Biodiversity Observations

http://bo.adu.org.za

An electronic journal published by the Animal Demography Unit at the University of Cape Town

The scope of Biodiversity Observations consists of papers describing observations about biodiversity in general, including animals, plants, algae and fungi. This includes observations of behaviour, breeding and flowering patterns, distributions and range extensions, foraging, food, movement, measurements, habitat and colouration/plumage variations. Biotic interactions such as pollination, fruit dispersal, herbivory and predation fall within the scope, as well as the use of indigenous and exotic species by humans. Observations of naturalised plants and animals will also be considered. Biodiversity Observations will also publish a variety of other interesting or relevant biodiversity material: reports of projects and conferences, annotated checklists for a site or region, specialist bibliographies, book reviews and any other appropriate material. Further details and guidelines to authors are on this website.

DISPLAYING CHANGES IN BIRD DISTRIBUTIONS BETWEEN SABAP1 AND SABAP2

Les G Underhill and Michael Brooks

Recommended citation format:
Underhill LG & Brooks M 2016. Displaying changes in bird distributions between SABAP1 and SABAP2**.** Biodiversity Observations 7.62: 1–13.

URL: http://bo.adu.org.za/content.php?id=255

Published online: 2 September 2016

– ISSN 2219-0341 –

PROJECT REPORT

DISPLAYING CHANGES IN BIRD DISTRIBUTIONS BETWEEN SABAP1 AND SABAP2

Les G Underhill and Michael Brooks

Animal Demography Unit, Department of Biological Sciences, University of Cape Town, Rondebosch, 7701 South Africa

Email: les.underhill@uct.ac.za

Introduction

One of the objectives of the Second Southern African Bird Atlas Project (SABAP2) was to be able to make comparisons with the distribution data collected during the First Southern African Bird Atlas Project (SABAP1); the spatial grid for SABAP2 was set up with a view to facilitating this (Underhill 2016). Strategies to make these comparisons have proved to be more difficult than were anticipated at the outset of SABAP2 in 2007 (Loftie-Eaton 2014, 2015). The objective of this paper is to document a procedure to construct maps which show the changes in bird distribution patterns between the two projects. There is a companion paper which provides a procedure for mapping distributions using only SABAP2 data (Underhill & Brooks 2016).

The challenges to making comparisons

Three main issues proved challenging for making comparisons between SABAP1 and SABAP2 distributions. (1) The change in spatial scale between the projects. SABAP1 was conducted using a 15-minute grid of longitude and latitude; the SABAP2 is five minutes of latitude and longitude, so that there are nine SABAP2 "pentads" in one SABAP1 "quarter degree grid cell" (Underhill 2016). (2) For SABAP1, there was no quantification of the amount of effort that was invested in a checklist of bird species recorded in a quarter degree grid cell; for SABAP2 observers undertake at least two hours of intensive birding for a "full-protocol" checklist (Underhill 2016). (3) For SABAP1 there was no instruction for the observer to attempt to get a complete checklist of species occurring in the quarter degree grid cell, an instruction which would have been unrealistic, given the size of the grid cell. Whereas in the 9.2×8.3 km pentads used for SABAP2, it is technically feasible to visit all habitats during two or more hours of fieldwork, and the instruction to observers is to make the species list which is as comprehensive as possible (Loftie-Eaton 2014, Underhill 2016).

Loftie-Eaton (2014, 2015) demonstrated that the mean lengths of checklists for SABAP1 and SABAP2 were similar. She suggested that what this meant was that the area effectively searched for birds within a SABAP1 quarter degree grid cell must have been roughly a SABAP2 pentad in extent.

Clearly, maps which aim to show changes in bird distributions have to be at the spatial resolution of the coarser project, in this case SABAP1. This means that the data for the nine SABAP2 pentads within a quarter degree grid cell need to combined in some way. This paper uses the simple expedient of pooling all the checklists for the nine pentads in the SABAP2 a quarter degree grid cell. Effectively, this means that each checklist for a pentad is treated as a checklist for the quarter degree grid cell into which it falls. We compute the SABAP2 reporting rate for a species in a pentad as the ratio "number of times the species was recorded in the quarter degree grid cell" divided by the "total number of checklists received for the pentads in the quarter degree grid cell". The pitfalls of this approach, and indeed the pitfalls of other approaches are considered below in the section **Pitfalls of the expedient approach to pooling**.

Loftie-Eaton (2014, 2015) provided a comprehensive discussion of the hazards associated with the interpretation of reporting rates. In spite

of all the caveats, she concluded: “Reporting rates remain a valuable tool to give broad-brush measures of changes in species’ geographic ranges. Reporting rates can be, and are being, used as an early warning system to detect range changes. Once these changes are detected, further investigation can be done on a species by species level” (Loftie-Eaton 2015).

Statistical approach

We denote the reporting rate for a species in a grid cell by R. Statistically, we can think of R as being dependent on two factors. (1) It varies with the number of birds of the species in the pentad, which we will denote by n. Our belief is that as n increases, so does the reporting rate R. (2) It varies with the detectability of the species. Suppose that there is just one bird of a species in the grid cell and that the probability that an atlaser encounters and identifies this bird is p. For skulking species this probability is relatively small and for conspicuous species it is relatively large.

With R, n and p defined, we establish the relationships between them. The theory underpinning the statistical arguments used here is contained in Underhill & Bradfield (2013, chapter 3). Suppose there is a single bird of a species in the grid cell. Then the probability that the atlaser does not record the species is $1-p$. If there are n birds of the species in the grid cell, then, assuming statistical independence, the probability that the atlaser misses all of them is $(1-p)^n$. The probability that the atlaser encounters at least one bird of the species is then $1-(1-p)^n$. This is precisely what we mean by reporting rate; the species is encountered at least once. Thus the fundamental relationship between reporting rate, detectability and species abundance is given by

$$R = 1-(1-p)^n$$

In reality, the only one of these three quantities for which we have data is the reporting rate. Detectability is what is known in statistics as a “nuisance parameter”; our real interest focuses on the species abundance in the grid cell, denoted n. We rearrange the terms in the fundamental relationship above so that n becomes the subject of the formula: $n = \frac{\log(1-R)}{\log(1-p)}$. The term in the denominator is effectively a conspicuousness factor which adjusts the log-transformed reporting rate to give the population size.

