Distribution Agreement

Signature:

_______________________ ___________

Charles Harless Date

Using Wearable GPS Technology to Quantify Occupation-Related Human Movement in a Remote Riverine Region of Hyperendemic Malaria Transmission

By

Charles Harless
Master of Public Health

Global Health

____

Dr. Juan Leon
Committee Chair

____

Dr. Julie Clennon
Committee Member

Using Wearable GPS Technology to Quantify Occupation-Related Human Movement in a Remote Riverine Region of Hyperendemic Malaria Transmission

By

Charles Harless

BS, BSW
North Carolina State University
2015

Thesis Committee Chair: Juan Leon, PhD, MPH

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in Global Health
2017

ABSTRACT

Using Wearable GPS Technology to Quantify Occupation-Related Human Movement in a Remote Riverine Region of Hyperendemic Malaria Transmission

By Charles Harless

Background: Though most of Amazonian Peru remains seasonally endemic for malaria, a number of high-risk remote hotspots have been identified in the primarily riverine region. It has been hypothesized that asymptomatic riverboat workers from these hotspots may be responsible for the occurrence of seasonal transmission and regional reintroduction.
Goal: Therefore, the primary goal of this study was to characterize the normal regional movement of residents of a remote rural hotspot village in order to assess the plausibility of this hypothesis. **Methods**: Wearable GPS technology was used in this setting to accurately record this regional and fine-scale movement. Local spatial statistics were then used to determine areas of significant clustering; results were further analyzed using time and profession as differentiating variables. **Results**: Local spatial statistics revealed a highly heterogeneous spatial distribution around neighboring remote villages, within the regional urban centers of Nauta and Iquitos, and along the riverine networks which connect these features. Diurnal and nocturnal sub-analyses revealed further heterogeneity within the data, which interpreted with knowledge of vector behavior, assess differences in potential for disease transmission as mediated by either *Anopheles* or *Aedes* mosquitos.
Conclusion: Given that Santa Emilia is an identified malaria hotspot in remote Amazonian Peru, the results of this human movement study suggest that local riverboat transporters and remote-based government school teachers are most likely involved in the ongoing regional transmission and reintroduction of malaria which is annually reported throughout the region.
Implications: These findings suggest that less regional transmission and reintroduction of malaria might occur, on an annual basis in the Peruvian Amazon, if riverboat transporters and remote-based government teachers throughout the Loreto Department were more actively screened and treated for malaria.

Using Wearable GPS Technology to Quantify Occupation-Related Human Movement in a Remote Riverine Region of Hyperendemic Malaria Transmission

By

Charles Harless

BS, BSW
North Carolina State University
2015

Thesis Committee Chair: Juan Leon, PhD, MPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in Global Health
2017

ACKNOWLEDGMENTS

The author would like to thank the Universidad Peruana Cayetano Heredia and their Unit for Malaria and Leishmaniasis for facilitating this process. Thanks to Dr. Alejandro Llanos Cuentas, Dra. Dionicia Gamboa Vilela, Dr. Mitchel Guzmán, and Gabriel Carrasco Escobar for the professional mentorship and financial sponsorship, without which this study could not have occurred. Special thanks to Daniel Wong for the assistance, patience, and friendship during our 2 months of field work in Santa Emilia. Finally, enumerable thanks to the people of Santa Emilia for opening their community to a true outsider – the conversations, food, and adventure made data collection a truly memorable experience.

The author would also like to thank Emory University, the Rollins School of Public Health, the Hubert Department of Global Health, and the Global Field Experience committee for sponsoring this research. Thanks goes Dr. Kenneth Castro for being an outstanding personal advisor and mentor, and for specifically assisting in the creation of this practicum/thesis opportunity. Thanks goes to Dr. Juan Leon and Dr. Julie Clennon for committing to be academic advisors on this project. Your all's patience and willingness to teach, guide, and assist in this lengthy and laborious task was enormously appreciated – I simply could not have finished this project without you both.

Finally, the author would like to thank his parents, Dr. Matt Harless and Jeanne Alexander, and younger siblings, Casey and Layla, for their eternal patience, understanding, and support – without you all I simply would not be where I am today. Special final thanks must go to Neha Acharya for being the best friend and partner a person could ask for. Thank you for having always helped me find a healthy balance between thesis and life – I am so proud of us for pulling this all off together.

TABLE OF CONTENTS

CHAPTER 3 – Results

CHAPTER 4 – Discussion

CHAPTER 5 – Recommendations and Implications

CHAPTER 1
Literature Review

Malaria & Public Health

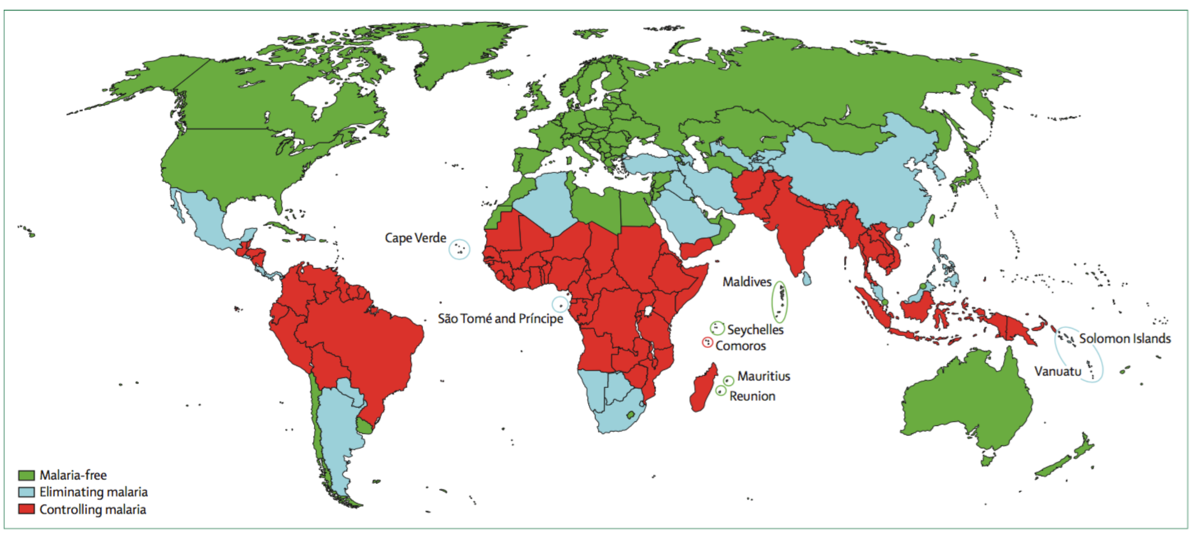
1 – Disease and Burden:

Despite recent successes, globally, malaria still presents a heavy global burden of disease – one which is not evenly distributed. The global burden of malaria has been steadily dropping over the last few decades; an estimated 212 million cases occurred in 2015, a 14% decrease in incidence rates since 2010 and a 22% drop since the year 2000[1]. However, it is still estimated that 429,000 deaths occurred in 2015 due to malaria – the majority of whom were children under the age of 5. Today's malaria prevalence is not equally distributed throughout the world. Around 90% of cases continue to occur in sub-Saharan Africa, followed by 7% in Southeast Asia, and 2% in the Eastern Mediterranean region.

Likewise, the 5 types of malaria are not evenly distributed globally and do not receive equal attention by infectious disease experts. Currently 96% of global malaria cases are *P. falciparum*, with only around 4% being from *P. vivax*[1]. When excluding the Africa region however, *P. vivax* accounts for over 40% of all malaria cases. Global trends can mask the immense burden felt by local regions. In endemic countries in South America, *P. vivax* makes up the majority of the burden, and presents unique issues in control and elimination efforts. In this context, the Co-Editor-in-Chief of *PLOS Neglected Tropical Diseases* journal recently penned an editorial expressing his professional belief that *P. vivax* in South America is very much neglected by the global public health community in such a way that it ought to be considered a part of the popular 'Neglected Tropical Diseases (NTDs)' grouping[2].

Countries with low incidence of malaria are often presented with unique challenges as they work to eliminate the disease. As global control efforts have increased their efficacy, an ever-growing number of countries are transitioning their national malaria program agendas from *control*[a] to *elimination*[b] efforts. This global distribution can be seen in the figure below by Feachem et al. In low incidence countries, the spatial distribution of malaria often appears patchy, with local "hotspots" providing a continuous source of new infections[3]. In these situations, a greater understanding of temporal spatial clustering is imperative to the continued success of malaria control programs, and the overall national success towards elimination. This study attempts to work towards this end.

Figure 1 – Categorization of countries by phase of malaria control efforts[4]



[a] ***Control*** – reduction in disease incidence, prevalence, and morbidity or mortality to what is deemed an acceptable level, as the result of deliberate country effort[4]
[b] ***Elimination*** – reduction to zero incidence of disease causing microorganism in a limited geographic region, where continued efforts are required[4]

Vectors: *Anopheles* mosquito versus *Aedes* mosquito:

Malaria is not the only infectious disease mediated by a mosquito vector. In an increasingly globalized world, the number of emerging infectious diseases is on the rise. Therefore, it is important to understand that different genius, and even species, of mosquitos are competent vectors for different diseases. The parasite which causes malaria is vectorially transmitted by the *Anopheles* genus of mosquito, while the *Aedes* genus is responsible for arbovirus diseases which also present a significant global health burden[5]. Within South America these arboviruses commonly include Yellow Fever virus, Dengue virus, Chikungunya, and most recently Zika virus[2].

Intuitively, different species of mosquitos express different behaviors. In the control of vector mediated diseases, it is especially important to note that different mosquito species exhibit different preferences for breeding locations and feeding behaviors[1]. Understanding these differences is vital to the effective control of these different diseases. For example, while the *Anopheles* mosquitos largely prefers feeding at night, the *Aedes* mosquito is preferential to day feeding[6].

2 – Transmission:

In the epidemiology of infectious diseases, the average movement patterns of the vector(s), host(s), and reservoir(s) are key to understanding the spread and persistence of disease causing pathogens across broad geographical areas[7].

Human Movement versus Vector Movement:

In comparison to humans, mosquitos have a limited range of travel. The *Anopheles* mosquito completes its entire life cycle within only a couple of hundred meters of where it was born[8]. Therefore, when patterns of incidence appear in patches, it must be human movement that is responsible for the long-distance transmission of the disease.

Several previous studies have designed models to evaluate the importance of exposure variance caused by human movement and how it related to pathogen transmission. Stripped mobile phone data was used in Kenya to compare human movement patterns to monthly regional prevalence information[9]. With this information, a number of regional *source settlements* [c] and *sink settlements*[d] for malaria were identified. Though successful, this method would not be feasible in those remote locations, such as much of the Amazon basin, that do not have access to mobile phone services. Stoddard et al.[8], working in Peru, developed a *conceptual* model which stressed the critical nature of studying individual human movement as an underlying factor influencing the risk of exposure to vector-borne illnesses and subsequent disease transmission. They concluded that understanding such movements would help to identify the key sites and individuals to be focused on for future surveillance, treatment intervention, and prevention control programs.

Though conceptual models can be useful, studies using fine-scale movement have illuminated additional challenges. Vazquez-Prokopec et al. used GPS data loggers to assess the impact of fine-scale human movement on mosquito-mediated infectious diseases[10]. The

[c] *Source* settlements are those inhabited areas endemic with malaria that emitted new cases[9]
[d] *Sink* settlements are those areas that receive persons[9]

authors found that less than 40% of tracked participants displayed regularity or predictability in their daily routine. This is in sharp contrast to the high predictability of individual daily routines in higher-income countries. Such high percentages of temporally unstructured daily routines have real consequences on the transmission of infectious diseases. The author's model predicted a 20% larger epidemic than would be expected in a setting with higher percentages of structured individual routines.

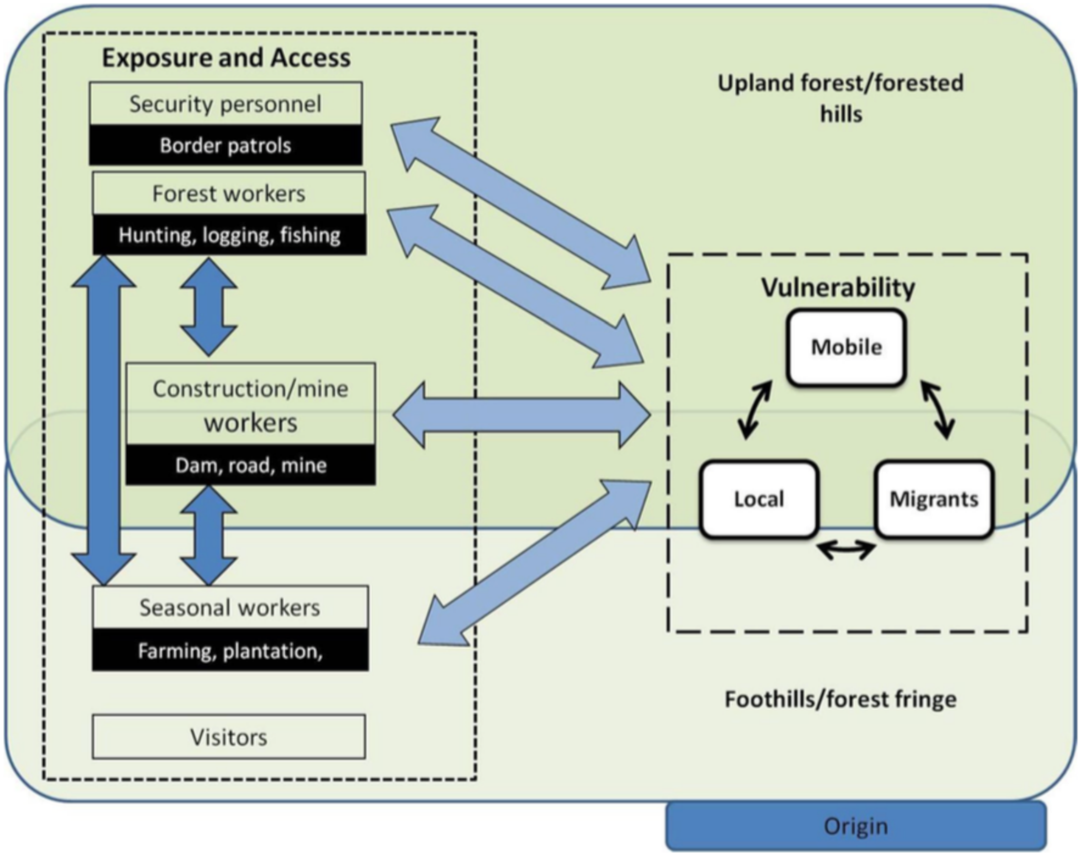
3 – Control Efforts:

In low resource settings, and especially in low incidence countries, the preponderance of public control efforts are focused on vector control strategies[1]. However, as posited above, the importance of individual human movement, as it relates to the transmission dynamics of pathogens, should not be overlooked in studies pertaining to the control of vector-mediated infectious diseases.

<u>Issues of Spread and Reintroduction</u>:

Human population movement (HPM) has been found to be important to the control of malaria. In a literature review conducted by Guyant et al.[11] concerning malaria in the greater Mekong region of Southeast Asia, it was found that HPM corresponds to the long distant transmission of malaria, which can lead to the novel introduction or reintroduction of disease to a region. The primary demographic of this effect was found to be the mobile and migrant population (MMP), composed of those who worked and traveled in seasonal patterns throughout the Mekong region. From their review of the literature, the authors designed a population movement framework (PMF) to visualize these findings.

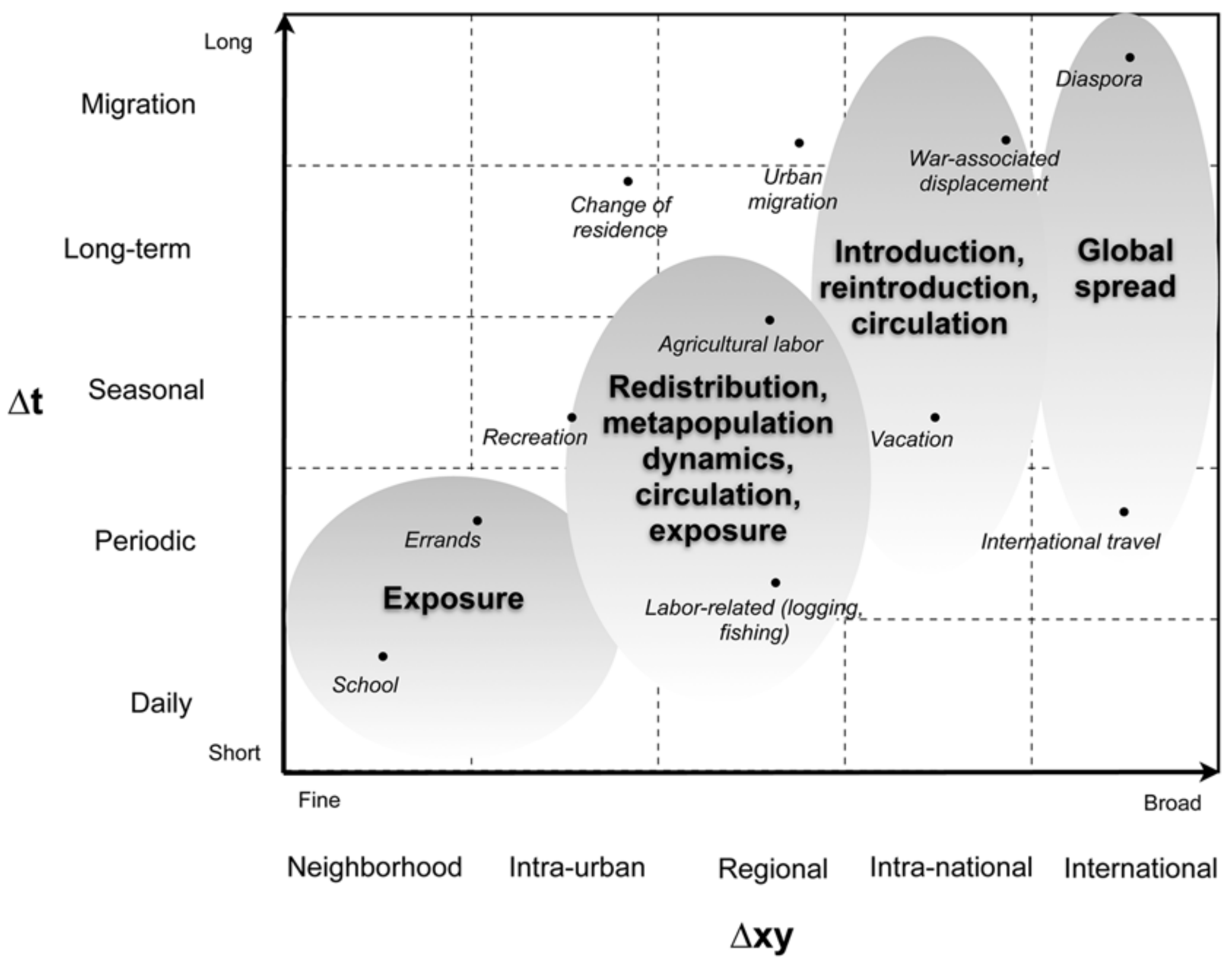
Figure 2 – Population Movement Framework (PMF)[11]



Use of population movement frameworks can help identify specific targets for more cost-effective control interventions**.** Census and/or surveillance data have been put together with spatially referenced malaria incidence data to highlight the differences in mobility of different demographic groups in three separate highly-riverine settings – in an East African setting by Pindolia[12]; in Rondonia, Brazil by Kasuragawa et al.[13]; and in the Ratanakiri province of Cambodia by Sluydts et al.[3]. In all cases, these demographically-stratified movement estimates provided evidence to inform targeted treatment options. These interventions led to more efficient uses of the limited local resources than was previously possible without the knowledge of how human population movement effects the regional transmission dynamics of malaria.

Visually summarizing the role of human movement in the transmission and reintroduction of malaria, Stoddard et al.[8] presented the framework below. The interaction of distance traveled and time spent (frequency and duration) are categorically visualized. Notice that movements of greater distance normally correspond to greater time spent, though this is not always the case with the increasing ease of international travel. The seasonal redistribution and reintroduction of malaria occurring in Amazonian Peru, discussed in more detail later, fall in the middle space of this framework.

Figure 3 – Framework for human movement and vector-borne pathogen transmission[8]



Use of GPS and GIS Technologies:

Accurately quantifying movement has traditionally presented many logistical challenges, but improving modern technology presents a novel solution to these issues. Historically, the tracking of individual movement has relied heavily on that individual's recall of their own movement[14]. This has always been, and inherently will always be, imperfect because of bias. The limitations of the human memory, interviewer error, changes in behavior, and associated privacy are all issues. Likewise, the alternative of directly observing movement is infeasible, especially over longer periods of time. Thankfully ever improving technology now presents an alternative. Wearable GPS devices can provide all essential information to quantify human movement as it applies to infectious disease (i.e. – positions of visited locations, duration of visit, and routes followed to access said locations)[15]. Simply put, GPS technology increasingly can provide precise readings at cheap prices.

GPS devices, though a relatively new technology based solution, prove especially useful in low resource settings. After using wearable GPS data loggers to characterize the fine-scale movement of people within the tropical urban center of Iquitos, Peru, Vazquez-Prokopec et al.[10] concluded that the usefulness of such data is of special relevance to low and middle income countries, as they compose most the globe's urban populations and are disproportionally burdened by most infectious diseases. Similarly, in 2010, Khan et al.[16] reviewed all tropical medicine and public health related examples of GIS use in the literature over the last 2 decades (1990-2010). They similarly concluded that "it is probably among the world's poorest that the vast potential of GIS has most to be realized. Beyond the mere description of a problem, such geographical tools could be used to track and respond and to improve, sustain and scale-up interventions"[16].

