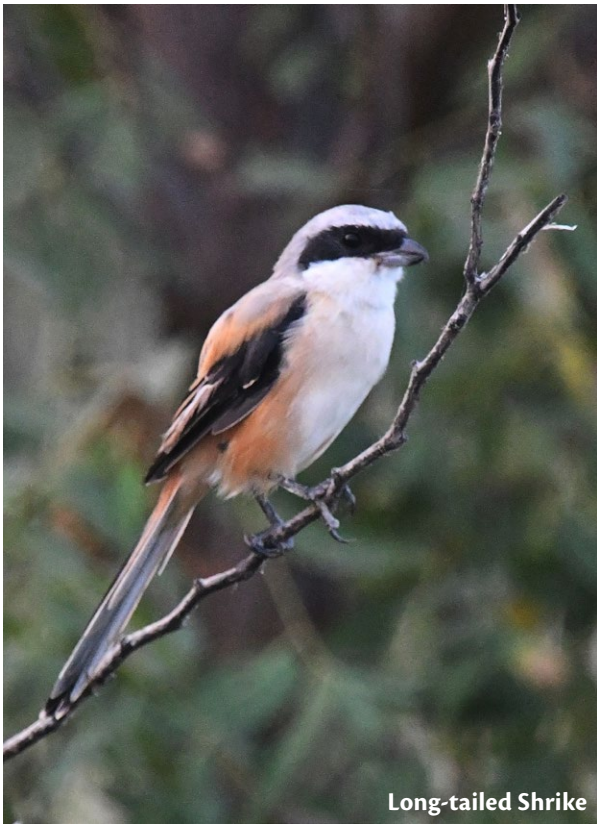
Evaluation of eBird data for Long-tailed Shrike *Lanius schach* and Bay-backed Shrike *Lanius vittatus* from Gujarat: a case study

Prasad Ganpule: C/o Parshuram Pottery Works, Opp. Nazarbaug Station, Morbi 363642. Email: prasadganpule@gmail.com



Photos: Prasad Ganpule

'The Bay-backed Shrike depicted here is not a full adult; note that the mask above the eye is not broad across the forehead, and such birds can be easily mistaken for Long-tailed Shrikes. But note the different structure (especially the longer tail in Long-tailed Shrike) and plumage (large white primary patch and darker chestnut-maroon mantle in Bay-backed Shrike). Identification in such cases remains tricky.

Introduction

Citizen science projects like 'eBird' (www.eBird.org) are fast gaining popularity in India. eBird is now one of India's most widely used citizen science databases and is also popular with bird watchers in Gujarat. More than 122000 checklists are now submitted by more than 5000 eBird users for Gujarat (eBird 2024). This database is a very useful source of information and is used in several ways. For example, eBird data was used to assess the conservation status of most species occurring in India in the 'State of India's Birds' report, wherein more than 30 million observations formed the basis of the analyses (SoIB 2023). The data from eBird is also used to assess the distribution of bird species; the winter distribution of Forest Wagtail *Dendronanthus indicus* in India was assessed based mainly on sightings reported on eBird (Kannan et al. 2018). The eBird website lists over 930 publications (https://science.ebird.org/en/research-and-conservation/publications), highlighting how eBird data is used. However, since it is a citizen science project, whether the data is reliable and accurate must be verified. A critical assessment of reported records/observations from time to time can shed more light on the accuracy of the data posted on eBird.

Here, I evaluate the records of two resident shrike species, the Long-tailed Shrike *Lanius schach* and the Bay-backed Shrike *Lanius vittatus* from Gujarat.

