

Tripartite classification and adaptive frequency classification sampling plans for monitoring
population density through time

J. P. NYROP

Cornell University, NYSAES, Geneva, NY 14456 USA

& W. VAN DER WERF

Wageningen Agricultural University, P.O. Box 430, 6700 AK Wageningen, the Netherlands

SUMMARY

Two methods have been developed for monitoring population density through time. One is based on cascading sampling plans that sequentially classify density into one of three categories (tripartite sequential classification, TSC). The other protocol is constructed by cascading sampling plans based on a combination of sequential classification and estimation of density (adaptive frequency classification, AFC). When these two types of protocols were compared, protocols based on AFC sampling plans required fewer sampling resources than protocols based on TSC plans. Furthermore, protocols based on AFC plans scheduled intervention at appropriate times as well as protocols based on TSC plans.

INTRODUCTION

Classical sampling and decision making in IPM deals with the question of whether or not a pest population at a given point in time requires control. The frequency with which sampling events (bouts) are scheduled is often empirically chosen and is not part of the theoretical background of the decision scheme. This is not of consequence if only one sample at a critical time is sufficient to reach a management decision. However, there is often no such single critical point in time and the pest population must be monitored through time. For example, spider mites (Acari, Tetranychidae) infesting apples must be monitored during the complete growing season to ensure that densities do not exceed levels that cause injury.

We use the term "monitoring protocol" to denote one or more sampling plans used to assess population density through time. Sampling costs for a monitoring protocol are dependent on the number of sampling bouts as well as the number of samples taken at each sampling bout. Design of a monitoring protocol must consider the costs of executing the protocol in terms of sampling effort, gains due to avoiding unnecessary pesticide applications or to timely intervention in situations that would have resulted in pest induced injury, and losses due to unnecessary pesticide

applications or to tardy intervention when densities exceed thresholds.

We have developed two schemes that can be used to monitor populations through time. Both of these procedures schedule sample bouts more or less frequently depending on the population dynamics process. The first of these monitoring protocols is based on tripartite sequential classification (TSC) of density (Nyrop et al., 1994). The tripartite classification sampling plans are constructed by overlaying two sequential probability ratio tests (SPRT) (Wald, 1947). The first test is constructed around a density (intervention threshold) which, if exceeded, indicates the need for management intervention. The second test is constructed around a density which, if allowed to grow at an expected population growth rate for the minimum time until the next sample bout, would not exceed the intervention threshold. The sequential procedures are truncated so that if a maximum sample size is reached, the total count is compared to the midpoints of the upper and lower stop limits and classifications are made accordingly.

Each time a population is sampled, density is classified into one of three categories: 1) low density indicating that damaging pest levels are unlikely to occur in the near future and hence resampling can be delayed 2 δ time periods where δ is the minimum waiting time before sampling again; 2) intermediate density showing that densities are not currently at a damaging level, but the population should be sampled again soon, δ time periods in the future; and 3) high density prompting immediate intervention. If the first or second decisions are reached, the population is sampled again using either the same or a different tripartite classification protocol. One or more TSC procedures are thus cascaded through time to monitor a population trajectory.

The second scheme that can be used to monitor density through time is based on sampling plans that we have labeled “adaptive frequency classification” (AFC) (Nyrop and van der Werf, 1993). Wald's SPRT is used to construct a classification stop line for determining whether density exceeds an intervention threshold. The mean for the alternate hypothesis of the SPRT is set equal to the intervention threshold. The mean for the null hypothesis of the SPRT is set equal to a density which, if allowed to grow at an expected growth rate for the minimum time until the next sample bout (δ), would not exceed the intervention threshold appropriate for the future sampling time. A decision is made to not intervene if cumulative sample counts do not exceed the upper stop limit of the SPRT and if the total count when a predetermined maximum sample size is reached is less than the lower stop limit. If this occurs, the time to the next sample bout is based on an upper confidence limit for an estimated density obtained via the fixed sample size. While a waiting time to the next sample bout could be determined for each possible estimated density, this is not practical. Instead, we specify waiting times as multiples of the minimum resample interval (δ) and then determine ranges of the total count that correspond to each waiting time.