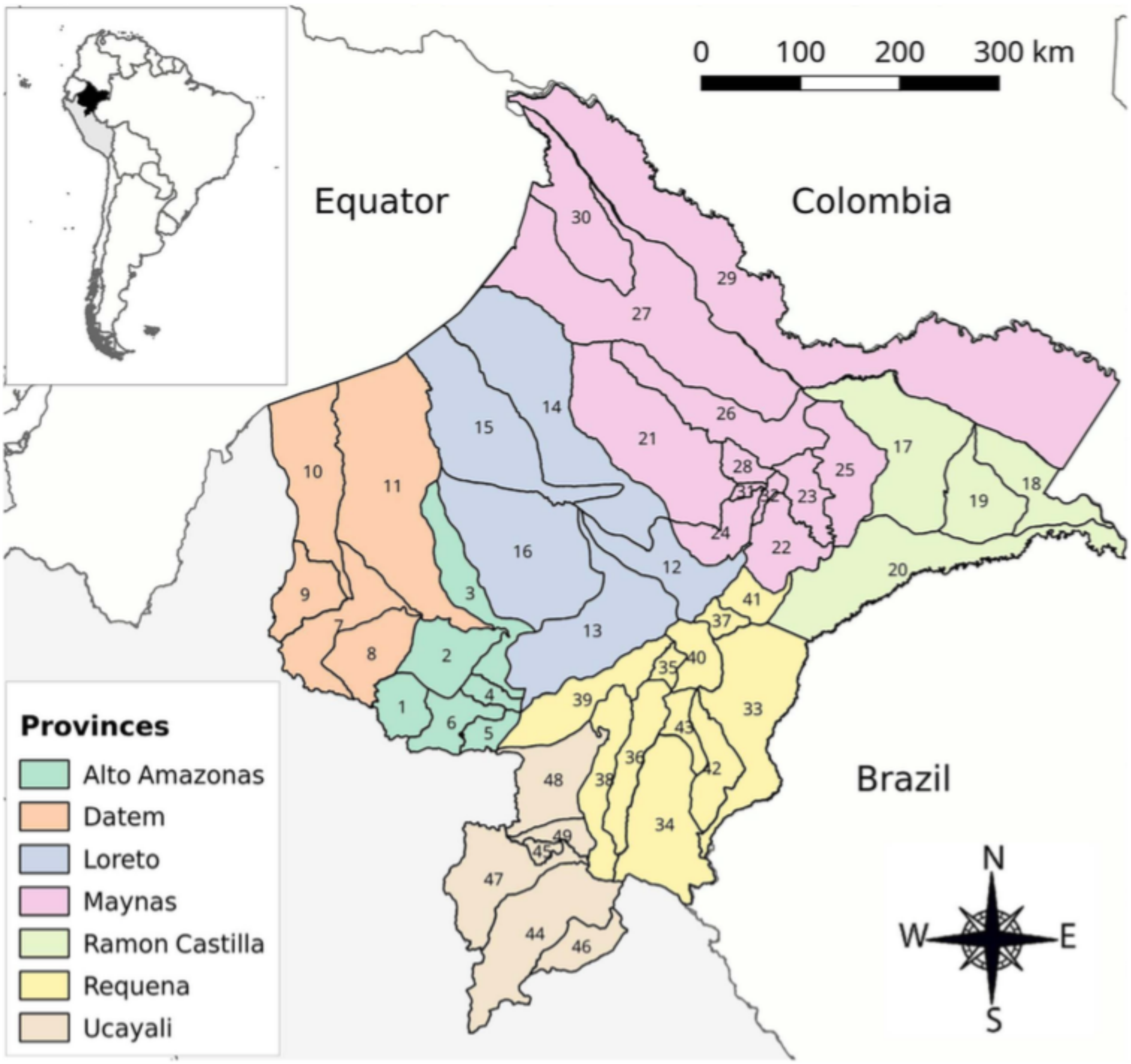
Some pros and cons of using GPS data loggers in vector-borne disease research are summarized in **Table 1**. The second and third points in the 'Pro' column enabled the field methods conducted for this study. As discussed further in the strengths and limitations section of the conclusions chapter, point three in the 'cons' column below proved to be a methodological limitation of this study.

Table 1 – Aspects of GPS technology use in vector-borne disease research[8]

Pros:	Cons:
Only require receivers (satellites are freely available)	Data sets are large & require intense computing
Work everywhere (unlike cell phones)	Commercial options are not tailored for research
Provide precise spatial information	Short battery life
Devices are becoming smaller & less expensive	Custom devices are still expensive

5 – Amazon Region:

Figure 4 – Political boundaries of the Loreto Department, Peru[17]



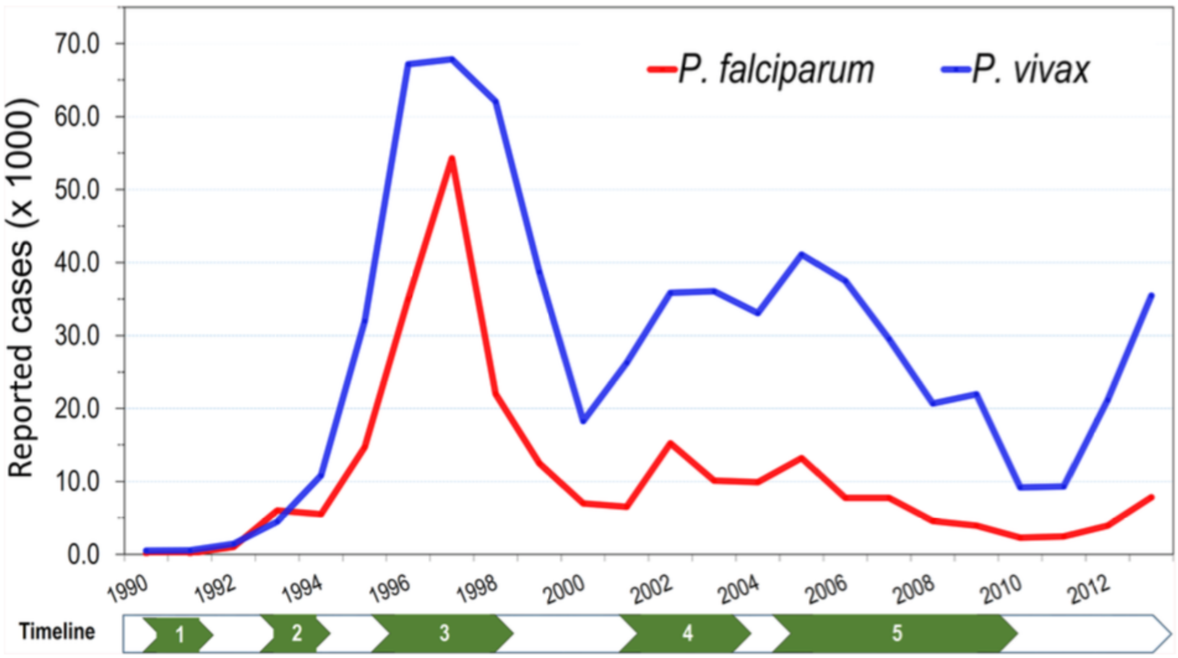
The Loreto Department of Peru, shown above, comprises the majority of the Amazon rainforest which falls within Peru's borders[17]. The region is divided into 7 provinces, which are further divided into districts. With the exception of the peri-urban districts surrounding the city of Iquitos (districts 23, 24, 28, 31, and 32), the region of Loreto is extremely rural and vastly lacking in all forms of infrastructure. Iquitos is the largest population center in the Peruvian Amazon, containing approximately 380,000 inhabitants[18]. In the last 30 years, as the commerce of natural-resources has expanded, transportation pathways including paved and dirt roads have also increased. To date, one substantial highway has been constructed to connect Iquitos with the smaller city of Nauta (population: ~17,000). This highway is paved and measures 95km. Nonetheless, river networks remain the predominant transportation pathway for the rest of the Loreto region[19].

Those communities which lie beyond Nauta are described as 'remote'[18]. They have populations ranging from 100 to 6,000 inhabitants, are geographically isolated from larger cities (i.e. Iquitos), and significantly limited access to communication networks such as cellular tower signals. The people of these communities survive on sustenance farming, hunting, and fishing. The water supply varies in these villages from non-potable piped systems, rainwater collection, ground wells, and a continued heavy reliance on water straight from the region's rivers. The containers used for these water collection methods are often lacking lids or have fallen into disrepair, leading to unintentional rainwater accumulation – ideal reproduction sites for vectors such as mosquitos, especially those of the *Aedes* genus.

Regional weather patterns largely drive the seasonal-endemicity of malaria and other vector-mediated infectious diseases. River levels fluctuate up to 10 meters, as a result of Andean snowmelt and a rainy season[20]. During the rainy season, 400 mm of rainfall a month are average. The population density of vectors mirror these changes. Of recent discovery, *Anopheles darlingi* subtypes have emerged, specially adapted for both seasons, and are likely helping to promote the recently observed year-long regional endemicity.

Regional Vectors:

Figure 5 – Annual malaria incidence, Loreto department: 1990-2013[21]



Anopheles darlingi is the primary vector of malaria in the Loreto District today, and Lainhart et al. have previously summarized its history in the region[21]. Thriving in riverine habitats, it has achieved uninterrupted population spread from Argentina to Mexico. *An. darlingi* were first observed in Peru (Loreto District) in 1933, but it composed only a small percentage of the *Anopheles* species found in the region. After an intense government run insecticide program starting in 1957, malaria was nearly eliminated from the entirety of Peru (95,000 cases nationally in 1944 fell to only 1,500 cases in 1965). By 1988 no malaria cases were detected in Iquitos, and for the next half decade *Anopheles* mosquitos were not found in the urban center. However, in 1996 *An. darlingi* were rediscovered, which corresponded to a district-wide epidemic of malaria throughout Loreto. By the early 1990's *An. darlingi* had displaced all other *Anopheles* species in the greater Amazon Basin. Malaria cases spiked to

150,000 in 1997, after which government programs fought to regain control over the once eliminated disease using a wider array of vector control methods. By 2000 regional cases had decreased back to ~22,000. Unfortunately, in the past decade a resurgence has been detected, with over 60,000 cases being detected regionally in 2014[19]. The reason for this current uptick remained unclear and had not been extensively studied.

In 2015, a population replacement study was conducted to answer these questions[21]. The study aimed to measure the natural, or unintended, population replacement that had occurred in the Peruvian Amazon within this last decade. Human biting rates (HBR) and entomological inoculation rates (EIRs) for *An. darlingi* throughout the Amazon river basin were found to vary greatly. In peri-urban and deforested zones a HBR of 6.5 bites/hour was measured, while a HBR of 41.13 bites/hour and an EIR of 5.3 infective bites/night was calculated in smaller riverside villages situated amongst pristine jungle. Study results showed that these rates were directly associated with the availability of biomass (population size), and not forest cover (pristine vs deforested). It was ultimately concluded that a population replacement event had occurred at some point between 2006 and 2014, which coincided with both the regional arrival of long lasting insecticide treated bed nets (LLIN) and an El Niño. Though successful in answering the main research question, this study was unable to determine what had caused or enabled the rapid and widespread nature of this total population replacement event.

6 – Regional Disease & Burden:

Though largely centralized around the city of Iquitos, regional commerce necessitates continued regular interaction with remote regions of the department. The Loreto Department accounts for 25% of the geographical size of Peru, but only 3% of the country's total population[22]. Despite this, the department represents 80% of all malaria cases reported in Peru today. Of all those who live in the department, only 35% live in rural communities outside of the sparse urban centers. This percentage is largely misleading however, as the main industries (agriculture, fishing, logging, and petroleum) necessitate that laborers work in rural villages and areas outside of the urban centers where they and their families permanently reside.

Malaria "hotspots" have been identified throughout the region, which likely play a key role in regional transmission dynamics and present unique surveillance needs[22]. Bautista et al.[23] conducted a retrospective surveillance study and found that constant high-risk hotspots were present in the Loreto region during periods of both localized high and low malaria incidence (i.e. – rainy and dry seasons). These findings, along with others[24], suggest that even modest control efforts targeted at these hotspots would highly impact overall regional transmission patterns. However, it has been shown that active case detection (ACD) is necessary to accurately facilitate such efforts[25]. Additionally, it has been posited[26] that PCR methods, rather than traditional blood-smear microscopy, will be essential to accurately diagnosing symptomatic and asymptomatic disease in Loreto since transmission endemicity is low and seasonal overall. Despite all of this, passive case detection (PCD) and blood-smear microscopy continue to be the current government facilitated malaria protocol throughout much of Loreto[19].

7 – River-Boat Workers & Riverine Communities:

Riverways compose the sole transportation network for most of Amazonian Peru. A variety of vehicles transport cargo and passengers, including[18]:

- ‘Lanchas’ – large barges for cargo and passengers
- ‘Lanchitas’ – medium-sized barges
- ‘Rápidos’ – speed boats, only for passengers, uncommon because of price
- ‘Peque-Peques’ – canoes with outboard motors, most common in remote areas

Remote riverine communities have been shown to be of central importance to the long-distance transmission and reintroduction of malaria in the Loreto region, despite their size and remoteness. Parker et al.[22] hypothesized that most malaria cases reported in the peri-Iquitos zone were contracted from remote riverine hotspot communities and brought back via traveling seasonal workers (primarily loggers and fisherman). Along with general lack of public health resources, vector capacity for malaria was found to be significantly higher in remote riverine communities than in similar communities closer to Iquitos. Surprisingly, these rates surpassed anything described in South America before, and held pace with holo-endemic areas of Sub-Saharan Africa. “These results further implicate human travelers as a potentially major and highly-mobile disease reservoir in the region, identifying a local target population involved in sporozoite movement and suggesting important considerations for malaria control efforts in populated communities to which malaria-carrying travelers return”[22]. There are however some comparability issues to be noted in interpreting this study’s finding in the context of this study involving Santa Emilia. First, while all 21 sites were rural and purely riverine in transportation networks, all were much closer to the urban zone of Iquitos than the field site for this study. Additionally, all 21 sites were worker campsites; Santa Emilia in contrast is a remote village that accepts seasonal workers but also has a permanent population year-round[27].

Biological Plausibility of Vector Migration:

In the case of malaria and other vector-borne diseases, discussions of rapid or heterogeneous transmission of disease across long distances is always focused on individual human movement[11]. This is intuitive as most mosquito vectors do not move more than 500 meters during their lifetime. Therefore, it is normally concluded that the rapid spread of vector-mediated diseases, across great distances, is facilitated by diseased humans traveling to new localities and subsequently inoculating the local vector population[22]. With this, it has largely been assumed that vector expansion occurs slowly as populations naturally grow[21].

However, in novel 2014 findings, Guagliardo et al.[18] characterized the rapid spread of mosquito vectors along the riverine networks of the Amazon basin. Their study revealed linear patterns of *Ae. aegypti* spread along the Iquitos-Nauta highway, and scattered spread between riverine communities. They thus concluded that in the Peruvian Amazon, the geographic spread of the *Ae. aegypti* vector must be driven by regional human transportation networks. They hypothesized that boat traffic is likely the driver of this long-distance dispersal via unintentional transport on these boats. This hypothesis led to a collections-based study[28] where common vehicle types were extensively studied for their role in unintentional vector spread. Their results were profound. Nine species of local mosquitos were found to be completing their reproductive life cycle in the water hidden beneath the floor boards of the transport vehicles[29]. The culmination of their results showed that 1) large and medium transport boats (for both humans and cargo) were heavily infested with mosquitos, 2) the vast majority of villages in the district of Loreto are only connected by rivers, and 3) the observed spatial pattern of *Ae. aegypti* expansion mirrors previously characterized riverine-mode spread. They therefore concluded that river boats were responsible for the significant regional spread of *Ae. aegypti* through the Peruvian Amazon, and that aquatic transit throughout the Peruvian Amazon is responsible for the long-distance transmission of both infected humans and the necessary vectors[18].

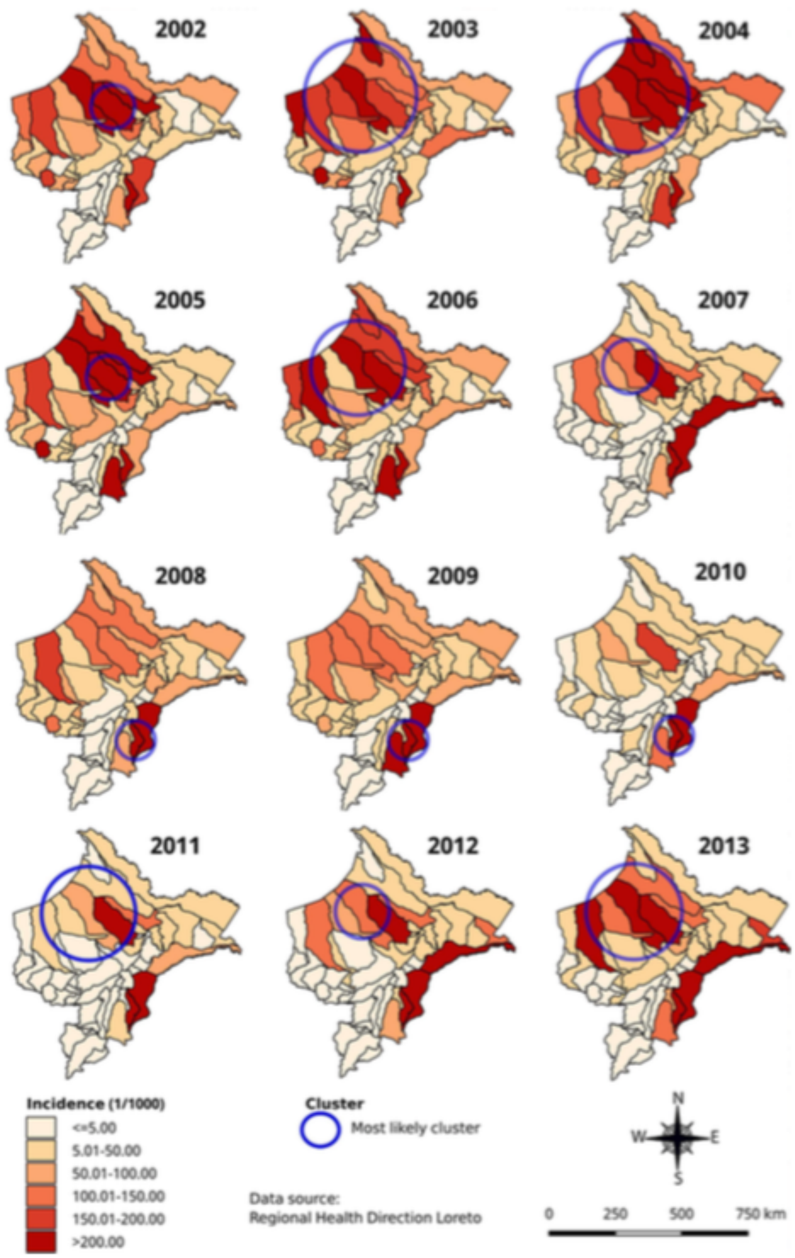
These findings are of interest to the scope of this study, despite being focused on the vector *Aedes aegypti*, because they establish the plausibility that traditionally urban mosquitos could become established in a remote setting such as Santa Emilia. With this potential population expansion, additional public health concern for novel arbovirus burden should not be ignored when studying regional human movement, as this study does.

8 – Regional uses of GPS and GIS:

Despite the call for increased use of GPS and GIS technologies in infectious disease transmission studies[10,15], and the general stance that this increased spatial understanding is critical for successful future control and elimination efforts[14,26], regional use of these technologies remains surprisingly low.

Focused on the changing regional distribution of malaria, Soto-Calle et al.[17] used GIS to visualize a retrospective analysis on annual malaria incidence. They used 12 years of publicly available data and performed a variety of temporal, spatial, and space-time analyses. Their primary results are presented below in **Figure 6**.

Figure 6 – Annual district malaria incidence in the Loreto Region: 2002-2013[17]



<u>Arbovirus and Iquitos-centric Uses</u>:

Most of the studies that have used GPS or GIS technology in the Loreto District of Peru have largely been used either 1) in the urban center of Iquitos, and/or 2) to study arboviruses and not malaria. Studies were performed in Iquitos in 2010[14] to assess the acceptability and barriers to utilizing wearable GPS devices. Results included concerns over financial responsibility for the units, whether the units were video/audio recording them, confidentiality of their movements, and health side effects. Following this acceptability study[15], strengths and weaknesses of using GPS loggers to study fine-scale human movement were conducted in the context of Dengue transmission within the city. Weaknesses of using the units in research included that 1) units could be turned off (both on purpose and by accident), 2) units were easy to forget at home, and 3) units were easy to forget to charge (resulting in them often being left at home on purpose). Another study[10], looked at the feasibility of using GPS units in the Iquitos population (again characterizing fine-scale human movement associated with Dengue virus), and concluded that 15 days of tracking by a wearable GPS unit was sufficient to fully characterize short-term spatial routines. To date, GIS has been used to model and simulate potential Dengue outbreaks in Iquitos[7], and assess the temporal stability of *Ae. aegypti* fine scale clusters within the different neighborhoods of Iquitos[30].

In conclusion, to date there are no examples in the literature of GPS or GIS technologies being used to accurately characterize the intersection of human movement and malaria hotspots in the Loreto District of Amazonian Peru. The studies discussed above, however, lay the ground work for such studies to be done more efficiently.

Conclusion

Need:

Patterns of human movement for the river-boat workers in the Peruvian deep-Amazon have never been described in the context of transmission, reintroduction, and control efforts of vector-born infectious diseases such as Malaria.

Goal:

Characterize the normal regional movement of Santa Emilia residents in relationship to the ongoing transmission and reintroduction of Malaria throughout this deep-Amazon region.

Aims:

1) Describe spatiotemporal patterns in collected GPS data with GIS analysis software
2) Interpret patterns of potential transmission by *Anopheles* mosquitos and *Aedes* mosquitos
3) Characterize the variance of professions amongst Santa Emilia residents, regarding their movement patterns and potential disease outcomes

Significance:

This research will fill a previous knowledge gap as regional migration of this population has never been quantitatively described. Further, these novel data will inform a more accurate understanding of the transmission dynamics of malaria within this unique region. The characterization of these migration patterns may lead to a better understanding of the role played by hotspot villages, and may assist in the control of malaria and other infectious diseases. It is hypothesized that hotspots play a key role in the ongoing regional transmission of malaria, and that river-boat workers play a key role in the reintroduction of malaria to peri-urban villages.

