Quantifying taxonomic richness in terms of the level of rarity assessed by a fixed count

Kerry J. Ritter, David S. Birkes[1], and N. Scott Urquhart[2]

ABSTRACT - A common measure of taxonomic richness used in freshwater macroinvertebrate inventories is numerical taxonomic richness (NTR), i.e., the number of taxa observed for a fixed (usually small) number of individuals subsampled from a larger collection. However, rare taxa tend to be excluded from the metric for smaller subsamples due to their low inclusion probabilities. NTR, then, essentially targets the number of non-rare taxa, rather than the total number of taxa. The level of rarity assessed by the metric depends on the size of the subsample. By specifying the target parameter that NTR is actually estimating, researchers may understand more precisely what is being assessed and compared. Furthermore, determining a target parameter allows the consideration of alternatives for estimation that may be used to increase resolution or to reduce cost. Here, we provide a means for determining the parameter targeted by NTR. Specifically, we show that for a fixed count of size n, NTR is nearly unbiased for the number of taxa that occurs in the collection with relative frequency $\geq 1/2n$. Further, the jackknife adjustment to observed taxa count is shown to enhance the level of rarity assessed by NTR for the same-sized subsample.

INTRODUCTION

Macroinvertebrate taxonomic richness is a commonly used measure of habitat quality and environmental health in stream biomonitoring. Because complete enumeration of all macro-invertebrates for a given stream is not feasible, the total number of taxa must be estimated from a representative sample. Unfortunately, heterogeneity, clustering of individuals and large numbers of rare taxa pose serious challenges for estimation (Chazdon *et al.* 1998). Large diverse populations require

sampling across many sites in order to acquire adequate representation. In addition, large samples must be taken at each site in order to ensure the stability and accuracy of the estimates. The effort and cost associated with acquiring a suitable sample and evaluating the collection can be prohibitive in estimating taxonomic richness for the population.

One popular approach for reducing cost is fixed-count subsampling. Rather than attempt to estimate the total number of taxa for a particular station, fixed-count procedures employ numerical taxonomic richness (NTR), or the number of taxa observed for a fixed count, as a measure of taxonomic richness. NTR has been used extensively in large-scale biomonitoring efforts to monitor trends and measure degrees of impairment for aquatic ecosystems. Rapid bioassessment protocols (RBPs) developed for the U.S. Environmental Protection Agency (U.S. EPA) and adopted by numerous large-scale biomonitoring agencies have incorporated NTR into their monitoring and assessment programs (Resh and Jackson 1993). The RBP III, for example, recommends a fixed count of 100 to 300 individuals to detect impact (Plafkin *et al.* 1989). When the number of individuals counted and the area sampled are standardized, NTR then provides an index of stream health that may be compared among multiple collections (Larsen and Herlihy 1998). Because the counting and identifying of individuals collected constitutes the major cost in evaluating macroinvertebrate inventories (Resh and Price 1984), NTR allows a greater number of sites to be monitored at considerable reduction in cost.

Although fixed-count subsampling may be practical for satisfying economic constraints, its ability to provide an accurate representation of the collection has been questioned. Because few rare taxa are likely to be included in smaller subsamples, there is concern that differences in taxonomic richness may be greatly underestimated. Some have argued that

[1]*Oregon State University, Department of Statistics, Corvallis, OR 97331*

[2]*Colorado State University, Department of Statistics, Fort Collins, CO 80523*

disturbances that reduce overall taxonomic richness are likely to manifest themselves first and most substantially with the removal of rare taxa (Gaston 1998). Taxonomic richness measures that rely only on the more abundant taxa, then, may not provide strong enough signals to detect small changes in environmental conditions and therefore may affect our ability to apply remedy. Because one of the primary challenges of biomonitoring is to detect disturbance, the level of rarity assessed by the metric is particularly important.

While the importance of rare taxa in biomonitoring has been the subject of debate in the literature, not all types and magnitudes of change require the same resolution in terms of the metric's ability to detect differences in taxa composition. Severe or large-scale impacts may be indicated by reductions in the abundance of the more common taxa. Disturbances such as mild sedimentation do not pose as serious a health risk to the larger community as in the case of heavy metal toxins, and therefore the need for immediate detection is not as important. Barbour and Gerritsen (1996) defended the fixed-count subsampling methods and have found them to be effective in practice for discriminating among large macroinvertebrate collections reflecting high, medium, and low diversity. Plafkin *et al.* (1989) cite other researchers who have supported 100-count subsampling as sufficient for detecting impact (Nuzzo 1986, Bode 1988, Shackleford 1988).

