PLATINUM OPEN ACCESS

The Journal of Threatened Taxa (JoTT) is dedicated to building evidence for conservation globally by publishing peer-reviewed articles online every month at a reasonably rapid rate at www.threatenedtaxa.org. All articles published in JoTT are registered under Creative Commons Attribution 4.0 International License unless otherwise mentioned. JoTT allows unrestricted use, reproduction, and distribution of articles in any medium by providing adequate credit to the author(s) and the source of publication.

Journal of Threatened Taxa

Building evidence for conservation globally

www.threatenedtaxa.org

ISSN 0974-7907 (Online) | ISSN 0974-7893 (Print)

Communication

Evaluating performance of four species distribution models using Blue-tailed Green Darner *Anax guttatus* (Insecta: Odonata) as model organism from the Gangetic riparian zone

Kritish De, S. Zeeshan Ali, Niladri Dasgupta, Virendra Prasad Uniyal, Jeyaraj Antony Johnson & Syed Ainul Hussain

26 October 2020 | Vol. 12 | No. 14 | Pages: 16962–16970
DOI: 10.11609/jott.6106.12.14.16962-16970

Journal of Threatened Taxa

For Focus, Scope, Aims, Policies, and Guidelines visit https://threatenedtaxa.org/index.php/JoTT/about/editorialPolicies#custom-0
For Article Submission Guidelines, visit https://threatenedtaxa.org/index.php/JoTT/about/submissions#onlineSubmissions
For Policies against Scientific Misconduct, visit https://threatenedtaxa.org/index.php/JoTT/about/editorialPolicies#custom-2
For reprints, contact <ravi@threatenedtaxa.org>

The opinions expressed by the authors do not reflect the views of the Journal of Threatened Taxa, Wildlife Information Liaison Development Society, Zoo Outreach Organization, or any of the partners. The journal, the publisher, the host, and the partners are not responsible for the accuracy of the political boundaries shown in the maps by the authors.

Publisher & Host

WILD ZOOREACH Threatened Taxa

Member

FJN

Journal of Threatened Taxa | www.threatenedtaxa.org | 26 October 2020 | 12(14): 16962–16970
ISSN 0974-7907 (Online) | ISSN 0974-7893 (Print)
DOI: https://doi.org/10.11609/jott.6106.12.14.16962-16970
#6106 | Received 06 May 2020 | Final received 19 August 2020 | Finally accepted 10 October 2020
PLATINUM OPEN ACCESS

COMMUNICATION

Evaluating performance of four species distribution models using Blue-tailed Green Darner *Anax guttatus* (Insecta: Odonata) as model organism from the Gangetic riparian zone

Kritish De [1], S. Zeeshan Ali [2], Niladri Dasgupta [3], Virendra Prasad Uniyal [4], Jeyaraj Antony Johnson [5] & Syed Ainul Hussain [6]

[1–6] Wildlife Institute of India, Chandrabani, Dehradun, Uttarakhand 248001, India.
[1] kritish.de@gmail.com (corresponding author), [2] zeeshanearth@gmail.com, [3] niladri4all@gmail.com, [4] uniyalvp@wii.gov.in, [5] jaj@wii.gov.in, [6] hussain@wii.gov.in

Abstract: In this paper we evaluated the performance of four species distribution models: generalized linear (GLM), maximum entropy (MAXENT), random forest (RF) and support vector machines (SVM) model, using the distribution of the dragonfly Blue-tailed Green Darner *Anax guttatus* in the Gangetic riparian zone between Bijnor and Kanpur barrage, Uttar Pradesh, India. We used forest cover type, land use, land cover and five bioclimatic variable layers: annual mean temperature, isothermality, temperature seasonality, mean temperature of driest quarter, and precipitation seasonality to build the models. We found that the GLM generated the highest values for AUC, Kappa statistic, TSS, specificity and sensitivity, and the lowest values for omission error and commission error, while the MAXENT model generated the lowest variance in variable importance. We suggest that researchers should not rely on any single algorithm, instead, they should test performance of all available models for their species and area of interest, and choose the best one to build a species distribution model.

Keywords: Generalized linear model, Kappa statistic, maximum entropy model, omission and commission error, random forest model, receiver operating characteristic curve, sensitivity, specificity, support vector machines model, true skill statistic

Editor: Neelesh Dahanukar, Indian Institute of Science Education and Research, Pune, India. **Date of publication:** 26 October 2020 (online & print)

Citation: De, K., S.Z. Ali, N. Dasgpta, V.P. Uniyal, J.A. Johnson & S.A. Hussain (2020). Evaluating performance of four species distribution models using Blue-tailed Green Darner *Anax guttatus* (Insecta: Odonata) as model organism from the Gangetic riparian zone. *Journal of Threatened Taxa* 12(14): 16962–16970. https://doi.org/10.11609/jott.6106.12.14.16962-16970

Copyright: © De et al. 2020. Creative Commons Attribution 4.0 International License. JoTT allows unrestricted use, reproduction, and distribution of this article in any medium by providing adequate credit to the author(s) and the source of publication.

Funding: This work is funded by the National Mission for Clean Ganga, Ministry of Jal Shakti, Department of Water Resources, River development and Ganga Rejuvenation, Government of India (Grant No. B-02/2015-16/1259/NMCG-WIIPROPOSAL).

Competing interests: The authors declare no competing interests.

Author details: Kritish De is working as project fellow at Wildlife Institute of India. His research interests are biodiversity and ecology. Sk. Zeeshan Ali is working as spatial analyst at Wildlife Institute of India. His research interests are geospatial modelling and spatial ecology. Niladri Dasgupta is working as project coordinator at Wildlife Institute of India. His research interests are river ecology and aquatic wildlife conservation. Virendra Prasad Uniyal is working as Scientist G at Wildlife Institute of India. His research interests are ecology and systematics of insects, bioindicators, biodiversity surveys and ecological monitoring. Jeyaraj Antony Johnson is working as Scientist E at Wildlife Institute of India. His research interests are ecology and monitoring of aquatic ecosystem. Syed Ainul Hussain worked as Scientist G at Wildlife Institute of India. His research interests are aquatic ecology and conservation biology.

Author contribution: KD—conceptualization, field work, formal analysis, writing original draft; SZA—field work, formal analysis, writing original draft; ND—editing the draft; VPU—supervision, review and editing the draft; JAJ—supervision, review and editing the draft; SAH—supervision, review and editing the draft, funding acquisition.

Acknowledgements: Authors are thankful to the National Mission for Clean Ganga, Ministry of Jal Shakti, Department of Water Resources, River development and Ganga Rejuvenation, Government of India for sponsoring the work under the project "Biodiversity conservation and Ganga Rejuvenation". Authors express gratitude to the Director and Dean, Wildlife Institute of India for their administrative support for the study. Authors acknowledge the Environment, Forest and Climate Change Department, Government of Uttar Pradesh for necessary support during fieldwork.