CHAPTER 2
Methods

---

1 – Introduction:

Santa Emilia has been classified as a hotspot for regional malaria by both the Universidad Peruana Cayetano Heredia (UPCH) and the regional health authorities. Unlike the rest of the region, seroprevalence levels of *Plasmodium vivax* and *Plasmodium falciparum* were initially found to nearly equal. Equally unusual for the region, cases of mixed malaria (individuals positive for both *Pl. falciparum* and *Pl. vivax*) have been encountered in Santa Emilia (UPCH, unpublished data).

UPCH is currently conducting phase 2 of a study entitled "Transmission of Malaria in Remote Communities of the Amazon Region of Peru", in which they are evaluating the characteristics and epidemiology of transmission longitudinally in remote Santa Emilia, compared to two other rural riverine villages which are within the peri-urban zone of Iquitos. UPCH hypothesizes that this comparison will result in important findings of transmission, maintenance, and reintroduction of malaria between remote and peri-urban villages. This ongoing study is complementary to a project entitled, "The Impact of Asymptomatic Riverboat Workers on the Epidemiology and Control of Malaria throughout the Peruvian Amazon." This project is being done in partnership between UPCH and the Amazon Center for Malaria Research (ICEMR Amazónico), which has received 5 years of funding from both the Peruvian government and the NIH of the United States. This study has been written as a sub-study within both of the aforementioned ongoing UPCH studies.

Therefore, though Santa Emilia is an identified active hotspot for malaria and the importance of accurately understanding the regional movement of individuals who work and/or live in such hotspots has been adequately described, the normal local and regional movement of Santa Emilia inhabitants has yet to be characterized. Thus, it is the goal of this study to accurately quantify the normal movement patterns of Santa Emilia residents, using wearable GPS technology, in order to fill this important knowledge gap.

2 – Population and Sample:

Setting:

Santa Emilia Nahuapa (74.209°W, 4.195°S, 202m above sea level) is a small, rural, riverine village in a part of the Loreto region of Peru classified as 'Remote Amazon'[22]. The trip from Iquitos, the region's capital and closest airport, to Santa Emilia takes 12-16 hours and normally is completed over 2-3 days. Utilizing local transport methods, a public mini-bus is first taken from Iquitos to Nauta via the Iquitos-Nauta highway (2 hours travel time). From Nauta, the last major urban-center in the region, the remainder of the journey is accomplished in a small outboard motorized canoe (Locally: peque-peque). From Nauta it is upstream the Marañon river towards its juncture with the Tigre river. At this juncture exists the town of MiraFlores, which contains the closest health outpost to Santa Emilia (6hour travel time). From MiraFlores, it is up the Tigre river, and then up the Nahuapa quebrada (English: mountain stream). Santa Emilia is the fifth of six villages on the Nahuapa quebrada (counting from mouth to origin). Total river-based travel time between Santa Emilia and Nauta depends on the direction of travel (i.e. – upstream or downstream), but via the average peque-peque the journey takes 8-12 hours. This journey is visualized in **Figure 2** of the Results chapter.

The population of Santa Emilia varies throughout the year, as individuals travel for seasonal work, but average inhabitance is between 300-400 persons[31]. This number can be further numerated into family clusters, which total between 12-15. Santa Emilia has no electricity or running water, though the government installed a borehole-well around a decade ago. Gasoline powered generators are becoming increasingly common, able to power the few lights, refrigerators, and stereo systems in the village – though gasoline is rare and mostly saved for special occasions. In 2012, the village built a hut for the express purpose of malaria surveillance studies, which contains a light microscope, materials for blood smears, and a cooler for polymerase chain reaction (PCR) samples[31]. A UPCH employed field-worker/microscopist had been living and working in this research hut continuously for the 18 months prior to the start of this study.

Malaria is hypo-endemic in Santa Emilia, though it still exhibits some seasonality. However, the seasonality of malaria in Santa Emilia has been characterized as relating to river height fluctuation, rather than rainfall fluctuations[27].

Population:

This study focused on characterizing the normal movement of those Santa Emilia residents who frequently utilize riverine networks to travel throughout the region. To this end, men who regularly use boats to work were the focus population. However, in Santa Emilia, as is common in such remote-rural settings, professions are not easily definable. Most individuals spend some time as agriculturalists in private chacras (English: small land plots), fishing in the river and nearby shallow lake, and occasionally hunting wild game in the jungle. Individuals do however display preferences for these activities. For example, some individuals spend most of each day fishing, and only a sparse amount of time per week as agriculturalists. Likewise, though everyone hunts occasionally, only a few go hunting multiple times per week. Overall, fishing and hunting are always done for local consumption; agriculture in Santa Emilia, however, serves both as subsistence farming and cash-crop cultivation of cassava root (*Manihot esculenta*), plantain fruit (*Musa paradisiaca*), and aguaje palm fruit (*Mauritia flexuosa*).

There are 3 'professions' that are noteworthy, as they significantly change the routine and regional movement of these individuals. The first is loggers, who travel well beyond the scope of the village to find and harvest specific fine woods, which are then transported to Nauta in bulk to be sold in the lumber mill. These trips occur once per month on average, though this transport is greatly hindered by low river levels. While in Nauta, these individuals normally spend several days enjoying city life with their recently acquired earnings. Secondly, a half dozen people in the village act as 'transporters.' These individuals own the largest canoes and most powerful outboard motors, and use these vehicles to transport the agricultural goods of Santa Emilia agriculturalists to be sold at the main market in Nauta. These trips occur more frequently, often weekly or biweekly. Thirdly, there are three government employed teachers who work in Santa Emilia. These individuals are not from Santa Emilia, and therefore travel back for 5-7 days each month to spend time with their families in Iquitos. This information comes from the travel-logs previously collected in Santa Emilia, which are described further below. What is not known at this point, 1) which intermediate villages these traveling professions, and others, stop at for food and overnight rest, 2) at what frequency, and 3) for how long.

Rationale:

UPCH has been conducting malaria research in Santa Emilia since 2012. At baseline, 22% of those tested (151 individuals) were positive for malaria by blood smear microscopy. That number increased to 51% when those samples were assayed with PCR technology. In that first year of monthly testing (2013), asymptomatic malaria ranged from 22% to 40% by microscopy and 41% to 60% by PCR. Of those asymptomatic individuals, microscopy methods failed to detect 39% of *Pl. falciparum* cases identified by PCR, and 47% of those with identified with *Pl. vivax* infections via PCR methods. Additionally, these 2013 findings showed that positive rates for both species of plasmodium were lowest from September to December[27]. As one of many co-occurring malaria studies being conducted in Santa Emilia, UPCH had recently attempted to characterize the movement patterns of inhabitants with the use of weekly self-reported travel logs[31]. After one year of collecting these travel logs, a preliminary review of the data caused concern in that little fine-scale detail was recorded in the reports. Additionally, reports were filed at a concerningly low frequency. It was the opinion of the UPCH investigators that such methods were not going to be sufficient to adequately inform their research questions concerning the roles that asymptomatic river-boat workers play in the transmission and reintroduction of malaria throughout the Loreto region.

3 – Research Design:

To answer these questions, active field research was conducted. Participants were asked to keep with them, for all hours of the day and night, a wearable GPS data logger for a period of 15 days. Though their participation was active, the intent of the observation was passive, and so all participants were asked to not change their normal behavior patterns in any way. The only data collected for this study was that of the GPS data loggers (date, time, elevation, latitude and longitude); no identifiable information was collected at any point for this study. Participants' data was recorded with random research identification numbers previously established by UPCH. Only the UPCH malaria research team had access to the codex used to translate these identification numbers back to the residents of Santa Emilia.

4 – Instruments:

Igot-U GT120 GPS data loggers (Mobile Action Technology Inc.), were used to continuously track individual movement patterns. These units have a 4.4 meter point accuracy, a 10.3 meter line accuracy, hold a single charge for many days, and only require 2 hours to fully recharge. They have also recently been described as culturally appropriate, logistically feasible, and efficacious devices by Vazquez Prokopec et al. in a not dissimilar setting and population[32,14,15]. Importantly, these devices are small (44.5 x 28.5 x 13 mm), lightweight (20g), water resistant, reasonably inexpensive ($50 each), and nondescript (devices emit no sound or light). Additionally, the GT120 units are easily programmable. For use in this study, units were set to collect one GPS data point (date, time, elevation, latitude and longitude) every 5 minutes, for 24 hours a day. In total, 20 GPS units were used in for this study.

5 – Ethical Considerations:

This study received ethical approval by the Comité Institucional de Ética en Investigación of the Universidad Peruana Cayetano Heredia (UPCH), as a sub-study, on May 23rd 2016 [SIDISI: 64024]. Additionally, this study was approved by Emory University's Institutional Review Board, under the Social/Behavioral protocols subsection, and received expedited approval on June 9th 2016 [IRB: 00088823]. See **Appendix 1-2** for copies of these documents.

6 – Procedures:

Recruitment:

Recruitment for this study was conducted by the UPCH field worker, who had been in Santa Emilia already for 16 months, for logistic and culturally sensitive reasons. Recruitment was limited to those individuals who reside in Santa Emilia and where 18 years or older. Recruitment was not random, and preference was given to those individuals who utilize river boats in their normal work routines. This led to 94% of participants being male.

The recruitment process began with an informal conversation with the whole community concerning the arrival of the PI and the topic and need for the new sub-study. Later, potential participants were then approached individually, where an official consent document was read aloud. Potential participants were given time to ask any questions, and then asked to briefly summarize what had just been read to them to assess their understanding. The recruitment process concluded with a verbal agreement to participate.

In total, 32 Santa Emilia residents were recruited for this study over a period of 2 months. Participants were recruited in 3 rounds. Group 1 included 15 participants, group 2 included 7 participants, and group 3 concluded with 10 participants. While participation was capped at 15 individuals per group, because of limited number of GPS units, rounds 2 and 3 had less than 15 individuals as recruitment saturation was reached within the community.

Field Methods:

Participants were asked to wear a GPS unit on a lanyard around their neck and to proceed about their lives as normal – i.e. no behavior change was desired. Participants were told expressly to only remove their unit for bathing (as units are water resistant and not water proof), and while sleeping if they found that the unit bothered them. It was reiterated that the unit should be kept on the participant's person always, especially when outside of their home.

Being set to record a data point every 5 minutes for 24 hours a day, the battery life for the units was found to be 5 days on average. Therefore, every 3 days the units were scheduled to be swapped out. During this swap, batteries were recharged, data was backed-up, and internal memory subsequently cleared. These meetings also afforded the PI the opportunity to check that each unit was being used correctly (i.e. not being left at home for long periods of time, no premature loss of battery life, etc.). Additionally, during these meetings the primary investigator asked each participant to self-report their movements, activities, and any related reasoning for their routines. All meetings occurred in the homes of the participants, during the preferred time of the participants – normally in the late evenings before meal time. All of this information was recorded and organized in an Excel worksheet (Microsoft Corporation).

At the end of 15 days, devices were collected and participants were given a final opportunity to ask questions. This concluded the individual's participation in this study.

7 – Data Analysis:

Data was downloaded from GPS units using the Igot-U proprietary software, "@trip PC." Files were individually labeled with a codex to their corresponding participant ID. Each file was cleaned of superfluous data points (those collected while charging or between participant use), and then exported to .CSV files. Each of these coded .CSV files was then combined in Microsoft Excel to form a single uninterrupted file (consisting of 15 days of data) per participant. These participant files were then individually loaded into ArcMap version 10.4 (ESRI, Redlands, California, USA), and subsequently projected from latitude and longitude values to the Universal Transverse Mercator, zone 18 south, World Geographic System 1984. The projected participant shapefiles for each participant were then manually checked for extraneous data points (i.e. a location change of 30 km in a 5 minute interval, likely caused by insufficient satellite triangulation). These implausible data points, composing less than 1% of any individual's data log, were then manually erased. Having been cleaned, all participant shapefiles were then merged into a single point shapefile, maintaining participant ID as a variable column.

ESRI basemap satellite imagery (30cm resolution) was then used to digitize areas of interest. These features included villages (point), urban centers (polygons), waterways (lines), roadways (lines), and forest (polygons). The village category included the home village of Santa Emilia, the 5 additional villages on the quebrada Nahuapa, the clinic containing town of MiraFlores, and all additional intermediate villages on the Marañon river. The urban centers category contained Nauta and Iquitos, with Iquitos being broken into its 4 principle neighborhoods. Waterways included the quebrada Nahuapa, Santa Emilia lake, Tigris river, and Marañon river. The Iquitos-Nauta highway composed the only roadway within this study area. All additional area was classified as forest.

Digitized shapefiles were then put through a buffer analysis to account for slight inaccuracies in recorded point locations from the GT120 GPS data loggers. Village and City shapefiles were given a 500 meter buffer, while all water and roadway shapefiles were given a

50 meter dissolved buffer. Villages buffers were subtracted from continuous water and roadway buffers, and all of these were collectively subtracted from the study area to create the final forest buffer zone. At the end of this process, all buffer zones were mutually exclusive.

GPS data points were then allotted to professions, and further subdivided by time. Professions were grouped together intuitively to best view variance in movement patterns. Group 1 was those individuals who primarily spend their work time as agriculturalists, fishermen, or hunters. These three professions were grouped based off the preliminary observation that these individuals did not frequently travel beyond the spatial scope of either neighboring village. Group 2 was for those individuals who primarily worked in lumber harvesting and lumber transportation. Group 3 was assigned to those who primarily transported commercial goods for others. Group 1-3 allotments were based off of participants' self-reported daily activities (a structured part of scheduled individual check-ins). Group 4 included the three teachers in Santa Emilia and the UPCH field researcher who has resided in Santa Emilia for more than 1 year. This final group was composed of the only participants with easily identifiable professions, as they were the only inhabitants of Santa Emilia with identified external employers.

Each of these profession-based groups was then further divided by time to approximate the biological plausibility of exposure to various mosquito-mediated infectious diseases. 'Day points' were categorized as any GPS point falling between 5:00am and 6:59pm, while those points falling between 7:00pm and 4:59am were classified as 'night points.' With these divisions, 'day points' represented average regional *Aedes aegypti* biting patterns (arboviruses such as Dengue, Chikungunya, Zika, etc.), and 'night points' represented the average recorded regional biting patterns of the *Anopheles darlingi* mosquito (*Plasmodium vivax and falciparum*). After these two divisions, the 164,316 total GPS data points were categorized as follows:

Table 2: Raw GPS Data Counts – Allotted by Time and Profession

Profession	Day	Night	Total:
Farm/Fish/Hunt	**44,628**	**33,626**	78,254
Wood	**15,379**	**12,156**	27,535
Transport	**26,910**	**20,223**	47,133
Teacher	**6,208**	**5,186**	11,394
Total:	93,125	71,191	164,316

These 8 data subsets were then spatially merged to the previously discussed buffer zone shapefile in order to determine both 1) total time spent, and 2) percentage of all time spent in each of the zones by each of these 8 subsets.

To initially visualize the GPS data points, point-density analyses were conducted, with symbology set to a heat-map color scale. This point-density analysis was conducted at both a fine and large resolution to allow for overall visualization of the entire study area (large resolution) and of sub-zones such as Santa Emilia, Nauta, or Iquitos (fine resolution). Needing to then accurately quantify the statistical significance of these point-density visualizations, the Anselin's Local Indicator of Spatial Autocorrelation (LISA) and Getis-Ord's $G_i^*(d)$ spatial statistics were performed[33,34]. Briefly, the LISA statistic detects a local clustering of spatial autocorrelation, and the $G_i^*(d)$ statistic assesses the strength of localized hotspots.

To complete both spatial statistics, each of these 8 data subsets was spatially joined to a grid hexagons measuring 250 meter across from flat side to flat side. After this hexagon join, a log conversion was completed on the count column of each hexagon, and both statistics were run for each of the 8 subsets using the continuous edges only analysis with false discovery rate corrections applied.

8 – Limitations and Delimitations:

There are several limiting and delimiting factors in the methods of this study that should be noted. First, the inclusion criteria for age (≥ 18yrs) meant that the movement of children were not included in the scope of this study. This is important as it was observed that many of the 'transporters' in this study often traveled all the way to Iquitos with their entire families. With this, it should be noted that children are at an increased likelihood for being both carriers of malaria disease and desirable feeding targets of mosquitos.

Second, the intent of this study was to characterize the normal regional movement of Santa Emilia residents. Because of this, a non-randomized sample of individuals was used. This sampling method decreases the generalizability of this study's findings.

Finally, the collection period per individual was capped at 15 days. This inherently limits the patterns observed to not include those movements made at less frequent intervals (for example, all Santa Emilia adults travel to Nauta once a year to vote under Peru's mandatory voting laws). Additionally, this study was conducted in a 2 month period between peak intensity of either the dry or wet seasons. It is understood that on occasion river levels during these peaks can impede boat passage and therefore normal travel patterns as well.

CHAPTER 3
Results

---

1 – Introduction:

The spatial distribution of locations visited by participating Santa Emilia residents was mapped using a point density function. To test the statistical significance of this visualization, local spatial statistics, the Local Indicator of Spatial Autocorrelation (LISA) and G_i*(d), were performed. In all cases, statistically significant results confirmed Point Density distribution patterns, though often less contiguous and in greater detail. Additionally, raw GPS points were spatially joined to digitized study area features of importance, and time counts were provided to numerically corroborate the visual patterns of the Point Density, LISA, and Getis maps.

2 – Findings:

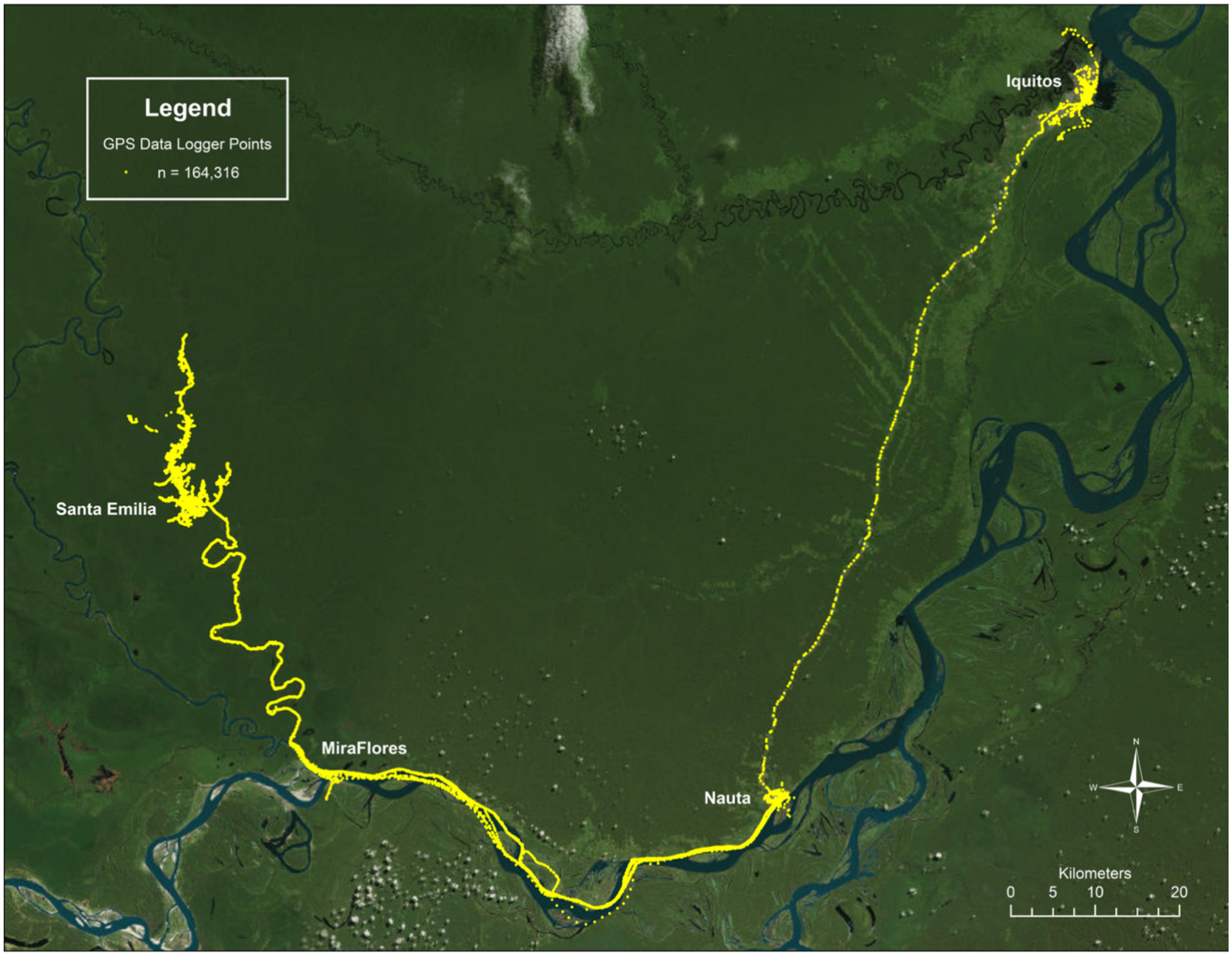
Important landscape features were digitized and then classified as either ‘village’, ‘urban center’, ‘roadway’, ‘waterway’, or ‘forest’. These categories of features were further broken down into ‘zones.’ Two zones were defined for the scope of this analysis, they included a ‘Remote Zone’ and an ‘Urban Centers Zone.’ These two zones together did not include the entirety of the study area, but instead focused on those areas of greatest interest.

For each of these grouped analyses (1- overall study area, 2- just the Remote Zone, 3- just the Urban Centers Zone), a point density function was first used to best visualize all of the raw data. To determine statistical significance of these density visualizations the LISA and G_i*(d) analyses were then performed (in that order). The primary intent of the two spatial statistics was to validate or contrast the initial patterns of the point density function. The secondary intent of conducting these statistics as pairs was to validate or contrast the results reported between the two different spatial statistics. Because of this intent, paragraphs following figures below are presented in a linear narrative format, and focus on highlighting the differences found through this process.

