Gmail

Hamzah Hasyim <hamzah.hasyim@gmail.com>

MALJ-D-17-00578

1 message

Murugappan, Magesh, Springer <magesh.murugappan@springer.com> 26 October 2017 at 11:44
To: Hamzah Hasyim <hamzah.hasyim@gmail.com>

Dear Dr. Hasyim,

I have forwarded your e-mail to the editor.

Thank you very much.

With best regards,

Magesh

Magesh Murugappan

Journal Editorial Office

BioMed Central

Web: www.biomedcentral.com

From: Hamzah Hasyim [mailto:hamzah.hasyim@gmail.com]
Sent: Wednesday, October 25, 2017 5:47 PM
To: Murugappan, Magesh, Springer
Subject: Re: Your submission to Malaria Journal - MALJ-D-17-00578

Dear Ms. Catherine Moyes,

Thank you for your email, I will discuss the feedback and nice comments from Reviewer of the journal with our team soon.

Since I quick read the comments, seem the line of comments a little bit different with the last archive article on my laptop. There is a single line increment. So to certain we discuss and revise the same article with the paper that has already feedback from reviewers, would you please send the last article in Ms word? How to do it if I want to download the last article that I had submitted from my account in http://www.editorialmanager.com/malj/default.aspx?

How to do if I want to change a little bit academic tiles one of the authors in authorship. Your prompt attention to this matter is greatly appreciated

Yours sincerely,

Hamzah Hasyim

Ph.D. Candidate in Department of Tropical Medicine and Public Health,

Institute of Occupational, Social and Environmental Medicine, Faculty of Medicine of the Goethe University in Frankfurt am Main

DEUTSCHLAND

http://www.med.uni-frankfurt.de/institut/occupational-medicine/

hamzah.hasyim@stud.uni-frankfurt.de

Phone number: +4915905821418

Occupational Health Safety and Environment Department, Faculty of Public Health, Sriwijaya University, South Sumatra, Palembang-Prabumulih, KM 32 Inderalaya (Ogan Ilir) 30662

INDONESIA

http://fkm.unsri.ac.id/id/

hamzah@fkm.unsri.ac.id

Phone number: +6281373178328

bit.ly/weM38G

On 25 October 2017 at 08:59, Malaria Journal Editorial Office <em@editorialmanager.com> wrote:

MALJ-D-17-00578
Spatial Modelling of Malaria Incidence in South Sumatra, Indonesia
Hamzah Hasyim, PhD candidate; Afi Nursafingi, M.Sc; Ubydul Haque, PhD; Meredian Alam, PhD candidate; Doreen Montag, DPhil; David Groneberg, Prof,Dr,PhD; Meghnath Dhimal, PhD; Ulrich Kuch, Dr; Ruth Müller, Dr
Malaria Journal

Dear Mr Hasyim,

Your manuscript "Spatial Modelling of Malaria Incidence in South Sumatra, Indonesia" (MALJ-D-17-00578) has been assessed by our reviewers. Although it is of interest, we are unable to consider it for publication in its current form. The reviewers have raised a number of points which we believe would improve the manuscript and may allow a revised version to be published in Malaria Journal.

Their reports, together with any other comments, are below and in an attachment provided by reviewer #2. Please also take a moment to check our website at http://malj.edmgr.com/ for any additional comments that were saved as attachments.

If you are able to fully address these points, we would encourage you to submit a revised manuscript to Malaria Journal. Once you have made the necessary corrections, please submit online by log onto the journal's website.

Your username is: Hamzah

If you forgot your password, you can click the 'Send Login Details' link on the EM Login page at http://malj.edmgr.com/.

Please include a point-by-point response within the 'Response to Reviewers' box in the submission system and highlight (with 'tracked changes'/coloured/underlines/highlighted text) all changes made when revising the manuscript. Please ensure you describe additional experiments that were carried out and include a detailed rebuttal of any criticisms or requested revisions that you disagreed with. Please also ensure that your revised manuscript conforms to the journal style, which can be found in the Submission Guidelines on the journal homepage.

Please note, if your manuscript is accepted you will not be able to make any changes to the authors, or order of authors, of your manuscript once the editor has accepted your manuscript for publication. If you wish to make any changes to authorship before you resubmit your revisions, please reply to this email and ask for a 'Request for change in authorship' form which should be completed by all authors (including those to be removed) and returned to this email address. Please ensure that any changes in authorship fulfil the criteria for authorship as outlined in BioMed Central's editorial policies (http://www.biomedcentral.com/about/editorialpolicies#authorship).

Once you have completed and returned the form, your request will be considered and you will be advised whether the requested changes will be allowed.
By resubmitting your manuscript you confirm that all author details on the revised version are correct, that all authors have agreed to authorship and order of authorship for this manuscript and that all authors have the appropriate permissions and rights to the reported data.

Please be aware that we may investigate, or ask your institute to investigate, any unauthorised attempts to change authorship or discrepancies in authorship between the submitted and revised versions of your manuscript.

The due date for submitting the revised version of your article is 22 Nov 2017.

I look forward to receiving your revised manuscript soon.

Best wishes,

Catherine Moyes
Malaria Journal
https://malariajournal.biomedcentral.com/

Reviewer reports:

Reviewer #1: MALJ-D-17-00578
Spatial modelling of malaria incidence in South Sumatra, Indonesia

General comments
Study description:
In South Sumatra in Indonesia, annual malaria in 2013 and environment relationship have been tested with two methods, a global linear regression, Ordinary least square methods and a geographically weighted regression.
This study is very interesting and the results maps are well done. But some details are missing and the interpretation could be improved.
Here are some major indications followed by minor suggested corrections.
Major:
1. The spatial analysis methods are not clearly explained and are sometimes a little bit confused.
a. OLS is a global linear regression method. It is often run before the GWR to select explanatory variables. Variables could be added through stepwise method. Each variable's contribution is validated if the AIC and AICc decrease (for at least 3 values). Moran'I of OLS residuals tests for autocorrelation. A residual map should show random distribution. If not, there is another spatial variable to add. The variance inflation factor (VIF) looks for explanatory variables redundancy. The disadvantage of this OLS method is that it doesn't consider spatial correlation and non-stationarity of data.
b. Then a GWR can be run. GWR is a modified regression model and calculate a local specific variance for each coordinate point. This a local regression that has the advantage to highlight local relationships between the dependant variable and the explanatory variables by addition of weighted parameters. These weights are automatically determined

for each location and can be mapped. The total regression for each location can also be mapped to identify the higher and lower regression coefficient. Then a map of residual should identify where other variable may be required.
c. Here you chose a kernel fixed type with Cross validation (CV) as a bandwidth method. The bandwidth controls the degree of smoothing in the model and identify an optimal fixed distance
d. Comparison between the two methods can be done, (even if they are not exactly the same analysis). If the regression coefficient are better and the AIC are lower, it is concluded that this method is the best.
e. The objective could be to predict but here it seems that it is very interesting to identify malaria-environment relation and according to the location.
f. A validated OLS can lead to a global policy and a validated relationship with GWR is more appropriate to lead to local policy.
g. In the case of this study, it an excellent spatial analysis to identify which parameter to look closer and where and how much it varies and where it would be more appropriate to do so and for example do.

2. The validation is not clear
3. Interpretation on the results could and be improved and better put in context.
a. Context with links with transmission and actually specific known ecological preferences of some Anopheles species (Ex: Anopheles found in forest for villages where distance to forest is a factor)
b. Context with other studies, it is better to refer to study with similar environment, latitude, health system
c. Discuss more the difference between districts, especially those with very high or very low local R2
d. Perspectives: not only add more parameters but explore at a finer time scale the relationship with parameters that vary intra-annually vary.
Minor:
* Title: add environment in the title. It could be "Spatial modelling of malaria incidence relationship with environment factors" or something like that
* Keywords: malaria, geographically weighted regression, GWR, Ordinary least squares regression, OLS, Sumatra, rainfall, elevation, distance to water
* It should be appreciated to name the main Anopheles vector species for each type of environment or district
* Cases number or incidence?
* Which georeference system is used in which units (meters or degrees)
* Maps 6 and 7 : Add units please.
* Which is the scale or resolution in time and space for each parameter?
* Forest:
o How old is the forest layer? Which year?
o Can we guess that some parts have been deforested since the forest cover has been recorded? Do we have information on the percentage of deforestation between this year and 2013?
o In the discussion links between your result and what you say about deforestation
* Rainfall: the rainfall-malaria relationship is probably a non linear relationship as it is written in the discussion. In this study annual rainfall is used. Is it average rainfall or total amount?
* Temperature
* Elevation: often described as an indirect factor: less humidity, lower temperature or suitable for different Anopheles species
* Results
o Present only the result without assuming cause between the variables.
* Interpretation
o Explain links with field data and known information
* References:
o Rainfall and malaria : you could add Botswana and Ethiopia works
Comments by Line
Background
* Along the background, when a reference is cited to state a link between malaria transmission and an environment factor, it should be better to mention in which country or environment type.
o Example: line 39: "lowland location", its depends where
o Line 48 "it proliferates faster under higher temperatures", it depends where.
* Line 63: very important to know which variables you have studied, please list them here.
* "performance of the OLS and GWR models in predicting.."

Methods
Study area
* Line 77 : a range of altitude would be appreciated, highest altitude for all the area or for each district
* Line 78: is it monthly rainfall amount by station?

Study population and data collection
* Lines 85-86: How many PHC? Just to have an idea of the density by district (or by population or by area)
* Line 92: 36 372 patients or presumed positive malaria cases? Some patients may come several times a year.
* Line 94 % (3578/36372 is around 10%

*   Line 97: precise which sort of villages or number (436 villages)
Preparation of spatial data
Data acquisition and selection
*   Line 106: How many stations? How many km are they close to each other? To have an idea of their density
Data preprocessing
*   Line 112 DEM which spatial resolution? Issued from which satellite data type?
*   Line 115 Which spatial interpolation method did you used for rainfall? Which classification and from which criteria did you use it?
*   Line 120: VIF? It should be useful here to describe the variance inflation factor, what is it and how it works.

Data processing
*   See major comment above
*   Miss comparison between the two methods OLS and GWR
*   Miss validation

Results
Environmental factors influencing malaria incidence at village level: local GWR model
*   Line 194: related to size of weights
*   Line 206: "..show that the environmental factors prevailing In these regions are less suitable for explaining the variance of malaria incidence in this area" need to explain why please.
Comparison between OLS and GWR: cf major
Discussion:
Cf major.
*   Line 289: Avoid "spatial epidemiology microscope"
*   Line 298: "The approach arbitrarily plots all of the cases in the settlements" I don't understand what you mean.
*   Line 305 - 311: Add seasonality studies, non linear relationship, time downscaling (to monthly rather than annual cases), etc.

Maps:
Figure 1 and 3: you have to choose the same methods for all the maps to code the districts, numbers or abbreviations.
Figure 3: scale and North is missing.
Figure 4: Legend (spatial representation map showing.. not needed) Each explanatory variable
Figure 5: Is multicolinnearity test applied also with the response variable?
Figure 6: reformulate the legend please. It should be something like " predicted value from GWR" .

Figure 7: ….. Significance percentage value for each explanatory variable by village location
Figure 8: Local regression coefficient (R2) from GWR method by village location

Hope that could help. Good luck.
Regards.

Reviewer #2: This manuscript applies spatial analysis to malaria data in a low-endemic and heterogeneous area. By focusing the analysis on routinely reported data as well as using spatial covariates accessible within the country, this provides an approach that is accessible to malaria programs within the country. Overall, this manuscript is well written and provides useful information to help better understand malaria epidemiology in this area. However, before recommending for publication, I have several comments that should be considered.

Reviewer #3: This manuscript presented an analysis of routine malaria surveillance data for 2013 to examine the spatial patterns of malaria in South Sumatra, Indonesia. Ordinary least squares and geographically weighted regression analyses were used to examine the potential role of environmental risk factors on the spatial patterns of malaria incidence. Findings indicated that rainfall and distance from the forest played a role in explaining the malaria incidence. While the paper contains results that could be of interest, major revisions are necessary in the language. The paper was not focused and included too much extraneous information, yet did not include important information with regard to the methods. There were also several concerns with the methods and interpretation of the findings.

1.   Abstract: From a statistical perspective, it is unclear how "having an R-squared value of 60%" indicates "that almost all independent variables were significant at certain locations at the village level."

2. Abstract: The conclusions do not match the stated aim of the paper and instead highlight the merits of methodological approach instead of how the findings "help in the development of local policies for malaria elimination" in South Sumantra.

3. Background: This section needs to be more concise and relevant to the study conducted and aims addressed. For instance, the authors exhaustively discuss the role of several variables (migration, population density, temperature, etc), none of which are considered in the present study. The authors need to focus on outlining the wider context, gaps in knowledge/evidence and then introduce the present research and how it addresses those gaps.

4. Methods: How many primary health centers reported malaria case data? And what is the level of completeness of this data? How does the malaria case data from the primary health centers become village level data? Was the analysis at the village or health facility level?

5. Methods: Authors state that "In the study region and period, 2,787,954 of the total population and 36,372 research participants visited hospitals or PHCs due to suspected malaria fever". Elsewhere, authors state "The study population was the number of participants who were suspected of having malaria while the sample was the number of participants with laboratory confirmed malaria." It is unclear what the authors mean by study population, sample, research participants, and total population.

6. Methods: Were multiple episodes from the same individual included? Or was the analysis based on single malaria episodes? As there can be potential biases from relapses especially from P.vivax.

7. Line 96: The authors discuss locations of cases in each district. Is this the location of the primary health center they sought care, or the location of their residence?

8. Methods: Authors included several distance variables - it is unclear whether these are distances from village of residence to the attribute of interest (river, forest, etc) or distances from primary health center.

9. Methods: Was any validation of the OLS or GWR models conducted. For example, cross validation or bootstrapping? And what was the impact on the results?

10. Methods: it is unclear how the outcome malaria incidence was defined as there was no mention of village size or population, and also unclear whether this was at the primary health center level or the village level?

11. Methods: one requirement for an OLS is that only statistically significant explanatory variables are included. However it seems that the OLS model used by authors included several variables that were not significant.

12. Methods: lines 123 - 169 go into an exhaustive explanation of the GWR and OLS approaches, while some detail is important, this much information seems to shift the focus of the paper to one on methodological approaches and distracts from the stated aim to "use global and local spatial modelling to analyse
the environmental risk factors for malaria in South Sumatra that vary geographically at the regional level."

13. Methods: why were other variables such as village size/health facility catchment area size, household density, distance to health facility, coverage of malaria interventions included? Also was seasonality accounted for?

14. Results: Lines 172 - 175, where is the incidence data presented? And it is still unclear what incidence refers to? Is the number of cases what is being referred to as incidence?

15. Table 1: From a statistical perspective, the OLS model should only include variables with significant coefficients and that are in the expected direction.

16. Table 1: Please provide units and scale the variables appropriately so the results are interpretable. For instance, distance from forest has a coefficient of 0.00 which cannot be interpreted.

17. Results: Lines 186 - 189: authors conclude that malaria incidence is more common in regions with high rainfall and areas adjacent to forest areas. However looking at the coefficients presented in Table 1, distance from forest area has a positive coefficient, meaning that as distance from forest area increases malaria incidence increases. Please clarify.

18. Lines 200 - 201: Authors state "The regression coefficients for malaria incidence at the local level range from 0.03 to 0.99 (Fig.8)." However, Fig 8 presents the R2 values which is different from the regression coefficients.

19. Lines 201-202: Authors state "The highest influence of environmental factors on malaria incidences was found in Lahat District." It is not clear where this conclusion came from especially considering Figure 8.

20. Lines 202 - 207. Authors have erroneously interpreted R2 values as values of regression coefficients.

21. Discussion: Authors state that their "analyses have identified Lahat as the South Sumatran district in which environmental factors were of greatest relevance for malaria incidence." Caution is needed in making such conclusions especially given that the small village level sample sizes (Fig 3). Inability to detect significant relationships may in fact be related to the small sample sizes.

22. Discussion: much of the discussion is very anecdotal and not directly related to the findings presented. For instance the authors discuss relevance of deforestation and distance to coal mines, none of which was assessed in the present study.

23. Discussion: authors should avoid introducing new data in the discussion. For instance, authors discuss distance between coal mines and local plantations and forests in Lahat District (lines 254-255). Elsewhere authors state "temperature was correlated with altitude and humidity…".

There is additional documentation related to this decision letter. To access the file(s), please click the link below. You may also login to the system and click the 'View Attachments' link in the Action column.
http://malj.edmgr.com/l.asp?i=56683&l=Y28W4LJN

If improvements to your figures have been requested or are needed, and you would like professional help, we can recommend our affiliates Peerwith for help with figure editing (https://bmc.peerwith.com/malj/figure-editing). Please note that use of any Peerwith service is neither a requirement nor a guarantee of publication.

Gmail

Hamzah Hasyim <hamzah.hasyim@gmail.com>

MALJ-D-17-00578

1 message

Murugappan, Magesh, Springer <magesh.murugappan@springer.com> 12 September 2017 at 15:21
To: Hamzah Hasyim <hamzah.hasyim@gmail.com>

Dear Dr. Hasyim,

Thank you very much for your e-mail.

Your paper looks fine and the same has been assigned to the editor.

With best regards,

Magesh

Magesh Murugappan

Journal Editorial Office

BioMed Central

Web: www.biomedcentral.com

From: Hamzah Hasyim <hamzah.hasyim@gmail.com>
Sent: Monday, September 11, 2017 2:54 PM
To: Murugappan, Magesh, Springer
Subject: Re: MALJ-D-17-00578 - Manuscript Sent Back

Dear

Magesh Murugappan

JEO Assistant,

Based on a message below, I would like to ask you, regarding the e-mail address of all authors on the title page. The message does mean the e-mail address should be rewritten directly under the Affiliations and contact data in list authors in the main paper. Isn't it?

Research article

Title:

Spatial Modelling of Malaria Incidence in South Sumatra, Indonesia

Authors:

Hasyim, H.[1,2], Nursafingi, A.[3], Haque, U.[4], Alam, M.[5], Montag, D.[6], Groneberg, D.A.[1], Dhimal, M.[1,7], Kuch, U.[1], Müller, R.[1]

Affiliations and contact data:

[1]Department of Tropical Medicine and Public Health, Institute for Occupational Medicine, Social Medicine and Environmental Medicine, Faculty of Medicine, Goethe University, Frankfurt am Main, Germany

[2]Faculty of Public Health, Sriwijaya University, Indralaya, South Sumatra, Indonesia

3Remote Sensing Program, Faculty of Geography, Gadjah Mada University, Yogyakarta, Indonesia

[4]Department of Public Health, Baldwin Wallace University, Berea, Ohio, USA

[5]Barts and the London School of Medicine, Centre for Primary Care and Public Health, Queen Mary University of London, London, UK

[6]School of Humanities and Social Sciences, University of Newcastle, Callaghan, NSW, Australia

[7]Nepal Health Research Council (NHRC), Ramshah Path, Kathmandu, Nepal

§ Corresponding author: hamzah.hasyim@stud.uni-frankfurt.de, hamzah@fkm.unsri.ac.id[1, 2]

afinursafingi@gmail.com[3], ubydul.kth@gmail.com[4], meredian.alam@uon.edu.au5,

d.montag@qmul.ac.uk[6], groneberg@med.uni-frankfurt.de1, meghdhimal@gmail.com[1, 7], kuch@med.uni-frankfurt.de[1], Ruth.Mueller@med.uni-frankfurt.de[1].

Thank You.

Warmest regards,

Hamzah Hasyim,

bit.ly/weM38G

Bitte denken Sie an die Umwelt, bevor Sie diese e-Mail ausdrucken

Please consider the environment before printing this e-mail

On 11 September 2017 at 08:50, Malaria Journal Editorial Office <em@editorialmanager.com> wrote:

MALJ-D-17-00578
Spatial Modelling of Malaria Incidence in South Sumatra, Indonesia
Hamzah Hasyim, PhD candidate; Afi Nursafingi, M.Sc; Ubydul Haque, PhD; Meredian Alam, PhD candidate; Doreen Montag, DPhil; David Groneberg, Prof,Dr,PhD; Meghnath Dhimal, PhD; Ulrich Kuch, Dr; Ruth Müller, Dr
Malaria Journal

Dear Mr Hasyim,

Your submission entitled "Spatial Modelling of Malaria Incidence in South Sumatra, Indonesia" has been received.

Before we can further process it you are kindly requested to make the following corrections to meet the journal's requirements (please also refer to the Submission Guidelines):

http://malariajournal.biomedcentral.com/submission-guidelines

Please include the e-mail address of all authors in the title page.

Please log onto Editorial Manager as an author.

Your username is: Hamzah

If you forgot your password, you can click the 'Send Login Details' link on the EM Login page at http://malj.edmgr.com/.

Please go to the menu item 'Submissions Sent Back to Author', and click on 'Edit Submission'. If no changes are to be made in the metadata, please go immediately to the last submission step 'attach files', and replace the appropriate files. Build the PDF, view your submission, and approve the changes.

Thank you for submitting your work to this journal.

With kind regards,

Magesh Murugappan
JEO Assistant

MALJ-D-17-00578

2 messages

Murugappan, Magesh, Springer <magesh.murugappan@springer.com> 22 November 2017 at 11:23
To: Hamzah Hasyim <hamzah.hasyim@gmail.com>

Dear Dr. Hasyim,

Please find below my responses for your questions. Hope this helps you.

1. You can include the revised title of your paper.
2. Whenever there is a change(addition/removal) of authors in the list, you should provide an authorship form signed by all the authors accepting to this change including the author who have been removed from the list.
3. You can upload the revised abstract here.
4. Please include a point by point response for all the reviewers here. You could also include a separate single file for response to reviewers, at your discretion.
5. You will be able to upload the new revised files in the system. Please click next in the attach files step when you see the previous version of your paper.
6. You can include a cover letter for this paper if you would like to provide any.

Looking forward to receiving your revised paper online.

Thank you very much.

With best regards,

Magesh

Magesh Murugappan

Journal Editorial Office

BioMed Central

Web: www.biomedcentral.com

From: Hamzah Hasyim [mailto:hamzah.hasyim@gmail.com]
Sent: Tuesday, November 21, 2017 11:44 PM
To: Murugappan, Magesh, Springer
Cc: Mr h|a|m|z|a|h
Subject: Re: Reminder: your revision for Malaria Journal is due soon - MALJ-D-17-00578

Dear

Editorial Office
Malaria Journal

Thank you for your friendly reminder. I want to submit revision manuscript "MALJ-D-17-00578R1".

However, due this is the first time submission my paper in malaria journal so, I would like to ask you regarding "revised submission", at the following link

http://www.editorialmanager.com/malj/Default.aspx

I saw the stages of Revised Submission, Should I fill one by one the step below?

Select Article Type

Enter Title

If I want to change the title of manuscript, according to with advising one of the anonymous reviewers

the old title

Spatial Modelling of Malaria Incidence in South Sumatra, Indonesia

the new title

Spatial Modelling of Malaria Incidence with environmental factors in South Sumatra, Indonesia

Current Author List

Save these changes to my user registration as well. What is the function if I remark this option?

Add/Edit/Remove Authors

If I want to change the authorship because one of the co-authors withdraws in authorship

Funding Information

Submit Abstract

Its mean I submit a new abstract in this windows. Isn't it?

Respond to Reviewers

where I send a file to giving feedback regarding the comments of three anonymous reviewers, in this stage or the next steps?

Should I split in three file response to the reviewer #1, #2, and #3, in this attachment file?

Attach Files

If I want to attach all new file, both revision paper and figures, I must delete the old file or not?

Respond to Reviewers is Required for Submission.

Please give your response to specific reviewer and editor comments in the box below. To see the comments, click the "View Decision Letter" link. You may select and copy the comments from there, and paste into the box below.

Should I make a new cover letter for this revision, or only send submission revision manuscript and figures?

Thank you, I sincerely appreciate your time and consideration, to answer my questions above.

Sincerely.

Hamzah Hasyim,

On 19 November 2017 at 09:56, Malaria Journal Editorial Office <em@editorialmanager.com> wrote:

MALJ-D-17-00578
Spatial Modelling of Malaria Incidence in South Sumatra, Indonesia
Hamzah Hasyim, PhD candidate; Afi Nursafingi, M.Sc; Ubydul Haque, PhD; Meredian Alam, PhD candidate; Doreen Montag, DPhil; David Groneberg, Prof,Dr,PhD; Meghnath Dhimal, PhD; Ulrich Kuch, Dr; Ruth Müller, Dr
Malaria Journal

Dear Mr Hasyim,

When checking our records, we noticed that the revised version of your manuscript MALJ-D-17-00578 is due soon on 22 Nov 2017.

If you are ready to submit, please access the manuscript by log onto the journal's website.

Your username is: Hamzah

If you forgot your password, you can click the 'Send Login Details' link on the EM Login page at http://MALJ.edmgr.com/.

We are looking forward to receiving your revision.

Best wishes,

Editorial Office
Malaria Journal
https://malariajournal.biomedcentral.com/

Hamzah Hasyim <hamzah.hasyim@gmail.com> 22 November 2017 at 13:43
To: Fadhilah Eka Maharani <fadhilah.em94@gmail.com>

Dear Dr

Please allow me to submit our final paper below to the malaria journal today due to the deadline. I have communicated with Journal Editorial Office regarding the submission and one of the point.

"Whenever there is a change (addition/removal) of authors in the list, you should provide an authorship form signed by all the authors accepting to this change including the author who has been removed from the list"

Is the official Official Statement - withdrawal of authorship from him, including to submission in revision paper?

I have friendly invite Meredian in our paper, however, I don't know why he is didn't replied my email. Please advise

Please forgive me for any inconvenience this may have caused

Sincerely,

Hamzah
[Quoted text hidden]

Gmail

Hamzah Hasyim <hamzah.hasyim@gmail.com>

MALJ-D-17-00578
35 messages

Hamzah Hasyim <hamzah.hasyim@gmail.com> 24 November 2017 at 14:50
To: Malaria Journal Editorial Office <magesh.murugappan@springer.com>

Dear Dr Magesh,

Malaria Journal Editorial Office
BioMed Central

Concerning our revision manuscript "MALJ-D-17-00578R1". In the last email, you send me information whenever there is a change (addition/removal) of authors in the list, I should provide an authorship form signed by all the authors accepting to this change including the author who has been removed from the list.

Meanwhile, our manuscript still reviewed by the referees of the journal, would you please, give me standard or sample an authorship form, if one of co-author in this current article will be removed from the list of authorship?

In another hand, the affiliation institutions of the co-authors in this current article come from some different countries, how to get their signature? Is possible by digital signature or is it enough If i forward "Official Statement - withdrawal of authorship" signed by the co-author directly who want to ask for revocation his name of authorship? May I submisson this form after the revision of manuscript declared accepted for published?

Hopefully, my revision will be accepted and published in Malaria Journal soon.

Thank you for your kind assistance in advance. I really appreciate your help.

Yours Sincerely,

Hamzah Hasyim,
PhD Candidate in Department of Tropical Medicine and Public Health,
Institute of Occupational, Social and Environmental Medicine, Faculty of Medicine of the Goethe University in Frankfurt am Main
DEUTSCHLAND
http://www.med.uni-frankfurt.de/institut/occupational-medicine/
hamzah.hasyim@stud.uni-frankfurt.de
Phone number: +4915905821418

Occupational Health Safety and Environment Department, Faculty of Public Health, Sriwijaya University, South Sumatra,
Palembang-Prabumulih, KM 32 Inderalaya (Ogan Ilir) 30662
INDONESIA
http://fkm.unsri.ac.id/id/
hamzah@fkm.unsri.ac.id
Phone number: +6281373178328

bit.ly/weM38G
Bitte denken Sie an die Umwelt, bevor Sie diese e-Mail ausdrucken
Please consider the environment before printing this e-mail

Murugappan, Magesh, Springer <magesh.murugappan@springer.com> 27 November 2017 at 10:25
To: Hamzah Hasyim <hamzah.hasyim@gmail.com>

Dear Dr. Hasyim,

You find the authorship form at, http://www.biomedcentral.com/submissions/editorial-policies#authorship. Please ensure that all authors, including those removed, sign this form. If any parts of the form are incorrectly completed then we will send this back to you for correction. The responses of all authors should be collated in one form.

Once you have completed the form, please send it back to us via email. Please be aware that we may ask your institution(s) to investigate any changes to authorship or discrepancies in authorship between the submitted, revised and accepted versions of your manuscript. In the event of an authorship dispute, the journal will refer the matter to your institution(s) to investigate.

Thank you very much.

With best regards,

Magesh

Magesh Murugappan

Journal Editorial Office

BioMed Central

Web: www.biomedcentral.com

[Quoted text hidden]

Hamzah Hasyim <hamzah.hasyim@gmail.com> 29 November 2017 at 01:49
To: "Dr. Ruth Müller" <Ruth.Mueller@med.uni-frankfurt.de>, "Dr. Ulrich Kuch" <kuch@med.uni-frankfurt.de>

Dear Dr Ruth and Dr Ulrich,

I would like to inform you, regarding change of authorship. I have obtained information from Journal Editorial Office BioMed Central below and I have filled the document concerning "change authorship request" form, kindly find attached.

I will send this document to all co-author, to request to them give a digital signature in image format GIF, PNG, JPEG, or TIFF. Is correct? Please advise, before I circulate this document to the others. I would like to express my sincere appreciation for your kind advice. Thank you for guiding me.

Sincerely yours,

Hamzah Hasyim

Note: I will make a letter below
Dear co-author,

I would like to express my sincere appreciation for your valuable contribution to our publication. One more stage to finalise this submission which needs your signature to finalise the authorship.

Kindly find attached the requested document, please sign and please send it back to me via email or please give you're a digital signature in image format GIF, PNG, JPEG, or TIFF and the signatures will collate in one authorship form

Again, thank you for your kind cooperation.

Sincerely,

Hamzah Hasyim
[Quoted text hidden]

v4.docx
34K

Dr. Ruth Müller <ruth.mueller@med.uni-frankfurt.de> 29 November 2017 at 02:32
Reply-To: ruth.mueller@med.uni-frankfurt.de
To: Hamzah Hasyim <hamzah.hasyim@gmail.com>, "Dr. Ulrich Kuch" <kuch@med.uni-frankfurt.de>

Dear Hamzah,

for me the form is fine. Please find attached my signature.

You can kindly ask from the coauthors a scan of signed form or a digital signature (which you copy in the form). Please give the authors their choice.

When you write to all authors, please mention the exact deadline (14 days as mentioned in the form translates to which date???).