Material and methods

I collected photographic reports of Long-tailed Shrikes and Bay-backed Shrikes posted on eBird from Gujarat till 28 May 2024. There are 460 photographs of Long-tailed Shrike and 657 photographs of Bay-backed Shrike on eBird until this date. Only confirmed sightings were selected. This data was then exported onto an Excel spreadsheet. Since multiple photographs of the same individual were on many checklists, the data was filtered, and the number of checklists submitted was obtained. Every checklist with more than one photograph was scrutinised to see whether more than one individual was photographed. Since most checklists

Table 1: Details of identification of Long-tailed Shrike and Bay-backed Shrike on eBird from Gujarat

	No. of photographs	No. of checklists	No. of individuals	No. of misidentifications	No. of correct identifications	Percentage of correct identification
Long-tailed Shrike	460	324	336	26	310	92.20%
Bay-backed Shrike	657	397	415	28	387	93.25%

had only one photograph, it was relatively easy to check all the individual checklists with multiple photographs. When more than one bird was present in the photos, each such bird was added to the total number of birds. It was not checked whether there was a duplication of records, i.e. whether the same bird was photographed more than once on different dates from the same area. The counts in the checklists were ignored, and only the photographs posted on each checklist were seen. All photographs were checked for identification, and the number of misidentified birds was counted. If a photograph was too poor to check for correct identification, then the bird present in such a photo was counted as misidentified; the number of such poor photos was very low (only a total of one or two in each species). I did not check if the same observer had made multiple mistakes in identification, and only the total number of misidentified birds was counted.

Results

I found that 310 out of 336 individuals of Long-tailed Shrike were correctly identified, and 387 out of 415 individuals of Bay-backed Shrike were correctly identified. Therefore, the identification accuracy was 92.20% in Long-tailed Shrike and 93.25% in Bay-backed Shrike. See Table 1 for details.

It was interesting to note that most of the misidentifications were reciprocal; the Long-tailed Shrike was misidentified as a Bay-backed Shrike and vice versa. The most extreme identification mistake was a prinia species misidentified as a shrike. Other misidentifications were Red-backed Shrike *Lanius collurio* and Brown Shrike *Lanius cristatus* misidentified as Bay-backed Shrike or Long-tailed Shrike. Some Bay-backed Shrikes, which were in moult and difficult to identify, were correctly identified. Some of the identification errors were in identifying juveniles. But, the overall skill in identifying these species was quite good. There is an option to use a slash – Bay-backed/Long-tailed Shrike – in eBird if a user is unsure of the identification. However, no photographs have been posted in this category.

Though my analysis was focused on identification errors, the data bias in the distribution of the Bay-backed Shrike and Long-tailed Shrike in Gujarat was evident from this eBird data. Though both these species are widespread and can be frequently found in suitable habitats all over the state, there are a few parts of the state from which there are no records or only some records in eBird. Further, there are a disproportionate number of records from 'hotspots'. This can be easily explained since more bird watchers visit these hotspots and upload their sightings while some of the interior parts of the state are not visited. However, this creates a situation where it would seem that these species are not present in some parts of the state. The fact is that these species occur in these areas, but this is not seen in the eBird data. Hence, the spatial (location) bias is a problem, and wider coverage is needed from bird watchers to get a correct distribution of these species.

Discussion

Checking the data on eBird showed that the misidentifications were 7.8% in Long-tailed Shrike and 6.75% in Bay-backed Shrikes in the state. The review system in eBird is based on filters which flag unusual sightings or large counts. These flagged sightings are then subjected to expert review. eBird users can also flag wrong identifications from public outputs. In the case of the Long-tailed Shrike and the Bay-backed Shrike, both these species are common in Gujarat, and hence, the filters do not flag reports of these species and only unusual counts are flagged. One or two birds of these species being reported would not be flagged since the filters are set considering the status and distribution of these species in Gujarat.

I am one of the eBird reviewers for Gujarat and have been associated with eBird as a reviewer since 2014. While reviewers and other birdwatchers try their best to keep the data accurate, the sheer volume of data in eBird makes it very difficult to maintain the accuracy of the data. Reviewers will look into and correct these identification errors for Long-tailed Shrike and Bay-backed Shrike in due course. However, new records will be added; correcting misidentified birds is a continuous process.

