Spatial-temporal distribution of *Anopheles* habitats in the Greater Mekong Subregion using remote sensing

Robert Zupko

Abstract: Malaria due to the *Plasmodium falciparum* parasite and transmitted by the *Anopheles* mosquitoes continues to be a grave concern. Within the Greater Mekong Subregion (GMS) of Southeast Asia, efforts towards malaria elimination have reduced the overall malaria burden, and a better understanding of where *Anopheles* mosquitoes may be could contribute to these efforts. One means of accomplishing this is through the use of remote sensing, coupled with bioclimatic envelopes, to determine habitat for *Anopheles* mosquitoes within the GMS. In this report, we explore how Google Earth Engine can be utilized in the capacity with an emphasis on the processing workflow used for habitat classification. Preliminary examination of how *Anopheles* habitat has changed over time suggests that all species studied in the GMS have seen an expansion in habitat, although more work is needed to ensure that the habitat identified is accurate and not the result of an artifact introduced during the processing workflow.

1 Introduction

Malaria due to the *Plasmodium falciparum* parasite was responsible for 241 million cases and 627 thousand deaths globally in 2020 (World Health Organization, 2021). While the majority of the disease burden is within Sub-Saharan Africa, the Greater Mekong Subregion (GMS) reported 65 thousand cases in 2021 and a total of 15 deaths in 2020 (The Mekong Malaria Elimination programme, 2022). The GMS region consists of Cambodia, Laos, Myanmar, Thailand, and Vietnam, along with Yunnan Province and the Guangxi Zhuang Autonomous Region of China and has seen declining malaria cases and mortality due to extensive efforts targeting malaria elimination by 2023. In 2021, China was certified as malaria free by WHO, and the remaining countries in the region continue to progress to elimination with Myanmar presently having the highest case burden with 53 thousand cases, or 82% of all cases in the GMS (The Mekong Malaria Elimination programme, 2022). While this progress is encouraging, the GMS region continues to be of broader concern due to identification of artemisinin resistance in the region (Ashley et al., 2014; Imwong et al., 2020). Presently, *P. falciparum* malaria within the GMS is commonly associated with forests and forested fringes (Vantaux et al., 2021), due to its associated *Anopheles* mosquito vector, thus allowing for remote sensing to be leveraged to identify possible areas of risk.

The *P. falciparum* parasite is an obligate parasite that requires both a human host and mosquito host as part of its lifecycle (Lang-Unnasch & Murphy, 1998). Infection of a human host occurs when a *Anopheles* mosquito carrying *Plasmodium* sporozoites takes a blood meal and the sporozoites migrate to the liver (Meibalan & Marti, 2017). During the liver stage, *Plasmodium* replicates until releasing merozoites into the blood stage, which in turn infect red blood cells and progress through ring, trophozoite, and schizont stages before new merozoites are released for further infection of red blood cells. During the blood stage, a limited number of parasites will proceed to the sexual gametocyte stage, at which point they migrate to the dermis allowing them to be taken up when an *Anopheles* mosquito takes a blood meal. While sporozoites may be cleared by the human host prior to the liver stage (Churcher et al., 2017), infections are typically cleared during the blood stage and may or may not present symptoms. In the case of uncomplicated, symptomatic cases, *P. falciparum* malaria typically presents with non-specific symptoms of fever, chills, muscle aches, loss of appetite, headache, and cough (Zekar & Sharman, 2023). Progression to severe or fatal cases occurs in 1% to 2% of infections and is more likely to occur in children under five (Milner et al., 2020).

Transmission of the *P. falciparum* parasite is solely through female *Anopheles* mosquitoes who take a blood meal; however, not all of the species are efficient at transmission as a result, only 30 to 40 *Anopheles* mosquito species are commonly associated with the parasite (Nicoletti, 2020). Within the GMS, the *Dirus* complex is the common, and occurs in forest and forest-fringes, although precisely distinguishing between members of the complex can be a challenge (Hii & Rueda, 2013; Obsomer et al., 2007). A growing concern is the possibility of range change in the *Anopheles* mosquito due to climate change (Karypidou et al., 2020), with rapid range shifts already having been demonstrated for African *Anopheles* mosquitoes (Carlson et al., 2023), and shifts in *Anopheles dirus* predicted for the South-East Asia and Western Pacific regions under various climate change scenarios (Liu et al., 2022). While the spatial distribution of *Anopheles* allows for significant heterogeneity in terms of where they can be found, specific *Anopheles* species are sensitive to their habitat (Castro, 2017). This leads to the association between *Anopheles* and forests in the GMS, when other bioclimatic factors are favorable (Durnez et al., 2013; Hii & Rueda, 2013; Obsomer et al., 2007, 2012).

Due to *Anopheles* mosquitoes being the sole vector for the *P. falciparum* parasite, it is possible to use remote sensing and geographic information systems (GIS) technologies to identify favorable habitat for the vector, thus isolating regions where *P. falciparum* malaria may be present as well. At broad level, remote sensing coupled with land cover classification has already proven effective at identifying forested regions and tracking forest cover change (Hansen et al., 2013). When narrowing focus to *Anopheles* mosquitoes, Ceccato et al. (2005) offered an early review of how remote sensing techniques were used to

predict the likely location of *Anopheles* mosquitoes in the African context. Centering on GMS, Obsomer et al. (2012) extended previous work in identifying the bioclimatic factors (Obsomer et al., 2007), to produce maps predicting the likely location of *Anopheles* mosquitoes. While their work was not dependent upon remote sensing, such data can be combined with artificial neural networks trained with historical patterns to develop a prediction model for malaria cases (Thakur & Dharavath, 2019). However, such models are limited since forward projections cannot consider malaria interventions (e.g., indoor residual spraying, long-lasting insecticide treated bed net distribution, etc.), although this is typically the case for all remote sensing applications. More recently, McMahon et al. (2021) used a combination of remote sensing and other geographic information (e.g., pre-computing land cover, topography, surface water, etc.) to assess the risk of malaria cases due to the *P. falciparum* and *P. vivax* parasites based upon the local geographic context.

This prior work suggests that remote sensing, partially when completed with regionally centered geographical models, may have applications in predicting the habitat of *Anopheles* mosquitoes which in turn has implications for addressing *P. falciparum* malaria. In situations where malaria survey data is limited or non-existent (Brousse et al., 2020), models produced using remotely sensed data may allow for better policy interventions, such as improving the distribution of mosquito nets (Acheson et al., 2015). In this report, my work in leveraging Google Earth Engine and remote sensing to identify *Anopheles* habitat within the GMS region will be reviewed, along with the current results, and future work to be conducting.

2 Methods

2.1 Bioclimatic Envelopes

Table 1. Table of parameters used based upon previous work by Obsomer et al. (2012)

	A. baimaii	*A. crascens*	*A. dirus s.s.*	*A. dirus s.l.*	*A. scanloni*
Annual Rainfall (mm)	≥ 1200	≥ 2000	≥ 1500	≥ 1500	≥ 1500
Minimum temperature, coldest month (°c)	12.5	20.0	12.5	11.0	15.0
Annual mean temperature (°c)	24.0 – 27.5	≥ 25.0	24.0 – 27.5	23.0 – 27.5	≥ 26.0
SD of annual mean temperature (°c)	0.5 – 2.5	0.0 – 1.0	0.5 – 2.5	1.0 – 5.0	0.0 – 2.0

	A. baimaii	*A. crascens*	*A. dirus s.s.*	*A. dirus s.l.*	*A. scanloni*
Maximum temperature, coldest month (°c)	28.0	28.0	26.0	24.0	28.0

The basis of the of the system is a bioclimatic envelope, which is used to predict the extremes in distributions of a species, based upon the parameterization for GMS *Anopheles* mosquitos by Obsomer et al. (2012), utilizing the parameters outlined in Table 1. While the authors examined eighteen different variables in identifying the ecological niche occupied by the *Anopheles* mosquito, five were highlighted as statistically significant. These include annual rainfall, the minimum temperature of the coldest month, the annual mean temperature, and the maximum temperature of the coldest month which defines the upper bounds of survivability for the mosquitoes. Given that a limited number of variables were found to be highly predictive of the preferred habitat for the various *Anopheles* species, it is possible to use a highly constrained number of data sources as part of the workflow (see Table 2). The *Anopheles* species selected as part of the workflow is based upon the most common vectors of *P. falciparum* malaria in the GMS region (Hii & Rueda, 2013; Obsomer et al., 2012) with *A. dirus sensu lato* allowing for a broader categorization than *A. dirus sensu stricto*.