INTRODUCTION

Species distribution models (SDMs) are tools that integrate information about species occurrence or abundance with environmental estimates of a landscape, used to predict distribution of a species across landscapes (Elith & Leathwick 2009). When applied in a geographic information system (GIS), SDMs can produce spatial predictions of occurrence likelihood at locations where information on species distribution was previously unavailable (Václavík & Meentemeyer 2009). Though various types of algorithms are used to build different SDMs (Elith et al. 2006), they share common and general approaches (Hirzel et al. 2002) such as: (i) at a specified resolution, the study area is divided into grid cells; (ii) species presence localities (and sometimes absence localities) data are used as the dependent variable; (iii) several environmental variables (e.g., temperature, precipitation, soil type, aspect, land cover type) are collected for each grid cell as predictor variables; and (iv) the suitability of each cell for the species distributions defined as a function of the environmental variables (Stanton et al. 2012). The species distribution prediction is central to applications in ecology, evolution and conservation science (Elith et al. 2006) across terrestrial, freshwater, and marine realms (Elith & Leathwick 2009). But it remains a question for researchers which model should be selected for particular organisms and habitats of interest, particularly when few samples are present for large under-sampled areas (Mi et al. 2017).

Riparian zones are broadly defined as terrestrial landscapes with characteristic vegetation associated with temporary or permanent aquatic ecosystems (Meragiaw et al. 2018). These areas are highly complex biophysical systems, and their ecological functions are maintained by strong spatio-temporal connectivity with adjacent riverine and upland systems (Décamps et al. 2009). It has been observed that species distribution models are used more often for terrestrial environments than for aquatic or riparian ecosystems. Globally, odonates are used as model organisms to study climate change, data simulation, environmental assessment and management, effects of urbanization, landscape planning, habitat monitoring and evaluation, and conservation of rare species (Bried & Samways 2015). To date, no work has been done on the comparative use of species distribution models in India using insects as model organisms in riparian or freshwater ecosystems. With this background, in the present work we evaluated the effectiveness of four species distribution models using odonates from the Gangetic riparian zone as model organisms.

MATERIALS AND METHODS

Study area and field data collection

For the study, we selected *Anax guttatus* (Burmeister, 1839) commonly called Blue-tailed Green Darner (Image 1) as the model insect species. It is a dragonfly (suborder Anisoptera Selys, 1854) under the family Aeshnidae Leach, 1815 and superfamily Aeshnoidea Leach, 1815 (Dijkstra et al. 2013). The species can be identified in the field due to its large size, highly active behaviour, green colour of the thorax & first, second, & third abdominal segments, and presence of turquoise blue colour on the dorsal part of the second abdominal segment (Subramanian 2005).

We conducted the study during May 2019 from Bijnor, Uttar Pradesh to Kanpur, Uttar Pradesh (Fig. 1). The river flows through alluvial plain and covers a length of about 450km in this stretch. For the study we selected four sites, and the distance between each two successive sites was about 150km. In each site we chose a 10km river stretch and observed the presence of Blue-tailed Green Darner. We collected a total of 10 sighting locations.

Data processing and analysis

We derived the thematic layer of LULC (N.R.S.C. 2016) from multi-temporal advanced wide field sensor (AWiFS) images with 56m spatial resolution using digital and rule-based image classification methods, and forest

Image 1. *Anax guttatus* (Burmeister, 1839) – Blue-tailed Green Darner

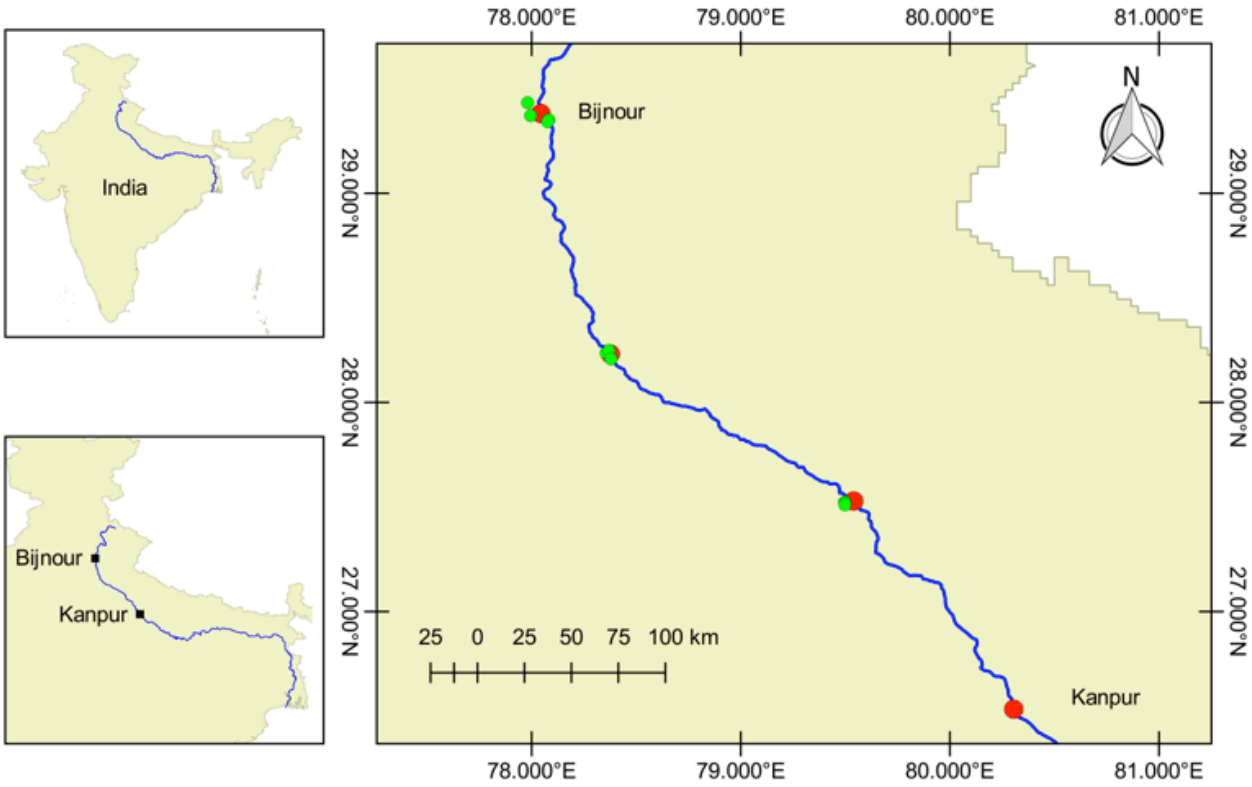

Figure 1. Study area on the river Ganga between Bijnour and Kanpur. The red circles represent sampling sites, and green circles represent species sighting locations.

cover type (F.S.I. 2009) from IRS P6 (Linear Imaging Self Scanning Sensor) LISS III with 23.5m spatial resolution using a combined method of digital and on-screen visual image classification and bioclimatic layers from worldclim gridded climatic data (Fick & Hijmans 2017) with 1km spatial resolution. For analysis, we took 2km buffer zones from the river bank and resampled all the layers to 1km spatial resolution.