Study findings are organized below by overall study area, and then by zones. Within each of these three subsections, figures and tables are presented in the same order for clarity. This order was 1) figure of map features, 2) table of corresponding GPS data counts, 3) table of GPS points broken down by profession and time variables, 4) point density visualization, 5) map of LISA analysis, and 6) map of G_i*(d) analysis. For brevity, not all six aspects are presented for each of these three subsections, as this process was conducted in a linear manner (i.e. – building off of previously found results).

Study Area-Level Analyses:

Figure 7: GPS Data Logger Points (Study Area)



Raw GPS locations (**164,316** points), obtained from tracking the movements of 31 individuals

Figure 7 displays the 164,316 raw GPS points that were included in the analysis of this study. The movement data is concentrated throughout the study area, largely following the riverine transport network previously described and to main roads/paths. One of the 32 individuals recruited lost their device, therefore 31 participants contributed to this total data number. The scope of the study area, and those features included for all further analysis, was defined by the spatial spread of data in **Figure 7**.

Figure 8: Study Area Features (Study Area)

Map of the Study Area, highlighting only the urban centers, roadways, waterways, and villages passed by Santa Emilia residents during their normal regional movement

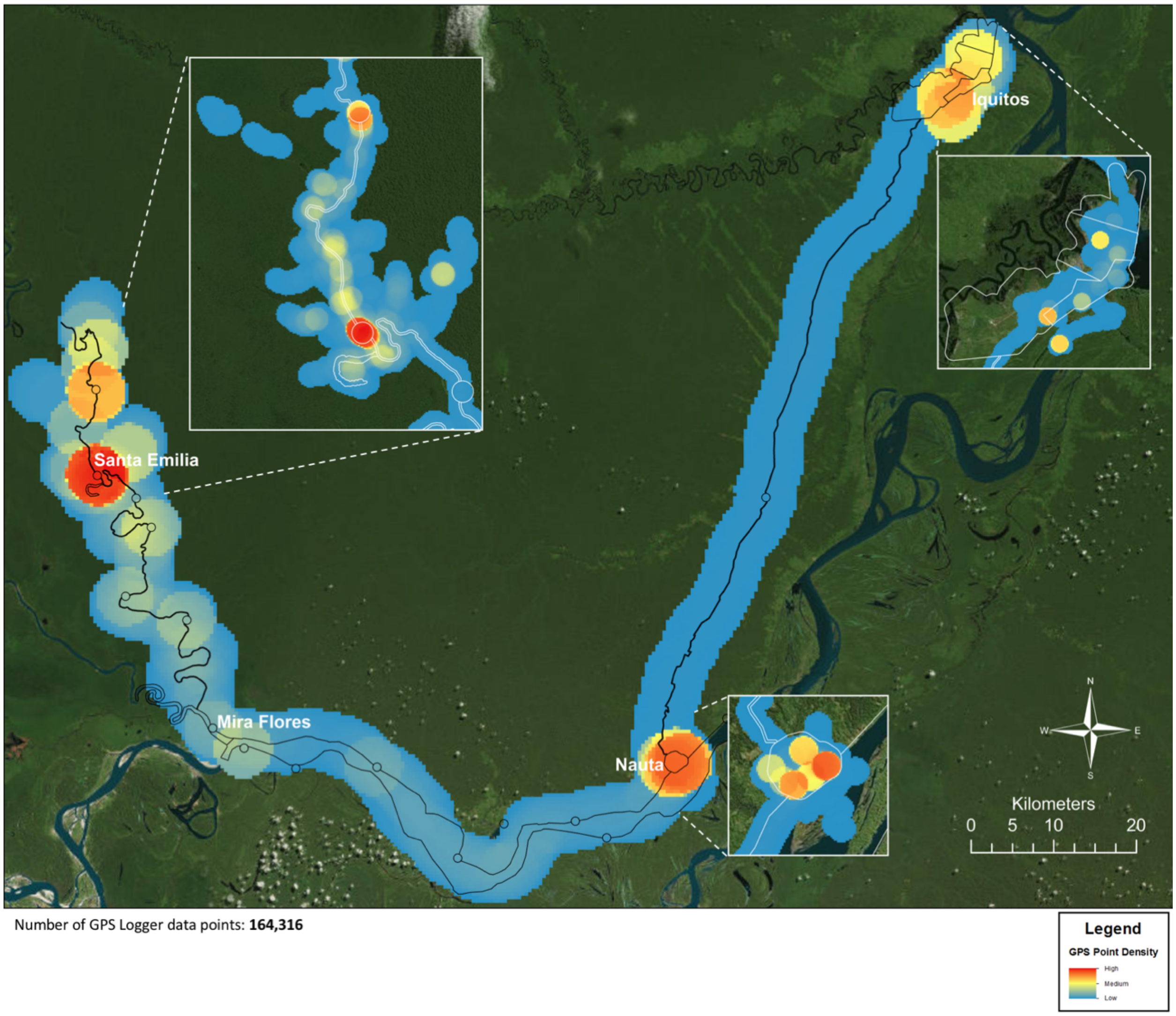
The resulting study area is described by **Figure 8** in greater depth. Map features are color-coded by categories, with villages as white, urban centers as orange, roadways as red, and waterways as blue. Many more villages and waterways exist within scope of this map window; however, only those features within proximity of the GPS data (**Figure 7**) are labeled here (**Figure 8**). Hence, the digitization of the Marañon, and Tigris rivers are cut short, and none of the villages located on riverways beyond these digitizations are displayed – despite their plentitude.

Table 3: Time Spent per Feature Category (Study Area)

Category	# of GPS points	Time (hours)	% of Total
Villages	119,022	9,919	72.4%
Urban Centers	21,637	9,451	13.2%
Roadways	213	18	0.1%
Waterways	12,052	1,004	7.3%
Forest	11,392	949	6.9%
Totals:	164,316	13,693	100%

The number of GPS points recorded per study feature is reported in **Table 3**. Numbers of GPS points were converted to a time variable with the knowledge that each GPS data logger was programmed to record a data point every 5 minutes. It is, therefore, plausible that these time calculations account for some transitory time (< 5minutes) into or out of said feature. Therefore, time as a variable should be read as an approximation rather than an absolute value. Results from **Table 3** show that the majority of participant time was spent within villages (9,919 hours). Including both villages and urban centers, 85.6% of all time was spent within the confines of a population center. The remainder of the time (14.4%) was spent on waterways, forest, or roadways (in decreasing order).

Figure 9: GPS Point Density, All Data (Study Area)

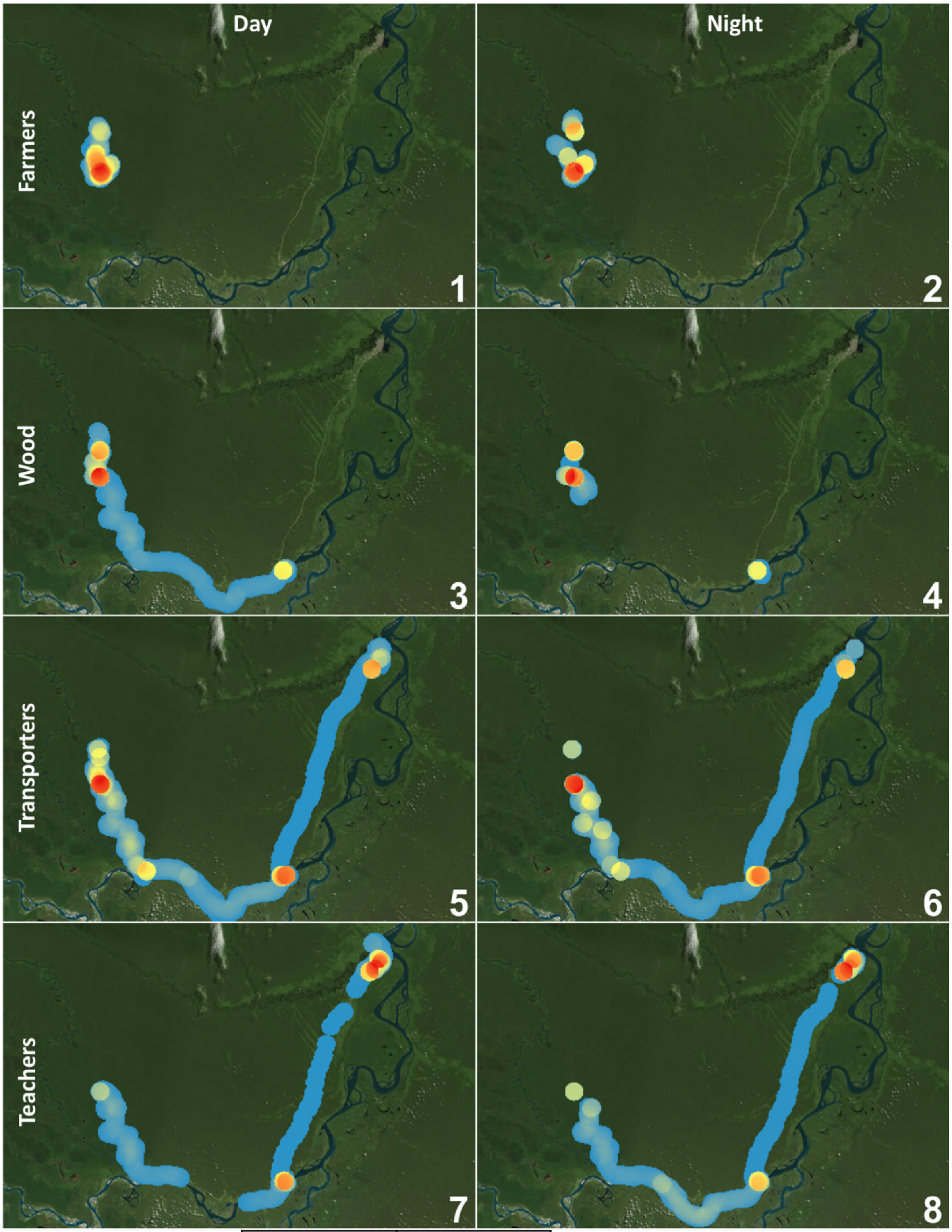


Number of GPS Logger data points: **164,316**

The spatial distribution of the raw data was mapped using a point density function in **Figure 9**. Color warmth corresponds with higher density values, whereas cool colors correspond to low data density. Data density was highest in the home village of Santa Emilia, neighboring villages on the quebrada Nahuapa, and in the urban centers of Nauta and Iquitos. Map insets display finer detail for those areas with the highest data density. Variation in movement data is more accurately displayed here than in the raw projection in **Figure 7**. Results of **Figure 9** established the need for the zone analyses that were performed.

Figures 7-9 and **Table 3**, above, each include all GPS data collected. These tables and figures established the heterogeneous spatial distribution of the raw data along the concentrated transportation routes and population centers, and helped to initially visualize the movement of Santa Emilia residents throughout the surrounding region. However, these visualizations do not account for additional variables of interest (time of day and participant profession). **Figures 10-24** therefore account for these additional variables. Tables and Figures below are displayed in grids for best comparability. These grids follow the layout of **Table 2**.

Figure 10: GPS Point Densities by Profession & Time (Study Area)



Number of GPS points per map:

1] **44,628**	2] **33,626**
3] **15,379**	4] **12,156**
5] **26,910**	6] **20,223**
7] **6,208**	8] **5,186**

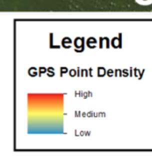
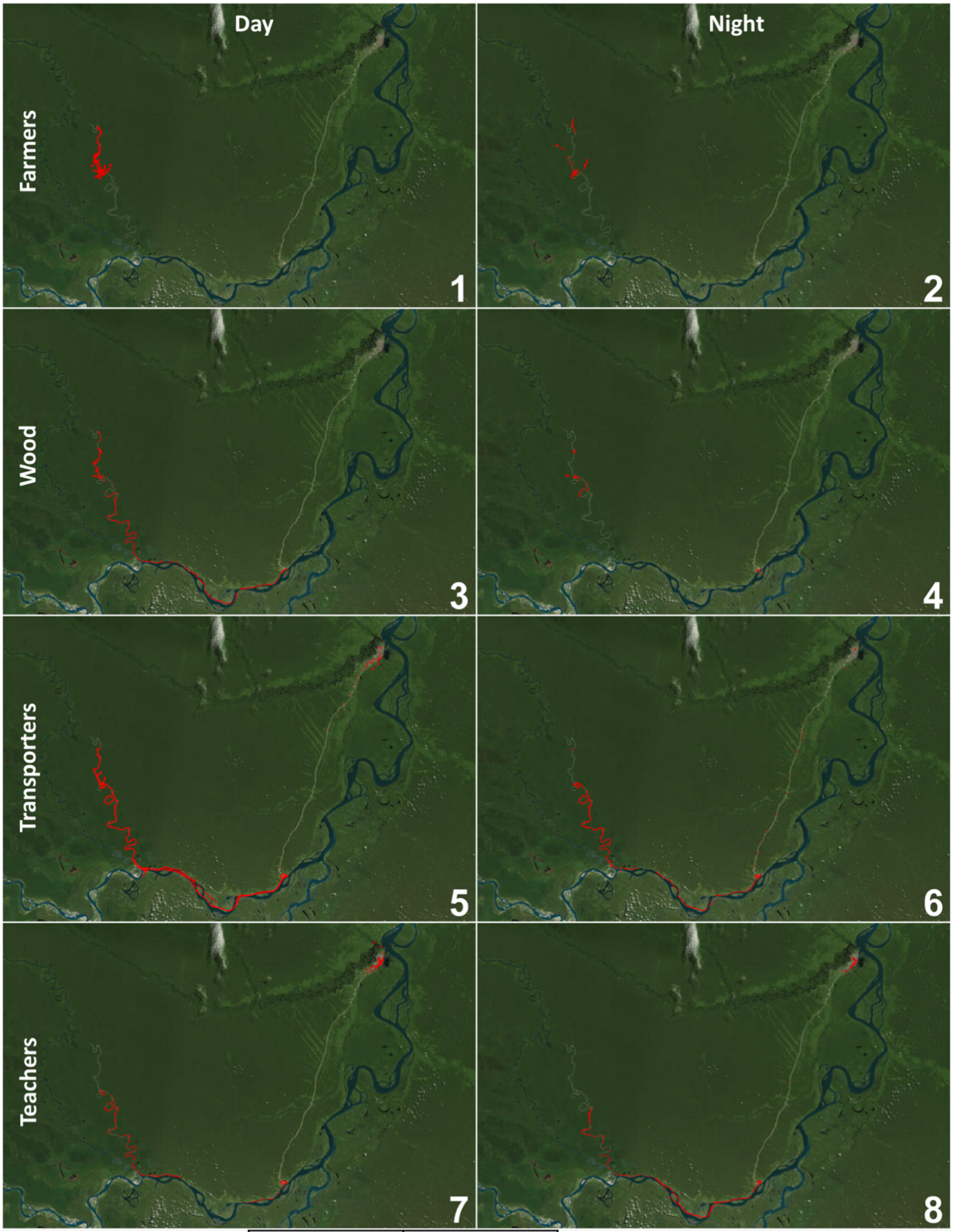


Figure 10 visualizes the same 164,316 data points as **Figure 9**, but accounts for both time and profession as differentiating variables. Counts for each of these data subsets are reported in the figure caption. Professional groups, as defined in **Table 2**, were abbreviated for visual clarity. These abbreviations were used consistently for the remainder of Tables and Figures.

The spatial heterogeneity of the data described in Figure 3 is further visible in **Figure 10**. Variation in data spread corresponds most heavily by profession, and secondarily by time of day. **Figures 10.1-10.2** (farmers, fishermen, and hunters) show the greatest variation in comparison to **Figures 10.3-10.8** (all other professional categories). Distribution of GPS points was not homogenous between profession-time allotments. This was most visible in **Figures 10.7-10.8** (teachers) where GPS counts were between one-half to one-seventh the number of the other profession-time allotments. Areas of highest point density remain highly homogeneous throughout all profession-time allotments, with highest densities reported in the home village and distant urban centers.

Figures 11-12, below, display the results of the local spatial statistics, LISA and G_i*(d), both accounting for profession-time data allotments. Only statistically significant areas (250 meter hexagons) are displayed. See **Appendix 3** for a map showing the extent of hexagons used to run these spatial statistics.

Figure 11: Local Moran's I Analysis by Profession & Time (Study Area)



Number of HH Clusters per map:

1] **574**	2] **201**
3] **711**	4] **113**
5] **1,554**	6] **742**
7] **536**	8] **583**

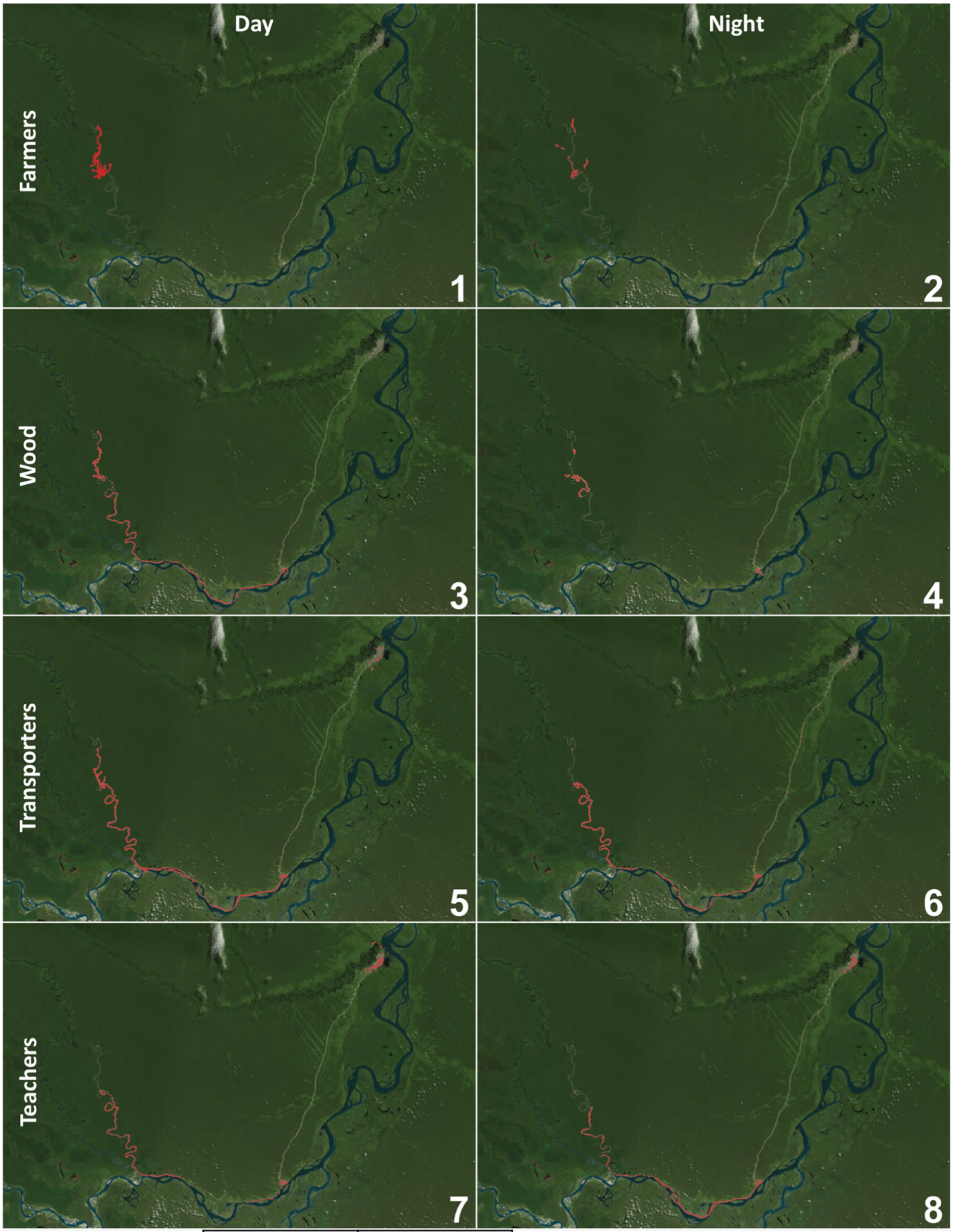
Legend

Local Morans I Cluster Analysis

- High-High Cluster
- Low-Low Cluster
- High-Low Outlier
- Low-High Outlier

Statistically significant high-high clusters of local autocorrelation were displayed red in **Figure 11**. No statistically significant low-low clusters, high-low outliers, or low-high outliers were reported. Again, variation in heterogeneity appeared mostly bound by profession rather than time; in agreeance with the results of **Figure 10**. The quantity of high-high clusters of local autocorrelation was higher during day hours than during night hours for all professions (**Figure 11** footnote table). High-high clustering of local autocorrelation (**Figure 11**) occurred in concordance to areas of highest point density (**Figure 10**). However, no correlation between number of significant clusters (**Figure 11** footnote table) and number of GPS points (**Figure 10** footnote table) was observed. For example, though the **Figure 10.1** accounted for the highest number of GPS points collected (44,628), the number of statistically significant clusters in **Figure 11.1** remained low (574). In contrast, **Figure 10.5** accounted for relatively low point totals (26,910), but **Figure 11.5** reported a greater number of significant spatial clusters (1,554) relative to other **Figure 11** profession-time allotments.

Figure 12: Getis Analysis by Profession & Time (Study Area)



Number of Hotspots (99%/95%/90%) per map:

1] **900/0/0**	2] **409/0/0**
3] **1,819/0/0**	4] **270/191/0**
5] **2,487/0/0**	6] **1,849/0/0**
7] **1,359/0/0**	8] **1,456/0/0**

Legend

Getis Hotspot Analysis

Cold Spot - 99% Confidence	Hot Spot - 90% Confidence
Cold Spot - 95% Confidence	Hot Spot - 95% Confidence
Cold Spot - 90% Confidence	Hot Spot - 99% Confidence

The local hotspot spatial statistic G_i*(d) results reported in **Figure 12** further confirmed the trends reported in **Figures 10-11** – mainly that heterogeneity appeared mostly bound by profession rather than time. No statistically significant cold spots were found. The majority of all hot spots reported were within a 99% confidence interval. Ratios of **Figure 12** hotspots to **Figure 10** point allocations approximately match ratios of **Figure 11** clusters to **Figure 10** point allocations, as described above, further demonstrating concordance in the visual results expressed by these two different local spatial statistics.

