Real-Time Targeted Vector Mosquito Monitoring

Best Practices Guide

Global Mosquito Alert Consortium

2020-02-17

Contents

1 Introduction

The Global Mosquito Alert Consortium's (GMAC's) best practices guides offer experiences gained from a variety of projects that use citizen science to better understand and combat disease-vector mosquitoes. The goal is to create a growing repository of information about how best to use and customize the GMAC's citizen science toolkit for local implementation.

The present guide encompasses Pillar 1 of the GMAC toolkit: real-time targeted vector mosquito monitoring. This is a set of tools that enable citizen scientists to identify and report adult mosquitoes, and that facilitate the subsequent validation and analysis of these reports. As references, this guide draws on several existing projects: Mosquito Alert, iNaturalist, Muggen Radar, and Abuzz.

2 Pillar Objectives

The objectives of this pillar are to:

- facilitate identification and reporting of targeted adult disease-vector mosquitoes by citizen scientists;
- validate and analyze this citizen science data;
- produce real time and reliable scientific data and models for management and risk evaluation.

In addition to these objectives, this pillar includes the cross-cutting objective of

- engaging citizen scientists in the fight against disease vector mosquitoes by helping them to better understand the problem and solutions;
- obtaining feedback from citizen scientists about monitoring and control strategies.

3 Citizen Scientist Roles and Motivations

This pillar involves citizen scientists primarily as collectors and validators of mosquito data. Citizen scientists form a massive network of sensors with the goal of expanding the geographic coverage of monitoring programs without the costs associated with traditional surveillance methods. Citizen scientists participate in data validation by reviewing others' reports, using practical identification tools and relying on both validator proficiency scores (each citizen scientist's proficiency in the validation process) and redundancy (multiple citizen scientists reviewing each report) to reduce errors. Experts are also involved in the validation stage to improve accuracy and provide a point of comparison for generating the citizen scientists' proficiency scores. The analysis is done by experts as well, but mechanisms for better involving citizen scientists as experts will be explored in the future.

Citizen scientists participating in this Pillar are likely to have a variety of motivations depending on local conditions and socio-demographic factors. There has yet to be a systematic study of participant motivations in the existing GMAC projects but a number of conclusions can be drawn from anecdotal evidence coming out of these projects.

The citizen scientists who participate in Mosquito Alert in areas of Spain with high mosquito prevalence appear to be motivated primarily by the annoyance of mosquitoes and a desire to "do something" in response to being bitten. Many are also motivated by concerns about the spread of mosquito-borne diseases, although this is more of an abstract concern in Spain as autochthonous transmission of such diseases has not yet been detected. It is likely that participants in areas of Spain with low mosquito prevalence are motivated more by this latter concern – by the desire to act as sentinels against the arrival of a potentially dangerous species. In addition, it is likely that many participate due to the more traditional citizen science motivations of curiosity and interest in science.

These motivations are likely to vary in other areas of the word and to shift depending on season, local mosquito distribution and biting patterns, changing disease risks, and experiences with the toolkit or with other citizen science projects. Understanding how this variation and shifting occurs, how different motivations may blend into one another is important both for ensuring sufficient participation levels and for correcting sampling bias. In addition, it should be stressed that the primary motivation of annoyance in high-prevalence areas makes mosquito-focused citizen science projects somewhat unique: At least in these areas, the target of the project itself provides participants with a continuous reminder to act.

4   Ethics

Ethical considerations take on heightened relevance when citizen scientists are involved in disease-vector monitoring – and particularly when they are in the vicinity of organisms that could potentially carry dangerous diseases. These concerns fall into two categories: health risks and privacy. In addressing these, autonomy must also be used as a guiding value.