Best regards,

Ruth

[Quoted text hidden]

GWR-paper_Authorship-change_signed-by_RM.docx
153K

Hamzah Hasyim <hamzah.hasyim@gmail.com> 29 November 2017 at 03:38
To: "Dr. Ruth Müller" <ruth.mueller@med.uni-frankfurt.de>
Cc: "Dr. Ulrich Kuch" <kuch@med.uni-frankfurt.de>

Dear Dr Ruth

Thank you so much for your prompt reply, I really appreciate it. I obtained the information concerning "change authorship request", from Journal Editorial Office BioMed Central on 27 Nov 2017, 1 day ago. It meant the exact deadline on Monday, on 11 December 2017 later. Is correct, Particularly for others co-authors, I will write a letter to them three days before the deadline on Friday, 8 December 2018. Isn't it?

Best regards

Hamzah Hasyim

Note: This my draft letter below

Dear co-author,

I would like to express my sincere appreciation for your valuable contribution to our publication. One more stage to finalise this submission which needs your signature to finalise the authorship.

Kindly find attached the requested document, please sign. Kindly choose a scan of signed form or a digital signature in image format GIF, PNG, JPEG, or TIFF (which I copy in the form). The signatures will collate in one authorship form.

Author's Response To Reviewer Comments

Close

Response to the Reviewer #1: MALJ-D-17-00578
Comments#: MALJ-D-17-00578.
Manuscript Title: Spatial Modelling of confirmed malaria cases in South Sumatra, Indonesia
Reviewer reports:
Reviewer #1: MALJ-D-17-00578
General comments
Study description :

In South Sumatra in Indonesia, annual malaria in 2013 and environment relationship have been tested with two methods, a global linear regression, Ordinary least square methods and a geographically weighted regression. This study is fascinating, and the results maps are well done. But some details are missing, and the interpretation could be improved.
Response: I highly appreciate your positive advice to improve our manuscript MALJ-D-17-00578 with revised title "Spatial Modelling of confirmed malaria cases in South Sumatra, Indonesia". Thank you so much for your constructive comments. Each comment has been carefully considered and responded point by point. Responses to the reviewer are made in italics.
Here are some major indications followed by minor suggested corrections.
Major comments:

1. The spatial analysis methods are not explained and are sometimes a little bit confused. Then a GWR can be run. GWR is a modified regression model and calculates a local specific variance for each coordinate point. This a local regression that has the advantage to highlight local relationships between the dependent variable and the explanatory variables by addition of weighted parameters. These weights are automatically determined for each location and can be mapped. The total regression for each location can also be mapped to identify the higher and lower regression coefficient. Then a map of residual should identify where another variable may be required.
Response: Thank you for your feedback. We revised the description of spatial analysis GWR was used to model predictive confirmed malaria cases based on a specific geographic area (geographical coordinates) by obtaining different regression coefficients for each location in the study area [1].

2. Here you chose a kernel fixed type with Cross-validation (CV) as a bandwidth method. The bandwidth controls the degree of smoothing in the model and identify an optimal fixed distance.
Response: The optimum distance threshold (also known as the bandwidth) or the optimum number of neighbours can be determined in two ways: by minimising the square of the residuals by cross-validation (CV) or by minimising the Akaike Information Criterion (AIC) [2]. In our study, we select the type of weighing (kernel type) and optimum bandwidth selection method based on selection criteria. In our case, we use AIC. Classic AIC tends to choose smaller bandwidths by which geographically varying coefficients are likely to be undersmoothed. CV is applicable only to Gaussian models [3]. Comparison between the two methods can be made, (even if they are not the same analysis). If the better regression coefficient is better and the AIC is lower, it is concluded that this method is the best.

Response: The best GWR model which used weighting function is 'Fixed' (Gaussian) fixed with the bandwidth selection method "Golden section search". Then we use AIC. It is a statistical measure, which quantifies the relative goodness-of-fit of various derived statistical models from a given sample dataset. The preferred model is that with the lowest AIC value. Geographically weighted regression (GWR) is the regression model that has been developed for data modelling with continuous response variable and considering the spatial or location aspect. The best bandwidth can be seen in the output table entitled bandwidth title and geographic ranges. We conclude the best bandwidth and criteria model goodness as stated in table below.

Table 1. GWR result using 'Fixed' (Gaussian)
Bandwidth and geographic ranges
Bandwidth size: 9184.47
Diagnostic information
Residual sum of squares: 33549.28
Classic AIC: 3482.17
BIC/MDL: 4198.30
CV: 178.92
R square: 0.687
Adjusted R square: 0.409

The best bandwidth generated 9184 neighbours that have significant spatial relationships with a region. In addition, we demonstrated the best model selection by the value of the residual sum of square, classic AIC, and the R square, like in table 2. The smaller the AIC, the better the model performed. Further, the AIC considers the simplicity of the established model. In addition, the better the model is created if the value of R2 increases. In the table below, we can see if the GWR model is better than OLS model.

Table 2. Comparison of GWR and OLS models by value RSS, Classic AIC, and R2
Value OLS GWR
Residual sum of square 100,625.26 33549.28
Classic AIC 3,625.82 3482.17
R2 0.062 0.687

As given in the table, we demonstrate residual sum of square (RSS), and Classic AIC, of GWR which are smaller than the OLS, whereas R2 of GWR is greater than OLS. These parameters or indicators prove that the GWR model is better fitting than OLS to investigate whether independent variables significantly vary spatially. The global OLS model explained 6.2% variation in confirmed malaria cases.by environmental factors ($R^2 = 0.062$). It implies that 93.8% of the confirmed malaria cases is caused by unknown factors not investigated in this study and may be related to local variation which is not taken into account in the OLS model [1]. The local GWR explained 68.7% variation in confirmed malaria cases (Y) by environmental factors ($R^2 = 0.687$).

Table 3: The result of global regression model and geographical variability test of local coefficients for six environmental factors.
Variables Global regression model output Geographical variability test
Estimate SE T value P value F DOF for F test DIFF of Criterion
Intercept 7.98 4.63 1.72 0.04 33.20 10.48 261.38 -347.99
"Altitude (X1)" -0.02 0.00 -4.03 0.00 0.24 12.02 261.38 19.19
"Aspect (X2)" -0.01 0.01 -1.60 0.05 0.55 22.68 261.38 24.91
"Distance from the river (X3)" 0.00 0.00 -0.84 0.24 1.84 18.15 261.38 -16.03
"Distance from lakes and pond (X4)" 0.00 0.00 0.39 0.71 0.90 15.04 261.38 7.99
"Distance from forest (X5)" 0.00 0.00 -3.69 0.00 2.99 14.61 261.38 -38.12
"Rainfall (X6)" 0.00 0.00 2.38 0.02 13.07 10.17 261.38 -158.91

We use GWR4 software to compare performance between global OLS and local GWR. Moran's I test is not available for the analysis. The value of DIFF criterion indicates that the independent variables have spatial variability or local spatial heterogeneity that are altitude, distance from lakes and pond, distance from forest, and rainfall have spatial variability. ANOVA in which the global model is compared with the GWR model. The ANOVA tests the null hypothesis that the GWR model represents no improvement over a global model. The results are shown below (Table 4).
Table 4: ANOVA testing the null hypothesis that the GWR model represents no improvement over a global model.

Source SS DF MS F

Global Residuals 100625.2620 429.0000
GWR Improvement 67075.981 197.736 339.220
GWR Residuals 33549.281 231.264 145.069 2.338336

The ANOVA test gives a brief guide to the improvement in model fit when we compare the local and global models. The GWR model could explain the relationship between the response variable „ confirmed malaria cases." and six explanatory variables significantly better than the global regression model OLS with F count (2.34) > F table (2.12), The locally weighed R2 between the observed and fitted values is a measure of how well the model replicates the local malaria incident values around each observation. GWR ANOVA Table is an integral part of result "Semiparametric Geographically Weighted Regression analysis", Release 1.0.90 (GWR 4.0.90)
3. The objective could be to predict, but here it seems that it is very interesting to identify malaria-environment relation and according to the location.
Response: We use ecology design study with the village that contains information of both attributes and location as unit analysis to predict confirmed malaria cases with potential environmental and geographic predictors of malaria.
4. A validated OLS can lead to a global policy and a validated relationship with GWR is more appropriate to drive to the local system.
Response: We completely agree and included this statement in our discussion chapter. Geographically Weighted Regression explores spatial varying impacts of these factors across the study area focusing attention on local variations in ecological associations. The set of selected environmental risk variables under consideration revealed significant associations with local confirmed malaria cases and these associations varied geographically across the study area. We observe and quantify different local factors driving confirmed malaria cases in different parts of the villages. A more in-depth understanding of local ecological factors influencing confirmed malaria cases may not only be used for developing sustainable regional malaria control programs but can also benefit malaria elimination efforts.
5. In the case of this study, it an excellent spatial analysis to identify which parameter to look closer and where and how much it varies and where it would be more appropriate to do so and for example do.
Response: Thank you for this positive comment on our study.
2. The validation is not clear.
Response: The model validation procedure conducted following steps: Step1: Preparation dataset. Step 2: Specify one regression type and the variable settings needed to determine the GWR model. We choose Geographical variability test, for model coefficient test obtained. Step 3: Currently, we use a geographic kernel type and its optimum bandwidth based on Selection Criteria. We demonstrated an "'Fixed' (Gaussian)" and selection bandwidth use "Golden section search" then use AIC criteria. It is a statistical model fit measure. It quantifies the relative goodness-of-fit of various derived statistical

models, giving a sample dataset. The preferred model is that with the lowest AIC value. Step 4: Specify filenames for the files storing the modelling results, and Step 5: Execute the session to compare necessary calculations and read results. Through the geographical variability tests, the AIC and 'Fixed' (Gaussian) kernel are enabled to find the size and select the optimal bandwidth if the model is fit. We demonstrated OLS assumptions for classical diagnostic regression as multicollinearity test has done before the modelling. The regression was computed with many variables, which potentially gave rise to multicollinearity. We used an index based on predictive modelling variance that is Variance inflation factor (VIF) [4] Multicollinearity could occur when one independent variable was a linear function of another independent variable and previously observed in GWR modelling [5].
The following 'rules-of-thumb' for evaluating these factors: VIF > 10 give evidence of multicollinearity. with VIF > 100 there is certainly multicollinearity among the variables. [6, 7]. We show in multicollinearity does not occur, because the VIF value is less than 10 and the tolerance value is higher than 0.1. So, in the OLS method obtained a regression equation to estimate the actual regression model.

Collinearity Diagnostics

Variable	VIF	SQRT VIF	Tolerance	R-Squared
altitude	1.42	1.19	0.7041	0.2959
aspect	1.00	1.00	0.9965	0.0035
distfriv	1.05	1.03	0.9497	0.0503
distflak	1.07	1.04	0.9335	0.0665
distffor	1.18	1.08	0.8502	0.1498
rainfall	1.17	1.08	0.8532	0.1468
Mean VIF	1.15			

The basic idea of GWR is that the parameters can be calculated in the study area with the dependent variable and one or more independent variables that it has been measured in places where the location is known. [8] In GWR, the sufficient number of degrees of freedom is a function of the bandwidth so the adjustment may be quite marked in comparison to a global model like OLS. For this reason, the AIC and R2 are preferred as a means of comparing models. So, we conclude, a valid GWR modelling is more appropriate to lead to local policy. In addition, the F test suggests that the GWR model is a significant improvement on the global model for confirmed malaria cases. In our case, these parameters prove that the GWR model is better than OLS that is a powerful tool for exploring spatial heterogeneity

6. Interpretation of the results could and be improved and better put in context.
Context with links with transmission and specific known ecological preferences of some Anopheles species (Ex: Anopheles found in the forest for villages where the distance to the forest is a factor).
Response: We revised the discussion chapter accordingly. See also line 290-295: An. (Cellia) leucosphyrus Dönitz is considered to be of epidemiological importance for malaria transmission in forested areas of Sumatra (McArthur, 1951). The Anopheles was reported from eight sources at 47 independent sites. The first record of Anopheles parangensis from Sumatra was reported by O'Connor and Sopa (1981), but with no details on location [9]. However, in current research, we did not investigate the main Anopheles vector diversity in each study area.
7. Context with other studies, it is better to refer to study with similar environment, latitude, health system.
Response: Determination of regional vulnerability using GWR in Purworejo Regency of Indonesia concludes that each region is considered to have a distinctive characteristic that is different from other regions. So, it is necessary to give individual calculation to get weight on each parameter determining the vulnerability of Malaria. We discussed this study and other studies related to the outcomes of our modelling (please see discussion).
8. Discuss more the difference between districts, especially those with very high or very low local R2
Response: See lines 207-213
The GWR model provides evidence for a locally different influence of environmental factors on confirmed malaria cases.as shown by varying R^2 (Fig. 6). "Altitude" and "distance from lake and pond" show a positive association and "aspect" a negative association with malaria case in the Northern study area (Musi Banyuasin). "Rainfall" and "distance from river" show a positive association with confirmed malaria cases.in the Eastern part of Musi Rawas and Lahat. The variables "aspect", distance from lake and pond" and "distance from forest" are positively associated with confirmed malaria cases in large parts of the study area and discussion chapter for discussing environmental parameters.
9. scale the relationship with parameters that vary intra-annually vary.
Response: Currently, we use secondary data 2013, due to current data limitations. Annual rain data is only available from some weather stations in South Sumatra and thus the interpolation of the 2013 rainfall data would result in bias. So, we use the five-year average data to spatially interpolate rain data throughout Sumatra.Minor comments:
1. Title: add environment to the title. It could be "Spatial modelling of malaria incidence relationship with environment factors" or something like that.
Response: Changed as suggested.
2. Keywords: malaria, geographically weighted regression, GWR, Ordinary least squares regression, OLS, Sumatra, rainfall, elevation, distance to water.
Response: Changed as suggested. Geographically weighted regression (GWR), Ordinary least squares (OLS), Physical environment, Local climate, Sumatra, rainfall, elevation, and distance to water.
3. It should be appreciated to name the primary Anopheles vector species for each type of environment or district
Response: An. nigerrimus is a confirmed malaria vector in Indonesia with the first evidence of Plasmodium infection reported by Overbeek from Palembang, South Sumatra in 1940 [9]. The distribution of malaria vectors amongst the

main islands is also not uniform Sumatra Island has six species, Papua (at least five species) and the Lesser Sundas archipelago (five species).
Figure: A map of the distribution of primary Anopheles malaria vectors in Indonesia

Currently, the primary vector of malaria which confirmed the main vector of malaria (found sporozoite) from the salivary glands as follows: An. letifer, An. nigerrimus, An. maculatus, An. sinensis, An. barbirostris, An. vagus, and An. sundaicus in South Sumatra Provinces region. The primary anopheles vector data are obtained from several studies, and particularly data from Vector and animal-borne disease control unit of Research and Development, National Institute of Health Research and Development (NIHRD), Ministry of Health (Indonesia) at Baruraja. Also, data are based on the regular reporting of malaria from South Sumatra Provincial Health Office, the kind of plasmodium was Plasmodium falciparum and Plasmodium vivax in this studies area. However, in current research, we did not investigate the primary vector Anopheles diversity in each study area.

10. Cases number or incidence?
Response:
The dependent variable is „confirmed malaria cases (Y).
Case, confirmed : Malaria case (or infection) in which the parasite has been detected in a diagnostic test, i.e. microscopy, a rapid diagnostic test or a molecular diagnostic test

Case, malaria : Occurrence of malaria infection in a person in whom the presence of malaria parasites in the blood has been confirmed by a diagnostic test
Note: A suspected malaria case cannot be considered a malaria case until parasitological confirmation. A malaria case can be classified as imported, indigenous, induced, introduced, relapsing or recrudescent (depending on the origin of infection); and as symptomatic or asymptomatic. In malaria control settings, a "case" is the occurrence of confirmed malaria infection with illness or disease. In settings where malaria is actively being eliminated or has been eliminated, a "case" is the occurrence of any confirmed malaria infection with or without symptoms
Incidence, malaria : Number of newly diagnosed malaria cases during a defined period in a specified population.
Ref :
Global Malaria Programme, WHO malaria terminology. World Health Organization 2016. Updated in August 2017. Retrieved from apps.who.int/iris/bitstream/10665/208815/1/WHO_HTM_GMP_2016.6_eng.pdf
Currently, from reporting of the new case of malaria, and these are confirmed malaria cases. Data for patients were positive for malaria parasites will entry in individual including (name, address, type of parasite, the treatment used). Monthly reporting is done in the first stages from puskesmas: the primary health care system in Indonesia at the village level continue to districts in the 2nd stage and then to provinces in the 3rd degree.
11. Which georeferenced system is used in which units (meters or degrees)
Response: The study area map (Figure 1) uses the World Geodetic System (WGS84) as its reference coordinate system (line 87-89).
12. Maps 6 and 7: Add units, please.
Response: The figures follow 3, 4, 6, 7 and eight deliberately do not display coordinate system due all these maps are meant to accentuate thematic information. The coordinate system can be seen in Figure 1.
13. Which is the scale or resolution in time and space for each parameter?
Response: Parameter distance from the river, distance from lake and pond, and distance from the forest are processed from River, Lakes, Ponds maps which derive from the topographic map which have 1: 50,000 scale. Forest cover maps obtained from Forest cover maps of South Sumatera 2013 on the scale of 1: 250.000. Rainfall parameter was calculated based on annual average rainfall over five years, and it was interpolated from several weather observations stations in studies area.
14. Forest: How old is the forest layer? Which year?
Response: The forest cover maps were extracted from the land cover map which made in 2013. This map is sourced from Ministry of Environment and Forestry, Indonesia.
15. Can we guess that some parts have been deforested since the forest cover has been recorded? Do we have information on the percentage of deforestation between this year and 2013?
Response: Indonesia contributes significantly to deforestation in Southeast Asia. However, much uncertainty remains over the relative contributions of various forest-exploiting sectors to forest losses in the country [14]. Forest is discussed first because one of variable research is the distance to the forest. Regarding of is studies area deforestation, we do not have information on the percentage of deforestation due to current data limitations.
16. In the discussion links between your result and what you say about deforestation.
Response:
Next to climatic and environmental factors, distance of houses to a forest are interrelated through anthropogenic activities influencing the local and regional climate [10, 11]. A cross-sectional view in Brazil revealed for example that malaria case across health districts is positively correlated with the percentage of aggregated deforestation [12]. These observations can be confirmed for the relationship of malaria case with distance to lake, pond and forest for South Sumatra. Anopheles (Cellia) leucosphyrus is considered to be of epidemiological importance for malaria transmission in forested areas of Sumatra [9]. Anopheles was reported from eight sources at 47 independent sites. The first record of Anopheles parangensis from Sumatra was reported by O'Connor and Sopa (1981), but with no details on location [9]. In current research, the main Anopheles vector diversity in each study area was not investigated.
17. Rainfall: the rainfall-malaria relationship is probably a nonlinear relationship as it is written in the discussion. In this annual study, rainfall is used. Is it average rainfall or total amount?
Response: Average annual rainfall period 2007-2013 in South Sumatra has been used for analysis.
18. Temperature
Response: See limitations of the study: Due to limited data, some explanatory variable were not investigated like

temperature. However, the temperature is connected with altitude and aspect or direction of the slope. In the same way, land use may be associated with distance from the river and distance from lakes and pond. Thus, although these parameters (temperature, humidity, land use) were eliminated from analysis, these environmental factors were indirectly represented by our chosen set of variables.

19. Elevation: often described as an indirect factor: less humidity, lower temperature or suitable for different Anopheles species.
Response: Thank you for your advice. The global OLS model revealed that altitude, distance from lakes and pond, and distance to the forest have a negative coefficient and rainfall has a positive coefficient, significantly influence malaria case. It meant confirmed malaria cases is more common in regions with high rainfall, lowland and areas adjacent to forest areas. Elevation often described as an indirect factor: less humidity, lower temperature or suitable for different Anopheles species. See also discussion chapter.
20. Results. Present only the result without assuming cause between the variables.
Changed as suggested.
21. Interpretation: Explain links with field data and known information.

The highest malaria case with 1,449 cases spread over 124 villages was found in Lahat District. Our analyses have identified Lahat as the South Sumatran district in which environmental factors were of greatest relevance for confirmed malaria cases. Lahat District has both lowland and mountain regions and is home to diverse ethnic groups such as the Gumai who live along the rivers of the highland areas. Lahat district is located between 3.25 to 4.15 degrees south latitude, 102.37 to 103.45 degrees east longitude. Lahat Regency has a climate tropical and wets with rainfall variation between 267,375 to 222,175 millimetres per month, with as many rainy days 145.25 days or an average of 12.10 days per month. Air temperature varies between 22,16 to 30.47 Celsius. The average air humidity is 78.50 with an average wind speed of 4.66 Km per hour.

22. References: Rainfall and malaria: you could add Botswana and Ethiopia works
Response: added in the discussion as suggested.
Comments by Line
23. Background: In the background, when a reference is cited to state a link between malaria transmission and an environment factor, it should be better to mention in which country or environment type. Example: line 39: "lowland location."
Response: added in the discussion as suggested.
24. Line 48 "it proliferates faster under higher temperatures", it depends where.
Response: The sentence has been rephrased: Vectors and parasites are both highly sensitive to any temperature changes, for example, the parasite proliferation depends on temperature.
Please read also: … "Higher temperatures also quicken the digestion of the blood meal and maturation of its developing eggs, thus increasing vectorial biting frequency. Given these well-established, mainly laboratory determined climate sensitivities, malaria has long been identified as the infectious disease most vulnerable to climate change (WHO, 1990)"…. And "…Higher temperatures may prolong the malaria transmission window and reduce the incubation period required for replication of the parasite in an infected mosquito. As regional temperatures change across India, the transmission window for malaria is likely to increase by 2–3 months in the northern states of Punjab, Haryana and Jammu and Kashmir, but to decrease in more southerly Odisha, Andhra Pradesh and Tamil Nadu as temperatures exceed 40°C (Dhiman et al., 2010)…."
And in another reference :
Plasmodium falciparum has Threshold (0C), Minimum for transmission was 16–19 and Maximum for survival was 33–39. P. vivax has Threshold (0C), Minimum for transmission 14.5–15 and Maximum for survival 33–39. Then Lower threshold (0C) both of the Plasmodium was 8–10 for biological activity. [13]
25. Line 63: very important to know which variables you have studied, please list them here. "performance of the OLS and GWR models in predicting.."
Response: We compare global OLS and local GWR modelling to analyse the environmental risk factors for malaria in South Sumatra that vary geographically at the regional level. The locally different ecological factors studied to potentially predict the response variable „ confirmed malaria cases." (Y) are altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4), distance from the forest (X5), and rainfall (X6).
26. Methods: Study area: Line 77: a range of altitude would be appreciated, highest altitude for all the area or for each district
Response: The elevation in the study area varies between 0 to 3,159 metres above sea level. [14]
27. Line 78: is it monthly rainfall amount by station?
Response: This is the amount of average monthly rainfall taken from the weather stations conducted by Indonesian Agency for Meteorological, Climatological and Geophysics (BMKG) Palembang.
28. Study population and data collection: Lines 85-86: How many PHC? Just to have an idea of the density by district (or by population or by area)
Response: Based on the dataset, totally of The Primary Health Centre (PHC) was 140. Some PHC in each district is varying. Lines 68-69
29. Line 92: 36 372 patients or presumed positive malaria cases? Some patients may come several times a year.
Response: There is a special form of malaria case to reporting whether the patient is a new or a relapse.
30. Line 94 % (3578/36372 is around 10%
Response: Almost 10% of those participants were tested positive for malaria.
31. Line 97: precise which sort of villages or number (436 villages)
Response: The number of person who does have a positive malaria test are spread across 436 villages.
32. Preparation of spatial data: Data acquisition and selection, Line 106:

How many stations? How many km are they close to each other? To have an idea of their density
Response: The distance between weather observation stations were 50-100 km in flat topography and 10 km in hilly terrain.
33. Data pre-processing, Line 112 DEM which spatial resolution? Issued from which satellite data type?
Response: DEM data is processed from a contour map or a topographic map scale 1:50.000 with high relief; the contour interval is 25 m.
34. Line 115 Which spatial interpolation method did you used for rainfall? Which classification and from which criteria did you use it?
Response: Rainfall data has been interpolated by Indonesian Agency for Meteorological, Climatological and Geophysics (BMKG).
35. Line 120: VIF? It should be useful here to describe the variance inflation factor, what is it and how it works.
Response: Please check in line 131
36. Data processing; See major comment above
Response: see revised manuscript, lines 126 to 185
37. Miss comparison between the two methods OLS and GWR
Response: see revised manuscript, lines 218-222, and table 2 : and table below:[1]

Ref: Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships/cA. Stewart Fotherington, Chris Brunsdon, and Martin Charlton. Wiley, 2002.
38. Miss validation
Response: see our response to comment 9.
39. Results; Environmental factors influencing malaria incidence at village level: local GWR model; Line 194: related to size of weights
Response: The regression coefficients were not directly related to the size of weights in the GWR model, but rather to the estimates of the values of all explanatory variables.
40. Line 206: "..show that the environmental factors prevailing In these regions are less suitable for explaining the variance of malaria incidence in this area" need to explain why, please.
Response: Based on the value of regression, the confirmed malaria cases caused by environmental factors is most dominant than others. So, that environmental factors are more appropriate to explain its contribution to the variation of confirmed malaria cases.in the location.
41. Comparison between OLS and GWR: cf major
Response: see revised manuscript, lines 218-222
42. Discussion; Cf major., Line 289: Avoid "spatial epidemiology microscope."
Response: Based on our findings, GWR is a diagnostic model discovering spatially varying relationships. and local GWR analysis can, therefore, serve as a 'spatial epidemiology microscope.'
43. Line 298: "The approach arbitrarily plots all of the cases in the settlements" I don't understand what you mean.
Response: The availability of malaria case data is the number of positive malaria per village, and it is not the coordinates of each malaria positive so, the case is placed in the centre of the settlement.
44. Line 305 - 311: Add seasonality studies, non-linear relationship, time downscaling (to monthly rather than annual cases), etc.
Response: Thank you for your advice. See below and discussion chapter:
Climate data are frequently used to account for the spatial, seasonal and interannual variation for Malaria transmission. Modelling numerical evaluations by time and space show connection with malaria prevalence.[15]
In the future, additional explanatory variables should be addressed to provide a comprehensive analysis of confirmed malaria cases. This should comprise, for example, the behaviour of mosquito vectors and that of community members, access to and delivery of health services, and other eco-bio-social factors that affect the confirmed malaria cases.
Despite these limitations, our study sheds light on relevant local and regional realities regarding environmental variation and sociocultural practice which might interplay with vector-host relationships and provide a suitable environment for malaria mosquitoes.
45. Maps; Figure 1 and 3: you have to choose the same methods for all the maps to code the districts, numbers or abbreviations.
Response: Thank you for your advice. Revised.
46. Figure 3: scale and North are missing.
Response: Revised as suggested.
47. Figure 4: Legend (spatial representation map showing.. not needed) Each explanatory variable
Response: Figure description revised.
48. Figure 5: Is multicollinearity test also applied with the response variable?
Response: Multicollinearity test is applied both in explanatory and in response variable
49. Figure 6: reformulate the legend, please. It should be something like " predicted value from GWR".
Response: Revised as follows: Figure 6: Predicted value from GWR for parameter estimates of explanatory variables of confirmed malaria cases in the study area.
50. Figure 7: ….. Significance percentage value for each explanatory variable by village location
Response: Revised as follows: Figure 7. Student's test significance (95% and 99% confidence interval) for each explanatory variable and village location.
51. Figure 8: Local regression coefficient (R2) from GWR method by village location
Response: Revised as follows: Figure 8. Goodness-of-fit of GWR model (local R2) for confirmed malaria cases associated with environmental factors in South Sumatra, Indonesia.
Reference:
1. Fotheringham AS, Brunsdon C, Charlton M: Geographically weighted regression the analysis of spatially varying relationships. University of Newcastle, UK: John Wiley & Sons; 2002.

2. Rodrigues M, de la Riva J, Fotheringham S: Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. Applied Geography 2014; 48:52-63.
3. Nakaya T: GWR4.09 User Manual. GWR4 Windows Application for Geographically Weighted Regression Modelling. Kyoto: Ritsumeikan University, Department of Geography24 March 2016.
4. Halimi M, Farajzadeh M, Delavari M, Takhtardeshir A, Moradi A: Modelling spatial relationship between climatic conditions and annual parasite incidence of malaria in southern part of Sistan & Balouchistan Province of Iran using spatial statistic models. Asia Pac J Public Health 2014; 4:S167-S172.
5. Wheeler D, Tiefelsdorf M: Multicollinearity and correlation among local regression coefficients in geographically weighted regression. J Geogr Syst 2005; 7:161-187.
6. Rabe-Hesketh S, Everitt B: Handbook of statistical analyses using stata. CRC Press; 2003.
7. Kothari P: Data Analysis with Stata. Packt Publishing Ltd; 2015.
8. Brunsdon C, Fotheringham A, Charlton M: Geographically weighted summary statistics—a framework for localised exploratory data analysis. Comput Environ Urban Syst 2002; 26:501-524.
9. Elyazar IR, Sinka ME, Gething PW, Tarmidzi SN, Surya A, Kusriastuti R, Winarno, Baird JK, Hay SI, Bangs MJ: The distribution and bionomics of anopheles malaria vector mosquitoes in Indonesia. Adv Parasitol 2013; 83:173-266.
10. Widayani P, Danoedoro P, Mardihusodo Sj: Pemodelan Spasial Kerentanan Wilayah Terhadap Penyakit Menular Terkait Lingkungan Berbasis Penginderaan Jauh (Kasus Malaria, Leptospirosis Dan Tuberkulosis Di Sebagian Wilayah Provinsi Jawa Tengah Dan DIY). Universitas Gadjah Mada, 2016.
11. Kleinschmidt I, Bagayoko M, Clarke G, Craig M, Le Sueur D: A spatial statistical approach to malaria mapping. Int J Epidemiol 2000; 29:355-361.
12. Olson SH, Gangnon R, Silveira GA, Patz JA: Deforestation and malaria in Mancio Lima county, Brazil. Emerg Infect Dis 2010; 16:1108.
13. Chihanga S, Haque U, Chanda E, Mosweunyane T, Moakofhi K, Jibril HB, Motlaleng M, Zhang W, Glass GE: Malaria elimination in Botswana, 2012–2014: achievements and challenges. Parasit Vectors 2016; 9:99.
14. Midekisa A, Beyene B, Mihretie A, Bayabil E, Wimberly MC: Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia. Parasit Vectors 2015; 8:339.
15. Chirebvu E, Chimbari MJ, Ngwenya BN, Sartorius B: Clinical Malaria Transmission Trends and Its Association with Climatic Variables in Tubu Village, Botswana: A Retrospective Analysis. PLoS One 2016; 11:e0139843.
16. Thomson MC, Ukawuba I, Hershey CL, Bennett A, Ceccato P, Lyon B, Dinku T: Using rainfall and temperature data in the evaluation of national malaria control programs in Africa. Am J Trop Med Hyg 2017; 97:32-45.
17. Patz JA, Olson SH: Climate change and health: global to local influences on disease risk. Ann Trop Med Parasitol 2006; 100:535-549.
18. Hoshen MB, Morse AP: A weather-driven model of malaria transmission. Malar J 2004; 3:32.