Now suppose we have two atlas projects, and two reporting rates for a species in a grid cell R_1 and R_2. We assume that the detectability of an individual bird of a species in the grid cell remains the same for both projects. Our interest focuses on how the abundance of the species has changed in the grid cell; i.e. we want to estimate n_1 and n_2. Because we do not know the detectability of the species, we cannot estimate these quantities, but we can estimate their ratio.

Thus if $n_1 = \frac{\log(1-R_1)}{\log(1-p)}$ and if $n_2 = \frac{\log(1-R_2)}{\log(1-p)}$, then

$$C = \frac{n_2}{n_1} = \frac{\log(1-R_2)}{\log(1-p)} \Big/ \frac{\log(1-R_1)}{\log(1-p)} = \frac{\log(1-R_2)}{\log(1-R_1)}$$

where C is the estimate of the relative change in density between SABAP1 and SABAP2. In this approach, the neat advantage of the use of the ratio is that the detectability parameter cancels out.

This relationship, $C = \frac{n_2}{n_1} = \frac{\log(1-R_2)}{\log(1-R_1)}$, provides an estimate of the relative change in abundance for a species which has changed in reporting rate in a grid cell from R_1 to R_2. If $C = 1$, then there has been no change in density, if $C < 1$, then the density has decreased during the intervening period and if $C > 1$, then the density has increased.

This relationship is a more defendable "change statistic" than any other functions of the two reporting rates in a grid cell R_1 and R_2. In particular, the difference in reporting rates, $R_2 - R_1$, should not be used, nor should use be made of the relative change in reporting rate, $(R_2 - R_1)/R_1$, and variations of these.

This change statistic can also be derived from the approach used by Griffioen (2001). His method also related abundance to reporting rate, but had an entirely different theoretical starting point, embedded in a spatial process for the distribution of animals, developed by Nachman (1981).

There are multiple possible criticisms relating to the assumptions which underpin the relationship, $C = \frac{n_2}{n_1} = \frac{\log(1-R_2)}{\log(1-R_1)}$, but the biggest potential source of error is the reality that the two reporting rates are frequently based on relatively small numbers of checklists, and are therefore subject to small sample variability. For the purposes of this paper, we only used grid cells for which the number of checklists on which both R_1 and R_2 were based was four or more. If the number of checklists is four and the observed reporting rate is $R = 0.5$, then the 95% confidence interval for the "true" reporting rate is approximately $(0.25, 0.75)$ (Underhill & Bradfield 2013, chapter 11). If there are 10 checklists, then the 95% confidence interval for the "true" reporting rate is $(0.35, 0.65)$ and if there are 30 , then the 95% confidence interval for the "true" reporting rate is $(0.41, 0.59)$. Thus even with samples as large as 30 checklists per grid cell, the sampling variability of the observed reporting rate remains substantial. For more details, see the section entitled **Sampling variability** below.

Ideally, the sample sizes on which reporting rates to estimate relative changes in abundance should be based on many more than four checklists. The reality is that, for many grid cells, the sample sizes are small. This is particularly true for SABAP1, and it is no longer possible to increase these sample sizes. The consequences of the decision to use a minimum of four checklists per grid cell for both SABAP1 and SABAP2 becomes clear in the examples below.

Colour shading protocol for range-change maps

For each species, the value of C is computed for each quarter degree grid cell. These values are colour-coded and plotted on the "range-change map" for the species. Five cut points are used to generate a six-colour map: 0 < RED <0.33 < ORANGE < 0.67 < YELLOW < 1 < LIGHT GREEN < 1.5 < DARK GREEN < 3 < BLUE. In other words, grid cells are shaded RED if the abundance in SABAP2 is estimated to be less than one-third of the abundance in SABAP1, ORANGE if the abundance in SABAP2 is between one-third and two-thirds of the abundance in SABAP1, YELLOW if the abundance is between unchanged and a decrease of one-third. LIGHT GREEN indicates an increase of up to 1.5 times of the SABAP1 population; DARK GREEN represents an increase of between 1.5-fold and three-fold, BLUE represents a more than three-fold increase in abundance between SABAP1 and SABAP2. In broad brush terms, RED, ORANGE and YELLOW represent grid cells with large, moderate and small relative decreases, and BLUE, DARK GREEN and LIGHT GREEN represent grid cells with large, moderate and small relative increases.

A key factor in interpreting these range-change maps is that the value of C is to grasp that it is a **relative** value. A grid cell may correctly be shaded BLUE, and the change in relative abundance between SABAP1 and SABAP2 may be genuinely large, but the increase is off a low baseline, a tiny reporting rate in SABAP1. So the fact that a grid cell is shaded BLUE does not lead to the conclusion that the species is now abundant in the grid cell. The only inference is that there has been a large increase in relative abundance in the grid cell. This is true even in grid cells with none of the difficulties caused by sampling variation in the reporting rates for SABAP1 and SABAP2.

