# Reply to the comment on ‘Working with population totals

in the presence of missing data comparing imputation methods in terms of bias and precision’ by Bogaart et al.

Thierry Onkelinx1 • Koen Devos1 • Ivy Jansen1 • Hans Van Calster1 •

Paul Quataert1

Keywords Commentary · Missing data · Monitoring of biodiversity · Survey design and analysis

Bogaart et al. ([2017](#_bookmark2)) indicate in their comment that Onkelinx et al. ([2017b](#_bookmark0)) misinterpreted some aspects of the Underhill index (UIndex), Species Trends Analysis Tool for birds (birdSTATs) and TRends and Indices for Moni- toring data (TRIM), and, as a consequence, do not suffi- ciently acknowledge the quality of those methods. We agree that some operational choices and underlying assumptions were not fully clear to us. However, if the documentation was incomplete and/or if variants existed, we made choices and filled in gaps, always in favour of Underhill and TRIM, to test our approach as thoroughly as possible and to guarantee a balanced comparison.

In fact, in our paper we acknowledge that under certain circumstances, Underhill and TRIM can work properly (see the results and discussion of our paper), but our main point is that multiple imputation covers a broader range of situations and assumptions, and hence it is more flexible and robust. In practice, we cannot always be sure that the assumptions of Underhill or TRIM are valid. In these circumstances, a method that proves to be more robust is preferable. For instance, Bogaart et al. ([2017](#_bookmark2)) mention in their fifth point that it is still an open question within the eco-statistical community whether a negative binomial or a quasi-Poisson distribution is more appropriate. However, with multiple

Communicated by F. Bairlein.

& Thierry Onkelinx thierry.onkelinx@inbo.be

1 Research Institute for Nature and Forest (INBO), Kliniekstraat 25, 1070 Anderlecht, Belgium

imputation, you can make your own choice according to the context or theoretical insight. With an analytical approach, for a new model, the source code needs to be adapted. In addition, in our paper, we caution that an appropriate model must be carefully selected, and we demonstrate what hap- pens with a less appropriate model.

In the following, we reply in greater depth to the statements in Bogaart et al. ([2017](#_bookmark2)). Each number corre- sponds to their numbering.

1. We agree that Onkelinx et al. ([2017b](#_bookmark0)) misinterpreted the imputation step of the UIndex. What we describe as an ‘alternative’ UIndex actually is equivalent to the algorithm described in Underhill and Prys-Jones ([1994](#_bookmark3)). This ‘alternative’ version was used in Fig. 2 of Onkelinx et al. ([2017b](#_bookmark0)).

Onkelinx et al. ([2017b](#_bookmark0)) did not consider incomplete counts sensu Underhill and Prys-Jones ([1994](#_bookmark3)), only complete counts and missing counts. Meanwhile, we have implemented incomplete counts in our multimput R package (Onkelinx et al. [2017a](#_bookmark5)). Our implementa- tion treats incomplete counts as missing counts during the imputation step. Before aggregation, the incom- plete counts are compared with the imputed values, and the highest value is retained.

1. Onkelinx et al. ([2017b](#_bookmark0)) implemented the UIndex algorithm through a generalised linear model. This kind of model assumes that the response variable matches the selected distribution. The negative bino- mial, Poisson and quasi-Poisson distribution all assume non-negative integer values. Although the glm func- tion in R technically handles non-integer values, they are rounded internally (with a warning) before they are processed. Hence, a priori rounding to the nearest integer is sensible.
2. Bogaart et al. ([2017](#_bookmark2)) write that Meij ([2013](#_bookmark6)) does not suggest or endorse the summation of the counts over the year and using 1 divided by the number of counts as weights. However, the paragraph that Bogaart et al. ([2017](#_bookmark2)) mention starts with ‘TRIM works with one count value for every year. When producing TRIM input files, the tool sums the counts in a year and produces a weight factor of 1, divided by the number of counts in this year.’ The paragraph ends with ‘If another method of calculating yearly count values is needed (for instance maximum count value per year), these count values have to be calculated in advance.’ (van der Meij [2013](#_bookmark6),v

2.03 p. 19). And a similar statement on p. 24: ‘When counts are selected, the number of counts per plot per year is determined, a weighing factor is produced dividing 1/number of counts per year and the counts are summed per year.’ (van der Meij [2013](#_bookmark6), v 2.03 p. 24). Furthermore, we read that ‘an option is provided for shifting the count dates. This option is useful for winter projects with a count season starting before new years day. By default, Birdstats uses calendar years as closed count periods for counts. All counts within one calendar year will be taken together. Since this may lead to splitting counts from one winter season into two consecutive years, the count dates can be shifted forward or backward into only one calender year. The number of days to shift is optional between -100 and

?100’ (van der Meij [2013](#_bookmark6), v 2.03 p. 4). Hence we conclude that van der Meij ([2013](#_bookmark6)) at least gives the impression that data from an entire winter season can be used. Onkelinx et al. ([2017b](#_bookmark0)) describe the default procedure of birdSTATs as an average ornithologist is likely to use it.

