assumptions.

A further major goal of the present work is to study the effect of a systematic bias affecting both the lot and lab measurements and the historic data, respectively, in order to get a better understanding of what happens in this case. Indeed, such a sensitivity analysis with respect to the sampling distribution – depending on the viewpoint it can be regarded as a robustness study as well – can provide valuable insights into the stability of a procedure. We will study both a

model where a constant bias is present and a model where a asymptotically vanishing bias is assumed. The latter approach has a nice interpretation in terms of an *asymptotic learning effect* and shows that a local bias which is sufficiently small in larger control samples has no effect on the sampling plan. Although it is not surprising that a constant bias has such an effect, our findings show that the bias does not affect the required sample size, which greatly simplifies its treatment in practice.

The organization of the article is as follows. In Section 2, we provide some information on the photovoltaic background of the problem. Section 3 reviews the derivation of acceptance sampling plans and introduces the sampling plans proposed for the above setting including results on their asymptotic optimality. The study on the effects of a systematic bias and asymptotic learning is presented in Section 4. The result on the asymptotic normality is given in Section 5. Finally, Section 6 provides numerical results from a Monte Carlo study to assess the accuracy of the proposed sampling plans. We reveal an interesting and surprisingly strong effect of the algorithm used to calculate a sample quantile. Proofs of the results are postponed to an appendix.

2. Background on and Application to Photovoltaics

In the present section, we give a brief account of the photovoltaic background which motivated our research and the way how we approached the problem.

Photovoltaics represents one of the key technologies having the potential to provide a substantial contribution to the world's energy problem. Presumably, the main reason why the market share of solar energy is still relatively small compared to its potential and benefits is the fact that the costs per watt are still rather high. Although the costs have been substantially decreased in recent years, research still focuses on further reductions in costs, either by increasing the efficiency of a given solar cell technology or by developing new technologies, e.g. by employing cheaper materials and chemicals.

The economic life time of a photovoltaic system (PV system) ranges between 20 and 30 years. Thus, the quality in terms of the power output of the modules at delivery is a crucial parameter for the profitability of such an investment. Even small departures from the nominal power output accumulate to considerable losses over the years. Assessing the quality of PV modules, which is done under standard conditions (STC) in a lab, is therefore an important issue for quality control.

PV modules are an interconnected assembly of solar cells. To protect the cells from damage during manufacturing, delivery and usage, they are embedded between a tedlar plate on the bottom, a tempered glass on the top, and framed, usually with an aluminium frame. Since a

single module can produce only a limited amount of electricity, around 200 watts under STC, a PV system consists of many connected PV modules.

There are two common technologies to manufacture PV modules: Crystalline modules use silicon solar cells produced from solid Si wafers, whereas the CIS thin-film technology applies copper (C), indium (I) and selenum (S) in a layer construction of around 2 μm onto a substrate. The electrical properties such as the spectrum of the sun light transformed into electricity, the loss of efficiency when exposed to heat, a serious issue for systems installed in Southern Europe or Africa, or the efficiency when there are clouds as it is often the case in Northern Europe, heavily depend on the technology and various other physical parameters of the chosen module type.

Calibrating a PV module in a testing laboratory is also a different problem, since a couple of factors may complicate collecting measurements and also may lead to considerable difference of indoor and outdoor measurements. As reviewed and experimentally analyzed in [14], the following effects matter in practice:

(i) Measurement related sweep-time effects refering to the influence of the duration to complete an IV scan. Depending on the selected flash tester, which generates a pulse of calibrated light, such a scan can be based on up to 100 flashes. The duration of each flash is typically 100 ms. For details on accurate testing of PV panels we refer to [10].

(ii) Spectral mismatch arising when using a reference cell with a spectral response different from that of the device under test; its size depends on the spectral irradiance distribution of the spectral simulator with respect to the reference spectrum AM 1.5G.

(iii) Finally, thin-film modules are affected by the effect of light soaking, since the performance (even under standard test conditions) depends on the history of module (exposure to light or storage in the dark). The effect is in effect at the time of delivery but disappears when the modules are exposed to sun light for several days. The light soaking effect was first reported in [9] and is addressed to the tunneling of electrons trapped in deep states of CdS to holes in the CIS layer valence band under illumination, resulting in an increase in the open-circuit voltage and fill factor. During that light soaking period the performance can increase by $2-5\%$, cf. [7]. In industrial practice, it can even be larger.

Finally, it is common practice in industry to classify the produced PV modules in classes. As a consequence of the above discussion, when analyzing *comparable* modules, i.e. modules of the same technology and power rating satisfying additional criteria for inclusion or exclusion in a study, the true distribution of measurements may have any form. Especially, measurements are typically non-normal, thus violating the classic assumptions in statistical acceptance sampling.

Relying on ad hoc proposals such as forming subgroups and then applying variables sampling to the subgroup means to ensure approximatively the normal assumption, is not feasible due to the high costs of taking control measurements, since this procedure leads to enormous sample sizes. However, taking the control measurements is very expensive. For the same reason, applying attribute sampling is no reasonable solution.

Solar cells are manufactured in a production line. The performance of each module is measured in a sun simulator using short flashes. These measurements are therefore called flash measurements and form the flash data tables. In present days, they are routinely collected by manufacturers, thus often large samples are available. However, these cheap measurements may differ from the measurements taken in a photovoltaic laboratory. One should check carefully, whether a given flasher report table follows the same distribution as the shipment before applying the methods discussed in the present article. Some standard tests and their application to real photovoltaic data are described in [3]. For a new approach to the problem we refer to the recent work [12].

3. The Acceptance Sampling Problem

Acceptance sampling is a well established field of statistics and quality control – at least for the classic distributional assumptions. For basic notions we therefore refer to [11]. Our goal is to find an acceptance sampling plan (n,c), i.e. a sample size n for a control sample and a critical value c, such that

$$P(T_n' > c) \geq 1-\alpha, \qquad p \leq AQL, \tag{1}$$

and

$$P(T_n' > c) \leq \beta, \qquad p \geq RQL. \tag{2}$$

Here α is an upper bound on the probability that the shipment is rejected when it is of high quality, thus controlling the producer's risk, whereas β is the consumer's risk that the shipment is accepted although it is of low quality. Our derivations below will show that $P(T_n' > c)$, the operating characteristic (OC), is indeed a function of the quality level p. Approximations based on large sample theory will then allow us to solve the problem to construct appropriate sampling plans. We first discuss the unrealistic case that the underlying distribution is known and then proceed to a general solution for an arbitrary unknown distribution of the measurements.