The risk of participants' health being harmed by being bitten by the mosquitoes they are reporting is the most obvious concern in this pillar. Although people in high-prevalence areas are likely exposed to mosquito biting regardless of participation, projects falling within this pillar must be fully transparent about the risks and must be careful not to encourage increased risk-taking. Indeed, projects should ideally help participants reduce their risks, not only by harnessing their reports to provide better control, but also by teaching them how best to avoid bites and remove mosquito breeding sites.

Privacy concerns also take on new dimensions when disease vectors are involved. Participants' locations, while inherent in reporting and important for correcting sampling bias, can reveal a wide variety of information that individuals may prefer to keep private – including disease risk. It is critical that people be fully informed about the location information they are revealing when they consent to participation and that they have control over how much information they share. The Mosquito Alert project, for example, relies on background location tracking to correct sampling bias, but it explains this to participants when they register and it allows them to turn the feature off or on at any time. It is also important to limit the amount of information collected: The Mosquito Alert app, for instance, collects only 5 background locations per day from each participant who has not opted out, and it obscures these locations by placing them in predefined cells of approximately 10 sq. kilometers and sharing only the cell identifiers with the central server. The iNaturalist app collects no background tracking information and it gives participants the option of obscuring report locations on the public webmap.

5   Data Collection

Data collection for this pillar involves, at a minimum, citizen scientists reporting adult mosquitoes as they observe them. This may be as simple as providing information about the date, time, and location of observations. Of greater use, however, are reports that (1) are limited to targeted species, (2) include the observer's taxonomic identification, (3) include the observer's description, through a brief questionnaire, of key features necessary for identification, and (4) are accompanied by a photograph of the mosquito. The last two pieces of information are particularly important, as they allow for subsequent data validation by other citizen scientists or experts. At the same time, not every report needs to include all this information to have value: As long as complete information is included in some reports, they can be used to estimate individual participants' proficiency, which can then serve as a basis for assessing the reliability of their other reports.

The most efficient mechanism for transmitting reports is currently through smart phones or other mobile devices. This is the approach taken by all of the existing projects in this pillar and it allows for near real-time collecting and analysis of data. The applications used for transmitting reports also serve as tools to help participants identify mosquitoes and take protective measures, to validate others' reports, and to view global project information and otherwise interact with project managers. At the same time, other reporting mechanisms, including SMS or even simply paper and regular mail, are clearly possible and might be explored for use in areas with low smartphone penetration.

6   Data Processing and Validation

Data collected by participants must be processed and validated before it can be usefully analyzed. Processing entails, at a minimum, combining reports into a centralized database; ideally it should also include some mechanism for validation and error-correction. Mosquito Alert relies on a central server with a set of

Django/Python-based web applications that handle this process as well as providing front-end portals for interface by various types of participants (general public, expert validators, etc.).

6.1 Participant Error-Checking and Revision

One component related to error-checking is ensuring that participants have the ability to change or even delete reports after sending them. Participants may at times accidentally mark the wrong location or enter some other erroneous information and the system should enable them to make changes. Mosquito Alert does this by allowing multiple versions of each report; all are stored for future reference but only the most recent is used in analysis and dissemination.

6.2 Sampling Effort Collection

Another component is the collection and analysis of information about participants' sampling effort. This information can be used to correct biases resulting from the uneven distribution of sampling activity across space and time. It is important to be able to determine, for example, whether no reports have been received from a particular town because there are no target species there or because there are no participants looking there; conversely, one must determine if a town that has lots of reports has elevated mosquito presence or simply many participants.

In Mosquito Alert, this is done by collecting a small amount of location information from participants. Unless they opt out of this feature, the Mosquito Alert application uses the network and satellite location services on participants' mobile devices to detect their locations 5 times per day. The times are randomly selected independently by each device each day during the hours when targeted species are most likely to be biting. As noted above in the ethics section, the device does not share the detected location itself with the central server, but instead shares only the identifier of the pre-defined sampling cell into which it falls. For computational efficiency (to reduce battery drain on the device), the sampling cell grid is defined simply by evenly spaced latitude and longitude divisions (initially 0.05 degrees each; currently 0.025 degrees each).