1. We cannot find any reference in Pannekoek and Van Strien ([2005](#_bookmark4)) on how the standard error on the imputed time totals is calculated, which hampers the ability to check the correctness of TRIM’s ‘analytical approach’ (Bogaart et al. [2017](#_bookmark2)). Moreover, the output examples display no standard error for the imputed time totals (Pannekoek and Van Strien [2005](#_bookmark4), p. 36, 39, 41 and 47). The output files for the TRIM version used by Onkelinx et al. ([2017b](#_bookmark0)) did contain a standard error for the imputed time totals; hence we cannot verify the statement: ‘TRIM does take the effect into account when computing the standard errors of totals and indices based on the fitted or imputed counts’ (Bogaart et al. [2017](#_bookmark2)). However we can state that TRIM uses the predicted value when imputing missing observations (Pannekoek and Van Strien [2005](#_bookmark4), Eq 2.17 p. 17), making TRIM a single imputation method sensu

Onkelinx et al. ([2017b](#_bookmark0)).

We stated that ‘multiple imputation easily allows the incorporation of alternative, more complex imputation models (e.g. interactions, covariates,…)’ (Onkelinx et al. [2017b](#_bookmark0)). We did not claim that it is not possible to use covariates with single imputation. In fact, ‘TRIM allows for additive effects of up to ten categorical covariates on trends and time-point parameters.’ (Pan-

nekoek and Van Strien [2005](#_bookmark4), p. 11), and ‘… the maximum number of categories of a covariate is 90 …’ (Pannekoek and Van Strien [2005](#_bookmark4), p. 19). Pannekoek

and Van Strien ([2005](#_bookmark4)) do not mention interactions. The implementation of the UIndex by Bell ([1995](#_bookmark1)) allows only strata; the implementation used by Onkelinx et al. ([2017b](#_bookmark0)) allows no covariates. We conclude that the current possibilities for using covariates is quite limited in the implementations of UIndex/birdSTATs/ TRIM as examined by Onkelinx et al. ([2017b](#_bookmark0)). Another limitation is that TRIM uses the same model structure for imputation as for analysis, whereas it can be appropriate to use different sets of covariates for imputation and analysis.

1. First of all, we stated that UIndex, as described by Underhill and Prys-Jones ([1994](#_bookmark3)), uses a quasi-Poisson likelihood. Onkelinx et al. ([2017b](#_bookmark0)) implemented it using a negative binomial distribution. We have two reasons for doing so. (1) The simulated data were generated under a negative binomial distribution (Onkelinx et al. [2017b](#_bookmark0), Eq 7). (2) The MI model uses a negative binomial distribution. Using the same distribution for both MI and UIndex eliminates differ- ences due to the selected distribution. We agree with Bogaart et al. ([2017](#_bookmark2)) that TRIM can handle a Poisson distribution with overdispersion (Pannekoek and Van Strien ([2005](#_bookmark4)) do not mention quasi-Poisson explicitly). TRIM is limited to a Poisson distribution with or without overdispersion (Onkelinx et al. ([2017b](#_bookmark0)) selected overdispersion). Figure 5 of Onkelinx et al. ([2017b](#_bookmark0)) should have read ‘overdispersion’ and ‘no overdispersion’ instead of ‘negative binomial’ and ‘Poisson’. birdSTATs would have been placed in the ‘overdispersion’ category.

Summarizing, we agree that Onkelinx et al. ([2017b](#_bookmark0)) misinterpreted some points of UIndex, birdSTATs and TRIM. However, in the case of uncertainty in the docu- mentation, we always chose the variant most favorable for these methods. Hence, we still believe that their methods are at risk of underperforming, even in situations for which they were designed.

References

Bell MC (1995) UINDEX4. a computer program for estimating population index numbers by the Underhill method. The Wildfowl & Wetland Trust, Slimbridge

Bogaart P, van der Meij T, Pannekoek J, Soldaat L, Van Strien AJ, Underhill LG (2017) Comment on ‘Working with population totals in the presence of missing data comparing imputation methods in terms of bias and precision’ by Onkelinx et al. (2016). J Ornithol. doi:[10.1007/s10336-017-1456-5](http://dx.doi.org/10.1007/s10336-017-1456-5)

Onkelinx T, Devos K, Quataert P (2017a) Multimput: using multiple imputation to address missing data. [http://doi.org/10.5281/](http://dx.doi.org/10.5281/zenodo.48423) [zenodo.266429](http://dx.doi.org/10.5281/zenodo.48423). Accessed 15 Feb 2017

Onkelinx T, Devos K, Quataert P (2017b) Working with population totals in the presence of missing data comparing imputation methods in terms of bias and precision. J Ornithol. doi:[10.1007/](http://dx.doi.org/10.1007/s10336-016-1404-9) [s10336-016-1404-9](http://dx.doi.org/10.1007/s10336-016-1404-9)

Pannekoek J, Van Strien A (2005). TRIM 3 Manual (TRends & Indices for Monitoring data). Voorburg

Underhill LG, Prys-Jones RP (1994) Index numbers for waterbird populations. I. Review and methodology. J Appl Ecol 31(3):463–480. doi:[10.2307/2404443](http://dx.doi.org/10.2307/2404443)

van der Meij T (2013) birdSTATs. Species Trends Analysis Tool (STAT) for European bird data. Bioland Informatie, Oegstgeest