3.1. **The Case of a Known Distribution.** Let us assume that F and therefore $\mu = \int x\,dF(x)$ as well as $\sigma^2 = \int x^2\,dF(x) - \mu^2$ are known. Using the crucial relationship $F(\tau) = p \Leftrightarrow \tau = F^{-1}(p)$, we obtain

$$T_n' > c \quad \Leftrightarrow \quad \sqrt{n}\frac{\overline{X}_n' - \mu}{\sigma} > c + \frac{\sqrt{n}(F^{-1}(p) - \mu)}{\sigma}.$$

By virtue of the central limit theorem, we get the approximation

$$P\left(\sqrt{n}\,\frac{\overline{X_n'} - \tau}{\sigma} > c\right) \approx 1 - \Phi\left(c + \frac{\sqrt{n}(F^{-1}(p) - \mu)}{\sigma}\right). \tag{3}$$

Notice that this approximation requires n to be large. However, since statistical inference should never be based on too few observations, assuming that the central limit theorem provides a sufficiently accurate approximation should not be too restrictive for many distributions F. Further, one may check the accuracy of the above approximation after calculating the sampling plan, such that c is fixed, using historic data which is available by assumption. This could be done, for instance, by estimating the Berry-Esséen upper bound or by means of a simulation study; the latter approach being preferable.

Thus, (1) and (2) are approximately satisfied, if we select (n,c) such that

$$1 - \Phi\left(c + \frac{\sqrt{n}(F^{-1}(AQL) - \mu)}{\sigma}\right) \geq 1 - \alpha \tag{4}$$

and

$$1 - \Phi\left(c + \frac{\sqrt{n}(F^{-1}(RQL) - \mu)}{\sigma}\right) \leq \beta \tag{5}$$

hold true. Since Φ is strictly increasing, these inequalities are equivalent to

$$\Phi^{-1}(\alpha) - \frac{\sqrt{n}(F^{-1}(AQL) - \mu)}{\sigma} \geq c \geq \Phi^{-1}(1-\beta) - \frac{\sqrt{n}(F^{-1}(RQL) - \mu)}{\sigma}.$$

Consider the left and right sides of the above chain of inequalities as functions in the real variable n. Then we arrive at the following proposition.

The optimal sampling plan (n,c) is obtained as the intersection of the mappings

$$n \longmapsto \Phi^{-1}(\alpha) - \frac{\sqrt{n}(F^{-1}(AQL) - \mu)}{\sigma}$$

and

$$n \longmapsto \Phi^{-1}(1-\beta) - \frac{\sqrt{n}(F^{-1}(RQL) - \mu)}{\sigma}.$$

Equating the above mappings leads to

$$\Phi^{-1}(\alpha)-\Phi^{-1}(1-\beta)=\frac{\sqrt{n}}{\sigma}(F^{-1}(AQL)-F^{-1}(RQL)).$$

By assumption, F is strictly increasing, which allows us to solve the equation for n and c leading the following result.

For known distributional parameters μ and σ, the optimal sampling plan is given by

$$n_\infty(\mu,\sigma) = \frac{\sigma^2\left(\Phi^{-1}(\alpha)-\Phi^{-1}(1-\beta)\right)^2}{\left(F^{-1}(AQL)-F^{-1}(RQL)\right)^2}, \tag{6}$$

$$c(\mu,\sigma) = \Phi^{-1}(\alpha)-\frac{\sqrt{n}(F^{-1}(AQL)-\mu)}{\sigma}. \tag{7}$$

Particularly, the optimal sampling plan depends on the unknown d.f. F of the measurements.

Theorem 3.1 shows that the asymptotically optimal sampling plan requires knowledge of the location μ and dispersion σ as well as knowledge of the distribution of the quality measurements. But those quantities are unknown to us, such that the sampling plan can not be applied in practice.

3.2. **The Case of an Unknown Distribution.** Clearly, the case that F is known to us is only of theoretical interest. Thus, we shall now assume that F is an arbitrary d.f. such that the fourth moment is finite. The basic idea is now to estimate the unknown quantities in the formulas derived in the previous subsection. However, various mathematical problems now arise. One has to establish an approximation of the OC curve leading to the same sampling plans one obtains when estimating unknown quantities in formulas (6) and (7). It turns out that within the framework of the present article, a transparent proof of such an approximation can be given without relying on empirical process theory as in [13]. Further, we present a new result establishing strong consistency of the estimated sample size and weak consistency of the estimated critical value under fairly weak assumptions.

We shall now derive the sampling plan in the case of unknown parameters μ and σ when additional data $Y_1,\dots,Y_m$ are given. Statistical intuition suggests to use the modified decision rule

$$T_n := \sqrt{n}\frac{\overline{X'_n}-\tau}{S_m} > c,$$

where $S_m^2 = \frac{1}{m-1}\sum_{i=1}^m (Y_i - \overline{Y}_m)^2$ with $\overline{Y}_m = \frac{1}{m}\sum_{i=1}^m Y_i$. Let us introduce the empirical distribution function of the historic sample defined by

$$F_m(y) = \frac{1}{m}\sum_{i=1}^{m} 1(Y_i \le y), \qquad y \in \mathbb{R}.$$

Here $1(A) = 1$, if the expression A (defining an event) is true, and $1(A) = 0$, if A is false. As usual,

$$F_m^{-1}(p) = \inf\{t \in \mathbb{R} : F_m(t) \ge p\}, \qquad p \in (0,1), \tag{8}$$

denotes the left continuous empirical quantile function, i.e. $F_m^{-1}(p) = Y_{(\lfloor np \rfloor + 1)}$, where $Y_{(1)} \le \cdots \le Y_{(m)}$ denotes the order statistic.

In the sequel, we need the following regularity assumption on the sample sizes n and m.

Assumption (A): $\frac{n}{m} \to 0$, as $n, m \to \infty$.

The reasoning behind that assumption is the following: The construction of the asymptotically optimal procedure requires certain approximations as in (3). Now both n and m have to be large, but, in addition, n/m has to be small. However, that condition is not restrictive in practice, as long as the historic data set is large enough. This is the typical case in photovoltaics and presumably in other areas of application as well.

The approximation of the operating characteristic $P(T_n > c)$ is now more involved and given in the following theorem, which is an analog of [13, Theorem 3.1].