Reviewer #2: MALJ-D-17-00578
Overall Comments:
This manuscript applies spatial analysis to malaria data in a low-endemic and different area. By focusing the analysis on routinely reported data as well as using spatial covariates accessible within the country, this provides an approach that is accessible to malaria programs within the country. Overall, this manuscript is well written and provides useful information to help better understand malaria epidemiology in this area. However, before recommending for publication, I have several comments that should be considered.
Response: We appreciate very much for the reviewers' valuable comments, constructive criticisms, and insightful feedback. We carefully considered reviewer's suggestion and tried our best to improve the manuscript based on their explanation below. Each comment has been carefully considered point by point. Responses are made in italics.
Major comments:
1. The study population needs to be better defined. The authors seem to use the terms surveillance and research population interchangeably, but the latter suggests that some of the data were collected outside of the routinely collected data. The details were given (abstract and methods (LL 83-98)) on the difference between the total population and 'research participants' needs to be simplified. As the total population has the potential to go to the facility and be captured as part of routine surveillance, it is not clear what the distinction is. It would improve clarity to use terms consistently throughout (e.g. X positive for malaria or XX suspected of having malaria)

Please clarify and use only one terminus
Response: Based on methodology, the population of our study was the total number of villages in 8 districts in South Sumatra, Indonesia and the sample was the village in the study area where positive malaria case is found. Currently, the village is a unit of analysis, and in our dataset, we call it "Toponymy". We noted 3,578 patient who has laboratory diagnosis of malaria. The cases were spread over 436 out of 1,613 villages. The village that contains information for both, attributes and location, is a unit analysis. We investigated potential ecological predictors of confirmed malaria cases in the different regions by performing global Ordinary Least Squares (OLS) and local Geographically Weighted Regression (GWR).
Please see also lines 92-93 for Toponymy", and lines70-73 for information The cases spread over 436 out of 1,613 villages that were used for unit analysis
2. Mosquitoes and their ecosystems are significant spatial drivers for malaria transmission. You mention that 25 different species of Anopheles have been identified in the country (LL 29) but are all of these found in your study area? Furthermore, the results of the spatial heterogeneity in risk should be discussed in the context of the spatial heterogeneity of the vectors in the reason. This is an important confounding factor to address as different species may have different ecological niches and therefore different factors may be important in different places.
Response: Actually, between 20-25 different species of Anopheles have been identified in Indonesia. There are 25 species Anopheles mosquitoes that have been confirmed to be malaria vectors in Indonesia, which are spread and

divided into two zones of geographic dispersal of the Australian and Oriental zones.[16] “Approximately 230 million people live in Indonesia. The country is also home to over 20 anopheline vectors of malaria which transmit all four of the species of Plasmodium that routinely infect humans.” [9, 17]
Currently, in South Sumatra Provinces region, the main vector of malaria which confirmed were An. letifer, An. nigerrimus, An. maculatus, An. sinensis, An. barbirostris, An. vagus, and An. sundaicus.
Pleasee also read line 360-365 in the manuscript: Anopheles (Cellia) leucosphyrus is considered to be of epidemiological importance for malaria transmission in forested areas of Sumatra [11]. Anopheles was reported from eight sources at 47 independent sites. The first record of Anopheles parangensis from Sumatra was reported by O'Connor and Sopa (1981), but with no details on location [9]. However, in current research, we did not investigate the main Anopheles vector diversity in each study area was not investigated.
3. The terminology of modelling methods: in the malaria spatial modelling field, the terms global and local typically refer to different scales of spatial autocorrelation both of which are present in malaria transmission (e.g. broader temperature bands vs mosquito flight range). Your description of the models used is very clearly articulated and accessible to non-spatial/statistical people. However, re-framing this as a ‘non-spatial’ and ‘spatial’ regression instead of global/OLS and local/GWR would help clarify the important differences between the approaches being compared.

Response: In our study, we use term global for OLS and local for GWR. OLS and GWR are regression methods that both consider spatial factors. The difference of the regressions are :
- OLS, the parameter estimate has the same value at all locations so that the relationship between the response variable and explanatory variable is considered homogeneous (stationary).
- GWR, the parameter estimation value at each location varies so that the relationship between a response variable and explanatory variable is heterogeneous (non-stationary).
Thus, the term OLS cannot be substituted by 'non-spatial' and GWR cannot be substituted by 'spatial'. Additional explanations, it can be read below.

4. Some important references on spatial modelling of malaria are missing. See works by A. Noor, P. Getting, I. Kleinschmidt, E. Giorgi for example.

Response: Added as suggested.

A global database of malaria parasite prevalence using Malaria Atlas Project (MAP) to collect relational databases and related GIS. The documentation will help to improve the global spatial of malaria which demands investment in the collection of epidemiological intelligence. [18]

Analysis spatial, multiple regression analysis and spatially adjusted, are important implications for malaria control programs in a certain area with a method of adjusting the regression analysis was undertaken to identify factors that might explain very strong heterogeneity in the rates in South Africa. The results indicated strong spatial correlation in the rates by using generalized linear mixed models and variograms that malaria case was significantly positively connected with higher winter rainfall, a higher average maximum temperature and significantly negatively associated with increasing distance from water bodies. [19]

A simple two-stage procedure for producing maps of predicted malaria risk that is OLS analysis modelling on a larger scale to determine the relationship between Malaria prevalence in children under ten within the interval 0 to 1 and geo-statistical (‘kriging') approaches used residual spatial dependence in the data to improve prediction at the local level. Some ecological potential predictors of malaria using climatic, population and topographic variables and investigated spatial pattern in the residuals of the model which is an important tool for malaria control in Mali. [20] A malaria risk map of the West African region uses on malariometric data survey to predict parasite prevalence for the whole of West Africa as a useful tool for health planners. It provides the opportunity for producing empirical models and maps of malaria distribution at a regional and eventually at a continental level. [21] A standard geostatistical model is important to prevalence mapping which relies on empirical prevalence data of this kind is a generalized linear mixed model with binomial error distribution, logistic link and a combination of explanatory variables and a Gaussian spatial stochastic process in the linear predictor.[22] Malaria endemicity within defined stable spatial limits of P. falciparum transmission has been investigated by a model-based geostatistical procedure was implemented within a Bayesian statistical framework.[23] Maps of transmission malaria and the impact of malaria on human populations not only contribute to a rational basis for control and elimination decisions but also are necessary to identify populations at different levels of risk and to evaluate options for disease control objectively.[24]
Advances geo-statistics are modelling, and malaria parasite prevalence data assemblies can be used to insert plasmodium falciparum risk distributions. A map of infection and disease risks is an appropriate strategy for the control of malaria requires Kenya. [25]

5. There is some repetition of concepts in the methods section that could be better organized or the difference between the multiple usages is not clear enough to appreciate the need for duplication. For example, the first two sections of preparation of spatial data both discuss data interpolation. Similarly, the discussion of testing for multicollinearity is in pre-processing and processing sections. It might be helpful to distinguish when the data in question is spatial/a map and non-spatial. The addition of figure 5 clarifies the flow of information, but this lucidity should be reflected in the text.
Response: We revised the method description as suggested. Please also see our response to your comment 3.
6. In the results section, a lot of emphases is placed on regression coefficients and less on the interpretation of these coefficients.

Response: We added some interpretation in result section.

7. The approach used for comparing the two models are missing in the methods section. Please add the testing approach.

Response: Added as suggested. See line 218-222

8. Minor comments; Abstract: How many villages didn't report any malaria (436 of X villages)?
Response: The cases were spread over 436 of 1,613 villages. It mean's villages without malaria cases were 1,177 villages in the study area.
9. LL 2-6 – a reference to figure 1 would be helpful here to give readers some spatial context of the places being mentioned.
Response: See lines 87-88 The study area is located between 1°46' and 4°55' of southern latitude and between 102°4' and 104°41' of eastern longitude and has a total surface area of 46,377.40 km2 (Fig. 1). It covers eight endemic malaria districts of South Sumatra, Indonesia, namely Lahat, Muara Enim, Musi Banyuasin, Musi Rawas, North Musi Rawas, Ogan Komering Ulu, South Ogan Komering Ulu, and Lubuk Linggau. The topography of the area varies from lowland to mountainous landscapes. The elevation in the study area varies between -6 to 3.150 metres above sea level (Fig. 4). The climate is tropical and wet. In 2013 in South Sumatra, the lowest rainfall was 31 mm (August) in Lahat district, and the highest rainfall was 613 mm (March) in Palembang City. Monthly average temperatures ranged from 26.6 to 28.3°C and relative humidity from 81% to 88% in 2013 [26].
10. LL 21- "recent developments" needs to be elaborated on to ensure that those not familiar with the area can understand the context. Recent political? Economic? Social?
Response:
Please read 59-65, and 285-287

Indonesia contributes significantly to deforestation in Southeast Asia. Recent developments of deforestation have led to unsustainable practices which have resulted in a high frequency of deforestation in some regions and are an important factor influencing malaria incidence [27]. Deforestation is connected with malaria incidence in the county (município) of Mâncio Lima, Acre State, Brazil. The cross-sectional study shows 48% increase in malaria incidence associated with cumulative deforestation within respective health districts in 2006. [12]
11. LL 94 – what diagnostic test was used?
Response:
Either Rapid Diagnostic Tests (RDT) or microscopic assessment or both were used to confirm a malaria case. Please read 76-78

12. LL 102 – more details on spatial input parameters are needed – what is the resolution of the different surfaces? Is it commercially available (e.g. landsat imagery) or did the government commission the images to be created? What year was it captured?

Response: see lines 100-105; The topographic map consists of a collection of geographic data presented as thematic layers on a sketch done by The Indonesian Geospatial Information Agency (BIG). Researchers are not involved in this process. Topography data source: RBI (Rupa Bumi Indonesia) Bakosurtanal which is updated in 2014 in the location of study area.

13. The topographic wetness index (Cohen et al.) was shown to be a significant predictor of malaria and is a metric that can be derived from available data. The authors should consider adding to their analysis.
Response: The topographic wetness index (Cohen et al.) will be considered for future research, though its use needs to be further discussed in our team.
https://en.wikipedia.org/wiki/Topographic_wetness_index
"The index was designed for hillslope catenas. Accumulation numbers in flat areas will be very large, so TWI will not be a relevant variable." This may be a disadvantage in our study.
"The TWI has been developed to study spatial scale effects on hydrological processes and characterize biological processes such as annual net primary production, vegetation patterns, and forest site quality." One may assume that the analysis of topographic wetness index (Cohen et al.) may not reveal new patterns in our study given that it integrates all six variables that we studied,
14. It would be helpful to highlight in the methods (LL 102 – 107) that the malaria data inputted into the model is aggregated village level data with the village centroid (?) used as the spatial unit.
Response: Added as suggested. The malaria input data is aggregated village level data with the village centroid used as the spatial unit.
15. LL 115 – "The rainfall map…..obtained from the scanned maps" – which maps?
Response:
See lines 93-110; A precipitation map (annual average) was obtained by interpolating the data of annual average rainfall from BMKG Climatological Station Class I in Palembang, South Sumatra, Indonesia 2007-2013 period. The interpolation process is done by BMKG then it classified into 7 rainfall classes. We obtained a map in JPG format. Further, the map rectified both in georeferencing and digitizing to create a map of precipitation vector format. Rectification is a process of transforming data from a single grid system using a geometric transformation. The result of digitization process can be seen in Figure 2 (rainfall variable).

16. LL 118 – "GWR should have a normal distribution" – is this that variables used for GWR should normally be distributed? It would be helpful to have the untransformed distributions as a supplementary table to show the non-normality and the transformed version to support this.
Spatial data contains information with both attributes and location. The Geographically weighted regression (GWR) model, a local regression, was developed from an Ordinary Linear Regression (OLS) model based on nonparametric regression [28]. A non-parametric test does not assume anything about the underlying that the data comes from a normal distribution. GWR is a local regression that emphases 2nd order variation whereas OLS is a first order model. GWR is a varying-coefficient modelling technique. The general model in running both is to draw inference about first (global) and second (local) order process but, more directly GWR is specified to account for nonstationarity. GWR is a method for exploring spatial nonstationarity. This then produces a set of parameter estimates at each point in the defined geographical area. In this case, we run OLS using robust regression. Robust statistical tests operate well across a wide variety of distributions. The basic GWR method may be regarded as generalisations of the basic method where the core notion of a spatially non-stationary OLS regression model is enhanced [28].
17. LL 120 – VIF should be defined at first instance
Response: see lines 130-131, and 191-192 ;Done as suggested. The Variance Inflation Factor (VIF) and tolerance are both widely used measures of the degree of multicollinearity. [29]
18. LL 201 – How much is the results in Lahat having the highest influence of environmental factors due to the higher case numbers and therefore more predictive power?
see lines 51 and 215-216
The regression coefficients for malaria incidence at the local level range between 0.18 - 1 (Fig. 8) The highest influence of environmental factors on malaria incidences was found in Lahat District. We will discuss after we re-calculated the models which suggested of Reviewer #1

19. LL 200-207 – the term regression coefficients are typically used to denote the covariates and their corresponding constant that represents the rate of the linear change in the association with the outcome variable. Whereas R2 is a statistical measure of how close the data are to the fitted regression line and is interpreted as to how much of the variability is explained by the covariates. They are different measures with very different meanings, and therefore different terminology should be used to denote the two.

Response: Thank you for your feedback.
see lines 170-173, 178-181, and 215-216
R2 is the coefficient of determination (R Squared): indicates the kindness of the model or the contribution of the independent variable to confirmed malaria cases . R-squared is a statistical measure of how close the data are to the fitted regression line. It is also known as the coefficient of multiple determination for multiple regression. The regression coefficient is the constant (a) that represents the rate of change of one variable (y) as a function of changes in the other (x); it is the slope of the regression line. GWR4 provides almost same results for traditional GWR modelling. A few corrections have been made with regards to calculation methods for local diagnostic statistics, including local sigma and local R square. In output GWR4 Windows analysis, R square found in both in Global regression and GWR result. t represents the fraction of variability in response that can be explained by the variability in predictor variables. R2 is a statistical measure of how close the data are to the fitted regression line and is interpreted as to how much of the variability is explained by the covariates. In the simple linear regression case, R2 is simply the square of the correlation coefficient. The best model selection can be seen not only from the residual sum of square, and classic AIC but also in R square values. R2 in OLS was 68.7% and in GWR was 6.15% using 'Fixed' (Gaussian).
Reference:
1. Kementerian Kesehatan Republik Indonesia: Pedoman manajemen malaria. Jakarta: Direktur Jenderal Pengendalian Penyakit dan Penyehatan Lingkungan, Kementerian Kesehatan Republik Indonesia,; 2014.
2. Elyazar IR, Sinka ME, Gething PW, Tarmidzi SN, Surya A, Kusriastuti R, Winarno, Baird JK, Hay SI, Bangs MJ: The distribution and bionomics of anopheles malaria vector mosquitoes in Indonesia. Adv Parasitol 2013; 83:173-266.
3. Elyazar IR, Hay SI, Baird JK: Malaria distribution, prevalence, drug resistance and control in Indonesia. Adv Parasitol 2011; 74:41-175.
4. Guerra CA, Hay SI, Lucioparedes LS, Gikandi PW, Tatem AJ, Noor AM, Snow RW: Assembling a global database of malaria parasite prevalence for the Malaria Atlas Project. Malar J 2007; 6:17.
5. Kleinschmidt I, Sharp BL, Clarke GP, Curtis B, Fraser C: Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in Kwazulu Natal, South Africa. Am J Epidemiol 2001; 153:1213-1221.
6. Kleinschmidt I, Bagayoko M, Clarke G, Craig M, Le Sueur D: A spatial statistical approach to malaria mapping. Int J Epidemiol 2000; 29:355-361.
7. Kleinschmidt I, Omumbo J, Briet O, Van De Giesen N, Sogoba N, Mensah NK, Windmeijer P, Moussa M, Teuscher T: An empirical malaria distribution map for West Africa. Trop Med Int Health 2001; 6:779-786.
8. Diggle PJ, Giorgi E: Model-Based Geostatistics for Prevalence Mapping in Low-Resource Settings. Journal of the American Statistical Association 2016; 111:1096-1120.
9. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, Kabaria CW, Manh BH, Elyazar IR, Brooker S: A world malaria map: Plasmodium falciparum endemicity in 2007. PLoS Med 2009; 6:e1000048.
10. Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IR, Johnston GL, Tatem AJ, Hay SI: A new world malaria map: Plasmodium falciparum endemicity in 2010. Malar J 2011; 10:378.
11. Noor AM, Gething PW, Alegana VA, Patil AP, Hay SI, Muchiri E, Juma E, Snow RW: The risks of malaria infection in Kenya in 2009. BMC Infect Dis 2009; 9:180.
12. Huisman O, De By R: Principles of geographic information systems. ITC Educational Textbook Series 2009; 1.
13. Brunsdon C, Fotheringham A, Charlton M: Geographically weighted summary statistics—a framework for localised exploratory data analysis. Comput Environ Urban Syst 2002; 26:501-524.
14. Abood SA, Lee JSH, Burivalova Z, Garcia-Ulloa J, Koh LP: Relative contributions of the logging, fiber, oil palm, and

mining industries to forest loss in Indonesia. Conserv Lett 2015; 8:58-67.
15. Olson SH, Gangnon R, Silveira GA, Patz JA: Deforestation and malaria in Mancio Lima county, Brazil. Emerg Infect Dis 2010; 16:1108.
16. Fotheringham AS, Brunsdon C, Charlton M: Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons; 2003.
17. O'brien RM: A caution regarding rules of thumb for variance inflation factors. Quality & Quantity 2007; 41:673-690.

Reviewer #3: MALJ-D-17-00578
This manuscript presented an analysis of routine malaria surveillance data for 2013 to examine the spatial patterns of malaria in South Sumatra, Indonesia. Ordinary least squares and geographically weighted regression analyses were used to examine the potential role of environmental risk factors on the spatial patterns of malaria incidence. Findings indicated that rainfall and distance from the forest played a role in explaining the malaria incidence. While the paper contains results that could be of interest, major revisions are necessary for the language. The paper was not focused and included too much extraneous information, yet did not include important information about the methods. There were also several concerns with the methods and interpretation of the findings.
Response: We thank anonymous reviewers for providing us very insightful and constructive comments. We have tried our best to carefully consider and respond to all the comments raised by the reviewer 3. We revised the manuscript substantially to improve the language and the presentation of our data as outlined below.
1. Abstract: From a statistical perspective, it is unclear how "having an R-squared value of 60%" indicates "that almost all independent variables were significant at certain locations at the village level."
Response:. We rephrased the result part of abstract as following: The importance of different environmental and geographic parameters for malaria was shown at global and village-level in South Sumatra, Indonesia. The independent variables altitude, distance from forest, and rainfall in global OLS were significantly associated with malaria cases. However, as shown by GWR model and in line with recent reviews, the relationship between malaria and environmental factors in South Sumatra was found to strongly vary spatially in different regions.
Abstract: The conclusions do not match the stated aim of the paper and instead highlight the merits of methodological approach instead of how the findings "help in the development of local policies for malaria elimination" in South Sumatra.
Response: We rephrased the conclusions as following: A more in-depth understanding of local ecological factors influencing malaria confirmed malaria case as shown in present study may not only be usedful for developing sustainable regional malaria control programmes, but can also benefit malaria elimination efforts at village level.

2. Background: This section needs to be more concise and relevant to the study conducted and aims addressed. For instance, the authors exhaustively discuss the role of several variables (migration, population density, temperature, etc.), none of which are considered in the present study. The authors need to focus on outlining the wider context, gaps in knowledge/evidence and then introduce the present research and how it addresses those gaps.
Response: Thank you for your advice. We revised this section as suggested. There is an overall very diverse malaria prevalence distribution with remote areas showing the highest prevalence [30]. Different factors affect malaria transmission within the province [16, 17, 31], and it is important to differentiate between factors that influence the vector, the parasite and the host-vector relationship since specific meteorological, environmental factors are at interplay [32]. Atieli et al. have demonstrated that topographic variables such as elevation, slope, and aspect are influencing the development of Anopheles mosquitoes [33] There is a significant association between local spatial variations like population density, lowland location in north-eastern Venezuela, and proximity to aquatic environments with malaria transmission [34]. In our study, the ecological potential to predict the response variable „malaria incidence" (Y) are altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4). Also, distance from the forest (X5) and rainfall (X6) that locally different as a variable of research.
3. Methods: How many primary health centres reported malaria case data? And what is the level of completeness of this data? How does the malaria case data from the primary health centres become village level data? Was the analysis at the village or health facility level?
Response: The malaria diagnostic data were obtained from the regular health information reporting system of the Provincial Health Office of South Sumatra. The data had been collected during 12 months (January to December 2013) at the village level from patients seeking treatment in PHC, locally called Pusat Kesehatan Masyarakat ("puskesmas"), and that were reported monthly to the Provincial Health Office via the malaria programs in the District Health Offices.
The analysis is based on village level.
We noted 140 primary health care reported malaria case data in the study area. The patients are categorised into "clinical diagnosis", "suspected malaria" and "positive malaria". Categories "clinical diagnosis" or "suspected malaria" are based on the patient's symptoms and physical findings at examination. "Positive malaria" is a case of malaria diagnosed microscopically (examination of blood specimen/preparation) or rapid diagnosis test (RDT) with positive results for Plasmodium. Either RDT or microscopic assessment or both were used to confirm the diagnosis of malaria.
Reporting of malaria incidence allowed the calculation of Annual Parasitic Incidence (API) that is the number of positive cases per 1,000 total population.
4. Methods: Authors state that "In the study region and period, 2,787,954 of the total population and 36,372 research participants visited hospitals or PHCs due to suspected malaria fever". Elsewhere, authors state "The study population was the number of participants who were suspected of having malaria while the sample was the number of participants with laboratory-confirmed malaria." It is unclear what the authors mean by study population, sample, research participants, and total population.
Response: The population of our study was a total village in 8 provinces South Sumatra, Indonesia. The sample was a village with a malaria case and together with location, this village is our unit ("toponym").
In total, 3,578 patients were laboratory positive for malaria. The malaria cases were spread over 436 out of 1,613

villages in 8 endemic malaria districts of South Sumatra Province.
5. Methods: Were multiple episodes from the same individual included? Or was the analysis based on single malaria episodes? As there can be potential biases from relapses especially from P.vivax.
Response: For each patient who visit a PHC, there is a unique patient data form which was filled out. So, there could be make a decision if cases were new or relapsed. Based on policy from the ministry of health, each patient who has diagnosed malaria positive should have had an epidemiology investigation. Case-finding activities were carried out passively (patients arrived at health-care facilities) and actively by mass blood survey and contact surveys r epidemiological investigations.
6. Line 96: The authors discuss locations of cases in each district. Is this the location of the primary health centre they sought care, or the location of their residence?

Response:
The malaria case data entered into the model has been aggregated to village level data with the village centroid used as spatial unit.
7. Methods: Authors included several distance variables - it is unclear whether these are distances from the village of residence to the attribute of interest (river, forest, etc.) or distances from the primary health centre.
Response: The distance in this paper meant was the distance of case (village) to the variable.
8. Methods: Was any validation of the OLS or GWR models conducted. For example, cross-validation or bootstrapping? And what was the impact on the results?
Response: The model validation procedure was conducted as following: Step1: Preparation of dataset. Step 2: Specify one regression type and the variable settings needed to determine the GWR model. We chose Geographical variability test, for model coefficient test obtained. Step 3: Choosing a geographic kernel type and its optimum bandwidth based on Selection Criteria. In this paper, we demonstrated an “Adaptive bi-square kernel” and selection bandwidth use “Golden section search” then use AIC criteria and residual sum of square. Step 4: Specify filenames for the files storing the modelling results, and Step 5: Execute the session to compare necessary calculations and read results. When the model is fit with the geographical variability test, the adaptive kernel function, the golden section search for finding the optimal bandwidth size, and AIC as the model indicator for selecting the optimal bandwidth. We demonstrated OLS assumptions with Durbin Watson coefficient, and we found value .092, hence the assumption of independence was fulfilled. Besides, diagnostic regression multicollinearity has been done before the modelling. We show that multicollinearity does not occur, because the VIF value is less than 10 and the tolerance value is greater than 0.1.
9. Methods: it is unclear how the outcome malaria incidence was defined as there was no mention of village size or population, and also unclear whether this was at the primary health centre level or the village level?

Response: In this study, a village was the geographic unit..

10. Methods: one requirement for an OLS is that only statistically significant explanatory variables are included. However, it seems that the OLS model used by authors included several variables that were not significant.
Response: Although the variable studies that are Distance from the river (X3) and Distance from lakes and pond (X4) is not statistically significant, we choose to investigate a full OLS model for following reason: The independent variable has a relationship in substance with the dependent variable. The independent variables are significant at some specific places at the local level analysis. That means if the independent variable is not involved in GWR analysis, we will lose critical information. We show in OLS model that a decline of altitude, aspect and distance to forest and an incline of rainfall are risk factors for getting malariacase. In this model, we show Akaike's information criterion and Bayesian information criterion better than the full model.

```
. *Full model
. regress cases altitude aspect distfriv distflak distffor rainfall , vce(robust) level(95)
Linear regression Number of obs = 436
F(6, 429) = 5.30
Prob > F = 0.0000
R-squared = 0.0615
Root MSE = 15.315

------------------------------------------------------------------------------
| Robust
cases | Coef. Std. Err. t P>|t| [95% Conf. Interval]
-------------+----------------------------------------------------------------
altitude | -.0154009 .0032999 -4.67 0.000 -.0218869 -.0089148
aspect | -.0137931 .0070601 -1.95 0.051 -.0276698 .0000835
distfriv | -.0011115 .0009416 -1.18 0.238 -.0029622 .0007392
distflak | .0000782 .0002092 0.37 0.709 -.000333 .0004893
distffor | -.0004079 .0001369 -2.98 0.003 -.000677 -.0001388
rainfall | .0038088 .0015844 2.40 0.017 .0006946 .006923
_cons | 7.976743 3.896792 2.05 0.041 .3175618 15.63592
------------------------------------------------------------------------------
r; t=0.43 13:22:10
. *partial model
. regress cases altitude distffor rainfall , vce(robust) level(95)

Linear regression Number of obs = 436
```

F(3, 432) = 8.92
Prob > F = 0.0000
R-squared = 0.0540
Root MSE = 15.323

```
---------------------------------------------------------------------------
| Robust
cases | Coef. Std. Err. t P>|t| [95% Conf. Interval]
-------------+--------------------------------------------------------------
altitude | -.0151679 .0031585 -4.80 0.000 -.0213758 -.0089599
distffor | -.0004035 .0001374 -2.94 0.003 -.0006736 -.0001334
rainfall | .0040303 .0015094 2.67 0.008 .0010637 .0069969
_cons | 4.680717 3.512194 1.33 0.183 -2.222397 11.58383
---------------------------------------------------------------------------
r; t=0.02 13:22:10
```

Akaike's information criterion and Bayesian information criterion

```
---------------------------------------------------------------------------
Model | Obs ll(null) ll(model) df AIC BIC
-------------+-------------------------------------------------------------
full | 436 -1818.753 -1804.908 7 3623.816 3652.359
sub | 436 -1818.753 -1806.657 4 3621.314 3637.625
---------------------------------------------------------------------------
Note: N=Obs used in calculating BIC; see [R] BIC note.
r; t=0.01 13:22:10.
end of do-file
```

11. Methods: lines 123 - 169 go into an exhaustive explanation of the GWR, and OLS approaches, while some detail is important, this much information seems to shift the focus of the paper to one on methodological approaches and distracts from the stated aim to "use global and local spatial modelling to analyses.

Response: Thank you for your advice. We comprehensively revised the method section.

12. The environmental risk factors for malaria in South Sumatra that vary geographically at the regional level."

Response: We deleted this sentence.

13. Methods: why were other variables such as village size/health facility catchment area size, household density, distance to health facility, coverage of malaria interventions included? Also was seasonality accounted for?
Response: We investigated physical environment variables as independent variables. Other non-physical environmental variables were not explored. We will consider other eco-bio-social variables in future studies.
14. Results: Lines 172 - 175, where is the incidence data presented? And it is still unclear what incidence refers to? Is the number of cases what is being referred to as incidence?

Response: Malaria case has been diagnosed microscopically (examination of blood specimen/preparation) or rapid diagnosis test (RDT) with positive results for Plasmodium. Either RDT or microscopic assessment or both were used to confirm the diagnosis of malaria. The malaria diagnostic data were obtained from the regular health information reporting system of the Provincial Health Office of South Sumatra. The data had been collected during 12 months (January to December 2013) at the village level from patients seeking treatment in PHC, locally called Pusat Kesehatan Masyarakat ("puskesmas"), and that were reported monthly to the Provincial Health Office via the malaria programs in the District Health Offices.

15. Table 1: From a statistical perspective, the OLS model should only include variables with significant coefficients, and that are in the expected direction.
Response: Thank you for your comment. The tables were changed accordingly.
16. Table 1: Please provide units and scale the variables appropriately, so the results are interpretable. For instance, distance from the forest has a coefficient of 0.00 which cannot be interpreted.

Response: Thank you for your comment. The tables were changed accordingly.