We used 'stack' function of package 'raster' (Hijmans 2019) to stack all the 19 available bioclimatic variable, forest cover and land use land cover (LULC) layers. After that we used 'pairs' function of the package 'raster' (Hijmans 2019) to find the correlation coefficient between stacked layers. Then we selected the variables which had a correlation coefficient less than 0.60 (Pozzobom et al. 2020), and again stacked the selected layers with 'stack' function of package 'raster' (Hijmans 2019). These selected layers were LULC, forest cover and five bioclimatic layers: annual mean temperature (Bio 1), isothermality (Bio 3), temperature seasonality (Bio 4), mean temperature of driest quarter (Bio 9), and precipitation seasonality (Bio 15).

We built four species distribution models: generalized linear model (GLM), maximum entropy (MAXENT) model, random forest (RF) model, and support vector machines (SVM).

GLM is an extension of classic linear regression modeling, where the iterative weighted linear regression technique is used to estimate maximum-likelihood of the parameters, with observations distributed in terms of an exponential family and systematic effects made linear by the suitable transformation that allow for analysis of non-linear effects among variables and non-normal distributions of the independent variables (McCullagh & Nelder 1989; Chefaoui & Lobo 2008; Shabani et al. 2016).

RF modeling is a machine learning technique which is a bootstrap-based classification and regression trees method (Cutler et al. 2007). It is used to model species distributions from both the abundance and the presence-absence data (Howard et al. 2014). It is insensitive to data distribution (Hill et al. 2017) and also takes a large number of potentially collinear variables; it is robust to over-fitting which makes it very useful for prediction (Prasad et al. 2006; Segal 2004).

MAXENT modeling is a general-purpose machine learning method to estimate a target probability distribution by finding the probability distribution of

maximum entropy and it has several aspects that make it well-suited for species distribution modelling (Phillips et al. 2006). It is relatively less sensitive to the spatial errors associated with location data and needs few locations to build useful models (Baldwin 2009) and it is one of the most accurate and trusted modelling methods for presence-only distribution data (Huerta & Peterson 2008; Srinivasulu & Srinivasulu 2016).

SVM modeling is developed from the theory of statistical learning, in which the error involved with sample size is minimized and the upper limit of the error involved in model generalization is narrowed, which solve the problems of nonlinearity, over-learning and the curse of dimensionality during modelling (Fielding & Bell 1997; Howley & Madden 2005; Huang & Wang 2006). It can be used on small data sets as it is independent of any distributional assumptions or asymptotic arguments (Wilson 2008).

We used 'load_var' function to normalize and load environmental variables, then used 'load_occ' function to load species occurrence data and then used 'modelling' function to build the models with 100 iterations by the package 'SSDM' (Schmitt et al. 2017) to plot the models.

We evaluated and compared four models by comparing values of area under the receiver operating characteristic curve (AUC), Kohen's Kappa, true skill statistic (TSS), model sensitivity, model specificity, and omission error.

The area under the receiver operating characteristic curve or AUC measures the ability of a model to discriminate between the sites where a species is present and the sites where a species is absent (Fielding & Bell 1997; Elith et al. 2006) and it provides a single measure of overall accuracy that is independent of a particular threshold (Fielding & Bell 1997). The evaluation criteria for the AUC statistic are as follows: excellent (0.90–1.00), very good (0.8–0.9), good (0.7–0.8), fair (0.6–0.7), and poor (0.5–0.6) (Swets 1988; Duan et al. 2014).

The Kappa statistic is based on the optimal threshold, measure the performance of the model by using the best of the information in the mixed matrix (Duan et al. 2014) ranges from −1 to +1, where +1 indicates perfect agreement and values of zero or less than zero indicate a performance no better than random (Allouche et al. 2006; Cohen 1960) and the evaluation criteria for the Kappa statistic are as follows: excellent (0.85–1.0), very good (0.7–0.85), good (0.55–0.7), fair (0.4–0.55), and fail (<0.4) (Duan et al. 2014; Monserud & Leemans 1992).

The true skill statistic (TSS) is expressed as Sensitivity + Specificity – 1 (Allouche et al. 2006) and ranges from −1 to +1, where +1 indicates a perfectly performing model with no error, 0 indicates the model with totally random error and -1 indicates the model with total error (Marcot 2012; Ruete & Leynaud 2015).

The model sensitivity denotes the proportion of correctly predicted presences, thus quantifying omission errors (Ward 2007; Shabani et al. 2016) and model specificity denotes the proportion of correctly predicted presences, thus quantifying commission errors (Shabani et al. 2016).

Omission error (1- sensitivity) is the under-prediction or false-negative result in areas being classified as unsuitable when they are not and commission error (1-specificity) is the over-prediction or false-positive result in areas being classified as suitable when they are not (Ward 2007) and for a good SDM, both of the omission error and commission error should be low.

For evaluation of model performance and variable importance we used 'knitr::kable(Modelname@evaluation)' function and 'knitr::kable(Modelname@variable.importance)' function of the package 'SSDM' (Schmitt et al. 2017), respectively.

We chose five probability classes (0 to <0.20, 0.20 to <0.40, 0.40 to <0.60, 0.60 to <0.80 and 0.80 to 1.00) to know what percentage of the area is being declared the best and worst by each of the models by 'ratify' function of package 'raster' (Hijmans 2019)

We performed all the analysis in the ArcMap 10.3.1, QGIS 2.14.7 and in R language and environment for statistical computing (R Core Team 2019).

RESULT

The plot for each of the four models is given in Fig. 2. We found that the AUC value was highest for GLM (0.983), followed by RF (0.833), MAXENT (0.829) and SVM (0.667); the value of Kappa statistic was highest for RF (0.667), followed by GLM (0.356), SVM (0.333) and MAXENT (0.049); the value of TSS was highest for GLM (0.965), followed by RF (0.666), MAXENT (0.658) and SVM (0.334); the value of model sensitivity was 1 for GLM, 0.833 for both MAXENT and RF and 0.667 for SVM; the value of model specificity was maximum for GLM (0.965), followed by RF (0.833), MAXENT (0.825) and SVM) (0.667); the omission error was lowest for GLM (0.00), for both MAXENT and RF models it was 0.167 and for SVM it was 0.333; the commission error was lowest for GLM (0.035), followed by RF model (0.167), MAXENT (0.175) and SVM (0.333) (Table 1, Fig. 3)