As with most estimators, the ability of NTR to discriminate among collections depends on subsample size. As the subsample size increases, so does the resolution of NTR to detect differences in taxa composition and hence changes in the biological community. The likelihood that a rare taxon will be observed in the subsample is greater for larger subsamples. Because the observed number of taxa increases with subsample size, it can be argued that a given subsample size allows a reasonable chance of detection only for those taxa that occur above a certain level of rarity. Resolution for NTR based on a particular fixed-count protocol, then, may be quantified in terms of the level of rarity assessed by the metric.

By determining the level of rarity likely to be observed for a particular fixed count, a new parameter may be defined which more closely reflects what is being estimated by NTR. Specifically, the number of "non-rare" taxa is introduced as the parameter effectively targeted by NTR where the level of rarity

is determined by the size of the subsample. With this parameter, practitioners may gain better insight into the ability of NTR to detect certain types and magnitudes of change. The size of the subsample then may be weighed against the level of rarity required to detect change. Further, defining the target parameter allows us to consider alternatives for estimation. Efforts can be directed toward developing estimators which may be used to increase resolution for detecting impact, to reduce cost, or to compare measures based on different-sized subsamples. Therefore, determining a target parameter for NTR offers several advantages for biomonitors.

This article quantifies the degree of rarity assessed by NTR as a function of subsample size. A new parameter, $C^*(\xi)$, is defined as the number of taxa that occur in a collection with relative frequency $\geq \xi$. An approximation is provided based on subsample size for the value of x such that NTR is nearly unbiased for $C^*(\xi)$. As an alternative measure, the jackknife adjustment to the observed number of taxa is examined for improving the level of rarity assessed by NTR. Performance with respect to estimating $C^*(\xi)$ in terms of accuracy (bias), stability (standard deviation, SD), and reliability (root mean square error, RMSE) are investigated for each of the two candidates, as well as recommended subsample sizes for minimizing bias. Results are based on simulation studies in which estimates were calculated from subsamples from large composites of macroinvertebrates from seven streams in Oregon.

RESOLUTION OF NUMERICAL TAXONOMIC RICHNESS IN TERMS OF LEVEL RARITY

This section offers a method for specifying the parameter targeted by NTR in terms of the number of taxa that occur in a collection with relative frequency greater than or equal to some cutoff value, ξ. An approximation for this target parameter, $C^*(\xi)$, is given as a function of subsample size and is shown to be robust to varying distributions of taxa sizes. Biases for NTR, both with respect to $C^*(\xi)$ and the total number of taxa, C, are investigated for several subsample sizes to illustrate the effectiveness of this approximation with regard to bias reduction. These results are based on simulations of simple random subsamplings of individuals from large composite samples of macroinvertebrates collected in 1992 and 1993 by personnel in the Oregon State University (OSU) Department of Fisheries and Wildlife, as part

of a cooperative agreement between OSU and EPA (CR824682). Composites were composed of 8 to 45 **Surber samples (0.09m²**, 500-μm mesh size) collected in seven streams in the Willamette Valley and Cascade region in Oregon (Li *et al.* 2001).

Concern about NTR often stems from the fact that NTR severely underestimates the total number of taxa and that the bias depends both on subsample size and the relative abundance of taxa in the assemblage (Walther and Morand 1998, Colwell and Coddington 1994, Courtemanch 1996). This is especially problematic when comparing NTR among heterogeneous collections or with respect to a reference value. Bias with respect to C will be larger for collections having more rare taxa than for those with fewer rare taxa. Comparisons based on NTR, then, may impose unfair penalties on those streams dominated by a few taxa and having large numbers of rare taxa (Cao *et al.* 1998). Fair comparisons require that estimates be approximately unbiased for the same parameter. Therefore, it is crucial to determine the parameter being targeted by NTR in order to understand what is being estimated and compared.

