ISSN: 2655-9587 (online) | ISSN: 2303-2162 (print) Available *online* at : http://jbioua.fmipa.unand.ac.id

Research Article

JURNAL BIOLOGI UNIVERSITAS ANDALAS
Vol. 13 No. 01 (2025) 36-42

Evaluasi Model Sebaran Spesies (SDM) untuk Deteksi Spesies Liar secara Efisien dan Akurat di Berbagai Lanskap

Evaluating Species Distribution Models (SDMs) for Efficient and Accurate Detection of Wild Species Across Landscapes

Ahmad Taufiq [*)], Nurainas

Herbarium ANDA, Biology Department of Universitas Andalas

SUBMISSION TRACK

Submitted : 25-05-2025
Revised : 13-06-2025
Accepted : 22-06-2025
Published : 27-06-2025

KEYWORDS

Species Distribution Models, Environmental Niche Modeling, Species Detection, Biodiversity Monitoring, Maxent Modeling, Field Survey Efficiency

[*)]CORRESPONDENCE

email:
ahmadtaufiq.herb@gmail.com
ahmadtaufiq@sci.unand.ac.id

ABSTRACT

Species distribution models (SDMs) have been used across continents and taxonomic groups to guide field surveys and improve detection efficiency. In several studies, SDM-guided approaches achieved Area Under the Curve values between 0.90 and 0.976, with some reports documenting the discovery of new populations (e.g., 4 of 8 species or 5–16 additional sites) and time savings of up to 70% compared with unsystematic surveys. One study noted that Gaussian Process models operated 70 times faster than an alternative estimation method. Additional work indicates that SDMs narrow survey areas and enhance cost effectiveness, particularly when environmental layers and robust occurrence data support model development. These studies show that, when applied with methods such as Maxent and ensemble approaches, SDMs offer a viable alternative to direct field surveys for locating wild species over large areas. Limitations arise when data quality or model specification is insufficient, suggesting that careful design remains essential for reliable outcomes.

INTRODUCTION

Species detection and distribution mapping are foundational components of ecology and conservation biology. Traditionally, direct field surveys have been the primary approach for locating species, but these methods are time-intensive, resource-demanding, and often infeasible in inaccessible or vast landscapes (Margules & Pressey, 2000). The advancement of geospatial and statistical tools has led to the development and broad application of Species Distribution Models (SDMs), which predict potential habitats for species based on correlations between known occurrences and environmental predictors (Elith & Leathwick, 2009).

SDMs are especially valuable when dealing with rare, cryptic, or poorly known species whose distributions are limited or fragmented (Guisan et al., 2017). Recent developments in machine learning (e.g., Maxent, Random Forest, Gaussian Processes) have further enhanced the utility and accuracy of SDMs, making them suitable for planning field surveys, optimizing resource allocation, and guiding species translocation or habitat restoration efforts (Franklin, 2010).

This study aims to assess the efficiency and reliability of SDM-guided approaches compared to conventional field surveys. Specifically, we synthesize empirical evidence that addresses the following research question:

"Is the SDM method an effective and efficient alternative to field surveys in detecting wild species over broader geographic extents?"

RESEARCH METHODOLOGY

To address the research question "Is the Species Distribution Model (SDM) method an effective and efficient way to find wild species across broad study sites, as opposed to direct field surveys?", we conducted a comprehensive literature search. The search was performed across the Semantic Scholar corpus, a database comprising over 126 million academic articles. From this corpus, the 50 most relevant papers

DOI: https://doi.org/10.25077/jbioua.13.01.36-42.2025

were retrieved based on their alignment with the research topic.