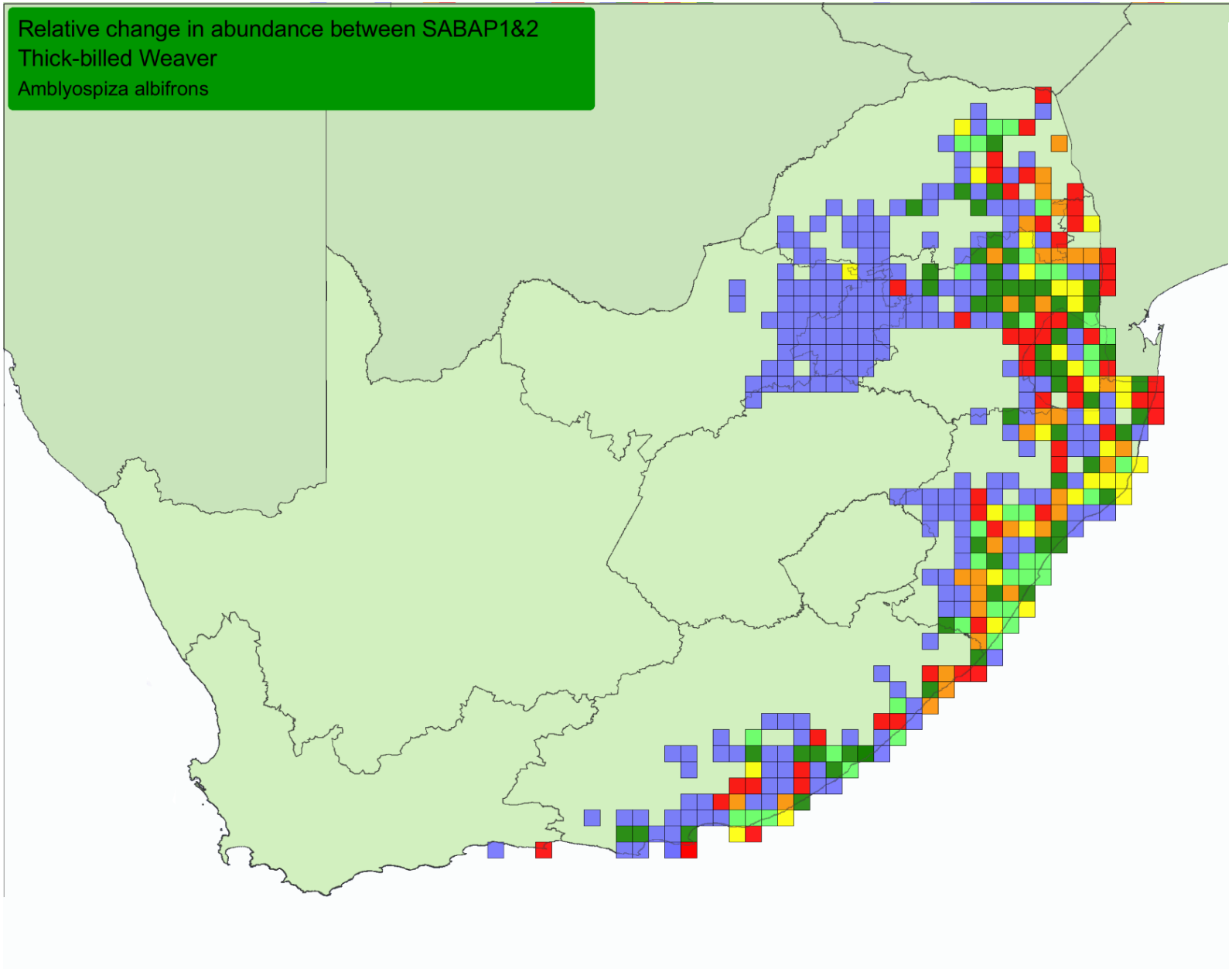

Figure 1. Range-change map between SABAP1 and SABAP2 for the Thick-billed Weaver in South Africa, Lesotho and Swaziland. RED, ORANGE and YELLOW represent quarter degree grid cells with large, moderate and small relative decreases, and BLUE, DARK GREEN and LIGHT GREEN represent grid cells with large, moderate and small relative increases

Figure 2. Thick-billed Weaver, Oribi Gorge Nature Reserve, KwaZulu-Natal 24 December 2014. Photograph: © Lia Steen.
http://vmus.adu.org.za/?vm=BirdPix-12960

Examples

The examples are developed to illustrate both the strengths and pitfalls of these new-generation range-change maps.

Thick-billed Weaver

The Thick-billed Weaver *Amblyospiza albifrons* is a species which started expanding its range westwards in the early 1960s (Craig 1997) (Figures 1 and 2). It was first recorded in what was to become Gauteng at the Melrose Bird Sanctuary in 1961; it was initially thought that the founder population consisted of escapees from aviaries, but it was subsequently considered more likely that this was a natural range expansion, with the Olifants River being used as a corridor from the Lowveld (Tarboton et al. 1987). The SABAP1 distribution map supported this hypothesis (Craig 1997). The SABAP1-SABAP2 range-change map for the species increases almost everywhere along the western edge of its range, and that the remarkable expansion across Gauteng into the adjacent provinces has continued (Figure 1). This outcome is unambiguous from the spatially continuous blue shading of a large number of contiguous quarter degree grid cells.

Along the eastern edge of the range, in the Kruger National Park, Swaziland and northern KwaZulu-Natal, grid cells shaded red, orange and yellow predominate. A quick interpretation will suggest that the species has decreased in abundance in these regions but this needs to be considered (1) in relation to the number of checklists per grid cell and the possibility that the pattern is a consequence of sampling variability due to small sample sizes (Figure 3) and (2) in relation to the actual reporting rates for SABAP1 and SABAP2 in the grid cells (Figure 3).

