

# Considerations for monitoring population trends of colonial waterbirds using the effective number of breeders and census estimates

Fagner M. da Silva<sup>1</sup>  | Carolina I. Miño<sup>2</sup>  | Rafael Izbicki<sup>3</sup> | Silvia N. Del Lama<sup>1</sup> 

<sup>1</sup>Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, São Paulo, Brazil

<sup>2</sup>Instituto de Biología Subtropical (IBS), Universidad Nacional de Misiones, CONICET, Posadas, Misiones, Argentina

<sup>3</sup>Departamento de Estatística, Universidade Federal de São Carlos, São Carlos, São Paulo, Brazil

## Correspondence

Fagner M. da Silva, Laboratório de Genética de Aves, Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, São Paulo, Brazil.  
Email: fagner.miguel.silva@gmail.com

## Funding information

Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Grant/Award Number: 2010/52315-7 and 2016/01673-7

## Abstract

Detecting trends in population size fluctuations is a major focus in ecology, evolution, and conservation biology. Populations of colonial waterbirds have been monitored using demographic approaches to determine annual census size ( $N_a$ ). We propose the addition of genetic estimates of the effective number of breeders ( $N_b$ ) as indirect measures of the risk of loss of genetic diversity to improve the evaluation of demographics and increase the accuracy of trend estimates in breeding colonies. Here, we investigated which methods of the estimation of  $N_b$  are more precise under conditions of moderate genetic diversity, limited sample sizes and few microsatellite loci, as often occurs with natural populations. We used the wood stork as a model species and we offered a workflow that researchers can follow for monitoring bird breeding colonies. Our approach started with simulations using five estimators of  $N_b$  and the theoretical results were validated with empirical data collected from breeding colonies settled in the Brazilian Pantanal wetland. In parallel, we estimated census size using a corrected method based on counting active nests. Both in simulations and in natural populations, the approximate Bayesian computation (ABC) and sibship assignment (SA) methods yielded more precise estimates than the linkage disequilibrium, heterozygosity excess, and molecular coancestry methods. In particular, the ABC method performed best with few loci and small sample sizes, while the other estimators required larger sample sizes and at least 13 loci to not underestimate  $N_b$ . Moreover, according to our  $N_b/N_a$  estimates (values were often  $\leq 0.1$ ), the wood stork colonies evaluated could be facing the loss of genetic diversity. We demonstrate that the combination of genetic and census estimates is a useful approach for monitoring natural breeding bird populations. This methodology has been recommended for populations of rare species or with a known history of population decline to support conservation efforts.

## KEYWORDS

conservation genetics, effective size, microsatellites, *Mycteria americana*, simulations, single-sample estimators

## 1 | INTRODUCTION

Assessing the size of natural populations is a major focus of population monitoring programs aimed at estimating population trends and determining priority species or areas for conservation (Lindenmayer & Likens, 2010). Changes in the census size of natural populations either over time ( $N_t$ ) or annually ( $N_a$ ) can indicate susceptibility to stochastic processes (Lande, 1988). However, the responsiveness of populations to evolutionary forces depends not only on population size, but also on genetic factors (Frankham, 1996). Therefore, measures of effective population size ( $N_e$ ) and effective number of breeders ( $N_b$ ) can be more informative than the census size itself (Trail, Brook, Frankham, & Bradshaw, 2010).  $N_e$  is the size of a Wright-Fisher ideal population (Fisher, 1930; Wright, 1931) affected by genetic drift at the same rate as a real population (Crow & Kimura, 1970).  $N_b$  is a measure of  $N_e$  for a single breeding season, reflecting the parental contribution and changes in the inbreeding rate (Waples, 1990, 2002).

In natural populations, however,  $N_e$  cannot be estimated accurately most of the time, which is a major limitation to applying such approaches for population monitoring. The genetic and demographic parameters that affect  $N_e$  are often unknown in natural populations and most key assumptions of the available estimators may be violated (Wang, 2016). In populations of age-structured species, estimates of  $N_b$  ( $\hat{N}_b$ ), which require data from a single cohort, are often more feasible to obtain than estimates of  $N_e$  ( $\hat{N}_e$ ) (Waples, 1990, 2002).  $\hat{N}_b$  has been proposed to assess population trends and can provide crucial information for wildlife management and conservation, such as levels of genetic variation and the effects of ecological factors on population size (Hinkson & Richter, 2016; Schwartz, Luikart, & Waples, 2007; Whiteley et al., 2015).

Single-sample methods using microsatellite information (for a review, see Luikart, Ryman, Tallmon, Schwartz, & Allendorf, 2010) have piqued the interest of researchers to estimate  $N_b$ . However, these estimators are affected differently by population processes (e.g., immigration and genetic substructure) and methodological aspects (e.g., genotyping errors and sample size) (Belmar-Lucero et al., 2012; Wang, 2016; Whiteley et al., 2012). Thus, evaluations of how these different factors influence  $\hat{N}_b$  can determine the conditions under which this parameter is more reliable with regard to detecting changes in the size of natural populations (Fraser et al., 2007; Menéndez, Álvarez, Fernandez, Menéndez-Arias, & Goyache, 2016). For example, the bias in  $\hat{N}_b$  resulting from immigration depends on whether it is incidental or recurrent and on the genetic differentiation between focal and source populations (Fraser et al., 2007; Gomez-Uchida, Palstra, Knight, & Ruzzante, 2013; Waples & England, 2011).

Many colonial waterbirds breed in sites where they depend on very specific environmental conditions and are subject to human-induced changes that can result in variations in population size (Tsai, Reichert, Frederick, & Meyer, 2016; Vásquez-Carrillo, Henry, Henkel, & Peery, 2013). Population sizes of colonial waterbirds

have been monitored to evaluate changes in breeding colonies in space and time and are indicators of regional biodiversity, restoration success, and wetland quality (Atkinson et al., 2006; Kushlan, 1992; Péron, Ferrand, Leray, & Gimenez, 2013; Tavares, Guadagnin, Moura, Siciliano, & Merico, 2015). It would be ideal to count all birds in a colony at a given time, but census techniques most often do not enable the detection of real variations in the number of individuals. Colonial waterbirds, in particular, are difficult to monitor using conventional census because breeding colonies can be composed of thousands of breeding pairs, with large numbers of nonbreeding individuals, cryptic nests, and asynchronous reproduction. These problems can be minimized using genetic methods to estimate  $N_b$  from representative samples of the colonies. Likewise, monitoring approaches combining genetic and census estimates can provide insights into the role of life history on  $N_b/N_a$  ratios and are recommended for the evaluation of the risk of stochasticity for population persistence (Belmar-Lucero et al., 2012; Palstra & Fraser, 2012; Trail et al., 2010). However, prior to applying combined estimates in monitoring programs, an understanding of the relationship between  $\hat{N}_b$  and  $\hat{N}_a$  and the limits of such an approach is required (Bernos, Yates, & Fraser, 2018; Ferchaud et al., 2016). In this context, it is crucial to evaluate how genetic methods behave when estimating  $N_b$  under the conditions affecting natural waterbird colonies, which has not previously been investigated, and how to obtain robust census estimates in breeding colonies.

In this study, we used the wood stork, *Mycteria americana* (Linnaeus 1758; Ciconiiformes: Ciconiidae), as a model to evaluate the competence of estimates of  $N_a$  ( $\hat{N}_a$ ) and  $\hat{N}_b$ , and, indirectly,  $\hat{N}_b/\hat{N}_a$  ratios for monitoring population trends. Characteristics of the wood stork make it a good model for these purposes: It is very sensitive to environmental changes in the wetlands where it inhabits and reproduces (Frederick, Gawlik, Ogden, Cook, & Lusk, 2009; Tsai et al., 2016); it exhibits colonial behaviour during the breeding season and is philopatric to its breeding sites, although to a lesser extent than are other species of Ciconiiformes (Frederick & Ogden, 1997); and it has a long pre fledgling period (Bryan, Snodgrass, Robinette, & Hopkins, 2005), which enables sampling distinct generations. The fact that breeding colonies are often established in the same locations over time enables the possible assessment of population trends and indirect inferences regarding changes in such areas.

Our primary aim was to provide baseline methodological guidelines for monitoring changes in population size of colonial waterbirds by assuming breeding colonies as attractive target units for conservation, using both census and genetic methods. Using simulations, we analyzed different empirical datasets to illustrate the common problems encountered when estimating  $N_b$  in natural populations, particularly regarding sample size and number of loci. We assumed that a suitable genetic estimator of  $N_b$  for monitoring changes in population size of colonial waterbirds should (a) yield estimates proportional to  $\hat{N}_a$ , (b) have the power to yield precise estimates from datasets based on a limited number of genetic markers, and (c)

require a sample size that can be collected during the short period of the breeding season at several sites. Furthermore, we apply our findings on the estimation of  $N_b$  to natural populations and we test the suitability of a counting method with an adjustment for the heterospecific composition of breeding colonies.