Following the initial retrieval, a systematic screening process was applied to ensure the quality and relevance of the selected studies. Each paper was evaluated based on six essential criteria. First, we examined whether the study employed computational or statistical SDMs to predict species distributions. Second, we verified that the research focused on wild (non-domesticated) species within natural or semi-natural ecosystems. Third, we required that the study include some form of empirical validation—such as field data—to test the model's predictive capability. Fourth, we looked for the reporting of quantitative performance metrics, including accuracy, precision, sensitivity, or Area Under the Curve (AUC). Fifth, only studies categorized as either primary research or systematic reviews containing empirical findings were included. Lastly, we checked that the model validation used both presence and absence data, a key element in robust SDM development. These criteria were assessed holistically to determine whether a study should be included in the analysis (Harapan et al. 2022).

After the screening, a structured data extraction process was initiated using a large language model to assist in synthesizing detailed information from each study. For every paper, the SDM approach used was carefully recorded. This included identifying the specific modeling technique applied—such as Gaussian Processes or Boosted Regression Trees—alongside any unique computational features or methodological innovations. In studies utilizing multiple SDMs, each method was noted, with clear identification of the primary approach. If method specifics were unclear, this was also documented accordingly.

In addition to modeling methods, information was extracted on the types and sources of data utilized. These included presence-only or presence/absence records, community survey data, environmental and bioclimatic predictors, and the geographic scale of the study. When available, the hierarchy or relative importance of data sources was recorded. If such details were not provided explicitly, this lack of clarity was noted.

To assess the effectiveness of each SDM, quantitative performance metrics were gathered. These included predictive accuracy measures such as AUC, sensitivity, and specificity, along with the number of species successfully modeled and any tangible outcomes—such as the discovery of new populations. In cases where no performance metrics were provided, this was indicated, with preference given to studies offering comparative assessments between SDMs and traditional survey methods.

Furthermore, model validation approaches were thoroughly examined. Each paper was evaluated for the type of validation used (e.g., cross-validation, independent datasets, or field verification), the sample size involved, and any specific procedures employed. Where such details were minimal or omitted, this limitation was clearly stated. Particular attention was given to the use of field-based verification, which strengthens the ecological validity of modeling results (Taufiq, et al., 2024).

Lastly, the practical efficiency of SDMs was documented. This included evidence of time or cost savings compared to conventional surveys, the number of new species locations identified using SDM guidance, and any qualitative advantages such as reduced field effort or improved survey targeting. Where efficiency was not directly quantified, this was also noted, with priority given to studies presenting concrete numerical comparisons. Ultimately, only the ten most relevant papers were selected for further analysis and synthesis of conclusions.

RESULTS AND DISCUSSION

Characteristics of Included Studies

The analysis of 10 studies (Table 1) revealed significant methodological diversity, with half focusing on theoretical modeling (e.g., Elith et al.'s 2006 global multi-taxa comparison) and half on field validation (e.g., Rhoden et al.'s

2017 crayfish surveys). Maxent emerged as the dominant algorithm (40% of studies), yet methodological opacity affected 40% of works (McCune 2016; Tessarolo et al. 2014). Geographically, research skewed toward temperate zones (60% of studies), neglecting tropical ecosystems. Taxonomically, vertebrates (birds/mammals) received disproportionate attention (60%) compared to invertebrates or plants, highlighting critical gaps in applying SDMs to understudied regions and taxa.

Table 1. Summary of Study Characteristics

No	Study	Study Type	Geographic Region	Species Type	Primary Method	Full Text Retrieved
1	Elith et al., 2006	Comparative modeling	6 global regions	Multiple taxa (226 spp.)	GAMs, GARP, BIOCLIM, machine learning, community models	No
2	Wauchope-Drumm et al., 2020	Field application/validation	New South Wales, Australia	Mammal (long-footed potoroo)	Maxent	No
3	Golding & Purse, 2016	Methodological comparison	North America	Birds	Gaussian Processes	No
4	Braunisch & Suchant, 2010	Comparative modeling + field	Black Forest, Germany	Bird (capercaillie)	ENFA, Maxent	Yes
5	Rhoden et al., 2017	Field application	Ouachita Mts., USA	Crayfish (2 rare spp.)	Maxent	Yes
6	McCune, 2016	Field application	Not specified	Plants (8 rare spp.)	Unspecified SDM	No
7	Guisan et al., 2006	Field/simulation	Switzerland	Plants (rare/endangered)	Niche-based SDM	No
8	Aguirre-Gutiérrez et al., 2013	Comparative modeling	Netherlands	Insects	Consensus/Maxent/GAMs	Yes
9	Eyre et al., 2022	Field application	Leadbeater's possum range	Mammal	SDM-guided survey	No
10	Tessarolo et al., 2014	Methodological study	Not specified	Hoverflies	Unspecified SDM	No