A decision is made to wait the shortest time to the next sample bout if the total sample count when the maximum number of samples (n_{max}) have been examined is less than the lower stop limit for the SPRT but greater than or equal to $td_2 \cdot n_{max}$. Here, td_2 is an estimated density whose

upper confidence limit, if allowed to grow at the expected rate of increase, would not exceed the intervention threshold appropriate for the time $t + 2 \cdot \text{delta}$. In general, a decision is made to wait $i \cdot \text{delta}$ time periods before resampling if the total count is $< td_i \cdot n_{max}$ and $\geq td_{i+1} \cdot n_{max}$. Note that AFC sampling plans take into account that future intervention thresholds may be different from current values whereas TSC sampling plans use only current threshold values.

If the total count exceeds the lower stop limit for the SPRT at n_{max} , a decision to intervene is made. Therefore, a horizontal line can be drawn from the lower stop limit at n_{max} to the upper stop limit and this line now defines part of the “intervene decision” stop limit. Stop lines for TSC and AFC sampling plans are illustrated in Fig. 1. These sampling plans were constructed using parameters for monitoring fruit tree red spider mite (*Panonychus ulmi*) as described later in the paper.

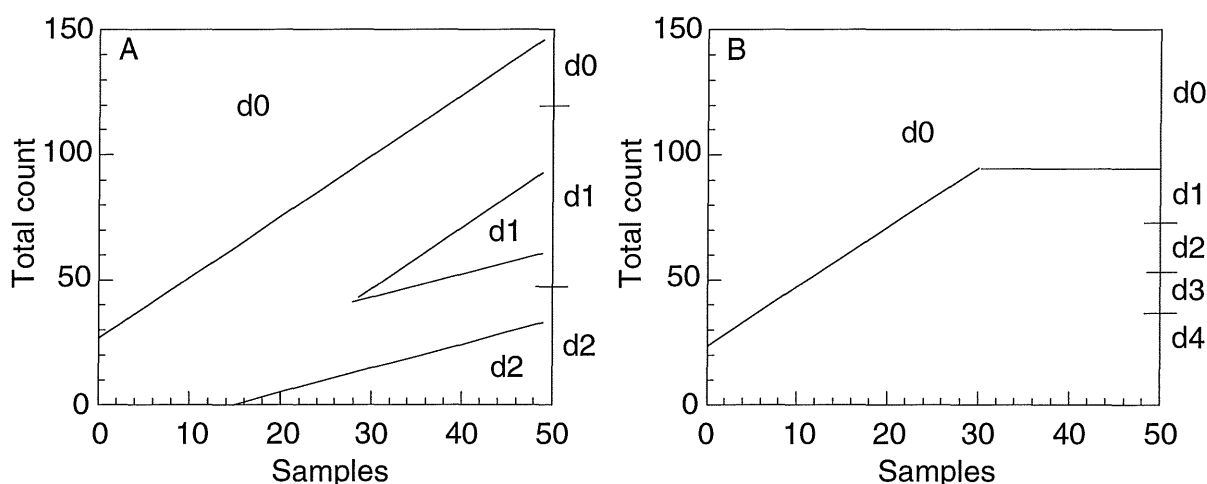


Fig. 1. Stop lines for a tripartite sequential classification (TSC) sampling plan (A) and an adaptive frequency classification (AFC) sampling plan (B). For the TSC plan, decisions d0, d1, and d2 correspond to intervening and resampling after 7 and 14 days. For the AFC plan, decisions d0 through d4 correspond to intervening and resampling after 7, 14, 21, and 28 days.

The performance of monitoring protocols based on TSC and AFC sampling plans can be evaluated using at least five criteria: 1) The probability of intervening (pi_p), 2) the expected number of occasions (bouts) on which samples were taken (eb_p), 3) the expected total sample size taken in all bouts (ess_p), 4) the expected density at which intervention occurs and an α percentile for this density, and 5) the expected area beneath the population trajectory (el_p) and an α percentile for this cumulative density. The α percentile indicates the value of density at intervention or cumulative density that $\alpha \cdot 100$ percent of possible densities or cumulative densities are less than.