For GLM, RF, and SVM models the forest had

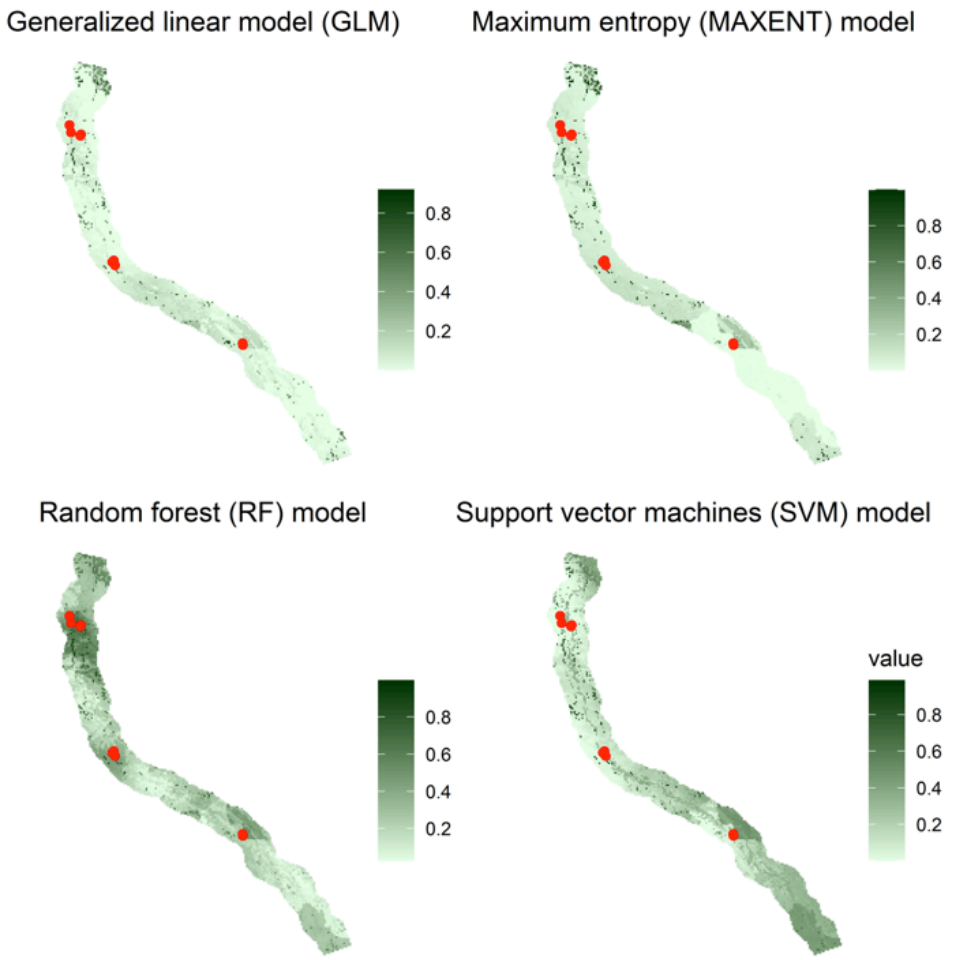

Figure 2. Comparative account of four species distribution models – generalized linear model (GLM), maximum entropy (MAXENT), random forest (RF), and support vector machines (SVM). The legend in each model shows probability of occurrence and the red dotes represents the occurrence location of the species.

the highest importance but for MAXENT model the Precipitation seasonality (Bio 15) had the highest importance (Table 2, Fig. 4). For GLM and SVM models the Precipitation seasonality (Bio 15) had lowest importance, for MAXENT forest had lowest importance, while for RF model Isothermality (Bio 3) had lowest importance (Table 2, Fig. 4). Overall, the variation in the variable importance was lowest in MAXENT model (SD = 3.367), followed by GLM (SD = 24.344), RF (SD = 30.868) and SVM (SD = 37.071) (Fig. 5).

By comparative analysis, we found that GLM showed 1.62% of total area as the best (occurrence probability, 0.80 to 1) and 65.50% of total area as the worst (occurrence probability, 0 to 0.20) for suitable habitat. MAXENT model showed 10.08% of total area as the best and 77.70% of total area as the worst for suitable habitat. RF model showed 5.39% of total area as the best and 23.79% of total area as the worst for suitable habitat. SVM model showed 4.53% of total area as the best and 27.68% of total area as the worst for suitable habitat (Table 3, Fig. 6).

Table 1. Values of AUC, Kappa statistic, TSS, sensitivity, specificity, omission error, and commission error generated by generalized linear model (GLM), maximum entropy (MAXENT) model, random forest (RF) model, and support vector machines (SVM) model.

	GLM	MAXENT	RF	SVM
AUC	0.983	0.829	0.833	0.667
Kappa statistic	0.356	0.049	0.667	0.333
True skill statistic	0.965	0.658	0.666	0.334
Sensitivity	1	0.833	0.833	0.667
Specificity	0.965	0.825	0.833	0.667
Omission error	0	0.167	0.167	0.333
Commission error	0.035	0.175	0.167	0.333

Table 2. Comparative importance (%) of seven variables from generalized linear model (GLM), maximum entropy (MAXENT) model, random forest (RF) model, and support vector machines (SVM) model.

	GLM	MAXENT	RF	SVM
Annual mean temperature (Bio 1)	11.831	16.352	2.254	0.198
Isothermality (Bio 3)	8.062	14.789	0.513	0.337
Temperature seasonality) (Bio 4)	5.709	15.405	4.076	0.239
Mean temperature of driest quarter (Bio 9)	3.241	13.638	0.907	0.069
Precipitation seasonality (Bio 15)	1.103	16.417	2.817	0.019
Forest	68.799	7.014	84.186	98.353
Land use land cover	1.252	16.384	5.247	0.785

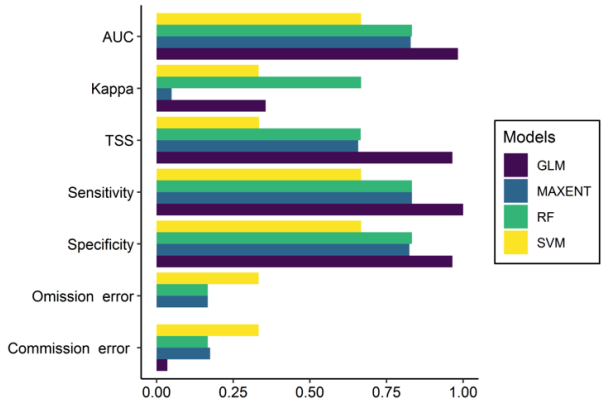

Figure 3. Bar diagram showing comparative account of values of AUC, Kappa statistic, TSS, sensitivity, specificity , omission error and commission error for four models – generalized linear model (GLM), maximum entropy (MAXENT) model, random forest (RF) model, and support vector machines (SVM) model.

Table 3. Comparison of percentage of total area obtained from each model for five occurrence probability classes,

Occurrence probability class	Models			
	GLM	MAXENT	RF	SVM
0 to <0.20	65.50	77.70	23.79	27.68
0.20 to <0.40	7.94	3.93	35.61	42.55
0.40 to <0.60	19.58	4.04	17.97	18.04
0.60 to <0.80	5.35	4.24	17.24	7.19
0.80 to 1.00	1.62	10.08	5.39	4.53

DISCUSSION

Freshwater ecosystems, which include rivers, lakes, peat lands, swamps, fens, and springs, are highly dynamic and host a great diversity of life forms, particularly freshwater endemic species (He et al. 2019; Tickner et al. 2020). They are among the most threatened ecosystems (He et al. 2019), as globally wetlands are vanishing more rapidly than forests and freshwater species are declining faster than terrestrial or marine populations (Tickner et al. 2020). Therefore, for proper conservation management, we should understand the distribution of plants and animals inhabiting aquatic ecosystems. Species distribution models can play an important role on such efforts, because they can produce credible, defensible and repeatable information and provide tools for mapping habitats to inform decisions (Sofaer et al. 2019). Species distribution models can forecast the potential impacts of future environmental changes (Howard et al. 2014) and predict how species will respond (Buckley et al. 2010). Yet debate remains over the most robust species distribution modelling approaches for making projections (Howard et al. 2014), because these models have sensitivity to data inputs and methodological choices. This makes it important to assess the reliability and utility of the model predictions (Sofaer et al. 2019).