Initial observations based on small composite samples of macroinvertebrates collected by OSU/ EPA as part of a pilot study revealed that NTR based on a subsample of size n tended to estimate $C^*(1/2n)$, or the number of taxa in the collection having relative frequency greater than $1/2n$. We decided to test the consistency of this observation by conducting a simulation study of NTR for estimating $C^*(1/2n)$ in additional composite samples.

Simulation Study 1:

Investigation of NTR as a function of sub sample size focused on seven large composite samples collected by OSU/EPA as part of a 1992 and 1993 pilot study. From the 51 composites collected, these seven were chosen to reflect a range of values for coefficient of variation of taxa sizes and of percent of rarity for several levels of rarity (see Tables 1 and 2). All individuals in each composite were counted and identified so that the true distribution of taxa frequencies was known.

Biases for NTR were considered with respect to both $C^*(\xi)$ and the total number of taxa, C, for each of the seven collections. Subsamples of sizes 50, 100, 250, and 500 were investigated and represent a range of plausible fixed-count protocols used in rapid assessment inventories. Estimates for bias were based on simulations of 10,000 subsamples where individuals within each subsample were selected randomly without replacement and subsamples were taken randomly with replacement. The relative biases (bias divided by the parameter) are provided in Figure 1 and measure average deviation of the NTR estimates from their targets, C and $C^*(\xi)$, in terms of the percentage of the parameter value.

As expected, NTR produced large underestimates for C. Averaging NTR estimates across simulations, we found that NTR tends to underestimate C by more than 20% of the parameter value for all stream composites and subsample sizes. The median across streams for the absolute value of the relative bias ranged between 27% and 66% of the target parameter, depending on subsample size. For $n = 50$, underestimation exceeded 57% for all seven streams. For each composite, the magnitude of bias was increased with smaller subsample sizes. Also, there is evidence to support the conclusion that magnitudes of relative bias increased as the proportion of rare taxa (as defined relative to $= \xi \ 1/2n$) increased. Tests of no association between the proportion of rare taxa and relative magnitude of bias using Spearman's rank correlation coefficient yielded approximate p-values .0062, .0014, .0014, and .0240 for fixed-counts of 50, 100, 250, and 500, respectively.

In contrast, the magnitude of bias for NTR with respect to $C^*(1/2n)$ were smaller than that for estimating C. The median across streams for the absolute value of the relative bias ranged from around 6% to 10% of the target parameter, depending on subsample size. Among the 28 cases (7 streams ξ 4 subsample sizes), in only two cases did magnitudes of relative bias for estimating the proposed parameter exceed 15%. These exceptions were composites which had high proportions of rare taxa (> 70%), and estimates were based on the smallest subsamples considered ($n = 50$). These cases also had low total numbers of "non-rare" taxa and magnitudes of relative bias are inflated by small values of the parameter. Taking a relative bias of less than 10% as the criterion for "nearly unbiased," we found that NTR was nearly unbiased for $C^*(1/2n)$ in 4 out of 7 streams for $n = 50$, 6 out of 7 streams for $n = 100$, all 7 streams for $n = 250$, and 6 out of 7 for $n = 500$.

From our simulation studies, we conclude that $C^*(1/2n)$ provides a more accurate representation of the parameter effectively targeted by NTR than does C. We summarize our observations with the following statements.

Table 1. Taxa sizes, number of individuals, CV, and percent of rare taxa.

Stream	C	N	CV	Percent of Rare Taxa $\xi=.001$	$\xi=.002$	$\xi=.005$	$\xi=.01$
ORC0693	57	5,060	442.3	40.4	56.1	68.4	75.4
ORV0292	31	3,088	320.6	32.3	41.9	54.9	77.4
ORV2593	25	11,013	176.2	20.0	32.0	48.0	50.0
ORC0992	66	2,917	163.6	24.2	39.4	48.5	60.6
ORC0292	59	5,289	258.2	20.3	33.9	55.9	69.5
ORV2493	26	5,293	240.8	7.7	30.8	50.0	73.0
ORV0692	36	5,841	164.2	22.2	29.8	39.0	56.1

C = Number of taxa, N = Number of individuals, CV = Coefficient of variation.

Table 2. Non-rare taxa sizes.