17. Results: Lines 186 - 189: authors conclude that malaria incidence is more common in regions with high rainfall and areas adjacent to forest areas. However looking at the coefficients presented in Table 1, distance from forest area has a positive coefficient, meaning that as the distance from forest area increases malaria incidence increases. Please clarify.
Response: The explanatory variables altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4) are locally different. Also, distance from the forest (X5) and rainfall (X6) have different strengths to predict the response variable „malaria case‟ (Y). The global OLS model revealed that altitude, distance from lakes and pond, and distance to forest have a negative coefficient and rainfall has a positive coefficient, and significantly influence malaria case

Table 3: The result of global regression model and geographical variability test of local coefficients for six environmental factors.
Variables Global regression model output Geographical variability test
Estimate SE T value P value F DOF for F test DIFF of Criterion
Intercept 7.98 4.63 1.72 0.04 33.20 10.48 261.38 -347.99
"Altitude (X1)" -0.02 0.00 -4.03 0.00 0.24 12.02 261.38 19.19
"Aspect (X2)" -0.01 0.01 -1.60 0.05 0.55 22.68 261.38 24.91
"Distance from the river (X3)" 0.00 0.00 -0.84 0.24 1.84 18.15 261.38 -16.03
"Distance from lakes and pond (X4)" 0.00 0.00 0.39 0.71 0.90 15.04 261.38 7.99
"Distance from forest (X5)" 0.00 0.00 -3.69 0.00 2.99 14.61 261.38 -38.12
"Rainfall (X6)" 0.00 0.00 2.38 0.02 13.07 10.17 261.38 -158.91

18. Lines 200 - 201: Authors state "The regression coefficients for malaria incidence at the local level range from 0.03 to 0.99 (Fig. 8)." However, Fig 8 presents the R2 values which are different from the regression coefficients.
Response: We have corrected the local coefficient of determination (R squared) for malaria cases at the local level range between < 0.20 - 0.78.
19. Lines 201-202: Authors state "The highest influence of environmental factors on malaria incidences was found in Lahat District." It is not clear where this conclusion came from especially considering Figure 8.
Response: The statement is not related to Figure 8, however, to Figure 7.
20. Lines 202 - 207. Authors have erroneously interpreted R2 values as values of regression coefficients.

Response: Thank you for your feedback. In our understanding, R^2 is: The coefficient of determination (R Squared) that indicates the kindness of the model or the contribution of the independent variable to confirmed malaria cases. R^2 is a statistical measure of how close the data are to the fitted regression line. It is also known as the coefficient of multiple determination for multiple regression. The regression coefficient is the constant (a) that represents the rate of change of one variable (y) as a function of changes in the other (x); it is the slope of the regression line.

21. Discussion: Authors state that their "analyses have identified Lahat as the South Sumatran district in which environmental factors were of greatest relevance for malaria incidence." Caution is needed in making such conclusions especially given that the small village level sample sizes (Fig 3). Inability to detect significant relationships may, in fact, be related to the small sample sizes.

Response: Thank you for your advice. The highest confirmed malaria cases with 1,449 cases spread over 124 villages were found in Lahat District. Based on local geographical variability tests of coefficients, we demonstrated that the independent variables significantly revealed spatial variability or local spatial heterogeneity (altitude, distance from lakes and pond). The global OLS model revealed that altitude, distance from lakes and pond, and distance to forest and rainfall significantly influence confirmed malaria cases.

22. Discussion: much of the discussion is very anecdotal and not directly related to the findings presented. For instance, the authors discuss the relevance of deforestation and distance to coal mines, none of which was assessed in the present study.

Response: Thank you for your advice. We consider your comment no. 23 to 25 and revised text in discussion chapter accordingly. In accordance, we now focus on locally different altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4), distance from forest (X5) and rainfall (X6) that different strengths to predict the response variable „confirmed malaria cases (Y).

23. Discussion: authors should avoid introducing new data in the discussion. For instance, authors discuss distance between coal mines and local plantations and forests in Lahat District (lines 254-255). Elsewhere authors state "temperature was correlated with altitude and humidity…".
Response: Thank you for advice. We revised the text in discussion chapter and focus on our explanatory variables.
24. The topic distance between coal mines and local plantations and forests in Lahat District (lines 254-255).

Response: On average, we observed distances of 200-700 m between the coal mines and local plantations and forests in Lahat District (M. Alam, unpublished data). Elsewhere, the distance of households from a forest and the borders of swamps have often been associated with the risk of malaria infection [35].

References.
1. Fotheringham A, Brunsdon C, Charlton M. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. University of Newcastle, UK: John Wiley & Sons. Retrieved from https://www.wiley.com/, isbn:0-471-49616-2
2. Rodrigues M, de la Riva J, Fotheringham S: Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. Applied Geography 2014; 48:52-63. doi:
3. Nakaya T. (24 March 2016). GWR4 Windows Application for Geographically Weighted Regression Modelling. Kyoto: Ritsumeikan University, Department of Geography. Retrieved from https://raw.githubusercontent.com/gwrtools/gwr4/master/GWR4manual_409.pdf,
4. Halimi M, Farajzadeh M, Delavari M, Takhtardeshir A, Moradi A: Modelling spatial relationship between climatic conditions and annual parasite incidence of malaria in southern part of Sistan & Balouchistan Province of Iran using spatial statistic models. Asia Pac J Public Health 2014; 4:S167-S72. doi: 10.1016/S2222-1808(14)60434-5

5. Wheeler D, Tiefelsdorf M: Multicollinearity and correlation among local regression coefficients in geographically weighted regression. J Geogr Syst 2005; 7:161-87. doi: 10.1007/s10109-005-0155-6
6. Rabe-Hesketh S, Everitt B: Handbook of statistical analyses using stata. edn.: CRC Press; 2003, isbn:1439832358
7. Kothari P: Data Analysis with Stata. edn.: Packt Publishing Ltd; 2015, isbn:1782173188
8. Brunsdon C, Fotheringham A, Charlton M: Geographically weighted summary statistics—a framework for localised exploratory data analysis. Comput Environ Urban Syst 2002; 26:501-24. doi: 10.1016/S0198-9715(01)00009-6
9. Elyazar IR, Sinka ME, Gething PW, Tarmidzi SN, Surya A, Kusriastuti R, et al. (2013). The distribution and bionomics of anopheles malaria vector mosquitoes in Indonesia. (Vol. 83, pp. 173-266). Academic Press, Elsevier. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23876873, isbn:978-0-12-407705-8
10. Steffen W, Persson Å, Deutsch L, Zalasiewicz J, Williams M, Richardson K, et al: The anthropocene: from global change to planetary stewardship. Ambio 2011; 40:739-61. doi: 10.1007/s13280-011-0185-x
11. Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li W, Nobre CA: Climate change, deforestation, and the fate of the Amazon. Science 2008; 319:169-72. doi: 10.1126/science.1146961
12. Olson SH, Gangnon R, Silveira GA, Patz JA: Deforestation and malaria in Mancio Lima county, Brazil. Emerg Infect Dis 2010; 16:1108. doi:
13. Patz JA, Olson SH: Climate change and health: global to local influences on disease risk. Ann Trop Med Parasitol 2006; 100:535-49. doi:
14. Badan Pusat Statistik. (2014). Sumatera Selatan dalam angka (Sumatera Selatan in figures) 2013. Sumatera Selatan: BPS – Statistics Sumatera Selatan Retrieved from http://sumsel.bps.go.id.
15. Hoshen MB, Morse AP: A weather-driven model of malaria transmission. Malar J 2004; 3:32. doi: 10.1186/1475-2875-3-32
16. Ministry of Health Republic of Indonesia. (2014). Malaria management: guideline. Jakarta: Directorate of Vector Borne Disease and Zoonosis Control, Directorate General of Disease Prevention and Control, Ministry of Health (MoH) of Indonesia Retrieved from http://eliminasimalaria.blogspot.de/p/download.html.
17. Elyazar IR, Hay SI, Baird JK. (2011). Malaria distribution, prevalence, drug resistance and control in Indonesia. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21295677, isbn:2163-6079
18. Guerra CA, Hay SI, Lucioparedes LS, Gikandi PW, Tatem AJ, Noor AM, et al: Assembling a global database of malaria parasite prevalence for the Malaria Atlas Project. Malar J 2007; 6:17. doi: 10.1186/1475-2875-6-17
19. Kleinschmidt I, Sharp BL, Clarke GP, Curtis B, Fraser C: Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in Kwazulu Natal, South Africa. Am J Epidemiol 2001; 153:1213-21. doi: doi.org/10.1093/aje/153.12.1213
20. Kleinschmidt I, Bagayoko M, Clarke G, Craig M, Le Sueur D: A spatial statistical approach to malaria mapping. Int J Epidemiol 2000; 29:355-61. doi:
21. Kleinschmidt I, Omumbo J, Briet O, Van De Giesen N, Sogoba N, Mensah NK, et al: An empirical malaria distribution map for West Africa. Trop Med Int Health 2001; 6:779-86. doi:
22. Diggle PJ, Giorgi E: Model-Based Geostatistics for Prevalence Mapping in Low-Resource Settings. J Am Stat Assoc 2016; 111:1096-120. doi: 10.1080/01621459.2015.1123158
23. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, et al: A world malaria map: Plasmodium falciparum endemicity in 2007. PLoS Med 2009; 6:e1000048. doi:
24. Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IR, Johnston GL, et al: A new world malaria map: Plasmodium falciparum endemicity in 2010. Malar J 2011; 10:378. doi: 10.1186/1475-2875-10-378
25. Noor AM, Gething PW, Alegana VA, Patil AP, Hay SI, Muchiri E, et al: The risks of malaria infection in Kenya in 2009. BMC Infect Dis 2009; 9:180. doi: 10.1186/1471-2334-9-180
26. Indonesian Agency for Meteorological Climatological and Geophysics: Informasi Hujan Bulanan dan Informasi Unsur Iklim Rata-Rata 2007 - 2013. BMKG, Climatology Station Class I Kenten Palembang2013.
27. Abood SA, Lee JSH, Burivalova Z, Garcia-Ulloa J, Koh LP: Relative contributions of the logging, fiber, oil palm, and mining industries to forest loss in Indonesia. Conserv Lett 2015; 8:58-67. doi: 10.1111/conl.12103
28. Fotheringham AS, Brunsdon C, Charlton M: Geographically weighted regression: the analysis of spatially varying relationships. edn.: John Wiley & Sons; 2003, isbn:0470855258
29. O'brien RM: A caution regarding rules of thumb for variance inflation factors. Quality & Quantity 2007; 41:673-90. doi:
30. World Health Organization. (2017). Regional action framework for malaria control and elimination in the Western Pacific: 2016-2020. Manila, Philippines WHO Regional Office for the Western Pacific. Retrieved from http://iris.wpro.who.int/handle/10665.1/13578, isbn:9789290618157
31. WHO Country Office for Indonesia and Ministry of Health Indonesia. (2013). National malaria control programme review: Republic of Indonesia. Jakarta: WHO Country Office for Indonesia and Ministry of Health, Indonesia. Retrieved from http://www.who.int/iris/handle/10665/253960, isbn:978 979 19477 4 9
32. Loha E, Lindtjorn B: Model variations in predicting incidence of Plasmodium falciparum malaria using 1998-2007 morbidity and meteorological data from south Ethiopia. Malar J 2010; 9:166. doi: 10.1186/1475-2875-9-166
33. Atieli HE, Zhou G, Lee M-C, Kweka EJ, Afrane Y, Mwanzo I, et al: Topography as a modifier of breeding habitats and concurrent vulnerability to malaria risk in the western Kenya highlands. Parasit Vectors 2011; 4:241. doi: 10.1186/1756-3305-4-241
34. Grillet ME, Barrera R, Martinez JE, Berti J, Fortin MJ: Disentangling the effect of local and global spatial variation on a mosquito-borne infection in a neotropical heterogeneous environment. Am J Trop Med Hyg 2010; 82:194-201. doi: 10.4269/ajtmh.2010.09-0040
35. Herdiana H, Cotter C, Coutrier FN, Zarlinda I, Zelman BW, Tirta YK, et al: Malaria risk factor assessment using active and passive surveillance data from Aceh Besar, Indonesia, a low endemic, malaria elimination setting with Plasmodium knowlesi, Plasmodium vivax, and Plasmodium falciparum. Malar J 2016; 15:468. doi: 10.1186/s12936-016-1523-z

Close

Gmail

Hamzah Hasyim <hamzah.hasyim@gmail.com>

MALJ-D-17-00578R3

4 messages

Hamzah Hasyim <hamzah.hasyim@gmail.com> 10 February 2018 at 12:00
To: Malaria Journal Editorial Office <magesh.murugappan@springer.com>

MALJ-D-17-00578R3
Spatial modelling of malaria case associated with environmental factors in South Sumatra, Indonesia
Hamzah Hasyim, PhD candidate; Afi Nursafingi, M.Sc; Ubydul Haque, PhD; Doreen Montag, DPhil; David Groneberg, Prof, Dr, PhD; Meghnath Dhimal, PhD; Ulrich Kuch, Dr; Ruth Müller, Dr
Malaria Journal

Dear Dr Magesh,
Journal Editorial Office
BioMed Central
Web: www.biomedcentral.com

I have succeeded build pdf manuscript and approval it but, I forgot to attach one important document that is track changes to the script. Please let me know, how to send the file to Malaria journal?

Thank you for your kind information.

Sincerely,

Hamzah

Murugappan, Magesh, Springer <magesh.murugappan@springer.com> 10 February 2018 at 13:08
To: Hamzah Hasyim <hamzah.hasyim@gmail.com>

Dear Dr. Hasyim,

Please provide me with the updated file so that I could upload it in the system.

Thank you very much.

With best regards,

Magesh

Magesh Murugappan (Mr.)

Journals Editorial Office (JEO)

[Quoted text hidden]

Hamzah Hasyim <hamzah.hasyim@gmail.com> 10 February 2018 at 14:59
To: Malaria Journal Editorial Office <magesh.murugappan@springer.com>

Dear Dr Magesh
Journals Editorial Office (JEO),

Thank you very much for your prompt response. I appreciate. Would you please upload in your system both clean and track changes version as an attachment

I saw in clean version previously, two double references is
Olson SH, Gangnon R, Silveira GA, Patz JA. 2010. Deforestation and Malaria in Mâncio Lima County, Brazil. Emerg Infect Dis 16:1108-1115. doi: 10.3201/eid1607.091785 and I found three of the references is not appears "doi."

Sincerely,

Hamzah,
[Quoted text hidden]

2 attachments

Track changes_Manuscript.docx
164K

Clean_Manuscript.docx
130K

Hamzah Hasyim <hamzah.hasyim@gmail.com>
To: Malaria Journal Editorial Office <magesh.murugappan@springer.com>

MALJ-D-17-00578R3
Spatial modelling of malaria case associated with environmental factors in South Sumatra, Indonesia
Hamzah Hasyim, PhD candidate; Afi Nursafingi, M.Sc; Ubydul Haque, PhD; Doreen Montag, DPhil; David Groneberg, Prof, Dr, PhD; Meghnath Dhimal, PhD; Ulrich Kuch, Dr; Ruth Müller, Dr
Malaria Journal

Dear Dr Magesh,
Journal Editorial Office
BioMed Central
Web: www.biomedcentral.com

I just saw the progress manuscript which editor assigned.

However, the document is old files meanwhile, I have attached a newspaper through your email.

The oldest file without track changes manuscript, and I saw the document still double references that are Olson appears like in

26. Olson SH, Gangnon R, Silveira GA, Patz JA: Deforestation and Malaria in Mâncio Lima 453 County, Brazil. Emerg Infect Dis 2010; 16:1108-15. doi: 10.3201/eid1607.091785, and in
54. Olson SH, Gangnon R, Silveira GA, Patz JA: 538 Deforestation and malaria in Mancio Lima 539 county, Brazil. Emerg Infect Dis 2010; 16:1108. doi:

Would you please update the document anymore in your system?

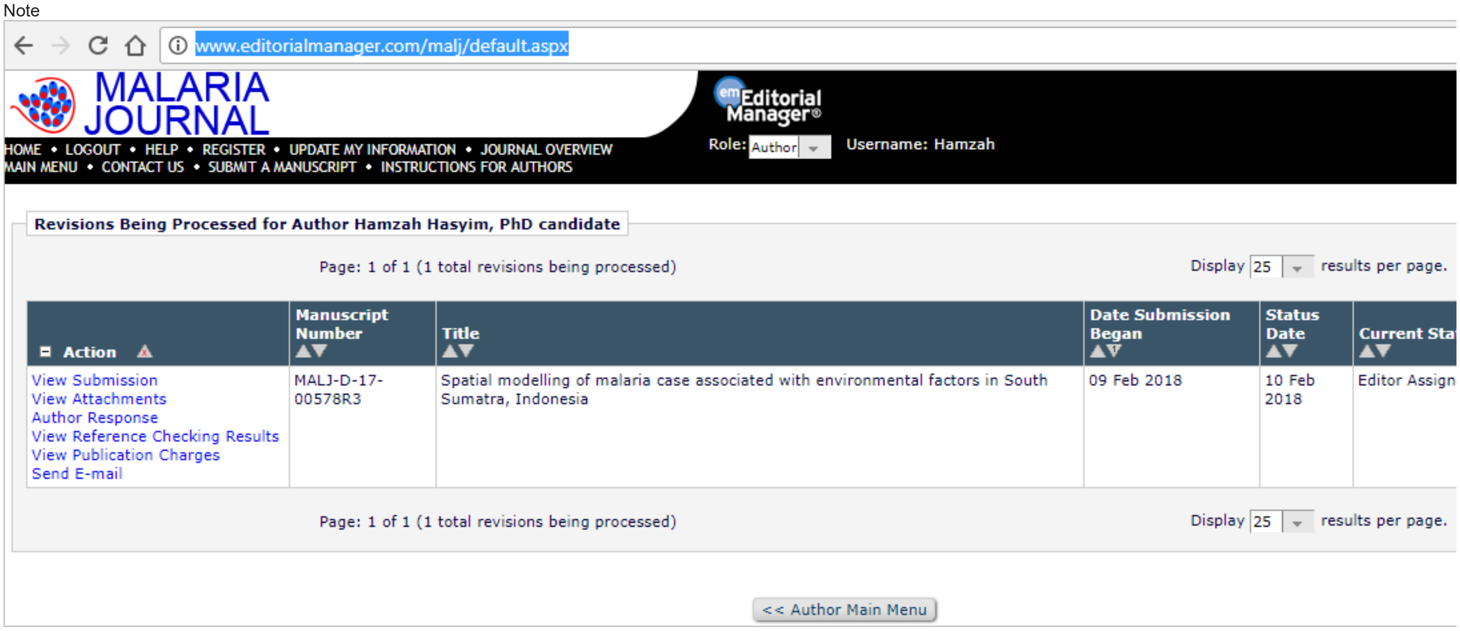
Thank you in advance for your attention to this matter and for any help you can provide.

Respectfully,

Hamzah Hasyim,
http://www.med.uni-frankfurt.de/institut/occupational-medicine/
hamzah.hasyim@stud.uni-frankfurt.de
Phone number: +4915905821418

bit.ly/weM38G
Bitte denken Sie an die Umwelt, bevor Sie diese e-Mail ausdrucken
Please consider the environment before printing this e-mail

Note



[Quoted text hidden]

MALJ-D-17-00578_R3.pdf
2706K

Spatial modelling of malaria incidence associated with environmental factors in South Sumatra, Indonesia

Hamzah Hasyim[1,2,*], Ali Nursafingi[3], Ubydul Haque[4], Doreen Montag[5], David A. Groneberg[1], Meghnath Dhimal[1,6], Ulrich Kuch[1], Ruth Müller[1]

[1]Department of Tropical Medicine and Public Health, Institute for Occupational Medicine, Social Medicine and Environmental Medicine, Faculty of Medicine, Goethe University, Frankfurt am Main, Germany

[2]Faculty of Public Health, Sriwijaya University, Indralaya, South Sumatra, Indonesia

[3]Remote Sensing Program, Faculty of Geography, Gadjah Mada University, Yogyakarta, Indonesia

[4]Department of Public Health, Baldwin Wallace University, Berea, Ohio, USA

[5]Barts and the London School of Medicine, Centre for Primary Care and Public Health, Queen Mary University of London, London, UK

[6]Nepal Health Research Council (NHRC), Ramshah Path, Kathmandu, Nepal

[§]Corresponding author: hamzah.hasyim@stud.uni-frankfurt.de, hamzah@fkm.unsri.ac.id

Formatted: Indent: Left: -0.16 cm

Abstract

Background

Malaria, a parasitic infection, is a life-threatening disease in South Sumatra Province, Indonesia. This study aimed to explore the spatial relationships between malaria occurrence and environmental risk factors which vary in studies area villages.

Methods

We analysed [PLEASE NOTE THAT MALARIA JOURNAL DOES NOT USE FIRST PERSON FORMAT; ADJUST ALL SENTENCES ACCORDINGLY]malaria incidence for the year 2013 from the routine reporting of the Provincial Health Office of South Sumatra. The cases were spread over 436 out of 1,613 villages. We investigated six potential ecological predictors of malaria incidence in the different regions. A model comparison of between ordinary least square (OLS) and geographically weighted regression (GWR) was performed to explore the global pattern and spatial variability of relationships between malaria incidence and the selected potential ecological predictors.

Results

The OLS revealed that malaria incidence is more common in regions with high rainfall, lowland and areas adjacent to forest areas. We demonstrated, by comparison, that the GWR model explains the contribution of environmentally explanatory variables to the response variable "malaria incidence" significantly better (68.7%) than the global OLS (6.2%). The GWR model could explain variance in malaria incidence at specific locations at the village level.

Conclusions

The independent variables altitude, distance from forest, and rainfall in global OLS are significantly associated with malaria incidence. In line with recent reviews, the relationship between malaria events and environmental factors in South Sumatra was found to vary spatially in different regions of each village. The GWR model is a more powerful tool for exploring spatial heterogeneity of malaria incidence in South Sumatra than OLS. A more in-depth understanding of local ecological factors influencing malaria incidence may not only be used for developing sustainable regional malaria control programmes, but can also benefit malaria elimination efforts.

Keywords Geographically weighted regression (GWR), Ordinary least squares (OLS), Akaike information criterion (AIC), Physical environment, Local climate, Sumatra, rainfall, elevation, and distance to water.

Formatted: Space Before: 0 pt

Background

Malaria is a significant public health concern worldwide, including Indonesia [1]. The Indonesian government has set a national goal to be malaria-free by 2030. Consequently, a malaria elimination programme will be carried out in the island of Sumatra, in Aceh and the Riau Island Province by 2020 [1]. Currently, 24 out of 576 districts in Indonesia classified as being malaria endemic, and an estimated 45% of Indonesia's total population are living at risk of contracting malaria. In South Sumatra Province, the malaria incidence was 0.46 per 1,000 people in 2013. In this province, the proportion of children under five years of age who applied mosquito nets was 32.7%, and the percentage of children under five who treated for fever with antimalarial medication was 89.8% in 2013 [2]. Malaria elimination has been a priority in the Millennium Development Goals (MDGs) [3], and since then has continued to be central to the Sustainable Development Goals (SDGs), supporting Indonesia's malaria elimination commitments [4]. It is now essential to generate the knowledge that is necessary to develop lasting policies for the national malaria elimination programme. Indonesia's South Sumatra Province is home to 7,828,700 inhabitants. In 2013, the Gross Regional Domestic Product (GRDP) with oil and gas was IDR 231.68 trillion (17.32 billion USD) [5], based on IDR to USD exchange rates at the time of writing. South Sumatra is an ethnically highly diverse province and home to different local languages and diverse cultural and socioeconomic practices [2]. Local people engage in coffee, rubber and palm oil plantation activities or work in the industrial mining area, which shapes not only people's lives but also the environment [6].–Indonesia contributes significantly to deforestation in Southeast Asia. Recent developments of deforestation have led to unsustainable practices which have resulted in a high frequency of deforestation in some regions and are an important factor influencing malaria incidence [7]. Deforestation has been shown to be connected with malaria incidence in the county (Município) of Mâncio Lima, Acre State, Brazil. There, a cross-sectional study shows 48%

increase in malaria incidence are associated with cumulative deforestation within respective health districts in 2006 [8].

Malaria is still an endemic problem throughout most of the tropics region. At Indonesia, malaria epidemiology is very complex because of various determinants, such as diverse mosquito bionomics, context-dependent vector behaviour, and a high diversity of local ecosystems for maintaining transmission cycles [9]. Around 20-25 species of *Anopheles* mosquitoes have been confirmed to be malaria vectors transmitting the *Plasmodium* parasite species that routinely infect humans in Indonesia [3] [10]. Currently, the main vectors of malaria are as follows: *Anopheles*~~*An.*~~ *letifer*, *Anopheles*~~*An.*~~ *nigerrimus*, *Anopheles*~~*An.*~~ *maculatus*, *Anopheles*~~*An.*~~ *sinensis*, *Anopheles*~~*An.*~~ *barbirostris*, *Anopheles*~~*An.*~~ *vagus*, and *Anopheles*~~*An.*~~ *sundaicus* in South Sumatra Provinces region. In addition, a first record of *Anopheles*~~*An.*~~ *paragenesis* from Sumatra was reported by O'Connor and Sopa ~~(1981)~~ [11].

Several meteorological and environmental variables are risk factors for malaria [12]. Since specific meteorological, environmental factors are at interplay and different factors can affect malaria transmission within a given province [3, 9, 10], it is important to differentiate between factors that influence the vector, the parasite and the host-vector relationship [13]. Atieli *et al.* have demonstrated that the topographic variables elevation, slope, and aspect are influencing the development of *Anopheles* mosquitoes [14]. In north-eastern Venezuela there is a significant association of malaria transmission with local spatial variations like population density, lowland location, and proximity to aquatic environments [15]. Elsewhere (e.g., Ethiopia and Senegal) spatial relationships between climatic variability like rainfall and malaria occurrence have been demonstrated [16]. Rainfall indirectly benefits *Anopheles* mosquitoes by increasing relative humidity which prolongs adult longevity [17], and the number of breeding places which in turn

Formatted: Font: Italic

favours population growth [18]. Temperature and the extent of water availability for larval breeding are crucial factors in the vector life-cycle, affecting transmission [3]. Vectors and parasites are both highly sensitive to any temperature changes, for example, the parasite proliferation depends on temperatures [19]. Temperatures above 28°C have been shown to reduce malaria incidence in Africa [20]. In Indonesia, the optimum temperature for malaria mosquitoes ranges between 25-27°C [3]. For the vector-host relationship, factors such as the distance of people's houses from a river, lakes, pond, distance to the regional urban centre [21-23] distance to forest [24, 25] were shown to be significant predictors. In Venezuela, GWR analysis revealed that ecological interactions that act on different scales play a role in malaria transmission and that modelling enhances the understanding of relevant spatiotemporal variability [15].

Spatial nonstationary is a condition in which a simple "global" model cannot define the relationship amongst several sets of variables [26]. Thus, we compare global OLS and local GWR modelling to analyse the environmental risk factors for malaria in South Sumatra that vary geographically at the regional level. The locally different ecological factors studied to potentially predict the response variable ~~„~~‘malaria incidence~~"~~’ (Y) are altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4), distance from the forest (X5), and rainfall (X6).

Formatted: Highlight

Methods

Study area

Formatted: Space After: 0 pt

The study area is located between 1°46' and 4°55' of southern latitude and between 102°4' and 104°41' of eastern longitude and has a total surface area of 46,377.40 km^2 (Fig. 1). It covers eight endemic malaria districts of South Sumatra, Indonesia, namely Lahat, Muara Enim, Musi Banyuasin, Musi Rawas, North Musi Rawas, Ogan Komering Ulu, South Ogan Komering Ulu, and Lubuk Linggau. The topography of the area varies from lowland to mountainous landscapes. The elevation in the study area varies between 0 to 3.150 metres above sea level (Fig. 4)(please number figures in order of citation in text). The climate is tropical and wet. In 2013 in South Sumatra, the lowest rainfall was 31 mm (August) in Lahat district, and the highest rainfall was 613 mm (March) in Palembang City. Monthly average temperatures ranged from 26.6 to 28.3°C and relative humidity from 81% to 88% in 2013 [27].

Study population and data collection

We noted 36,372 patients seeking treatment due to suspected malaria fever in 140 primary health centres (PHC) in the study region South Sumatra during January to December 2013. Among them, 3,578 were laboratory positive for malaria. The cases spread over 436 out of 1,613 villages that were used for unit analysis. The detailed number of malaria cases in different provinces are presented in Fig. 2. The spatial distribution of participants who had confirmed cases of malaria is shown in Fig. 3.

The patients are categor~~ise~~ized into "clinical diagnosis", "suspected malaria" and "positive malaria". Categories "clinical diagnosis" or "suspected malaria" are based on the patient's symptoms and physical findings at examination. "Positive malaria" is a case of malaria diagnosed microscopically (examination of blood specimen/preparation) or rapid diagnosis test (RDT) with positive results for *Plasmodium*. Either RDT or microscopic assessment or both were used to confirm the diagnosis of malaria. The malaria diagnostic data were obtained from the regular health information reporting system of the Provincial Health Office of South Sumatra. The data had been collected during 12 months (January to December 2013) at the village level from patients seeking treatment in PHC, locally called Pusat Kesehatan Masyarakat ("puskesmas"), and that were reported monthly to the Provincial Health Office via the malaria programmes in the District

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Health Offices. Data for patients who were positive for malaria included individual information (name, address, type of parasite, the treatment used) and whether the patient is a new or a relapse. Reporting of malaria incidence allowed the calculation of Annual Parasitic Incidence (API) that is the number of positive cases per 1,000 total population.

Geographic information

The study area map (Fig. 1) uses the World Geodetic System (WGS84) as its reference coordinate system. As shown in Fig. 5, we distinguish three stages of working with geographic information: data acquisition and processing, data analysis and data presentation [28]. We used GWR 4.0 version 4.0.90 and Arc GIS 10.3 for data processing, analysis, and visual~~isa~~ization. We used data from the Provincial Health Department, Ministry of Health (malaria incidence, see previous paragraph) as well as topographic (toponymy map, hypsographic map, hydrographic maps, land cover map) and climate data (rainfall map). The primary spatial data were obtained from a topographical map of Indonesia (cartographic material) which has a scale of 1:50,000 and consists of several layers of plots grouped. The malaria input data is aggregated village level data with the village centroid used as the spatial unit. This map consisted of a collection of geographic data presented as thematic layers for land cover, hydrographic data and a sheet of hypsography. Indonesian topographic map known as Peta Rupabumi Indonesia (RBI) was updated in 2014. In 2013, topographic data visual~~isa~~ization has been changed into geodatabase cartography to reduce the steps of creating cartography visual~~isa~~ization in topographic mapping activity [29]. These maps were obtained from the Geospatial Information Agency (BIG) of Indonesia. The forest cover maps were extracted from the land cover map in 2013 on the scale of 1:250.000. The map was sourced from Ministry of Environment and Forestry, Indonesia. The precipitation map (annual average) was obtained by inserting the data of average yearly rainfall from BMKG Climatological Station

Formatted: Space After: 0 pt

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Class I in Palembang, South Sumatra, Indonesia. The distance between weather observations stations was 50-100 km in flat topography and 10 km in hilly terrain.