Table 2. Data sources and source resolution

Data Source	**Resolution**	**Description**
CHIRPS Pentad	5566 meters, mm / pentad	Precipitation data derived from 0.05° resolution satellite imagery with *in situ* station data.
Landsat 7	Visible, NIR, SWIR: 30 meters Panchromatic: 15 m Thermal: 100 m	Used for imagery from 2000 to 2013
Landsat 8		Used for imagery from 2014 onwards
Terra Land Surface Temperature (MOD11A1.061)	1000 meters	Daytime land surface temperature in Kelvin

2.2 Google Earth Engine

Google Earth Engine is cloud-based platform that hosts satellite imagery and other global-scale geographic data sets, with additional analysis and visualization tools that can be used for data analysis. As

part of Google Earth Engine, Landsat and Sentinel-2 collections from Google Cloud Storage are exposed for use, and other tools such as Google Compute Engine and Google Cloud Machine Learning are leveraged for use by Google Earth Engine Users.

Within the Google Earth Engine envelopment, there are three options for executing code. First is on-demand from within the Google Earth Engine development environment,[1,2] which provides both JavaScript development tools tailored to the Google Earth Engine environment, but also a map view and task manager for submitting longer running jobs to the Google Cloud environment. From the development environment it is possible to deploy applications (apps) that may be run by other users without needing access to the Google Earth Engine, offering the second means of access. However, users of apps are limited to the functionality that is implemented by the developer, similar to other web applications and desktop applications. Finally, Google Earth Engine offers a Python application programming interface (API) that allows access to the same computational functionality that can be accessed via the JavaScript development environment, including a means bulk queuing longer running jobs for processing within Google Cloud.[3]

2.3 Processing Workflow

While the same fundamental workflow (Figure 1) is used regardless of how the application is used, how the workflow is invoked is dependent upon how the application is accessed. When the web application is loaded (see Plate 1), data processing and the habitat raster are generated in real time – subject to timeout constraints – allowing the user to control the selected species, year, and location within the bounds of the GMS. However, production of raster data for more in-depth analysis is done via the command line interface and requests for rasters are queued for generation in the Google Cloud environment.[4] In order to ensure consistency in the data used by both applications (e.g., training data, Landsat swaths covering the GMS, etc.) the JavaScript to Python functionality in the `geemap` library is used to convert relevant JavaScript files to Python with additional post-processing to ensure syntactically correct Python code.

---

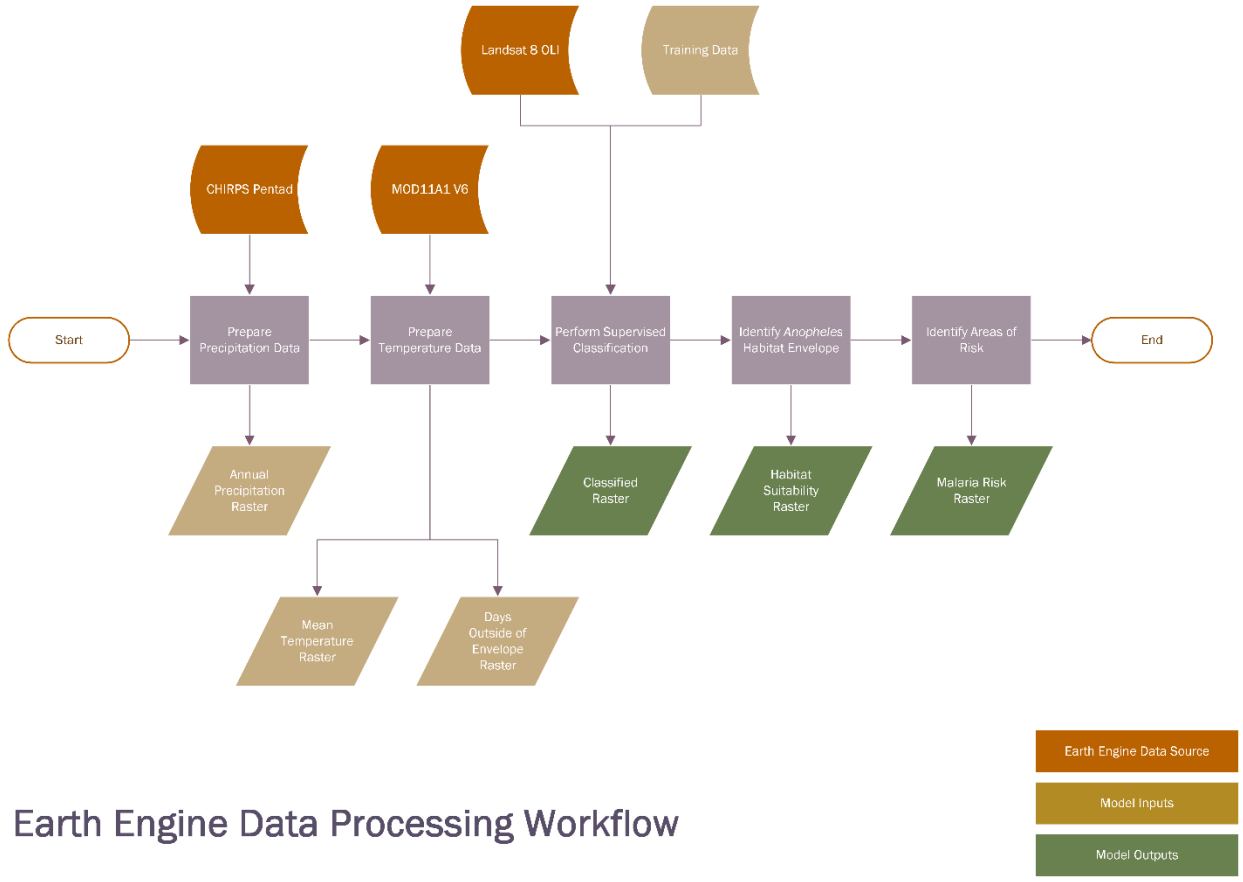
[1] https://code.earthengine.google.com/

[2] In order to use the Google Earth Engine development environment is necessary to have a Google account that has been granted access, either as a commercial or non-commercial (e.g., academic, non-profit, etc.) account

[3] While beyond the scope of this report, configuring the Python API can be a challenge and requires some working knowledge of Google Cloud configuration. For this project, a Google Cloud service account was created to queue the Google Earth Engine jobs, necessitating additional scripting to retrieve data from the Google Drive associated with the service account.

[4] The total jobs are limited 3,000 as of April 2023 and the number of jobs that are run simultaneously is dependent upon Google Cloud's resource allocation with paid subscribers receiving priority over academic users.

Figure 1. Data processing workflow implemented in Google Earth Engine. When the user uses the application, or a processing request is queued, imagery data is processed using the same workflow which utilizes CHIRPS Pentad precipitation, Terra Land Surface Temperature (MOD11A1.061) for the daytime, and Landsat imagery to habitat suitability.