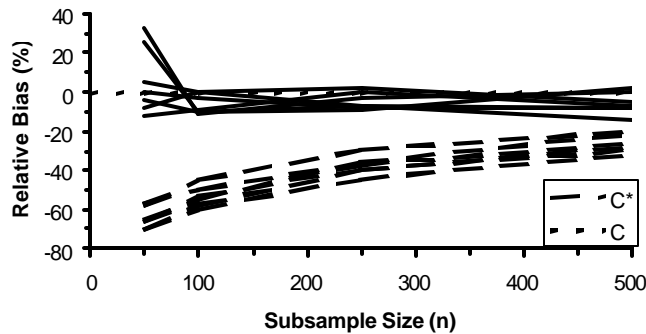| Stream	$C^*(1/2\,n)$ n = 50	n = 100	n = 250	n = 500	$C^*(1/5n)$ n = 50	n = 100	n = 250	n = 500
ORC0693	14	18	25	34	18	25	34	41
ORV0292	7	14	18	21	15	18	23	25
ORV2593	10	13	17	20	14	17	20	24
ORC0992	26	34	40	50	34	40	50	54
ORC0292	18	26	39	47	28	39	49	52
ORV2493	7	13	18	24	14	18	24	26
ORV0692	16	22	28	28	24	28	29	33

Figure 1. Relative bias of observed numerical taxonomic richness (NTR) as an estimate of C (long dashes) and NRT* (solid lines), as a function of subsample size, over 10,000 simulations; curves represent results using data from different Oregon streams. (Short dashed line locates bias = 0.)

Finding 1.

Numerical taxonomic richness based on a subsample of size *n* is generally nearly unbiased for the number of taxa which occur in the collection with relative frequency $\geq 1/2n$.

The converse of Finding 1 may be used to determine a subsample size that provides near unbiasedness for NTR when the target parameter is specified. This is given below.

Finding 1a.

For estimating the number of taxa in a large collection with relative frequency $\geq \xi$, a nearly unbiased estimate is given by NTR based on a subsample of approximate size $1/2\xi$.

JACKNIFE ADJUSTMENT FOR INCREASING THE LEVEL OF RARITY ASSESSED FOR A FIXED COUNT

Disagreements concerning the fixed-count procedure among researchers often focus on whether or not the evaluation of collections based on small subsample sizes can provide the necessary information for detecting important changes in the physical and biological community. In particular, is the level of rarity assessed by NTR for a given subsample size sufficient to monitor important differences in taxa composition for large macroinvertebrate collections? This section reports on the jackknife adjustment to the observed taxa count as a way to enhance the level of rarity assessed for a particular fixed count. We provide a value of *x*, as a function of subsample size, such that the jackknife is nearly unbiased for $C^*(\boldsymbol{x})$.

The jackknife procedure is a commonly used technique to reduce the bias of a given estimator. Introduced by Quenouille (1949), jackknifing adjusts

the bias of an estimator by adding a correction term based on estimates from subsamples with observations deleted. Following the description given by Schucany and Gray (1971) for the jackknife procedure, let $c^{(n)}$ be the total number of taxa observed in a sample of size n and let $c_{(i)}^{(n-1)}$ be the total number of taxa observed in the collection with the ith observation removed. Then $J_i = nc^{(n)} - (n-1)c_{(i)}^{(n-1)}$ is called the ith pseudo-value for the jackknife. Averaging these pseudo-values over all possible point deletions gives the jackknife estimator. Burnham and Overton (1979) gave a closed form for the jackknife estimator. Although intended to estimate population size through mark-recapture studies, the estimator can be used to estimate the total number of taxa in an assemblage. Within this context, the jackknife estimator for C is given by

$$\text{Jack} = c^{(n)} + \frac{n-1}{n}c_1$$

where n is the size of the sample and is the number of singletons (i.e., number of taxa represented by exactly one individual) in the subsample.

By adding a positive correction factor, the jackknife attempts to correct for the negative bias of $c^{(n)}$ by adjusting for those taxa not represented in the subsample. While many have found the jackknife procedure to be successful in reducing bias, there is a trade-off in that the added correction factor introduces extra variability.