This is a useful measure for risk to a grower. The subscript p for each performance criterion refers to a population trajectory because specific values of each performance criterion pertains to a specific population dynamic. Mathematical descriptions of these performance criteria are provided by Nyrop et al. (1994).

To evaluate the performance of a monitoring protocol the performance of each constituent sampling plan in the protocol must first be determined. Performance criteria for TSC and AFC plans consist of probability of decision functions ($pdeci$) and average sample size (asn) functions. For TSC sampling plans there are three $pdeci$ functions corresponding to; 1) waiting the longest period to resample, 2) waiting the shorter time period to resample, and 3) intervening. For AFC plans the $pdeci$ functions are similar except that there may be more corresponding to the increased number of δ time periods sampling can be delayed. Probability of decision and average sample size functions are constructed using simulation. Computer programs are available from the authors that perform these calculations as well as those for assessing the performance of the monitoring protocols.

To date we have not made a detailed comparison of monitoring protocols based on TSC and AFC sampling plans. The purpose of this paper is to present such a comparison when the two protocols are used to monitor two sets of hypothetical population trajectories. We include in this comparison the effect of different maximum sample sizes on monitoring scheme performance.

MATERIALS AND METHODS

The monitoring protocols compared were designed for monitoring fruit tree red spider mite infesting apple trees in New York, USA. Three intervention thresholds have been established for this purpose; 2.5 motile mites (i.e. anything but eggs) per leaf for the period 1 June to 30 June, 5.0 motiles per leaf for 1 July to 31 July, and 7.5 motiles per leaf thereafter. These thresholds are expert judgements based on the need to prevent a cumulative density (mitedays) of more than 500-600 from occurring during this period (Nyrop et al., 1989), while also minimizing the risk of letting dense populations of more than approximately 15 mites per leaf establish, as such dense populations are sometimes difficult to control. Counts of fruit tree red spider mite on apple leaves can be described by negative binomial distributions where the parameter k increases with mean density (Nyrop and Binns, 1992). This dependency was modeled by describing the variance as a function of the mean using Taylor's variance-mean model ($s^2 = am^b$ with $a = 4.32$ and $b = 1.42$) (Taylor, 1961) and then calculating k via moments: $k = m^2/(s^2-m)$.

Monitoring protocols based on TSC sampling plans were constructed using three TSC plans each with one SPRT based on an intervention threshold (cd_2) and the other SPRT based on a density (cd_1) that was a factor 2.5 less than the intervention threshold. The factor 2.5 results from calculating cd_1 as the density that would result in the intervention threshold density after 14 days,

assuming the population grew exponentially with a growth rate of 0.065 d^{-1} ($cd_2 = cd_1 \cdot e^{0.065 \cdot 14}$). The growth rate of 0.065 d^{-1} was an average determined by fitting an exponential model to 14 data sets on *P. ulmi* dynamics (Nyrop unpubl. data). The resample interval if the mean was classified as $\leq cd_1$ being 14 days, we accordingly set the minimum time interval to the next sample (*delta*) to 7 days. Three monitoring protocols were constructed with constituent sampling plans having maximum sample sizes (*nmax*) of 25, 50 and 100. Alpha and beta of the SPRTs were set equal to 0.20, 0.15, and 0.075 when *nmax* was 25, 50, and 100 respectively. Values of *k* for use in constructing SPRT stop limits were calculated via Taylor's variance-mean model by setting the mean to the midpoint of cd_1 and cd_2 . Means for the null and alternate hypotheses for the SPRTs were equidistant from cd_1 and cd_2 and were adjusted so that maximum *asn* values for each SPRT were approximately equal to the maximum sample size (Table 1).

Table 1. Means for null and alternate hypotheses for SPRTs used to construct tripartite classification sampling plans.