The overarching data processing workflow can thus be summarized as follows:

1. Data specific to the year is generated:
   1.1. The CHIRPS Pentad within the GMS for the year is load and the total rainfall for the year is calculated, this is the *annual precipitation raster*.
   1.2. The Terra Land Surface Temperature within the GMS for the year is loaded and mean is calculated, this is the *mean temperature raster*.
   1.3. All available Landsat swaths within the GMS are loaded (e.g., date filter 2020-01-01 to 2020-12-31), cloud masking is applied, and median of the image stack is taken.
   1.4. A CART classifier is trained, and used to classify the cloud masked, median Landsat imagery, this produces the *classified raster*.
2. Data specific to the *Anopheles* species is generated:

2.1. The annual temperature raster is used to determine the days outside the temperature envelope for the species, this produces the *days outside of envelope raster*.

2.2. The preferred habitat for the species is calculated based upon the total rainfall, mean temperature, and days outside of the temperature envelop using raster algebra (Listing 1), this produces the *habitat suitability raster*.

3. After determination of habitat suitability, the *malaria risk raster* can be calculated based upon the proximity of development to possible *Anopheles* habitat.

```
// Use raster algebra to score the best habitat
exports.getHabitat = function(variables) {
  // Find the terrain that is within the basic bounds for the species
  var habitat = ee.Image(0).expression('(totalRainfall >= speciesRainfall)
&& (daysOutsideBounds <= 30)', variables);

  // Improve the score if terrain has the appropriate landcover (forest or
heavy vegetation)
  // and is within the mean annual temperature bounds
  if (variables.speciesMeanLower == variables.speciesMeanUpper) {
      habitat = habitat.expression('b(0) + \
        ((b(0) == 1) && \
         (landcover == 11 || landcover == 12) && \
         (meanTemperature >= speciesMeanLower))', variables);
  } else {
      habitat = habitat.expression('b(0) + \
        ((b(0) == 1) && \
         (landcover == 11 || landcover == 12) && \
         ((speciesMeanLower <= meanTemperature) && (meanTemperature <=
speciesMeanUpper)))', variables);
  }

  // Rename the band and return
  return habitat.rename('scored_habitat');
};
```

Listing 1. Source of the Earth Engine JavaScript raster algebra used to determine the bioclimatic envelope.

A major component of the workflow is landcover classification of the Landsat imagery through the use of Classification and Regression Trees (CART) analysis implemented in Google Earth Engine as `ee.Classifier.smileCart`. A CART is a decision tree algorithm that utilizes training data to predict the appropriate class for unknown values (Bittencourt & Clarke, 2004) and is one of the

techniques employed in land cover classification (Bittencourt & Clarke, 2003; Xu et al., 2005). In order to train the CART, five Landsat swaths were selected that were imaged by both the Landsat 7 and 8 platforms in 2020 and the cloud masked median of all over flights in the year was used as the basis of the training site. The imagery was then manually classified into one of nine classifications (snow, shadow / occulted, water, forest, vegetation, vegetation / scrub, barren, developed, and agricultural) based upon the broad spectral characteristics of a location.

Development of the workflow proceeded by first implementing it using the JavaScript-based Google Earth Engine development environment. In addition to allowing a web application to be deployed, thus allowing users to explore the possible habitat for *Anopheles* vectors within the GMS region (see Plate 1),[5] it also allowed for the workflow to be evaluated for functional correctness before it was replicated in Python in order to use the Google Earth Engine API for job queueing. A command line tool was then developed that permutates on the standard deviation range of the mean of mean temperature, in steps of 0.5°C, thus exploring the sensitivity of the predicted habitat to changes in the variable to be explored. Upon completion of the tool, jobs were then queued and downloaded for aggregation into tabular format and preparation of results.

3 Results

When comparing the 20-year change in habitat based upon the bioclimatic envelope using a 1 km^2 pixel resolution, all *Anopheles* species saw an increase with *A. baimaii* increasing by 3.5% to 219,488 km^2, *A. crascens* by 88.3% to 148,461 km^2, *A. dirus* s.l. by 2.7% to 292,193 km^2, *A. dirus* s.s. by 21.9% to 232,553 km^2, and *A. scanloni* by 126.6% to 132,386 km^2 (Results). While an expansion in *Anopheles* habitat has been forecast as a result of climate change (Karypidou et al., 2020; Liu et al., 2022), more rapid expansion by *A. crascens* and *A. scanloni* suggest that additional work is needed for verify that the bioclimatic envelope is being calculated correctly given that the species prefer higher temperatures.

[5] https://rzupko.users.earthengine.app/view/gms-malaria

Figure 2. Change in *Anopheles* habitat over time. Note that the change in imaging platform from Landsat 7 to Landsat 8 did not appear to have an appreciable difference in the total habitat (in contrast to landcover classifications [**Error! Not a valid bookmark self-reference.**]). This suggests that habitat is likely constrained by climatic factors as opposed to landcover.

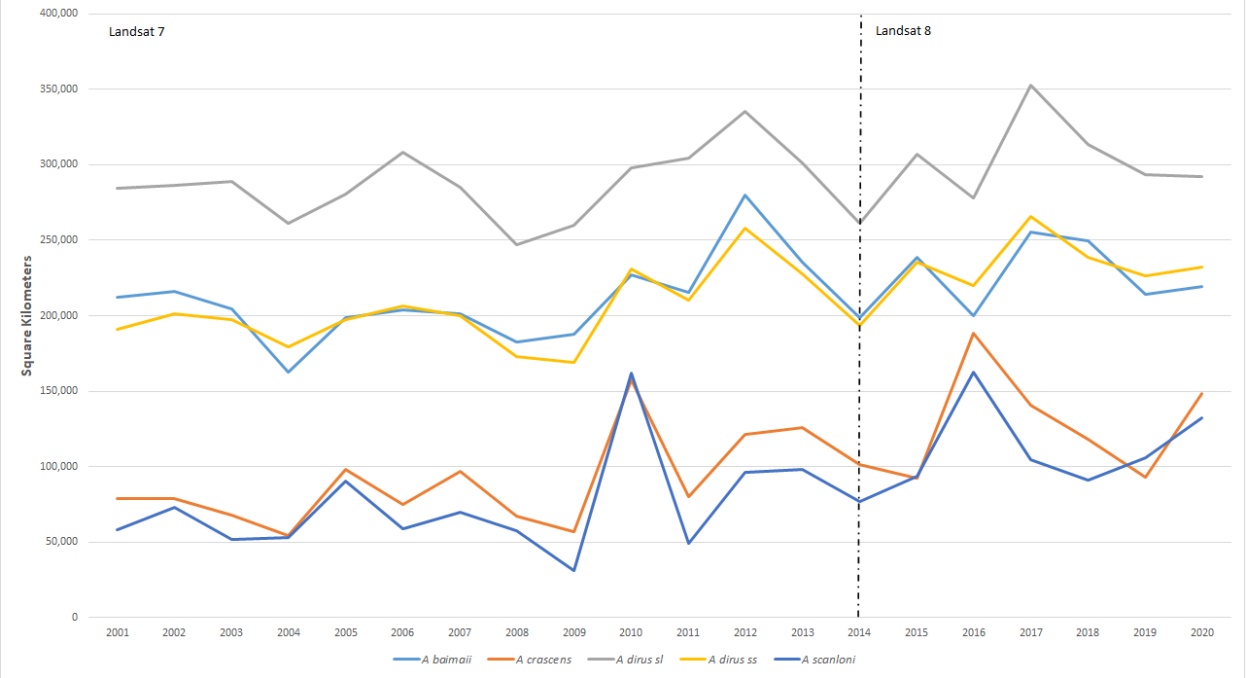
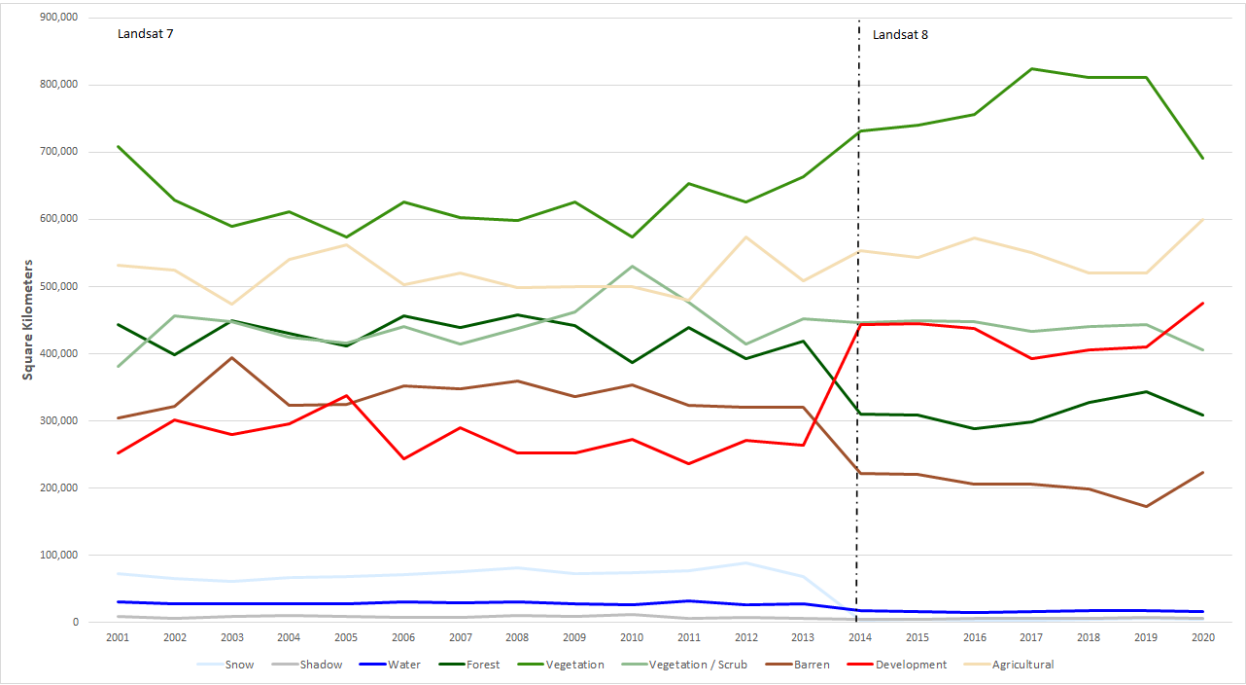