Suppose F is a continuous and strictly increasing d.f. with a finite second moment. If $X_1', \ldots, X_n'$, the control sample, and $Y_1, \ldots, Y_m$, the historic data set, are random samples satisfying Assumptions (A), then there exists a sequence $\delta_n(p)$, $n \in \mathbb{N}$, of random variables with $\delta_n(o) = o_P(1)$, such that for all $c \in \mathbb{R}$

$$P(T_n > c) = P\left(\sqrt{n}\frac{\overline{X}_n' - \mu}{\sigma} + \delta_n(p) > c + \frac{\sqrt{n}(F_m^{-1}(p) - \overline{Y}_m)}{S_m}\right)$$

leading to the approximation

$$P(T_n > c) \approx 1 - \Phi\left(c + \frac{\sqrt{n}(F_m^{-1}(p) - \overline{Y}_m)}{S_m}\right).$$

A proof of this result is given in the appendix. Notice that the result holds true under the weak assumption of a finite second moment of the underlying distribution. Repeating the derivations of the previous subsection, we obtain the analog of Proposition 3.1. The optimal sampling plan

(n,c) is obtained as the intersection of the mappings

$$n \longmapsto \Phi^{-1}(\alpha) - \frac{\sqrt{n}(F_m^{-1}(AQL) - \overline{Y}_m)}{S_m}$$

and

$$n \longmapsto \Phi^{-1}(1-\beta) - \frac{\sqrt{n}(F_m^{-1}(RQL) - \overline{Y}_m)}{S_m}.$$

One has to take into account that F_m (and therefore F_m^{-1}) is not strictly increasing and $F_m^{-1}(AQL) < F_m^{-1}(RQL)$ may not hold. However, the latter condition holds true if m is large enough, which allows us to solve algebraically the equation as in the previous subsection leading to formulas for n and c depending on F_m^{-1}. Further, the empirical quantiles $F_m^{-1}(p)$ converge to $F^{-1}(p)$ with probability 1, since that property is well known to be equivalent to $F_m(y) \to F(y)$ with probability 1, as $m \to \infty$, which of course holds true, and $\overline{Y}_m$ to μ as well as S_m to σ, as $m \to \infty$, with probability 1, as long as the underlying d.f. F has a finite second moment. A more refined argument yields the following result, which is proved in the appendix.

Suppose F is a d.f. with a finite fourth moment. For unknown distributional parameters, the estimated sampling plan is given by

$$\begin{aligned} n_m &= \frac{S_m^2\left(\Phi^{-1}(\alpha) - \Phi^{-1}(1-\beta)\right)^2}{\left(F_m^{-1}(AQL) - F_m^{-1}(RQL)\right)^2}, \\ c_m &= \Phi^{-1}(\alpha) - \frac{\sqrt{n}(F_m^{-1}(AQL) - \overline{Y}_m)}{S_m}. \end{aligned}$$

It converges to the optimal sampling plan in the sense that

$$n_m/n_\infty(\mu,\sigma) \xrightarrow{a.s.} 1 \quad \text{and} \quad c_m - c(\mu,\sigma) \xrightarrow{P} 0,$$

as $m \to \infty$, provided Assumption (A) holds true as well as $F'(F^{-1}(AQL)) > 0$ and $F'(F^{-1}(RQL)) > 0$.

It is worth mentioning that the asymptotic optimality holds true under the weak regularity assumption of a finite fourth moment, the minimal assumption under which the statistic S_m^2 is meaningful in the sense of strong consistency and asymptotic normality. The proof given in the appendix shows that the strong consistency of n_m even holds under the weaker condition of a finite second moment. However, although the assumption of finite higher moments is not regarded an issue in photovoltaics, since measurements are often even bounded by definition of the sampling process, it may be in other areas of applications.

4. The Effect of a Bias and Asymptotic Learning

In practice, it may happen that the distribution of the historic data set, $Y_1,\dots,Y_m$, and the distribution of control measurements $X_1',\dots,X_n'$, which are made in a laboratory, do not coincide. Therefore the present section is devoted to a study of the effect of departures from the assumption $X_i \sim F$. We are interested in the effect of a systematic bias as it may happen when using a differently calibrated measurement system to measure modules of the lot and of the control sample. We will see that in this setting the optimal sampling plan depends on the bias, even asymptotically. That result is interesting in its own right, but can be used in practice as well, in order to correct for such a systematic bias, provided an estimate of the bias is available.

The next step of our analysis is then to model the bias as a function of the sample size n which tends to 0, as n approaches ∞, i.e. to consider local alternatives, and to ask under which conditions on n and m there is no asymptotic effect of that disturbance. Considering a sequence of distributional models indexed by the optimal sample size n has two interpretations, both of which are meaningful. Firstly, one may conduct a sequence of experiments where the risk probabilities α and β are decreased from experiment to experiment leading to larger sample sizes n. Then a bias of the order $o(1)$ can be interpreted as a *learning effect* when conducting more and more experiments. Secondly, when conducting one experiment, a model with a bias of order $o(1)$ captures the fact that presumably more efforts are spent on obtaining better measurements with smaller bias, when analyzing large samples to get very precise results, which are more expensive than conducting small experiments.

4.1. The Effect of a Fixed Systematic Bias.

Let us assume that the random variables $X_1,\dots,X_N$ representing the shipment (lot) satisfy

$$X_1+\gamma,\dots,X_N+\gamma \sim F, \qquad N\in\mathbb{N},$$

for a given constant shift γ, whereas the historic sample is not affected by the shift, i.e. $Y_1,\dots,Y_m \sim F$. This means that the distribution of the shipment measurements is equal in distribution to the historical measurements after adding γ. Since the control measurements are selected from the shipment, we have $X_1'+\gamma,\dots,X_n'+\gamma \sim F$ as well. Equivalently, we could assume that $X_1,\dots,X_N \sim F$ and $Y_1-\gamma,\dots,Y_m-\gamma \sim F$, but the above formulation simplifies the derivations.

Now the fraction of non-conforming modules satisfies the equation

$$p = E\left(\frac{1}{N}\sum_{i=1}^{N} 1(X_i\le\tau)\right) = F(\tau+\gamma)$$

or, equivalently, $\tau = F^{-1}(p) - \gamma$. It turns out that when going through all derivations given in the previous section and the proof of Theorem 3.2 in the appendix, we arrive at the following approximation of the operating characteristic

$$P(T_n' > c) \approx 1 - \Phi\left(c + \frac{\sqrt{n}(F_m^{-1}(p) - \overline{Y}_m)}{S_m} - \frac{\sqrt{n}\gamma}{S_m}\right)$$

leading to the optimal sampling plan

$$n_\gamma = \frac{S_m^2(\Phi^{-1}(\alpha) - \Phi^{-1}(1-\beta))^2}{(F_m^{-1}(AQL) - F_m^{-1}(RQL))^2}, \tag{9}$$

$$c_\gamma = \Phi^{-1}(\alpha) - \frac{\sqrt{n}(F_m^{-1}(AQL) - \overline{Y}_m)}{S_m} + \frac{\sqrt{n}\gamma}{S_m}. \tag{10}$$

The interpretation of these results is as follows: The bias (location shift) γ has only an effect on the critical value, but the optimal sample size remains the same. As a consequence, we may formulate the following rule of thumb: If the measurements of the shipment are shifted by γ compared to the historic sample measurements, one can apply the optimal sampling plan derived in the previous section when correcting the critical value by the additive term $\sqrt{n}\gamma/S_m$.

4.2. **The Effect of Asymptotic Learning.** Suppose now that

$$X_1 + \gamma_n, \ldots, X_N + \gamma_n \sim F, \qquad N \in \mathbb{N}, \tag{11}$$

for some sequence $\{\gamma_n\}$ of real numbers converging to 0 as $n \to \infty$.