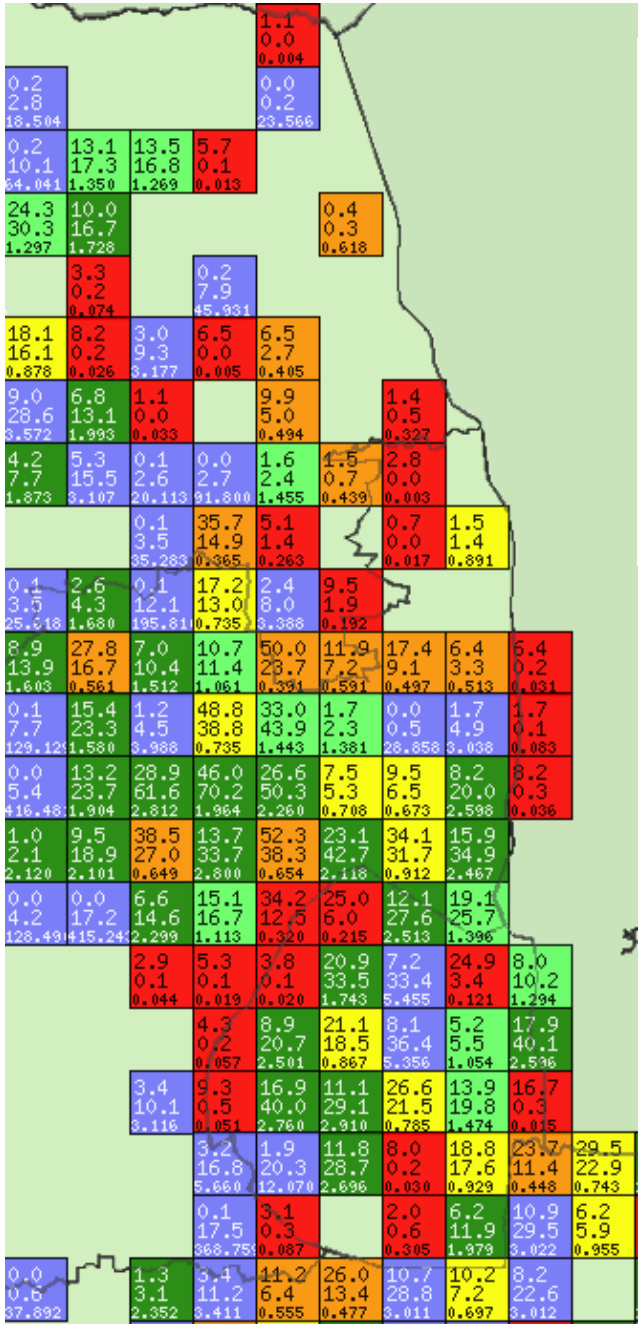

Figure 3. Annotated range change map for the Thick-billed Weaver in north-eastern South Africa, including the Kruger National Park. In each quarter degree grid cell, the top number is the SABAP1 reporting rate, the middle number is the SABAP2 reporting rate, and the bottom number is the associated value of C, the estimated relative change in abundance C between the two projects

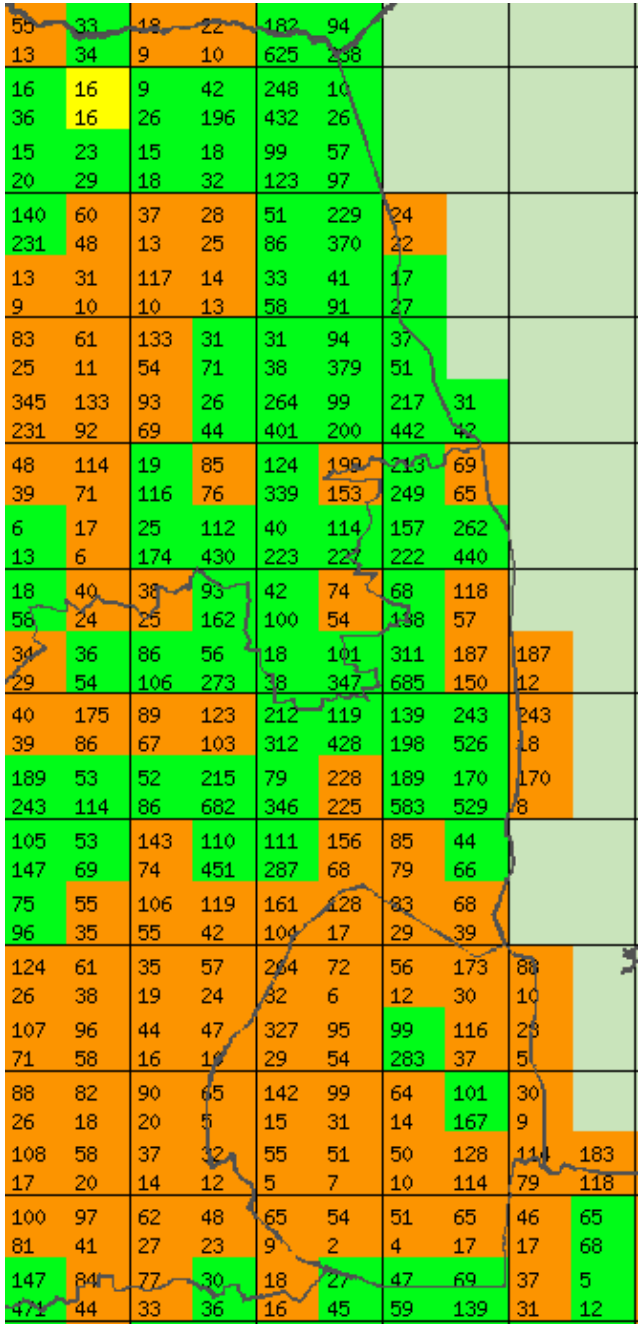

Figure 4. North-eastern South Africa, the same region as Figure 3. The top number in each quarter degree grid cell is the number of checklists submitter for the quarter degree grid cell during SABAP1. The bottom number is the total number of SABAP2 checklists submitted for the (usually) nine pentads in the quarter degree grid cell. Grid cells shaded green have more SABAP2 lists in total for the pentads in the grid cell than there were SABAP1 lists. Vice versa for pentads shaded orange