Plan 1		Plan 2				Plan 3					
First SPRT		Second SPRT		First SPRT		Second SPRT		First SPRT		Second SPRT	
H_0	H_1	H_0	H_1	H_0	H_1	H_0	H_1	H_0	H_1	H_0	H_1
maximum sample size = 25, $\alpha = \beta = 0.2$											
0.7	1.3	1.9	3.1	1.5	2.5	4.1	5.9	2.35	3.65	6.3	8.7
maximum sample size = 50, $\alpha = \beta = 0.15$											
0.75	1.25	2.0	3.0	1.6	2.4	4.2	5.8	2.45	3.55	6.45	8.55
maximum sample size = 100, $\alpha = \beta = 0.075$											
0.75	1.25	2.0	3.0	1.6	2.4	4.2	5.8	2.45	3.55	6.45	8.55

Monitoring protocols based on AFC sampling plans allowed sampling to be delayed a maximum of 28 days ($4 \cdot \text{delta}$, $\text{delta} = 7$ days). Three monitoring protocols were constructed with maximum sample sizes for each constituent sampling plan of 25, 50 and 100. The intervention thresholds used in the AFC plans were linear interpolations of a time dependent function defined by the means for the alternate hypotheses built about cd_2 for the TSC plans. These thresholds were used in place of the actual intervention thresholds (2.5, 5.0, and 7.5) for two reasons. First, to make the AFC and TSC plans more similar for comparison purposes. Second, because prior work had shown that monitoring protocols based on AFC plans too readily scheduled intervention when the alternative hypotheses for the SPRTs was set equal to the nominal intervention thresholds. Alpha and beta for the SPRTs used in the AFC plans were identical to those used in the TSC plans.

Values of k for use in constructing SPRT stop lines were determined using the midpoint of the means for the null and alternate hypotheses. Confidence intervals used to determine how long to wait before sampling again were computed using $\alpha = 0.25$.

Probability of decision and average sample size functions were estimated for each sampling plan by simulating sampling 1000 times for a range of mean densities. The parameter k of the negative binomial distribution was computed using Taylor's variance - mean model.

Performance of the monitoring schemes was studied by applying the monitoring protocols to two sets of hypothetical population dynamics for a period of 98 days. To cover this time period, three TSC sampling plans were required in a monitoring protocol and ten AFC sampling plans were required (plans 10-15 were identical). The first set of populations consisted of 15 trajectories described by logistic growth which were meant to broadly represent mite population growth under a range of conditions and where the shape of the population trajectories were similar. The shape restriction was required so that we could use cumulative density as an index for each population dynamic. All populations had initial densities of 0.1, the growth rate r ranged from 0.03 to 0.14 per day, and the maximum density was set to 60 motiles per leaf. When density approached the maximum it did not decline as would happen in the real world due to decline in plant quality. We purposely excluded this realism because we wanted cumulative density of populations that reached the maximum density earliest to be greater than the cumulative density of populations that reached the maximum density later. This was because high density populations early in the growing season are most damaging and we wanted cumulative density to reflect this.

The second set of populations consisted of a trajectory that exactly followed the nominal intervention threshold values of 2.5, 5.0, and 7.5 and a set of population trajectories that were 10, 20, 30, 40, and 50 percent less than these values. These dynamics were used to determine how the monitoring protocols behaved when population levels were close to levels requiring control but the intervention thresholds had not been exceeded. This situation could arise when *P. ulmi* population growth was being constrained by natural enemies.

RESULTS AND DISCUSSION

When used to monitor the populations with logistic growth, all protocols with constituent TSC or AFC sampling plans scheduled intervention before densities exceeded approximately 15 mites per leaf and before cumulative density exceeded 500 (Fig. 2). Monitoring protocols based on AFC sampling plans scheduled intervention at slightly lower densities compared with protocols based on TSC sampling plans. This happened for two reasons. First, the sequential test in the AFC sampling plans used to determine whether density was in excess of an intervention threshold used means for the null hypotheses that were less than similar tests in the TSC sampling plans (means for the alternate hypotheses were identical in both sampling plans). Second, when a terminal decision was required with an AFC plan, a decision was made to intervene if the total count

exceed the lower stop limit of the SPRT. In the TSC sampling plans, a decision was made to intervene if the cumulative count exceeded the midpoint of the null and alternate hypotheses of the SPRT.