## 2 | MATERIALS AND METHODS

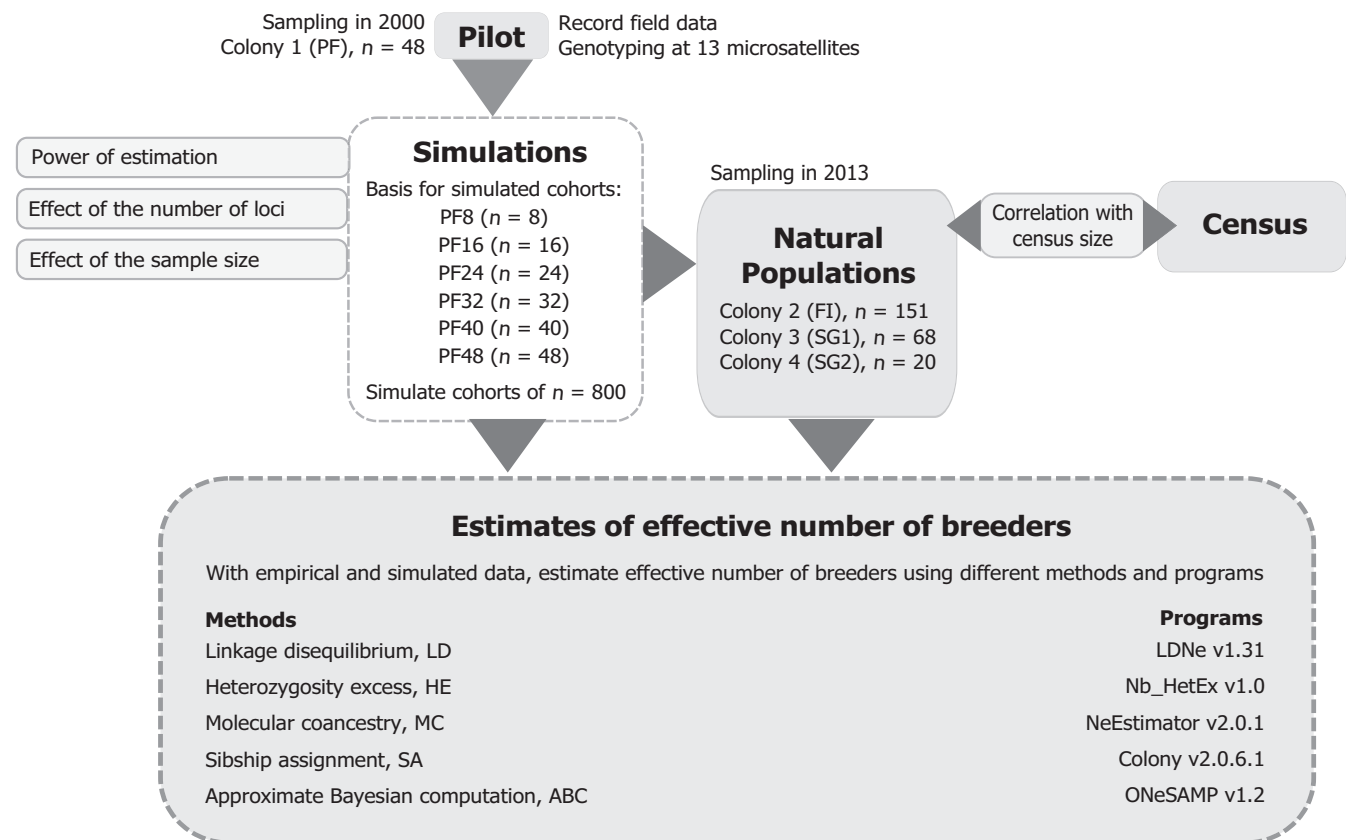
### 2.1 | Sample collection

Wood stork nestlings were randomly sampled at breeding colonies established in the Brazilian Pantanal wetland: colony Porto da Fazenda (PF, latitude  $-16.46873^\circ$ , longitude  $-56.1258^\circ$ ,  $n = 48$ ) was sampled in 2000, colony Fazenda Ipiranga (FI, latitude  $-16.42736^\circ$ , longitude  $-56.62183^\circ$ ,  $n = 151$ ), colony Sangradouro 1 (SG1, latitude  $-16.30992^\circ$ , longitude  $-57.04925^\circ$ ,  $n = 68$ ), and colony Sangradouro 2 (SG2; latitude  $-16.31898^\circ$ , longitude  $-57.04716^\circ$ ,  $n = 20$ ) were sampled in 2013 (Figure 1). The nests were accessed using ladders and climbing techniques. The nestlings were placed in bird bags, moved from the nests to a table on the ground where banding and blood collection took place, and returned safely to their respective nests. Blood ( $\sim 0.2$  ml) was collected

from the brachial vein, using syringes rinsed with anticoagulant (EDTA 0.3%). The number of nestlings from each accessed nest containing at least one nestling ( $k$ ) was recorded and adherence to Poisson's distribution ( $\alpha < 0.05$ ) was tested. The mean and the variance of the recorded values of  $k$  ( $\bar{k}$  and  $V_k$ , respectively) were calculated for each colony and used to determine the index of variability ( $V_k/\bar{k}$ ) (Crow & Morton, 1955).

### 2.2 | Genetic analyses

Genomic DNA was extracted using the DNeasy<sup>®</sup> Blood & Tissue kit (Qiagen). Samples were amplified at 13 species-specific microsatellite loci (Supporting Information Table S1) and genotyped in a MegaBACE<sup>®</sup>1000 sequencer (GE Healthcare). Genotypes were recorded using FRAGMENT PROFILER<sup>®</sup> v1.2 (GE Healthcare), checked for evidence of null alleles using MICRO-CHECKER v2.2.3 (van Oosterhout, Hutchinson, Wills, & Shipley, 2004), and checked for identical multilocus genotypes using COLONY v2.0.6.1 (Jones & Wang, 2010). Locus WS6 was reliably recorded only in PF and maintained in the simulation dataset, but had an average of 49.79% of missing data in other breeding colonies and was excluded from subsequent analyses. Identical multilocus genotypes were maintained in the datasets to test their effects on the estimation of  $N_b$  (Supporting Information Table S2).



**FIGURE 1** Flowchart outlining the methodological procedure followed. Original multilocus genotypes from Colony 1 and its subsets were used as basis to simulate large cohorts, in which predictions and effects of methodological issues in estimation of effective number of breeders ( $N_b$ ) were evaluated using different approaches. Estimates of  $N_b$  for natural populations (Colonies 2, 3, and 4) were obtained following guidelines determined using simulations, and effects of methodological issues were evaluated using census ( $N_a$ )

**TABLE 1** Genetic diversity indices for simulated cohorts and wood stork populations from Pantanal wetland. Mean genetic diversity indices were showed for wood stork population of Porto da Fazenda (PF) and subsets of different numbers of individuals genotyped (PF8, PF16, PF24, PF32 and PF40) as well as for the natural populations from Fazenda Ipiranga (FI), Sangradouro 1 (SG1) and Sangradouro 2 (SG2). Indices obtained for different numbers of individuals genotyped ( $n$ ): observed number of alleles ( $A_O$ ), allelic richness ( $A_R$ ), and expected heterozygosity ( $H_E$ )

Simulated genotypes												
Dataset	$n$	$A_O$			$A_R^a$			$H_E$				
		7 loci	10 loci	13 loci	7 loci	10 loci	13 loci	7 loci	10 loci	13 loci	13 loci <sup>b</sup>	
PF8	8	2.14	2.20	2.54	2.31	2.14	2.54	2.31	0.41	0.44	0.49	0.45
PF16	16	2.57	2.50	2.77	2.54	2.29	2.57	2.35	0.40	0.41	0.46	0.42
PF24	24	2.57	2.60	2.85	2.62	2.21	2.55	2.34	0.40	0.40	0.45	0.42
PF32	32	2.57	2.90	3.15	2.92	2.16	2.62	2.41	0.41	0.43	0.47	0.44
PF40	40	2.86	3.10	3.31	3.08	2.20	2.62	2.42	0.43	0.43	0.47	0.44
PF	48	3.00	3.20	3.31	3.15	2.29	2.60	2.39	0.42	0.43	0.43	0.44
Natural populations												
Pop. <sup>c</sup>	$n$	$A_O$			$A_R^a$			$H_E$				
FI	151			3.92			2.43				0.40	
SG1	68			3.25			2.45				0.43	
SG2	20			2.50			2.40				0.41	

<sup>a</sup>Estimates of  $A_R$  were obtained by rarefaction method in HP-RARE (Kalinowski, 2005), using smallest number of gene copies at single locus ( $n = 16$  for set of genotypes used for simulations and  $n = 22$  for natural populations). <sup>b</sup>Without adjusting allele frequencies for presence of null alleles. <sup>c</sup>Obtained for genotypic datasets at 12 microsatellite loci.

### 2.3 | Composition of datasets and common assumptions for estimation of $N_b$

The informative content of each locus was estimated as 1—probability of identity using GENALEX v6.5 (Peakall & Smouse, 2012). The mean number of alleles and mean expected heterozygosity was computed in ARLEQUIN v3.5.2.2 (Excoffier & Lischer, 2010), and mean allelic richness was computed in HP-RARE v1.0 (Kalinowski, 2005); these indices were compared among all datasets. Linkage disequilibrium between loci and departures from Hardy-Weinberg (HW) expectations were assessed in ARLEQUIN v3.5.2.2 (Excoffier & Lischer, 2010), adjusting  $p$ -values ( $\alpha = 0.05$ ) with the Holm-Bonferroni correction (Holm, 1979).

Compliance with common assumptions of estimators was evaluated in the datasets. Selective neutrality was verified in LOSITAN v1.0 (Antao, Lopes, Lopes, Beja-Pereira, & Luikart, 2008). Genetic differentiation within and among colonies was investigated by analysis of molecular variance (AMOVA) in ARLEQUIN v3.5.2.2 (Excoffier & Lischer, 2010). Population subdivision was further investigated using the Bayesian clustering algorithm in STRUCTURE v2.3.4 (Pritchard, Stephens, & Donnelly, 2000). The parameters for these analyses were as follows: sampling location as prior, an admixture model with correlated allele frequencies, five runs per  $K$  (1 to 4),  $2 \times 10^5$  runs of MCMC and  $10^4$  of burn-in. The Delta  $K$  evaluation (Evanno, Regnaut, & Goudet, 2005) was applied in STRUCTURE HARVESTER v0.6.94 (Earl & vonHoldt, 2012). Migration patterns were inferred using a Bayesian method (Rannala & Mountain, 1997) implemented in GENECLASS v2.0 (Piry et al., 2004) to compute the probabilities of multilocus genotypes in a given colony being descendants of immigrant ( $\alpha = 0.01$ , re-sampling procedure with 10,000 simulated individuals).