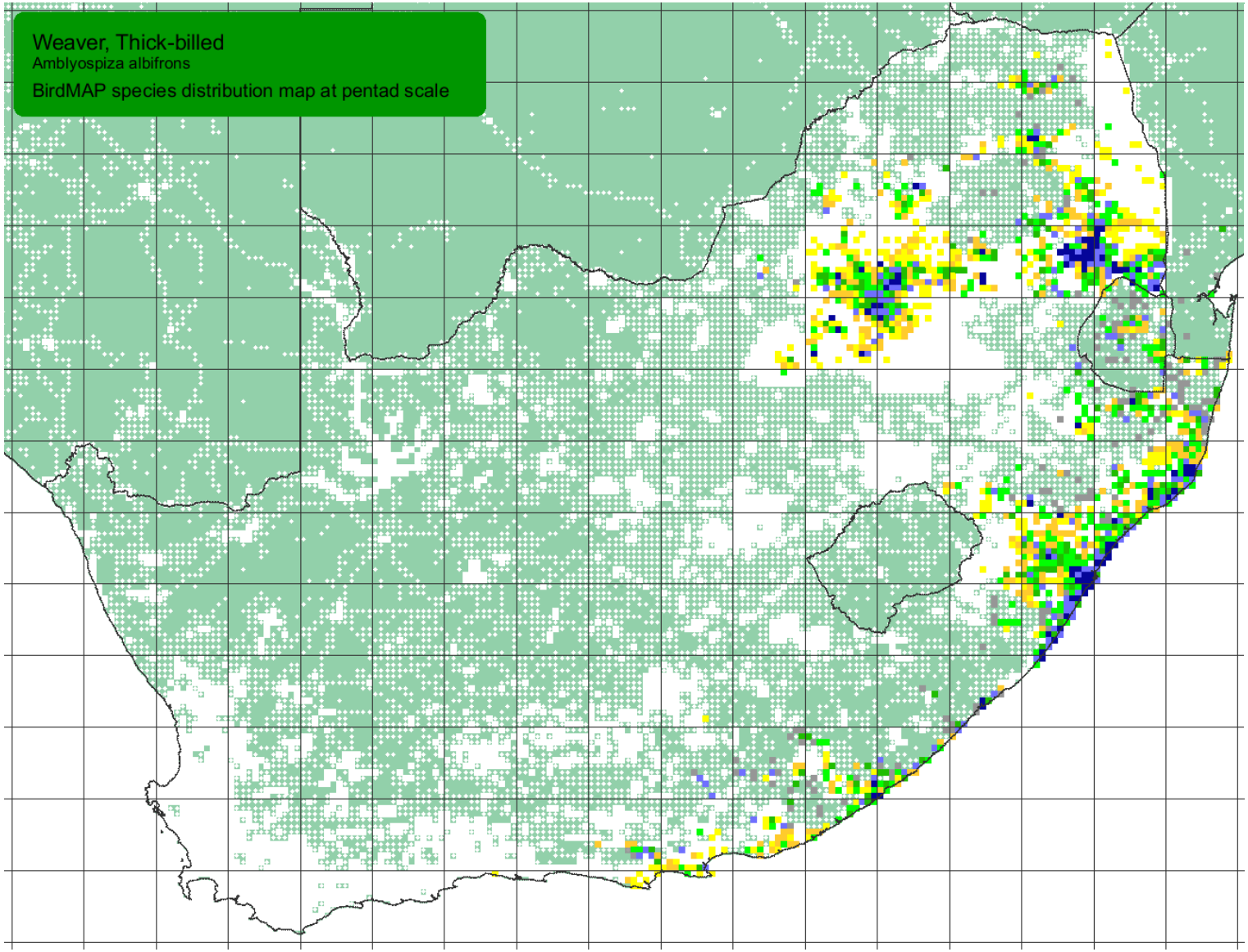

Figure 5. Pentad-scale distribution map for Thick-billed Weaver. The interpretation guidelines are in Underhill & Brooks (2016). Pentads with four or more checklists are in WHITE (not recorded), or in colours ranging from YELLOW, through ORANGE, LIGHT GREEN, DARK GREEN, LIGHT BLUE and DARK BLUE, indicating increasing reporting rates. Pentads shaded GREY have the species present, but there are too few checklists to justify a reporting rate. There are bird records from pentads with WHITE DOTS, but Thick-billed Weaver was not recorded

It is also informative to interpret the range-change map in relation to the pentad-scale distribution map for the species (Underhill & Brooks 2016); the most striking feature of this map for the Thick-billed Weaver is that the actual distribution map (Figure 5) looks far smaller than the distribution shown in the range-change map (Figure 1), even though in this case the species is expanding its range. This is a by-product of changing the mapping scale from the quarter degree gird cell to the pentad.

Pied Crow

For the Pied Crow *Corvus albus* (Figure 6) we first present the SABAP2 pentad scale distribution map (Figure 7). The pentads shaded light blue and dark blue represent the core of the range of the species (Underhill & Brooks 2016). There are not yet four checklists for the pentads shaded grey, and reporting rates are not calculated for them; at least part of this area is likely to prove to be core range (Underhill & Brooks 2016). The striking feature of this distribution map is the high reporting rates of Pied Crow in the Swartland north of Cape Town, and the relatively low reporting rates in the Overberg east of Cape Town (Figure 7). Overall the Pied Crow distribution is a remarkable patchwork of discrete area of high and low reporting rates. This pattern was also noted in SABAP1 (Jenkins & Underhill 1997).