One lesson learned by the Mosquito Alert project is that participant location provides only part of the picture in terms of sampling effort. If it also important to know when participants are actually in a position to observe and report targeted mosquitoes. Our experience is that most people install the application on their device and use it briefly but stop interacting with it after a relatively short period of time. There is also large variation among participants in the amount of time they spend using the application. We therefore model what we call "reporting propensity" as a function of time elapsed since installation of the application and intrinsic motivation. We then adjust sampling effort based on the reporting propensity of each participant. The process is complicated by the fact that we do not link background tracking information with reporting information (for privacy reasons). The results, however, have proven to be effective.

6.3 Report Validation

The validation stage is important as a way to directly check that the reported mosquitoes are targeted species and to provide a basis for assessing participants' proficiency. The latter outcome improves the possibilities for making accurate inferences about the reliability of reports from these participants that are not validated: Many participants will send some reports that include a photograph that can be used for validation and others that do not (either no photograph or none in which the specimen can be clearly seen). The expert validation of the first category of report facilitates inferences about the second.

There are a number of different ways of carrying out validation, and the choice among them will depend on local circumstances as well as evolving research on what works best. The primary approach of Mosquito Alert is to rely on a team of 9 entomologists who review reports from citizen scientists through a special expert validation portal. Each report is reviewed independently by 3 of these entomologists. In addition to selecting a category indicating their level of confidence in the report being of a targeted species, the

entomologists are also able to write internal notes and notes to the citizen scientist. They are also able to flag the report for review by the entomology team's leader, who can override any final decision.

Another approach to validation is to rely on other citizen scientists to review photographs. Mosquito Alert also uses this approach, sending each photograph to 30 different citizen scientists using the Crowd Crafting platform (which is accessed through the application directly on through a web browser).

Another project that has had success with this type of crowd-based validation approach is iNaturalist. That platform also allows much more interaction between citizen-science validators and the person making the original report (known as an observation) in a threaded conversation at the record level. Validated records arise from agreement between 'identifiers' leading the record to gain a data quality classification. It also provides an effective mechanism for cultivating citizen scientists to develop expertise in identifying certain specifies and thus improve their validation proficiency.

7 Data Presentation and Use

7.1 End-users and requirements

The pillar and the data generated can be of interest to 4 main types of end-users: the general public, public health managers, educators, and academics. Each targeted audience requires a different mode of data presentation and use.

7.2 Data presentation to public and specific end users

7.2.1 General Public:

- General informative contents on main web and social networks, and also through app notifications (entomology, public health, distribution, prevention measures, etc.). Content should be distributed in a way that maximizes potential reuse while also protecting participants' privacy and respecting their preferences. Where possible, this should be done by placing it in the public domain with the Creative Commons "no rights reserved" mark (CC0). Where attribution is required, the Creative Commons Attribution License (CC BY) is recommended. This can be combined with privacy protections by listing the author as an "anonymous" citizen scientist, as is done for participant photos shared in Mosquito Alert. Another option is to give participants the choice of multiple licenses and let them choose how to list themselves, as is done in iNaturalist.
- Annual reports (freely downloadable from web in pdf): Summary of the project current results and achievements within the year, encompassing science, mosquito management, education and communication.
- Interactive Map: A public map with all the reports and validations, with filters (temporal and spatial) and from which anyone can directly download data in CSV, KML, and other formats. This can be done directly on project websites and/or by publishing the data on GBIG.org, which provides a web map, download options, dataset metrics, a download-DOI, and statistics on downloads and citations. As explained in the first bullet-point above, content should be distributed in a way that maximizes potential reuse while also protecting participants' privacy and respecting their preferences.
- Direct access to data download may also be provided to the public separately from the map (for example, as is done in iNaturalist and Mosquito Alert). As explained in the first bullet-point above, content should be distributed in a way that maximizes potential reuse while also protecting participants' privacy and respecting their preferences.