Assumption (B): Suppose the sample sizes n, m and the sequence $\{\gamma_n\}$ are selected such that

$$\frac{n}{m} = o(1) \quad \text{and} \quad \sqrt{m}\gamma_n = o(1),$$

as $n, m \to \infty$.

Suppose γ_n is chosen as

$$\gamma_n = \Gamma n^{-\xi}, \qquad n \geq 1,$$

for positive constants Γ and ξ. Then it is easily seen that Assumption (B) is satisfied if in addition to $n/m = o(1)$ the condition

$$\frac{m}{n^{2\xi}} = o(1)$$

holds true. For brevity of presentation, we omit the proof of the following result and refer to [8] for details.

Under the sampling model (11) and Assumption (B), the approximation of the operating characteristic obtained in Theorem 3.2 still holds true, such that the sampling plan (n_m, c_m) given in Theorem 3.2 is asymptotically optimal in this case.

5. Asymptotic Normality

The present section is devoted to a study of the asymptotic distribution of the optimal sample size, as the sample size m of the historic data set tends to ∞. It turns out that in the present setting a proof of the asymptotic normality can be based on the Bahadur-Kiefer representation of sample quantiles and the delta method in $\mathbb{R}^3$ combined with the multivariate central limit theorem.

For simplicity of presentation, we use the abbreviations

$$p_\alpha = AQL \qquad \text{and} \qquad p_\beta = RQL$$

in what follows.

Suppose the historic data set as well as the lot and control measurements are distributed according to a strictly increasing and continuous d.f. F such that $\int x^4 dF(x) < \infty$ and

$$F'(F^{-1}(p_\alpha)) > 0 \quad \text{as well as} \quad F'(F^{-1}(p_\beta)) > 0.$$

Let $n_\infty = n_\infty(\mu, \sigma)$ and n_m be as in Theorem 3.1 and Theorem 3.2. If Assumption (A) holds true, n_m is asymptotically normal,

$$\sqrt{m}(n_m - n_\infty) \xrightarrow{d} N(0, \eta^2), \quad m \to \infty,$$

for $\eta^2 = \mathbf{g}\Sigma\mathbf{g}'$, where

$$\mathbf{g} = \frac{C(\alpha,\beta)}{(F^{-1}(p_\alpha) - F^{-1}(p_\beta))^3} \cdot (-2\sigma^2, 2\sigma^2, (F^{-1}(p_\alpha) - F^{-1}(p_\beta))'$$

and

$$\Sigma = \begin{pmatrix} \Sigma_{\alpha\alpha} & \Sigma_{\alpha\beta} & \zeta_{Y\alpha} \\ \Sigma_{\alpha\beta} & \Sigma_{\beta\beta} & \zeta_{Y\beta} \\ \zeta_{Y\alpha} & \zeta_{Y\beta} & \zeta_{YY} \end{pmatrix}$$

with entries

$$\begin{aligned}
\Sigma_{\alpha\alpha} &= \frac{p_\alpha(1-p_\alpha)}{(F'(F^{-1}(p_\alpha)))^2}, \qquad \Sigma_{\beta\beta} = \frac{p_\beta(1-p_\beta)}{\left(F'(F^{-1}(p_\beta))\right)^2},\\
\Sigma_{\alpha\beta} &= \frac{p_\alpha(1-p_\beta)}{(F'(F^{-1}(p_\alpha)))\left(F'(F^{-1}(p_\beta))\right)}\\
\zeta_{Y\alpha} &= \frac{E((Y_1-\mu)^2-\sigma^2)(p_\alpha - 1(Y_1 \le F^{-1}(p_\alpha)))}{F'(F^{-1}(p_\alpha))},\\
\zeta_{Y\beta} &= \frac{E((Y_1-\mu)^2-\sigma^2)(p_\beta - 1(Y_1 \le F^{-1}(p_\beta)))}{F'(F^{-1}(p_\beta))},\\
\zeta_{YY} &= \mathrm{Var}\,((Y_1-\mu)^2),
\end{aligned}$$

leading to

$$\begin{aligned}
\eta^2 \;=\; & \frac{4\sigma^4(\Phi^{-1}(\alpha)-\Phi^{-1}(1-\beta))^4}{(p_\alpha-p_\beta)^6}\big[4\sigma^4(\Sigma_{\alpha\alpha}-2\Sigma_{\alpha\beta}+\Sigma_{\beta\beta})\\
& +4\sigma^2(F^{-1}(p_\alpha)-F^{-1}(p_\beta))(\zeta_{Y\beta}-\zeta_{Y\alpha})+\zeta_{YY}(F^{-1}(p_\alpha)-F^{-1}(p_\beta))^2\big].
\end{aligned}$$

It is worth mentioning that the above formulas are more transparent than those obtained in [13]. Although it is interesting that the estimated sample size is asymptotically normal, the result of limited value for practical purposes, since simulations have shown that the convergence is rather slow. Consequently, it is not clear whether the construction of asymptotic confidence intervals based on the above result would yield intervals with accurate coverage probabilities.

6. SIMULATIONS

We conducted small-scale simulations in order to investigate to some extent the accuracy of the new sampling plans under some models. We were also interested in the effect of the method used to calculate a sample quantile, as standard statistical software usually offers several methods.

As a kind of benchmark model, we selected the normal distribution with mean 220 and variance 4. The reason is that photovoltaic modules are often traded with a nominal power output of 220 watts. The variance, 4, captures to some extent the variability observed in practice, although that varies with technology. Two kinds of departures from the normality assumption were studied: One-sided contaminations inducing skewness in the data samples and symmetric contaminations inducing, e.g., a different kurtosis. The amount of contamination to induce these

effects was chosen as 20%, and the mean and the variance of the contaminating subpopulations was chosen between $210-240$ and $4-8$, respectively.

For each parameter combination given by the sample size m of the historic data set and the parameters of the above mixture model, we calculated Monte Carlo estimates for the expected required sample size, $E(n_m)$, and the associated standard deviation of n_m based on 50000 replications. In addition, the quartiles $q_{0.25}, q_{0.5}$ and $q_{0.75}$ of the distribution of n_m are reported which enables us to judge the skewness of the distribution of n_m. Finally, the expected critical value, $E(c_m)$, and its standard deviation are provided.

6.1. **One-Sided Contaminations.** Data sets according to the following models were simulated:

$$\begin{aligned} &\text{Model 1:} && X_i \sim F_1 = N(220,4), \\ &\text{Model 2:} && X_i \sim F_2 = 0.1N(210,6) + 0.9N(230,4), \\ &\text{Model 3:} && X_i \sim F_3 = 0.9N(220,4) + 0.1N(230,8). \end{aligned} \tag{12}$$

The parameters were specified as $AQL = 2\%$, $RQL = 5\%$ and $\alpha = \beta = 5\%$. Tables 1-3 show the simulation results for these models.