Taken in isolation of Figure 7, for the western half of South Africa, the range-change map for the Pied Crow (Figure 8), is predominantly the three colours associated with increases: light green, dark green and blue (Figure 6). In contrast, in northeastern South Africa, the pattern is more complex with large regions apparently showing consistent decreases and other regions consistent increases. The largest area of decrease is from the border of Botswana across North West Province and the Free State to the Lesotho border. The largest increases appear to be in the grasslands of the eastern Free State and adjacent Mpumalanga.

Figure 6. Pied Crow on the shore of Lake Malawi, Malawi. 26 April 2015. Photograph: © Gary Brown. http://vmus.adu.org.za/?vm=BirdPix-17240

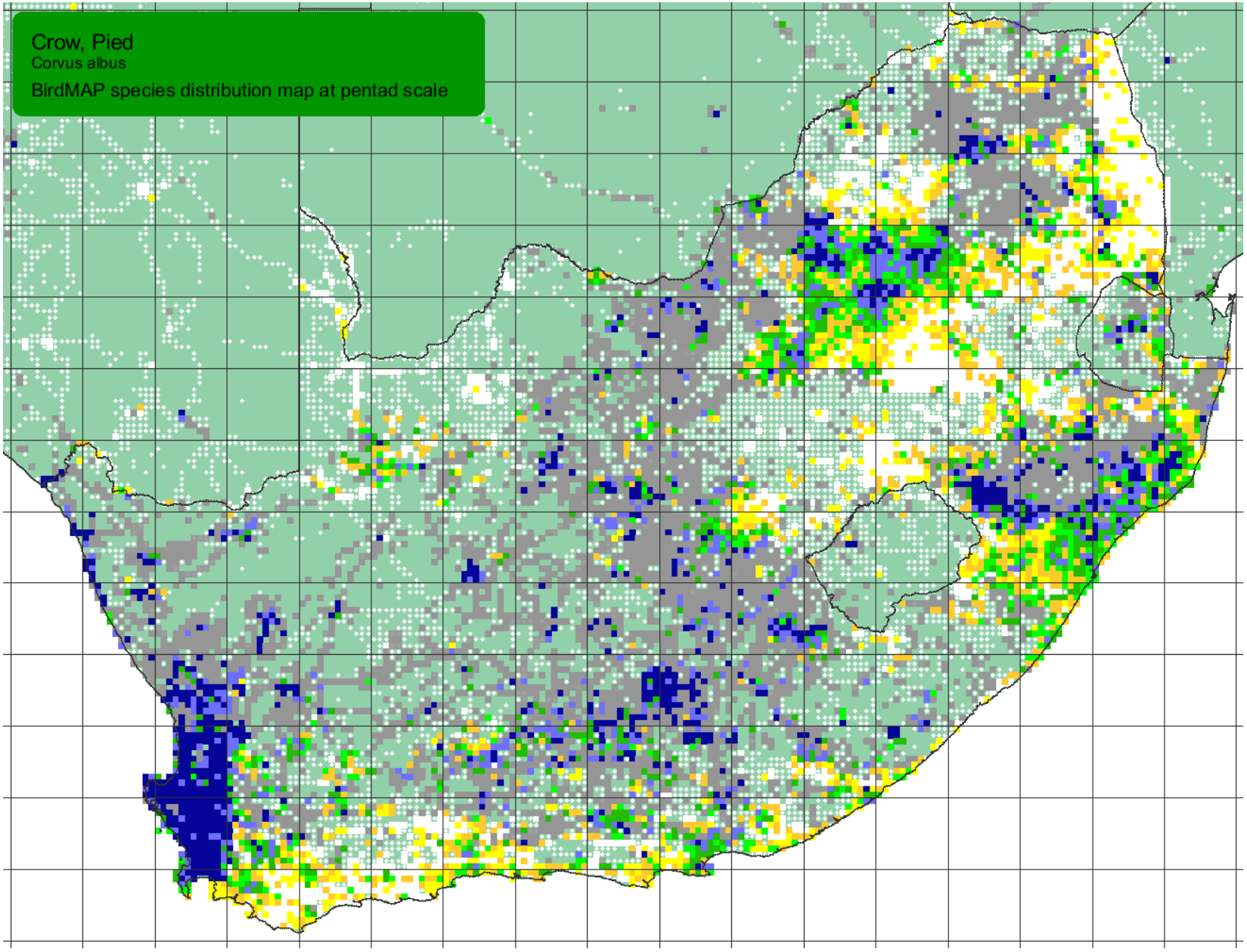

Figure 7. Pentad-scale distribution map for Pied Crow. The interpretation guidelines are in Underhill & Brooks (2016). Pentads with four or more checklists are in WHITE (not recorded), or in colours ranging from YELLOW, through ORANGE, LIGHT GREEN, DARK GREEN, LIGHT BLUE and DARK BLUE, indicating increasing reporting rates. Pentads shaded GREY have the species present, but there are too few checklists to justify a reporting rate. There are bird records from pentads with WHITE DOTS, but Pied Crow was not recorded