7.2.2 Public Health Managers:

- "Enrollment Kit" for Managers: An open and downloadable document with all the necessary information for stakeholders surveilling and controlling mosquito populations to use the mosquito alert

platform on their own benefit. From the most simple type of information exploitation to more complex and committed ones.

- Private Portal: A digital platform with a private account where managers can see real-time information by citizens (without even being expert-validated). Data classification is more exhaustive than in the public map. The interface includes a set of temporal and spatial filters to manipulate data, a system to write notes to single or groups of participants, and the possibility to incorporate private geo-located management information (water drain cartography, risk areas, epidemiologically relevant information, etc).
- Interactive Map: A public map with all the reports and validations, with filters (temporal and spatial) and from which anyone can directly download data in CSV, KML and other formats. As explained above in the first bullet point under General Public, content should be distributed in a way that maximizes potential reuse while also protecting participants' privacy and respecting their preferences.

7.2.3   Educators:

- Educational contents in the form of Webquests or other digital formats, for teachers to use in the classroom.
- Leaflets, power points, and other graphically designed materials for teachers to use in the classrooms
- Interactive Map: A public map with all the reports and validations, with filters (temporal and spatial) and from which anyone can directly download data in CSV, KML and other formats. As explained above in the first bullet point under General Public, content should be distributed in a way that maximizes potential reuse while also protecting participants' privacy and respecting their preferences.

7.2.4   Academics:

- Interactive Map: A public map with all the reports and validations, with filters (temporal and spatial) and from which anyone can directly download data in CSV, KML and other formats. As explained above in the first bullet point under General Public, content should be distributed in a way that maximizes potential reuse while also protecting participants' privacy and respecting their preferences.
- Daily data transfers to open access repository like zenodo.
- Make code freely available in a public repository like GitHub under an open source license like the GNU General Public License (GPL) or the MIT license.
- Data sharing through GBIF. (To get started with this, contact your national GBIF node or, in the absence of one, register as a GBIF publisher and find a partner.)

7.3   Data use for vector management and research

There are a variety of options for facilitating data use by vector managers and researchers apart from the mechanisms for presentation and distribution described above. For example, iNaturalist allows anyone to create a project to aggregate any data at any level (for example, some set of vector species at a selected geographic location). Mosquito Alert cooperates directly with vector managers and other public stakeholders through various mechanisms. We draw on these in proposing the following models of cooperation.

Cooperation type 1: I want to use Global Mosquito Alert's data for monitoring and control purposes.

We offer the following options:

- View data: view sightings of targeted mosquitoes and their breeding sites in your territory, filtering by sighting type and date (months and years).
- Share selected data: have you found one or more sightings of interest in your territory? A breeding site you were previously unaware of, for example? Share the information with those involved in monitoring and control activities in your territory (municipal personnel, pest control companies, town or city councils, etc.).

- Export data in a report: export sightings from a map view, with all their details (photo, coordinates, etc.), in the form of a report. Share it with those involved in monitoring and control activities in your territory (municipal personnel, pest control companies, town or city councils, etc.).
- Communicate with citizens via the app: tell us about your territory's monitoring and control activities. We can send app users in your town or city messages of your choosing via the notification system.
- Generate your own maps: create a hashtag (e.g. #LocalCouncil) for your territory's citizens and municipal personnel to include in the "Notes" section when they report sightings. It is possible to use a filter to export all the data corresponding to a hashtag from the public online map in the form of a list.

Cooperation type 2: I want to carry out informative, educational or prevention activities or campaigns to combat targeted mosquitoes, (a) using Global Mosquito Alert's free resources or (b) arranging a face-to-face activity (education, rise awareness action).