Remote Zone Analyses:

Following the analysis of the entire study area, two subareas were defined (referred to as 'zones') to further investigate spatial and temporal heterogeneity. These zonal analyses benefit from smaller denominators, and, therefore, result in finer detail. Remote Zone findings are described below, followed by the Urban Centers Zone results.

Figure 13: Study Area Features (Remote Zone)

On the Quebrada Nahuapa, Santa Emilia is the fifth of six rural remote permanent villages

The extent of the 'Remote Zone' is shown in **Figure 13**. The Remote Zone, as defined here, began with the mouth of the quebrada Nahuapa and included all villages, additional waterways, and surrounding forested area along the length of the stream. Six remote permanent villages were included within this zone, the study site (Santa Emilia) is the fifth along the quebrada Nahuapa.

Table 4: Time Spent (hours) per Study Feature (Remote Zone)

Day Hours	Farmers	Wood	Transporters	Teachers
Santa Emilia	3,016	675	1,196	25
Victor Raul	5	174	24	0
Downstream Villages (4)	0	2	13	2
Waterways	247	145	207	26
Forest	451	99	57	2
Night Hours	**Farmers**	**Wood**	**Transporters**	**Teachers**
Santa Emilia	2,674	682	1,059	19
Victor Raul	0	195	0	0
Downstream Villages (4)	0	.3	36	7
Waterways	55	10	82	19
Forest	73	20	6	1

Time spent within the Remote Zone, by each profession-time allotment, is reported in **Table 4**. Time was reported in units of hours, rather than as percentages of total time, in order to convey the biological plausibility of participants having been bitten by mosquitos during their time in each study feature. In all cases, except teachers (day and night) and transporters (night only), more time was spent in the one village (Victor Raul) upstream of Santa Emilia than the sum of time spent in the four villages downstream of the home village. This is interesting considering the next closest village downstream (28 de Julio) is closer in distance to Santa Emilia than Victor Raul (**Figure 13**).

It should be noted that because the intent of this study was to characterize the normal regional movement of individuals living in Santa Emilia, *a priori* knowledge of teacher movement was used in the enrollment of these individuals. Village teachers were enrolled the day before their monthly planned weeklong trip outside of Santa Emilia. Because of this, movement data on teachers is heavily skewed to reflect time outside of Santa Emilia.

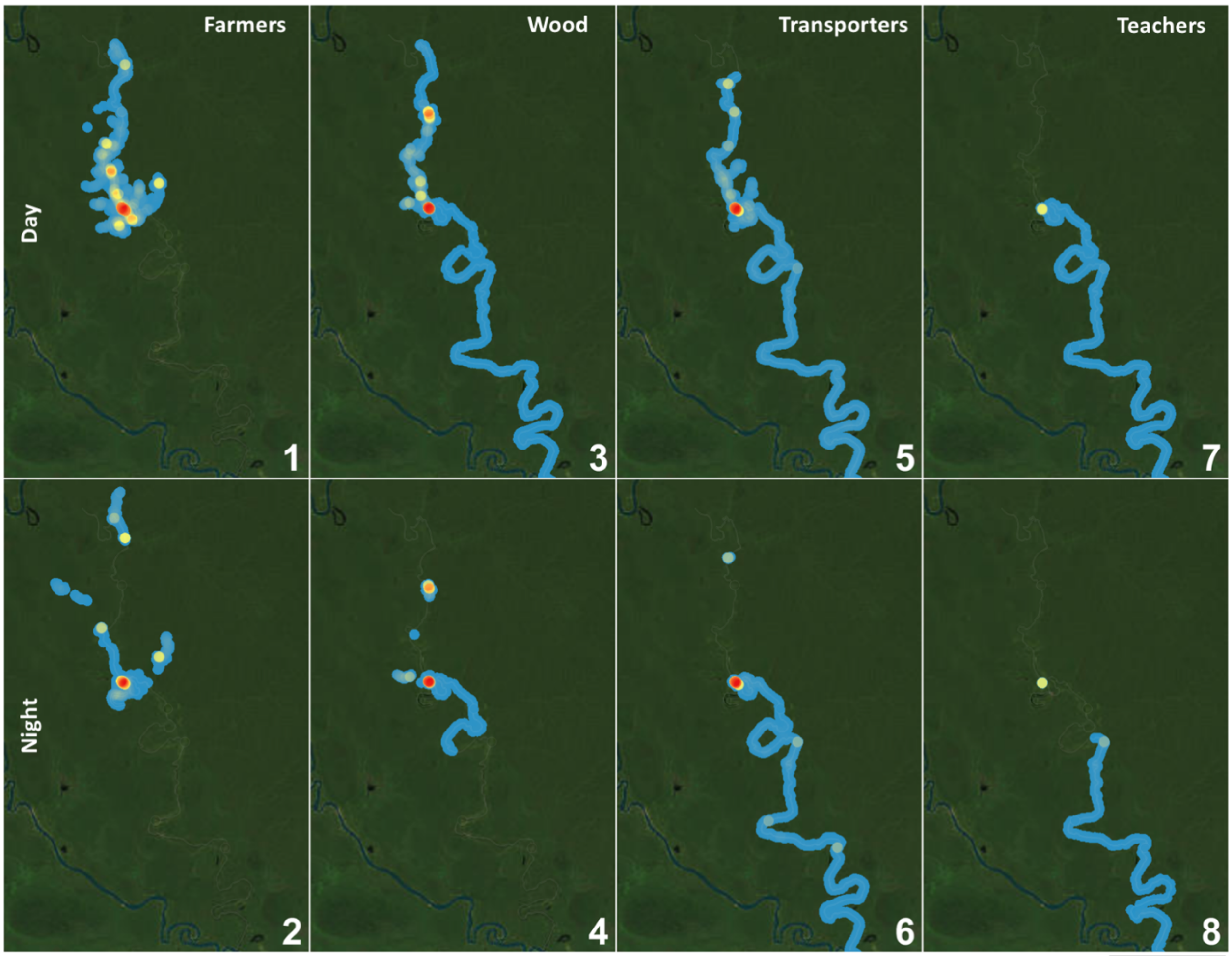
Table 5: Percentage of Time Spent Within and Beyond the Remote Zone

Profession-Time	**Within Remote Zone**	**Outside of Remote Zone**
1 – Farmers-Day	100%	0%
2 – Farmers-Night	100%	0%
3 – Wood-Day	85%	15%
4 – Wood-Night	90%	10%
5 – Transporters-Day	67%	33%
6 – Transporters-Night	70%	30%
7 – Teachers-Day	6%	94%
8 – Teachers-Night	5%	95%

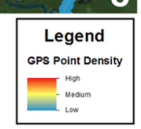
As reported in the footnote tables of **Figure 10** (point density of Study Area) and **Figure 14** (point density of Remote Zone), the number of GPS points recorded within the Remote Zone versus in the overall study area varied greatly by profession-time allotments. Percentages of this variation are displayed above in **Table 5**. These percentages are reported vertically using the same professional-time allotments, and are labeled using the same numbers as the professional-time map grids used throughout. Though this table is a restatement of data, it serves as a reference for the following Remote Zone based results subsection.

Again, it should be noted that village teachers were intentionally enrolled the day before their monthly planned weeklong trip outside of Santa Emilia. Therefore, it is not unexpected that >90% of this profession's recorded data fell outside of the Remote Zone (**Table 5**, rows 7 and 8).

Figure 14: GPS Point Densities by Profession & Time (Remote Zone)



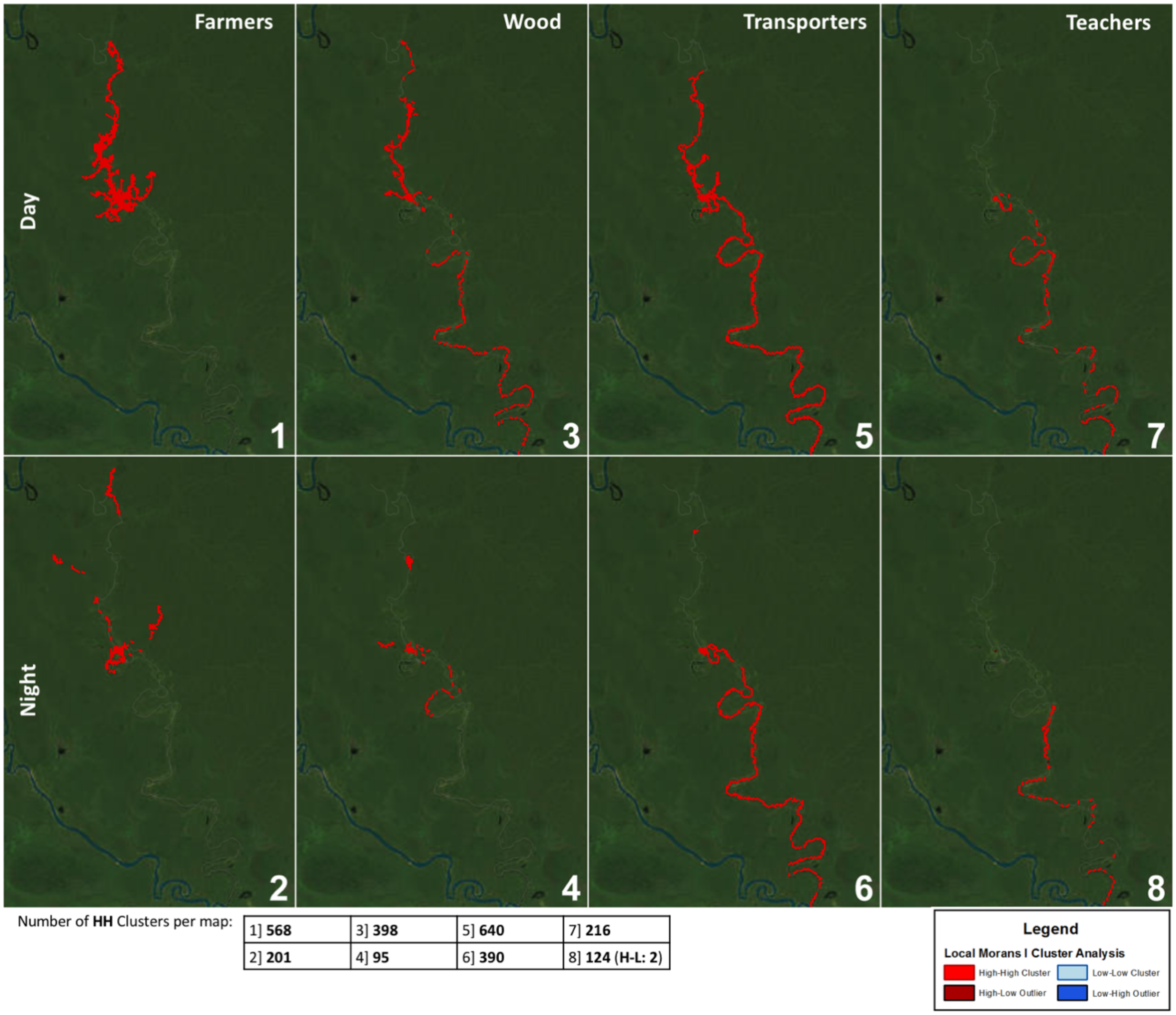
Number of GPS points per map:				
	1] **44,628**	3] **13,128**	5] **17,961**	7] **656**
	2] **33,626**	4] **10,887**	6] **14,197**	8] **546**



The spatial distribution of the Remote Zone data was mapped using a point density function in **Figure 14**. As in the overall study area, within the Remote Zone spatial heterogeneity varied by profession-time allotments. The highest density throughout all eight maps was the home village Santa Emilia. In **Figures 14.1-14.5** (farmers, lumber, and transporters-day) the next highest point densities all existed upstream from Santa Emilia. **Figures 14.6-14.8** (teachers and transporters-night) show mostly downstream movement. **Figures 14.1-14.4** show scattered movement throughout much of the Remote Zone, while **Figures 14.5-14.8** show movement mostly constrained to the quebrada Nahuapa. **Figure 14** daytime maps (**14.1, 14.3, 14.5, 14.7**) all display continuity in spatial distribution. In contrast, **Figure 14** nighttime maps (**14.2, 14.4, 14.6, 14.8**) all display non-contiguous densities within the scope of the Remote Zone.

Figures 15-16, below, report results from the LISA and G_i*(d) statistics repeated for Remote Zone data. The decrease in study area correlated to a smaller number of hexagons included in the denominator of these functions, resulting in finer-scale results. Again, in both figures only statistically significant areas (250 meter hexagons) are displayed. See **Appendix 4** for a map of the hexagon extent used to run both of these spatial statistics for the Remote Zone subset of GPS data.

Figure 15: Local Moran's I Analysis by Profession & Time (Remote Zone)

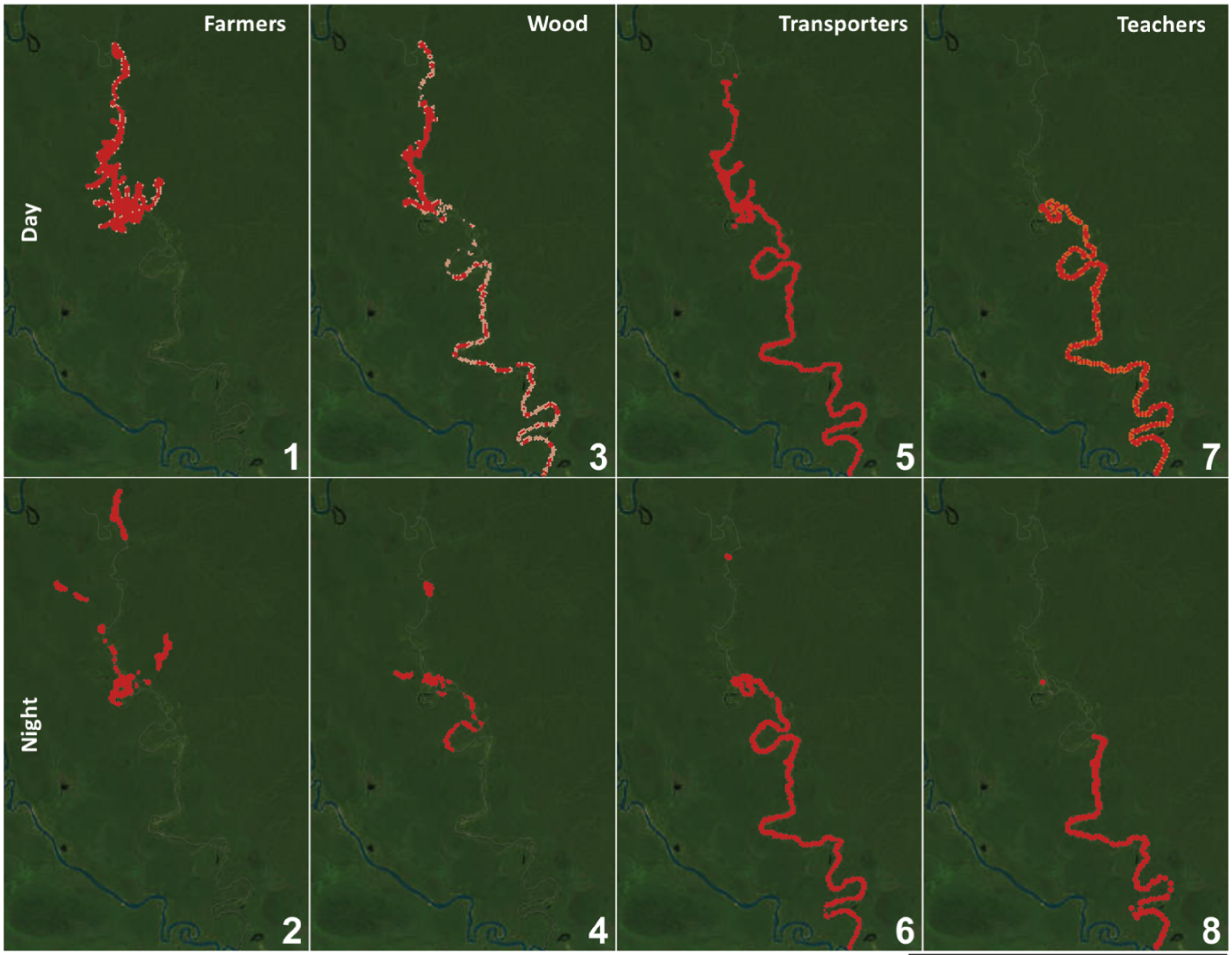


Number of **HH** Clusters per map:

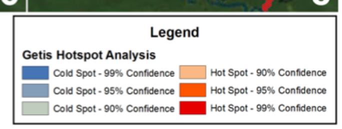
1] **568**	3] **398**	5] **640**	7] **216**
2] **201**	4] **95**	6] **390**	8] **124 (H-L: 2)**

Results of the LISA analysis for the Remote Zone are reported in **Figure 15**. Maps **15.1-15.4** (farmers and lumber) had pockets of significant local autocorrelation both on the quebrada Nahuapa waterway and into the dense forest. By contrast, maps **15.5-15.8** (transporters and teachers) reported high-high clustering only on and along the banks of the quebrada Nahuapa riverine transport system. In all profession-time allotments, high-high clustering was more contiguous during day hours when compared to continuity of nighttime clustering. Map **15.8** did report two statistically significant high-low outliers, though these are most likely due to GPS loss of signal as they both fell on the waterway track.

Figure 16: Getis Analysis by Profession & Time (Remote Zone)



Number of Hotspots (99%/95%/90%) per map:				
	1] **786/0/114**	3] **522/2/493**	5] **1,211/0/0**	7] **624/396/0**
	2] **409/0/0**	4] **234/0/0**	6] **919/0/0**	8] **669/0/0**



Overall spatial patterns of variation in significant hotspots (**Figure 16**) mirror variation patterns in significant clusters (**Figure 15**). Hotspot analysis resulted in more contiguous patterns, especially along the quebrada. **Figure 15.8** and **Figure 16.8** show this most clearly.

In all profession-time allotment maps, the number of significant hotspots reported (**Figure 16** footnote table) was higher than the number of significant clusters reported (**Figure**

15 footnote table), despite these two statistics using the same number of profession-time allotted points (**Figure 14** footnote table). This suggests that the G_i*(d) analysis better distinguishes the highly heterogeneously distributed spatial data (i.e. data highly confined to waterway transport networks).

Again, as reported in **Figures 11-12**, maps **15.5-15.6** and maps **16.5-16.6** reported more statistically significant clusters of autocorrelation and hotspots (respectively) than maps **15.1-15.2** and maps **16.1-16.2**. This is despite 100% of farmer movement having existed within the Remote Zone and ~30% of transporter data having fallen outside of the Remote Zone (**Table 4**).

Urban Centers Zone Analyses:

The second zone used for sub-area analysis was the 'Urban Centers Zone'. This second zone included the urban centers of Nauta and Iquitos, and all study features between them (Iquitos-Nauta highway, 1 village were stops are frequently made, and surrounding forest). Because 100% of Farmer/Fisher/Hunter data collected fell within the scope of the Remote Zone (**Table 4**), this profession based category of data was excluded from the Urban Centers Zone analyses.

Table 6: Time Spent (hours) per Feature Category (Urban Centers Zone)

Day Hours	Wood	Transporters	Teachers
Nauta	144	388	140
Roadway	0	7	2
Iquitos – San Juan	0	7	7
Iquitos – Belen	0	29	153
Iquitos – Iquitos	0	4	140
Iquitos – Punchana	0	7	2
Night Hours	**Wood**	**Transporters**	**Teachers**
Nauta	105	338	88
Roadway	0	6	3
Iquitos – San Juan	0	9	1
Iquitos – Belen	0	0	143
Iquitos – Iquitos	0	.3	96
Iquitos – Punchana	0	6	0

Time spent within the Urban Centers Zone, by each profession-time allotment, is reported in **Table 6**. Time spent in each area was reported in units of hours instead of percentages of total time to better express the potential for participants' exposure to mosquitos during the study's time period in relation to different features. During the scope of this study, 100% of time spent in the Urban Centers Zone by wood harvesters was spent in the town of Nauta. Those in the transporters profession category did travel via the highway to Iquitos, but the majority of their time (day and night) was still recorded in Nauta. This was not the case with Santa Emilia teachers. As previously described, the teachers and field microscopist who work and stay in Santa Emilia, are all not from this village. All four of these individuals have families and permanently reside in the city of Iquitos. As reported in **Table 6**, these participants spent nearly all of their time in Iquitos within only two of its four neighborhoods – Belen and Iquitos central.