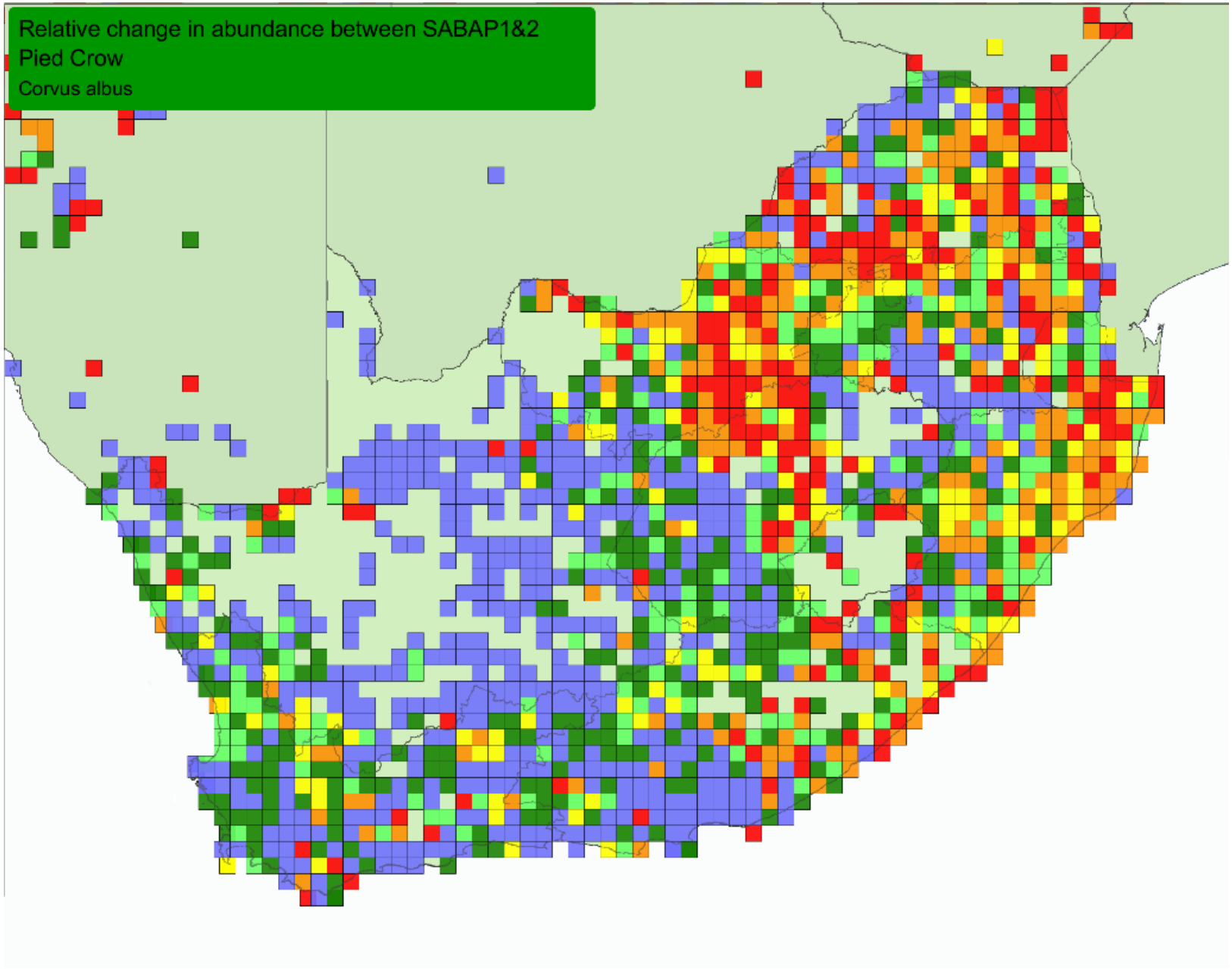

Figure 8. Range-change map between SABAP1 and SABAP2 for the Pied Crow. RED, ORANGE and YELLOW represent quarter degree grid cells with large, moderate and small relative decreases, and BLUE, DARK GREEN and LIGHT GREEN represent grid cells with large, moderate and small relative increases

At first glance, Figures 7 and 8 do not obviously relate to the same species. The range-change map (Figure 8) suggests that there have been increases in relative abundance in both the Swartland and Overberg, but in the Overberg, and along the southern coast as far as into the Eastern Cape, this increase has been off a low baseline (Figures 7 and 8). The increase in the Northern Cape, in the Upington region, also appears to have been off a low base, because the pentads in this region with four or more checklists are shaded mostly orange and green (Figure 7), indicating that reporting rates have increased to values close to the median reporting rate for the species (Underhill & Brooks 2016). Understanding how the range of the Pied Crow has changed requires careful and thoughtful comparison of both the distribution map (Figure7) and the range-change map (Figure 8).

Pitfalls of the expedient approach to pooling

The expedient approach, described above, and also used in the examples presented here, simply pools all the SABAP2 checklists for the quarter degree cell. With this approach, some pentads have many checklists, and others only a few. Loftie-Eaton (2015) synthesized some of the pitfalls of this approach, based on four studies published in *Ornithological Observations* (McKenzie 2011, Carter 2012, de Swardt 2012, Retief 2013). The main insight from these papers is that the comparison narrative for each of these quarter degree grid cells raised different concerns in making comparisons. More papers, undertaking comparisons between SABAP1 and SABAP2 for a particular quarter degree grid cell, would be an extremely valuable contribution to our understanding of the variety of factors that are involved.

Probably one of the worst examples of the problems associated with the expedient approach is for the quarter degree grid cell Cape Town 3318CD. The quarter degree grid cell contains six pentads (three are entirely in the ocean). The only land in one of these six pentads, 3345_1820, is Robben Island. In August 2016, this pentad has 184 checklists. The other five pentads had 79, 62, 119, 44 and 574 checklists. Robben Island was inaccessible at the time of SABAP1, but the colony of African Penguins *Spheniscus demersus* had already started. The SABAP1 reporting rate for African Penguin in the quarter degree grid cell was 4.0% (this would have been generated by penguins observed from the mainland shore, which occurs under good visibility from vantage points such as the Sea Point Promenade, and also birds found dead on the shoreline of the grid cell); the SABAP2 reporting rate was 18.5%, influenced by the large number of checklists for SABAP2 from Robben Island (generated mainly by ADU postgraduate students as a by-product of their own research fieldwork). If the Robben Island pentad is omitted, the SABAP2 reporting rate was 1.7%, but this was biased downwards because pentad 1825_3355 has no coastline, and none of its 574 checklists record African Penguins. The reporting rate for the four mainland coastal pentads was 4.9% (15 records on 304 checklists). This problem does not only affect African Penguins; every species which is common on the mainland, but does not occur on Robben Island, will have its reporting rate depressed by the large set of lists from the Robben Island pentad.