We offer you the tools and resources listed below, which will be available on Global Mosquito Alert's website, free of charge, for your own informative, educational or prevention activities or campaigns to combat tiger mosquitoes:

- Informative leaflet in different languages, to distribute online or print*. You can add your institution's logo using the version in .ppt format. If you need to adapt the leaflet further, please contact us.
- Global Mosquito Alert poster in different languages, to distribute online or print.
- Images and drawings of targeted mosquitoes and their breeding sites (see distribution licence in each case). Global Mosquito Alert images and logos.
- Extensive informative content, all up to date and revised by experts in scientific communication, covering entomology, public health, distribution, prevention measures, etc. Unless otherwise indicated, these may be reused under a Creative Commons licence (CC BY).
- Global Mosquito Alert's communication channels are available to you to give your campaign a boost online. Contact us to discuss ways of doing so.
- Web blog for brief communiqués or news items.
- Social media accounts (Twitter, Facebook, Weibo, WeChat).
- App notifications sent directly to the public participants.
- YouTube channel.

Cooperation type 3: I want to establish a different type of cooperation or formalize cooperation through an agreement or protocol.

Formalizing cooperation entails many benefits for both parties. It is a way of establishing a joint plan, specifying each party's rights and obligations, and pursuing optimal results in relation to common and individual goals alike, so that cooperation pays greater dividends. It is also a way for an institution to become an official partner of Global Mosquito Alert and have access to the private managers portal and the app notification system from which managers can freely communicate with citizens their surveillance and control actions.

7.4 Data Licensing

Data should be released with an open access license or placed in the public domain (CC0). The latter option is best for maximizing potential reuse, but there may be situations in which CC-BY (open access with attribution required) is preferable, Restrictions beyond CC-BY are best avoided, although there may be value in giving individual citizen scientists a choice of multiple license options for their individual reports (as is done, for example, by iNaturalist).

8 Data Structures and Repository Links

The Mosquito Alert data is stored in a SQL database (PostgreSQL/PostGIS) that is managed by the Django/Python server layer. It is exported daily to Zenodo at http://doi.org/10.5281/zenodo.597466. The

project has also set up a mechanism to share the data regularly with GBIF (see https://www.gbif.org/es/dataset/1fef1ead-3d02-495e-8ff1-6aeb01123408). It is also exploring the possibility of linking the application to iNaturalist so that participants' observations can be viewed in real time by the iNaturalist community.

The primary unit of observation for the Mosquito Alert data is a report-version. Each time a participant creates a new report or edits or deletes an existing report, a new report-version record is created. Different versions of the same report are linked by a unique report UUID automatically assigned to each report when it is created. Each report-version is linked to the set of photographs that the participant included with the report, and to the participant's answers to the three taxonomic questions designed as a first check of validity. One important point here is that the application sends both the answer to the question and the question, both written as posed to the participant in whatever language the participant was using. Both pieces of information (question and answer in original language) are stored in the database. This makes it possible to deal with updates over time in the language used in the applications and to identify possibly errors or sources of confusion that might be affecting data quality.

The most recent version of each report is sent to the expert validation system (and if new versions come in after validation, they are sent again). The validation results are then linked to the report-version (each report version also has its own unique UUID).

Reports are also linked to participants through a unique reporting UUID assigned by the application to each participant when they register. Data on participants includes registration time, which is used in the sampling effort model.

The background tracking data is stored in its own table, which contains a unique user UUID for background tracking purposes (different from the reporting UUID).

9 Existing Tools

The Mosquito Alert system can be used by anyone in the world and it is possible for specific projects to make use of this directly, tracking their own data through the use of hashtags entered in the participant note section or through other mechanisms. All of the Mosquito Alert code for mobile device applications as well as server-side processes is free and open source (GPLv3), available at https://github.com/MoveLab.

10 Case Studies

Existing projects most relevant to this pillar include:

- Mosquito Alert (http://mosquitoalert.com)
- iNaturalist (http://www.inaturalist.org/) with example projects including (https://www.inaturalist.org/projects/mosquitoes-in-hawaii).