The jackknife has less bias than NTR for estimating C, but the jackknife still tends to underestimate C to some extent. So it is not unbiased for C, but just as NTR has been found to be nearly unbiased for $C^*(1/2n)$, it seems plausible that the jackknife may be nearly unbiased for estimating $C^*(\xi)$ for some value of ξ less than $1/2n$ (recall Finding 1). In other words, the jackknife adjustment may improve the level of rarity assessed by NTR. A simulation study was conducted to examine the relationship between the bias with respect to $C^*(\xi)$ and subsample size for the jackknife estimator. Initial examination focused on three composite samples of macroinvertebrates collected from three different streams in Oregon. We observed the following.

Finding 2.

The jackknife adjustment for the observed taxa count based on a subsample of size n is generally nearly unbiased for the number of taxa in the collection with relative frequency $\geq 1/5n$.

Simulation Study 2:

Consistency of Finding 2 across a variety of taxa distributions was evaluated via simulation of jackknife estimates from subsampling the same 7 composites used to evaluate Finding 1 for NTR. One of these 7 composites, ORCO693, was also one of the three used in our initial investigation. Bias, SD, and RMSE of the jackknife for estimating $C^*(1/5n)$ were approximated for subsample sizes $n = 50, 100, 250,$ and 500 in order to assess potential trade-offs in accuracy and stability. According to Finding 2, the jackknife estimator based on these subsample sizes should target the number of taxa occurring in the collection with relative frequency [3] x where x =.004, .002, .0008, and .0004, respectively. Performances (with respect to bias, SD, and RMSE) of the jackknife for estimating $C^*(1/5n)$ were also compared with those for NTR for estimating $C^*(1/2n)$ based on the same fixed counts. Results are based on simulations of 10,000 subsamples where individuals within each subsample were selected randomly without replacement and subsamples were taken randomly with replacement from each of the 7 composites. Summaries are given on the relative scale in Figures 2a –2c.

To measure the trade-off between accuracy and stability, the RMSE was estimated for the two estimators. In most cases, NTR resulted in lower RMSE than did the jackknife, although differences in the relative RMSEs were generally less than 10% (percentage points) in all but one case. Depending on the subsample size, medians across streams of the relative RMSE for the jackknife ranged from around 12 to 20%, while those for NTR ranged from 9 to 15%.

As expected, the larger RMSE for the jackknife was primarily due to its larger variability, although differences in relative SDs between the two estimators were less than 10% (percentage points) in all cases. Depending on subsample size, the medians across streams of the relative SDs for the jackknife ranged from around 12 to 19%, while those for NTR ranged from around 7 to 14%.

In most cases, the jackknife estimator for $C^*(1/5n)$ resulted in smaller magnitudes of relative bias than did NTR for $C^*(1/2n)$. However, differences in magnitudes between the two estimators were generally smaller than 5% (percentage points) and exceeded 13% in only 2 out of 28 cases. The medians across streams for the absolute value of the relative bias for the jackknife for estimating $C^*(1/5n)$ ranged from around 2 to 6% of $C^*(1/5n)$ while those for
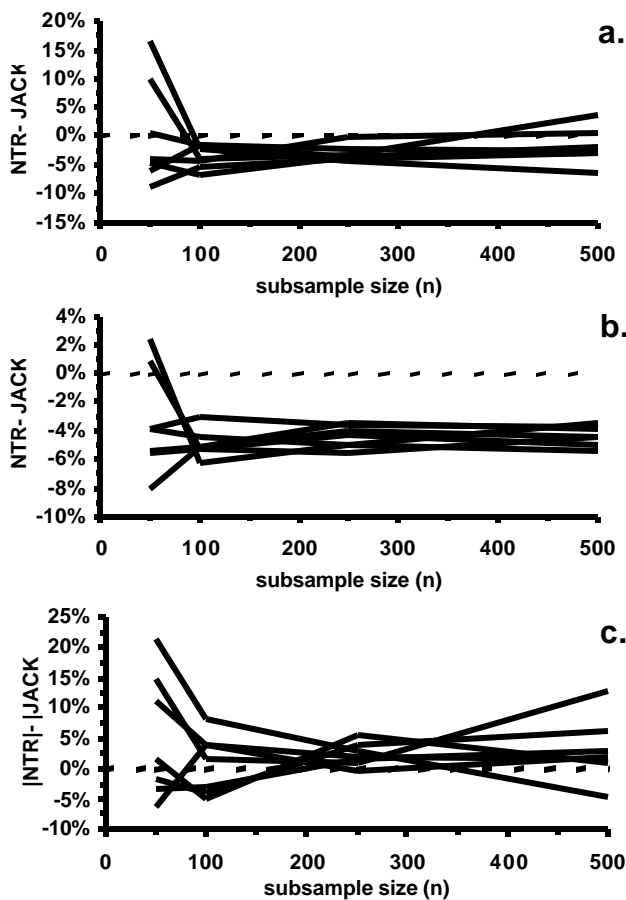
Figure 2. Difference comparisons of the performance of the Jackknife Estimator (JACK) for estimating C*(1/5n) to numerical taxonomic richness (NTR) for estimating C*(1/2n), as a function of subsample size, over 10,000 simulations; curves represent results using data from different Oregon streams: (a.) Relative mean square error, (b.) Relative standard deviation, and (c.) Relative bias. (Short dashed line locates the difference = 0.)