In the present study we compared the GLM, MAXENT, RF, and SVM approaches. We found that GLM generated the highest values for AUC, TSS, specificity and sensitivity, and the lowest values for omission error and commission error. The value of Kappa statistic was highest for RF modelling. The MAXENT model used roughly all variables equally, which is not true of the other models which put more emphasis on forest cover.

The success of a model depends on many factors, such as sample size, spatial extent of the study area, and number of ecological and statistical significant variables which affect the distribution of species of interest. We acknowledge that there were some limitations to the current work, such as that our sample size was small (only 10 presence locations), we used only seven variables, we tested only four species distribution models, and we selected a species whose distribution depends on other factors, such as the physiochemical parameters of water and availability of resources. We did not include such

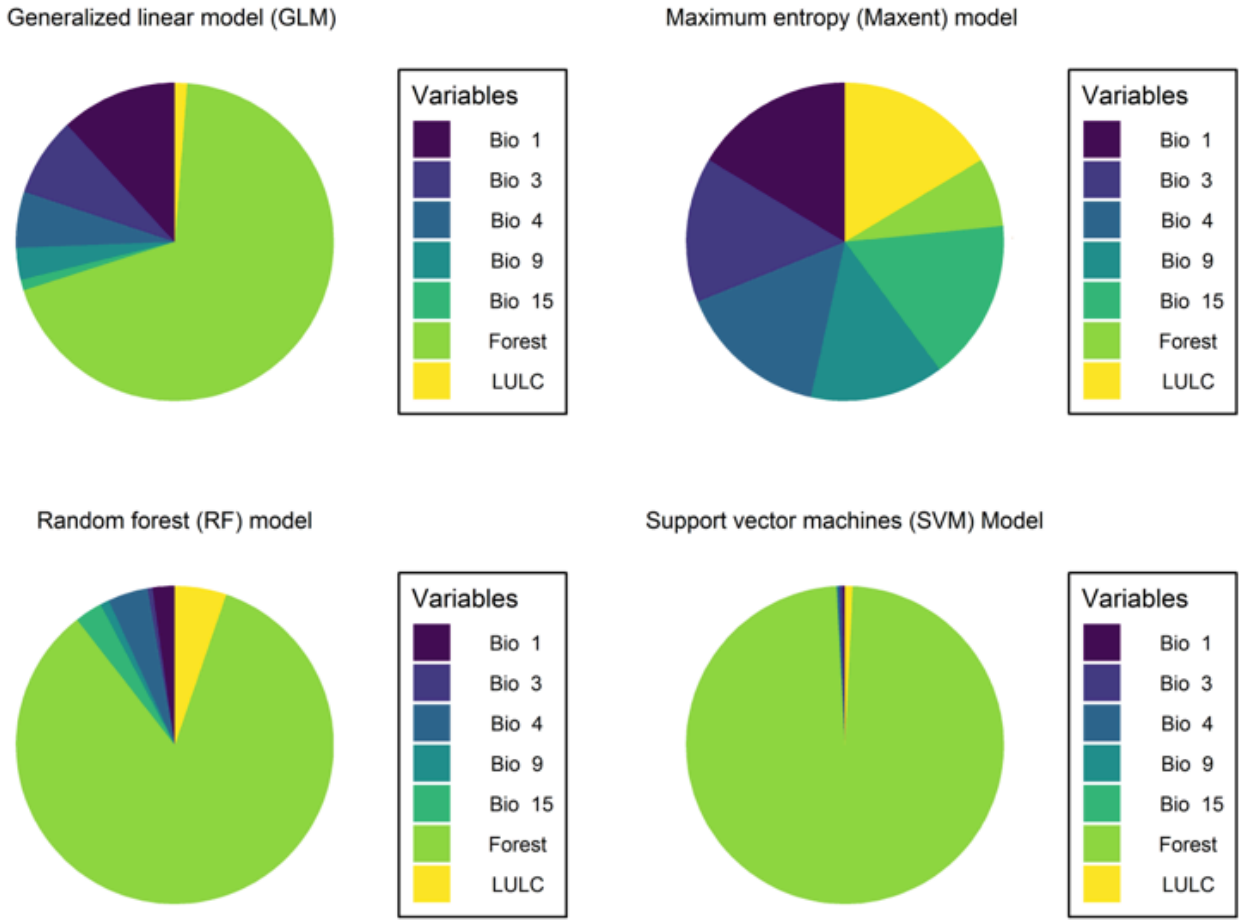

Figure 4. Pie diagram showing percentage importance for seven variables for four models – generalized linear model (GLM), maximum entropy (MAXENT) model, random forest (RF) model, and support vector machines (SVM) model.

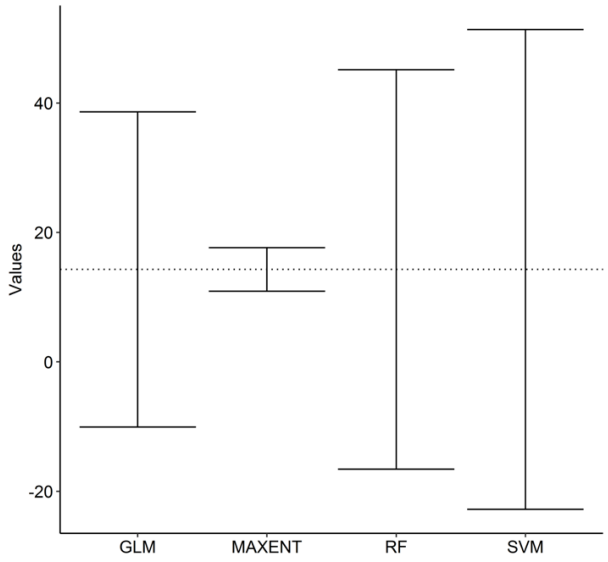

Figure 5. Comparison of standard deviation of variable importance for generalized linear model (GLM), maximum entropy (MAXENT) model, random forest (RF) model, and support vector machines (SVM) model. The black dotted line represents the mean value.

variables as this study was preliminary.