If there are only occasional errors in the data, then the impact of such errors may be negligible when this data is used in different analyses. However, it is well known that in similar-looking species, the errors in identifications are higher, meaning that the error rate increases when birds co-occur

and appear similar (Hull et al. 2010). If there are abundant and systematic identification errors, problems arise (Costa et al. 2015). It is well known that there are data biases in eBird; see Tang et al. (2021) for modelling spatially biased citizen science efforts. Similarly, Jhonston et al. (2019) suggest best practices for making reliable inferences from citizen science data to estimate species distributions. Ramesh et al. (2017) and Praveen (2017) discuss the sampling and spatial bias in species distribution and home ranges. In some cases, the misidentifications can bias the phenological estimates; Gorleri & Areta (2022) found that phenological estimates for two hard-to-identify *Elaenia* flycatchers in South America were biased due to a large number of misidentifications between these species. The Kerala Bird Atlas 2015-2020 used statistical methods to minimise temporal, spatial or taxonomic bias but recognised that some difficult-to-identify species pairs like Green Warbler *Phylloscopus nitidus* / Greenish Warbler *Phylloscopus trochiloides* were likely to be misidentified. Their solution was to merge the records into a single taxon; Greenish Warbler records were considered to be of Green Warbler (doi: 10.18520/cs/v122/i3/298-309). Similarly, SoIB (2023) combined records of 17 species pairs, which were considered hard to identify; in this case, Green Warbler records were merged into Greenish Warbler!

Since eBird data is now widely used, its reliability and accuracy should be periodically checked. The misidentifications in these two shrike species were less than 10% in checked photographs. In this case, to inexperienced birders, both species are similar in appearance, so the chances of making mistakes in identification are higher. The error percentage for those sightings that did not have photographs is impossible to check, but it can be assumed to be generally less than 10% based on the examined photographic data and assuming a similar error rate. The error percentages may be lower in easier-to-identify species. Conversely, likely, the error percentages in difficult-to-identify species like *Aquila* eagles, *Phylloscopus* and *Iduna* warblers, Pipit species, and many other groups would be higher. Smart filters may be able to catch the errors, but identification remains challenging in many species, even from photographs.

It should be noted that this is only a rudimentary analysis of two shrike species from Gujarat. This analysis can be further refined by checking if an observer had made repeated mistakes in identification or whether the errors were evenly distributed across observers. It would be interesting to see whether the error percentage is similar if the data for the whole country is checked; there are 13915 photographs of Long-tailed Shrike and 7670 photographs of Bay-backed Shrike from India till 31 May 2024 and checking all the photographs for correct identification would be a herculean task! But it might be easier to do it state-wise and see if there are similar or better results. Similar species can be selected, and the identification errors in such species pairs can be analysed to determine the reliability of eBird data for difficult-to-identify species. More complex analyses to identify errors can be undertaken on a wider scale to evaluate the accuracy and reliability of eBird data.

The bigger issue is whether the data from eBird would give correct outputs if there are identification errors on a large scale and what would be the best way to tackle such errors. What is the percentage of error after which the data would generate biased outputs? What would be the best way to minimise and correct identification errors? These are some of the questions which need more studies. Since eBird data is used in trends analysis, species distribution, conservation applications, population estimates and other related fields, it becomes imperative to check whether the data is reliable and accurate. Different statistical models may be useful in mitigating identification errors, and the data's accuracy can be improved by imparting basic identification skills to beginners and regular evaluation of photographic records by a dedicated team of reviewers.

Conclusion

This is a first-of-its-kind study on evaluating photographic data from eBird in India. The identification accuracy for both Long-tailed Shrike and Bay-backed Shrike was more than 90% based on photographs, but identification of both these species is challenging and results in misidentifications. Maintaining the accuracy and reliability of eBird data is difficult due to the volume of the data being added almost daily. Regular reviews of photographs of confusing species posted on eBird are recommended. Further, it is crucial to check for large-scale errors in identifying some species and take corrective measures before the data is used in various analyses. More studies on data evaluation of similar-looking species from India need to be initiated.

Acknowledgements

I am grateful to Maulik Varu for reviewing the draft manuscript.