NTR ranged from around 6 to 13% of $C^*(1/2n)$, depending on the subsample size. In fact, in only 2 cases out of 28 did the magnitude of relative bias for the jackknife exceed 10%, supporting our initial finding that the jackknife nearly unbiasedly estimates $C^*(1/5n)$. These exceptions were composites with relatively high proportions of rare taxa ($C^*(1/5n) = 14$ and 15) and estimates were based on the smallest subsample size n = 50.

The Jackknife as a Cost-Reducing Alternative

In addition to assessing a higher level of rarity for fixed counts, the jackknife method may also be useful for comparing with NTR when smaller counts are used or to reduce cost. For example, suppose in one

year NTR based on a fixed count protocol of 100 individuals was used. Suppose, however, the following year budget constraints required that smaller subsamples be used. According to Finding 1 in the previous section, NTR based on a subsample of size $n = 100$ is nearly unbiased for $C^*(.005)$ ($\xi = 1/(2*100)$). However, for $n \ll 100$, NTR is no longer unbiased for the same parameter. Therefore, an alternative is needed for a fair comparison with the previous year.

As with NTR, the converse of Finding 2 can be used to specify the subsample size that offers nearly unbiased estimation of $C^*(\xi)$.

Finding 2a.

The jackknife adjustment for observed taxa count based on a subsample of $n \cong 1/5\xi$ is generally nearly unbiased for the number of taxa in the collection with relative frequency $\geq \xi$.

Simulation Study 3:

To investigate the effect of reducing the size of the subsample and applying the jackknife method for estimating $C^*(\xi)$, we compared the bias, standard deviation, and RMSE for the jackknife at $n = 1/5\xi$ (i.e. Jack(1/5ξ)) with those for NTR at $n = 1/2\xi$ (i.e., NTR(1/2ξ)). Estimates were based on simulation studies similar to the ones above only subsample sizes for each of the estimators were chosen according to Findings 1a and 2a, where values for x were chosen to be .001, .002, .005, and .01. For NTR, these values for x corresponded to subsample sizes $n = 500, 250, 100,$ and 50, respectively. For the jackknife, corresponding subsamples sizes were $n = 200, 100, 40,$ and 20, respectively. Bias, SD, and RMSE for NTR and jackknife were approximated with respect to $C^*(\xi)$, based on simulations of 10,000 random subsamplings. Results are given in Figures 3a-3c.