The protocols based on AFC sampling plans required fewer sample bouts but more samples at each bout to monitor the populations with logistic growth compared with the protocols with constituent TSC sampling plans. Our intuitive sense is that travelling to a field to sample is more costly than taking additional samples once at a field site. Therefore, the monitoring protocol with constituent AFC sampling plans is probably more parsimonious than the protocol based on TSC sampling plans.

The maximum sample size had little effect on performance of either of the two protocols when used to monitor the populations with logistic growth.

When applied to the populations that followed the nominal intervention thresholds or were proportionally less than these thresholds, monitoring protocols based on TSC sampling plans and on AFC sampling plans scheduled intervention too frequently (Fig. 3). For example, when the maximum sample size was 50 and a population that was 30% less than the intervention thresholds was monitored, the probability of intervention with the TSC based protocol was 0.42 and for the AFC based protocol this probability was 0.87. The reason both types of monitoring protocols schedule intervention too frequently is because a decision to not intervene in one sample bout leaves open the possibility that a decision to intervene will be made in the following sample bout. For example, suppose a sampling plan can only lead to two decisions, intervene or sample again, each with probability 0.5. Then in three sample bouts the overall probability of intervening is 0.875.

The monitoring protocol based on TSC sampling plans generally required more sample bouts but approximately the same samples per bout compared with monitoring protocols based on AFC sampling plans. Thus, the protocol with constituent AFC plans produced lower sampling costs when used to monitor populations with slow growth and low to moderate densities.

Maximum sample size had a greater influence on protocol performance when monitoring the low density populations compared with monitoring the populations with logistic growth. As maximum sample size increased, intervention was scheduled less frequently. However, the change when sample size was increased from 50 to 100 is probably not enough to justify the added sampling costs.

Previous work (Nyrop et al., 1994) has shown that monitoring protocols based on TSC sampling plans better schedule intervention when necessary compared with monitoring protocols in which populations are sampled every second week and are more parsimonious than monitoring protocols in which sampling occurs every week. For the populations described in this paper, monitoring protocols based on AFC sampling plans performed better than those based on TSC sampling plans. The only discrepancy in this conclusion is that AFC based monitoring protocols scheduled intervention somewhat too frequently. This deficiency can be overcome by increasing the

mean densities used for the alternate hypothesis of the SPRT. We constructed a monitoring protocol using AFC sampling plans identical to the ones presented here with a maximum sample size of 50, but the intervention thresholds were increased by 20% (3.6, 6.96, 10.26). When used to monitor the populations with logistic growth, performance was nearly identical to the protocol that used lower thresholds with the exception that the 95th percentile for density at intervention exceeded 15 for two populations (17.21 and 18.0). When used to monitor the low density populations, the probability of intervention was greatly reduced (e.g. for the population that was 30% less than the threshold, p_i was reduced from 0.872 to 0.295.) Average sample sizes and bouts increased; however, these were approximately equal to the monitoring protocol based on TSC sampling that used the original thresholds (2.5, 5.0, 7.5).

In summary, monitoring protocols based on AFC sampling plans are superior to those based on TSC plans and provide an effective way to monitor population density through time.

REFERENCES

- NYROP, J. P., AGNELLO, A. M., KOVACH, J. & RIESSIG, W. H. (1989) Binomial sequential classification sampling plans for European red Mite (Acari: Tetranychidae) with special reference to their performance criteria. *Journal of Economic Entomology* **82**, 482-490.
- NYROP, J. P., BINNS, M. R., VAN DER WERF, W. & KOVACH, J. (1994) Cascading tripartite binomial classification plans to monitor European red mite (Acari, Tetranychidae) through a season; development and evaluation of a new methodology for pest monitoring. *Experimental and Applied Acarology* submitted.
- NYROP, J. P. & BINNS, M. R. (1992) Algorithms for computing operating characteristic and average sample number functions for sequential sampling plans based on binomial count models. *Journal of Economic Entomology* **85**, 1253-1273.
- NYROP, J. P. & VAN DER WERF, W. (1993) Sampling to predict or monitor biological control. in: *Handbook of Sampling Methods for Arthropod Pests in Agriculture*, pp 245-336. Eds. L. Pedigo and G. D. Buntin. CRC Press, Boca Raton.
- TAYLOR, L.R. (1961) Aggregation, variance, and the mean. *Nature (London)* **189**, 732-735.
- WALD, A (1947) *Sequential analysis*. Wiley, New York.