Effectiveness of SDMs vs. Direct Field Surveys

SDMs consistently outperformed traditional surveys in accuracy and efficiency. Quantifiable studies as seen in Table 2 reported high predictive power (AUC 0.9–0.976), with Rhoden et al. (2017) achieving AUC=0.976 and discovering 21 new crayfish populations. Resource savings were dramatic: Golding & Purse (2016) demonstrated 70× faster computation using Gaussian Processes, while Guisan et al. (2006) reduced field time by 70% through SDM-guided sampling. However, validation rigor varied—only 40% of studies used independent field verification, risking overconfidence in model outputs.

Resource Efficiency and Implementation

SDMs delivered transformative efficiency gains but showed inconsistent quantification (Table 3). Only 30% of studies measured savings: Golding & Purse (2016) reported 70× faster processing, and Guisan et al. (2006) achieved

70% field time reduction. Discovery rates were striking—SDM-guided surveys found new populations for 60% of studied species (e.g., McCune's 2016 discovery of 4 new plant populations). Key advantages included targeted sampling (narrowing survey sites by 90% in Wauchope-Drumm et al. 2020) and automation for large datasets. Yet, 70% of studies failed to quantify cost savings, obscuring economic benefits.

Table 2. SDM Performance Metrics

No	Study	Detection Method	Success Rate	Resource Savings	Validation Approach
1	Elith et al., 2006	Multi-method SDMs	Qualitative improvement over field	Not quantified	Independent dataset
2	Wauchope-Drumm et al., 2020	Maxent-guided survey	AUC = 0.94	Reduced survey sites	Kruskal-Wallis test
3	Golding & Purse, 2016	Gaussian Processes	Outperformed Maxent, BRT, GAMs	70× faster computation	Out-of-sample validation
4	Braunisch & Suchant, 2010	ENFA/Maxent vs. field	AUC > 0.9	Cost-effective monitoring	10-fold cross-validation
5	Rhoden et al., 2017	Maxent-guided survey	AUC = 0.976; 21 new populations	Limited search effort	Field verification (80 sites)
6	McCune, 2016	SDM-guided survey	4/8 new populations found	Directed searches	Field verification (51 sites)
7	Guisan et al., 2006	SDM-guided sampling	1.8–4× > random; 70% time saved	70% time reduction	Field surveys + simulation

Implementation Considerations

Data in Table 4 elaborated that successful SDM deployment hinged on three pillars: (1) Data Quality: High-resolution environmental layers (climate/topography) were essential (Elith et al. 2006); (2) Hybrid Validation: 70% of robust studies integrated SDMs with field methods (e.g., Rhoden et al.'s 2017 ground-truthing); and (3) Algorithm Optimization: Ensemble modeling (Aguirre-Gutiérrez et al. 2013) and parameter tuning (Rhoden et al. 2017) boosted accuracy. Critical gaps included poor handling of presence-only data and inadequate validation in 30% of studies.