TABLE 1. Characteristics of the distribution of n_m and c_m for Model 1

m	$E(n_m)$	$sd(n_m)$	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	$E(c_m)$	$sd(c_m)$
100	195.4	1144.0	30	61	143	19.4	16.9
250	79.1	77.5	37	58	94	15.3	5.6
500	74.9	44.5	46	64	91	15.5	3.9
5000	65.6	10.5	58	65	72	14.9	1.1
50000	64.8	3.2	63	65	67	14.9	0.3

TABLE 2. Characteristics of the distribution of n_m and c_m for Model 2

m	$E(n_m)$	$sd(n_m)$	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	$E(c_m)$	$sd(c_m)$
100	317.6	2568.7	44	98	231	38.2	34.3
250	123.7	133.9	54	89	148	30.6	11.8
500	118.6	72.7	70	101	145	31.4	8.1
5000	103.6	17.5	91	102	114	30.4	2.3
50000	102.5	5.5	99	102	106	30.4	0.7

TABLE 3. Characteristics of the distribution of n_m and c_m for Model 3

m	$E(n_m)$	$sd(n_m)$	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	$E(c_m)$	$sd(c_m)$
100	679.1	5724.1	92	193	455	23.6	22.2
250	256.2	253.1	117	186	303	18.7	7.0
500	241.7	143.6	148	206	293	19.0	4.8
5000	211.3	33.7	188	208	232	18.2	1.4
50000	209.2	10.8	202	209	216	18.2	0.4

6.2. **Symmetric Contaminations.** Let us now study symmetric contaminations according to the models

$$
\begin{aligned}
&\text{Model 4:} \quad X_i \sim F_4 = 0.2N(210,8) + 0.6(220,4) + 0.2N(230,8),\\
&\text{Model 5:} \quad X_i \sim F_5 = 0.2N(200,8) + 0.6(220,4) + 0.2N(240,8),\\
&\text{Model 6:} \quad X_i \sim F_6 = 0.2N(210,4) + 0.6(220,4) + 0.2N(230,4),\\
&\text{Model 7:} \quad X_i \sim F_7 = 0.2N(200,4) + 0.6(220,4) + 0.2N(240,4).
\end{aligned} \tag{13}
$$

Notice that in all models the observations have a mean of 220. In Model 4, we consider the case that one fifth of the contaminated data have the mean 210, whereas another fifth scatters around 230. The variance for these two subpopulations is 8. Model 6 is similar to Model 4 except that the variance is fixed at 4. In Models 5 and 7 the contaminating subpopulations have means 200 and 240. The empirical results for these models are provided in Tables 4-7.

TABLE 4. Characteristics of the distribution of n_m and c_m for Model 4

m	$E(n_m)$	$sd(n_m)$	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	$E(c_m)$	$sd(c_m)$
100	624.8	23533.2	79	161	364	31.7	34.2
250	203.9	191.6	96	150	242	25.2	9.3
500	193.7	111.4	121	167	234	25.5	6.4
5000	169.2	26.3	150	167	185	24.5	1.8
50000	167.8	8.2	162	168	173	24.5	0.6

It is interesting to note that $q_{0.5}$ for $m = 250$ is typically smaller then $q_{0.5}$ for $m = 100$ since the use of empirical distribution leads to rounding (discretion). Indeed, for a sample of size 250, the 0.02-quantile is the 5-order statistics of the sample ($250 \times 0.02 = 5$) and the 0.05-quantile is the 13-order statistics of the sample ($250 \times 0.05 = 12.5 \neq 13$)

TABLE 5. Characteristics of the distribution of n_m and c_m for Model 5

m	$E(n_m)$	$sd(n_m)$	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	$E(c_m)$	$sd(c_m)$
100	1951.6	22814.7	284	578	1319	55.9	52.9
250	727.9	677.7	347	539	872	44.6	16.3
500	705.7	405.9	438	606	849	45.5	11.5
5000	616.9	96.3	550	607	676	43.8	3.3
50000	608.6	29.6	588	608	629	43.6	1.0

TABLE 6. Characteristics of the distribution of n_m and c_m for Model 6

m	$E(n_m)$	$sd(n_m)$	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	$E(c_m)$	$sd(c_m)$
100	1023.9	19714.2	152	309	709	41.6	39.2
250	392.8	390.0	186	287	461	33.3	12.5
500	374.3	215.0	233	321	453	33.8	8.5
5000	327.3	51.6	291	322	358	32.5	2.4
50000	323.5	15.8	312	323	334	32.4	0.8

TABLE 7. Characteristics of the distribution of n_m and c_m for Model 7

m	$E(n_m)$	$sd(n_m)$	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	$E(c_m)$	$sd(c_m)$
100	3897.3	49612.1	572	1160	2643	76.7	72.8
250	1453.1	1435.4	686	1059	1718	60.8	22.8
500	1406.5	817.9	872	1201	1698	62.2	15.9
5000	1221.0	190.6	1085	1202	1337	59.7	4.5
50000	1204.8	58.5	1164	1204	1244	59.5	1.4

6.3. **Effect of Quantile Algorithms.** Statistical software such as R or SAS implements various standard procedures to calculate sample quantiles. R implements nine algorithms discussed in [5]. ALGORITHM 1 corresponds to the left inverse of the empirical distribution function, the definition we use in this article, see (8). However, R's default algorithm is ALGORITHM 7. SAS's PROC UNIVARIATE also provides several methods; the default is to use the average of the npth and $(np+1)$-th order statistic, if np is an integer, and the $(\lfloor np \rfloor + 1)$th order statistic, otherwise, corresponding to ALGORITHM 2.

We used the parameters $\alpha = \beta = 0.05$, $AQL = 0.02$ and $RQL = 0.05$ and simulated data according to the model

$$X_i \sim \mathrm{N}(220,4), \quad i = 1,\ldots,m.$$

For each algorithm and sample size $m = 100, 250, 500, 5000$, the same statistical quantities as above were estimated using $50,000$ simulation runs. For better comparison, for each algorithm the same random numbers were used by initializing the random number generator using the statement set.seed(17). The results are presented in Tables 8 and 9.

TABLE 8. Results for Algorithms 1-5

m	$E(n_m)$	$sd(n_m)$	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	$E(c_m)$	$sd(c_m)$
			ALGORITHM 1				
100	195.4	1144.0	30	61	143	19.4	16.9
250	79.1	77.5	37	58	94	15.3	5.6
500	74.9	44.5	46	64	91	15.5	3.9
5000	65.6	10.5	58	65	72	14.9	1.1
			ALGORITHM 2				
100	162.8	424.6	40	75	156	19.2	12.3
250	95.1	93.3	45	70	112	16.7	6.2
500	78.1	44.8	49	67	95	15.8	3.9
5000	65.9	10.5	59	65	72	15	1.1
			ALGORITHM 3				
100	195.4	1144.0	30	61	143	19.4	16.9
250	101.6	114.4	44	70	118	17.3	7.0
500	74.9	44.5	46	64	91	15.5	3.9
5000	65.6	10.5	58	65	72	14.9	1.1
			ALGORITHM 4				
100	195.4	1144.0	30	61	143	19.4	16.9
250	88.2	88.9	40	64	104	16.2	6.1
500	74.9	44.5	46	64	91	15.5	3.9
5000	65.6	10.5	58	65	72	14.9	1.1
			ALGORITHM 5				
100	162.8	424.6	40	75	156	19.2	12.3
250	95.1	93.3	45	70	112	16.7	6.2
500	78.1	44.8	49	67	95	15.8	3.9
5000	65.9	10.5	59	65	72	15	1.1

For ALGORITHM 1, we can see that the median $q_{0.5}$ is close to 65 for small m but the standard deviation of n_m is very large. In general, results for ALGORITHMS 2–6 and 8–9 are rather close to the results for ALGORITHM 1; in some cases even identical to the results for ALGORITHM 1. However, the results for ALGORITHM 7, which is used in R by default, are worse than the results when using ALGORITHM 1. In particular, the median $q_{0.5}$ is substantially larger than 65. For $m \geq 5000$, there are no notable differences.