### 2.4 | Simulations to estimate $N_b$ and design of natural population study

To evaluate the variation in  $\hat{N}_b$  under different conditions and define the experimental design to properly sample colonies during the second sampling effort, cohorts were simulated using a method based on empirical data (Figure 1). The PF population was chosen as our base sample to perform the simulations because it is larger and more stable in number of individuals over the years than other studied wood stork populations (e.g., colonies FI, SG1 and SG2). Cohorts were simulated by resampling the genotypes from PF to compose five subdatasets with different allele frequencies; then, sets of 30 independent large cohorts ( $n = 800$ ) were simulated from each subdataset in KINGROUP v2.08 (Konovalov, Manning, & Henshaw, 2004), setting a 1:1 adult sex ratio. To better reflect the situation found in colonial waterbird populations, the simulation model was based directly on allele frequencies of the PF genotypic dataset and simulated genotypes were created by randomly drawing alleles. Given a set of population frequencies, alleles are drawn using the internal number generator from the computer and matching the random number to the frequencies; for example, given alleles A, B and C with frequencies of 0.4, 0.5, and 0.1, respectively, the simulation model chooses a random number ( $R$ ) from 0 to 1 and chooses an

allele as follows:  $0 \leq R \leq 0.4$ : choose A,  $0.4 \leq R \leq 0.9$ : choose B, and  $0.9 \leq R \leq 1$ : choose C (Dmitry A. Konovalov, personal communication). Therefore, the genotype frequencies in simulated cohorts are in concordance with HW equilibrium and with the provided allele frequencies (Dmitry A. Konovalov, personal communication). Frequencies of null alleles were adjusted using the Brookfield 1 method (Brookfield, 1996).

Further simulations were carried out to investigate how genetic estimators of  $N_b$  would behave under the properties of the genotypic datasets encountered in waterbird colonies (Figure 1). The simulations included a low number of loci, to account for the low presence of microsatellites in avian genomes (Ellegren, 2013; Primmer, Raudsepp, Chowdhary, Møller, & Ellegren, 1997), and low allele frequencies, to account for the intrinsically low levels of polymorphisms in microsatellite found in waterbirds (Campanini, Sanches, Hatanaka, & Del Lama, 2012; Nunes, Efe, Freitas, & Bugoni, 2017; Peters, Omland, & Johnson, 2007). The effect of a small increase in the number of loci was evaluated using sets of 7, 10, and 13 loci for each set of simulated cohorts: the 10-locus dataset was achieved by removing three randomly selected loci ( $WS_{\mu}$  03,  $WS_{\mu}$  09 and  $WS_{\mu}$  20); subsequently, another three randomly selected loci ( $WS_1$ ,  $WS_{\mu}$  14 and  $WS_{\mu}$  24) were removed to obtain the seven-locus dataset. To investigate the influence of sample size on  $\hat{N}_b$ , random samples of 5% ( $n = 40$ ), 10% ( $n = 80$ ), 15% ( $n = 120$ ), 20% ( $n = 160$ ) and 50% ( $n = 400$ ) were taken from the simulated cohorts derived from the original PF dataset.  $N_b$  was estimated for all simulations using five single-sample methods (Figure 1). The linkage disequilibrium (LD) method was employed in LDNE v1.31 (Hill, 1981; Waples, 2006; Waples & Do, 2008), assuming random mating and using minimum allele frequencies of 0.02, 0.05, and 0.10. The heterozygosity excess (HE) method was used in  $N_b\_HETEX$  v1.0 (Pudovkin, Zaykin, & Hedgecock, 1996; Zhdanova & Pudovkin, 2008) with 10,000 bootstrap iterations. For the approximate Bayesian computation (ABC) method in ONESAMP v1.2 (Tallmon, Koyuk, Luikart, & Beaumont, 2008), the prior range was set as 2-100 to save computational time (changing the upper prior limit to 500 did not significantly change the resulting estimates). The molecular coancestry (MC) method was run in NEESTIMATOR v2.0.1 (Do et al., 2014; Nomura, 2008). The sibship assignments (SA) method implemented in COLONY v2.0.6.1 (Jones & Wang, 2010; Wang, 2009) was used assuming monogamy and a medium sibship prior of 2.0 for each parent.

### 2.5 | Estimation of $N_a$ , $N_b$ , and $N_b/N_a$ from natural populations

An improved field technique based on counting the number of active nests and using  $\bar{k}$  values was used to estimate  $N_a$  and cohort size ( $N_t$ ) of colonies FI, SG1, and SG2. At breeding colonies, nests were classified as “with content confirmed visually from the ground” ( $N_{CG}$ ) or “with content not confirmed visually from the ground” ( $N_{NG}$ ).  $N_{CG}$  were classified as active (i.e., those containing eggs, nestlings or occupied by adults) or inactive. Active nests were classified as belonging to wood storks ( $N_{WD}$ ) or other species.

To estimate the number of nests with contents that could not be confirmed from the ground, but were indeed active wood stork nests, the status of nests was visually checked from tree tops in randomly selected sub-areas of the colonies. The frequency of active wood stork nests ( $F_{WD}$ ) within  $N_{NG}$  was estimated using the total number of nests with contents that could not be confirmed from the ground ( $N_{NGSA}$ ) and the number of active wood stork nests confirmed only when checking from tree tops ( $N_{WDSA}$ ) in the sub-areas (i.e.,  $F_{WD} = N_{WDSA}/N_{NGSA}$ ). Using this frequency, the total number of active nests per wood stork population ( $TN_{WD}$ ) was estimated as  $\widehat{TN}_{WD} = N_{WD} + (N_{NG}F_{WD})$ .  $\widehat{N}_a$  was estimated as  $2(\widehat{TN}_{WD})$  considering an independent pair of reproductive adults per active nest, and the number of nestlings in cohorts ( $\widehat{N}_1$ ) was estimated as  $\bar{k}(TN_{WD})$ . Based on Sahai and Khurshid (1995), the 95% confidence interval (CI) for  $\widehat{N}_a$  was estimated as:

$$95\%CI_{(\widehat{N}_a)} = 2N_{WD} + 2N_{NG}(2F_{WD}) \pm 1.96\sqrt{(4N_{NG}^2) \left(\frac{N_{NG}-N_{NGSA}}{N_{NG}-1}\right) (F_{WD}) \left(\frac{1-F_{WD}}{N_{NGSA}}\right)}$$

and the 95% CI for  $\widehat{N}_1$  was estimated as  $95\%CI_{(\widehat{N}_1)} = (95\%CI_{(\widehat{N}_a)}\bar{k})/2$ .  $N_b$  was estimated for cohorts sampled from the FI, SG1, and SG2 colonies as well as the pooled dataset of the three colonies using the genetic estimators described above in item 2.4, and  $\widehat{N}_b/\widehat{N}_a$  ratios were calculated. To estimate  $N_b$  using the SA method, empirical  $\bar{k}$  values were set as the medium sibship prior for each parent (J. Wang, personal communication).

## 2.6 | Evaluation of $N_b$ estimators

Although simulations under ideal conditions are more appropriate for assessing the accuracy of  $\widehat{N}_b$ , when the objective is to provide genetic monitoring guidelines for natural populations, one should prioritize the use of a precise estimator under conditions similar to those encountered in nature (i.e., nonideal conditions). Estimates with infinite values could compromise the monitoring of population sizes, making it impossible to compare estimates in time and space. Therefore, the applicability of the single-sample methods for monitoring changes in population size was initially evaluated based on the precision of the estimates measured by the mean percentage of

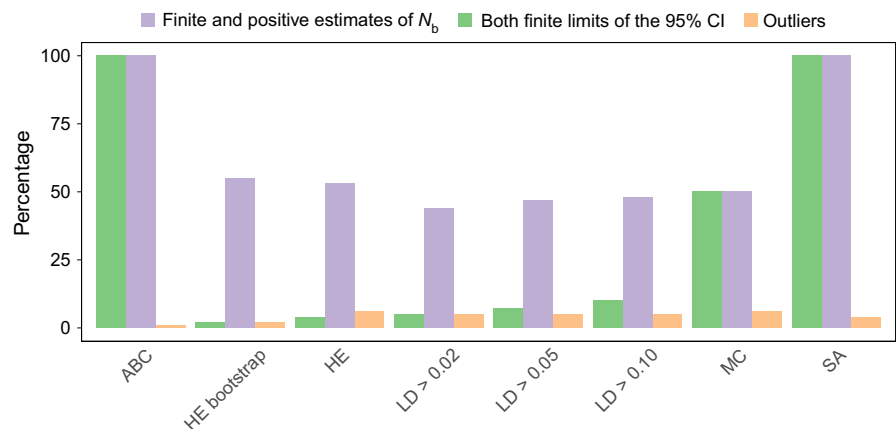
outliers (for simulations), the mean percentage of finite  $\widehat{N}_b$  and the mean percentage of  $\widehat{N}_b$  with narrower 95% CIs.

The minimum sample size needed for each method was defined as the smallest sample size yielding  $\widehat{N}_b$  that differed significantly with an error rate  $\leq 10\%$  from those obtained by sampling all individuals from the simulated cohort. The average error rate of the estimates was determined by computing the percentage of the differences between harmonic means of  $\widehat{N}_b$  obtained using different sample sizes, and the entire set of simulated cohorts and the standard errors were computed using 10,000 bootstrap iterations. For real cohorts, correlations between  $\widehat{N}_b$  versus  $\widehat{N}_a$  and sample sizes ( $n$ ) were assessed. The precision of  $\widehat{N}_b$  was equated to its log root mean squared error ( $\log RMSE$ ) for simulated cohorts and its variance ( $V$ ) for real cohorts (Beebe, 2009). All statistical analyses were performed in R v3.3.1 (R Development Core Team, 2016).