Figure 3. Difference in landcover classification over time. Note the difference in landcover classification after the shift from Landsat 7 to Landsat 8 in 2014.



Landcover classification is largely consistent with expectations with areas projected as forest aligning with that identified by Obsomer et al. (2012) as such (see Plate 2). Likewise, agricultural land use appears to be consistent with other imaging. However, the CART appears to have issues distinguishing between developed and barren land, with the more mountainous Yunnan Province and Guangxi Zhuang Autonomous Region of China in the northern GMS being misclassified as developed land. Contrasting the landcover classification of Landsat 7 (see Plate 3), this appears to be more common with imagery from Landsat 8, although more work is required to isolate the cause of the misclassification. Furthermore, while the *Anopheles* habitat demonstrates an increase over time, the landcover classifications are largely consistent within the context of a specific sensor (i.e., Landsat 7 or Landsat 8); however, when the system switches from Landsat 7 to Landsat 8 in 2014 there is a marked change in the area classified as developed, barren, and forest; while vegetation shows a general upward trend (Figure 2. Change in *Anopheles* habitat over time. Note that the change in imaging platform from Landsat 7 to Landsat 8 did not appear to have an appreciable difference in the total habitat (in contrast to landcover classifications [**Error! Not a valid bookmark self-reference.**]). This suggests that habitat is likely constrained by climatic factors as opposed to landcover.

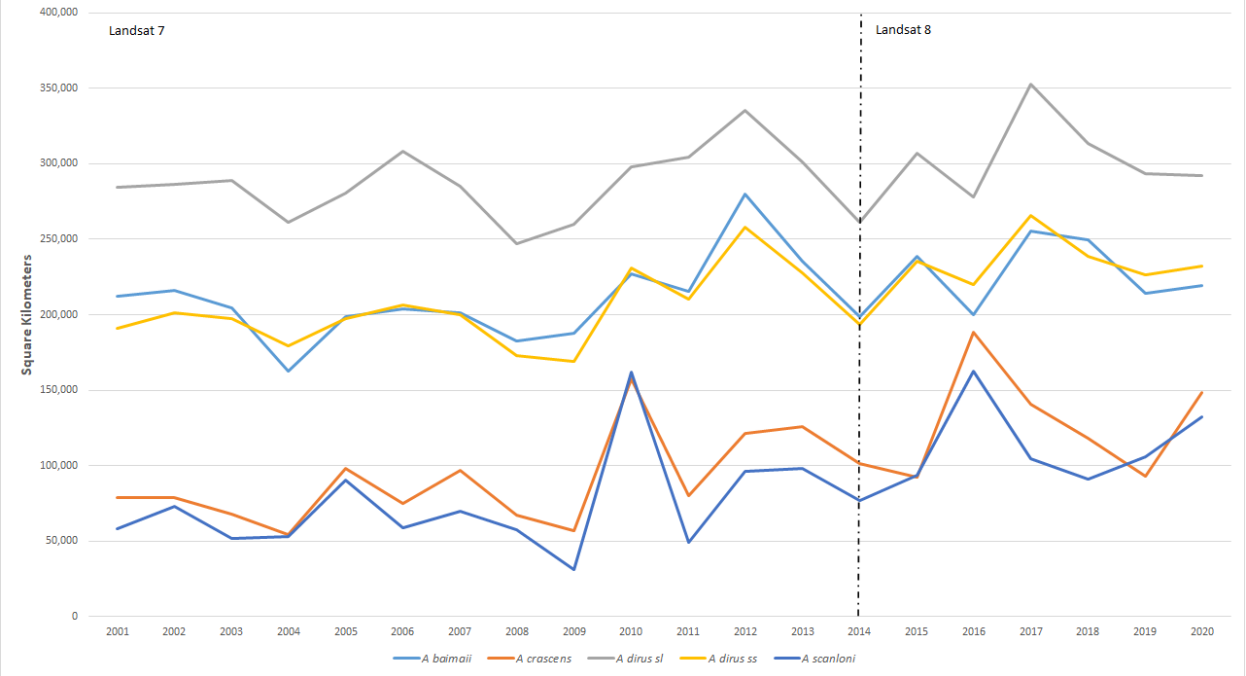


Figure 3). This is consistent with the misclassification noted; however, given that there is not a dramatic change in the *Anopheles* habitat during this time it is possible that the discrepancies in landcover may be isolated to regions that are outside the climatic boundaries of the species.

5 Discussion and Future Work

The processing workflow developed and deployed Google Earth Engine application is effective at producing habitat projections at 1 km^2 (see Plate 4) and demonstrates the effectiveness of leveraging cloud-based resources. However, three major areas of concern arose over the course of this project that

require additional exploration. The overarching objective of future work will be to address all of these concerns in the process of preparing a manuscript for peer review.

First, while most of the bioclimatic variables identified by Obsomer et al. (2012) offer loose boundary conditions (i.e., rainfall greater than a given amount), the mean annual temperature has an associated standard deviation, indicating that additional sensitivity analysis is required as part of the process of identifying the *Anopheles* habitat. This is demonstrated by Plate 5 which illustrates the habitat for *A. dirus s.s.* when the annual mean temperature is increased to ±2.5°C and further reinforced by Plate 6 which contains the difference after the two rasters subtracted from each other (i.e., habitat when ±2.5°C is added to the annual mean temperature range versus the identified range of 24 – 27.5°C).