TABLE 9. Results for Algorithms 6-9

m	$E(n_m)$	$sd(n_m)$	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	$E(c_m)$	$sd(c_m)$
ALGORITHM 6							
100	174	689.3	30	61	140	19.0	15.0
250	87.7	87.9	40	64	103	16.2	6.1
500	74.7	44.2	46	64	91	15.5	3.9
5000	65.5	10.5	58	65	72	14.9	1.1
ALGORITHM 7							
100	301.6	1405.4	51	104	239	23.1	19.1
250	108.5	107	50	79	129	17.5	6.6
500	83.2	49.3	51	71	101	16.1	4.1
5000	66.2	10.6	59	65	73	15.0	1.1
ALGORITHM 8							
100	157.1	437.8	36	69	147	18.9	12.5
250	92.0	90.1	43	68	109	16.5	6.1
500	76.8	44.3	48	66	93	15.7	3.8
5000	65.8	10.5	58	65	72	15.0	1.1
ALGORITHM 9							
100	157.8	431.0	37	71	149	18.9	12.4
250	92.7	90.7	44	68	110	16.5	6.1
500	77.1	44.4	48	66	94	15.7	3.8
5000	65.8	10.5	58	65	72	15.0	1.1

7. DISCUSSION

Sampling plans have been proposed for variables sampling when the true but unknown distribution is completely unknown. These plans require additional historic sampling information, which is usually available in photovoltaics, the key area of application we have in mind where the methods are already in active use. Our theoretical results show that the proposed sampling plans are consistent and asymptotically optimal under very weak assumptions. Moreover, the estimated sample size satisfies a central limit theorem. Whether or not those results remain valid when the measurements are dependent is still an open issue. However, our results do not require independence of the historic sample, the control measurements and the shipment. Indeed, *any* kind of dependence between the samples is allowed. Hence our results are valid both when drawing randomly the control measurements from the shipment and when using observations independent from the shipment, which is, e.g., the case when the modules used for the control measurements are not returned to the customer.

Our simulations indicate, firstly, that the presented methodology provides accurate sampling plans for a wide range of distributions, provided the size of the historic data set is sufficiently large. Otherwise, the estimated sample size is affected by a substantially large variance which hinders practical application. Here further research is in order to develop procedures with reduced variability, which would lead to improved sampling plans which can be used for smaller historic data sets.

The simulation study on the effect of the algorithm used to calculate sample quantiles reveals a striking and unexpected strong effect for small sample sizes. All algorithms estimate quantiles by calculating a function of at most two successive order statistics, i.e. they perform a smoothing operation in some cases. The effect on the results is surprisingly strong and points to the conjecture that, in general, smoothing may lead to better sampling plans. Again, future research is in order to reveal to which extent improved sampling plans can be constructed by using refined quantile estimation algorithms based on smoothing techniques.

The authors acknowledge financial support from the German Federal Ministery of the Environment, Nature Conservation and Nuclear Safety (grant no. 0325226).

APPENDIX: PROOFS

In this mathematical appendix, we use the following notations. For a function f defined on some subset of the p-dimensional Euclidean space $\mathbb{R}^p$, $\dot{f}$ denotes the gradient of f. When there is no danger of confusion, we identify column and row vectors. The symbol $\xrightarrow{P}$ denotes convergence in probability and $\xrightarrow{d}$ the convergence in distribution of a sequence of random variables and random vectors. For brevity of presentation, we use the stochastic o and O symbols: $o_P(1)$ stands for a random sequence R_n with $R_n \xrightarrow{P} 0$, as $n \to \infty$, whereas $O_P(1)$ denotes a sequence R_n such that for any $\varepsilon > 0$ one may find a constant C with $P(|R_n| > C) < \varepsilon$ for all n. As well known, $R_n = O_P(1)$ holds true, if R_n converges in distribution. We also frequently use rules such as $O_P(1) \cdot o(1) = o_P(1)$.

We need the following auxiliary results on the Bahadur representation of sample quantiles, a classical result dating back to [2] and [6], which in turn implies their joint asymptotic normality. For the reader's convenience, we provide those results in some detail.

(i) Let $p \in (0,1)$ and suppose $F'(F^{-1}(p)) > 0$. Then

$$F_m^{-1}(p) = F^{-1}(p) + \frac{F(F^{-1}(p)) - F_m(F^{-1}(p))}{F'(F^{-1}(p))} + o_P(m^{-1/2}).$$

(ii) Suppose $0 < p_1 < \cdots < p_k < 1$ and $F'(F^{-1}(p_i)) > 0$ for $i = 1, \ldots, k$. Then

$$\sqrt{m}(F_m^{-1}(p_1) - F^{-1}(p_1), \ldots, F_m^{-1}(p_k) - F^{-1}(p_k)) \xrightarrow{d} N(0, S_k),$$

where $S_k = (s_{ij})$ is a symmetric $k \times k$ matrix with entries

$$s_{ij} = \frac{p_i(1-p_i)}{F'(F^{-1}(p_i))},$$

for $1 \leq i, j \leq k$.