Data pre-processing

We created the malaria distribution map (Fig. 3) and plotted six selected explanatory variables (Fig. 4). The altitude map was obtained by interpolation and contouring of the map into a Digital Elevation Model (DEM). Subsequently, the DEM data was converted into a map containing the direction of the slope (aspect). The parameter distance from the river, and distance from lake and pond processed from river, lakes, and ponds maps which were derived from the topographic map whereas distance from the forest processed from forest cover map. These variables were analysed using Euclidean distances. Rainfall parameter was calculated based on annual average rainfall over five years, and it was interpolated from several weather observation stations in study area. The rainfall map (isohyets map) was obtained from the scanned maps which are the result of interpolation and classified into several classes. The map needed to be rectified and digit~~ise~~ized to get a digital rainfall map.

Formatted: Highlight

Data processing and modelling

The response variable “distribution of malaria cases” and explanatory variables “altitude/aspect”, “distance from river”, “distance from lake and pond”, “distance from forest” and “rainfall” were tested for multicollinearity. Therefore, the values of all explanatory variables were extracted for each case location. We used an index based on predictive modelling variance, the variance inflation factor (VIF) [30]. Multicollinearity could occur when one independent variable was a linear function of another independent variable and previously observed in GWR modelling [31].

Formatted: Highlight

The pattern of connection between malaria incidence and environmental factors was expressed by the OLS method. Here, OLS model is called global regression model because the existence of local variation had not taken into account in regression so that the estimate of the regression remained constant. Thus, the regression parameters had the same value for each point within the study area. If spatial heterogeneity occurred in regression parameters, then the information that could not be processed by the global regression model was seen as an error. In such cases, the global regression model was less able to explain the actual data phenomenon [32]. A global regression coefficient value close to zero indicated that the explanatory variables had a small effect on the response variable.

As alternative, we used the GWR model. The local GWR was used to investigate the relationships between response and explanatory variables since our study area was ~~characterised~~ characterized by highly variable geography [33]. We carried out a semiparametric GWR4.09 which is a new release of the windows application software tool for modelling spatially varying relationships among variables by calibrating GWR. GWR4.09 for Windows was developed and programmed by Professor Tomoki Nakaya and team.

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

The estimated parameter of the GWR model uses the least squares given the location coordinates as a weighting factor. The influence of the points in this neighbourhood varies according to the distance to the central point [34]. The optimum distance threshold (also known as the bandwidth) or the optimum number of neighbours determined in two ways: by ~~minimising~~ minimizing the square of the residuals cross-validation (CV) or by ~~minimising~~ minimizing the Akaike Information Criterion (AIC) [35]. At this stage, we select the type of weighing (kernel type) and optimum bandwidth selection method based on selection criteria AIC. Classic AIC chooses smaller bandwidths in geographically varying coefficients are possible to be under smoothed~~.~~ [33]. In a

Formatted: Highlight

GWR context, the measurement of utility is the AIC or the corrected Akaike Information Criterion (AICc) to know whether a global regression model or GWR is most useful [34].

The local GWR model as earlier described is as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \qquad (1~~1~~)$$

Formatted: Indonesian, Do not check spelling or grammar

Based on the model, y ou, x_{ik}, (u_i, v_i), $\beta_k(u_i, v_i)$, and ε_i were sequentially the response and explanatory variables k to location i, location coordinates to i, ~~realisation~~ realization of the continuous function $\beta_k(u_i, v_i)$ at point i, and Gaussian error to location i. It is noteworthy that the kernel Fixed Gaussian function was used which highlights the optimal bandwidth found by using the Golden section search with the selection criteria AIC. Also, the Gaussian kernel supported the constant weight, and the value became less from the centre of the kernel but never touched zero. The kernel was suitable for fixed kernel because it could prevent the risk of the absence of data in the kernel. The Fixed Gaussian kernel earlier described [34] is as follows:

$$w_{ij} = exp\left[-\left(d_{ij}/b\right)^2\right] \qquad (2~~2~~)$$

Formatted: Indonesian, Do not check spelling or grammar

Also, w_{ij} was the weight value observed at the location j to approximate the calculation of the coefficients on area i, d_{ij} was the Euclidean distance between i and j, and b was the size of fixed bandwidth given by the size of metric. The Golden section automatically searched the optimal frequency range value by comparing indicators of the model with the bandwidth size. A positive R^2 indicates a positive correlation. A positive coefficient means X and Y changed in the same direction and if the environmental risk factor increased, then malaria incidence increased. Conversely, a negative coefficient means X (explanatory variable) and Y (the response variable) changed in opposite directions. Student's t distribution that had values outside the range of -1.97 and 1.97 formed a critical region with a 0.05 (95% CI) level of significance, whereas values outside the range of -2.59 and 2.59 formed critical regions with a 0.01 (99% CI) level of significance. Step-wise computation performed with these data is shown in the flowchart Fig. 5.

The locally weighed R^2 between the observed and fitted values has been calculated to measure how well the model replicates the local malaria incident values around each observation. A variable is correctly clarified for each location by the model if $R^2 = 1$ with values ranging from 0 to 1.

To compare the performance between global OLS and local GWR, we also used GWR4 software. We performed an ANOVA testing the null hypothesis that the GWR model represents no improvement over a global model. For local GWR, the sufficient number of degrees of freedom was a function of the bandwidth.

Formatted: Highlight

Formatted: Highlight

Results

Formatted: Space After: 0 pt

Data pre-processing

Multicollinearity does not occur, because the VIF value is less than 10 and the tolerance value is higher than 0.1.

Formatted: Space After: 0 pt

Environmental factors influencing malaria incidence at global level: OLS model

The global OLS model reveals that altitude and distance to the forest (negative coefficients) and rainfall (positive coefficient) significantly influence malaria incidence. Malaria incidence is more common in regions with high rainfall, lowland and areas adjacent to forest. On the other hand, environmental factors such as aspect or direction towards the slope, distance from the river, and the distance from lakes and pond do not have any significant association with malaria incidence. Based on OLS full model, each the variables used to assess dependent variable where each factor has a different predictor of malaria incident preferences in GWR model stage.

Formatted: Space After: 0 pt

Environmental factors influencing malaria incidence at local level: GWR model

We show the results of GWR using Fixed Gaussian in Table 1. The best bandwidth generates 9,184 neighbours and a significant spatial relationship with a specific region has been found. The GWR model provides evidence for a locally different influence of environmental factors on malaria incidence as shown by varying parameter estimate value (Fig. 6). “Altitude” and “distance from lake and pond” show a positive association and “aspect” a negative association with malaria incidence in the Northern study area (Musi Banyuasin). “Rainfall” and “distance from river” show a positive association with malaria incidence in the Eastern part of Musi Rawas and Lahat. The variables “aspect”, distance from lake and pond” and “distance from forest” are positively associated with malaria incidence in large parts of the study area. The significance thresholds of explanatory variables according to Student's t-test in the GWR model are shown in Fig. 7.

Formatted: Highlight

The local coefficient of determination (local R^2) for malaria incidence at the local level range between 0.18 - 1 (Fig. 8).

Comparison between the two methods OLS and GWR

Like OLS, GWR is a statistical model that provides insights into the relationship between the dependent variable malaria incidence and six independent explanatory variables. GWR is selected as best model based on the residual sum of square, classic and corrected AIC, and the R^2 as stated in Table 2.

The global regression model indicates that the variables have some influence on the study area (Table 3). The global OLS model explains 6.2% variation of malaria incidences by environmental factors (R^2 = 0.06). This implies that 93.8% of the malaria incidence is caused by unknown environmental factors related to local variation which are not taken into account in the OLS model [34]. The local GWR explained 68.7% variation in malaria incidences (Y) by environmental factors (R^2 = 0.69). The DIFF criterion indicates that the spatial distribution of malaria incidence is associated with the independent variables “altitude”, “distance from lakes and pond”, “distance from forest”, and “rainfall” with local spatial heterogeneity (Table 3). Though the testing of local coefficients for “aspect” and “distance from river” suggests no spatial variability (Table 3).

The GWR model explains the relationship between the response variable „‘malaria incidence”’ and six explanatory variables significantly better than the global regression model OLS (F = 2.12, $P < 0.05$). The best model weights are automatically determined for each location and are mapped in Figure 7.

Discussion

Formatted: Space After: 0 pt

The global OLS model revealed that altitude, distance to forest, and rainfall significantly influence malaria incidence in South Sumatra. GWR technique extends the traditional use of global regression models by allowing calculation of local regression parameters and estimation of spatial heterogeneity [36]. GWR has indeed been selected as best model to explain the association of malaria incidence with environmental factors in South Sumatra.

The significant environmental factors to malaria incidence malaria vary strongly at the local level. This finding is consistent with those obtained in studies in Ethiopia (Addis Ababa), the Amazon region of Brazil (Rondôia), and Cambodia [16, 37, 38]. Similarly, independent variables land use, humidity, altitude and rainfall have been identified by GWR to determine the regional vulnerability to malaria in Purworejo, Indonesia [39]. In Mali, the analysis of the residuals of the geo-statistical model in order to identify potential ecological predictors of spatial pattern of malaria at a local level using climatic, population and topographic variables is an important tool for malaria prediction locally [40]. In the Like in the highlands of western Kenya, topographic parameters could be used to identify the risk of malaria and thereby help to improve malaria monitoring or targeted malaria control activities [14].

The topographic wetness index (Cohen et al.) method will be considered for further research [41]. Globally, *Anopheline* species diversity and density decline from the lowlands to highlands [42]. Accordingly, poor villagers living in forested lowland areas in Papua, Indonesia, were found to be at higher risk of malaria infection than those in the highlands [43]. In contrast, a positive correlation between altitude and the abundance of *Anopheles* mosquitoes has observed in the highlands of Ethiopia, Colombia and Ecuador, particularly in warmer years [44-46]. This observation may be related to the direction towards the slopes as the distribution and density of mosquito populations may be affected by wind direction [47]. Interestingly in an Ethiopian study, minimum temperatures were significantly associated with malaria cases in cold areas, while precipitation was associated with transmission in hot areas [48]. Next to climatic and environmental factors, distance of houses to a forest are interrelated through anthropogenic activities influencing the local and regional climate [49, 50]. This can be confirmed for distance to lake, pond and forest by our data for South Sumatra. *Anopheles (Cellia) leucosphyrus* ~~Dönitz~~ is considered to be of epidemiological importance for malaria transmission in forested areas of Sumatra (McArthur, 1951). References should be cited in the text using consecutive numbers in square brackets starting from [1] then [2] etc. The *Anopheles* species was

Formatted: p1, Left
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Highlight
Formatted: Font: 12 pt
Formatted: Font: 12 pt, Highlight
Formatted: Font: 12 pt

reported from eight sources at 47 independent sites. The first record of *Anopheles parangensis* from Sumatra was reported by O'Connor and Sopa (1981), but with no details on location [11]. However, in current research, we did not investigate the main *Anopheles* vector diversity in each study area.

Formatted: Highlight

Formatted: Font: Helvetica

In accordance to many studies, malaria incidence was significantly associated with rainfall in villages of South Sumatra. Rainfall showed correlation with the incidence of clinical malaria cases in Tubu village, Botswana [51]. Variations in monthly rainfall in rural Tanzania were largely associated with malaria [52]. Rainfall creates oviposition sites for female mosquitoes, whereas humidity is a key parameter for adult mosquito daily survival [53]. *Anopheline* mosquitoes require stagnant water to complete their larval and pupal development. Thus, rainfall affects the transmission of malaria by providing water to create aquatic habitats. Malaria incidence was significantly positively connected with higher winter rainfall, but also with a higher average maximum temperature and significantly negatively associated with increasing distance from water bodies in South Africa [54]. Southern Africa Development Community estimates the positive correlation between increasing rainfall and the number of cases in Botswana during 2013 and 2014 [55].

The statistical approaches, both global and local analysis, have been implemented to predict malaria incidence in connection with potential ecological predictors. For example OLS analysis modelling and geostatistics with climatic, population and topographic variables have been applied in Mali [40]. A ~~generalised~~ generalized linear mixed model and a Gaussian spatial model was used for prevalence mapping in Nyanza Province, Kenya [56]. Geostatistical Bayesian model [57] is an appropriate strategy for the control of malaria in Kenya [58].

Our analyses have identified Lahat as the South Sumatran district in which environmental factors were of greatest relevance for malaria incidence. Lahat District has both lowland and mountain regions and is home to diverse ethnic groups, such as the Gumai who live along the rivers of the highland areas [59]. Climate data are frequently used to predict for the spatial, seasonal and interannual variation for malaria transmission, for example the dynamic malaria model forecasting malaria prevalence with seasonal climate published by Hoshen and Morse [60].

One of the key activities for malaria elimination should be the establishment of systems and tools to reduce disease burden where local transmission is high. By comparing the local GWR model with the global OLS model, we found that GWR yielded new information about the spatial variation of malaria incidence and thereby better explanations for local phenomena. The variability of predicted malaria rates in our study was due to climatic and local differences, and malaria incidence did not follow a single model in all locations [13]. Based on our findings, GWR should be used as a diagnostic model discovering spatially varying relationships between malaria prevalence and environmental factors. The use of GWR allows the uncovering of significant environmental variation for malaria incidence, which has previously been unobservable in a specific location [61].

Limitations of ~~Research~~research

Due to practical constraints, this study was unable to encompass the entirety of environmental factors, particularly climate parameters, temperature and humidity, for which only limited data were available and hence not-representative data could not be included. Also the factor land use was eliminated and malaria location information was plotted using a village centre approach which ignored all other locations where actual infections may have occurred (e.g., forests, plantations, lakes, rivers). The number of positive malaria per village, did not include the specific coordinates of each positive malaria case and thus, each positive case was placed in the centre of the settlement. Therefore, if land use variables would be involved, there will very likely be a strong bias. However, these eliminated or uninvestigated variables may be correlated with existing variables, for example, the temperature connected with altitude and with aspect or direction of the slope. In the same way, land use may be associated with the distance from the river and the distance from lakes and ponds. Thus, although these parameters (temperature, humidity, land use) had eliminated before analysis, these environmental factors were represented by our chosen set of variables. In the future, additional explanatory variables should be addressed to provide a comprehensive review of malaria incidence in our study area. It should comprise, for example, the ~~behavior~~ behaviour of mosquito vectors and that of community members, the access to and the delivery of health services, and other eco-bio-social factors that affect the incidence of malaria. Despite these limitations, our study sheds light on relevant, not only in regional but also local realities regarding environmental variation and sociocultural practice which might interplay with vector-host relationships and provide a suitable environment for malaria mosquitoes.

Conclusion

In the present study, we applied two different statistical approaches to study the importance of environmental parameters for malaria incidence. We conclude that the independent variables altitude, distance from forest, and rainfall in global OLS are significantly associated with malaria incidence. In line with recent reviews, the relationship between malaria events and environmental factors in South Sumatra was found to vary spatially in different regions of each village. A more in-depth understanding of local ecological factors influencing malaria incidence may not only be

Formatted: Highlight

Formatted: Space After: 0 pt

Formatted: Highlight

Formatted: Highlight

used for developing sustainable regional malaria control ~~programs~~programmes, but can also benefit malaria elimination efforts.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We express our gratitude to the Ministry of Research, Technology and Higher Education of the Republic of Indonesia for supporting the first author (HH) with a PhD scholarship in the context of the cooperation between Sriwijaya University and Goethe University. The authors wish to thank the Head of the Geospatial Information Agency (BIG) Indonesia for access to the digit~~ise~~ized map, and the Head of the Indonesian Agency for Meteorological, Climatological and Geophysics (BMKG) climatology station Class I, in Kenten Palembang, for providing interpolated data of annual rainfall averages. We are also grateful to the Head and staff of the Health Office of South Sumatra Province who kindly permitted for us to do further analyses on the malaria data from the study area.

Ethics approval and consent to participate

Not applicable.

Availability of data and materials

We analysed data for the year 2013 from the routine reporting of the Provincial Health Office of South Sumatra. A precipitation map was obtained by interpolating the data of annual average rainfall which got from BMKG, Station Class I, in Kenten, Palembang. We got authorization for the use of the topographical map of Indonesia from the Indonesian Geospatial Information Agency. However, restrictions were put to use the availability of these data and therefore are not publicly available. Data were collected by creating a research protocol which is used under license for the current study. The data that backs the findings of the research are served in the main paper.

Authors' contributions

HH was responsible for the management of this study, design and collection of data. Under the supervision of HH, NA performed the data analysis and was responsible for data acquisition, pre-processing, and processing. HH, UH, DM, MD, DAG, UK and RM contributed to the interpretation and visual~~isa~~ization of the results. HH, UH, DM, MD, NA, UK and RM wrote the paper. All authors read and approved the final version of the manuscript.

References

1. Kementerian Kesehatan Republik Indonesia. (2009). Keputusan Menteri Kesehatan Republik Indonesia Nomor 293/MENKES/SK/IV/2009 tentang Eliminasi malaria di Indonesia Jakarta: Kementerian Kesehatan Republik Indonesia. Retrieved from http://perpustakaan.depkes.go.id/.

I wonder if it would not be better to provide an English translation of the title, while indicating that the text is in Indonesian ?

2. Kementerian Kesehatan Republik Indonesia. (2014). Profil kesehatan Provinsi Sumatera Selatan tahun 2013. Palembang: Dinas kesehatan Provinsi Sumatera Selatan Retrieved from http://www.depkes.go.id/.
3. Kementerian Kesehatan Republik Indonesia. (2014). Pedoman manajemen malaria. Jakarta: Direktur Jenderal Pengendalian Penyakit dan Penyehatan Lingkungan, Kementerian Kesehatan Republik Indonesia, Retrieved from http://eliminasimalaria.blogspot.de/p/download.html.
4. ~~World Health Organization~~WHO and UNICEF. ~~(2015).~~ Achieving the malaria MDG target: reversing the incidence of malaria 2000-2015. Geneva: World Health Organization, 2015. Retrieved from http://www.who.int/malaria/, isbn:978 92 4 150944 2
5. Badan Pusat Statistik. (2014). Sumatera Selatan dalam angka 2013. Sumatera Selatan: BPS – Statistics Sumatera Selatan Retrieved from http://sumsel.bps.go.id.
6. Lee JSH, Abood S, Ghazoul J, et al: (please provide name of 6 first authors, then et al.) Environmental impacts of large-scale oil palm enterprises exceed that of smallholdings in Indonesia. Conserv Lett 2014; 7:25-33. doi: 10.1111/conl.12039
7. Abood SA, Lee JSH, Burivalova Z, et al: Relative contributions of the logging, fiber, oil palm, and mining industries to forest loss in Indonesia. Conserv Lett 2015; 8:58-67. doi: 10.1111/conl.12103
8. Olson SH, Gangnon R, Silveira GA, et al: Deforestation and Malaria in Mâncio Lima County, Brazil. Emerg Infect Dis. 2010; 16:1108-~~11~~15. doi: 10.3201/eid1607.091785

Formatted: Indent: Left: -1.27 cm

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

9. WHO Country Office for Indonesia and Ministry of Health Indonesia. (2013). National malaria control programme review: Republic of Indonesia. Jakarta: WHO Country Office for Indonesia and Ministry of Health, Indonesia. Retrieved from http://www.who.int/iris/handle/10665/253960, isbn:978 979 19477 4 9

10. Elyazar IR, Hay SI, Baird JK: Malaria distribution, prevalence, drug resistance and control in Indonesia. Adv Parasitol 2011; 74:41-175. doi: 10.1016/B978-0-12-385897-9.00002-1

11. Elyazar IR, Sinka ME, Gething PW, et al. (2013). The distribution and bionomics of anopheles malaria vector mosquitoes in Indonesia. (Vol. 83, pp. 173-266). Academic Press, Elsevier. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23876873, isbn:978-0-12-407705-8

12. World Health Organization. (2017). Regional action framework for malaria control and elimination in the Western Pacific: 2016-2020. Manila, Philippines WHO Regional Office for the Western Pacific. Retrieved from http://iris.wpro.who.int/handle/10665.1/13578, isbn:9789290618157

13. Loha E, Lindtjorn B: Model variations in predicting incidence of Plasmodium falciparum malaria using 1998-2007 morbidity and meteorological data from south Ethiopia. Malar J 2010; 9:166. doi: 10.1186/1475-2875-9-166

14. Atieli HE, Zhou G, Lee M-C, et al: Topography as a modifier of breeding habitats and concurrent vulnerability to malaria risk in the western Kenya highlands. Parasit Vectors 2011; 4:241. doi: 10.1186/1756-3305-4-241

15. Grillet ME, Barrera R, Martinez JE, et al: Disentangling the effect of local and global spatial variation on a mosquito-borne infection in a neotropical heterogeneous environment. Am J Trop Med Hyg 2010; 82:194-201. doi: 10.4269/ajtmh.2010.09-0040

16. Alemu A, Abebe G, Tsegaye W, et al: Climatic variables and malaria transmission dynamics in Jimma town, South West Ethiopia. Parasit Vectors 2011; 4:30. doi: 10.1186/1756-3305-4-30

17. Yamana TK, Eltahir EA: Incorporating the effects of humidity in a mechanistic model of Anopheles gambiae mosquito population dynamics in the Sahel region of Africa. Parasit Vectors 2013; 6:235. doi: 10.1186/1756-3305-6-235

18. Abiodun GJ, Maharaj R, Witbooi P, et al: Modelling the influence of temperature and rainfall on the population dynamics of Anopheles arabiensis. Malar J 2016; 15:364. doi: 10.1186/s12936-016-1411-6

19. Weaver HJ: Climate change and human parasitic disease. In Climate change and global health. CABI Nosworthy Way Wallingford UK: ©CAB International; 2014: 95. Retrieved from https://www.cabi.org/cabebooks/ebook/20143328432.

20. Mordecai EA, Paaijmans KP, Johnson LR, et al: Optimal temperature for malaria transmission is dramatically lower than previously predicted. Ecol Lett 2013; 16:22-30. doi: 10.1111/ele.12015

21. Villarreal-Treviño C, Penilla-Navarro RP, Vázquez-Martínez MG, et al: Larval habitat characterization of Anopheles darlingi from its northernmost geographical distribution in Chiapas, Mexico. Malar J 2015; 14:517. doi: 10.1186/s12936-015-1037-0

22. Ernst KC, Lindblade KA, Koech D, et al: Environmental, socio-demographic and behavioural determinants of malaria risk in the western Kenyan highlands: a case-control study. Trop Med Int Health 2009; 14:1258-1265. doi: 10.1111/j.1365-3156.2009.02370.x

23. Haque U, Scott LM, Hashizume M, et al: Modelling malaria treatment practices in Bangladesh using spatial statistics. Malar J 2012; 11:63. doi: 10.1186/1475-2875-11-63

24. Dhimal M, O'Hara RB, Karki R, et al: Spatio-temporal distribution of malaria and its association with climatic factors and vector-control interventions in two high-risk districts of Nepal. Malar J 2014; 13:457. doi: 10.1186/1475-2875-13-457

25. Haque U, Glass GE, Bomblies A, et al: Risk factors associated with clinical malaria episodes in Bangladesh: a longitudinal study. Am J Trop Med Hyg 2013; 88:727-732. doi: 10.4269/ajtmh.12-0456

26. Brunsdon C, Fotheringham A, Charlton M: Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. 1996; 28:281-298. doi: 10.1111/j.1538-4632.1996.tb00936.x

27. BMKG. (2013). Annual average rainfall 2007 - 2013 Indonesian Agency for Meteorological, Climatological and Geophysics (BMKG) Climatology Station Class I Kenten Palembang Retrieved from www.bmkg.go.id

28. Huisman O, De By R. (2009). Principles of geographic information systems. The Netherlands: The International Institute for Geo-Information Science and Earth Observation. Retrieved from http://www.gdmc.nl/oosterom/PoGISHyperlinked, isbn:978–90–6164–269–5

29. Perdana AP, Juniati E, Hidayat J. (2013). Transformation of Topographic Data Visualization From Freehand Drawing to Cartographic Representation. Jurnal Ilmiah Geomatika, 19(2), Retrieved from http://jurnal.big.go.id/.

30. Halimi M, Farajzadeh M, Delavari M, et al: Modelling spatial relationship between climatic conditions and annual parasite incidence of malaria in southern part of Sistan & Balouchistan Province of Iran using spatial statistic models. Asia Pac J Public Health 2014; 4:S167-S172. doi: 10.1016/S2222-1808(14)60434-5

31. Wheeler D, Tiefelsdorf M: Multicollinearity and correlation among local regression coefficients in geographically weighted regression. J Geogr Syst 2005; 7:161-187. doi: 10.1007/s10109-005-0155-6

32. Yasin H: Pemilihan variabel pada model geographically weighted regression. Media Statistika 2011; 4:63-72. doi: 10.14710/medstat.4.2.63-72

33. Nakaya T. (24 March 2016). GWR4 Windows Application for Geographically Weighted Regression Modelling. Kyoto: Ritsumeikan University, Department of Geography. Retrieved from https://raw.githubusercontent.com/gwrtools/gwr4/master/GWR4manual_409.pdf,

34. Fotheringham A, Brunsdon C, Charlton M. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. University of Newcastle, UK: John Wiley & Sons. Retrieved from https://www.wiley.com/, isbn:0-471-49616-2

35. Rodrigues M, de la Riva J, Fotheringham S: Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. Applied Geography 2014; 48:52-63. doi: 10.1016/j.apgeog.2014.01.011

36. Brunsdon C, Fotheringham A, Charlton M: Geographically weighted summary statistics—a framework for localised exploratory data analysis. Comput Environ Urban Syst 2002; 26:501-524. doi: 10.1016/S0198-9715(01)00009-6

37. de Castro MC, Monte-Mór RL, Sawyer DO, et al: Malaria risk on the Amazon frontier. Proc Natl Acad Sci U S A 2006; 103:2452-2457. doi: 10.1073/pnas.0510576103

38. Dysoley L, Kaneko A, Eto H, et al: Changing patterns of forest malaria among the mobile adult male population in Chumkiri District, Cambodia. Acta Trop 2008; 106:207-212. doi: 10.1016/j.actatropica.2007.01.007

39. Widayani P, Danoedoro P, Mardihusodo Sj: Pemodelan spasial kerentanan wilayah terhadap penyakit menular terkait lingkungan berbasis penginderaan jauh (kasus malaria, leptospirosis dan tuberkulosis di sebagian wilayah Provinsi Jawa Tengah dan DIY). Universitas Gadjah Mada, Fakultas Geografi; 2016. Retrieved from http://etd.repository.ugm.ac.id/.

40. Kleinschmidt I, Bagayoko M, Clarke G, et al: A spatial statistical approach to malaria mapping. Int J Epidemiol 2000; 29:355-361. doi:

41. Cohen JM, Ernst KC, Lindblade KA, et al: Local topographic wetness indices predict household malaria risk better than land-use and land-cover in the western Kenya highlands. Malar J 2010; 9:328. doi: 10.1186/1475-2875-9-328

42. Ovadje L, Nriagu J: Malaria as an Environmental Disease. In Encyclopedia of Environmental Health. Burlington: Elsevier; 2011: 558–567 Retrieved from

https://www.researchgate.net/publication/268515612_Ovadje_L_Nriagu_J_2011_Malaria_as_an_Environmental_Disease_In_Nriagu_JO_ed_Encyclopedia_of_Environmental_Health_volume_3_pp_558-567_Burlington_Elsevier.

43. Hanandita W, Tampubolon G: Geography and social distribution of malaria in Indonesian Papua: a cross-sectional study. Int J Health Geogr 2016; 15:13. doi: 10.1186/s12942-016-0043-y

44. Siraj AS, Santos-Vega M, Bouma MJ, et al: Altitudinal changes in malaria incidence in highlands of Ethiopia and Colombia. Science 2014; 343:1154-1158. doi: 10.1126/science.1244325

45. Pinault LL, Hunter FF: New highland distribution records of multiple Anopheles species in the Ecuadorian Andes. Malar J 2011; 10:236. doi: 10.1186/1475-2875-10-236

46. Alimi TO, Fuller DO, Qualls WA, et al: Predicting potential ranges of primary malaria vectors and malaria in northern South America based on projected changes in climate, land cover and human population. Parasit Vectors 2015; 8:431. doi: 10.1186/s13071-015-1033-9

47. Messina JP, Taylor SM, Meshnick SR, et al: Population, behavioural and environmental drivers of malaria prevalence in the Democratic Republic of Congo. Malar J 2011; 10:161. doi: 10.1186/1475-2875-10-161

48. Midekisa A, Beyene B, Mihretie A, et al: Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia. Parasit Vectors 2015; 8:339. doi: 10.1186/s13071-015-0954-7

49. Steffen W, Persson Å, Deutsch L, et al: The Anthropocene: From Global Change to Planetary Stewardship. Ambio 2011; 40:739-761. doi: 10.1007/s13280-011-0185-x

50. Malhi Y, Roberts JT, Betts RA, et al: Climate change, deforestation, and the fate of the Amazon. Science 2008; 319:169-172. doi: 10.1126/science.1146961

51. Chirebvu E, Chimbari MJ, Ngwenya BN, et al: Clinical Malaria Transmission Trends and Its Association with Climatic Variables in Tubu Village, Botswana: A Retrospective Analysis. PLoS One 2016; 11:e0139843. doi: 10.1371/journal.pone.0139843

52. Thomson MC, Ukawuba I, Hershey CL, et al: Using Rainfall and Temperature Data in the Evaluation of National Malaria Control Programs in Africa. Am J Trop Med Hyg 2017; 97:32-45. doi: 10.4269/ajtmh.16-0696

53. Day JF: Mosquito Oviposition Behavior and Vector Control. Insects 2016; 7:65. doi: 10.3390/insects7040065

54. Kleinschmidt I, Sharp BL, Clarke GPY, et al: Use of Generalized Linear Mixed Models in the Spatial Analysis of Small-Area Malaria Incidence Rates in KwaZulu Natal, South Africa. Am J Epidemiol 2001; 153:1213-1221. doi: 10.1093/aje/153.12.1213

55. Chihanga S, Haque U, Chanda E, et al: Malaria elimination in Botswana, 2012–2014: achievements and challenges. Parasit Vectors 2016; 9:99. doi: 10.1186/s13071-016-1382-z

56. Diggle PJ, Giorgi E: Model-Based Geostatistics for Prevalence Mapping in Low-Resource Settings. J Am Stat Assoc 2016; 111:1096-1120. doi: 10.1080/01621459.2015.1123158

57. Hay SI, Guerra CA, Gething PW, et al: A world malaria map: Plasmodium falciparum endemicity in 2007. PLoS Med 2009; 6:e1000048. doi:

58. Noor AM, Gething PW, Alegana VA, et al: The risks of malaria infection in Kenya in 2009. BMC Infect Dis 2009; 9:180. doi: 10.1186/1471-2334-9-180

59. Sakai M: Remembering origins ancestors and places in the Gumai society of South Sumatra. In The Poetic Power of Place comparative perspectives on Austronesian ideas of locality Australia: ANU Press; 2006: 42-55. Retrieved from http://www.jstor.org/stable/j.ctt2jbjrm.2.