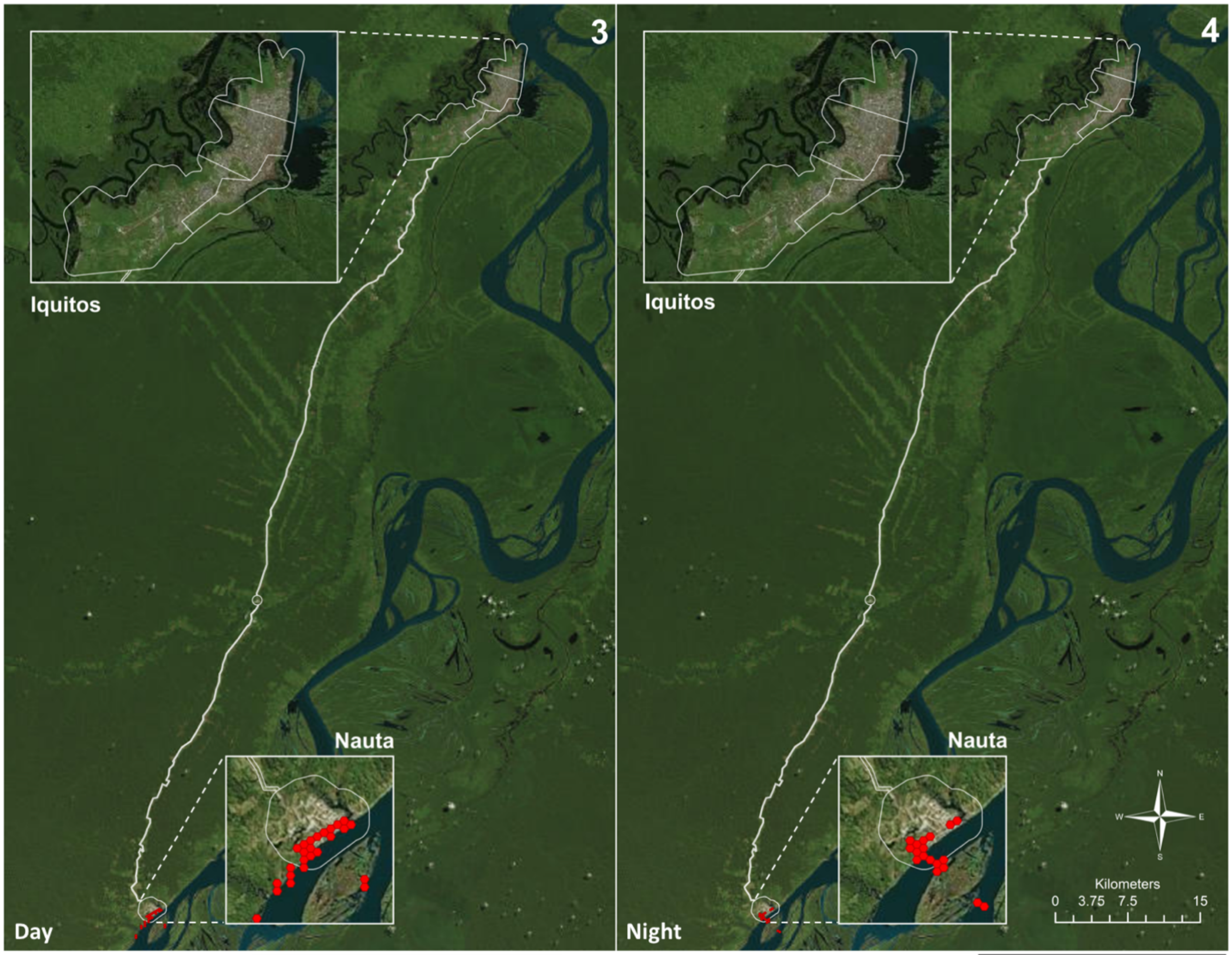
Table 7: Percentage of Time Spent Within and Beyond the Urban Centers Zone

Profession-Time	**Within Urban Centers Zone**	**Outside of Urban Centers Zone**
3 – Wood-Day	11%	89%
4 – Wood-Night	10%	90%
5 – Transporters-Day	20%	80%
6 – Transporters-Night	21%	79%
7 – Teachers-Day	85%	15%
8 – Teachers-Night	77%	23%

Time spent within and outside of the Urban Centers Zone is reported as a percentage in **Table 7**. Despite the hundreds of hours spent inside the Urban Centers Zone by participants (**Table 6**), the majority of total study time was spent outside of the Urban Centers Zone (**Table 7**). This holds true for all professions and during both time periods, apart from teachers, who's special study recruitment strategy has already been described above.

Below, **Figures 17-22** report the LISA and G_i*(d) statistics results for the six profession-time allotments included in this Urban Centers Zone analysis. The decrease in study area correlated to a smaller number of hexagons included in the denominator of these functions, resulting in finer-scale results. Again, in both figures only statistically significant areas (250 meter hexagons) are displayed. See **Appendix 5** for a map of the hexagon extent used to run both of these spatial statistics for the Urban Centers Zone subset of GPS data.

Figure 17: Local Moran's I Analysis for **Lumber Workers**, by Time (Urban Centers Zone)



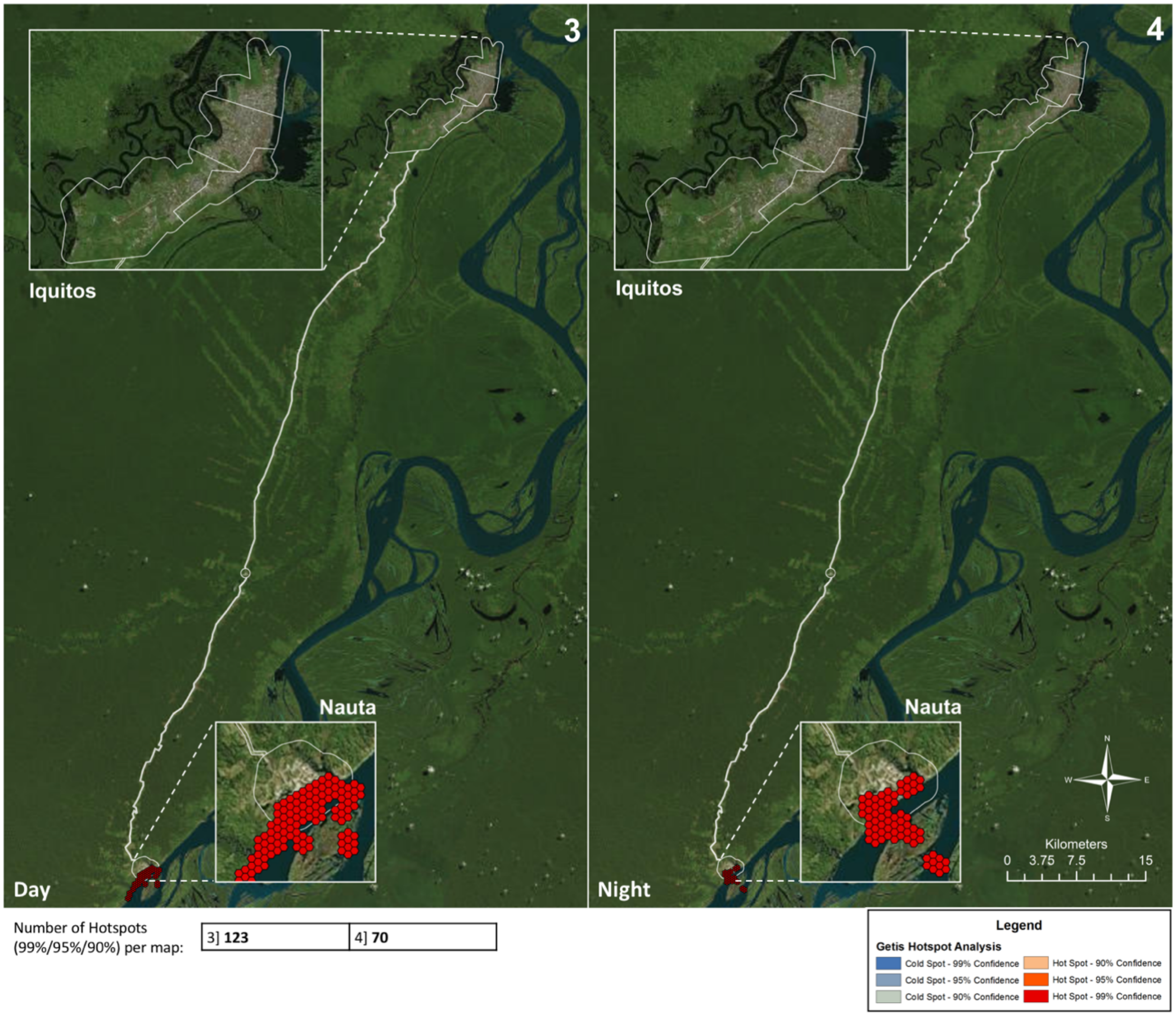
Number of **HH** Clusters per map: 3] **26** 4] **18**

Legend

Local Morans I Cluster Analysis

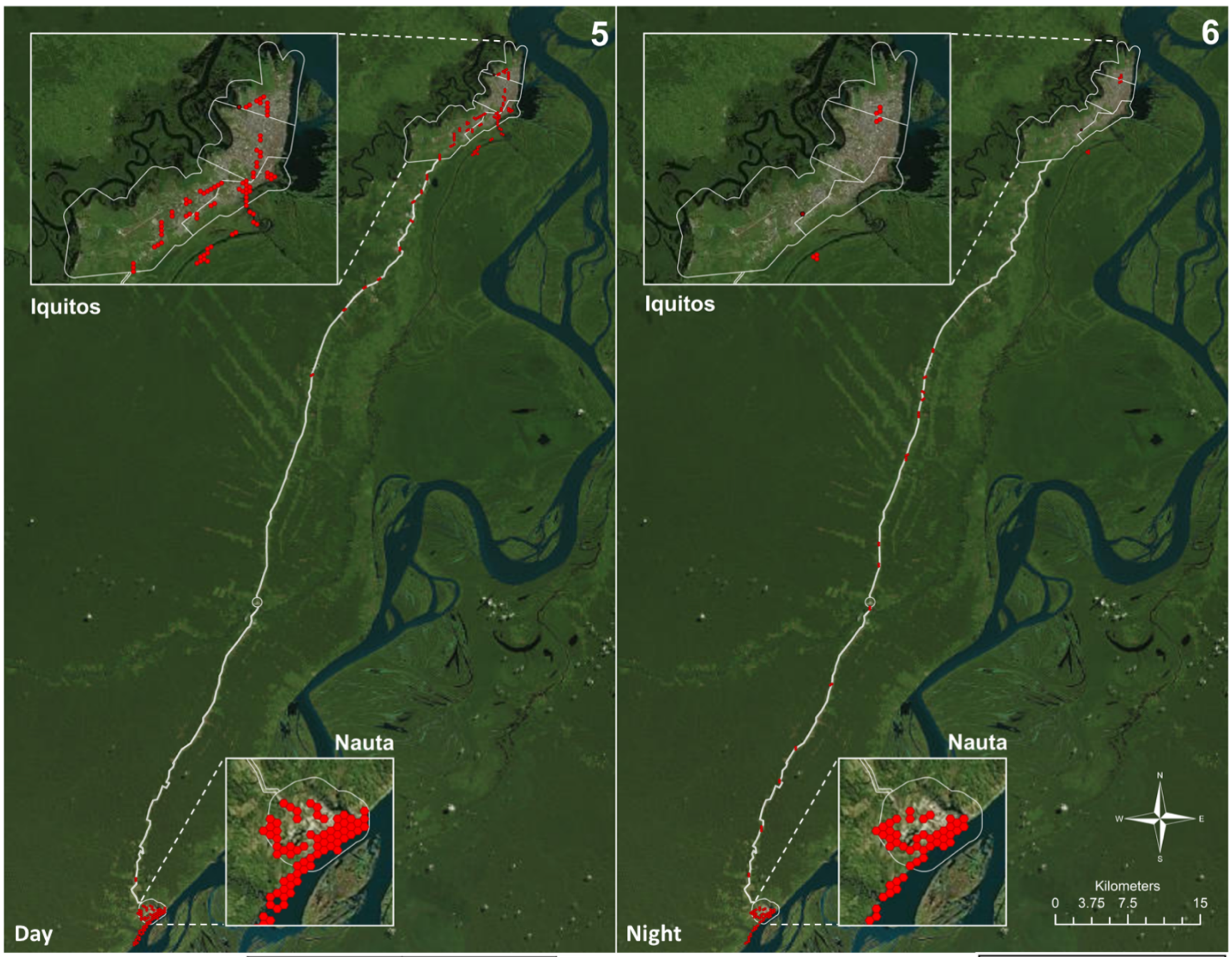
High-High Cluster | Low-Low Cluster
High-Low Outlier | Low-High Outlier

Figure 18: Getis Analysis for **Lumber Workers**, by Time (Urban Centers Zone)



Within the scope of data catchment for this study, Santa Emilia lumber workers (**Figures 17-18**) were recorded as only traveling to the urban center of Nauta. This is logical, as *a priori* knowledge informed that Nauta contains a regional lumber mill. Little variation was reported in movement between day and night hours, though less spatial spread was reported during nighttime hours. In both time allotments, statistically significant clusters are limited to the coastal edge of Nauta, with the majority of clusters falling on the southwest edge of the town. The Nauta lumber mill is located in this southwest corner, possibly explaining this observed spatial distribution. Results from the LISA statistic (**Figure 17**) and those from $G_i^*(d)$ (**Figure 18**) are spatially concordant, though the number of statistically significant hotspot clusters (**Figure 18** footnote table) was nearly quadruple the number of statistically significant high-high autocorrelation clusters (**Figure 17** footnote table).

Figure 19: Local Moran's I Analysis for **Transporters**, by Time (Urban Centers Zone)



Number of **HH** Clusters per map:	5] **177** (H-L: **1**)	6] **91** (H-L: **1**)

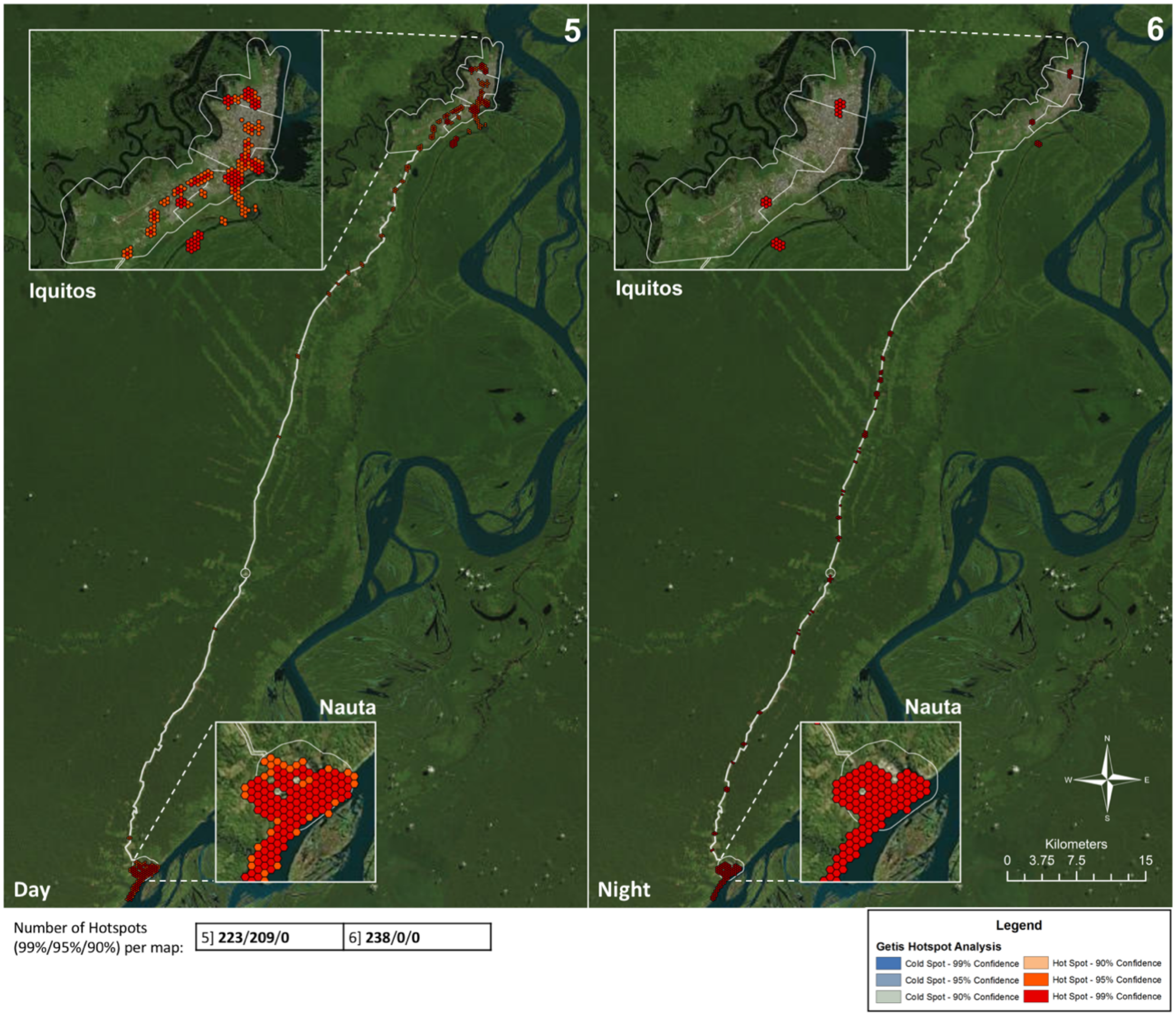
Legend

Local Morans I Cluster Analysis

High-High Cluster | Low-Low Cluster

High-Low Outlier | Low-High Outlier

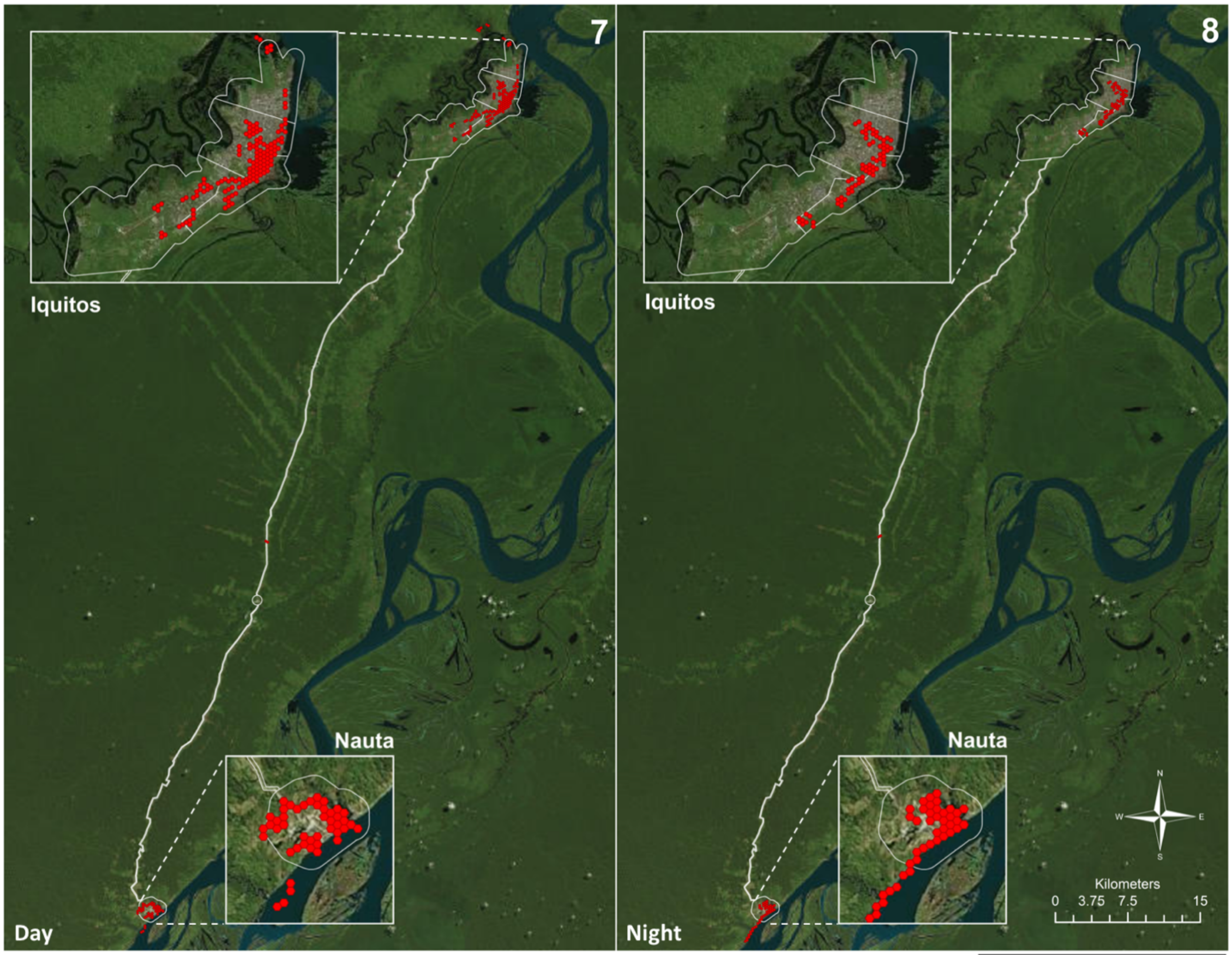
Figure 20: Getis Analysis for **Transporters**, by Time (Urban Centers Zone)



Santa Emilia participants professionally classified as 'transporters' traveled to both Nauta and Santa Emilia during the scope of this study. The distribution of high-high clusters of autocorrelation for these participants was highly heterogeneous, with clear patterns especially found within the city limits of Iquitos (**Figure 19**). These heterogeneous patterns were mirrored in the G_i*(d) hotspot analysis (**Figure 20**), though number of hotspots was more robust for these same spatial patterns. Both of these statistically significant spatial distributions agree with number of hours recorded within each neighborhood of Iquitos, as reported above in

Table 6. In both **Figure 20** and **Figure 21**, spread of significant movement is more disperse during day hours, and more spatially concentrated during night hours.