□

Proof. (*Theorem 3.2*)

To prove the assertion, we follow arguments used in [13]. Straightforward algebra leads to

$$T_n > c \quad \Leftrightarrow \quad \sqrt{n}\left(\frac{\overline{X'_n} - \mu}{\sigma} + T_1(n) + T_2(n) - T_3(n)\right) > c + \frac{\sqrt{n}(F_m^{-1}(p) - \overline{Y}_m)}{S_m}$$

where

$$\begin{aligned} T_1(n) &= \frac{\overline{X'}_n - \mu}{\sigma} \cdot \frac{\sigma - S_m}{S_m}, \\ T_2(n) &= \frac{F_m^{-1}(p) - F^{-1}(p)}{S_m}, \\ T_3(n) &= \frac{\overline{Y}_m - \mu}{S_m}. \end{aligned}$$

The assertion follows, if we show that

$$\delta_n(p) = \sqrt{n}(T_1(n) + T_2(n) - T_3(n)) = o_P(1),$$

as $n \to \infty$. Clearly, $\sqrt{n}T_1(n) = o_P(1)$ by Slutzky's lemma, since $\sqrt{n}(\overline{X}'_n - \mu)/\sigma \xrightarrow{d} N(0,1)$, as $m \to \infty$, and $(\sigma - S_m)/\sigma \xrightarrow{P} 0$, as $n \to \infty$. Next notice that

$$\sqrt{n}T_2(n) = \frac{1}{S_m}\frac{\sqrt{n}}{\sqrt{m}}\sqrt{m}(F_m^{-1}(p) - F^{-1}(p)).$$

The first factor converges to $1/\sigma$, in probability, the second one is $o(1)$ by Assumption (A) and the third factor is asymptotically normal by Theorem 7, as $m \to \infty$. Thus, $T_2(n) = o_P(1)$. Similarly, we have $\sqrt{n}T_3(n) = \frac{\sqrt{n}}{\sqrt{m}} \frac{\sigma}{S_m} \sqrt{m} \frac{\overline{Y}_m - \mu}{\sigma} = o_P(1)$ by Assumption (A). Let us now check the approximation for $P(T_n > c)$. Clearly,

$$U_n = \sqrt{n}(\overline{X}'_n - \mu)/\sigma + \delta_n(p) \xrightarrow{d} N(0,1),$$

as $n \to \infty$. This implies that $\sup_{z \in \mathbb{R}} |P(U_n \le z) - \Phi(z)| = o(1)$, by virtue of Glivenko-Cantelli theorem, which completes the proof, since

$$\begin{aligned} |P(T_n > c) - [1 - \Phi(c + \sqrt{n}(F_m^{-1}(p) - \overline{Y}_m))]| \\ = |P(U_n > c) - [1 - \Phi(c + \sqrt{n}(F_m^{-1}(p) - \overline{Y}_m))]| \\ \le \sup_{z \in \mathbb{R}} |P(U_n > z) - [1 - \Phi(z)]| = o(1). \end{aligned}$$

□ □

Proof. (*Theorem 3.2*)

Obviously,

$$\frac{n_m}{n_\infty(\mu,\sigma)} = \frac{S_m^2}{\sigma^2} \frac{(F^{-1}(AQL) - F^{-1}(RQL))^2}{(F_m^{-1}(AQL) - F_m^{-1}(RQL))^2}.$$

The first factor on the right-hand side converges to 1, almost surely, as $m \to \infty$, as well as the second one, since $F_m^{-1}(AQL) - F_m^{-1}(RQL) \to F^{-1}(AQL) - F^{-1}(RQL)$, almost surely, as $m \to \infty$. To show the second assertion, recall

$$\begin{aligned} c_m &= \Phi^{-1}(\alpha) - \frac{\sqrt{n}(F_m^{-1}(AQL) - \overline{Y}_m)}{S_m}, \\ c(\mu,\sigma) &= \Phi^{-1}(\alpha) - \frac{\sqrt{n}(F^{-1}(AQL) - \mu)}{\sigma}, \end{aligned}$$

and notice that

$$c_m - c(\mu,\sigma) = \frac{\sigma}{S_m} W_{n,m},$$

where

$$W_{n,m} = \left[-\frac{\sqrt{n}(F_m^{-1}(AQL) - \overline{Y}_m)}{\sigma} + \frac{\sqrt{n}(F^{-1}(AQL) - \mu)}{\sigma} \frac{S_m}{\sigma} \right].$$

Clearly, $\sigma/S_m = 1 + (\sigma - S_m)/S_m \to 1$, as $m \to \infty$, a.s. Thus, it suffices to show that $W_{n,m}$ is $o_P(1)$. Using $S_m/\sigma = 1 + (S_m - \sigma)/\sigma$, we see that the second summand of $W_{n,m}$ can be written as

$$\frac{\sqrt{n}(F^{-1}(AQL) - \mu)}{\sigma} + \frac{F^{-1}(AQL) - \mu}{\sigma} \sqrt{n} \frac{S_m - \sigma}{\sigma}.$$

Rearranging terms, we see that $W_{n,m}$ can be written as

$$\sqrt{\frac{n}{m}} \left[-\sqrt{m} \frac{F_m^{-1}(AQL) - F^{-1}(AQL)}{\sigma} + \sqrt{m} \frac{\overline{Y}_m - \mu}{\sigma} + \frac{F^{-1}(AQL) - \mu}{\sigma} \sqrt{m} \frac{S_m - \sigma}{\sigma}. \right]$$

By Theorem 7 (ii),

$$\sqrt{m}(F_m^{-1}(AQL) - F^{-1}(AQL)) = O_P(1).$$

Further, since $EY_1^4 < \infty$ by assumption, we also have

$$\sqrt{m}(\overline{Y}_m - \mu) = O_P(1) \quad \text{and} \quad \sqrt{m}(S_m - \sigma) = O_P(1),$$

as $m \to \infty$; for the second statistic also confer the proof of Theorem 5. Since by Assumption (A) $n/m = o(1)$, as $n, m \to \infty$, we obtain $W_{n,m} = o_P(1)$ follows, which completes the proof. □ □

Proof. (*Theorem 5*)

For brevity of notation, we put $p_\alpha = AQL$ and $p_\beta = RQL$. The Bahadur representation yields for $p \in \{p_\alpha, p_\beta\}$

$$\sqrt{m}[F_m^{-1}(p) - F^{-1}(p)] = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \frac{p - 1(Y_i \le F^{-1}(p))}{F'(F^{-1}(p))} + o_P(1).$$

Further, with $\widetilde{Y}_i = Y_i - \mu$ and $\overline{\widetilde{Y}} = m^{-1}\sum_{i=1}^m \widetilde{Y}_i$ we obtain the asymptotic linearization

$$\begin{aligned}\frac{1}{\sqrt{m}}\sum_{i=1}^m [(Y_i-\overline{Y})^2-\sigma^2] &= \frac{1}{\sqrt{m}}\sum_{i=1}^m (\widetilde{Y}_i^2-\sigma^2) - 2\overline{\widetilde{Y}}\sqrt{m}\overline{\widetilde{Y}} + \sqrt{m}(\overline{\widetilde{Y}})^2 \\ &= \frac{1}{\sqrt{m}}\sum_{i=1}^m (\widetilde{Y}_i^2-\sigma^2) + o_P(1),\end{aligned}$$

since $\sqrt{m}\overline{\widetilde{Y}} = O_P(1)$ and $\overline{\widetilde{Y}} = o_P(1)$, which, of course, carries over to $\sqrt{m}(S_m^2-\sigma^2)$. This gives

$$U_m = \sqrt{m}\begin{pmatrix} F_m^{-1}(p_\alpha) - F^{-1}(p_\alpha) \\ F_m^{-1}(p_\beta) - F^{-1}(p_\beta) \\ S_m^2 - \sigma^2 \end{pmatrix} = \frac{1}{\sqrt{m}}\sum_{i=1}^m Z_i + o_P(1),$$

where for $i = 1,\dots,m$

$$Z_i = \begin{pmatrix} (p_\alpha - 1(Y_i \le F^{-1}(p_\alpha)))/F'(F^{-1}(p_\alpha)) \\ (p_\beta - 1(Y_i \le F^{-1}(p_\beta)))/F'(F^{-1}(p_\beta)) \\ (Y_i-\mu)^2-\sigma^2 \end{pmatrix}.$$