Next, it is unclear as to how significant an effect an increase, or decrease, in pixel resolution will have upon the total habitat. As suggested by Plate 3, decreasing the pixel resolution from 1 km^2 to 25 km^2 results in an aggregation of underlying pixels, as well as an expansion of water pixels along clipping boundaries. In order to evaluate this, sensitivity analysis to the pixel resolution can be conducted to quantify the difference.

Finally, the last major area of concern regards the CART classification of landcover and the inconsistent classification between Landsat 7 and Landsat 8, coupled with the misclassification of landcover. While misclassifications are generally to be expected and can be addressed as part of the communication of results (Sandler & Rashford, 2018), the intent of the processing workflow was for the same basis in training data to be used for both Landsat 7 and Landsat 8. Therefore, it is reasonable to conclude that consistent results should be produced when transition from Landsat 7 to Landsat 8 imagery. However, the notable change in classifications for development, barren, forest, and snow (Figure 3) indicates that it necessary to isolate the cause of the problem. However, it is important to highlight the extent to which misclassification may or may not be a concern. Given that the bioclimatic envelope model, and the processing workflow both assume that *Anopheles* mosquitoes will be in forested regions, misclassification of forest as vegetation (or vice versa) would have a larger impact than barren land being misclassified as developed (or vice versa).

6 Conclusion

This report examined how Google Erath Engine was leveraged along with bioclimatic envelopes for *Anopheles* mosquitoes in the GMS region to identify their possible habitat. In addition to demonstrating the effectiveness of Google Earth Engine as a platform with multiple approaches to data processing (i.e.,

the developer environment, Python API, and deployed apps), a data processing workflow was produced and evaluated. While additional work is needed to assess the sensitivity of the habitat to the standard deviation in the range of the annual mean temperature determined by Obsomer et al. (2012), as well as the pixel resolution, the deployed web application is capable of allowing users to identify possible habitat at 1 km2 resolution from 2000 onwards. The next steps are to continue the sensitivity analysis as well as examination of the inconsistent classification and misclassification of landcover in preparation of a manuscript to be submitted to *Remote Sensing of the Environment.*

Availability of data and materials

All source code associated with the project can be found on GitHub at https://github.com/rjzupkoii/gms-malaria, the deployed Google Earth Engine application can be found at https://rzupko.users.earthengine.app/view/gms-malaria.

7 References

Acheson, E. S., Plowright, A. A., & Kerr, J. T. (2015). Where have all the mosquito nets gone? Spatial modelling reveals mosquito net distributions across Tanzania do not target optimal Anopheles mosquito habitats. *Malaria Journal*, *14*(1), 322. https://doi.org/10.1186/s12936-015-0841-x

Ashley, E. A., Dhorda, M., Fairhurst, R. M., Amaratunga, C., Lim, P., Suon, S., Sreng, S., Anderson, J. M., Mao, S., Sam, B., Sopha, C., Chuor, C. M., Nguon, C., Sovannaroth, S., Pukrittayakamee, S., Jittamala, P., Chotivanich, K., Chutasmit, K., Suchatsoonthorn, C., … White, N. J. (2014). Spread of Artemisinin Resistance in Plasmodium falciparum Malaria. *New England Journal of Medicine*, *371*(5), 411–423. https://doi.org/10.1056/NEJMoa1314981

Bittencourt, H. R., & Clarke, R. (2004). Feature selection by using classification and regression trees (CART). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.

Bittencourt, H. R., & Clarke, R. T. (2003). Use of classification and regression trees (CART) to classify remotely-sensed digital images. *IGARSS 2003. 2003 IEEE International Geoscience and Remote*

Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477), *6*, 3751–3753 vol.6. https://doi.org/10.1109/IGARSS.2003.1295258

Brousse, O., Georganos, S., Demuzere, M., Dujardin, S., Lennert, M., Linard, C., Snow, R. W., Thiery, W., & van Lipzig, N. P. M. (2020). Can we use local climate zones for predicting malaria prevalence across sub-Saharan African cities? *Environmental Research Letters*, *15*(12), 124051. https://doi.org/10.1088/1748-9326/abc996

Carlson, C. J., Bannon, E., Mendenhall, E., Newfield, T., & Bansal, S. (2023). Rapid range shifts in African Anopheles mosquitoes over the last century. *Biology Letters*, *19*, 20220365. https://doi.org/10.1098/rsbl.2022.0365

Castro, M. C. (2017). Malaria Transmission and Prospects for Malaria Eradication: The Role of the Environment. In D. F. Wirth & P. L. Alsonso (Eds.), *Malaria: Biology in the Era of Eradication* (pp. 1–12). Cold Spring Harbor Laboratory Press.

Ceccato, P., Connor, S., Jeanne, I., & Thomson, M. (2005). Application of geographical information systems and remote sensing technologies for assessing and monitoring malaria risk. *Parassitologia*, *47*(1), 81–96.

Churcher, T. S., Sinden, R. E., Edwards, N. J., Poulton, I. D., Rampling, T. W., Brock, P. M., Griffin, J. T., Upton, L. M., Zakutansky, S. E., Sala, K. A., Angrisano, F., Hill, A. V. S., & Blagborough, A. M. (2017). Probability of Transmission of Malaria from Mosquito to Human Is Regulated by Mosquito Parasite Density in Naïve and Vaccinated Hosts. *PLOS Pathogens*, *13*(1), e1006108. https://doi.org/10.1371/journal.ppat.1006108

Durnez, L., Mao, S., Denis, L., Roelants, P., Sochantha, T., & Coosemans, M. (2013). Outdoor malaria transmission in forested villages of Cambodia. *Malaria Journal*, *12*(1), 329. https://doi.org/10.1186/1475-2875-12-329

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.

O., & Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, *342*(6160), 850. https://doi.org/10.1126/science.1244693

Hii, J., & Rueda, L. M. (2013). Malaria vectors in the Greater Mekong Subregion: Overview of malaria vectors and remaining challenges. *The Southeast Asian Journal of Tropical Medicine and Public Health*, *44 Suppl 1*, 73–165; discussion 306-7. PubMed.

Imwong, M., Dhorda, M., Myo Tun, K., Thu, A. M., Phyo, A. P., Proux, S., Suwannasin, K., Kunasol, C., Srisutham, S., Duanguppama, J., Vongpromek, R., Promnarate, C., Saejeng, A., Khantikul, N., Sugaram, R., Thanapongpichat, S., Sawangjaroen, N., Sutawong, K., Han, K. T., … White, N. J. (2020). Molecular epidemiology of resistance to antimalarial drugs in the Greater Mekong subregion: An observational study. *The Lancet Infectious Diseases*, *20*(12), 1470–1480. https://doi.org/10.1016/S1473-3099(20)30228-0

Karypidou, M. C., Almpanidou, V., Tompkins, A. M., Mazaris, A. D., Gewehr, S., Mourelatos, S., & Katragkou, E. (2020). Projected shifts in the distribution of malaria vectors due to climate change. *Climatic Change*, *163*(4), 2117–2133. https://doi.org/10.1007/s10584-020-02926-9

Lang-Unnasch, N., & Murphy, A. D. (1998). Metabolic Changes of the Malaria Parasite During the Transition from Human to Mosquito Host. *Annual Review of Microbiology*, *52*(1), 561–590. https://doi.org/10.1146/annurev.micro.52.1.561

Liu, X., Song, C., Ren, Z., & Wang, S. (2022). Predicting the Geographical Distribution of Malaria-Associated Anopheles dirus in the South-East Asia and Western Pacific Regions Under Climate Change Scenarios. *Frontiers in Environmental Science*, *10*. https://www.frontiersin.org/articles/10.3389/fenvs.2022.841966

McMahon, A., Mihretie, A., Ahmed, A. A., Lake, M., Awoke, W., & Wimberly, M. C. (2021). Remote sensing of environmental risk factors for malaria in different geographic contexts. *International Journal of Health Geographics*, *20*(1), 28. https://doi.org/10.1186/s12942-021-00282-0

Meibalan, E., & Marti, M. (2017). Biology of Malaria Transmission. In D. F. Wirth & P. L. Alsonso (Eds.), *Malaria: Biology in the Era of Eradication* (pp. 27–41). Cold Spring Harbor Laboratory Press.