The RMSE for the jackknife at $n = 1/5\xi$ tended to be about 1.5 times that of NTR based on subsample sizes $n = 1/2\xi$ and average differences ranged between 5% and 10% (percentage points) depending on ξ. Similarly, the SDs for Jack(1/5ξ) were between 1.5 and 2.5 times those of NTR (1/2ξ) in all 28 cases. Average absolute differences in SDs were between 6% and 11% (percentage points) depending on x. In terms of bias, the two methods performed similarly at their respective subsample sizes. Absolute differences between the magnitudes of relative bias for Jack(1/5ξ) and NTR(1/2ξ) were considered for each collection.
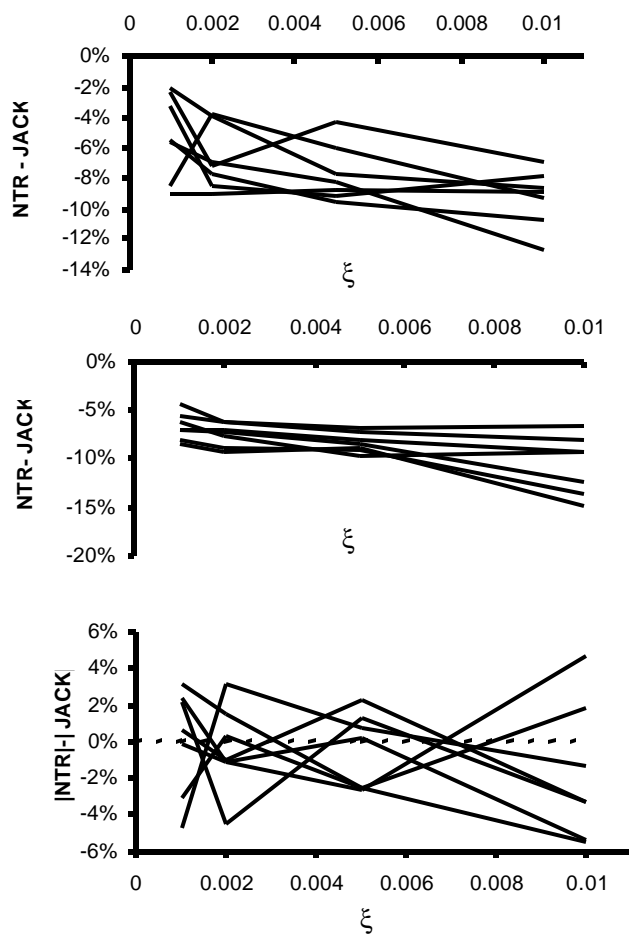
Figure 3. Difference comparisons of the performance of the Jackknife Estimator (Jack (1/ 5ξ)) to numerical taxonomic richness (NTR*(1/ 2ξ)), as a function of the level of rarity, ξ, over 10,000 simulations; curves represent results using data from different Oregon streams: (a.) Relative mean square error, (b.) Relative standard deviation, and (c.) Magnitudes of relative bias. (Short dashed line locates the difference = 0.)

These differences averaged across the 7 stream composites were around 2% to 4% (percentage points), depending on ξ. Whether or not the benefits of subsample size reduction outweigh the decrease in stability depends on the priorities and objectives of the study. However, the jackknife offers a cost-reducing alternative for NTR for estimating *C*.

DISCUSSION

This study examined the level of rarity assessed by numerical taxonomic richness (NTR) as a function of subsample size. While most researchers are well aware that NTR underestimates the total number of taxa in the collection and that the bias depends on subsample size, the extent is not well understood

quantitatively. With this study, a quantitative approach is offered for estimating the portion of the biotic assemblage, in terms of the level of rarity, that is accurately estimated by a count of "*n*". Determining the level of rarity targeted by NTR has important ecological and environmental implications for large-scale biomonitoring programs.

Our investigation has shown that NTR based on a fixed count of size *n* is in general nearly unbiased for the number of taxa in a collection with relative frequency $\geq 1/2n$. The degree of rarity assessed for a particular collection, however, may be increased with the jackknife correction to the observed taxa count. Our simulation studies on 7 large composites suggest that the jackknife estimator based on a subsample of size *n* effectively targets the number of taxa in a collection with relative frequency $^{3} 1/5n$. There is a trade-off with the jackknife in that the bias adjustment term decreases the estimator's stability. If resolution for detecting change increases with our ability to monitor the more rare taxa, as some have argued, then the jackknife method may be a more effective measure of taxa richness than NTR.

From an ecological standpoint, assessing the degree of rarity targeted by the metric provides insight into how changes to the environment influence the distribution of individuals in the assemblage. Our study has allowed the potential for linking type and/or degree of disturbance with degree of rarity required for impact detection. These results bring us closer to identifying and understanding those disturbances which result in the depletion of only rare taxa and which seem to impact the community as a whole. Defining the target parameter of NTR, then, may allow for discriminating among possible candidates to determine disturbances that result in meaningful changes to the community structure. Further, identifying disturbances may provide insight into what is driving the distribution of taxa in that environment.