60. Hoshen MB, Morse AP: A weather-driven model of malaria transmission. Malar J 2004; 3:32. doi: 10.1186/1475-2875-3-32

61. Barati M, Keshavarz-Valian H, Habibi-Nokhandan M, et al: Spatial outline of malaria transmission in Iran. Asian Pac J Trop Med 2012; 5:789-795. doi: 10.1016/s1995-7645(12)60145-x

Formatted: Indent: Left: -1.27 cm, Space After: 0 pt

Figure legends

Formatted: Indent: Left: -1.25 cm

Fig. 1: Map of the study area covering one city and seven districts of South Sumatra Province, Indonesia.

Fig. 2: Malaria cases and their geographical locations in the study area.

Fig. 3: Malaria incidence at village level.

Fig. 4: Each explanatory variable mapped in the study area.

Fig. 5: Flow chart of the research strategy.

Fig. 6: Predicted value from GWR for parameter estimates of explanatory variables of malaria incidence in the study area.

Fig. 7: Student's test significance (95% and 99% confidence interval) for each explanatory variable and village location.

Fig. 8: Goodness-of-fit of GWR model (local Coefficient of determination R^2) for malaria incidence associated with environmental factors in South Sumatra, Indonesia.

Table 1: GWR result based on Fixed Gaussian (distance) kernel function for geographical weighting.

Bandwidth and Geographic Ranges	Value
Bandwidth size:	9,184.47
Diagnostic information	
Residual sum of squares:	33,549.28
Classic AIC:	3,482.17
AICc:	3,721.35
BIC/MDL:	4,198.30
CV:	178.92
R square:	0.69
Adjusted R square:	0.41

Formatted: Space After: 0 pt

Table 2: Comparison between Global OLS and Local GWR models

Value	OLS	GWR
Residual sum of square	100,625.26	33,549.28
Classic AIC	3,625.82	3,482.17
AICc	3,626.15	3,721.35
R^2	0.06	0.69
Adjusted R^2	0.05	0.41

Table 3: The result of global regression model and geographical variability test of local coefficients for six environmental factors.

Variables	Global regression model output				Geographical variability test			
	Estimate	SE	T value	P value	F	DOF for F test		DIFF of Criterion
Intercept	7.98	4.63	1.72	0.04	33.20	10.48	261.38	-347.99
"Altitude (X1)"	-0.02	0.00	-4.03	0.00	0.24	12.02	261.38	19.19
"Aspect (X2)"	-0.01	0.01	-1.60	0.05	0.55	22.68	261.38	24.91
"Distance from the river (X3)"	0.00	0.00	-0.84	0.24	1.84	18.15	261.38	-16.03
"Distance from lakes and pond (X4)"	0.00	0.00	0.39	0.71	0.90	15.04	261.38	7.99
"Distance from forest (X5)"	0.00	0.00	-3.69	0.00	2.99	14.61	261.38	-38.12
"Rainfall (X6)"	0.00	0.00	2.38	0.02	13.07	10.17	261.38	-158.91

Table 4: ANOVA testing the null hypothesis that the GWR model represents no improvement over a global model.

Source	SS	DF	MS	F Count	F Table
Global Residuals	100,625.26	429.00			
GWR Improvement	67,075.98	197.74	339.22		
GWR Residuals	33,549.28	231.26	145.07	2.34	2.12

Formatted: Indent: Left: -1.25 cm

<u>Malaria Journal Review Comments</u>

Manuscript Title: Spatial Modelling of Malaria Incidence in South Sumatra, Indonesia

Authors: Hasyim et al.

Overall Comments:
This manuscript applies spatial analysis to malaria data in a low-endemic and heterogeneous area. By focusing the analysis on routinely reported data as well as using spatial covariates accessible within the country, this provides an approach that is accessible to malaria programs within the country. Overall, this manuscript is well written and provides useful information to help better understand malaria epidemiology in this area. However, before recommending for publication, I have several comments that should be considered.

Major:

- The study population needs to be better defined. The authors seem to use the terms surveillance and research population interchangeably, but the latter suggests that some of the data was collected outside of the routinely collected data. The details given (abstract and methods (LL 83-98)) on the difference between the total population and 'research participants' needs to be simplified. As the total population has the potential to go to the facility and be captured as part of routine surveillance it is not clear what the distinction is. It would improve clarity to use terms consistently throughout (e.g. X positive for malaria or XX suspected of having malaria)
- Mosquitoes and their ecosystems are significant spatial drivers for malaria transmission. You mention that 25 different species of *Anopheles* have been identified in the country (LL 29) but are all of these found in your study area? Furthermore, the results of the spatial heterogeneity in risk should be discussed in the context of the spatial heterogeneity of the vectors in the reason. This is an important confounding factor to address as different species may have different ecological niches and therefore different factors may be important in different places.
- Terminology of modeling methods: in the malaria spatial modeling field, the terms global and local typically refer to different scales of spatial autocorrelation both of which are present in malaria transmission (e.g. broader temperature bands vs. mosquito flight range). Your description of the models used is very clearly articulated and accessible to non-spatial/statistical people. However, re-framing this as a 'non-spatial' and 'spatial' regression instead of global/OLS and local/GWR would help clarify the important differences between the approaches being compared.
- Some important references on spatial modeling of malaria are missing. See works by A. Noor, P. Gething, I. Kleinschmidt, E. Giorgi for example.
- There is some repetition of concepts in the methods section that could be better organized or the difference between the multiple usages is not clear enough to appreciate the need for duplication. For example, the first two sections of preparation of spatial data both discuss data

interpolation. Similarly, the discussion of testing for multicollinearity is in pre-processing and processing sections. It might be helpful to distinguish when the data in question is spatial/a map and non-spatial. The addition of figure 5 clarifies the flow of information, but this lucidity should be reflected in the text.
- In the results section a lot of emphasis is placed on regression coefficients and less on the interpretation of these coefficients
- The approach used for comparing the two models are missing in the methods section.

Minor:
- Abstract: How many villages didn't report any malaria (436 of X villages)?
- LL 2-6 – a reference to figure 1 would be helpful here to give readers some spatial context of the places being mentioned.
- LL 21- "recent developments" needs to be elaborated on to ensure that those not familiar with the area can understand the context. Recent political? Economic? Social?
- LL 94 – what diagnostic test was used?
- LL 102 – more details on spatial input parameters are needed – what is the resolution of the different surfaces? Is it commercially available (e.g. landsat imagery) or did the government commission the images to be created? What year was it captured?
- The topographic wetness index (Cohen et al.) was shown to be a significant predictor of malaria and is a metric that can be derived from available data. The authors should consider adding to their analysis.
- It would be helpful to highlight in the methods (LL 102 – 107) that the malaria data inputted into the model is aggregated village level data with the village centroid (?) used as the spatial unit.
- LL 115 – "The rainfall map…..obtained from the scanned maps" – which maps?
- LL 118 – "GWR should have a normal distribution" – is this that variables used for GWR should be normally distributed? It would be helpful to have the untransformed distributions as a supplementary table to show the non-normality and the transformed version to support this.
- LL 120 – VIF should be defined at first instance
- LL 201 – How much is the results in Labat having the highest influence of environmental factors due to the higher case numbers and therefore more predictive power?
- LL 200-207 – the term regression coefficients are typically used to denote the covariates and their corresponding constant that represents the rate of the linear change in the association with the outcome variable. Whereas R^2 is a statistical measure of how close the data are to the fitted regression line and is interpreted as to how much of the variability is explained by the covariates. They are different measures with very different meanings and therefore different terminology should be used to denote the two.

Malaria Journal

Spatial modelling of malaria case associated with environmental factors in South Sumatra, Indonesia

--Manuscript Draft--

Manuscript Number:	MALJ-D-17-00578R3
Full Title:	Spatial modelling of malaria case associated with environmental factors in South Sumatra, Indonesia
Article Type:	Research
Funding Information:	
Abstract:	Background Malaria, a parasitic infection, is a life-threatening disease in South Sumatra Province, Indonesia. This study aimed to investigate the spatial association between malaria occurrence and environmental risk factors. Methods The number of confirmed malaria cases was analysed for the year 2013 from the routine reporting of the Provincial Health Office of South Sumatra. The cases were spread over 436 out of 1,613 villages. Six potential ecological predictors of malaria cases were analysed in the different regions using ordinary least square (OLS) and geographically weighted regression (GWR). The global pattern and spatial variability of associations between malaria cases and the selected potential ecological predictors was explored. Results The importance of different environmental and geographic parameters for malaria was shown at global and village-level in South Sumatra, Indonesia. The independent variables altitude, distance from forest, and rainfall in global OLS were significantly associated with malaria cases. However, as shown by GWR model and in line with recent reviews, the relationship between malaria and environmental factors in South Sumatra was found to strongly vary spatially in different regions. Conclusions A more in-depth understanding of local ecological factors influencing malaria disease as shown in present study may not only be useful for developing sustainable regional malaria control programmes, but can also benefit malaria elimination efforts at village level.
Corresponding Author:	Hamzah Hasyim, PhD candidate Goethe-Universitat Frankfurt am Main Fachbereich 16 Medizin Frankfurt am Main, Hessen GERMANY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Goethe-Universitat Frankfurt am Main Fachbereich 16 Medizin
Corresponding Author's Secondary Institution:	
First Author:	Hamzah Hasyim, PhD candidate
First Author Secondary Information:	
Order of Authors:	Hamzah Hasyim, PhD candidate
	Afi Nursafingi, M.Sc
	Ubydul Haque, PhD
	Doreen Montag, DPhil
	David Groneberg, Prof,Dr,PhD
	Meghnath Dhimal, PhD
	Ulrich Kuch, Dr

	Ruth Müller, Dr
Order of Authors Secondary Information:	
Response to Reviewers:	Response to the Reviewer #1: MALJ-D-17-00578 Comments#: MALJ-D-17-00578. Manuscript Title: Spatial Modelling of confirmed malaria cases in South Sumatra, Indonesia Reviewer reports: Reviewer #1: MALJ-D-17-00578 General comments Study description : In South Sumatra in Indonesia, annual malaria in 2013 and environment relationship have been tested with two methods, a global linear regression, Ordinary least square methods and a geographically weighted regression. This study is fascinating, and the results maps are well done. But some details are missing, and the interpretation could be improved. Response: I highly appreciate your positive advice to improve our manuscript MALJ-D-17-00578 with revised title “Spatial Modelling of confirmed malaria cases in South Sumatra, Indonesia”. Thank you so much for your constructive comments. Each comment has been carefully considered and responded point by point. Responses to the reviewer are made in italics. Here are some major indications followed by minor suggested corrections. Major comments: 1.The spatial analysis methods are not explained and are sometimes a little bit confused. Then a GWR can be run. GWR is a modified regression model and calculates a local specific variance for each coordinate point. This a local regression that has the advantage to highlight local relationships between the dependent variable and the explanatory variables by addition of weighted parameters. These weights are automatically determined for each location and can be mapped. The total regression for each location can also be mapped to identify the higher and lower regression coefficient. Then a map of residual should identify where another variable may be required. Response: Thank you for your feedback. We revised the description of spatial analysis GWR was used to model predictive confirmed malaria cases based on a specific geographic area (geographical coordinates) by obtaining different regression coefficients for each location in the study area [1]. 2.Here you chose a kernel fixed type with Cross-validation (CV) as a bandwidth method. The bandwidth controls the degree of smoothing in the model and identify an optimal fixed distance. Response: The optimum distance threshold (also known as the bandwidth) or the optimum number of neighbours can be determined in two ways: by minimising the square of the residuals by cross-validation (CV) or by minimising the Akaike Information Criterion (AIC) [2]. In our study, we select the type of weighing (kernel type) and optimum bandwidth selection method based on selection criteria. In our case, we use AIC. Classic AIC tends to choose smaller bandwidths by which geographically varying coefficients are likely to be undersmoothed. CV is applicable only to Gaussian models [3]. Comparison between the two methods can be made, (even if they are not the same analysis). If the better regression coefficient is better and the AIC is lower, it is concluded that this method is the best. Response: The best GWR model which used weighting function is 'Fixed' (Gaussian) fixed with the bandwidth selection method "Golden section search". Then we use AIC. It is a statistical measure, which quantifies the relative goodness-of-fit of various derived statistical models from a given sample dataset. The preferred model is that with the lowest AIC value. Geographically weighted regression (GWR) is the regression model that has been developed for data modelling with continuous response variable and considering the spatial or location aspect. The best bandwidth can be seen in the output table entitled bandwidth title and geographic ranges. We conclude the best bandwidth and criteria model goodness as stated in table below. Table 1. GWR result using ‘Fixed’ (Gaussian)

Bandwidth and geographic ranges
Bandwidth size: 9184.47
Diagnostic information
Residual sum of squares: 33549.28
Classic AIC: 3482.17
BIC/MDL: 4198.30
CV: 178.92
R square: 0.687
Adjusted R square: 0.409

The best bandwidth generated 9184 neighbours that have significant spatial relationships with a region. In addition, we demonstrated the best model selection by the value of the residual sum of square, classic AIC, and the R square, like in table 2. The smaller the AIC, the better the model performed. Further, the AIC considers the simplicity of the established model. In addition, the better the model is created if the value of R2 increases. In the table below, we can see if the GWR model is better than OLS model.

Table 2. Comparison of GWR and OLS models by value RSS, Classic AIC, and R2
ValueOLSGWR
Residual sum of square100,625.2633549.28
Classic AIC 3,625.823482.17
R20.0620.687

As given in the table, we demonstrate residual sum of square (RSS), and Classic AIC, of GWR which are smaller than the OLS, whereas R2 of GWR is greater than OLS. These parameters or indicators prove that the GWR model is better fitting than OLS to investigate whether independent variables significantly vary spatially. The global OLS model explained 6.2% variation in confirmed malaria cases.by environmental factors ($R^2 = 0.062$). It implies that 93.8% of the confirmed malaria cases is caused by unknown factors not investigated in this study and may be related to local variation which is not taken into account in the OLS model [1]. The local GWR explained 68.7% variation in confirmed malaria cases (Y) by environmental factors ($R^2 = 0.687$).

Table 3: The result of global regression model and geographical variability test of local coefficients for six environmental factors.
VariablesGlobal regression model outputGeographical variability test
EstimateSET valueP valueFDOF for F testDIFF of Criterion
Intercept7.984.631.720.0433.2010.48261.38-347.99
"Altitude (X1)"-0.020.00-4.030.000.2412.02261.3819.19
"Aspect (X2)"-0.010.01-1.600.050.5522.68261.3824.91
"Distance from the river (X3)"0.000.00-0.840.241.8418.15261.38-16.03
"Distance from lakes and pond (X4)"0.000.000.390.710.9015.04261.387.99
"Distance from forest (X5)"0.000.00-3.690.002.9914.61261.38-38.12
"Rainfall (X6)"0.000.002.380.0213.0710.17261.38-158.91

We use GWR4 software to compare performance between global OLS and local GWR. Moran's I test is not available for the analysis. The value of DIFF criterion indicates that the independent variables have spatial variability or local spatial heterogeneity that are altitude, distance from lakes and pond, distance from forest, and rainfall have spatial variability.
ANOVA in which the global model is compared with the GWR model. The ANOVA tests the null hypothesis that the GWR model represents no improvement over a global model. The results are shown below (Table 4).
Table 4: ANOVA testing the null hypothesis that the GWR model represents no improvement over a global model.

Source SS DF MS F

Global Residuals 100625.2620429.0000
GWR Improvement 67075.981197.736339.220
GWR Residuals 33549.281231.264145.0692.338336

The ANOVA test gives a brief guide to the improvement in model fit when we compare the local and global models. The GWR model could explain the relationship between the response variable „ confirmed malaria cases." and six explanatory variables significantly better than the global regression model OLS with F count (2.34) > F table (2.12), The locally weighed R2 between the observed and fitted values is a measure of how well the model replicates the local malaria incident values around each observation. GWR ANOVA Table is an integral part of result "Semiparametric Geographically Weighted Regression analysis", Release 1.0.90 (GWR 4.0.90)

3.The objective could be to predict, but here it seems that it is very interesting to identify malaria-environment relation and according to the location.

Response: We use ecology design study with the village that contains information of both attributes and location as unit analysis to predict confirmed malaria cases with potential environmental and geographic predictors of malaria.

4.A validated OLS can lead to a global policy and a validated relationship with GWR is more appropriate to drive to the local system.

Response: We completely agree and included this statement in our discussion chapter. Geographically Weighted Regression explores spatial varying impacts of these factors across the study area focusing attention on local variations in ecological associations. The set of selected environmental risk variables under consideration revealed significant associations with local confirmed malaria cases and these associations varied geographically across the study area. We observe and quantify different local factors driving confirmed malaria cases in different parts of the villages. A more in-depth understanding of local ecological factors influencing confirmed malaria cases may not only be used for developing sustainable regional malaria control programs but can also benefit malaria elimination efforts.

5.In the case of this study, it an excellent spatial analysis to identify which parameter to look closer and where and how much it varies and where it would be more appropriate to do so and for example do.

Response: Thank you for this positive comment on our study.

2.The validation is not clear.

Response: The model validation procedure conducted following steps: Step1: Preparation dataset. Step 2: Specify one regression type and the variable settings needed to determine the GWR model. We choose Geographical variability test, for model coefficient test obtained. Step 3: Currently, we use a geographic kernel type and its optimum bandwidth based on Selection Criteria. We demonstrated an "'Fixed' (Gaussian)" and selection bandwidth use "Golden section search" then use AIC criteria. It is a statistical model fit measure. It quantifies the relative goodness-of-fit of various derived statistical models, giving a sample dataset. The preferred model is that with the lowest AIC value. Step 4: Specify filenames for the files storing the modelling results, and Step 5: Execute the session to compare necessary calculations and read results. Through the geographical variability tests, the AIC and 'Fixed' (Gaussian) kernel are enabled to find the size and select the optimal bandwidth if the model is fit. We demonstrated OLS assumptions for classical diagnostic regression as multicollinearity test has done before the modelling. The regression was computed with many variables, which potentially gave rise to multicollinearity. We used an index based on predictive modelling variance that is Variance inflation factor (VIF) [4] Multicollinearity could occur when one independent variable was a linear function of another independent variable and previously observed in GWR modelling [5].

The following 'rules-of-thumb' for evaluating these factors: VIF > 10 give evidence of multicollinearity. with VIF > 100 there is certainly multicollinearity among the variables. [6, 7]. We show in multicollinearity does not occur, because the VIF value is less than 10 and the tolerance value is higher than 0.1. So, in the OLS method obtained a regression equation to estimate the actual regression model.

Collinearity Diagnostics

Variable	VIF	SQRT VIF	Tolerance	R-Squared
altitude	1.42	1.19	0.7041	0.2959
aspect	1.00	1.00	0.9965	0.0035
distfriv	1.05	1.03	0.9497	0.0503
distflak	1.07	1.04	0.9335	0.0665
distffor	1.18	1.08	0.8502	0.1498
rainfall	1.17	1.08	0.8532	0.1468

Mean VIF 1.15

The basic idea of GWR is that the parameters can be calculated in the study area with the dependent variable and one or more independent variables that it has been measured in places where the location is known. [8] In GWR, the sufficient number of degrees of freedom is a function of the bandwidth so the adjustment may be quite marked in comparison to a global model like OLS. For this reason, the AIC and R2 are preferred as a means of comparing models. So, we conclude, a valid GWR modelling is more appropriate to lead to local policy. In addition, the F test suggests that the GWR model is a significant improvement on the global model for confirmed malaria cases. In our case, these parameters prove that the GWR model is better than OLS that is a powerful tool for exploring spatial heterogeneity

6.Interpretation of the results could and be improved and better put in context. Context with links with transmission and specific known ecological preferences of some Anopheles species (Ex: Anopheles found in the forest for villages where the distance to the forest is a factor).
Response: We revised the discussion chapter accordingly. See also line 290-295: An. (Cellia) leucosphyrus Dönitz is considered to be of epidemiological importance for malaria transmission in forested areas of Sumatra (McArthur, 1951). The Anopheles was reported from eight sources at 47 independent sites. The first record of Anopheles parangensis from Sumatra was reported by O'Connor and Sopa (1981), but with no details on location [9]. However, in current research, we did not investigate the main Anopheles vector diversity in each study area.
7.Context with other studies, it is better to refer to study with similar environment, latitude, health system.
Response: Determination of regional vulnerability using GWR in Purworejo Regency of Indonesia concludes that each region is considered to have a distinctive characteristic that is different from other regions. So, it is necessary to give individual calculation to get weight on each parameter determining the vulnerability of Malaria. We discussed this study and other studies related to the outcomes of our modelling (please see discussion).
8.Discuss more the difference between districts, especially those with very high or very low local R2
Response: See lines 207-213
The GWR model provides evidence for a locally different influence of environmental factors on confirmed malaria cases.as shown by varying R^2 (Fig. 6). “Altitude” and “distance from lake and pond” show a positive association and “aspect” a negative association with malaria case in the Northern study area (Musi Banyuasin). “Rainfall” and “distance from river” show a positive association with confirmed malaria cases.in the Eastern part of Musi Rawas and Lahat. The variables “aspect”, distance from lake and pond” and “distance from forest” are positively associated with confirmed malaria cases in large parts of the study area and discussion chapter for discussing environmental parameters.
9.scale the relationship with parameters that vary intra-annually vary.
Response: Currently, we use secondary data 2013, due to current data limitations. Annual rain data is only available from some weather stations in South Sumatra and thus the interpolation of the 2013 rainfall data would result in bias. So, we use the five-year average data to spatially interpolate rain data throughout Sumatra.Minor comments:
1.Title: add environment to the title. It could be "Spatial modelling of malaria incidence relationship with environment factors" or something like that.
Response: Changed as suggested.
2.Keywords: malaria, geographically weighted regression, GWR, Ordinary least squares regression, OLS, Sumatra, rainfall, elevation, distance to water.
Response: Changed as suggested. Geographically weighted regression (GWR), Ordinary least squares (OLS), Physical environment, Local climate, Sumatra, rainfall, elevation, and distance to water.
3.It should be appreciated to name the primary Anopheles vector species for each type of environment or district
Response: An. nigerrimus is a confirmed malaria vector in Indonesia with the first evidence of Plasmodium infection reported by Overbeek from Palembang, South Sumatra in 1940 [9]. The distribution of malaria vectors amongst the main islands is

also not uniform Sumatra Island has six species, Papua (at least five species) and the Lesser Sundas archipelago (five species).
Figure: A map of the distribution of primary Anopheles malaria vectors in Indonesia

Currently, the primary vector of malaria which confirmed the main vector of malaria (found sporozoite) from the salivary glands as follows: An. letifer, An. nigerrimus, An. maculatus, An. sinensis, An. barbirostris, An. vagus, and An. sundaicus in South Sumatra Provinces region. The primary anopheles vector data are obtained from several studies, and particularly data from Vector and animal-borne disease control unit of Research and Development, National Institute of Health Research and Development (NIHRD), Ministry of Health (Indonesia) at Baruraja. Also, data are based on the regular reporting of malaria from South Sumatra Provincial Health Office, the kind of plasmodium was Plasmodium falciparum and Plasmodium vivax in this studies area. However, in current research, we did not investigate the primary vector Anopheles diversity in each study area.

10.Cases number or incidence?
Response:
The dependent variable is „confirmed malaria cases (Y).
Case, confirmed : Malaria case (or infection) in which the parasite has been detected in a diagnostic test, i.e. microscopy, a rapid diagnostic test or a molecular diagnostic test

Case, malaria : Occurrence of malaria infection in a person in whom the presence of malaria parasites in the blood has been confirmed by a diagnostic test
Note: A suspected malaria case cannot be considered a malaria case until parasitological confirmation. A malaria case can be classified as imported, indigenous, induced, introduced, relapsing or recrudescent (depending on the origin of infection); and as symptomatic or asymptomatic. In malaria control settings, a "case" is the occurrence of confirmed malaria infection with illness or disease. In settings where malaria is actively being eliminated or has been eliminated, a "case" is the occurrence of any confirmed malaria infection with or without symptoms
Incidence, malaria : Number of newly diagnosed malaria cases during a defined period in a specified population.
Ref :
Global Malaria Programme, WHO malaria terminology. World Health Organization 2016. Updated in August 2017. Retrieved from apps.who.int/iris/bitstream/10665/208815/1/WHO_HTM_GMP_2016.6_eng.pdf
Currently, from reporting of the new case of malaria, and these are confirmed malaria cases. Data for patients were positive for malaria parasites will entry in individual including (name, address, type of parasite, the treatment used). Monthly reporting is done in the first stages from puskesmas: the primary health care system in Indonesia at the village level continue to districts in the 2nd stage and then to provinces in the 3rd degree.
11.Which georeferenced system is used in which units (meters or degrees)
Response: The study area map (Figure 1) uses the World Geodetic System (WGS84) as its reference coordinate system (line 87-89).
12.Maps 6 and 7: Add units, please.
Response: The figures follow 3, 4, 6, 7 and eight deliberately do not display coordinate system due all these maps are meant to accentuate thematic information. The coordinate system can be seen in Figure 1.
13.Which is the scale or resolution in time and space for each parameter?
Response: Parameter distance from the river, distance from lake and pond, and distance from the forest are processed from River, Lakes, Ponds maps which derive from the topographic map which have 1: 50,000 scale. Forest cover maps obtained from Forest cover maps of South Sumatera 2013 on the scale of 1: 250.000. Rainfall parameter was calculated based on annual average rainfall over five years, and it was interpolated from several weather observations stations in studies area.
14.Forest: How old is the forest layer? Which year?
Response: The forest cover maps were extracted from the land cover map which made in 2013. This map is sourced from Ministry of Environment and Forestry, Indonesia.
15.Can we guess that some parts have been deforested since the forest cover has been recorded? Do we have information on the percentage of deforestation between this year and 2013?
Response: Indonesia contributes significantly to deforestation in Southeast Asia.

However, much uncertainty remains over the relative contributions of various forest-exploiting sectors to forest losses in the country [14]. Forest is discussed first because one of variable research is the distance to the forest. Regarding of is studies area deforestation, we do not have information on the percentage of deforestation due to current data limitations.

16.In the discussion links between your result and what you say about deforestation.

Response:

Next to climatic and environmental factors, distance of houses to a forest are interrelated through anthropogenic activities influencing the local and regional climate [10, 11]. A cross-sectional view in Brazil revealed for example that malaria case across health districts is positively correlated with the percentage of aggregated deforestation [12]. These observations can be confirmed for the relationship of malaria case with distance to lake, pond and forest for South Sumatra. Anopheles (Cellia) leucosphyrus is considered to be of epidemiological importance for malaria transmission in forested areas of Sumatra [9]. Anopheles was reported from eight sources at 47 independent sites. The first record of Anopheles parangensis from Sumatra was reported by O'Connor and Sopa (1981), but with no details on location [9]. In current research, the main Anopheles vector diversity in each study area was not investigated.

17.Rainfall: the rainfall-malaria relationship is probably a nonlinear relationship as it is written in the discussion. In this annual study, rainfall is used. Is it average rainfall or total amount?

Response: Average annual rainfall period 2007-2013 in South Sumatra has been used for analysis.

18.Temperature

Response: See limitations of the study: Due to limited data, some explanatory variable were not investigated like temperature. However, the temperature is connected with altitude and aspect or direction of the slope. In the same way, land use may be associated with distance from the river and distance from lakes and pond. Thus, although these parameters (temperature, humidity, land use) were eliminated from analysis, these environmental factors were indirectly represented by our chosen set of variables.

19.Elevation: often described as an indirect factor: less humidity, lower temperature or suitable for different Anopheles species.

Response: Thank you for your advice. The global OLS model revealed that altitude, distance from lakes and pond, and distance to the forest have a negative coefficient and rainfall has a positive coefficient, significantly influence malaria case. It meant confirmed malaria cases is more common in regions with high rainfall, lowland and areas adjacent to forest areas. Elevation often described as an indirect factor: less humidity, lower temperature or suitable for different Anopheles species. See also discussion chapter.

20.Results. Present only the result without assuming cause between the variables.

Changed as suggested.

21.Interpretation: Explain links with field data and known information.

The highest malaria case with 1,449 cases spread over 124 villages was found in Lahat District. Our analyses have identified Lahat as the South Sumatran district in which environmental factors were of greatest relevance for confirmed malaria cases. Lahat District has both lowland and mountain regions and is home to diverse ethnic groups such as the Gumai who live along the rivers of the highland areas. Lahat district is located between 3.25 to 4.15 degrees south latitude, 102.37 to 103.45 degrees east longitude. Lahat Regency has a climate tropical and wets with rainfall variation between 267,375 to 222,175 millimetres per month, with as many rainy days 145.25 days or an average of 12.10 days per month. Air temperature varies between 22,16 to 30.47 Celsius. The average air humidity is 78.50 with an average wind speed of 4.66 Km per hour.

22.References: Rainfall and malaria: you could add Botswana and Ethiopia works

Response: added in the discussion as suggested.

Comments by Line

23.Background: In the background, when a reference is cited to state a link between malaria transmission and an environment factor, it should be better to mention in which country or environment type. Example: line 39: "lowland location."

Response: added in the discussion as suggested.