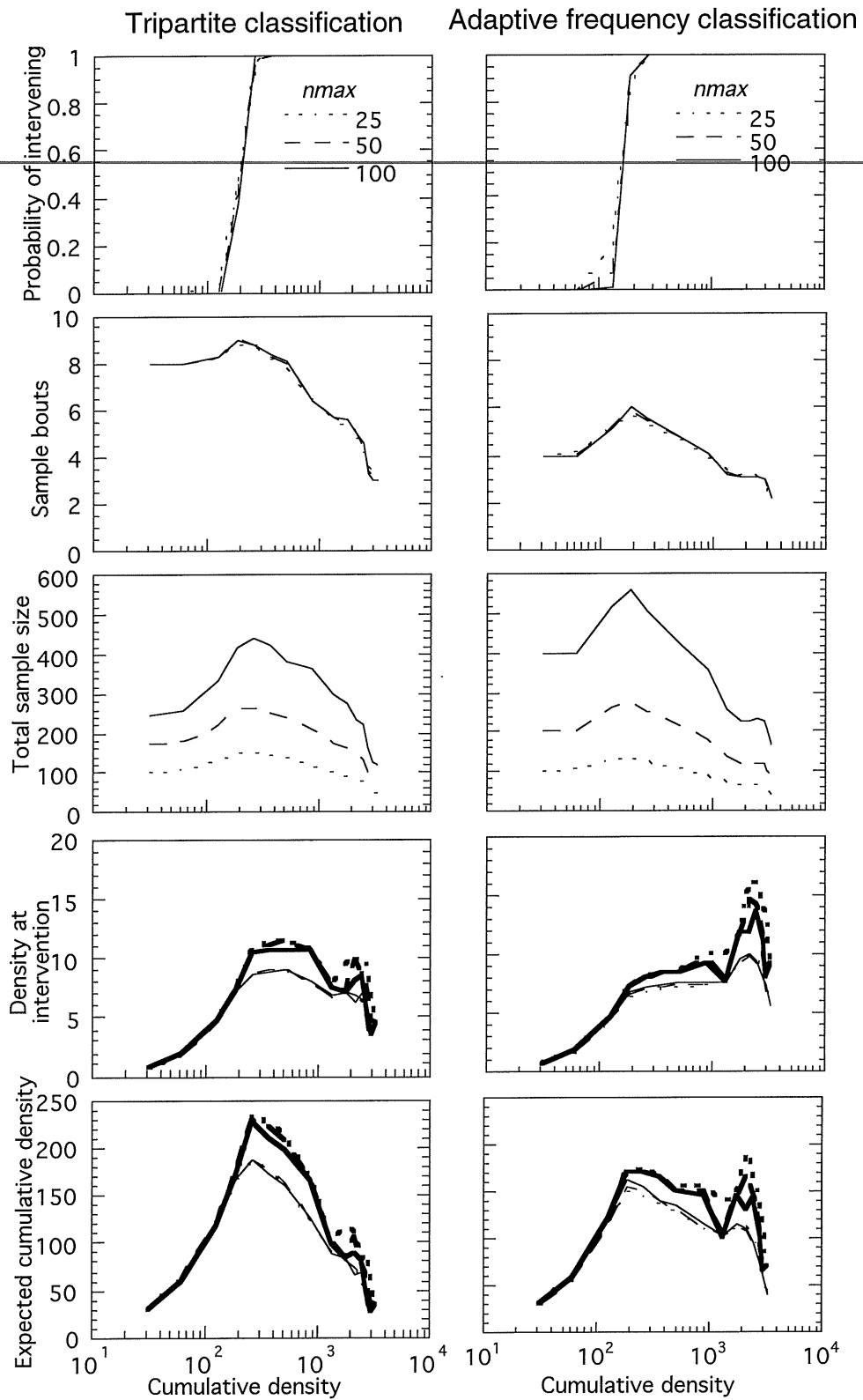


Fig. 2. Results of using protocols with constituent AFC and TSC sampling plans to monitor populations with logistic growth. Thin lines are averages, thick lines are 95 percentiles.