Milner, E. M., Kariger, P., Pickering, A. J., Stewart, C. P., Byrd, K., Lin, A., Rao, G., Achando, B., Dentz, H. N., Null, C., & Fernald, L. C. H. (2020). Association between Malaria Infection and Early Childhood Development Mediated by Anemia in Rural Kenya. *International Journal of Environmental Research and Public Health*, *17*(3). https://doi.org/10.3390/ijerph17030902

Nicoletti, M. (2020). Chapter Five—Three scenarios in insect-borne diseases. In M. Nicoletti (Ed.), *Insect-Borne Diseases in the 21st Century* (pp. 99–251). Academic Press. https://doi.org/10.1016/B978-0-12-818706-7.00005-X

Obsomer, V., Defourny, P., & Coosemans, M. (2007). The Anopheles dirus complex: Spatial distribution and environmental drivers. *Malaria Journal*, *6*(1), 26. https://doi.org/10.1186/1475-2875-6-26

Obsomer, V., Defourny, P., & Coosemans, M. (2012). Predicted Distribution of Major Malaria Vectors Belonging to the Anopheles dirus Complex in Asia: Ecological Niche and Environmental Influences. *PLOS ONE*, *7*(11), e50475. https://doi.org/10.1371/journal.pone.0050475

Sandler, A. M., & Rashford, B. S. (2018). Misclassification error in satellite imagery data: Implications for empirical land-use models. *Land Use Policy*, *75*, 530–537. https://doi.org/10.1016/j.landusepol.2018.04.008

Thakur, S., & Dharavath, R. (2019). Artificial neural network based prediction of malaria abundances using big data: A knowledge capturing approach. *Clinical Epidemiology and Global Health*, *7*(1), 121–126. https://doi.org/10.1016/j.cegh.2018.03.001

The Mekong Malaria Elimination programme. (2022). *Accelerating malaria elimination in the Greater Mekong* (WHO/UCN/GMP/MME/2022.01). World Health Organization. https://www.who.int/publications/i/item/WHO-UCN-GMP-MME-2022.01

Vantaux, A., Riehle, M. M., Piv, E., Farley, E. J., Chy, S., Kim, S., Corbett, A. G., Fehrman, R. L., Pepey, A., Eiglmeier, K., Lek, D., Siv, S., Mueller, I., Vernick, K. D., & Witkowski, B. (2021).

Anopheles ecology, genetics and malaria transmission in northern Cambodia. *Scientific Reports*, *11*(1), 6458. https://doi.org/10.1038/s41598-021-85628-1

World Health Organization. (2021). *World Malaria Report 2021*. World Health Organization. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021

Xu, M., Watanachaturaporn, P., Varshney, P. K., & Arora, M. K. (2005). Decision tree regression for soft classification of remote sensing data. *Remote Sensing of Environment*, *97*(3), 322–336. https://doi.org/10.1016/j.rse.2005.05.008

Zekar, L., & Sharman, T. (2023). Plasmodium Falciparum Malaria [Updated 2022 Aug 8]. In *StatPearls*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK555962/

Plate 1. Screenshot of the deployed Google Earth Engine web application.

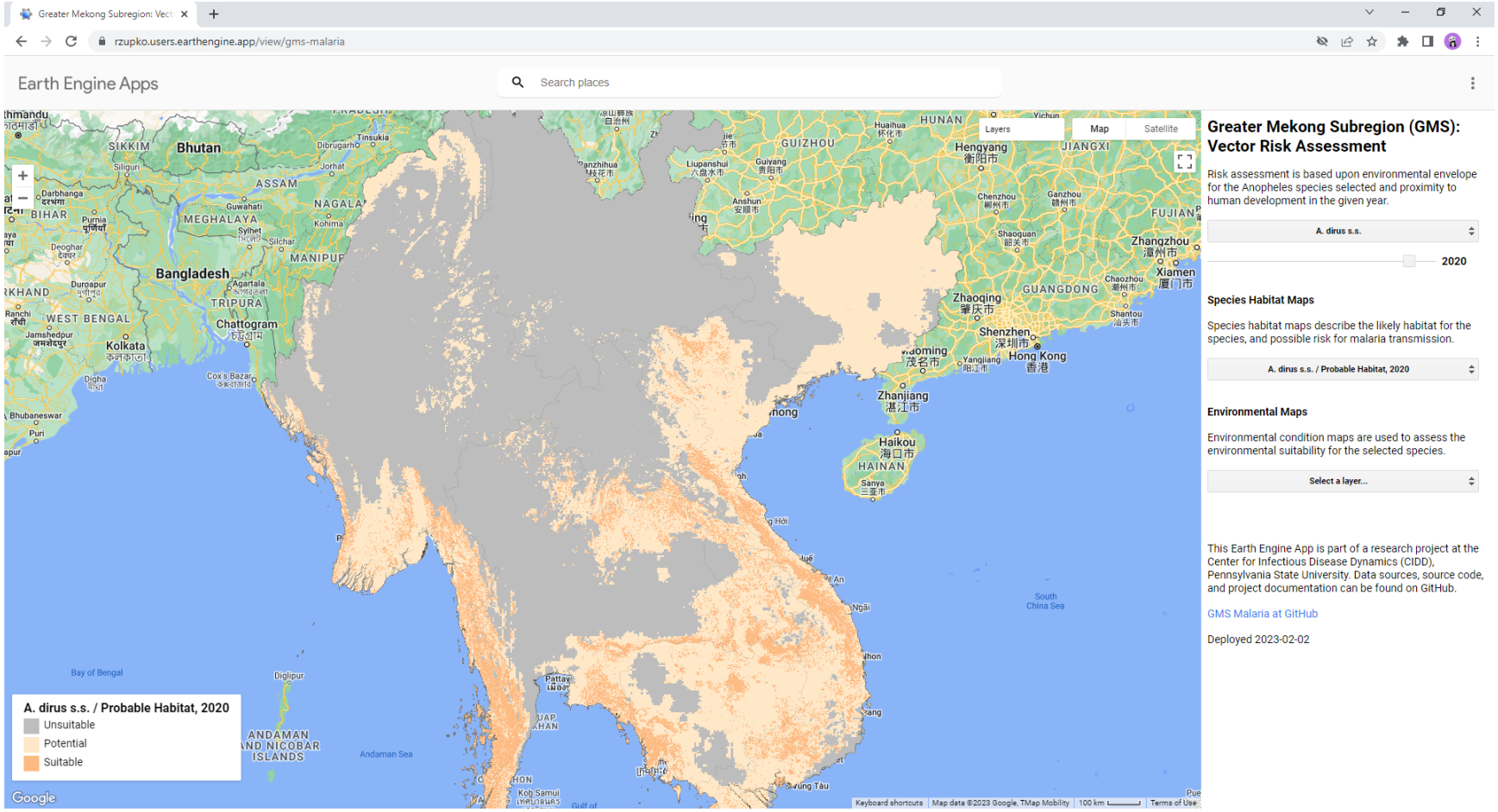


Plate 2. Landcover classification at 1km in 2020. Note that only a minimal amount of land is classified as snow in the north-north-east regions of the map and a significant portion of the mountains appear to be misclassified.