Figure 21: Local Moran's I Analysis for **Teachers**, by Time (Urban Centers Zone)



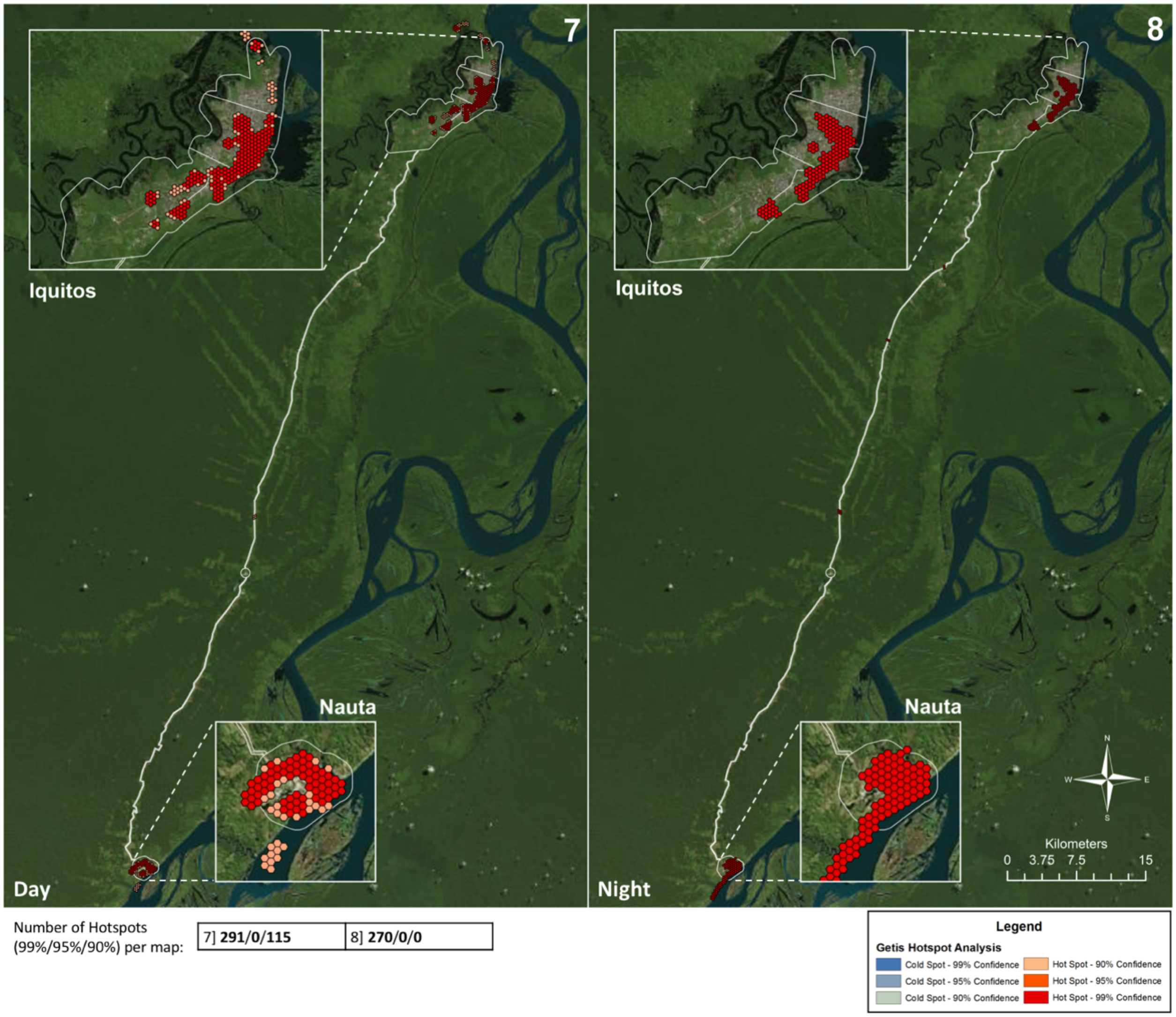
Number of **HH** Clusters per map:	7] **197**	8] **118**

Legend

Local Morans I Cluster Analysis

High-High Cluster, Low-Low Cluster, High-Low Outlier, Low-High Outlier

Figure 22: Getis Analysis for **Teachers**, by Time (Urban Centers Zone)



The scheduled regional movement of Santa Emilia teachers has already been noted. Where these individuals spend their time during these leaves is described in **Figures 21-22**. Significant high-high autocorrelation clustering (**Figure 21**) and hotspots (**Figure 22**) are highly heterogeneously distributed. Nearly all significant areas lie within either Nauta or Iquitos. In Nauta, distribution is clustered in the Northeast corner and along the non-river adjacent back of the town. In Iquitos, much of the distribution of significance is clustered within the neighborhoods of Belen and Iquitos central. The spread within Iquitos was greater during daytime hours, while nighttime hours reported the tight contiguous clustering within these two neighborhoods.

3 – Additional Findings:

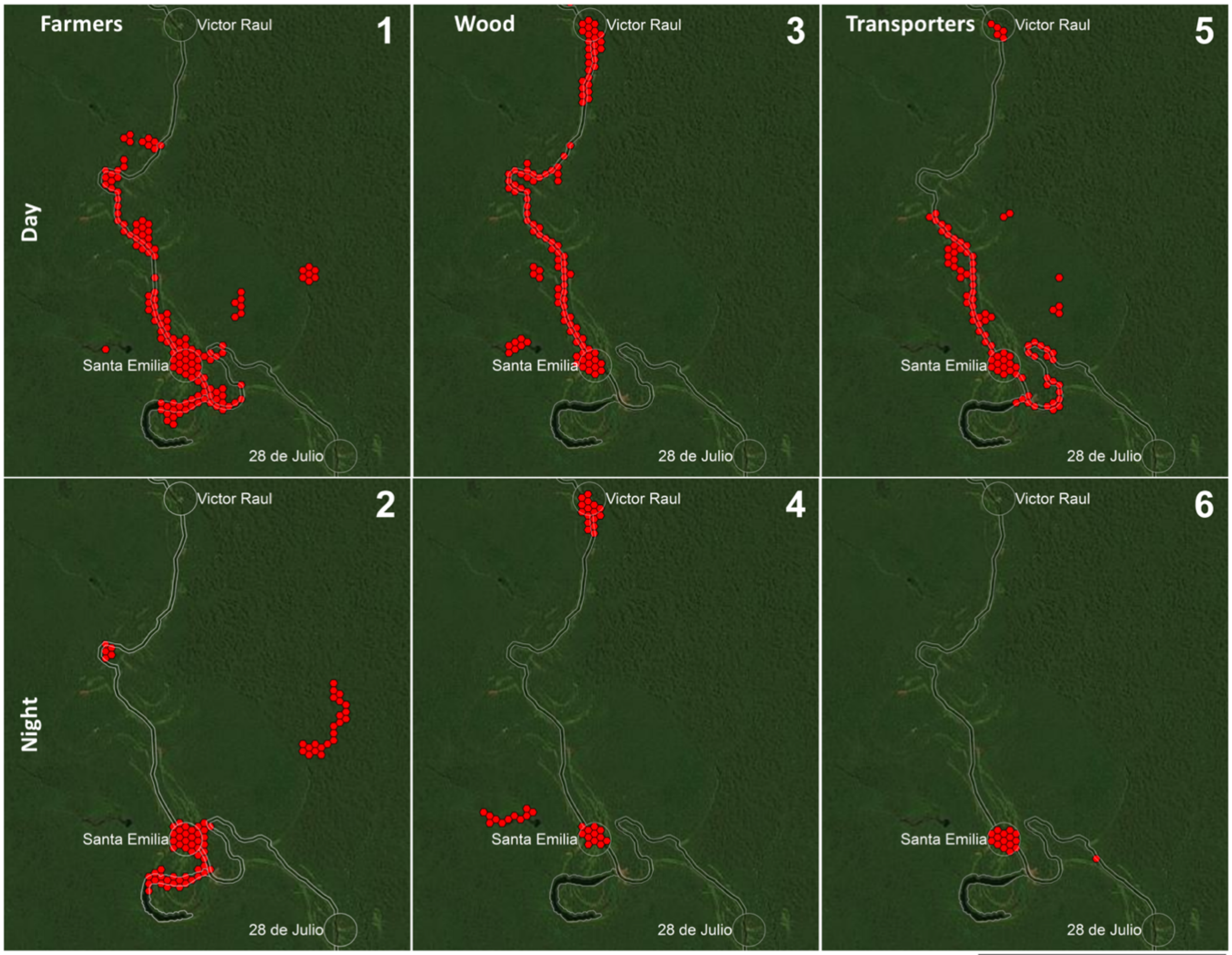
After completing the overall study area analyses, and then subsequently the two zonal analyses, it was determined that a still finer-scale analysis of the immediate surrounding of Santa Emilia was necessary. The scope of this final round of analyses was bound by the two neighboring villages up and downstream. In such remote rural areas village boundaries and land ownership are not well defined. Therefore, it is likely that the forest and waterway space between any two remote villages is shared on a regular basis. The scope of these final analyses, referred to as the 'Santa Emilia Subzone,' focused specifically on these potentially shared water, forest, and agricultural areas.

Santa Emilia Subzone Analyses:

Below, **Figures 23-24** report LISA and G_i*(d) statistic results for the six profession-time allotments included in this Santa Emilia Subzone analysis. Teachers were not included in this subzone analysis because they did not travel upstream during the scope of this study.

This extent of this final subzone greatly decreased the study area of focus, but the method of defining the hexagon grid used for these statistics was further refined as well. Statistically significant results from **Figure 16** were selected and expanded with a 300 meter buffer. This resulted in a unique plot of contiguous hexagons for each of the professional-time allotments, different from the blanket grids described in **Appendices 3-5**. This method resulted in a much smaller denominator, and thus even finer-scale results. See **Appendix 6** for maps of these varying hexagon extents, used to run the following LISA and $G_i^*(d)$ analyses. Once again, in both figures below, only statistically significant areas (250 meter hexagons) are displayed.

Figure 23: Local Moran's I Analysis by Profession & Time (Santa Emilia Subzone)



Number of **HH** Clusters per map:

1] **114**	3] **117**	5] **102**
2] **98**	4] **39**	6] **40**

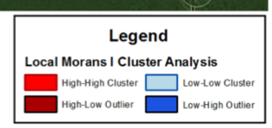
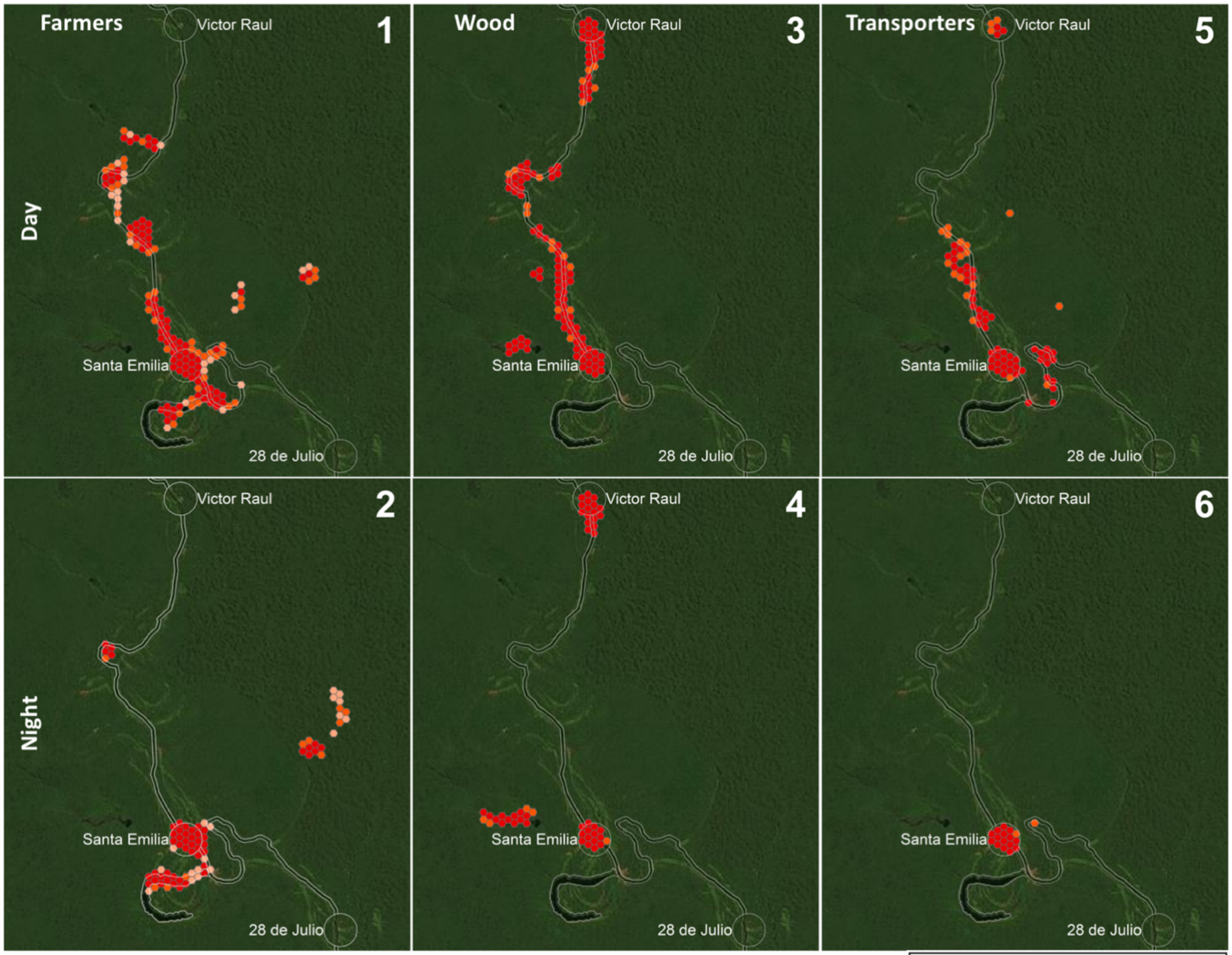


Figure 24: Getis Analysis by Profession & Time (Santa Emilia Subzone)



Number of Hotspots (99%/95%/90%) per map:

1] **85/41/25**	3] **105/17/0**	5] **62/22/0**
2] **61/14/25**	4] **43/5/0**	6] **29/11/0**

Legend

Getis Hotspot Analysis

- Cold Spot - 99% Confidence
- Cold Spot - 95% Confidence
- Cold Spot - 90% Confidence
- Hot Spot - 90% Confidence
- Hot Spot - 95% Confidence
- Hot Spot - 99% Confidence

Results of the Santa Emilia Subzone analysis are reported above in **Figures 23-24**. These maps show in fine detail the spatially heterogeneous distributions of each profession-time allotment. Professional group #1 (farmers, fishermen, and hunters) show the greatest variation in significant autocorrelation clusters (**Figure 23**) and hotspots (**Figure 24**) into the forest, and reported the only areas of significance on the lake south of Santa Emilia. Daytime data is less interrupted, and is focused linearly around the transport waterway. All nighttime data reported greater discontinuity in patterns of significant autocorrelation clusters and hotspots. Significant time was spent inside the village of Victor Raul by lumber harvesters during the day and night (**Figures 23.3-23.4** and **24.3-24.4**), and by transporters during the day (**Figures 23.5** and **24.5**).

4 – Summary:

Spatial distribution was highly heterogeneous along transportation routes and population centers regardless of profession-time allotment or scope of analysis (study area versus zones versus subzone). In all cases, statistically significant results from the local spatial statistics, Anselin's LISA clustering analyses and Getis-Ord's G_i*(d) hotspot analyses confirmed the point density function distribution patterns. Analysis on data separated by time and profession further explained the variation seen in the raw data. Performing these analyses in multiple rounds of decreasing size ultimately resulted in a more fine-scale understanding of the significant spatial clustering and hotspots within the data. Further informed by spatially joined GPS data time counts, the sum of these analyses adequately describes the natural variation in normal movement experienced by Santa Emilia residents.

CHAPTER 4
Discussion and Conclusions

---

1 – Introduction:

The primary goal of this study was to characterize the normal regional movement of Santa Emilia residents. This was done in context of the ongoing transmission and reintroduction of malaria occurring throughout the deep-Amazon region of Peru. Local spatial statistics revealed a highly heterogeneous spatial distribution around neighboring remote villages, within the regional urban centers of Nauta and Iquitos, and along the riverine networks which connect these features. Diurnal and nocturnal sub-analyses revealed further heterogeneity within the data, which interpreted with knowledge of vector behavior, assess differences in potential for disease transmission as mediated by either *Anopheles* or *Aedes* mosquitos. Finally, using participant profession as a variable revealed further variation in the spatial and temporal distribution of movement. These differences highlight those professions potentially key to the ongoing regional cycles of malaria transmission and reintroduction.

2 – Context:

Findings describing spatial patterns of regional movement for Santa Emilia residents, whom live in a previously identified hotspot[22,27], are novel amongst the currently existing literature. This study's comparisons of diurnal and nocturnal activity spaces applied to arboviruses as well as malaria plasmodia are novel for the remote region of Amazonian Peru, and hinge on Guagliardo's recent novel findings concerning the plausibility of region *Aedes* expansion via riverboats[18,28,29]. Finally, this study's findings identify three professions as potentially key in the documented ongoing regional transmission and reintroduction of malaria – 1) loggers, 2) riverboat transporters, and 3) government employed teachers. The identification of the first two professions is echoed in the literature[22,23,24], but the third profession of teachers is a novel finding in this arena.

3 – Discussion:

Local spatial statistics revealed a heterogeneous spatial distribution among collected movement data. This heterogeneity in hotspots was primarily congregated 1) around neighboring remote villages, 2) within the regional urban centers of Nauta and Iquitos, and 3) along the riverine networks which connect these features (Figures 11-12, 15-24). This confirms the hypothesis that in remote rural locations of the Loreto Department of Peru, human movement is largely bound by the naturally existing riverine networks[22]. With little to no transportation infrastructure in the region, most residents remain near waterbodies at all times[29]. The only participants with significant travel beyond the linear confines of the regional waterways are those categorized as farmers/hunters/fishermen or loggers (**Figure 23-24**). This is intuitive given that Santa Emilia chacras (small land plots) largely exist beyond the immediate flood banks of the quebrada Nahuapa[21], and the best hunting is found far beyond the boundaries of any village. For the most part, lumber harvesters stay close to waterways given that all harvested lumber is transported to lumber mills via these waterways[22].

When Santa Emilia residents leave the immediate confines of their village, they are more likely to travel upstream than downstream. This statistically significant upstream movement extended all the way into the village boundaries of Victor Raul – the terminal village on the quebrada Nahuapa (**Figures 15-16**, **23-24**). This is striking for two reasons. First, all access to markets and other goods and services exist downstream in the town of MiraFlores or in the urban center of Nauta. Therefore, it was hypothesized that most movement would be downstream[31]. This however was not the case (**Table 4**). Secondly, in cases where Santa Emilia residents travel to another neighboring village, but not travel beyond the quebrada Nahuapa, it was hypothesized that such travel would occur to the village closest in proximity[31]. Again, this was not observed. Instead, the upstream village of Victor Raul was visited significantly more than the downstream village of 28 de Julio. Considering the distance to Victor Raul by boat is double the distance to 28 de Julio (**Figure 13**), this was unexpected. Furthermore, field observations suggest that the village 28 de Julio is not only closer in distance, but also a much larger village than Victor Raul.

As was hypothesized, using participant profession as a differentiating variable revealed significant differences in the spatial and temporal distribution of human movement. These differences highlighted those individuals potentially key to the ongoing transmission and reintroduction of malaria throughout the region[22]. Participants classified as primarily farmers, fishermen, or hunters reported the least amount of spread throughout the overall study area, and totaled in the least number of significant clusters (**Figure 11**) or hotspots (**Figures 12**). Lumber harvesters traveled beyond the Remote Zone to Nauta, while transporters and teachers continued this journey to the city of Iquitos as well. Those professional classifications which traveled farthest throughout the overall study area also reported the highest number of statistically significant results (**Figures 11-12**).

Finally, sub-analyses including time as a variable revealed further heterogeneity within the data. These patterns could provide estimates on potential environmental contact with *Anopheles* mosquitos (nighttime biters) or *Aedes* mosquitos (daytime biters), given previously described local vector behavior[21]. Within the Remote Zone (**Table 4**), all professions spent hours of contact time in neighboring villages. Wood harvesters reported 175+ hours of contact during the day, and 195+ hours of contact time during the night, in other quebrada Nahuapa villages. Transporters and teachers reported far less contact hours in other villages, but still reported 35+ contact hours both during the day and night. Put in context of previously reported human biting rates (HBR) and entomological inoculation rates (EIRs)[21], and given the previous identification of Santa Emilia as a regional hotspot for malaria[27], these high sums of contact hours spent in other remote villages suggest ample opportunity for localized transmission to occur. Beyond malaria, daytime contact patterns could be applied to a potential future where *Aedes* mosquito populations extend into remote rural areas of the Loreto Department[18,28,29], such as Santa Emilia.

Urban Centers zonal analysis detected similar findings (**Table 6**). While wood harvesters reported the highest number of contact hours in other villages within the Remote Zone, it was transporters and teachers who reported high contact hours with other populations within the Urban Centers Zone. Transporters reported 300+ contact hours, both day and night, in the town of Nauta, 40+ day hours and 15+ night hours in the city of Iquitos. Similarly, teachers reported 140 daytime contact hours and 80+ nighttime hours in Nauta. As was hypothesized, Santa Emilia teachers spent more time in the city of Iquitos than any other Santa Emilia resident. Despite the 12-16 hour travel time, Santa Emilia teachers reported 300+ day hours and nearly

250 night hours inside the city of Iquitos. Nearly all contact hours were spent in the Iquitos neighborhoods of Belen and Iquitos central. Belen is lowest socio-economic status neighborhood in Iquitos, and is largely composed of floating houses on the Amazon River. With this ecology, the potential for Santa Emilia teachers to share infected bites with Belen residents is assumed particularly high. Given the current endemicity of arboviruses, such as Dengue, in Iquitos, these human travel patterns further suggest the biological plausibility of arbovirus infected residents returning to Santa Emilia to inoculate potentially future established *Aedes* mosquito populations.

4 – Strengths and Limitations:

To the knowledge of the author, this was the first study utilizing wearable GPS technology in the remote amazon region of Peru. However, this study only captured up to 15 days of contiguous movement data per participant. Therefore, the findings of this study cannot be used to generalize regional movement which may occur at lesser frequencies. Additionally, while GPS technology has become increasingly suited for such field applications, the units used still had limited battery life and required an AC power source to be charged. These charging needs, coupled with participation being limited to 15 days, meant that longer trips away from the home village could not be studied. For example, during data collection period of this study two young men from Santa Emilia went to Nauta to work in a factory and remained there for 2 week periods before returning. These individuals, and their movement, could not feasibly be recorded given the technological limitations of this study.