References

Costa, H., Foody, G. M., Jiménez, S., & Silva, L., 2015. Impacts of Species Misidentification on Species Distribution Modeling with Presence-Only Data. *ISPRS International Journal of Geo-Information* 4 (4): 2496-2518. https://doi.org/10.3390/ijgi4042496.

eBird. 2024. eBird: An online database of bird distribution and abundance [web application]. eBird, Ithaca, New York. Available: http://www.ebird.org (Accessed: 28 May 2024).

Gorleri, F. C., & Areta, J. I. 2022. Misidentifications in citizen science bias the phenological estimates of two hard-to-identify Elaenia flycatchers. *Ibis* 164(1): 13–26.

Hull, J. M., Fish, A. M., Keana, J. J., Mori, S. R., Sacks, B. N. & Hull, A. C. 2010. Estimation of Species Identification Error: Implications for Raptor Migration Counts and Trend Estimation. *The Journal of Wildlife Management* 74: 1326-1334. https://doi.org/10.1111/j.1937-2817.2010.tb01254.x

Johnston, A., Hochachka, W, M., Strimas-Mackey, M, E., Ruiz Gutierrez, V., Robinson, O, J., Miller, E, T., Auer, T., Kelling, S, T., & Fink, D. 2019. Best practices for making reliable inferences from citizen science data: case study using eBird to estimate species distributions. *Diversity and Distributions*. doi: 10.1111/ddi.13271.

Kannan, R., Santharam, V., Kannan, A., & Nagarajan, V. M. 2018. True winter distribution of the Forest Wagtail *Dendronanthus indicus* in India. *Indian BIRDS* 14 (2): 33–36.

Praveen, J. 2017. "On the geo-precision of data for modelling home range of a species – A commentary on Ramesh et al. (2017)." *Biological Conservation* 213: 245-246.

Ramesh, V., Gopalakrishna, T., Barve, S., & Melnick, D. J. 2017. "Finer spatial resolution improves accuracy of species distribution models in heterogeneous landscapes – A response to Praveen J." *Biological Conservation* 213: 247-248.

SoIB 2023. *State of India's Birds, 2023: Range, trends, and conservation status*. The SoIB Partnership. Pp. 119.

Tang, B., Clark, J. S., & Gelfand, A. E., 2021. Modelling spatially biased citizen science efforts through the eBird database. *Environmental and Ecological Statistics* 28 (3): 609–630. □

Sighting of Chestnut-winged Cuckoo *Clamator coromandus* at Vansda National Park – A second record for the State

Dharmesh Patel: 7/B Jai Jalaram Society, Opp. V. S Patel College, Bilimora 396321. Dist: Navsari. Email: drdharmesh202@gmail.com

Photo: Dharmesh Patel

On the morning of March 26th 2024, I was on a birding trip in the Bharadi area of Vansda National Park (20°46'05.7"N 73°28'10.0"E). Around 8 am, I observed a bird with reddish wings flying across the trail and perching on a bamboo branch. Initially, I considered it might be a Greater Coucal *Centropus senensis* due to its colours. But it appeared a bit smaller and to me, against the light, it resembled a Jacobin Cuckoo *Clamator acobinus*. However, upon reviewing my photographs, I concluded it as a Chestnut-winged Cuckoo *Clamator coromandus*. The bird remained perched for a few seconds before flying into the dense canopy. Additionally, I had spotted nearly 10 Forest Wagtails *Dendronanthus indicus*, in the vicinity.

Upon searching eBird database, I found no prior records of the Chestnut-winged Cuckoo in Gujarat. However, Mr. Ashok Mashru, a senior birdwatcher, informed me that this sighting is significant as it represents the second record of this species in Gujarat. The first sighting was documented by Prashant Desai on July 10th 2005 in Vapi, and the note was published in the Flamingo newsletter (Desai 2005). According to 'Field Guide to the Birds of Gujarat' (Ganpule 2022), status of Chestnut-winged Cuckoo in Gujarat is listed as 'vagrant' and there is only a single record in the state. This makes my observation particularly noteworthy, as it marks the bird's vagrant visit to the state after nearly two decades.

The Chestnut-winged Cuckoo is distinguished by its striking reddish-brown wings and a slender built. It is known to