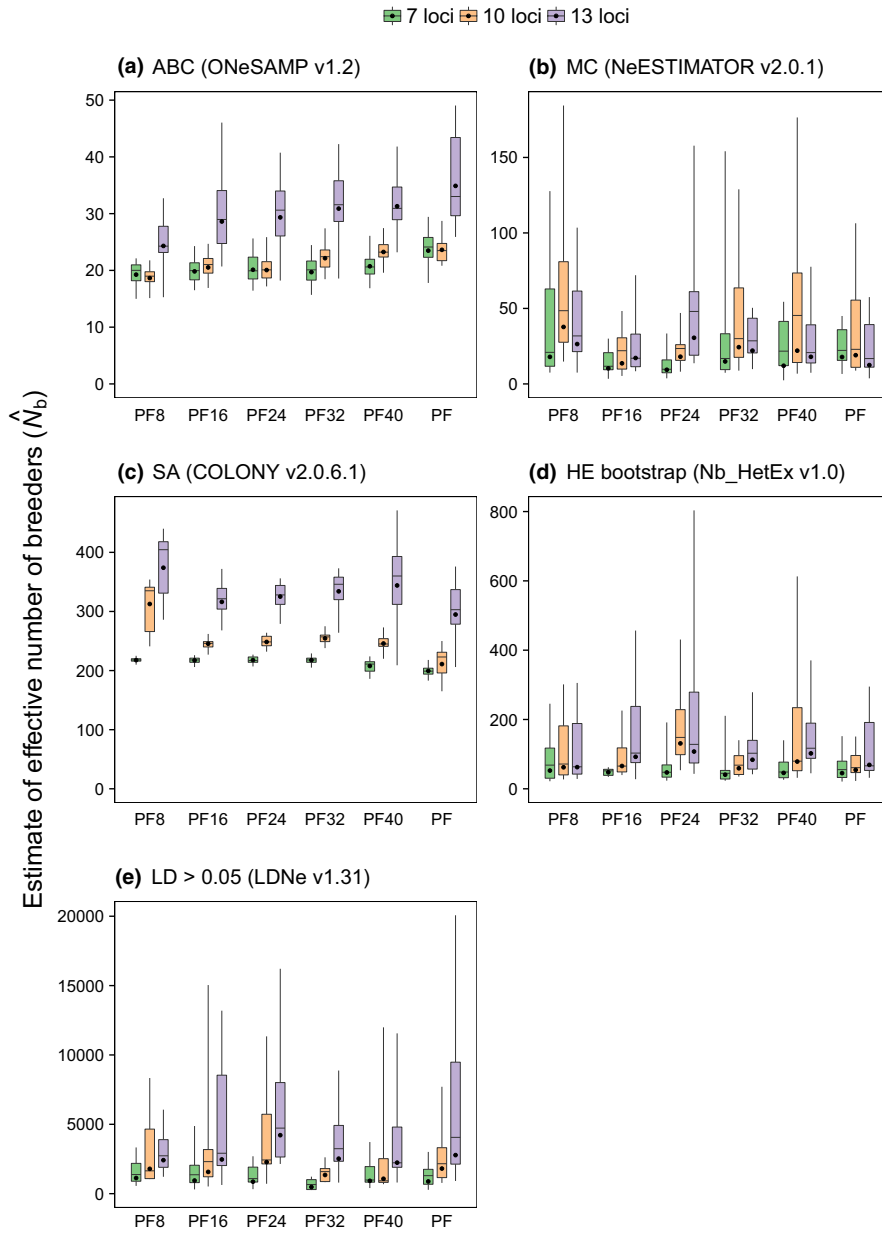
## 3 | RESULTS

### 3.1 | Description of microsatellite loci data and population parameters

Null alleles were detected at locus  $WS_{\mu} 03$  only in PF (frequency: 0.12) and locus  $WS_{\mu} 20$  in all colonies (frequencies: 0.13 in PF, 0.05 in FI, 0.14 in SG1 and 0.20 in SG2). However, loci  $WS_{\mu} 03$  and  $WS_{\mu} 20$  had high informative content (from 0.55 to 0.80 for  $WS_{\mu} 03$  and 0.73 to 0.85 for  $WS_{\mu} 20$ ), and these were maintained in the datasets with their allelic frequencies adjusted. After correcting for the presence of null alleles, only loci  $WS_{\mu} 08$  and  $WS_{\mu} 23$  in the FI colony and locus  $WS_{\mu} 23$  in the SG1 colony departed significantly from HW expectations. There was no evidence of significant linkage disequilibrium or deviations from selective neutrality at any loci. Genetic diversity indices (Table 1) differed moderately among the simulated cohorts and weakly among the real cohorts. For example, 55% of the 27 pairwise comparisons of  $A_O$  were significant, 26% of comparisons of  $A_R$  were significant and 41% of comparisons of  $H_E$  were significant. AMOVA indicated that most nuclear variation (99.09%) was contained within populations ( $F_{ST} = 0.009$ ,  $p < 0.001$ ). Likewise, Bayesian clustering analyses revealed low Delta  $K$  values, with all individuals admixed at  $K \geq 2$ , suggesting gene flow and no significant population subdivisions (Supporting Information Appendix S1).



**FIGURE 2** Power of different methods in estimating  $N_b$  for simulated cohorts. Mean percentage of all simulated cohorts that yielded finite and positive  $\widehat{N}_b$ , finite 95% confidence interval (95% CI), and outliers are presented



**FIGURE 3** Estimates of effective number of breeders ( $\hat{N}_b$ ) for cohorts simulated from different datasets.  $\hat{N}_b$  obtained for cohorts simulated for Porto da Fazenda (PF) population and subsets of different numbers of individuals genotyped (PF8, PF16, PF24, PF32, and PF40) are showed using following methods: approximate Bayesian computation (ABC; a), molecular coancestry (MC; b), sibship assignment with sibship size prior and sibship scaling (SA; c), heterozygote excess with 10,000 bootstrapping iterations (HE bootstrap; d), and unbiased linkage disequilibrium with allelic frequencies  $>0.05$  (LD  $>0.05$ ; e). Black dots within each bar indicate harmonic mean of  $\hat{N}_b$

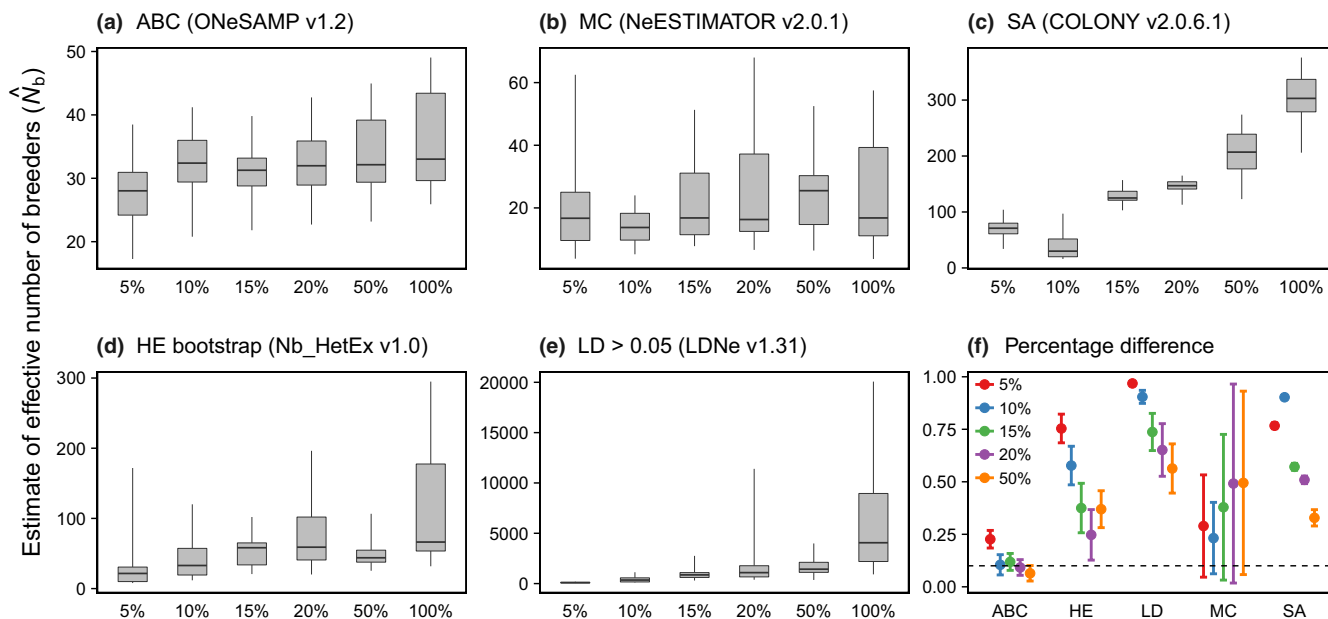
Evidence of low migration was found among wood stork colonies from the Pantanal wetland, as indicated by the assignment of three individuals sampled in FI and one individual sampled in SG1 as belonging to the SG2 colony. The mean number of nestlings from accessed nests ( $\bar{k}$ ) was 2.29 in FI, 2.13 in SG1, 2.12 in SG2, and 2 in PF. The seasonal variance in the number of nestlings ( $V_k$ ) was 0.48 in FI, 0.35 in SG1, 0.24 in SG2, and 0.51 in PF. Nonrandom variance in breeding success was also found: mean  $V_k/\bar{k}$  ranged from 0.11 to 0.25 and  $k$  values did not adhere to a Poisson distribution model in any population ( $\alpha < 0.05$ ).

### 3.2 | Estimates of $N_b$ from simulated data

The  $\hat{N}_b$  obtained for simulated cohorts revealed that the ABC and SA methods could be useful for estimating population sizes in waterbird colonies (Figures 2 and 3, Supporting Information Appendix S2). Only the ABC and SA methods yielded finite  $\hat{N}_b$  with finite 95%

CI for all cohorts. There was no significant difference between the methods in the percentage of outliers (1 to 6%) (Figure 2). By yielding higher and more precise  $\hat{N}_b$  when more loci were employed (Figure 3, Supporting Information Appendix S2), the ABC and SA methods also demonstrated a more consistent pattern of distribution of  $\hat{N}_b$  obtained using different numbers of loci. Moreover, the ABC method, followed by the SA method, had the narrowest 95% CIs and  $\log RMSE$  (Supporting Information Appendices S2 and S3). The HE and LD methods performed better with the re-sampling approach and assuming a threshold of 0.05 for minimum allele frequencies, respectively. Only the results of these best approaches were considered in further evaluations with the HE and LD methods using simulations.