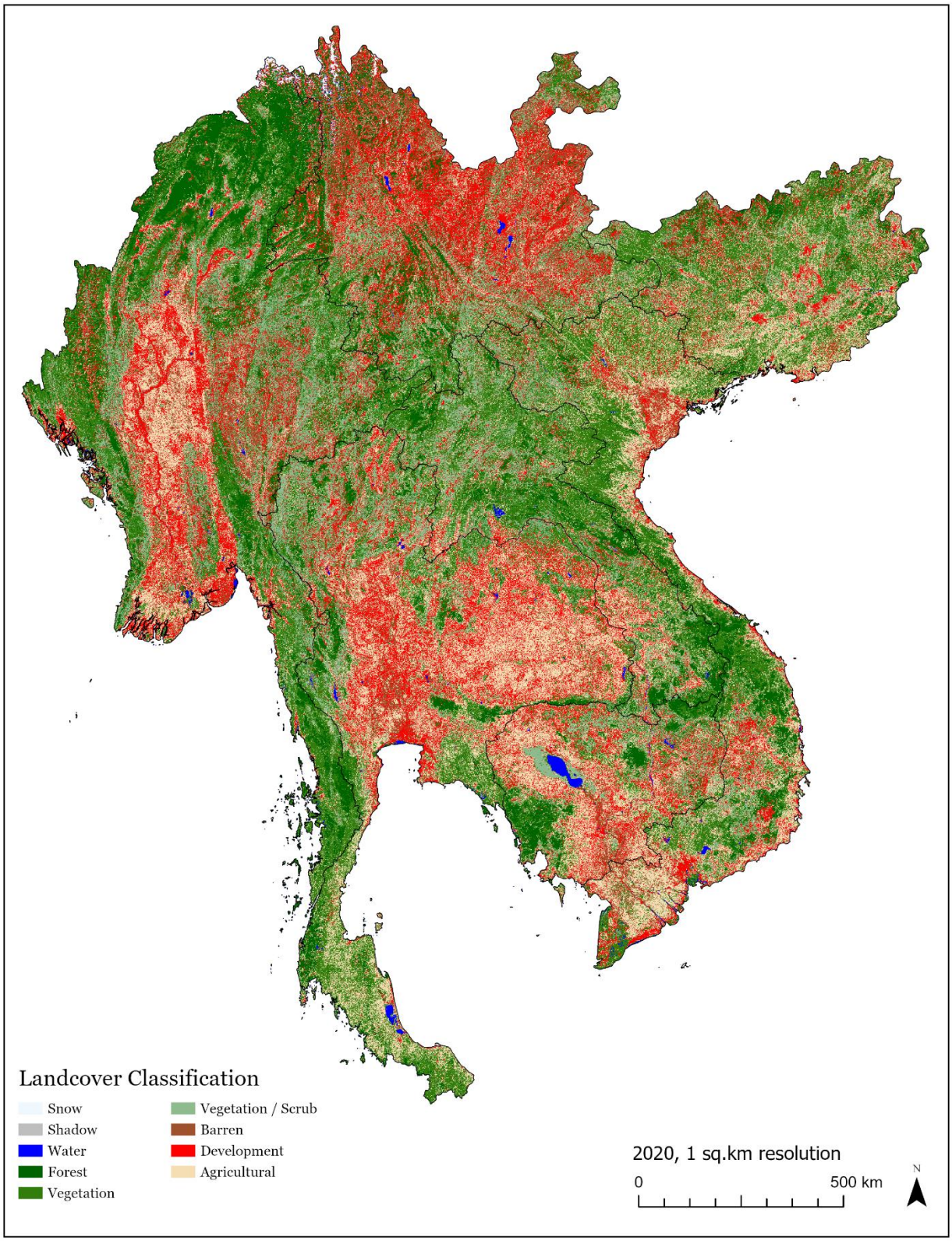


Plate 3. Illustration of how pixel resolution impacts landcover. Produced using classified Landsat 7 imagery from 2001, note that as the pixel resolution decreases from 1 km2 (left) to 25 km2 (right), Google Earth Engine aggregates classifications down to the lower resolution and additional water pixels may appear along clipping boundaries (i.e., coastlines).

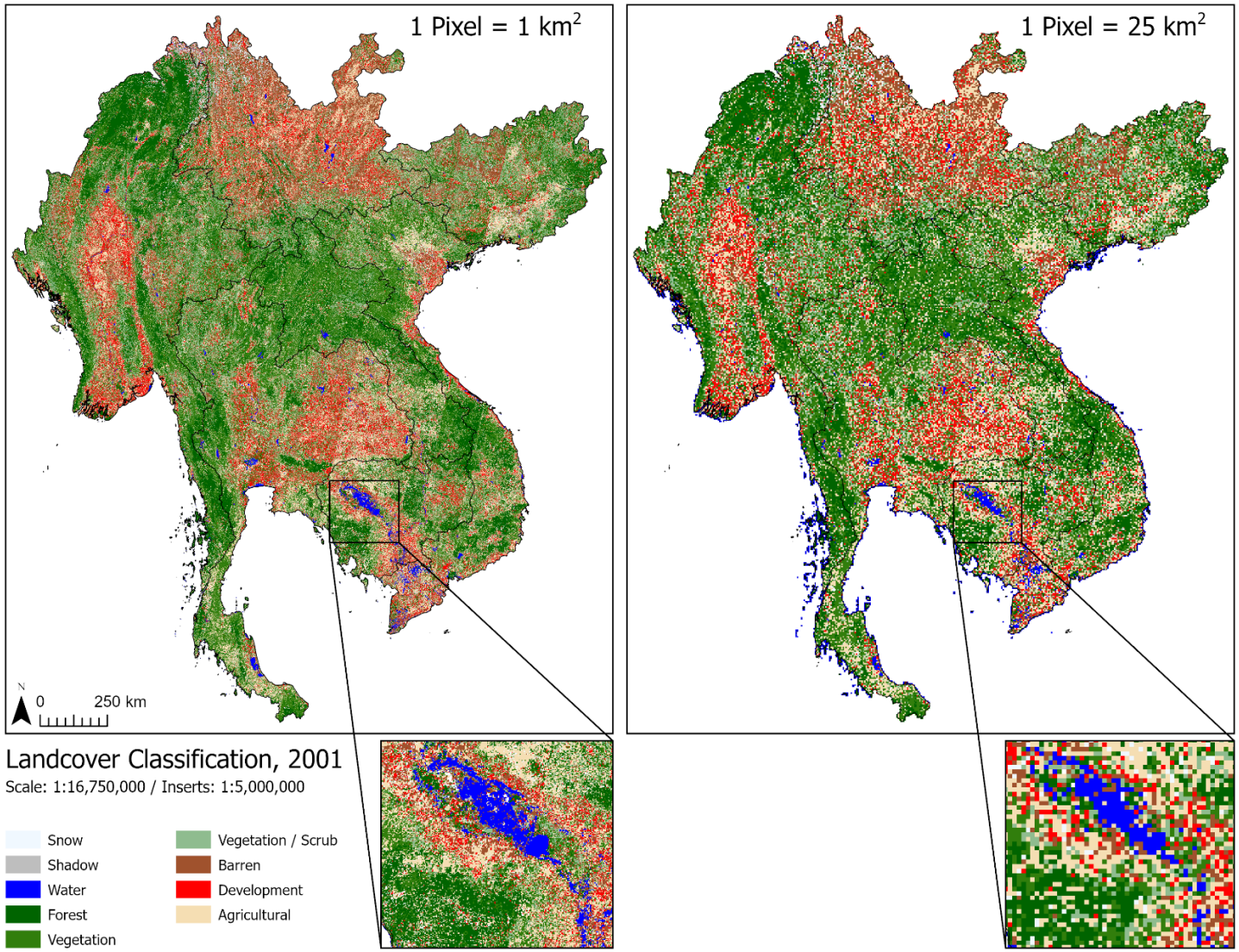


Plate 4. Habitat map for *A. dirus s.s.* at 1 km^2 resolution with no sensitivity to annual mean temperature applied.