Collins & McIntyre (2015) reviewed 30 studies on species distribution modelling of odonates across the world, and found that 43% used GLM, 33% MAXENT and 20% RF models. Other models used were BIOMOD, general additive model (GAM), generalized boosted model (GBM), artificial neural networks (ANN), multivariate adaptive regression splines (MARS), classified tree analysis (CTA), flexible discriminant analysis (FDA), boosted regression trees (BRT), surface range envelopes (SRE), and mixture discriminant analysis (MDA). Different species distribution models produce different results (Shabani et al. 2016), and the same model can give different results for different species and areas. We urge researchers not to rely on just one model, rather they should compare different available species distribution models and select the best one. Our study was in India where an insect was used for comparative evaluation of species distribution models in a riverine riparian zone. We recommend that further

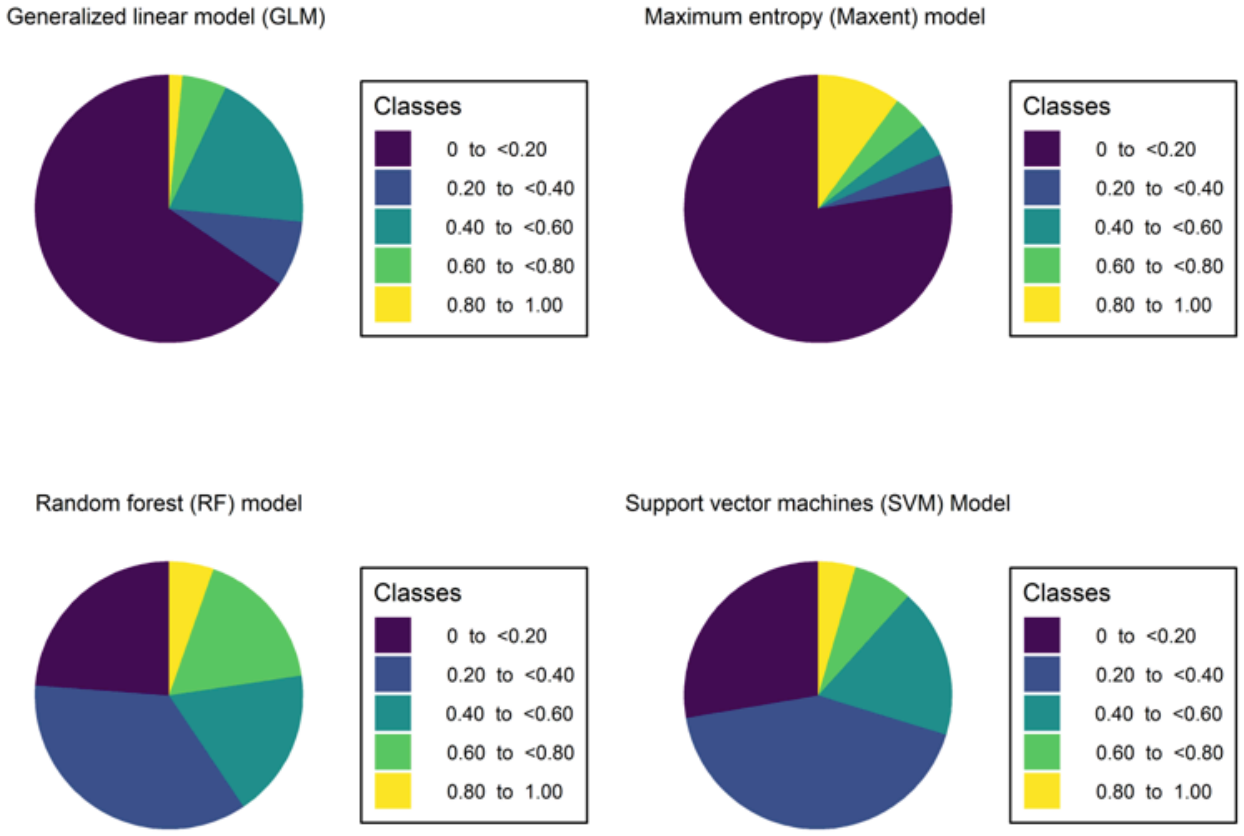

Figure 6. Pie diagram showing percentage area for five occurrence probability classes for four models – generalized linear model (GLM), maximum entropy (MAXENT) model, random forest (RF) model, and support vector machines (SVM) model.

studies on different species distribution models using different animals and ecological variables should be carried out in the riparian zones of Indian river systems for proper design and implementation of ecological habitat management plans.

REFERENCES

Allouche, O., A. Tsoar & R. Kadmon (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43(6): 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Baldwin R. (2009). Use of Maximum Entropy Modeling in Wildlife Research. *Entropy* 11(4): 854–866. https://doi.org/10.3390/e11040854

Bried, J.T. & M.J. Samways (2015). A review of odonatology in freshwater applied ecology and conservation science. *Freshwater Science*, 34(3): 1023–1031. https://doi.org/10.1086/682174

Buckley, L.B., M.C. Urban, M.J. Angilletta, L.G. Crozier, L.J. Rissler & M.W. Sears (2010). Can mechanism inform species' distribution models? *Ecology Letters* 13(8): 1041–1054. https://doi.org/10.1111/j.1461-0248.2010.01479.x

Chefaoui, R.M. & J.M. Lobo (2008). Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling* 210(4): 478–486. https://doi.org/10.1016/j.ecolmodel.2007.08.010

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20(1): 37–46. https://doi.org/10.1177/001316446002000104

Collins, S.D. & N.E. McIntyre (2015). Modeling the distribution of odonates: a review. *Freshwater Science* 34(3): 1144–1158. https://doi.org/10.1086/682688

Cutler, D.R., T.C. Edwards Jr, K.H. Beard, A. Cutler, K.T. Hess, J. Gibson & J.J. Lawler (2007). Random forests for classification in ecology. *Ecology* 88(11): 2783–2792. https://doi.org/10.1890/07-0539.1

Décamps, H., R.J. Naiman & M.E. McClain (2009). Riparian Zones. In: *Encyclopedia of Inland Waters* (pp. 396–403). https://doi.org/10.1016/b978-012370626-3.00053-3

Dijkstra, K.D.B., G. Bechly, S.M. Bybee, R.A. Dow, H.J. Dumont, G. Fleck, R.W. Garrison, M. Hämäläinen, V.J. Kalkman, H. Karube, M.L. May, A.G. Orr, D.R. Paulson, A.C. Rehn, G. Theischinger, J.W.H. Trueman, J.V. Tol, N.V. Ellenrieder & J. Ware (2013). The classification and diversity of dragonflies and damselflies (Odonata). *Zootaxa* 3703(1): 036–045. https://doi.org/10.11646/zootaxa.3703.1.9

Duan, R.Y., X.Q. Kong, M.Y. Huang, W.Y. Fan & Z.G. Wang (2014). The predictive performance and stability of six species distribution models. *PLoS ONE* 9(11): e112764. https://doi.org/10.1371/journal.pone.0112764

Elith, J. & J.R. Leathwick (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual review of Ecology, Evolution and Systematics* 40: 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Elith, J., C.H. Graham, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J.McC.M. Overton, A.T. Peterson, Steven J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberón, S. Williams, M.S. Wisz & N.E. Zimmermann (2006). Novel methods

improve prediction of species' distributions from occurrence data. *Ecography* 29(2): 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

F.S.I. (2009). India State of Forest Report – 2009. Forest Survey of India (Ministry of Environment Forests and Climate Change, Government of India), Dehradun.