The minimum acceptable sample sizes determined for each estimator considering precision and error rates varied widely (Figure 4). The ABC method maintained an error rate of  $\leq 10\%$  with the smallest minimum sample size ( $\leq 10\%$  of the population), whereas the LD



**FIGURE 4** Estimates of effective number of breeders ( $\hat{N}_b$ ) obtained with different sample sizes.  $\hat{N}_b$  values obtained by randomly sampling 5%, 10%, 15%, 20%, and 50% of entire cohorts simulated from original Porto da Fazenda (PF) population dataset.  $\hat{N}_b$  values obtained by following methods are presented: approximate Bayesian computation (ABC; a), molecular coancestry (MC; b), sibship assignment with sibship size prior and sibship scaling (SA; c), heterozygote excess with 10,000 bootstrap iterations (HE bootstrap; d), and unbiased linkage disequilibrium with allelic frequencies >0.05 (LD >0.05; e). Percentage differences (error rate) calculated among harmonic mean of  $\hat{N}_b$  obtained using different sample sizes versus harmonic mean of  $\hat{N}_b$  obtained using entire cohorts (f). Dashed line indicates threshold of acceptable error rate ( $\leq 10\%$ )

method required a large sample size to maintain a low error rate. With a threshold of 10% for the error rate, the minimum sample sizes required by the SA and LD methods were >50% of the population. The minimum sample sizes required by the MC and HE methods were >10% and >20%, respectively. Comparisons of the  $\hat{N}_b$  obtained by sampling 100% and 10% of the cohorts from the simulated populations revealed that the ABC method had the lowest error rate (10%), followed by the MC (23%), HE (58%), SA (90%), and LD (90%) methods (Figure 4).

### 3.3 | Estimates of $N_a$ and $N_b$ for natural populations

The number of active nests,  $N_a$ , and  $N_1$  are given in Table 2.  $\hat{N}_b$  computed using different methods varied widely (Table 2). The ABC and SA methods yielded finite  $\hat{N}_b$  values for all cohorts. The ABC method also yielded more precise estimates (narrower 95% CIs and lower V), followed by the SA, MC, LD, and HE methods (Table 2).  $\hat{N}_b$  obtained using the ABC, SA and MC methods were positively correlated with  $\hat{N}_a$  and sample sizes (correlation values = 1.0).  $\hat{N}_b/\hat{N}_a$  ratios ranged from 0.02 to 0.49 (Table 2).

## 4 | DISCUSSION AND CONCLUSIONS

### 4.1 | Assumptions for estimation of $N_b$

Genetic estimators of  $N_b$  make several simplifying assumptions, such as no immigration, no mutation, no selection, no population

subdivision, and samples composed of discrete generations (Luikart et al., 2010; Wang, 2016). The paired comparisons using genetic diversity indices showed that the set of loci used provided sufficient discriminatory power to compute parameters associated with the estimation of  $N_b$  (Table 1). Likewise, most of the assumptions made by genetic estimators of  $N_b$  were met in the analyses of wood stork breeding colonies. For example, samples were collected at random from single cohorts representing nonoverlapping generations. Departures from a 1:1 adult sex ratio were not seen and assortative mating has not been reported in the wood storks (Coulter, Rodgers, Ogden, & Depkin, 1999). Given the low mutation rate of microsatellites in short time intervals (Beaumont, 2003), the observed departures from HW expectations in FI are unlikely to result from mutation and no evidence of selection was found (non-significant neutrality tests). Furthermore, there was no evidence of significant population subdivision, as indicated by AMOVA and Bayesian clustering analyses (Supporting Information Appendix S1).

Evidence of migration was found among the studied colonies, albeit at a low rate, which is consistent with previous studies reporting regional philopatry with migration (Del Lama, Avelar, & Nascimento, 2015; Miño et al., 2017). However, if migrants are exchanged between genetically similar sources, the effects on  $\hat{N}_b$  are expected to be small (Waples & England, 2011). Thus, the effect of migration on the estimation of  $N_b$  can be considered negligible in the colonies studied.

The observed variance in breeding success deviated significantly from the  $V_k/\bar{k} \approx 1$  ratio assumed by the Wright-Fisher model (Waples,



**TABLE 2** Census data and estimates of effective number of breeders ( $\hat{N}_b$ ) in wood stork populations. Estimates and corresponding 95% confidence intervals (CI) for number of active nests, annual census of adults ( $\hat{N}_a$ ), and number of nestlings in cohorts ( $\hat{N}_1$ ) as well as  $\hat{N}_b$ , variance in  $\hat{N}_b$  ( $V$ ), and  $\hat{N}_b/\hat{N}_a$  ratio were showed

Pop.	Active nests (95% CI)	$\hat{N}_a$ (95% CI)	$\hat{N}_1$ (95% CI)	Method	$\hat{N}_b$ (95% CI)	$V^i$	$\hat{N}_b/\hat{N}_a$
Fazenda Ipiranga (FI)	674 (637 – 710)	1,347 (1,274–1,420)	1,543 (1,459–1,626)	ABC <sup>a</sup>	30.60 (23.98–43.78)	64.71	0.02
				MC <sup>b</sup>	78.00 (0.1–391.7)	502.05	0.06
				SA <sup>c</sup>	78.00 (56–107)	65.38	0.06
				LD >0.02 <sup>d</sup>	232.60 (110.9–1461.4)	580.61	0.17
				LD >0.05 <sup>e</sup>	170.90 (87.6–591.6)	294.91	0.13
				LD >0.10 <sup>f</sup>	149.50 (76.2–492.5)	278.46	0.11
Sangradouro 1 (SG1)	341 (327–354)	691 (654–708)	726 (696–754)	ABC <sup>a</sup>	24.70 (19.98–33.08)	53.04	0.04
				MC <sup>b</sup>	34.00 (0.9–125.6)	366.76	0.05
				SA <sup>c</sup>	52.00 (36–80)	84.62	0.08
				LD >0.02 <sup>d</sup>	234.00 (68.6–∞)	NA	0.34
				LD >0.05 <sup>e</sup>	332.10 (71.0–∞)	NA	0.49
				LD >0.10 <sup>f</sup>	267.90 (60.9–∞)	NA	0.39
				HE <sup>g</sup>	15.00 (5.5–∞)	NA	0.02
				HE BTSP. <sup>h</sup>	13.00 (6.6–∞)	NA	0.02
Sangradouro 2 (SG2)	85	170	180	ABC <sup>a</sup>	18.60 (14.99–24.80)	52.74	0.11
				SA <sup>c</sup>	48.00 (26–124)	204.17	0.28
Pool of FI, SG1 and SG2	1,100	2,208	2,449	ABC <sup>a</sup>	31.64 (25.52–43.18)	55.81	0.01
				SA <sup>c</sup>	225.00 (181–277)	42.66	0.10

Notes. ∞: infinite value; NA: not available.

<sup>a</sup>Approximate Bayesian computation (ONE-SAMP v1.2). <sup>b</sup>Molecular coancestry (NEESTIMATOR v2.0.1). <sup>c</sup>Sibship assignment without (COLONY v2.0.6.1). <sup>d,e,f</sup>

Unbiased linkage disequilibrium with allele frequencies >0.02, >0.05, or >0.10, respectively (LDNE v1.31). <sup>g,h</sup>Heterozygote excess with and without 10,000 bootstrap iterations, respectively (NB\_HETEX v1.0). <sup>i</sup>Difference between finite limits of 95% CI as percentage of  $\hat{N}_b$ .

2006). This is somewhat expected in the wood stork, because non-random variance in breeding success may be associated with failure to recruit individuals during the breeding season or the presence of age categories among breeders (Borkhataria, Frederick, Hylton, Bryan, & Rodgers, 2008). This variance in breeding success could reduce  $N_b$  in relation to the census size (e.g., Whiteley et al., 2015).

## 4.2 | Estimates of $N_b$ from simulated data

As reported for some methods in earlier studies (Pudovkin, Zhadanova, & Hedgecock, 2010; Wang, 2016; Waples & Do, 2010), the simulations made in this study suggest that single-sample methods perform differently depending on the sample size and number of loci employed (Figures 3 and 4; Supporting Information Appendices S2 and S3). Under the conditions used herein, the ABC method demonstrated greater precision in estimating  $N_b$  (Figures 2 and 3; Supporting Information Appendices S2 and S3), which may be associated with the amount of information assessed using this method and the adequate choice of priors (Holleley et al., 2014; Luikart et al., 2010). Although the exact contribution of each summary genetic statistics used by the ABC method is unclear (Wang, 2016),  $\hat{N}_b$  should be mainly influenced by

the number of breeders producing cohorts (David A. Tallmon, personal communication). In contrast, the LD method demonstrated the most infinite values and lowest precision, possibly due to large sampling errors or very large  $N_b$  (Waples & Do, 2010). The high percentage of infinite estimates obtained by the HE and MC methods is consistent with previous studies showing that these methods most often yield finite values when the real  $N_b$  is very small (Pudovkin et al., 2010; Wang, 2016). Moreover, the ABC and SA methods demonstrated reduced  $\log RMSE$ , hence increased precision, as more loci were used (Supporting Information Appendix S2). Few loci could lead to an increased effect of noise on the signal of genetic drift (Waples, 2006). The simulations made in this study are in line with this expectation, as few loci with low or moderate informative content originated some identical-by-state multilocus genotypes (Supporting Information Table S2) and resulted in lower  $\hat{N}_b$  (Figure 3).

The ABC method seems to be less influenced by change in sample size, as indicated by its low error rate even when sampling only 10% of the population and the low rate of change in mean  $\hat{N}_b$  with the increase in sample size (Figure 4). The SA method demonstrated more precise  $\hat{N}_b$  (narrower 95% CIs and low error rate) when the sample size increased (Figure 4), as expected for a method that requires

larger sample sizes when using markers with low informative content (Wang, 2009). The LD method also demonstrated greater precision with the increase in sample size (Figure 4). This is in agreement with the prediction that doubling the sample size improves the precision of this method more than doubling the number of loci (Antao, Pérez-Figueroa, & Luikart, 2011). Larger sample sizes increase the number of alleles per locus used in the analyses and lower the odds of over-sampling families (England, Luikart, & Waples, 2010; Whiteley et al., 2012).

### 4.3 | Estimates of $N_a$ and $N_b$ for natural populations

Estimates of the effective number of breeders can provide important information on a population and complement census measures in monitoring programs (Ferchaud et al., 2016). The method used in the present study to estimate  $N_a$  in the wood stork proved to be promising for the population monitoring of colonial waterbirds, as it enabled determining the size of breeding colonies previously assumed to have different sizes by field observations alone. Moreover,  $\hat{N}_a$  were useful for determining how many individuals corresponded to 10% of the sampled cohorts.