Economically, specifying the target parameter of NTR allows consideration for more cost-efficient surveys. The benefits of increased resolution can be weighed against the costs of increased subsampling. Resources can be redistributed to increase either sampling intensity or subsample size. By understanding the limitations of the metric, we may further redirect our efforts toward more feasible research alternatives. Finally, we can make more informed decisions about the viability and appropriateness of employing alternatives to NTR, such as the jackknife, in order to increase the degree of rarity assessed or to reduce costs.

Future research should focus on alternative taxonomic richness measures that reduce costs and increase resolution for detecting change. It may be possible to construct a measure that is nearly unbiased for a given arbitrary level of rarity with a given arbitrary subsample size. Such advancements could provide methods for measuring smaller changes in the biological community and allow increased flexibility in evaluation protocol design. Perhaps one of the most important applications for long-term or large-scale surveys is adjusting estimates so as to provide a means for fair comparisons among different-sized subsamples. While we have provided a jackknife correction that allows for comparison with NTR for a specific subsample size, comparisons among different surveys require correction for an arbitrary subsample size. Finally, future efforts are necessary to determine whether the findings reported here generalize to other distributions and types of organisms beyond those examined in this study.

LITERATURE CITED

Barbour, M.T. and J. Gerritsen. 1996. Subsampling of benthic samples: A defense of the fixed-count method. *Journal of American Benthological Society* 15: 386-391.

Bode, R.W. 1988. Quality assurance work plan for biological stream monitoring in New York State. New York State Department of Environmental Conservation. New York, NY.

Burnham, K.P. and W.S. Overton. 1979. Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* 65: 625-633.

Cao, Y., D. Williams and N.E. Williams. 1998. How important are rare species in aquatic community ecology and bioassessment? *Limnology and Oceanography* 43: 1403-1409.

Chazdon, R.L., R.K. Colwell, J.S. Denslow and M.R. Guariguata. 1998. Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of northeastern Costa Rica. pp. 285-309 *in:* F. Dallmeier and J.A. Comisky (eds.), Forest Biodiversity Research, Monitoring, and Modeling (Vol. 20). Parthenon Publishing Group. New York, NY.

Colwell, R.K. and J.A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. pp.101-118 *in:* Philosophical Transactions of the Royal Society London. Series B, 345. London, England.

Courtemanch, D.L. 1996. Commentary on the subsampling procedures used for rapid bioassessments. *Journal of the North American Benthological Society* 15: 381-385.

Gaston, K.J. 1998. Ecology: Rarity as double jeapardy. *Nature* 944: 229-230.

Larsen, D.P. and A.T. Herlihy. 1998. The dilemma of sampling streams for macroinvertebrate richness. *Journal of the North American Benthological Society* 17: 359-366.

Li, J., A.T. Herlihy, W. Gerth, P.R. Kaufmann, S.V. Gregory, N.S. Urquhart and D.P. Larsen. 2001. Quantifying variance in stream macroinvertebrates at multiple spatial scales. *Freshwater Biology* 46: 87–97.

Nuzzo, R. 1986. Macroinvertebrate rapid assessment methodology, Mass. Viv. WPC.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers. Benthic Macroinvertebrates and Fish. EPA/440/4-89/001. U.S. Environmental Protection Agency, Office of Water Regulations and Standards. Washington, DC.

Quenouille, M.H. 1949. Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society, Series. B.* 11: 68-84.

Resh, V.H. and J.K. Jackson. 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. pp. 195-233 *in:* D.M. Rosenberg and V.H. Resh (eds.), Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman and Hall. New York, NY.

Resh, V.H. and D.G. Price. 1984. Sequential sampling: A cost-effective approach for monitoring benthic macroinvertebrates in environmental impact assessments. *Environmental Management* 8: 75-80.

Schucany, W.R., H.L. Gray and D.B. Owen. 1971. On bias reduction in estimation. *Journal of American Statistical Association* 335: 524-533.

Shackleford, B. 1988. Rapid bioassessment of lotic macroinvertebrate communities: Biocriteria development. Arkansas Department of Pollution Control and Ecology. Little Rock, AR.

Walther, B.A. and S. Morand. 1998. Comparative performance of species richness estimation. *Parisitology* 116: 395-405.

ACKNOWLEDGEMENTS