Fick, S.E. & R.J. Hijmans (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37(12): 4302–4315. https://doi.org/10.1002/joc.5086

Fielding, A.H. & J.F. Bell (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24(1): 38–49. https://doi.org/10.1017/S0376892997000088

He, F., C. Zarfl, V. Bremerich, J.N.W. David, Z. Hogan, G. Kalinkat, K. Tockner & S.C. Jähnig (2019). The global decline of freshwater megafauna.*Global Change Biology* 25(11): 3883–3892. https://doi.org/10.1111/gcb.14753

Hijmans, R.J. (2019). raster: Geographic Data Analysis and Modeling. R package version 3.0-7. https://CRAN.R-project.org/package=raster

Hill, L., A. Hector, G. Hemery, S. Smart, M. Tanadini & N. Brown (2017). Abundance distributions for tree species in Great Britain: A two-stage approach to modeling abundance using species distribution modeling and random forest. *Ecology and Evolution* 7(4): 1043–1056. https://doi.org/10.1002/ece3.2661

Hirzel, A.H., J. Hausser, D. Chessel & N. Perrin (2002). Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data?. *Ecology* 83(7): 2027–2036. https://doi.org/10.1890/0012-9658(2002)083[2027:ENFAHT]2.0.CO;2

Howard, C., P.A. Stephens, J.W. Pearce-Higgins, R.D. Gregory & S.G. Willis (2014). Improving species distribution models: the value of data on abundance. *Methods in Ecology and Evolution* 5(6): 506–513. https://doi.org/10.1111/2041-210X.12184

Howley, T. & M.G. Madden (2005). The genetic kernel support vector machine: Description and evaluation. *Artificial Intelligence Review* 24(3–4): 379–395. https://doi.org/10.1007/s10462-005-9009-3

Huang, C.L. & C.J. Wang (2006). A GA-based feature selection and parameters optimizationfor support vector machines. *Expert Systems with Applications* 31(2): 231–240. https://doi.org/10.1016/j.eswa.2005.09.024

Huerta, M.A.O. & A.T. Peterson (2008). Modeling ecological niches and predicting geographic distributions: a test of six presence-only methods. *Revista Mexicana de Biodiversidad* 1(1): 205–216.

Marcot, B.G. (2012). Metrics for evaluating performance and uncertainty of Bayesian network models. *Ecological Modelling* 230: 50–62. https://doi.org/10.1016/j.ecolmodel.2012.01.013

McCullagh, P. & J.A. Nelder (1989). Generalized Linear Models. Chapman and Hall, London, 511pp

Meragiaw, M., Z. Woldu, V. Martinsen & B.R. Singh (2018). Woody species composition and diversity of riparian vegetation along the Walga River, Southwestern Ethiopia. *PLoS ONE* 13(10): e0204733. https://doi.org/10.1371/journal.pone.0204733

Mi, C., F. Huettmann, Y. Guo, X. Han & L. Wen (2017). Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ* 5: p.e2849. https://doi.org/10.7717/peerj.2849

Monserud, R.A. & R. Leemans (1992). Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling* 62(4): 275–293. https://doi.org/10.1016/0304-3800(92)90003-W

N.R.S.C. (2016). National Land Use and Land Cover (LULC) mapping using multi-temporal AWiFS Data (2015–16). National Remote Sensing Centre, Hyderabad, India.

Phillips, S.J., R.P. Anderson & R.E. Schapire (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190(3–4): 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pozzobom, U.M., J. Heino, M.T. da S. Brito & V.L. Landeiro (2020). Untangling the determinants of macrophyte beta diversity in tropical floodplain lakes: insights from ecological uniqueness and species contributions. *Aquatic Sciences* 82(3): 56. https://doi.org/10.1007/s00027-020-00730-2

Prasad, A.M., L.R. Iverson & A. Liaw (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9(2): 181–199. https://doi.org/10.1007/s10021-005-0054-1

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ruete, A. & G.C. Leynaud (2015). Goal-oriented evaluation of species distribution models' accuracy and precision: True Skill Statistic profile and uncertainty maps. *PeerJ PrePrints* 3: e1208v1. https://doi.org/10.7287/peerj.preprints.1208v1

Schmitt, S., R. Pouteau, D. Justeau, F. de Boissieu & P. Birnbaum (2017). ssdm: An r package to predict distribution of species richness and composition based on stacked species distribution models. *Methods in Ecology and Evolution* 8(12): 1795–1803. https://doi.org/10.1111/2041-210X.12841

Segal, M.R. (2004). Machine Learning Benchmarks and Random Forest Regression. UCSF: Center for Bioinformatics and Molecular Biostatistics. Retrieved from https://escholarship.org/uc/item/35x3v9t4

Shabani, F., L. Kumar & M. Ahmadi (2016). A comparison of absolute performance of different correlative and mechanistic species distribution models in an independent area. *Ecology and Evolution* 6(16): 5973–5986. https://doi.org/10.1002/ece3.2332

Sofaer, H.R., C.S. Jarnevich, I.S. Pearse, R.L. Smyth, S. Auer, G.L. Cook, T.C. Edwards Jr, G.F. Guala, T.G. Howard, J.T. Morisette & H. Hamilton (2019). Development and delivery of species distribution models to inform decision-making. *BioScience* 69(7): 544–557. https://doi.org/10.1093/biosci/biz045

Srinivasulu, A. & C. Srinivasulu (2016). All that glitters is not gold: A projected distribution of the endemic Indian Golden Gecko *Calodactylodes aureus* (Reptilia: Squamata: Gekkonidae) indicates a major range shrinkage due to future climate change. *Journal of Threatened Taxa* 8(6): 8883–8892. https://doi.org/10.11609/jott.2723.8.6.8883-8892

Stanton, J.C., R.G. Pearson, N. Horning, P. Ersts & H. ReşitAkçakaya (2012). Combining static and dynamic variables in species distribution models under climate change. *Methods in Ecology and Evolution* 3(2): 349–357. https://doi.org/10.1111/j.2041-210X.2011.00157.x

Subramanian, K.A. (2005). Dragonflies and Damselflies of Peninsular India-A Field Guide. E-Book of Project Lifescape. Centre for Ecological Sciences, Indian Institute of Science and Indian Academy of Sciences, Bangalore, 118pp.

Swets, J.A. (1988). Measuring the accuracy of diagnostic systems. *Science* 240(4857): 1285–1293. https://doi.org/10.1126/science.3287615

Tickner, D., J.J. Opperman, R. Abell, M. Acreman, A.H. Arthington, S.E. Bunn, S.J. Cooke, J. Dalton, W. Darwall, G. Edwards, I. Harrison, K. Hughes, T. Jones, D. Leclère, A.J. Lynch, P. Leonard, M.E. McClain, D. Muruven, J.D. Olden, S.J. Ormerod, J. Robinson, R.E. Tharme, M. Thieme, K. Tockner, M. Wright & L. Young (2020). Bending the curve of global freshwater biodiversity loss: an emergency recovery plan. *BioScience* 70(4): 330–342. https://doi.org/10.1093/biosci/biaa002

Václavík, T. & R.K. Meentemeyer (2009). Invasive species distribution modeling (iSDM): are absence data and dispersal constraints needed to predict actual distributions?. *Ecological Modelling* 220(23): 3248–3258. https://doi.org/10.1016/j.ecolmodel.2009.08.013

Ward, D.F. (2007). Modelling the potential geographic distribution of invasive ant species in New Zealand. *Biological Invasions* 9(6): 723–735. https://doi.org/10.1007/s10530-006-9072-y

Wilson, M.D. (2008). Support Vector Machines. pp. 3431–3437. In: Jørgensen, S.E. & B.D. Fath (eds.). *Encyclopedia of Ecology*. Elsevier, 3120pp. https://doi.org/10.1016/B978-008045405-4.00168-3

Journal of Threatened Taxa

www.threatenedtaxa.org

PLATINUM OPEN ACCESS

The Journal of Threatened Taxa (JoTT) is dedicated to building evidence for conservation globally by publishing peer-reviewed articles online every month at a reasonably rapid rate at www.threatenedtaxa.org. All articles published in JoTT are registered under Creative Commons Attribution 4.0 International License unless otherwise mentioned. JoTT allows allows unrestricted use, reproduction, and distribution of articles in any medium by providing adequate credit to the author(s) and the source of publication.