When the objective is to determine population trends using  $\hat{N}_b$ , it is desirable to have precise (even downward-biased) estimates that are positively correlated with  $N_a$ , but not with sample size (Beebee, 2009; Tallmon et al., 2010). In the present study, the ABC method yielded the most precise  $\hat{N}_b$ , followed by the SA method (Table 2), which is in line with previous studies on other species (e.g., Álvarez, Lourenço, Oro, & Velo-Antón, 2015). As presented herein, Beebee (2009) also found greater variance ( $V$ ) in  $\hat{N}_b$  obtained with the SA method than with the ABC method, probably because setting prior limits to  $\hat{N}_b$  in the ABC method leads to an increase in precision. On the other hand, the LD, MC, and HE methods demonstrated lower precision and none of these methods yielded finite estimates for SG1 or the pooled dataset of the three cohorts (Table 2). The ABC, SA, and MC methods fulfilled the criterion of positively correlating with  $\hat{N}_a$ , but also positively correlated with sample size. Using the SA method and setting  $\bar{k}$  as the sibship prior improved  $\hat{N}_b$  the with regard to reflecting the population size. Without the prior, however, larger populations yielded smaller estimates (data not shown). Given that natural populations with intrinsic moderate microsatellite variation were studied (Table 1), the results of the SA method could indicate that  $\hat{N}_b$  may correlate with population size when loci information is insufficient, but  $\bar{k}$  is known. The correlations with sample size found in the ABC method could probably arise because the sample size was set proportionally to  $\hat{N}_a$ , given that  $\hat{N}_b$  obtained with this method were the only ones not to increase proportionally to sample size in the simulated cohorts (Item 4.2).

Although  $\hat{N}_b/\hat{N}_a$  ratios varied widely, even in single colonies (Table 2), estimating these ratios in waterbird colonies may assist in long-time monitoring to investigate demographic changes that may mirror concurrent environmental changes. The  $\hat{N}_b/\hat{N}_a$  ratios (Table 2) were below the mean of 0.225 reported in a previous quantitative survey (Palstra & Fraser, 2012), except those calculated using  $\hat{N}_b$

computed with the SA method for SG2 and the LD method overall. Nonetheless,  $\hat{N}_b/\hat{N}_a$  ratios obtained in wood storks were often close to  $\hat{N}_e/\hat{N}_c$  ratios reported for rare and endangered bird species (often  $<0.1$ ) (e.g., Brekke, Bennett, Santure, & Ewen, 2011; Lopes, Miño, Rocha, Oliveira, & Del Lama, 2013; Ramstad, Colbourne, Robertson, Allendorf, & Daugherty, 2013).

### 4.4 | Guidelines for monitoring populations of colonial waterbirds

Many species of colonial waterbirds nest in accessible aggregations, which is a condition that at first seems favorable to the application of census methods based on counts, as used in the present study, for monitoring population sizes (e.g., Brown, Tims, Erwin, & Richmond, 2001; Rush et al., 2015; Torres et al., 2010). However, census estimates can be difficult to obtain accurately when thousands of individuals are recruited during the breeding season or when breeding colonies occupy large areas that are difficult to cover in a short period of time. We recommend some characteristics to evaluate prior to choosing a counting method and a target for counting (e.g., active nests or breeding adults): (a) whether there are non-breeding adults in the colony, (b) how easily nests and individuals can be observed for counting, (c) whether reproduction is asynchronous, which may require multiple visits to the colonies and the marking of individuals, (d) whether inactive nests that were abandoned or had persisted from previous breeding seasons may be confused with active nests by the presence of adults in the vicinity, and (e) the difficulty in observing the contents of nests from the ground. The use of genetic estimates, such as  $\hat{N}_b$ , is a valuable alternative for monitoring population size when there is insufficient information on the ecological traits and dynamics of breeding colonies, obtaining this information in the field would consume a prohibitively large amount of time, and when available information indicates that a census is not feasible.

Based on the workflow described in item 2 of this study, the monitoring using  $\hat{N}_a$  and/or  $\hat{N}_b$  can be divided into three consecutive steps. Step 1 (prior to field research) involves performing an extensive literature review, setting up a pilot study, and planning the field and laboratory research. Steps 2 and 3 (field and laboratory research) involve the execution of the planning to obtain the estimates of the census and effective number of breeders, following what was determined in the pilot study (i.e., applying the method(s) chosen for estimating  $N_b$  considering the available information on a given sample size and number of loci).

A literature review and an empirical enquiry near the location of the breeding colonies can eliminate some time-consuming steps from the field research itself, which can assist in the planning of the monitoring as a whole. Using the wood stork populations sampled in the present study as an example, personal communications on the onset of the establishment of the colonies were crucial to defining the period in which the breeding colonies would be visited, as the more adequate time for collecting blood from wood stork nestlings is the period between three and six weeks of age. All information on the reproductive biology, genetic population structure, population

dynamics, and impact of the human presence in the colonies (e.g., Bouton, Frederick, Rocha, Dos Santos, & Bouton, 2005; Del Lama et al., 2015; Frederick & Ogden, 1997; Miño et al., 2017) can contribute to answering important questions during field research planning, such as: 1—Which region and which colonies should primarily be sampled? 2—How much time can be spent visiting each colony? 3—What data can be recorded in the field? 4—What is the best way to count and record active nests considering the heterospecific composition of the colonies?

The results obtained in this study highlight the importance of conducting a pilot study to evaluate the quality of the genotypic dataset for the characterization of polymorphisms in loci, determine the smallest sample size needed to obtain a representative sample of the breeding colonies, and estimate the power of the marker set for confirming whether common assumptions for the estimation of  $N_b$  are fulfilled. A pilot study may start from either published data on populations of the species within the area of interest for monitoring, or from samples already collected and stored, as performed here using the PF sample and described in items 2.2, 2.3, and 2.4. In the absence of previous data or material, researchers can develop a pilot sampling expedition covering at least one colony in the focal region, which is a reasonable decision when considering long-term monitoring.

Detailed planning for the monitoring of breeding colonies can optimize the use of funds and time, especially during the field work, when time can be a limiting factor to visiting several colonies in a single breeding season. Another aspect to consider during the field work is the strengthening of field efforts to minimize human disturbance within colonies and maximizing the information obtained when accessing nests (e.g., including banding, measuring individuals, collecting blood, recording demographic data to estimate reproductive success and census). Depending on the degree of natal philopatry of individuals to either colonies or regions, it is possible to monitor specific colonies or a geographic region. If birds exhibit high fidelity to the breeding sites, a researcher can monitor focal colonies and, in this case, the target should be larger colonies, which are most likely to persist over time than smaller ones (Tsai et al., 2016). Colonies that are not established in consecutive breeding seasons can be included when the monitoring covers a larger geographic range and the birds are philopatric to the region rather than to the colonies, as occurs with the wood stork (Miño et al., 2017). Moreover, breeding colonies should not be considered separate populations without testing for significant genetic differentiation or migration from dissimilar population sources prior to monitoring population size using  $\hat{N}_b$ . Similarly, to pool genetic samples confidently from different colonies as if they belong to a single panmictic unit, prior testing is necessary, as performed in the present study. The effects of ignoring underlying genetic structuring could be more harmful than those of estimating  $N_b$  separately for each colony, even if colonies are not structured (Bernos et al., 2018; Holleley et al., 2014).

The present study pointed that the ABC method is suitable for genetically tracking demographic changes in wood stork populations, as it required smaller sample sizes and fewer loci to estimate

$N_b$  with greater precision. Results also showed that the SA method is promising for the estimation of  $N_b$  when a sibship size prior can be set with certain confidence, for example, using the mean number of nestlings from accessed nests ( $\bar{k}$ ). However, as demonstrated here, the SA method requires a larger sample size to yield precise estimates with a limiting number of loci. Likewise, findings from this study show that the sample size relative to  $N_a$  is an important and limiting factor to estimating  $N_b$ . Therefore, whenever possible, researchers may estimate  $N_a$  and  $\bar{k}$ , in order to reach adequate sample sizes to obtain precise, unbiased  $\hat{N}_b$ .  $N_a$  and  $\bar{k}$  should be estimated immediately before or concurrently with the blood collection during initial field work, as the number of nestlings in a colony may vary greatly over the breeding season (Borkhataria et al., 2008; Vergara & Aguirre, 2006).

This study demonstrates that  $N_a$  and/or  $N_b$  estimates are useful tools for monitoring natural breeding bird populations. However, as the estimates varied widely depending on the method used, researchers can follow the methodological workflow offered herein to explore genotypic datasets and choose the estimator most suited to the conditions of the focal natural populations prior to start the monitoring. Furthermore, this workflow may not provide useful for the effective monitoring of natural populations when it is not feasible in terms of funds and time to collect representative population samples or when monitoring populations of nomadic species. The methodology based on  $\hat{N}_a$  described herein can be used by researchers interested in genetic monitoring and studying populations of birds with colonial breeding where the sampling of discrete generations is possible. In particular, the workflow offered herein can be used by researchers studying colonial birds with a limited number of microsatellites and low levels of polymorphism. Combining demographic and genetic estimates is a feasible strategy for monitoring few small colonies with no major access difficulties and low human impact. Moreover, combining demographic and genetic approaches is highly recommended for populations of rare species or with a known history of population decline to support conservation efforts by enabling the informed, timely, precise monitoring of trends in population size.

## ACKNOWLEDGMENTS

The authors are grateful to the Brazilian environmental agency IBAMA for granting permission to handle the birds and collect biological material (Permit no. 12437-2/2013) and the SESC Pantanal National Management and *Secretaria do Meio Ambiente do Estado de Mato Grosso* (SEMA-MT) for providing logistical support during the sample collection. We thank Keith Goodnight for clarifications regarding the simulation model. F.M.S. was supported by a scholarship from *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES). During the initial drafting of this manuscript C.I.M. was a fellow of CAPES under the PRODOC program (Project no. 004515; AUXPE 2637/2010). This study was supported by the Brazilian fostering agency Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) through research grants to S.N.D.L. (2010/52315-7, 2016/01673-7).

## CONFLICT OF INTEREST

None declared.

## AUTHOR CONTRIBUTIONS

F.M.S., C.I.M., and S.N.D.L. designed the study and wrote the manuscript. F.M.S. and S.N.D.L. collected samples. F.M.S. conducted the genetic analyses and the census in natural populations. F.M.S. and C.I.M. led the computational work. F.M.S. and R.I. conducted the statistical analyses and created the figures with the results.