The GPS data loggers chosen for this study were small, lightweight, water resistant, could not be accidentally turned off, and did not emit any light or sound. These strengths also presented some limitations in their field use. Several units were destroyed while swimming because participants forgot they were wearing them. Also, without light indication it was difficult to know when battery life was waning. Although the manufacturer projected 5 days of life per charge, some units nevertheless were returned without battery after only 3 days of use. Because of this, there were spans of time for some participants when no data was collected. Additionally, issues with participant adherence to protocol did occur on occasion, though these instances were few thanks to the frequency of check-ins (every 3 days).

Finally, this research was designed as a substudy for an ongoing 5 year investigation being conducted by the Universidad Peruana Cayetano Heredia (UPCH). Because of this, the data collected expressly for this study included only movement data and limited participant demographics. While a useful and novel implementation of GPS technology, without participant seroprevalence results linked to this captured movement data, the discussion and conclusions of this study lean on a number of assumptions – primarily that, as previously recorded, human biting rates and entomological infection rates throughout the remote region remained high during the scope of this study[27,31]. Additionally, all non-malarial interpretations of the data were done under the biological plausibility established by Guagliardo et al.[18,28,29], and not by any current Santa Emilia *Aedes* population prevalence data.

CHAPTER 5
Recommendations & Implications

---

1 – Recommendations:

The results of this study suggest that some professions (transporters and teachers) are more likely to frequently travel beyond the scope of the Remote Zone and into peri-urban and urban population centers. It therefore could be recommended that these individuals be the primary focus of any future increase in active case detection or disease surveillance activities corresponding with the noted annual regional transmission and reintroduction of malaria. Additionally, if in the documented population spread of *Aedes* mosquitos continues into the more remote areas of the Loreto Department, these same professions should be the primary focus of contact tracing and other epidemiology methods for any future arbovirus outbreaks.

Along the quebrada Nahuapa, additional active case detection and malarial treatments should be made available to the village Victor Raul. Though smaller and even more distant from available public health resources, the findings of this study suggest that this village is at higher risk of becoming a malaria hotspot than other larger and more proximal villages downstream of Santa Emilia.

Finally, methodologically this study further established the feasibility and need for more integrated and novel uses of GPS technology both in 1) remote regions of the world, and 2) epidemiology studies on infectious disease spread. The methods for data collection and analysis described here should serve as a foundation for similar studies to be conducted in the future.

2 – Implications:

If the results-based recommendations of this study were readily implemented, it is plausible that less annual malaria transmission and reintroduction may occur in the remote region of the quebrada Nahuapa in coming years. More broadly applied, if riverboat transporters and remote-based government teachers throughout the Loreto Department were more actively screened and treated for malaria, then it is plausible that less regional transmission and reintroduction of malaria would occur on an annual basis in the Amazon.

3 – Conclusions:

Given that Santa Emilia is an identified malaria hotspot in remote Amazonian Peru, the results of this human movement study suggest that local riverboat transporters and remote-based government school teachers are most likely involved in the ongoing regional transmission and reintroduction of malaria which is annually reported throughout the region. The methods used are novel for such a setting, and provide statistically significant insights towards the greater hypothesis that some Santa Emilia residents travel greater distances, at higher frequencies, than other village residents. As a nested study, these findings successfully identify hotspots in other population centers where Santa Emilia residents traveling work and sleep. These locations provide the UPCH research team with a defensible hypothesis for where to next implement their numerous ongoing active PCR-based malaria surveillance studies.

References

1 *World Malaria Report 2016.* (2016). Geneva: World Health Organization.

2 Hotez, P. J. Ten global "hotspots" for the neglected tropical diseases. *PLoS Negl. Trop. Dis.* **8**, e2496, doi:10.1371/journal.pntd.0002496 (2014).

3 Sluydts, V. *et al.* Spatial clustering and risk factors of malaria infections in Ratanakiri Province, Cambodia. *Malar. J.* **13**, 387, doi:10.1186/1475-2875-13-387 (2014).

4 Feachem, R. G. *et al.* Shrinking the malaria map: progress and prospects. *Lancet* **376**, 1566-1578, doi:10.1016/S0140-6736(10)61270-6 (2010).

5 Labeaud, A. D., Bashir, F. & King, C. H. Measuring the burden of arboviral diseases: the spectrum of morbidity and mortality from four prevalent infections. *Popul Health Metr* **9**, 1, doi:10.1186/1478-7954-9-1 (2011).

6 Becker, N. *Mosquitoes and Their Control*. (Springer-Verlag Berlin Heidelberg, 2010).

7 Perkins, T. A. *et al.* Theory and data for simulating fine-scale human movement in an urban environment. *J R Soc Interface* **11**, doi:10.1098/rsif.2014.0642 (2014).

8 Stoddard, S. T. *et al.* The role of human movement in the transmission of vector-borne pathogens. *PLoS Negl. Trop. Dis.* **3**, e481, doi:10.1371/journal.pntd.0000481 (2009).

9 Wesolowski, A. *et al.* Quantifying the impact of human mobility on malaria. *Science* **338**, 267-270, doi:10.1126/science.1223467 (2012).

10 Vazquez-Prokopec, G. M. *et al.* Using GPS technology to quantify human mobility, dynamic contacts and infectious disease dynamics in a resource-poor urban environment. *PLoS One* **8**, e58802, doi:10.1371/journal.pone.0058802 (2013).

11 Guyant, P. *et al.* Malaria and the mobile and migrant population in Cambodia: a population movement framework to inform strategies for malaria control and elimination. *Malar. J.* **14**, 252, doi:10.1186/s12936-015-0773-5 (2015).

12 Pindolia, D. K. *et al.* The demographics of human and malaria movement and migration patterns in East Africa. *Malar. J.* **12**, 397, doi:10.1186/1475-2875-12-397 (2013).

13 Katsuragawa, T. H. *et al.* The dynamics of transmission and spatial distribution of malaria in riverside areas of Porto Velho, Rondonia, in the Amazon region of Brazil. *PLoS One* **5**, e9245, doi:10.1371/journal.pone.0009245 (2010).

14 Paz-Soldan, V. A. *et al.* Assessing and maximizing the acceptability of global positioning system device use for studying the role of human movement in dengue virus transmission in Iquitos, Peru. *Am. J. Trop. Med. Hyg.* **82**, 723-730, doi:10.4269/ajtmh.2010.09-0496 (2010).

15 Paz-Soldan, V. A. *et al.* Strengths and weaknesses of Global Positioning System (GPS) data-loggers and semi-structured interviews for capturing fine-scale human mobility: findings from Iquitos, Peru. *PLoS Negl. Trop. Dis.* **8**, e2888, doi:10.1371/journal.pntd.0002888 (2014).

16 Khan, O. A. *et al.* Geographical information systems and tropical medicine. *Ann. Trop. Med. Parasitol.* **104**, 303-318, doi:10.1179/136485910X12743554759867 (2010).

17 Soto-Calle, V. *et al.* Spatio-temporal analysis of malaria incidence in the Peruvian Amazon Region between 2002 and 2013. *Sci. Rep.* **7**, 40350, doi:10.1038/srep40350 (2017).

18 Guagliardo, S. A. *et al.* Patterns of geographic expansion of Aedes aegypti in the Peruvian Amazon. *PLoS Negl. Trop. Dis.* **8**, e3033, doi:10.1371/journal.pntd.0003033 (2014).
19 Rosas-Aguirre, A. *et al.* Epidemiology of Plasmodium vivax Malaria in Peru. *Am. J. Trop. Med. Hyg.* **95**, 133-144, doi:10.4269/ajtmh.16-0268 (2016).
20 Moreno, M. *et al.* Implications for changes in Anopheles darlingi biting behaviour in three communities in the peri-Iquitos region of Amazonian Peru. *Malar. J.* **14**, 290, doi:10.1186/s12936-015-0804-2 (2015).
21 Lainhart, W. *et al.* Evidence for temporal population replacement and the signature of ecological adaptation in a major Neotropical malaria vector in Amazonian Peru. *Malar. J.* **14**, 375, doi:10.1186/s12936-015-0863-4 (2015).
22 Parker, B. S. *et al.* Hyperendemic malaria transmission in areas of occupation-related travel in the Peruvian Amazon. *Malar. J.* **12**, 178, doi:10.1186/1475-2875-12-178 (2013).
23 Bautista, C. T. *et al.* Epidemiology and spatial analysis of malaria in the Northern Peruvian Amazon. *Am. J. Trop. Med. Hyg.* **75**, 1216-1222 (2006).
24 Pizzitutti, F. *et al.* A validated agent-based model to study the spatial and temporal heterogeneities of malaria incidence in the rainforest environment. *Malar. J.* **14**, 514, doi:10.1186/s12936-015-1030-7 (2015).
25 Branch, O. *et al.* Clustered local transmission and asymptomatic Plasmodium falciparum and Plasmodium vivax malaria infections in a recently emerged, hypoendemic Peruvian Amazon community. *Malar. J.* **4**, 27, doi:10.1186/1475-2875-4-27 (2005).
26 Rosas-Aguirre, A. *et al.* Hotspots of Malaria Transmission in the Peruvian Amazon: Rapid Assessment through a Parasitological and Serological Survey. *PLoS One* **10**, e0137458, doi:10.1371/journal.pone.0137458 (2015).
27 Chuquiyauri, R., Molina, D., Davies, H., Rodriguez, H., Llanos-Cuentas, A., Felgner, P., Vinetz, J. M. (2014). *Genome-Scale Protein Microarray Analysis of P. falciparum and P. vivax symptomatic and Asymptomatic Infections in a Remote Village of the Peruvian Amazon.* Paper presented at the The American Society of Tropical Medicine and Hygiene, New Orleans, LA USA.
28 Guagliardo, S. A. *et al.* Evidence for Aedes aegypti (Diptera: Culicidae) Oviposition on Boats in the Peruvian Amazon. *J. Med. Entomol.* **52**, 726-729, doi:10.1093/jme/tjv048 (2015).
29 Guagliardo, S. A. *et al.* River boats contribute to the regional spread of the dengue vector Aedes aegypti in the Peruvian Amazon. *PLoS Negl. Trop. Dis.* **9**, e0003648, doi:10.1371/journal.pntd.0003648 (2015).
30 LaCon, G. *et al.* Shifting patterns of Aedes aegypti fine scale spatial clustering in Iquitos, Peru. *PLoS Negl. Trop. Dis.* **8**, e3038, doi:10.1371/journal.pntd.0003038 (2014).
31 Llanos-Cuentas, A. *UPCH unpublished data* (Universidad Peruana Cayetano Heredia, 2016).
32 Vazquez-Prokopec, G. M. *et al.* Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. *Int J Health Geogr* **8**, 68, doi:10.1186/1476-072X-8-68 (2009).
33 Waller, L. A., Gotway, C.A. *Applied spatial statistics for public health data*. (John Wiley & Sons, 2004).

34 Ord, J. K., Getis, A. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis* **27**, 286-306, doi:10.1111/j.1538-4632.1995.tb00912.x (1995).

Appendices

Appendix 1

UNIVERSIDAD PERUANA
CAYETANO HEREDIA
Vicerrectorado de Investigación
Dirección Universitaria de Investigación, Ciencia y Tecnología (DUICT)

CONSTANCIA

El Presidente del Comité Institucional de Ética (CIE) de la Universidad Peruana Cayetano Heredia hace constar que el comité de ética aprobó de manera expedita la **ENMIENDA/MODIFICACIÓN** del proyecto de investigación señalado a continuación.

Título del Proyecto	**: "Transmisión de la malaria en comunidades lejanas de región Amazónica del Perú: Fase II"**
Código de inscripción	**: 64024**
Investigador principal	**: Dr. Alejandro Llanos Cuentas**

La **enmienda/modificación** corresponde a los siguientes documentos:

1. **Protocolo de investigación,** versión 1.0 de fecha 23 de mayo del 2016.
2. **Consentimiento informado (sub-estudio: "Estudio de la Migración en Santa Emilia",** versión 1.0 de fecha 23 de mayo del 2016.

Lima, 01 de junio del 2016

COMITÉ INSTITUCIONAL DE ÉTICA-CIE UNIVERSIDAD PERUANA CAYETANO HEREDIA DUICT

Dra. Frine Samalvides Cuba
Presidenta
Comité Institucional de Ética en Investigación

/smr

Av Honorio Delgado 430
Lima 31
P.O. Box 4314 Lima100

T. (51 1) 319-0000 A. 2271 / 2542
E. duict@oficinas-upch.pe

www.upch.edu.pe/vrinve/duict

Appendix 2

EMORY UNIVERSITY | Institutional Review Board

TO: Charles Harless
Principal Investigator
School of Public Health

DATE: June 9, 2016

RE: **Expedited Approval**
IRB00088823
Using GPS Technology to Quantify the Relationship between Human Mobility and Malaria Outbreaks in Rural Indigenous Peru

Thank you for submitting a new application for this protocol. This research is eligible for expedited review under 45 CFR.46.110 and/or 21 CFR 56.110 because it poses minimal risk and fits the regulatory categories F4 and F7 as set forth in the Federal Register. The Emory IRB reviewed it by expedited process on 6/8/2016 and granted approval effective from **6/8/2016** through **6/7/2017**. Thereafter, continuation of human subjects research activities requires the submission of a renewal application, which must be reviewed and approved by the IRB prior to the expiration date noted above. The following documents are approved for use or otherwise acknowledged:

- Study Protocol, version date 6/8/2015
- Consent documents, all version 1.0, all dated May 23, 2016:
  - Consent script, English
  - Information sheet, English
  - Consent script, Spanish
  - Information sheet, Spanish

Any reportable events (e.g., unanticipated problems involving risk to subjects or others, noncompliance, breaches of confidentiality, HIPAA violations, protocol deviations) must be reported to the IRB according to our Policies & Procedures at www.irb.emory.edu, immediately, promptly, or periodically. Be sure to check the reporting guidance and contact us if you have questions. Terms and conditions of sponsors, if any, also apply to reporting.

Before implementing any change to this protocol (including but not limited to sample size, informed consent, study design, you must submit an amendment request and secure IRB approval.

In future correspondence about this matter, please refer to the IRB file ID, name of the Principal Investigator, and study title. Thank you

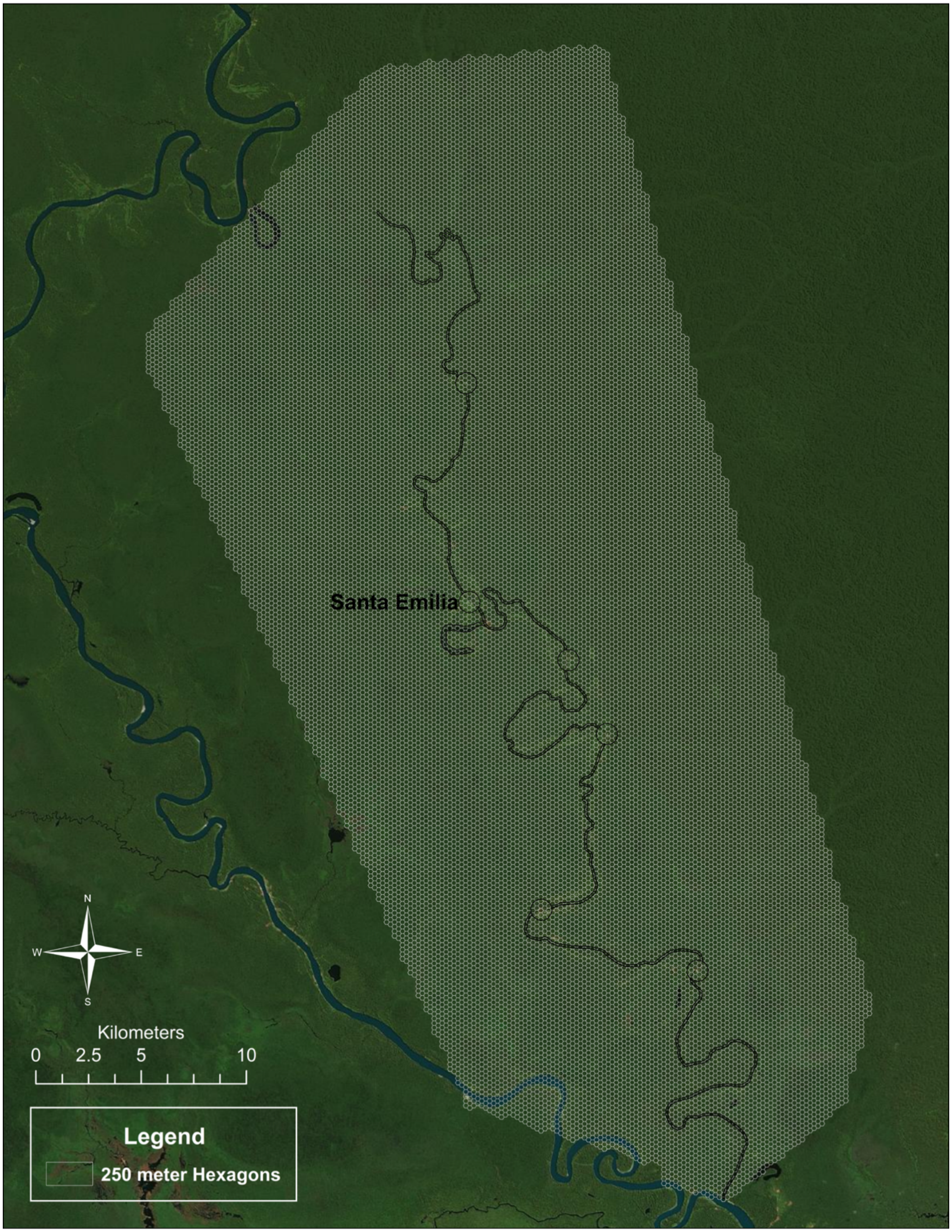
Samuel Roberts
Senior Research Protocol Analyst
This letter has been digitally signed

Emory University

Appendix 3: Hexagon Extent for Study Area Analyses

Number of contiguous 250meter hexagons used in the spatial analyses of the 'Study Area': **57,929**

Appendix 4: Hexagon Extent for Remote Zone Analyses

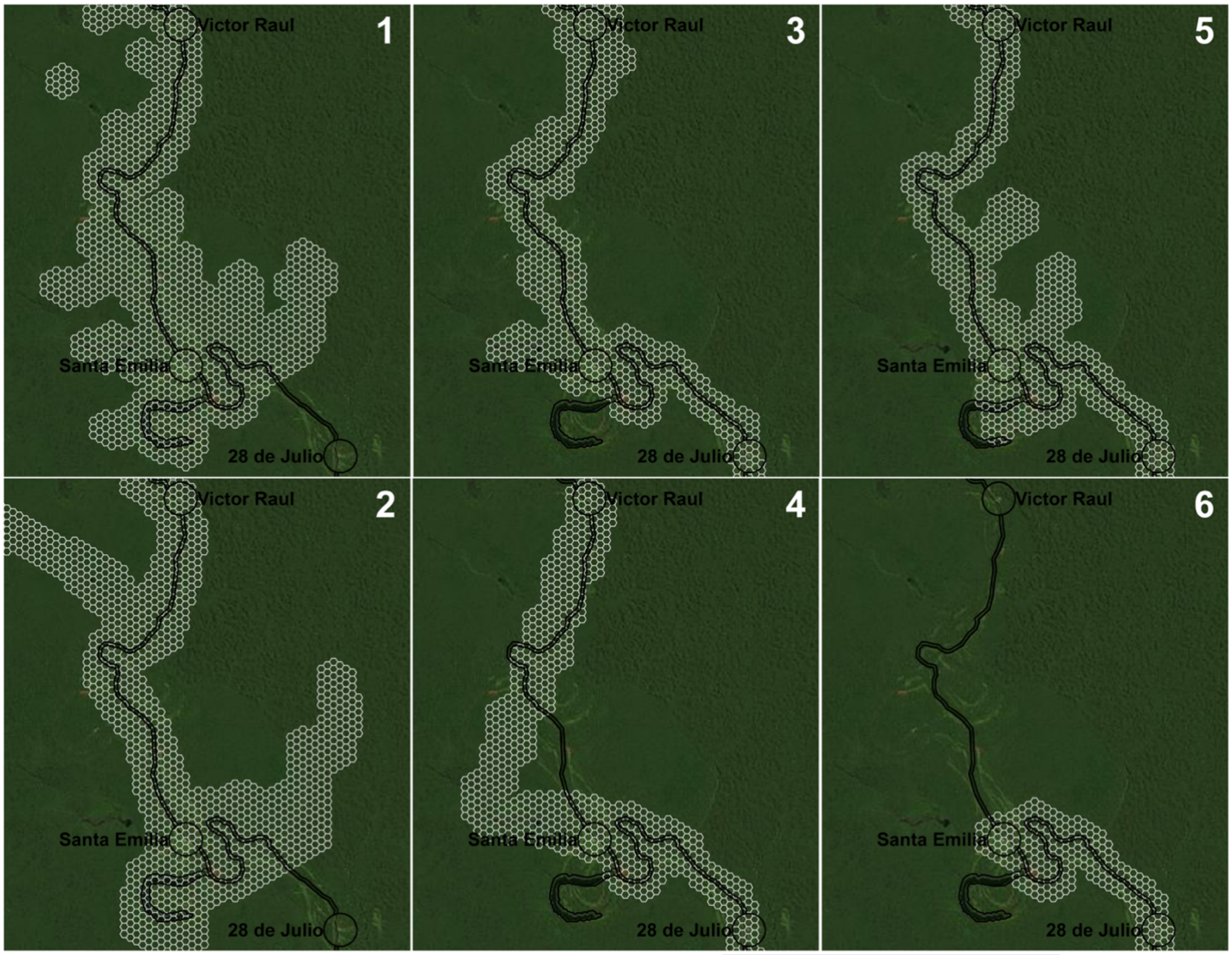


Number of contiguous 250meter hexagons used in the spatial analyses of the 'Remote Zone': **27,421**

Appendix 5: Hexagon Extent for Urban Centers Zone Analyses

Number of contiguous 250meter hexagons used in the spatial analyses of the 'Urban Centers Zone': **17,017**

Appendix 6: Hexagon Extent for Santa Emilia Subzone Analyses



Number of contiguous 250meter hexagons used in the spatial analyses of the 'Santa Emilia Subzone':

1] **1,609**	3] **2,435**	5] **2,517**
2] **1,705**	4] **986**	6] **1,827**