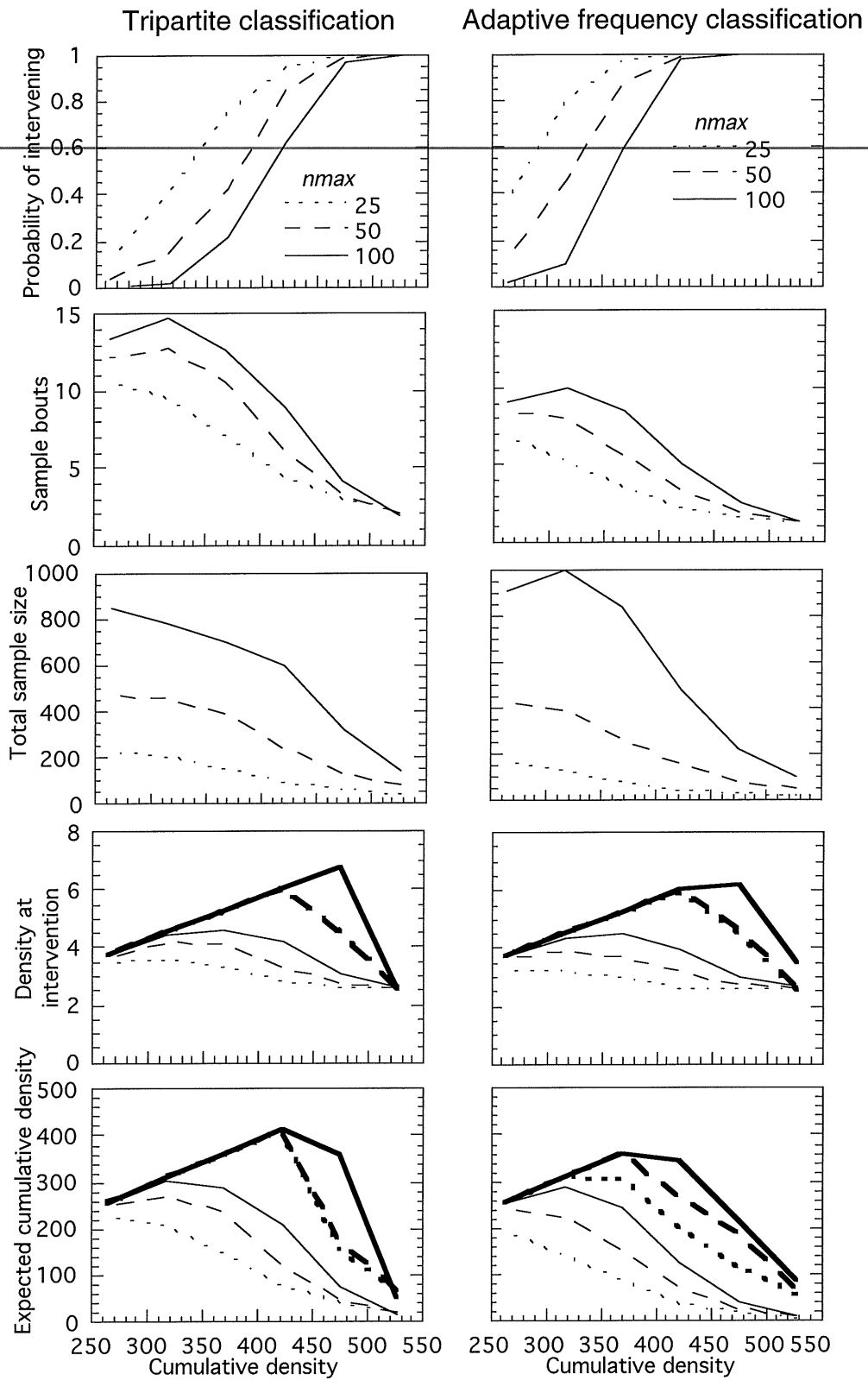


Fig. 3. Results of using protocols with constituent AFC and TSC sampling plans to monitor populations with slow growth and low to moderate densities close to the thresholds.