## ORCID

Fagner M. Silva  <http://orcid.org/0000-0002-2218-3122>

Carolina I. Miño  <http://orcid.org/0000-0002-9098-583X>

Silvia N. Del Lama  <http://orcid.org/0000-0001-8920-3442>

## REFERENCES

- Álvarez, D. A., Lourenço, A., Oro, D., & Velo-Antón, G. (2015). Assessment of census ( $N$ ) and effective population size ( $N_e$ ) reveals consistency of  $N_e$  single-sample estimators and a high  $N_e/N$  ratio in an urban and isolated population of fire salamanders. *Conservation Genetics Resources*, 7, 705–712. <https://doi.org/10.1007/s12686-015-0480-0>
- Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: A workbench to detect molecular adaptation based on a  $F_{st}$ -outlier method. *BioMed Central Bioinformatics*, 9, 323–327.
- Antao, T., Pérez-Figueroa, A., & Luikart, G. (2011). Early detection of population declines: High power of genetic monitoring using effective population size estimators. *Evolutionary Applications*, 4, 144–154. <https://doi.org/10.1111/j.1752-4571.2010.00150.x>
- Atkinson, P. W., Austin, G. E., Rehfish, M. M., Baker, H., Cranswick, P., Kershaw, M., ... Maclean, M. D. (2006). Identifying declines in waterbirds: The effects of missing data, population variability and count period on the interpretation of long-term survey data. *Biological Conservation*, 130, 549–559. <https://doi.org/10.1016/j.biocon.2006.01.018>
- Beaumont, M. A. (2003). Estimation of population growth or decline in genetically monitored populations. *Genetics*, 164, 1139–1160.
- Beebee, T. J. C. (2009). A comparison of single-sample effective size estimators using empirical toad (*Bufo calamita*) population data: Genetic compensation and population size-genetic diversity correlations. *Molecular Ecology*, 18, 4790–4797. <https://doi.org/10.1111/j.1365-294X.2009.04398.x>
- Belmar-Lucero, S., Wood, J. L. A., Scott, S., Harbicht, A. B., Hutchings, J. A., & Fraser, D. J. (2012). Concurrent habitat and life history influences on effective/census population size ratios in stream-dwelling trout. *Ecology and Evolution*, 2, 562–573. <https://doi.org/10.1002/ece3.196>
- Bernos, T. A., Yates, M. C., & Fraser, D. J. (2018). Fine-scale differences in genetic and census population size ratios between two stream fishes. *Conservation Genetics*, 19, 265–274. <https://doi.org/10.1007/s10592-017-0997-8>
- Borkhataria, R. R., Frederick, P. C., Hylton, R., Bryan, A. L. Jr, & Rodgers, J. A. Jr (2008). A preliminary model of wood stork population dynamics in the Southeastern United States. *Waterbirds*, 31(sp1), 42–49. [https://doi.org/10.1675/1524-4695\(2008\)31\[42:APMOWS\]2.0.CO;2](https://doi.org/10.1675/1524-4695(2008)31[42:APMOWS]2.0.CO;2)
- Bouton, S. N., Frederick, P. C., Rocha, C. D., Dos Santos, A. T. B., & Bouton, T. C. (2005). Effects of tourist disturbance on wood stork nesting success and breeding behavior in the Brazilian Pantanal. *Waterbirds*, 28, 487–497. [https://doi.org/10.1675/1524-4695\(2005\)28\[487:EOTDOW\]2.0.CO;2](https://doi.org/10.1675/1524-4695(2005)28[487:EOTDOW]2.0.CO;2)
- Brekke, P., Bennett, P. M., Santure, A. W., & Ewen, J. G. (2011). High genetic diversity in the remnant island population of hihi and the genetic consequences of re-introduction. *Molecular Ecology*, 20, 29–45. <https://doi.org/10.1111/j.1365-294X.2010.04923.x>
- Brookfield, J. F. Y. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, 5, 453–455. <https://doi.org/10.1111/j.1365-294X.1996.tb00336.x>
- Brown, K. M., Tims, J. L., Erwin, R. M., & Richmond, M. E. (2001). Changes in the nesting populations of colonial waterbirds in Jamaica Bay Wildlife Refuge, New York, 1974–1998. *Northeastern Naturalist*, 8, 275–292.
- Bryan, A. L. Jr, Snodgrass, J. W., Robinette, J. R., & Hopkins, L. B. (2005). Parental activities of nesting Wood Storks relative to time of day, tide level and breeding stage. *Waterbirds*, 28, 139–145. [https://doi.org/10.1675/1524-4695\(2005\)028\[0139:PAONWS\]2.0.CO;2](https://doi.org/10.1675/1524-4695(2005)028[0139:PAONWS]2.0.CO;2)
- Campanini, E. B., Sanches, A., Hatanaka, T., & Del Lama, S. N. (2012). Isolation and characterization of 11 microsatellite loci from cattle egret (*Bubulcus ibis*) and cross-amplification in other Ardeidae species. *Conservation Genetics Resources*, 4, 707–709. <https://doi.org/10.1007/s12686-012-9627-4>
- Coulter, M. C., Rodgers, J. A., Ogden, J. C., & Depkin, F. C. (1999). Wood Stork (*Mycteria americana*). In: A. Poole & F. Gill (Eds.), *The birds of North America*, no. 409 (pp. 1–28). Washington, DC: Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union.
- Crow, J. F., & Kimura, M. (1970). *An introduction to population genetics theory*. New York: Harper and Row.
- Crow, J. F., & Morton, N. E. (1955). Measurement of gene frequency drift in small populations. *Evolution*, 9, 202–214. <https://doi.org/10.1111/j.1558-5646.1955.tb01531.x>
- Del Lama, S. N., Avelar, L. H. S., & Nascimento, J. L. X. (2015). Post-breeding movements of Wood Storks in Brazil and Argentina. *Journal of Field Ornithology*, 86, 283–287. <https://doi.org/10.1111/jofo.12122>
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillet, B. J., & Ovenden, J. R. (2014). NeEstimator v2: Reimplementation of software for the estimation of contemporary effective population size ( $N_e$ ) from genetic data. *Molecular Ecology Resources*, 14, 209–214. <https://doi.org/10.1111/1755-0998.12157>
- Earl, D., & vonHoldt, B. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Ellegren, H. (2013). The evolutionary genomics of birds. *Annual Review of Ecology, Evolution and Systematics*, 44, 239–259. <https://doi.org/10.1146/annurev-ecolsys-110411-160327>
- England, P. R., Luikart, G., & Waples, R. S. (2010). Early detection of population fragmentation using linkage disequilibrium estimation of effective population size. *Conservation Genetics*, 11, 2425–2430. <https://doi.org/10.1007/s10592-010-0112-x>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Ferchaud, A. L., Perrier, C., April, J., Hernandez, C., Dionne, M., & Bernatchez, L. (2016). Making sense of the relationships between  $N_e$ ,  $N_b$  and  $N_c$  towards defining conservation thresholds in Atlantic salmon (*Salmon salar*). *Heredity*, 117, 268–278. <https://doi.org/10.1038/hdy.2016.62>

- Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford, UK: Oxford University Press. <https://doi.org/10.5962/bhl.title.27468>
- Frankham, R. (1996). Relationship of genetic variation to population size in wildlife. *Conservation Biology*, 10, 1500–1508. <https://doi.org/10.1046/j.1523-1739.1996.10061500.x>
- Fraser, D. J., Hansen, M. M., Østergaard, N. T., Tessier, N., Legault, M., & Bernatchez, L. (2007). Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. *Molecular Ecology*, 16, 3866–3889. <https://doi.org/10.1111/j.1365-294X.2007.03453.x>
- Frederick, P. C., Gawlik, D. E., Ogden, J. C., Cook, M. I., & Lusk, M. (2009). The white ibis and wood stork as indicators for restoration of the everglades ecosystem. *Ecological Indicators*, 9, 583–585. <https://doi.org/10.1016/j.ecolind.2008.10.012>
- Frederick, P. C., & Ogden, J. C. (1997). Philopatry and nomadism: Contrasting long-term movement behavior and population dynamics of White Ibises and Wood Storks. *Colonial Waterbirds*, 20, 316–323. <https://doi.org/10.2307/1521699>
- Gomez-Uchida, D., Palstra, F. P., Knight, T. W., & Ruzzante, D. E. (2013). Contemporary effective population and metapopulation size ( $N_e$  and meta- $N_e$ ): Comparison among three salmonids inhabiting a fragmented system and differing in gene flow and its asymmetries. *Ecology and Evolution*, 3, 569–580. <https://doi.org/10.1002/ece3.485>
- Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetical Research*, 38, 209–216. <https://doi.org/10.1017/S0016672300020553>
- Hinkson, K. M., & Richter, S. C. (2016). Temporal trends in genetic data and effective population size support efficacy of management practices in critically endangered dusky gopher frogs (*Lithobates sevosus*). *Ecology and Evolution*, 5, 2667–2678. <https://doi.org/10.1002/ece3.2084>
- Holleley, C. E., Nichols, R. A., Whitehead, M. R., Adamack, A. T., Gunn, M. R., & Sherwin, W. B. (2014). Testing single-sample estimators of effective population size in genetically structured populations. *Conservation Genetics*, 15, 23–35. <https://doi.org/10.1007/s10592-013-0518-3>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.
- Jones, O. R., & Wang, J. (2010). COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, 10, 551–555. <https://doi.org/10.1111/j.1755-0998.2009.02787.x>
- Kalinowski, S. T. (2005). HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, 5, 187–189. <https://doi.org/10.1111/j.1471-8286.2004.00845.x>
- Konovalov, D. A., Manning, C., & Henshaw, M. T. (2004). KINGROUP: A program for pedigree relationship reconstruction and kin group assignments using genetic markers. *Molecular Ecology Notes*, 4, 779–782. <https://doi.org/10.1111/j.1471-8286.2004.00796.x>
- Kushlan, J. A. (1992). Population biology and conservation of colonial wading birds. *Colonial Waterbirds*, 15, 1–7. <https://doi.org/10.2307/1521348>
- Lande, R. (1988). Genetics and demography in biological conservation. *Science*, 241, 1455–1460. <https://doi.org/10.1126/science.3420403>
- Lindenmayer, D. B., & Likens, G. E. (2010). The science and application of ecological monitoring. *Biological Conservation*, 143, 1317–1328. <https://doi.org/10.1016/j.biocon.2010.02.013>
- Lopes, I. F., Miño, C. I., Rocha, C. D., Oliveira, D. M. M., & Del Lama, S. N. (2013). Inferred kinship patterns reveal low levels of extra-pair paternity in the endangered Neotropical Jabiru Stork (*Jabiru mycteria*, Aves: Ciconiiformes). *Genetica*, 141, 195–203. <https://doi.org/10.1007/s10709-013-9718-5>
- Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K., & Allendorf, F. W. (2010). Estimation of census and effective population sizes: The increasing usefulness of DNA-based approaches. *Conservation Genetics*, 11, 355–373. <https://doi.org/10.1007/s10592-010-0050-7>
- Menéndez, J., Álvarez, I., Fernandez, I., Menéndez-Arias, N. A., & Goyache, F. (2016). Assessing performance of single-sample molecular genetic methods to estimate effective population size: Empirical evidence from the endangered Gochu Asturcelta pig breed. *Ecology and Evolution*, 6, 4971–4980. <https://doi.org/10.1002/ece3.2240>
- Miño, C. I., Avelar, L. H. S., da Silva, F. M., Perez, M. F., Menezes, L. F., & Del Lama, S. N. (2017). Genetic differentiation and historical demography of Wood Stork populations in Brazilian wetlands: Implications for the conservation of the species and associated ecosystems. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 27, 1313–1324. <https://doi.org/10.1002/aqc.2791>
- Nomura, T. (2008). Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications*, 1, 462–474. <https://doi.org/10.1111/j.1752-4571.2008.00015.x>
- Nunes, G. T., Efe, M. A., Freitas, T. R. O., & Bugoni, L. (2017). Conservation genetics of threatened Red-billed Tropicbirds and White-tailed Tropicbirds in the southwestern Atlantic Ocean. *The Condor*, 119, 251–260. <https://doi.org/10.1650/CONDOR-16-141.1>
- Palstra, F. P., & Fraser, D. J. (2012). Effective/census population size ratio estimation: A compendium and appraisal. *Ecology and Evolution*, 2, 2357–2365. <https://doi.org/10.1002/ece3.329>
- Peakall, R., & Smouse, P. E. (2012). GenA1Ex 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, 28, 2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Péron, G., Ferrand, Y. D. L., Leray, G., & Gimenez, O. (2013). Waterbird demography as indicator of wetland health: The French-wintering common snipe population. *Biological Conservation*, 164, 123–128. <https://doi.org/10.1016/j.biocon.2013.04.015>
- Peters, J. L., Omland, K. E., & Johnson, K. P. (2007). Population structure and mitochondrial polyphyly in North American Gadwalls (*Anas strepera*). *The Auk*, 124, 444–462. [https://doi.org/10.1642/0004-8038\(2007\)124\[444:PSAMPI\]2.0.CO;2](https://doi.org/10.1642/0004-8038(2007)124[444:PSAMPI]2.0.CO;2)
- Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L., & Estoup, A. (2004). GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, 95, 536–539. <https://doi.org/10.1093/jhered/esh074>
- Primmer, C. R., Raudsepp, T., Chowdhary, B. P., Møller, A. P., & Ellegren, H. (1997). Low frequency of microsatellites in the avian genome. *Genome Research*, 7, 471–482. <https://doi.org/10.1101/gr.7.5.471>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Pudovkin, A. I., Zaykin, D., & Hedgecock, D. (1996). On the potential for estimating effective number of breeders from heterozygote-excess in progeny. *Genetics*, 144, 383–387.
- Pudovkin, A. I., Zhadanova, O. L., & Hedgecock, D. (2010). Sampling properties of the heterozygote-excess estimator of the effective number of breeders. *Conservation Genetics*, 11, 759–771. <https://doi.org/10.1007/s10592-009-9865-5>
- R Development Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ramstad, K. M., Colbourne, R. M., Robertson, H. A., Allendorf, F. W., & Daugherty, C. H. (2013). Genetic consequences of a century of protection: Serial founder events and survival of the little spotted kiwi (*Apteryx owenii*). *Proceedings of the Royal Society London B: Biological Sciences*, 280, 20130576. <https://doi.org/10.1098/rspb.2013.0576>
- Rannala, B., & Mountain, J. L. (1997). Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 9197–9201. <https://doi.org/10.1073/pnas.94.17.9197>
- Rush, S. A., Pekarik, C., Weseloh, D. V., Cuthbert, F., Moore, D., & Wires, L. (2015). Changes in heron and egret populations on the Laurentian

- Great Lakes and connecting channels. *Avian Conservation and Ecology*, 10, 1977–2009.
- Sahai, H., & Khurshid, A. (1995). A note on confidence intervals for the hypergeometric parameter in analyzing biomedical data. *Computers in biology and medicine*, 25, 35–38. [https://doi.org/10.1016/0010-4825\(95\)98883-F](https://doi.org/10.1016/0010-4825(95)98883-F)
- Schwartz, M. K., Luikart, G., & Waples, R. S. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution*, 22, 25–33. <https://doi.org/10.1016/j.tree.2006.08.009>
- Tallmon, D. A., Gregovich, D., Waples, R. S., Baker, C. S., Jackson, J., Taylor, B. L., ... Schwartz, M. K. (2010). When are genetic methods useful for estimating contemporary abundance and detecting population trends? *Molecular Ecology Resources*, 10, 684–692. <https://doi.org/10.1111/j.1755-0998.2010.02831.x>
- Tallmon, D. A., Koyuk, A., Luikart, G., & Beaumont, M. A. (2008). ONE-SAMP: A program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources*, 8, 299–301. <https://doi.org/10.1111/j.1471-8286.2007.01997.x>
- Tavares, D. C., Guadagnin, D. L., Moura, J. F., Siciliano, S., & Merico, A. (2015). Environmental and anthropogenic factors structuring waterbird habitats of tropical coastal lagoons: Implication for management. *Biological Conservation*, 186, 12–21. <https://doi.org/10.1016/j.biocon.2015.02.027>
- Torres, R., Michelutti, P., Dominino, J., León, J., Mangeaud, A., Hilman, G. (2010). Effects of weather and water level on reproduction of colonial waterbirds in Laguna Mar Chiquita – Bañados del Río Dulce (Central Argentina). *Ornitología Neotropical*, 21, 383–396.
- Trail, L. W., Brook, B. W., Frankham, R. R., & Bradshaw, C. J. A. (2010). Pragmatic population viability targets in a rapidly changing world. *Biological Conservation*, 143, 28–34. <https://doi.org/10.1016/j.biocon.2009.09.001>
- Tsai, J. S., Reichert, B. E., Frederick, P. C., & Meyer, K. D. (2016). Breeding site longevity and site characteristics have intrinsic value for predicting persistence of colonies of an endangered bird. *Wetlands*, 36, 636–647.
- van Oosterhout, C., Hutchinson, W. F., Wills, D. P., & Shipley, P. (2004). Micro-checker: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Resources*, 4, 535–538.
- Vásquez-Carrillo, C., Henry, R. W., Henkel, L., & Peery, M. Z. (2013). Integrating populations and genetic monitoring to understand changes in the abundance of a threatened seabird. *Biological Conservation*, 167, 173–178. <https://doi.org/10.1016/j.biocon.2013.08.008>
- Vergara, P., & Aguirre, J. I. (2006). Age and breeding success related to nest position in a white stork *Ciconia ciconia* colony. *Acta Oecologica*, 30, 414–418. <https://doi.org/10.1016/j.actao.2006.05.008>
- Wang, J. (2009). A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular Ecology*, 18, 2148–2164. <https://doi.org/10.1111/j.1365-294X.2009.04175.x>
- Wang, J. (2016). A comparison of single-sample estimators of effective population sizes from genetic marker data. *Molecular Ecology*, 25, 4692–4711. <https://doi.org/10.1111/mec.13725>
- Waples, R. S. (1990). Conservation genetics of Pacific salmon. 2. Effective population size and the rate of loss of genetic variability. *Journal of Heredity*, 81, 267–276. <https://doi.org/10.1093/oxfordjournals.jhered.a110989>
- Waples, R. S. (2002). Effective size of fluctuating salmon populations. *Genetics*, 161, 783–791.
- Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, 7, 167–184. <https://doi.org/10.1007/s10592-005-9100-y>
- Waples, R. S., & Do, C. (2008). LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, 8, 753–756. <https://doi.org/10.1111/j.1755-0998.2007.02061.x>
- Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary  $N_e$  using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3, 244–262. <https://doi.org/10.1111/j.1752-4571.2009.00104.x>
- Waples, R. S., & England, P. R. (2011). Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics*, 189, 633–644. <https://doi.org/10.1534/genetics.111.132233>
- Whiteley, A. R., Coombs, J. A., Cembrola, M., O'Donnell, M. J., Hudy, M., Nislow, K. H., & Letcher, B. H. (2015). Effective number of breeders provides a link between interannual variation in stream flow and individual reproductive contribution in a stream salmonid. *Molecular Ecology*, 24, 3585–3602. <https://doi.org/10.1111/mec.13273>
- Whiteley, A. R., Coombs, J. A., Hudy, M., Robinson, Z., Nislow, K. H., & Letcher, B. H. (2012). Sampling strategies for estimating brook trout effective population size. *Conservation Genetics*, 13, 625–637. <https://doi.org/10.1007/s10592-011-0313-y>
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16, 0097–0159.
- Zhdanova, O. L., & Pudovkin, A. I. (2008). Nb\_HetEx: A program to estimate the effective number of breeders. *Journal of Heredity*, 99, 694–695. <https://doi.org/10.1093/jhered/esn061>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** da Silva FM, Miño CI, Izbicki R, Del Lama SN. Considerations for monitoring population trends of colonial waterbirds using the effective number of breeders and census estimates. *Ecol Evol*. 2018;8:8088–8101. <https://doi.org/10.1002/ece3.4347>