24.Line 48 "it proliferates faster under higher temperatures", it depends where.
Response: The sentence has been rephrased: Vectors and parasites are both highly sensitive to any temperature changes, for example, the parasite proliferation depends on temperature.
Please read also: … "Higher temperatures also quicken the digestion of the blood meal and maturation of its developing eggs, thus increasing vectorial biting frequency. Given these well-established, mainly laboratory determined climate sensitivities, malaria has long been identified as the infectious disease most vulnerable to climate change (WHO, 1990)"…. And "…Higher temperatures may prolong the malaria transmission window and reduce the incubation period required for replication of the parasite in an infected mosquito. As regional temperatures change across India, the transmission window for malaria is likely to increase by 2–3 months in the northern states of Punjab, Haryana and Jammu and Kashmir, but to decrease in more southerly Odisha, Andhra Pradesh and Tamil Nadu as temperatures exceed 40°C (Dhiman et al., 2010)…."
And in another reference :
Plasmodium falciparum has Threshold (0C), Minimum for transmission was 16–19 and Maximum for survival was 33–39. P. vivax has Threshold (0C), Minimum for transmission 14.5–15 and Maximum for survival 33–39. Then Lower threshold (0C) both of the Plasmodium was 8–10 for biological activity. [13]
25.Line 63: very important to know which variables you have studied, please list them here. "performance of the OLS and GWR models in predicting.."
Response: We compare global OLS and local GWR modelling to analyse the environmental risk factors for malaria in South Sumatra that vary geographically at the regional level. The locally different ecological factors studied to potentially predict the response variable „ confirmed malaria cases." (Y) are altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4), distance from the forest (X5), and rainfall (X6).
26.Methods: Study area: Line 77: a range of altitude would be appreciated, highest altitude for all the area or for each district
Response: The elevation in the study area varies between 0 to 3,159 metres above sea level. [14]
27.Line 78: is it monthly rainfall amount by station?
Response: This is the amount of average monthly rainfall taken from the weather stations conducted by Indonesian Agency for Meteorological, Climatological and Geophysics (BMKG) Palembang.
28.Study population and data collection: Lines 85-86: How many PHC? Just to have an idea of the density by district (or by population or by area)
Response: Based on the dataset, totally of The Primary Health Centre (PHC) was 140. Some PHC in each district is varying. Lines 68-69
29.Line 92: 36 372 patients or presumed positive malaria cases? Some patients may come several times a year.
Response: There is a special form of malaria case to reporting whether the patient is a new or a relapse.
30. Line 94 % (3578/36372 is around 10%
Response: Almost 10% of those participants were tested positive for malaria.
31. Line 97: precise which sort of villages or number (436 villages)
Response: The number of person who does have a positive malaria test are spread across 436 villages.
32.Preparation of spatial data: Data acquisition and selection, Line 106:
How many stations? How many km are they close to each other? To have an idea of their density
Response: The distance between weather observation stations were 50-100 km in flat topography and 10 km in hilly terrain.
33.Data pre-processing, Line 112 DEM which spatial resolution? Issued from which satellite data type?
Response: DEM data is processed from a contour map or a topographic map scale 1:50.000 with high relief; the contour interval is 25 m.
34. Line 115 Which spatial interpolation method did you used for rainfall? Which classification and from which criteria did you use it?
Response: Rainfall data has been interpolated by Indonesian Agency for Meteorological, Climatological and Geophysics (BMKG).
35. Line 120: VIF? It should be useful here to describe the variance inflation factor, what is it and how it works.
Response: Please check in line 131

36.Data processing; See major comment above
Response: see revised manuscript, lines 126 to 185
37.Miss comparison between the two methods OLS and GWR
Response: see revised manuscript, lines 218-222, and table 2 : and table below:[1]

Ref: Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships/cA. Stewart Fotherington, Chris Brunsdon, and Martin Charlton. Wiley, 2002.
38. Miss validation
Response: see our response to comment 9.
39.Results; Environmental factors influencing malaria incidence at village level: local GWR model; Line 194: related to size of weights
Response: The regression coefficients were not directly related to the size of weights in the GWR model, but rather to the estimates of the values of all explanatory variables.
40.Line 206: "..show that the environmental factors prevailing In these regions are less suitable for explaining the variance of malaria incidence in this area" need to explain why, please.
Response: Based on the value of regression, the confirmed malaria cases caused by environmental factors is most dominant than others. So, that environmental factors are more appropriate to explain its contribution to the variation of confirmed malaria cases.in the location.
41.Comparison between OLS and GWR: cf major
Response: see revised manuscript, lines 218-222
42.Discussion; Cf major., Line 289: Avoid "spatial epidemiology microscope."
Response: Based on our findings, GWR is a diagnostic model discovering spatially varying relationships. and local GWR analysis can, therefore, serve as a 'spatial epidemiology microscope.'
43. Line 298: "The approach arbitrarily plots all of the cases in the settlements" I don't understand what you mean.
Response: The availability of malaria case data is the number of positive malaria per village, and it is not the coordinates of each malaria positive so, the case is placed in the centre of the settlement.
44. Line 305 - 311: Add seasonality studies, non-linear relationship, time downscaling (to monthly rather than annual cases), etc.
Response: Thank you for your advice. See below and discussion chapter:
Climate data are frequently used to account for the spatial, seasonal and interannual variation for Malaria transmission. Modelling numerical evaluations by time and space show connection with malaria prevalence.[15]
In the future, additional explanatory variables should be addressed to provide a comprehensive analysis of confirmed malaria cases. This should comprise, for example, the behaviour of mosquito vectors and that of community members, access to and delivery of health services, and other eco-bio-social factors that affect the confirmed malaria cases. Despite these limitations, our study sheds light on relevant local and regional realities regarding environmental variation and sociocultural practice which might interplay with vector-host relationships and provide a suitable environment for malaria mosquitoes.
45.Maps; Figure 1 and 3: you have to choose the same methods for all the maps to code the districts, numbers or abbreviations.
Response: Thank you for your advice. Revised.
46.Figure 3: scale and North are missing.
Response: Revised as suggested.
47.Figure 4: Legend (spatial representation map showing.. not needed) Each explanatory variable
Response: Figure description revised.
48.Figure 5: Is multicollinearity test also applied with the response variable?
Response: Multicollinearity test is applied both in explanatory and in response variable
49.Figure 6: reformulate the legend, please. It should be something like " predicted value from GWR".
Response: Revised as follows: Figure 6: Predicted value from GWR for parameter estimates of explanatory variables of confirmed malaria cases in the study area.
50.Figure 7: ….. Significance percentage value for each explanatory variable by village location
Response: Revised as follows: Figure 7. Student's test significance (95% and 99% confidence interval) for each explanatory variable and village location.

51.Figure 8: Local regression coefficient (R2) from GWR method by village location
Response: Revised as follows: Figure 8. Goodness-of-fit of GWR model (local R2) for confirmed malaria cases associated with environmental factors in South Sumatra, Indonesia.
Reference:
1.Fotheringham AS, Brunsdon C, Charlton M: Geographically weighted regression the analysis of spatially varying relationships. University of Newcastle, UK: John Wiley & Sons; 2002.
2.Rodrigues M, de la Riva J, Fotheringham S: Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. Applied Geography 2014; 48:52-63.
3.Nakaya T: GWR4.09 User Manual. GWR4 Windows Application for Geographically Weighted Regression Modelling. Kyoto: Ritsumeikan University, Department of Geography24 March 2016.
4.Halimi M, Farajzadeh M, Delavari M, Takhtardeshir A, Moradi A: Modelling spatial relationship between climatic conditions and annual parasite incidence of malaria in southern part of Sistan & Balouchistan Province of Iran using spatial statistic models. Asia Pac J Public Health 2014; 4:S167-S172.
5.Wheeler D, Tiefelsdorf M: Multicollinearity and correlation among local regression coefficients in geographically weighted regression. J Geogr Syst 2005; 7:161-187.
6.Rabe-Hesketh S, Everitt B: Handbook of statistical analyses using stata. CRC Press; 2003.
7.Kothari P: Data Analysis with Stata. Packt Publishing Ltd; 2015.
8.Brunsdon C, Fotheringham A, Charlton M: Geographically weighted summary statistics—a framework for localised exploratory data analysis. Comput Environ Urban Syst 2002; 26:501-524.
9.Elyazar IR, Sinka ME, Gething PW, Tarmidzi SN, Surya A, Kusriastuti R, Winarno, Baird JK, Hay SI, Bangs MJ: The distribution and bionomics of anopheles malaria vector mosquitoes in Indonesia. Adv Parasitol 2013; 83:173-266.
10.Widayani P, Danoedoro P, Mardihusodo Sj: Pemodelan Spasial Kerentanan Wilayah Terhadap Penyakit Menular Terkait Lingkungan Berbasis Penginderaan Jauh (Kasus Malaria, Leptospirosis Dan Tuberkulosis Di Sebagian Wilayah Provinsi Jawa Tengah Dan DIY). Universitas Gadjah Mada, 2016.
11.Kleinschmidt I, Bagayoko M, Clarke G, Craig M, Le Sueur D: A spatial statistical approach to malaria mapping. Int J Epidemiol 2000; 29:355-361.
12.Olson SH, Gangnon R, Silveira GA, Patz JA: Deforestation and malaria in Mancio Lima county, Brazil. Emerg Infect Dis 2010; 16:1108.
13.Chihanga S, Haque U, Chanda E, Mosweunyane T, Moakofhi K, Jibril HB, Motlaleng M, Zhang W, Glass GE: Malaria elimination in Botswana, 2012–2014: achievements and challenges. Parasit Vectors 2016; 9:99.
14.Midekisa A, Beyene B, Mihretie A, Bayabil E, Wimberly MC: Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia. Parasit Vectors 2015; 8:339.
15.Chirebvu E, Chimbari MJ, Ngwenya BN, Sartorius B: Clinical Malaria Transmission Trends and Its Association with Climatic Variables in Tubu Village, Botswana: A Retrospective Analysis. PLoS One 2016; 11:e0139843.
16.Thomson MC, Ukawuba I, Hershey CL, Bennett A, Ceccato P, Lyon B, Dinku T: Using rainfall and temperature data in the evaluation of national malaria control programs in Africa. Am J Trop Med Hyg 2017; 97:32-45.
17.Patz JA, Olson SH: Climate change and health: global to local influences on disease risk. Ann Trop Med Parasitol 2006; 100:535-549.
18.Hoshen MB, Morse AP: A weather-driven model of malaria transmission. Malar J 2004; 3:32.

Reviewer #2: MALJ-D-17-00578
Overall Comments:
This manuscript applies spatial analysis to malaria data in a low-endemic and different area. By focusing the analysis on routinely reported data as well as using spatial covariates accessible within the country, this provides an approach that is accessible to malaria programs within the country. Overall, this manuscript is well written and provides useful information to help better understand malaria epidemiology in this area. However, before recommending for publication, I have several comments that should be considered.

	Response: We appreciate very much for the reviewers' valuable comments, constructive criticisms, and insightful feedback. We carefully considered reviewer's suggestion and tried our best to improve the manuscript based on their explanation below. Each comment has been carefully considered point by point. Responses are made in italics. Major comments: 1.The study population needs to be better defined. The authors seem to use the terms surveillance and research population interchangeably, but the latter suggests that some of the data were collected outside of the routinely collected data. The details were given (abstract and methods (LL 83-98)) on the difference between the total population and 'research participants' needs to be simplified. As the total population has the potential to go to the facility and be captured as part of routine surveillance, it is not clear what the distinction is. It would improve clarity to use terms consistently throughout (e.g. X positive for malaria or XX suspected of having malaria) Please clarify and use only one terminus Response: Based on methodology, the population of our study was the total number of villages in 8 districts in South Sumatra, Indonesia and the sample was the village in the study area where positive malaria case is found. Currently, the village is a unit of analysis, and in our dataset, we call it "Toponymy". We noted 3,578 patient who has laboratory diagnosis of malaria. The cases were spread over 436 out of 1,613 villages. The village that contains information for both, attributes and location, is a unit analysis. We investigated potential ecological predictors of confirmed malaria cases in the different regions by performing global Ordinary Least Squares (OLS) and local Geographically Weighted Regression (GWR). Please see also lines 92-93 for Toponymy", and lines70-73 for information The cases spread over 436 out of 1,613 villages that were used for unit analysis 2.Mosquitoes and their ecosystems are significant spatial drivers for malaria transmission. You mention that 25 different species of Anopheles have been identified in the country (LL 29) but are all of these found in your study area? Furthermore, the results of the spatial heterogeneity in risk should be discussed in the context of the spatial heterogeneity of the vectors in the reason. This is an important confounding factor to address as different species may have different ecological niches and therefore different factors may be important in different places. Response: Actually, between 20-25 different species of Anopheles have been identified in Indonesia. There are 25 species Anopheles mosquitoes that have been confirmed to be malaria vectors in Indonesia, which are spread and divided into two zones of geographic dispersal of the Australian and Oriental zones.[16] "Approximately 230 million people live in Indonesia. The country is also home to over 20 anopheline vectors of malaria which transmit all four of the species of Plasmodium that routinely infect humans." [9, 17] Currently, in South Sumatra Provinces region, the main vector of malaria which confirmed were An. letifer, An. nigerrimus, An. maculatus, An. sinensis, An. barbirostris, An. vagus, and An. sundaicus. Pleasee also read line 360-365 in the manuscript: Anopheles (Cellia) leucosphyrus is considered to be of epidemiological importance for malaria transmission in forested areas of Sumatra [11]. Anopheles was reported from eight sources at 47 independent sites. The first record of Anopheles parangensis from Sumatra was reported by O'Connor and Sopa (1981), but with no details on location [9]. However, in current research, we did not investigate the main Anopheles vector diversity in each study area was not investigated. 3.The terminology of modelling methods: in the malaria spatial modelling field, the terms global and local typically refer to different scales of spatial autocorrelation both of which are present in malaria transmission (e.g. broader temperature bands vs mosquito flight range). Your description of the models used is very clearly articulated and accessible to non-spatial/statistical people. However, re-framing this as a 'non-spatial' and 'spatial' regression instead of global/OLS and local/GWR would help clarify the important differences between the approaches being compared. Response: In our study, we use term global for OLS and local for GWR. OLS and GWR are regression methods that both consider spatial factors. The difference of the regressions are : - OLS, the parameter estimate has the same value at all locations so that the relationship between the response variable and explanatory variable is considered

homogeneous (stationary).
- GWR, the parameter estimation value at each location varies so that the relationship between a response variable and explanatory variable is heterogeneous (non-stationary).
Thus, the term OLS cannot be substituted by 'non-spatial' and GWR cannot be substituted by 'spatial'. Additional explanations, it can be read below.

4.Some important references on spatial modelling of malaria are missing. See works by A. Noor, P. Getting, I. Kleinschmidt, E. Giorgi for example.

Response: Added as suggested.

A global database of malaria parasite prevalence using Malaria Atlas Project (MAP) to collect relational databases and related GIS. The documentation will help to improve the global spatial of malaria which demands investment in the collection of epidemiological intelligence. [18]

Analysis spatial, multiple regression analysis and spatially adjusted, are important implications for malaria control programs in a certain area with a method of adjusting the regression analysis was undertaken to identify factors that might explain very strong heterogeneity in the rates in South Africa. The results indicated strong spatial correlation in the rates by using generalized linear mixed models and variograms that malaria case was significantly positively connected with higher winter rainfall, a higher average maximum temperature and significantly negatively associated with increasing distance from water bodies. [19]

A simple two-stage procedure for producing maps of predicted malaria risk that is OLS analysis modelling on a larger scale to determine the relationship between Malaria prevalence in children under ten within the interval 0 to 1 and geo-statistical ('kriging') approaches used residual spatial dependence in the data to improve prediction at the local level. Some ecological potential predictors of malaria using climatic, population and topographic variables and investigated spatial pattern in the residuals of the model which is an important tool for malaria control in Mali. [20] A malaria risk map of the West African region uses on malariometric data survey to predict parasite prevalence for the whole of West Africa as a useful tool for health planners. It provides the opportunity for producing empirical models and maps of malaria distribution at a regional and eventually at a continental level. [21] A standard geostatistical model is important to prevalence mapping which relies on empirical prevalence data of this kind is a generalized linear mixed model with binomial error distribution, logistic link and a combination of explanatory variables and a Gaussian spatial stochastic process in the linear predictor.[22] Malaria endemicity within defined stable spatial limits of P. falciparum transmission has been investigated by a model-based geostatistical procedure was implemented within a Bayesian statistical framework.[23] Maps of transmission malaria and the impact of malaria on human populations not only contribute to a rational basis for control and elimination decisions but also are necessary to identify populations at different levels of risk and to evaluate options for disease control objectively.[24]
Advances geo-statistics are modelling, and malaria parasite prevalence data assemblies can be used to insert plasmodium falciparum risk distributions. A map of infection and disease risks is an appropriate strategy for the control of malaria requires Kenya. [25]

5.There is some repetition of concepts in the methods section that could be better organized or the difference between the multiple usages is not clear enough to appreciate the need for duplication. For example, the first two sections of preparation of spatial data both discuss data interpolation. Similarly, the discussion of testing for multicollinearity is in pre-processing and processing sections. It might be helpful to distinguish when the data in question is spatial/a map and non-spatial. The addition of figure 5 clarifies the flow of information, but this lucidity should be reflected in the text.
Response: We revised the method description as suggested. Please also see our response to your comment 3.
6.In the results section, a lot of emphases is placed on regression coefficients and less

on the interpretation of these coefficients.

Response: We added some interpretation in result section.

7.The approach used for comparing the two models are missing in the methods section. Please add the testing approach.

Response: Added as suggested. See line 218-222

8.Minor comments; Abstract: How many villages didn't report any malaria (436 of X villages)?
Response: The cases were spread over 436 of 1,613 villages. It mean's villages without malaria cases were 1,177 villages in the study area.
9.LL 2-6 – a reference to figure 1 would be helpful here to give readers some spatial context of the places being mentioned.
Response: See lines 87-88 The study area is located between 1°46' and 4°55' of southern latitude and between 102°4' and 104°41' of eastern longitude and has a total surface area of 46,377.40 km2 (Fig. 1). It covers eight endemic malaria districts of South Sumatra, Indonesia, namely Lahat, Muara Enim, Musi Banyuasin, Musi Rawas, North Musi Rawas, Ogan Komering Ulu, South Ogan Komering Ulu, and Lubuk Linggau. The topography of the area varies from lowland to mountainous landscapes. The elevation in the study area varies between -6 to 3.150 metres above sea level (Fig. 4). The climate is tropical and wet. In 2013 in South Sumatra, the lowest rainfall was 31 mm (August) in Lahat district, and the highest rainfall was 613 mm (March) in Palembang City. Monthly average temperatures ranged from 26.6 to 28.3°C and relative humidity from 81% to 88% in 2013 [26].
10.LL 21- "recent developments" needs to be elaborated on to ensure that those not familiar with the area can understand the context. Recent political? Economic? Social?
Response:
Please read 59-65, and 285-287

Indonesia contributes significantly to deforestation in Southeast Asia. Recent developments of deforestation have led to unsustainable practices which have resulted in a high frequency of deforestation in some regions and are an important factor influencing malaria incidence [27]. Deforestation is connected with malaria incidence in the county (município) of Mâncio Lima, Acre State, Brazil. The cross-sectional study shows 48% increase in malaria incidence associated with cumulative deforestation within respective health districts in 2006. [12]
11.LL 94 – what diagnostic test was used?
Response:
Either Rapid Diagnostic Tests (RDT) or microscopic assessment or both were used to confirm a malaria case. Please read 76-78

12.LL 102 – more details on spatial input parameters are needed – what is the resolution of the different surfaces? Is it commercially available (e.g. landsat imagery) or did the government commission the images to be created? What year was it captured?

Response: see lines 100-105; The topographic map consists of a collection of geographic data presented as thematic layers on a sketch done by The Indonesian Geospatial Information Agency (BIG). Researchers are not involved in this process. Topography data source: RBI (Rupa Bumi Indonesia) Bakosurtanal which is updated in 2014 in the location of study area.

13.The topographic wetness index (Cohen et al.) was shown to be a significant predictor of malaria and is a metric that can be derived from available data. The authors should consider adding to their analysis.
Response: The topographic wetness index (Cohen et al.) will be considered for future research, though its use needs to be further discussed in our team.
https://en.wikipedia.org/wiki/Topographic_wetness_index
"The index was designed for hillslope catenas. Accumulation numbers in flat areas will be very large, so TWI will not be a relevant variable." This may be a disadvantage in our study.
"The TWI has been developed to study spatial scale effects on hydrological processes

	and characterize biological processes such as annual net primary production, vegetation patterns, and forest site quality." One may assume that the analysis of topographic wetness index (Cohen et al.) may not reveal new patterns in our study given that it integrates all six variables that we studied, 14.It would be helpful to highlight in the methods (LL 102 – 107) that the malaria data inputted into the model is aggregated village level data with the village centroid (?) used as the spatial unit. Response: Added as suggested. The malaria input data is aggregated village level data with the village centroid used as the spatial unit. 15.LL 115 – "The rainfall map…..obtained from the scanned maps" – which maps? Response: See lines 93-110; A precipitation map (annual average) was obtained by interpolating the data of annual average rainfall from BMKG Climatological Station Class I in Palembang, South Sumatra, Indonesia 2007-2013 period. The interpolation process is done by BMKG then it classified into 7 rainfall classes. We obtained a map in JPG format. Further, the map rectified both in georeferencing and digitizing to create a map of precipitation vector format. Rectification is a process of transforming data from a single grid system using a geometric transformation. The result of digitization process can be seen in Figure 2 (rainfall variable). 16.LL 118 – "GWR should have a normal distribution" – is this that variables used for GWR should normally be distributed? It would be helpful to have the untransformed distributions as a supplementary table to show the non-normality and the transformed version to support this. Spatial data contains information with both attributes and location. The Geographically weighted regression (GWR) model, a local regression, was developed from an Ordinary Linear Regression (OLS) model based on nonparametric regression [28]. A non-parametric test does not assume anything about the underlying that the data comes from a normal distribution. GWR is a local regression that emphases 2nd order variation whereas OLS is a first order model. GWR is a varying-coefficient modelling technique. The general model in running both is to draw inference about first (global) and second (local) order process but, more directly GWR is specified to account for nonstationarity. GWR is a method for exploring spatial nonstationarity. This then produces a set of parameter estimates at each point in the defined geographical area. In this case, we run OLS using robust regression. Robust statistical tests operate well across a wide variety of distributions. The basic GWR method may be regarded as generalisations of the basic method where the core notion of a spatially non-stationary OLS regression model is enhanced [28]. 17.LL 120 – VIF should be defined at first instance Response: see lines 130-131, and 191-192 ;Done as suggested. The Variance Inflation Factor (VIF) and tolerance are both widely used measures of the degree of multicollinearity. [29] 18.LL 201 – How much is the results in Lahat having the highest influence of environmental factors due to the higher case numbers and therefore more predictive power? see lines 51 and 215-216 The regression coefficients for malaria incidence at the local level range between 0.18 - 1 (Fig. 8) The highest influence of environmental factors on malaria incidences was found in Lahat District. We will discuss after we re-calculated the models which suggested of Reviewer #1 19.LL 200-207 – the term regression coefficients are typically used to denote the covariates and their corresponding constant that represents the rate of the linear change in the association with the outcome variable. Whereas R2 is a statistical measure of how close the data are to the fitted regression line and is interpreted as to how much of the variability is explained by the covariates. They are different measures with very different meanings, and therefore different terminology should be used to denote the two. Response: Thank you for your feedback. see lines 170-173, 178-181, and 215-216 R2 is the coefficient of determination (R Squared): indicates the kindness of the model or the contribution of the independent variable to confirmed malaria cases . R-squared

is a statistical measure of how close the data are to the fitted regression line. It is also known as the coefficient of multiple determination for multiple regression. The regression coefficient is the constant (a) that represents the rate of change of one variable (y) as a function of changes in the other (x); it is the slope of the regression line. GWR4 provides almost same results for traditional GWR modelling. A few corrections have been made with regards to calculation methods for local diagnostic statistics, including local sigma and local R square. In output GWR4 Windows analysis, R square found in both in Global regression and GWR result. t represents the fraction of variability in response that can be explained by the variability in predictor variables. R2 is a statistical measure of how close the data are to the fitted regression line and is interpreted as to how much of the variability is explained by the covariates. In the simple linear regression case, R2 is simply the square of the correlation coefficient. The best model selection can be seen not only from the residual sum of square, and classic AIC but also in R square values. R2 in OLS was 68.7% and in GWR was 6.15% using 'Fixed' (Gaussian).

Reference:

1.Kementerian Kesehatan Republik Indonesia: Pedoman manajemen malaria. Jakarta: Direktur Jenderal Pengendalian Penyakit dan Penyehatan Lingkungan, Kementerian Kesehatan Republik Indonesia,; 2014.
2.Elyazar IR, Sinka ME, Gething PW, Tarmidzi SN, Surya A, Kusriastuti R, Winarno, Baird JK, Hay SI, Bangs MJ: The distribution and bionomics of anopheles malaria vector mosquitoes in Indonesia. Adv Parasitol 2013; 83:173-266.
3.Elyazar IR, Hay SI, Baird JK: Malaria distribution, prevalence, drug resistance and control in Indonesia. Adv Parasitol 2011; 74:41-175.
4.Guerra CA, Hay SI, Lucioparedes LS, Gikandi PW, Tatem AJ, Noor AM, Snow RW: Assembling a global database of malaria parasite prevalence for the Malaria Atlas Project. Malar J 2007; 6:17.
5.Kleinschmidt I, Sharp BL, Clarke GP, Curtis B, Fraser C: Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in Kwazulu Natal, South Africa. Am J Epidemiol 2001; 153:1213-1221.
6.Kleinschmidt I, Bagayoko M, Clarke G, Craig M, Le Sueur D: A spatial statistical approach to malaria mapping. Int J Epidemiol 2000; 29:355-361.
7.Kleinschmidt I, Omumbo J, Briet O, Van De Giesen N, Sogoba N, Mensah NK, Windmeijer P, Moussa M, Teuscher T: An empirical malaria distribution map for West Africa. Trop Med Int Health 2001; 6:779-786.
8.Diggle PJ, Giorgi E: Model-Based Geostatistics for Prevalence Mapping in Low-Resource Settings. Journal of the American Statistical Association 2016; 111:1096-1120.
9.Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, Kabaria CW, Manh BH, Elyazar IR, Brooker S: A world malaria map: Plasmodium falciparum endemicity in 2007. PLoS Med 2009; 6:e1000048.
10.Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IR, Johnston GL, Tatem AJ, Hay SI: A new world malaria map: Plasmodium falciparum endemicity in 2010. Malar J 2011; 10:378.
11.Noor AM, Gething PW, Alegana VA, Patil AP, Hay SI, Muchiri E, Juma E, Snow RW: The risks of malaria infection in Kenya in 2009. BMC Infect Dis 2009; 9:180.
12.Huisman O, De By R: Principles of geographic information systems. ITC Educational Textbook Series 2009; 1.
13.Brunsdon C, Fotheringham A, Charlton M: Geographically weighted summary statistics—a framework for localised exploratory data analysis. Comput Environ Urban Syst 2002; 26:501-524.
14.Abood SA, Lee JSH, Burivalova Z, Garcia-Ulloa J, Koh LP: Relative contributions of the logging, fiber, oil palm, and mining industries to forest loss in Indonesia. Conserv Lett 2015; 8:58-67.
15.Olson SH, Gangnon R, Silveira GA, Patz JA: Deforestation and malaria in Mancio Lima county, Brazil. Emerg Infect Dis 2010; 16:1108.
16.Fotheringham AS, Brunsdon C, Charlton M: Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons; 2003.
17.O’brien RM: A caution regarding rules of thumb for variance inflation factors. Quality & Quantity 2007; 41:673-690.

Reviewer #3: MALJ-D-17-00578

This manuscript presented an analysis of routine malaria surveillance data for 2013 to

examine the spatial patterns of malaria in South Sumatra, Indonesia. Ordinary least squares and geographically weighted regression analyses were used to examine the potential role of environmental risk factors on the spatial patterns of malaria incidence. Findings indicated that rainfall and distance from the forest played a role in explaining the malaria incidence. While the paper contains results that could be of interest, major revisions are necessary for the language. The paper was not focused and included too much extraneous information, yet did not include important information about the methods. There were also several concerns with the methods and interpretation of the findings.
Response: We thank anonymous reviewers for providing us very insightful and constructive comments. We have tried our best to carefully consider and respond to all the comments raised by the reviewer 3. We revised the manuscript substantially to improve the language and the presentation of our data as outlined below.
1.Abstract: From a statistical perspective, it is unclear how "having an R-squared value of 60%" indicates "that almost all independent variables were significant at certain locations at the village level."
Response:. We rephrased the result part of abstract as following: The importance of different environmental and geographic parameters for malaria was shown at global and village-level in South Sumatra, Indonesia. The independent variables altitude, distance from forest, and rainfall in global OLS were significantly associated with malaria cases. However, as shown by GWR model and in line with recent reviews, the relationship between malaria and environmental factors in South Sumatra was found to strongly vary spatially in different regions.
Abstract: The conclusions do not match the stated aim of the paper and instead highlight the merits of methodological approach instead of how the findings "help in the development of local policies for malaria elimination" in South Sumatra.
Response: We rephrased the conclusions as following: A more in-depth understanding of local ecological factors influencing malaria confirmed malaria case as shown in present study may not only be usedful for developing sustainable regional malaria control programmes, but can also benefit malaria elimination efforts at village level.