ISSN 0974-7907 (Online) | ISSN 0974-7893 (Print)

October 2020 | Vol. 12 | No. 14 | Pages: 16927–17062
Date of Publication: 26 October 2020 (Online & Print)
DOI: 10.11609/jott.2020.12.14.16927-17062

Article

Elevational pattern and seasonality of avian diversity in Kaligandaki River Basin, central Himalaya
– Juna Neupane, Laxman Khanal, Basant Gyawali & Mukesh Kumar Chalise, Pp. 16927–16943

Communications

A highway to hell: a proposed, inessential, 6-lane highway (NH173) that threatens the forest and wildlife corridors of the Western Ghats, India
– H.S. Sathya Chandra Sagar & Mrunmayee, Pp. 16944–16953

Species diversity and feeding guilds of birds in Malaysian agarwood plantations
– Nor Nasibah Mohd Jamil, Husni Ibrahim, Haniza Hanim Mohd Zain & Nur Hidayat Che Musa, Pp. 16954–16961

Evaluating performance of four species distribution models using Blue-tailed Green Darner *Anax guttatus* (Insecta: Odonata) as model organism from the Gangetic riparian zone
– Kritish De, S. Zeeshan Ali, Niladri Dasgupta, Virendra Prasad Uniyal, Jeyaraj Antony Johnson & Syed Ainul Hussain, Pp. 16962–16970

Butterfly species richness and diversity in rural and urban areas of Sirajganj, Bangladesh
– Sheikh Muhammad Shaburul Imam, Amit Kumer Neogi, M. Ziaur Rahman & M. Sabbir Hasan, Pp. 16971–16978

Chroococcalean blue green algae from the paddy fields of Satara District, Maharashtra, India
– Sharada Jagannath Ghadage & Vaneeta Chandrashekhar Karande, Pp. 16979–16992

Short Communications

Avifaunal diversity along the riverine habitats of Papikonda National Park, Andhra Pradesh, India
– Paromita Ray, Giridhar Malla, Upma Manral, J.A. Johnson & K. Sivakumar, Pp. 16993–16999

Medetomidine may cause heart murmur in Cougars and Jaguars: case report
– Thiago Cavalheri Luczinski, Gediendson Ribeiro de Araújo, Matheus Folgearini Silveira, Murillo Daparé Kirnew, Roberto Andres Navarrete, Jorge Aparecido Salomão-Jr, Letícia Alecho Requena, Jairo Antonio Melo dos Santos, Marcell Hideki Koshiyama, Cristiane Schilbach Pizzutto & Pedro Nacib Jorge-Neto, Pp. 17000–17002

Description of a new species of *Omyomymar* Schauff from India with a key to Oriental species and first report of *Palaeoneura markhoddlei* Triapitsyn (Hymenoptera: Mymaridae) from the Indian subcontinent
– H. Sankararaman & S. Manickavasagam, Pp. 17003–17008

Incursion of the killer sponge *Terpios hoshinota* Rützler & Muzik, 1993 on the coral reefs of the Lakshadweep archipelago, Arabian Sea
– Rocktim Ramen Das, Chemmencheri Ramakrishnan Sreeraj, Gopi Mohan, Kottarathil Rajendran Abhilash, Vijay Kumar Deepak Samuel, Purvaja Ramachandran & Ramesh Ramachandran, Pp. 17009–17013

Contribution to the macromycetes of West Bengal, India: 63–68
– Rituparna Saha, Debal Ray, Anirban Roy & Krishnendu Acharya, Pp. 17014–17023

Notes

A rare camera trap record of the Hispid Hare *Caprolagus hispidus* from Dudhwa Tiger Reserve, Terai Arc Landscape, India
– Sankarshan Rastogi, Ram Kumar Raj & Bridesh Kumar Chauhan, Pp. 17024–17027

First distributional record of the Lesser Adjutant *Leptoptilos javanicus* Horsfield, 1821 (Ciconiiformes: Ciconiidae) from Sindhuli District, Nepal
– Badri Baral, Sudeep Bhandari, Saroj Koirala, Parashuram Bhandari, Ganesh Magar, Dipak Raj Basnet, Jeevan Rai & Hem Sagar Baral, Pp. 17028–17031

First record of African Sailfin Flying Fish *Parexocoetus mento* (Valenciennes, 1847) (Beloniformes: Exocoetidae), from the waters off Andaman Islands, India
– Y. Gladston, S.M. Ajina, J. Praveenraj, R. Kiruba-Sankar, K.K. Bineesh & S. Dam Roy, Pp. 17032–17035

A first distribution record of the Indian Peacock Softshell Turtle *Nilssonia hurum* (Gray, 1830) (Reptilia: Testudines: Trionychidae) from Mizoram, India
– Gospel Zothanmawia Hmar, Lalbiakzuala, Lalmuansanga, Dadina Zote, Vanlalhruaia, Hmar Betlu Ramengmawii, Kulendra Chandra Das & Hmar Tlawmte Lalremsanga, Pp. 17036–17040

A frog that eats foam: predation on the nest of *Polypedates* sp. (Rhacophoridae) by *Euphlyctis* sp. (Dicroglossidae)
– Pranoy Kishore Borah, Avrajjal Ghosh, Bikash Sahoo & Aniruddha Datta-Roy, Pp. 17041–17044

New distribution record of two endemic plant species, *Euphorbia kadapensis* Sarojin. & R.R.V. Raju (Euphorbiaceae) and *Lepidagathis keralensis* Madhus. & N.P. Singh (Acanthaceae), for Karnataka, India
– P. Raja, N. Dhatchanamoorthy, S. Soosairaj & P. Jansirani, Pp. 17045–17048

***Cirsium wallichii* DC. (Asteraceae): a key nectar source of butterflies**
– Bitupan Boruah, Amit Kumar & Abhijit Das, Pp. 17049–17056

***Hypecoum pendulum* L. (Papaveraceae: Ranunculales): a new record for the flora of Haryana, India**
– Naina Palria, Nidhan Singh & Bhoo Dev Vashistha, Pp. 17057–17059

Addendum

Erratum and addenda to the article 'A history of primatology in India'
– Mewa Singh, Mridula Singh, Honnavalli N. Kumara, Dilip Chetry & Santanu Mahato, Pp. 17060–17062

Member

FJN

Publisher & Host

WILD ZOOREACH

Threatened Taxa