Notice that $Z_1,\dots,Z_m$ are i.i.d. with a finite second moment. Thus, an application of the multivariate central limit theorem yields

$$U_m \xrightarrow{d} U \sim N(\mathbf{0},\Sigma),\ \text{with } \Sigma = E(Z_1^2) = \begin{pmatrix} \Sigma_{\alpha\alpha} & \Sigma_{\alpha\beta} & \zeta_{Y\alpha} \\ \Sigma_{\alpha\beta} & \Sigma_{\beta\beta} & \zeta_{Y\beta} \\ \zeta_{Y\alpha} & \zeta_{Y\beta} & \zeta_{YY} \end{pmatrix},$$

as $m \to \infty$, where the entries are as given in the theorem. Now consider

$$\begin{aligned}\sqrt{m}(n_m - n_\infty) = \sqrt{m}\Bigg(&\frac{S_m^2 \cdot \left(\Phi^{-1}(\alpha) - \Phi^{-1}(1-\beta)\right)^2}{\left(F_m^{-1}(p_\alpha) - F_m^{-1}(p_\beta)\right)^2} \\ &- \frac{\sigma^2 \cdot \left(\Phi^{-1}(\alpha) - \Phi^{-1}(1-\beta)\right)^2}{\left(F^{-1}(p_\alpha) - F^{-1}(p_\beta)\right)^2}\Bigg).\end{aligned} \tag{14}$$

Observe that we may write

$$\sqrt{m}(n_m - n_\infty) = \sqrt{m}[g(F_m^{-1}(p_\alpha), F_m^{-1}(p_\beta), S_m^2) - g(F^{-1}(p_\alpha), F^{-1}(p_\beta), \sigma^2)]$$

where the function $g : D \to \mathbb{R}$, $D = \{(x,y,z) \in \mathbb{R}^3 : x \neq y\}$, is given by

$$g(x,y,z) = C(\alpha,\beta) \cdot \frac{z}{(x-y)^2}, \quad C(\alpha,\beta) = \left(\Phi^{-1}(\alpha) - \Phi^{-1}(1-\beta)\right)^2$$

for $(x,y,z) \in D$. The function g is differentiable with

$$\dot{g}(x,y,z) = C(\alpha,\beta)(-2z(x-y)^{-3}, 2z(x-y)^{-3}, (x-y)^{-2}),$$

such that

$$\begin{aligned}\mathbf{g} &= \dot{g}(F^{-1}(p_\alpha), F^{-1}(p_\beta), \sigma^2) \\ &= \frac{C(\alpha,\beta)}{(F^{-1}(p_\alpha) - F^{-1}(p_\beta))^3} \cdot (-2\sigma^2, 2\sigma^2, (F^{-1}(p_\alpha) - F^{-1}(p_\beta))).\end{aligned}$$

The delta method now implies that

$$\sqrt{m}(n_m - n_\infty) \xrightarrow{d} \mathbf{g}U \sim N(\mathbf{0}, \eta^2),$$

as $m \to \infty$, where

$$\eta^2 = \mathbf{g}\Sigma\mathbf{g}'.$$

This completes the proof. □ □

REFERENCES

[1] Herrmann, W., Althaus, J., Steland, A. and Zaehle, H.: Statistical and experimental methods for assessing the power output specification of PV modules. Proceedings of the 21st European Photovoltaic Solar Energy Conference, 2416-2420 (2006)

[2] Bahadur, R. R.: A note on quantiles in large samples. Ann. Math. Statist. **37**, 577–580 (1966)

[3] Herrmann, W., Steland, A. and Herff, W.: Sampling procedures for the validation of PV module output specification. Proceedings of the 24th European Photovoltaic Solar Energy Conference, Hamburg, Germany, ISBN 3-936338-25-6, 3540-3547, DOI: 10.4229/24thEUPVSEC2009-4AV.3.70 (2010)

[4] Herrmann, W. and Steland, A.: Evaluation of photovoltaic modules based on sampling inspection using smoothed empirical quantiles. Progress in Photovoltaics **18**(1), 1–9 (2010)

[5] Hyndman R. J. and Fan Y.: Sample quantiles in statistical packages. The American Statistician **50**(4), 361–364 (1996)

[6] Kiefer, J.: On Bahadurs representation of sample quantiles. Ann. Math. Statist. **38**, 1323-1342 (1967)

[7] Kuurne, J., Tolvanen A., Hyvärinen J. and Oy, E.: Sweep time, spectral mismatch and ligh soaking in thin film module measurements. PVSC '08. 33rd IEEE, 1-3, ISSN 0160-8371 (2008)

[8] Meisen, S.: Ein nichtparametrisches Verfahren zur Annahmeprüfung. Diploma thesis, Institute of Statistics, RWTH Aachen University (2010)

[9] Ruberto M.N. and Rothwarf, J.: Time-dependent open-circuit voltage in $CuInSe_2$/CdS solar cells: Theory and experiment. J. Appl. Phys. **61**(9), 4662–4669 (1987)

[10] Roy, J.N., Gariki, G.R. and Nagalakhsmi, V.: *Reference module selection criteria for accurate testing of photovoltaic (PV) panels*. Solar Energy **84**, 32–36 (2010)

[11] Schilling, D.G. and Neubauer, D.V.: *Acceptance Sampling in Quality Control*, Chapman & Hall/CRC, Boca Raton (2009)

[12] Steland, A., Padmanabhan, A.R. and Akram, M.: *Resampling methods for the nonparametric and generalized Behrens-Fisher problems.* Sankhya A, tentatively accepted.

[13] Steland, A. and Zaehle, H.: Sampling inspection by variables: Nonparametric setting. Statist. Neerlandica **63**(1), 101–123 (2009)

[14] Virtuani, A., Muellejans, H., Ponti, F. and Dunlop, E.: Comparison of indoor and outdoor performance measurements of recent commercially available solar modules. Proceedings of the 23th European Photovoltaic Solar Energy Conference, Hamburg, Germany, ISBN 3-936338-25-6, DOI: 10.4229/24thEUPVSEC2009-3CO.12.1, 2379 - 2385 (2010)

E-mail address: SMEISEN@UKAACHEN.DE

E-mail address: PEPELYSHEV@STOCHASTIK.RWTH-AACHEN.DE

E-mail address: STELAND@STOCHASTIK.RWTH-AACHEN.DE