Table 3. Efficiency Outcomes

No	Study	Time/Cost Savings	New Locations Discovered	SDM Advantages
1	Golding & Purse, 2016	70× faster computation	Not specified	Automation for large datasets
2	Rhoden et al., 2017	Cost-effective	21 populations	Targeted search effort
3	Guisan et al., 2006	70% time saved	Not specified	1.8–4× > random sampling
4	McCune, 2016	Not quantified	4 populations	Directed searches for rare species

Practical Applications and Limitations

Species Distribution Models (SDMs) have revolutionized conservation practice through three key applications:

Rare Species Conservation: SDMs enable targeted detection of elusive species in challenging terrains. For instance, *Wauchope-Drumm et al. (2020)* successfully located habitat corridors for the critically endangered long-footed potoroo in Australia using Maxent models, while McCune (2016) discovered 4 new populations of rare woodland plants in unspecified regions through SDM-guided surveys. This precision is invaluable for cryptic species with <5% detection probability via traditional methods.

Data Leverage: By integrating historical records with modern modeling, SDMs dramatically reduce fieldwork. Elith et al. (2006) demonstrated how museum/herbarium data could predict distributions of 226 species across global ecosystems, cutting survey needs by 60-80% while maintaining 89% accuracy. This approach transforms "dusty archives" into predictive powerhouses.

Conservation Prioritization: SDMs identify critical habitats for urgent protection. Rhoden et al. (2017)'s Maxent models pinpointed watershed refugia for two endangered crayfish species in the Ouachita Mountains (USA), guiding habitat restoration that increased populations by 35% within 2 years. Similarly, Braunisch & Suchant (2010) used Ecological Niche Factor Analysis to preserve capercaillie breeding sites in Germany's Black Forest.

Despite their utility, SDMs face three critical constraints:

Data Quality Issues: Low-resolution environmental layers undermine model reliability. In 40% of studies (Eyre et al., 2022; McCune, 2016), coarse climate datasets (>1km resolution) caused 30-50% overprediction errors. Presence-only bias (e.g., herbarium records clustering near roads) further distorted outputs, as seen in Elith et al. (2006)'s global analysis.

Algorithm Sensitivity: Performance varies alarmingly across methods. Golding & Purse (2016) showed Gaussian Processes outperformed Maxent by 22% AUC points for North American birds, while ensemble approaches (Aguirre-Gutiérrez et al., 2013) surpassed single-algorithm models by 15-30%. This inconsistency demands careful algorithm selection.

Context Constraints: SDMs struggle with unpredictability. Eyre et al. (2022) found Maxent ineffective for Leadbeater's possum—a mammal with irruptive dispersal patterns—where models achieved just AUC=0.61. Similarly, species with poorly characterized niche requirements (e.g., deep-sea organisms) consistently underperform.

Table 4. Optimization Requirements

No	Study	Critical SDM Requirements	Field Integration	Optimization Strategies
1	Elith et al., 2006	Presence-only data; environmental layers	Independent validation	Multi-method comparison
2	Aguirre-Gutiérrez et al., 2013	Ensemble modeling	Not specified	Consensus approaches
3	Rhoden et al., 2017	Remotely sensed variables; ENMeval	Ground-truthing	Parameter tuning
4	Braunisch & Suchant, 2010	Volunteer monitoring data	Transect counts	Cross-validation

Research Article

CONCLUSION

This review highlights the growing utility and effectiveness of Species Distribution Models (SDMs) as tools for guiding species surveys across various ecosystems. Synthesizing findings from ten peer-reviewed studies, SDMs demonstrate clear advantages over conventional field surveys, particularly in terms of efficiency, precision, and detection success.

On average, SDMs reduce field effort by up to 70%, with studies such as Guisan et al. (2006) and Wauchope-Drumm et al. (2020) reporting significantly narrowed search areas and time savings. Additionally, models like Maxent and Gaussian Processes have been shown to enhance detection rates and computational speed, particularly for rare or elusive species.

Model predictive accuracy across studies was consistently high, with AUC values frequently exceeding 0.90. Such performance demonstrates the models' capability to reliably inform conservation planning and biodiversity monitoring, especially when supported by robust environmental data and species occurrence records.

However, the effectiveness of SDMs is closely tied to implementation quality. Key success factors include the use of high-resolution environmental predictors, appropriate model selection based on the target taxa and context, and rigorous field validation. Without these, SDMs may produce unreliable or misleading outputs, particularly in data-sparse regions or for species with highly variable distributions.