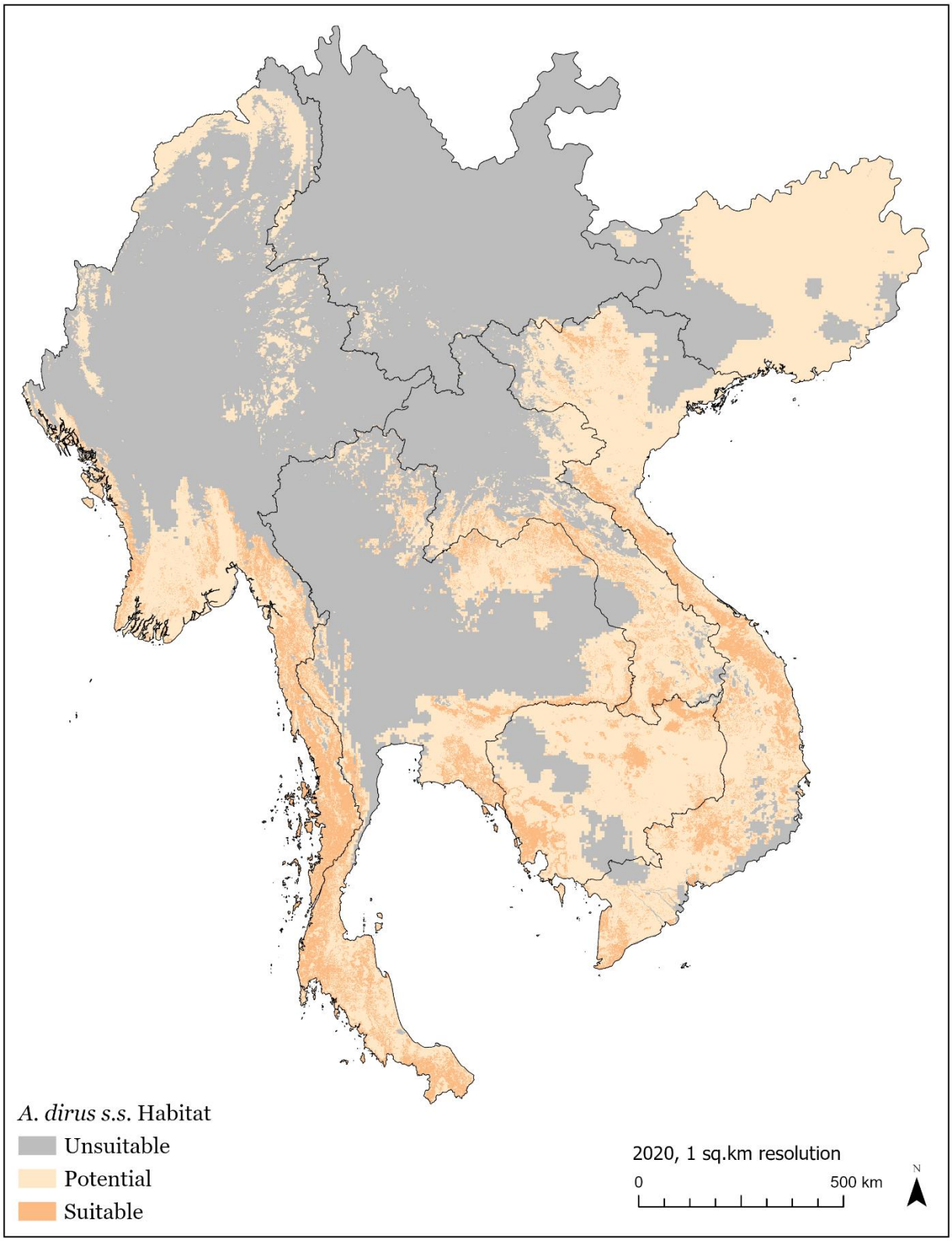


Plate 5. Habitat map for *A. dirus s.s.* at 1 km^2 resolution when the annual mean temperature applied is changed to ±2.5°C the identified range of 24 – 27.5°C.

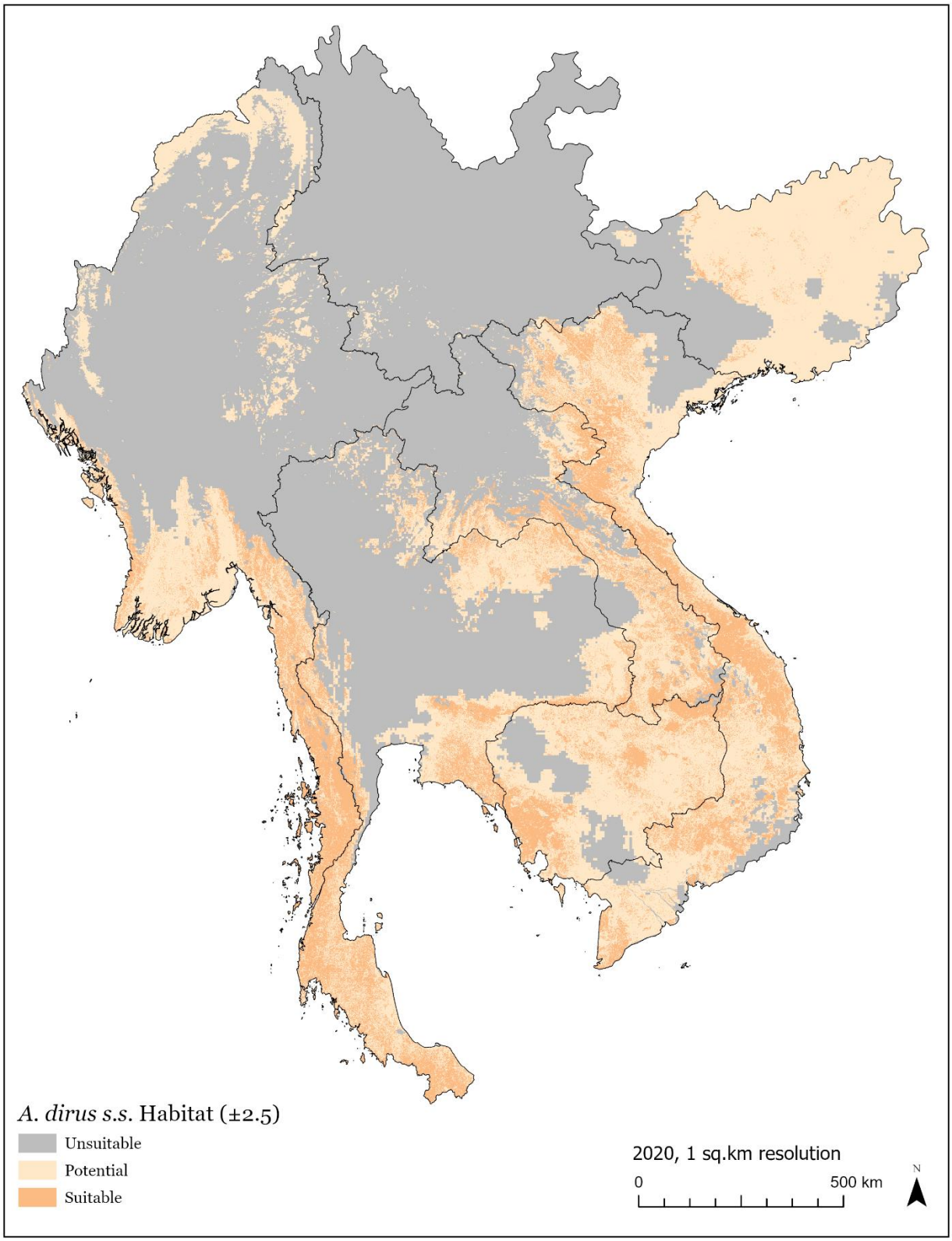


Plate 6. Raster algebra difference highlighting the increase in *A. dirus s.s.* habitat when the standard deviation is applied to the annual mean temperature range.