An alternative approach is to compute the reporting rates for each pentad in the quarter degree cell, and then to average these reporting rates for each species. The apparent advantage is that each pentad in the quarter degree grid cell then has equal weight. As attractive as this solution appears, it cannot be implemented at present, because many pentads only have a single checklist, and this “equality of pentads” algorithm gives too much emphasis to pentads with little data. A more fundamental reason why this approach is flawed is that the SABAP1 data were not collected giving equal weight to each section of the pentad. The more easily accessible sections effectively had more visits (and therefore weight) than the inaccessible sections. The expedient approach to smoothing problem probably captures the

behaviour of SABAP1 atlasers more closely than the equal weights approach.

Sampling variability

If sampling variability did not exist, every family of four would have two girls and two boys. But the reality is that some families of four have no boys, and some have no girls. In fact, the percentages of 0, 1, 2, 3 and 4 girls in a family of four are 6%, 25%, 38%, 25% and 6%. So if the "true" reporting rate for a species in a pentad is really 50%, and if we get four checklists, it is only 38% of the time that the observed reporting rate for the species will actually be 50%. That is what sampling variability does to us. It is not a consequence of poor fieldwork, it is just a fact of life. The impact of sampling variability gets smaller as the sample size increases. The observed reporting rate slowly gets closer to the true reporting rate as the sample size increases. The consequence of sampling variability is that there will be some grid cells in which the observed reporting rates for the two projects are very different when in reality the true underlying reporting rate has not changed at all.

In practical terms, what this means for the range-change maps, such as Figures 1 and 8, is that the results for a single grid cell cannot be taken as gospel for that grid cell. Especially if the number of checklists for either SABAP1 or SABAP2 for that grid cell is small, the result may simply be a casualty of sampling variability.

However, if the bulk of the grid cells in a region show an either upwards or downwards trend, then it is highly unlikely that all of them can be dismissed as being a consequence of sampling variability. The harsh reality is that a few red cells will be scattered in among a carpet of blue grid cells (and *vice versa*). The grid cell might be showing a different trend to its neighbours; a more likely explanation is that sampling variability impacted that particular grid cell, and that it is actually showing the same pattern as its neighbours.

Acknowledgements

SABAP2 is a partnership project between SANBI, BirdLife South Africa and the Animal Demography Unit at the University of Cape Town. The enthusiastic participation of 1000s of citizen scientists is fundamental to the production of these range-change maps. Without their contributions, there could not be a paper like this.

References

Carter J 2012. Somerset West – SABAP1 and SABAP2 compared. Ornithological Observations 3: 253–267. Available online at http://bo.adu.org.za/content.php?id=72

Craig AJFK 1997. Thickbilled Weaver *Amblyospiza albifrons*. In: Harrison JA, Allan DG, Underhill LG, Herremans M, Tree AJ, Parker V, Brown CJ (eds) The atlas of southern African birds. Vol. 2: Passerines. BirdLife South Africa, Johannesburg: 548–549.

de Swardt D 2012. 2925BD Hagesdam – SABAP1 and SABAP2 compared. Ornithological Observations 3: 109–122. Available online at http://bo.adu.org.za/content.php?id=51

Griffioen P 2001. Temporal changes in the distributions of bird species in eastern Australia. PhD thesis, La Trobe University, Bundoora, Victoria, Australia.

Jenkins AR, Underhill LG 1997. Pied Crow *Corvus albus*. In: Harrison JA, Allan DG, Underhill LG, Herremans M, Tree AJ, Parker

V, Brown CJ (eds) The atlas of southern African birds. Vol. 2: Passerines. BirdLife South Africa, Johannesburg: 104–105.

Loftie-Eaton M 2014. Geographic range dynamics of South Africa's bird species. MSc thesis, University of Cape Town, South Africa available online at http://www.adu.uct.ac.za/sites/default/files/image_tool/images/352/students/Loftie-Eaton_M_2014_MSc_thesis.pdf

Loftie-Eaton M 2015. Comparing reporting rates between the First and Second Southern African Bird Atlas Projects. Ornithological Observations 6: 1–11. Available online at http://bo.adu.org.za/content.php?id=163

McKenzie D 2012. Reporting rate comparisons for birds in the Nelspruit area – SABAP1 vs SABAP2. Ornithological Observations 2: 22–31. Available online at http://bo.adu.org.za/content.php?id=17

Nachman G 1981. A mathematical model of the functional relationship between density and the spatial distribution of a population. Journal of Animal Ecology 50: 453–460.

Retief E 2013. 2528AB Pienaarsrivier – Analysis of SABAP1 and SABAP2 data for the quarter degree square. Ornithological Observations 4: 107–118. Available online at http://bo.adu.org.za/content.php?id=92

Tarboton WR, Kemp MI, Kemp AC 1987. Birds of the Transvaal. Transvaal Museum, Pretoria.

Underhill LG 2016. The fundamentals of the SABAP2 protocol. Biodiversity Observations 7.42: 1–12. Available online at http://bo.adu.org.za/content.php?id=235

Underhill LG, Bradfield D 2013. IntroSTAT. Department of Statistical Sciences, University of Cape Town, Cape Town. Available online at https://open.uct.ac.za/bitstream/handle/11427/4150/INTROSTAT_ebook.pdf?sequence=1

Underhill LG, Brooks M 2016. Pentad-scale distribution maps for bird atlas data. Biodiversity Observations 7.52: 1–8. Available online at http://bo.adu.org.za/content.php?id=245