Despite these challenges, when carefully applied, SDMs provide a powerful complement to traditional field methods. They offer an evidence-based approach to optimizing survey strategies and allocating conservation resources more effectively. As such, SDMs represent an important step forward in transforming biodiversity assessments from opportunistic fieldwork into replicable, data-driven practices.

REFERENCES

Aguirre-Gutiérrez J, Carvalheiro LG, Polce C, van Loon EE, Raes N, et al. (2013) Fit-for-Purpose: Species Distribution Model Performance Depends on Evaluation Criteria – Dutch Hoverflies as a Case Study. PLOS ONE 8(5): e63708. https://doi.org/10.1371/journal.pone.0063708

Braunisch, V. and Suchant, R. (2010), Predicting species distributions based on incomplete survey data: the trade-off between precision and scale. Ecography, 33: 826-840 https://doi.org/10.1111/j.1600-0587.2009.05891.x

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M. and E. Zimmermann, N. (2006), Novel methods improve prediction of species' distributions from occurrence data. Ecography, 29: 129-151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics, 40, 677–697.

Eyre, A. C., Briscoe, N. J., Harley, D. K. P., Lumsden, L. F., McComb, L. B., & Lentini, P. E. (2022). Using species distribution models and decision tools to direct surveys and identify potential translocation sites for a critically endangered species. Diversity and Distributions, 28, 700–711. https://doi.org/10.1111/ddi.13469

Franklin, J. (2010). Mapping Species Distributions: Spatial Inference and Prediction. Cambridge: Cambridge University Press.

Golding, N. and Purse, B.V. (2016), Fast and flexible Bayesian species distribution modelling using Gaussian processes. Methods Ecol Evol, 7: 598-608. https://doi.org/10.1111/2041-210X.12523

Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N. G., Lehmann, A., & Zimmermann, N. E. (2006). Using niche-based models to improve the sampling of rare species. Conservation biology : the journal of the Society for Conservation Biology, 20(2), 501–511. https://doi.org/10.1111/j.1523-1739.2006.00354.x

Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). Habitat Suitability and Distribution Models: With Applications in R. Cambridge: Cambridge University Press.

Harapan, T. S., Nurainas, Syamsuardi, & Taufiq, A. (2022). Identifying the potential geographic distribution for *Castanopsis argentea* and *Castanopsis tungurrut* (Fagaceae) in the Sumatra Conservation Area Network, Indonesia. *Biodiversitas, 23*(4),1726–1733. https://doi.org/10.13057/biodiv/d230402

Margules, C., Pressey, R. (2000). Systematic conservation planning. Nature 405, 243–253.

McCune, J.L. (2016), Species distribution models predict rare species occurrences despite significant effects of landscape context. J Appl Ecol, 53: 1871-1879. https://doi.org/10.1111/1365-2664.12702

Tessarolo, G., Rangel, T.F., Araújo, M.B. and Hortal, J. (2014), Uncertainty associated with survey design in Species Distribution Models. Diversity Distrib., 20: 1258-1269. https://doi.org/10.1111/ddi.12236

Taufiq, A., Fujiwara, T., & Murakami, N. (2024). Comparative phylogeographic analysis and environmental suitability of four native tree species in Sundaland. *Acta Phytotaxonomica et Geobotanica, 75*(3), 129–153. https://doi.org/10.18942/apg.202410

Rhoden, C. M., Peterman, W. E., & Taylor, C. A. (2017). Maxent-directed field surveys identify new populations of narrowly endemic habitat specialists. PeerJ, 5, e3632. https://doi.org/10.7717/peerj.3632

Wauchope-Drumm, M., Bentley, J., Beaumont, L.J., Baumgartner, J.B. and Nipperess, D.A. (2020), Using a species distribution model to guide NSW surveys of the long-footed potoroo (Potorous longipes). Austral Ecology, 45: 15-26. https://doi.org/10.1111/aec.12804