2.Background: This section needs to be more concise and relevant to the study conducted and aims addressed. For instance, the authors exhaustively discuss the role of several variables (migration, population density, temperature, etc.), none of which are considered in the present study. The authors need to focus on outlining the wider context, gaps in knowledge/evidence and then introduce the present research and how it addresses those gaps.
Response: Thank you for your advice. We revised this section as suggested. There is an overall very diverse malaria prevalence distribution with remote areas showing the highest prevalence [30]. Different factors affect malaria transmission within the province [16, 17, 31], and it is important to differentiate between factors that influence the vector, the parasite and the host-vector relationship since specific meteorological, environmental factors are at interplay [32]. Atieli et al. have demonstrated that topographic variables such as elevation, slope, and aspect are influencing the development of Anopheles mosquitoes [33] There is a significant association between local spatial variations like population density, lowland location in north-eastern Venezuela, and proximity to aquatic environments with malaria transmission [34]. In our study, the ecological potential to predict the response variable „malaria incidence" (Y) are altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4). Also, distance from the forest (X5) and rainfall (X6) that locally different as a variable of research.
3.Methods: How many primary health centres reported malaria case data? And what is the level of completeness of this data? How does the malaria case data from the primary health centres become village level data? Was the analysis at the village or health facility level?
Response: The malaria diagnostic data were obtained from the regular health information reporting system of the Provincial Health Office of South Sumatra. The data had been collected during 12 months (January to December 2013) at the village level from patients seeking treatment in PHC, locally called Pusat Kesehatan Masyarakat ("puskesmas"), and that were reported monthly to the Provincial Health Office via the malaria programs in the District Health Offices.
The analysis is based on village level.
We noted 140 primary health care reported malaria case data in the study area. The patients are categorised into “clinical diagnosis”, “suspected malaria” and “positive

	malaria". Categories "clinical diagnosis" or "suspected malaria" are based on the patient's symptoms and physical findings at examination. "Positive malaria" is a case of malaria diagnosed microscopically (examination of blood specimen/preparation) or rapid diagnosis test (RDT) with positive results for Plasmodium. Either RDT or microscopic assessment or both were used to confirm the diagnosis of malaria. Reporting of malaria incidence allowed the calculation of Annual Parasitic Incidence (API) that is the number of positive cases per 1,000 total population. 4.Methods: Authors state that "In the study region and period, 2,787,954 of the total population and 36,372 research participants visited hospitals or PHCs due to suspected malaria fever". Elsewhere, authors state "The study population was the number of participants who were suspected of having malaria while the sample was the number of participants with laboratory-confirmed malaria." It is unclear what the authors mean by study population, sample, research participants, and total population. Response: The population of our study was a total village in 8 provinces South Sumatra, Indonesia. The sample was a village with a malaria case and together with location, this village is our unit ("toponym"). In total, 3,578 patients were laboratory positive for malaria. The malaria cases were spread over 436 out of 1,613 villages in 8 endemic malaria districts of South Sumatra Province. 5.Methods: Were multiple episodes from the same individual included? Or was the analysis based on single malaria episodes? As there can be potential biases from relapses especially from P.vivax. Response: For each patient who visit a PHC, there is a unique patient data form which was filled out. So, there could be make a decision if cases were new or relapsed. Based on policy from the ministry of health, each patient who has diagnosed malaria positive should have had an epidemiology investigation. Case-finding activities were carried out passively (patients arrived at health-care facilities) and actively by mass blood survey and contact surveys r epidemiological investigations. 6.Line 96: The authors discuss locations of cases in each district. Is this the location of the primary health centre they sought care, or the location of their residence? Response: The malaria case data entered into the model has been aggregated to village level data with the village centroid used as spatial unit. 7.Methods: Authors included several distance variables - it is unclear whether these are distances from the village of residence to the attribute of interest (river, forest, etc.) or distances from the primary health centre. Response: The distance in this paper meant was the distance of case (village) to the variable. 8.Methods: Was any validation of the OLS or GWR models conducted. For example, cross-validation or bootstrapping? And what was the impact on the results? Response: The model validation procedure was conducted as following: Step1: Preparation of dataset. Step 2: Specify one regression type and the variable settings needed to determine the GWR model. We chose Geographical variability test, for model coefficient test obtained. Step 3: Choosing a geographic kernel type and its optimum bandwidth based on Selection Criteria. In this paper, we demonstrated an "Adaptive bi-square kernel" and selection bandwidth use "Golden section search" then use AIC criteria and residual sum of square. Step 4: Specify filenames for the files storing the modelling results, and Step 5: Execute the session to compare necessary calculations and read results. When the model is fit with the geographical variability test, the adaptive kernel function, the golden section search for finding the optimal bandwidth size, and AIC as the model indicator for selecting the optimal bandwidth. We demonstrated OLS assumptions with Durbin Watson coefficient, and we found value .092, hence the assumption of independence was fulfilled. Besides, diagnostic regression multicollinearity has been done before the modelling. We show that multicollinearity does not occur, because the VIF value is less than 10 and the tolerance value is greater than 0.1. 9.Methods: it is unclear how the outcome malaria incidence was defined as there was no mention of village size or population, and also unclear whether this was at the primary health centre level or the village level? Response: In this study, a village was the geographic unit.. 10.Methods: one requirement for an OLS is that only statistically significant explanatory

variables are included. However, it seems that the OLS model used by authors included several variables that were not significant.
Response: Although the variable studies that are Distance from the river (X3) and Distance from lakes and pond (X4) is not statistically significant, we choose to investigate a full OLS model for following reason: The independent variable has a relationship in substance with the dependent variable. The independent variables are significant at some specific places at the local level analysis. That means if the independent variable is not involved in GWR analysis, we will lose critical information. We show in OLS model that a decline of altitude, aspect and distance to forest and an incline of rainfall are risk factors for getting malariacase. In this model, we show Akaike's information criterion and Bayesian information criterion better than the full model.

```
. *Full model
. regress cases altitude aspect distfriv distflak distffor rainfall , vce(robust) level(95)
Linear regression                          Number of obs     =        436
                                           F(6, 429)         =       5.30
                                           Prob > F          =     0.0000
                                           R-squared         =     0.0615
                                           Root MSE          =     15.315

------------------------------------------------------------------------------
             |               Robust
       cases |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    altitude |  -.0154009   .0032999    -4.67   0.000    -.0218869   -.0089148
      aspect |  -.0137931   .0070601    -1.95   0.051    -.0276698    .0000835
    distfriv |  -.0011115   .0009416    -1.18   0.238    -.0029622    .0007392
    distflak |   .0000782   .0002092     0.37   0.709     -.000333    .0004893
    distffor |  -.0004079   .0001369    -2.98   0.003     -.000677   -.0001388
    rainfall |   .0038088   .0015844     2.40   0.017     .0006946     .006923
       _cons |   7.976743   3.896792     2.05   0.041     .3175618    15.63592
------------------------------------------------------------------------------
r; t=0.43 13:22:10
 . *partial model
. regress cases altitude distffor rainfall , vce(robust) level(95)

Linear regression                          Number of obs     =        436
                                           F(3, 432)         =       8.92
                                           Prob > F          =     0.0000
                                           R-squared         =     0.0540
                                           Root MSE          =     15.323

------------------------------------------------------------------------------
             |               Robust
       cases |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    altitude |  -.0151679   .0031585    -4.80   0.000    -.0213758   -.0089599
    distffor |  -.0004035   .0001374    -2.94   0.003    -.0006736   -.0001334
    rainfall |   .0040303   .0015094     2.67   0.008     .0010637    .0069969
       _cons |   4.680717   3.512194     1.33   0.183    -2.222397    11.58383
------------------------------------------------------------------------------
r; t=0.02 13:22:10

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |        Obs  ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
        full |        436 -1818.753  -1804.908       7    3623.816   3652.359
         sub |        436 -1818.753  -1806.657       4    3621.314   3637.625
-----------------------------------------------------------------------------
               Note: N=Obs used in calculating BIC; see [R] BIC note.
r; t=0.01 13:22:10.
```

end of do-file

11.Methods: lines 123 - 169 go into an exhaustive explanation of the GWR, and OLS approaches, while some detail is important, this much information seems to shift the focus of the paper to one on methodological approaches and distracts from the stated aim to "use global and local spatial modelling to analyses.

Response: Thank you for your advice. We comprehensively revised the method section.

12.The environmental risk factors for malaria in South Sumatra that vary geographically at the regional level."

Response: We deleted this sentence.

13.Methods: why were other variables such as village size/health facility catchment area size, household density, distance to health facility, coverage of malaria interventions included? Also was seasonality accounted for?
Response: We investigated physical environment variables as independent variables. Other non-physical environmental variables were not explored. We will consider other eco-bio-social variables in future studies.
14.Results: Lines 172 - 175, where is the incidence data presented? And it is still unclear what incidence refers to? Is the number of cases what is being referred to as incidence?

Response: Malaria case has been diagnosed microscopically (examination of blood specimen/preparation) or rapid diagnosis test (RDT) with positive results for Plasmodium. Either RDT or microscopic assessment or both were used to confirm the diagnosis of malaria. The malaria diagnostic data were obtained from the regular health information reporting system of the Provincial Health Office of South Sumatra. The data had been collected during 12 months (January to December 2013) at the village level from patients seeking treatment in PHC, locally called Pusat Kesehatan Masyarakat ("puskesmas"), and that were reported monthly to the Provincial Health Office via the malaria programs in the District Health Offices.

15.Table 1: From a statistical perspective, the OLS model should only include variables with significant coefficients, and that are in the expected direction.
Response: Thank you for your comment. The tables were changed accordingly.
16.Table 1: Please provide units and scale the variables appropriately, so the results are interpretable. For instance, distance from the forest has a coefficient of 0.00 which cannot be interpreted.

Response: Thank you for your comment. The tables were changed accordingly.

17.Results: Lines 186 - 189: authors conclude that malaria incidence is more common in regions with high rainfall and areas adjacent to forest areas. However looking at the coefficients presented in Table 1, distance from forest area has a positive coefficient, meaning that as the distance from forest area increases malaria incidence increases. Please clarify.
Response: The explanatory variables altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4) are locally different. Also, distance from the forest (X5) and rainfall (X6) have different strengths to predict the response variable „malaria case“ (Y). The global OLS model revealed that altitude, distance from lakes and pond, and distance to forest have a negative coefficient and rainfall has a positive coefficient, and significantly influence malaria case

Table 3: The result of global regression model and geographical variability test of local coefficients for six environmental factors.
VariablesGlobal regression model outputGeographical variability test
EstimateSET valueP valueFDOF for F testDIFF of Criterion
Intercept7.984.631.720.0433.2010.48261.38-347.99
"Altitude (X1)"-0.020.00-4.030.000.2412.02261.3819.19
"Aspect (X2)"-0.010.01-1.600.050.5522.68261.3824.91

"Distance from the river (X3)"0.000.00-0.840.241.8418.15261.38-16.03
"Distance from lakes and pond (X4)"0.000.000.390.710.9015.04261.387.99
"Distance from forest (X5)"0.000.00-3.690.002.9914.61261.38-38.12
"Rainfall (X6)"0.000.002.380.0213.0710.17261.38-158.91

18.Lines 200 - 201: Authors state "The regression coefficients for malaria incidence at the local level range from 0.03 to 0.99 (Fig. 8)." However, Fig 8 presents the R2 values which are different from the regression coefficients.
Response: We have corrected the local coefficient of determination (R squared) for malaria cases at the local level range between < 0.20 - 0.78.
19.Lines 201-202: Authors state "The highest influence of environmental factors on malaria incidences was found in Lahat District." It is not clear where this conclusion came from especially considering Figure 8.
Response: The statement is not related to Figure 8, however, to Figure 7.
20.Lines 202 - 207. Authors have erroneously interpreted R2 values as values of regression coefficients.

Response: Thank you for your feedback. In our understanding, R^2 is: The coefficient of determination (R Squared) that indicates the kindness of the model or the contribution of the independent variable to confirmed malaria cases. R^2 is a statistical measure of how close the data are to the fitted regression line. It is also known as the coefficient of multiple determination for multiple regression. The regression coefficient is the constant (a) that represents the rate of change of one variable (y) as a function of changes in the other (x); it is the slope of the regression line.

21.Discussion: Authors state that their "analyses have identified Lahat as the South Sumatran district in which environmental factors were of greatest relevance for malaria incidence." Caution is needed in making such conclusions especially given that the small village level sample sizes (Fig 3). Inability to detect significant relationships may, in fact, be related to the small sample sizes.

Response: Thank you for your advice. The highest confirmed malaria cases with 1,449 cases spread over 124 villages were found in Lahat District. Based on local geographical variability tests of coefficients, we demonstrated that the independent variables significantly revealed spatial variability or local spatial heterogeneity (altitude, distance from lakes and pond). The global OLS model revealed that altitude, distance from lakes and pond, and distance to forest and rainfall significantly influence confirmed malaria cases.

22.Discussion: much of the discussion is very anecdotal and not directly related to the findings presented. For instance, the authors discuss the relevance of deforestation and distance to coal mines, none of which was assessed in the present study.

Response: Thank you for your advice. We consider your comment no. 23 to 25 and revised text in discussion chapter accordingly. In accordance, we now focus on locally different altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4), distance from forest (X5) and rainfall (X6) that different strengths to predict the response variable „confirmed malaria cases (Y).

23.Discussion: authors should avoid introducing new data in the discussion. For instance, authors discuss distance between coal mines and local plantations and forests in Lahat District (lines 254-255). Elsewhere authors state "temperature was correlated with altitude and humidity…".
Response: Thank you for advice. We revised the text in discussion chapter and focus on our explanatory variables.
24.The topic distance between coal mines and local plantations and forests in Lahat District (lines 254-255).

Response: On average, we observed distances of 200-700 m between the coal mines and local plantations and forests in Lahat District (M. Alam, unpublished data). Elsewhere, the distance of households from a forest and the borders of swamps have often been associated with the risk of malaria infection [35].

References.

1.Fotheringham A, Brunsdon C, Charlton M. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. University of Newcastle, UK: John Wiley & Sons. Retrieved from https://www.wiley.com/, isbn:0-471-49616-2
2.Rodrigues M, de la Riva J, Fotheringham S: Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. Applied Geography 2014; 48:52-63. doi:
3.Nakaya T. (24 March 2016). GWR4 Windows Application for Geographically Weighted Regression Modelling. Kyoto: Ritsumeikan University, Department of Geography. Retrieved from https://raw.githubusercontent.com/gwrtools/gwr4/master/GWR4manual_409.pdf,
4.Halimi M, Farajzadeh M, Delavari M, Takhtardeshir A, Moradi A: Modelling spatial relationship between climatic conditions and annual parasite incidence of malaria in southern part of Sistan & Balouchistan Province of Iran using spatial statistic models. Asia Pac J Public Health 2014; 4:S167-S72. doi: 10.1016/S2222-1808(14)60434-5
5.Wheeler D, Tiefelsdorf M: Multicollinearity and correlation among local regression coefficients in geographically weighted regression. J Geogr Syst 2005; 7:161-87. doi: 10.1007/s10109-005-0155-6
6.Rabe-Hesketh S, Everitt B: Handbook of statistical analyses using stata. edn.: CRC Press; 2003, isbn:1439832358
7.Kothari P: Data Analysis with Stata. edn.: Packt Publishing Ltd; 2015, isbn:1782173188
8.Brunsdon C, Fotheringham A, Charlton M: Geographically weighted summary statistics—a framework for localised exploratory data analysis. Comput Environ Urban Syst 2002; 26:501-24. doi: 10.1016/S0198-9715(01)00009-6
9.Elyazar IR, Sinka ME, Gething PW, Tarmidzi SN, Surya A, Kusriastuti R, et al. (2013). The distribution and bionomics of anopheles malaria vector mosquitoes in Indonesia. (Vol. 83, pp. 173-266). Academic Press, Elsevier. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23876873, isbn:978-0-12-407705-8
10.Steffen W, Persson Å, Deutsch L, Zalasiewicz J, Williams M, Richardson K, et al: The anthropocene: from global change to planetary stewardship. Ambio 2011; 40:739-61. doi: 10.1007/s13280-011-0185-x
11.Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li W, Nobre CA: Climate change, deforestation, and the fate of the Amazon. Science 2008; 319:169-72. doi: 10.1126/science.1146961
12.Olson SH, Gangnon R, Silveira GA, Patz JA: Deforestation and malaria in Mancio Lima county, Brazil. Emerg Infect Dis 2010; 16:1108. doi:
13.Patz JA, Olson SH: Climate change and health: global to local influences on disease risk. Ann Trop Med Parasitol 2006; 100:535-49. doi:
14.Badan Pusat Statistik. (2014). Sumatera Selatan dalam angka (Sumatera Selatan in figures) 2013. Sumatera Selatan: BPS – Statistics Sumatera Selatan Retrieved from http://sumsel.bps.go.id.
15.Hoshen MB, Morse AP: A weather-driven model of malaria transmission. Malar J 2004; 3:32. doi: 10.1186/1475-2875-3-32
16.Ministry of Health Republic of Indonesia. (2014). Malaria management: guideline. Jakarta: Directorate of Vector Borne Disease and Zoonosis Control, Directorate General of Disease Prevention and Control, Ministry of Health (MoH) of Indonesia Retrieved from http://eliminasimalaria.blogspot.de/p/download.html.
17.Elyazar IR, Hay SI, Baird JK. (2011). Malaria distribution, prevalence, drug resistance and control in Indonesia. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21295677, isbn:2163-6079
18.Guerra CA, Hay SI, Lucioparedes LS, Gikandi PW, Tatem AJ, Noor AM, et al: Assembling a global database of malaria parasite prevalence for the Malaria Atlas Project. Malar J 2007; 6:17. doi: 10.1186/1475-2875-6-17
19.Kleinschmidt I, Sharp BL, Clarke GP, Curtis B, Fraser C: Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in Kwazulu Natal, South Africa. Am J Epidemiol 2001; 153:1213-21. doi: doi.org/10.1093/aje/153.12.1213
20.Kleinschmidt I, Bagayoko M, Clarke G, Craig M, Le Sueur D: A spatial statistical approach to malaria mapping. Int J Epidemiol 2000; 29:355-61. doi:
21.Kleinschmidt I, Omumbo J, Briet O, Van De Giesen N, Sogoba N, Mensah NK, et al: An empirical malaria distribution map for West Africa. Trop Med Int Health 2001; 6:779-86. doi:
22.Diggle PJ, Giorgi E: Model-Based Geostatistics for Prevalence Mapping in Low-Resource Settings. J Am Stat Assoc 2016; 111:1096-120. doi:

10.1080/01621459.2015.1123158
23.Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, et al: A world malaria map: Plasmodium falciparum endemicity in 2007. PLoS Med 2009; 6:e1000048. doi:
24.Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IR, Johnston GL, et al: A new world malaria map: Plasmodium falciparum endemicity in 2010. Malar J 2011; 10:378. doi: 10.1186/1475-2875-10-378
25.Noor AM, Gething PW, Alegana VA, Patil AP, Hay SI, Muchiri E, et al: The risks of malaria infection in Kenya in 2009. BMC Infect Dis 2009; 9:180. doi: 10.1186/1471-2334-9-180
26.Indonesian Agency for Meteorological Climatological and Geophysics: Informasi Hujan Bulanan dan Informasi Unsur Iklim Rata-Rata 2007 - 2013. BMKG, Climatology Station Class I Kenten Palembang2013.
27.Abood SA, Lee JSH, Burivalova Z, Garcia-Ulloa J, Koh LP: Relative contributions of the logging, fiber, oil palm, and mining industries to forest loss in Indonesia. Conserv Lett 2015; 8:58-67. doi: 10.1111/conl.12103
28.Fotheringham AS, Brunsdon C, Charlton M: Geographically weighted regression: the analysis of spatially varying relationships. edn.: John Wiley & Sons; 2003, isbn:0470855258
29.O'brien RM: A caution regarding rules of thumb for variance inflation factors. Quality & Quantity 2007; 41:673-90. doi:
30.World Health Organization. (2017). Regional action framework for malaria control and elimination in the Western Pacific: 2016-2020. Manila, Philippines WHO Regional Office for the Western Pacific. Retrieved from http://iris.wpro.who.int/handle/10665.1/13578, isbn:9789290618157
31.WHO Country Office for Indonesia and Ministry of Health Indonesia. (2013). National malaria control programme review: Republic of Indonesia. Jakarta: WHO Country Office for Indonesia and Ministry of Health, Indonesia. Retrieved from http://www.who.int/iris/handle/10665/253960, isbn:978 979 19477 4 9
32.Loha E, Lindtjorn B: Model variations in predicting incidence of Plasmodium falciparum malaria using 1998-2007 morbidity and meteorological data from south Ethiopia. Malar J 2010; 9:166. doi: 10.1186/1475-2875-9-166
33.Atieli HE, Zhou G, Lee M-C, Kweka EJ, Afrane Y, Mwanzo I, et al: Topography as a modifier of breeding habitats and concurrent vulnerability to malaria risk in the western Kenya highlands. Parasit Vectors 2011; 4:241. doi: 10.1186/1756-3305-4-241
34.Grillet ME, Barrera R, Martinez JE, Berti J, Fortin MJ: Disentangling the effect of local and global spatial variation on a mosquito-borne infection in a neotropical heterogeneous environment. Am J Trop Med Hyg 2010; 82:194-201. doi: 10.4269/ajtmh.2010.09-0040
35.Herdiana H, Cotter C, Coutrier FN, Zarlinda I, Zelman BW, Tirta YK, et al: Malaria risk factor assessment using active and passive surveillance data from Aceh Besar, Indonesia, a low endemic, malaria elimination setting with Plasmodium knowlesi, Plasmodium vivax, and Plasmodium falciparum. Malar J 2016; 15:468. doi: 10.1186/s12936-016-1523-z

Additional Information:	
Question	**Response**

Spatial modelling of malaria case associated with environmental factors in South Sumatra, Indonesia

Hamzah Hasyim[1,2,*], Afi Nursafingi[3], Ubydul Haque[4], Doreen Montag[5], David A. Groneberg[1], Meghnath Dhimal[1,6], Ulrich Kuch[1], Ruth Müller[1]

[1]Department of Tropical Medicine and Public Health, Institute for Occupational Medicine, Social Medicine and Environmental Medicine, Faculty of Medicine, Goethe University, Frankfurt am Main, Germany

[2]Faculty of Public Health, Sriwijaya University, Indralaya, South Sumatra, Indonesia

[3]Remote Sensing Program, Faculty of Geography, Gadjah Mada University, Yogyakarta, Indonesia

[4]Department of Public Health, Baldwin Wallace University, Berea, Ohio, USA

[5]Barts and the London School of Medicine, Centre for Primary Care and Public Health, Queen Mary University of London, London, UK

[6]Nepal Health Research Council (NHRC), Ramshah Path, Kathmandu, Nepal

[§]Corresponding author: hamzah.hasyim@stud.uni-frankfurt.de, hamzah@fkm.unsri.ac.id

Abstract

Background

Malaria, a parasitic infection, is a life-threatening disease in South Sumatra Province, Indonesia. This study aimed to investigate the spatial association between malaria occurrence and environmental risk factors.

Methods

The number of confirmed malaria cases was analysed for the year 2013 from the routine reporting of the Provincial Health Office of South Sumatra. The cases were spread over 436 out of 1,613 villages. Six potential ecological predictors of malaria cases were analysed in the different regions using ordinary least square (OLS) and geographically weighted regression (GWR). The global pattern and spatial variability of associations between malaria cases and the selected potential ecological predictors was explored.

Results

The importance of different environmental and geographic parameters for malaria was shown at global and village-level in South Sumatra, Indonesia. The independent variables altitude, distance from forest, and rainfall in global OLS were significantly associated with malaria cases. However, as shown by GWR model and in line with recent reviews, the relationship between malaria and environmental factors in South Sumatra was found to strongly vary spatially in different regions.

Conclusions

A more in-depth understanding of local ecological factors influencing malaria disease as shown in present study may not only be useful for developing sustainable regional malaria control programmes, but can also benefit malaria elimination efforts at village level.

Keywords Geographically weighted regression (GWR), Ordinary least squares (OLS), Akaike information criterion (AIC), Physical environment, Local climate, Sumatra, rainfall, elevation, and distance to water.

Background

Malaria is a significant public health concern worldwide, including Indonesia [1]. The Indonesian government has set a national goal to be malaria-free by 2030. Currently, 24 out of 576 districts in Indonesia classified as being malaria endemic, and an estimated 45% of Indonesia's total population are living at risk of contracting malaria [2]. In South Sumatra Province, the malaria incidence was 0.46 per 1,000 people in 2013. In this province, the proportion of children under five years of age who applied mosquito nets was 32.7%, and the percentage of children under five who treated for fever with antimalarial medication was 89.8% in 2013 [2]. Malaria elimination has been a priority in the Millennium Development Goals (MDGs) [3], and since then has continued to be central to the Sustainable Development Goals (SDGs), supporting Indonesia’s malaria elimination commitments [4]. It is now essential to generate the knowledge that is necessary to develop lasting policies for the national malaria elimination programme.

Several meteorological and environmental variables are risk factors for malaria [5]. Since specific meteorological, environmental factors are at interplay and different factors can affect malaria transmission within a given province [3, 6, 7], it is important to differentiate between factors that influence the vector, the parasite and the host-vector relationship [8]. Atieli *et al.* have demonstrated that the topographic variables elevation, slope, and aspect are influencing the development of *Anopheles* mosquitoes [9]. In north-eastern Venezuela, there is a significant association of malaria transmission with local spatial variations like population density, lowland location, and proximity to aquatic environments [10]. Elsewhere (e.g., Ethiopia and Senegal) spatial relationships between climatic variability like rainfall and malaria occurrence have been demonstrated [11]. Rainfall indirectly benefits *Anopheles* mosquitoes by increasing relative humidity which prolongs adult longevity [12], and the number of breeding places which in turn favours population growth [13]. Temperature and the extent of water availability for larval breeding are crucial factors in the vector life-cycle, affecting transmission [3]. Vectors and parasites are both highly sensitive to any temperature changes, for example, the parasite proliferation depends on temperatures [14]. Temperatures above 28°C have been shown to reduce malaria incidence in Africa [15]. In Indonesia, the optimum temperature for malaria mosquitoes ranges between 25-27°C [3]. For the vector-host relationship, factors such as the distance of people's houses from a river, lakes, pond, distance to the regional urban centre [16-18] distance to forest [19, 20] were shown to be significant predictors.

Spatial nonstationary is a condition in which a simple "global" model cannot define the relationship amongst several sets of variables [21]. Thus, global OLS and local GWR modelling was performed to analyse the environmental risk factors for malaria in South Sumatra that vary geographically at the regional level. The locally different ecological factors studied to potentially predict the response variable 'confirmed malaria case' (Y) are altitude (X1), aspect (X2), distance from the river (X3), distance from lakes and pond (X4), distance from the forest (X5), and rainfall (X6).

Methods

Study area

The study area is located between 1°46' and 4°55' of southern latitude and between 102°4' and 104°41' of eastern longitude and has a total surface area of 46,377.40 km^2 (Fig. 1). It covers eight endemic malaria districts of South Sumatra, Indonesia, namely Lahat, Muara Enim, Musi Banyuasin, Musi Rawas, North Musi Rawas, Ogan Komering Ulu, South Ogan Komering Ulu, and Lubuk Linggau. The topography of the area varies from lowland to mountainous landscapes. The elevation in the study area varies between 0 to 3,159 metres above sea level. [22]. The climate is tropical and wet [22]. In 2013 in South Sumatra, the lowest rainfall was 31 mm (August) in

Lahat district, and the highest rainfall was 613 mm (March) in Palembang City. Monthly average temperatures ranged from 26.6 to 28.3°C and relative humidity from 81 to 88% in 2013 [23].

Indonesia's South Sumatra Province is home to 7,828,700 inhabitants. In 2013, the Gross Regional Domestic Product (GRDP) with oil and gas was IDR 231.68 trillion (17.32 billion USD) [22], based on IDR to USD exchange rates at the time of writing. South Sumatra is an ethnically highly diverse province and home to different local languages and diverse cultural and socioeconomic practices [2]. Local people engage in coffee, rubber and palm oil plantation activities or work in the industrial mining area, which shapes not only people's lives but also the environment [24]. Indonesia contributes significantly to deforestation in Southeast Asia. Recent developments of deforestation have led to unsustainable practices which have resulted in a high frequency of deforestation in some regions and are an important factor influencing malaria incidence [25]. Deforestation has been shown to be connected with malaria incidence in the county (Município) of Mâncio Lima, Acre State, Brazil. There, a cross-sectional study shows 48% increase in malaria incidence are associated with cumulative deforestation within respective health districts in 2006 [26].

Study population and data collection

36,372 patients seeked treatment due to suspected malaria fever in 140 primary health centres (PHC) in the study region South Sumatra during January to December 2013. Among them, 3,578 were laboratory positive for malaria. The cases spread over 436 out of 1,613 villages that were used for unit analysis. The detailed number of malaria cases in different provinces are presented in Fig. 3. The spatial distribution of participants who had confirmed cases of malaria is shown in Fig. 4.

The patients are categorised into “clinical diagnosis”, “suspected malaria” and “confirmed malaria cases”. Categories “clinical diagnosis” or “suspected malaria” are based on the patient's symptoms and physical findings at examination. “confirmed malaria cases” is a case of malaria diagnosed microscopically (examination of blood specimen/preparation) or rapid diagnosis test (RDT) with positive results for *Plasmodium*. Either RDT or microscopic assessment or both were used to confirm the diagnosis of malaria. The malaria diagnostic data were obtained from the regular health information reporting system of the Provincial Health Office of South Sumatra. The data had been collected during 12 months (January to December 2013) at the village level from patients seeking treatment in PHC, locally called Pusat Kesehatan Masyarakat ("puskesmas"), and that were reported monthly to the Provincial Health Office via the malaria programmes in the District Health Offices.

Geographic information

The study area map (Fig. 1) uses the World Geodetic System (WGS84) as its reference coordinate system. As shown in Fig. 5, three stages of working with geographic information were distinguished: data acquisition and processing, data analysis and data presentation [27]. GWR 4.0 version 4.0.90 and Arc GIS 10.3 were used for data processing, analysis, and visualization. Malaria case data were collected from the Provincial Health Department, Ministry of Health (see previous paragraph) as well as topographic (toponymy map, hypsographic map, hydrographic maps, land cover map) and climate data (rainfall map). The primary spatial data were obtained from a topographical map of Indonesia (cartographic material) which has a scale of 1:50,000 and consists of several layers of plots grouped. The malaria input data is aggregated village level data with the village centroid used as the spatial unit. This map consisted of a collection of geographic data presented as thematic layers for land cover, hydrographic data and a sheet of hypsography. Indonesian topographic map known as Peta Rupabumi Indonesia (RBI) was updated in 2014. In

2013, topographic data visualisation has been changed into geodatabase cartography to reduce the steps of creating cartography visualisation in topographic mapping activity [28]. These maps were obtained from the Geospatial Information Agency (BIG) of Indonesia. We got authorization for the use of the topographical map of Indonesia from the Indonesian Geospatial Information Agency. However, restrictions were put to use the availability of these data and therefore are not publicly available. Data were collected by creating a research protocol which is used under license for the current study. The data that backs the findings of the research are served in the main paper.

The forest cover maps were extracted from the land cover map in 2013 on the scale of 1:250.000. The map was sourced from Ministry of Environment and Forestry, Indonesia. The precipitation map (annual average) was obtained by inserting the data of average yearly rainfall from BMKG Climatological Station Class I in Palembang, South Sumatra, Indonesia. The distance between weather observations stations was 50-100 km in flat topography and 10 km in hilly terrain.

Data pre-processing

The malaria distribution map (Fig. 3) was created and six selected explanatory variables plotted (Fig. 2). The altitude map was obtained by interpolation and contouring of the map into a Digital Elevation Model (DEM). Subsequently, the DEM data was converted into a map containing the direction of the slope (aspect). The parameter distance from the river, and distance from lake and pond processed from river, lakes, and ponds maps which were derived from the topographic map whereas distance from the forest processed from forest cover map. These variables were analysed using Euclidean distances. Rainfall parameter was calculated based on annual average rainfall over five years, and it was interpolated from several weather observation stations in study area. The rainfall map (isohyets map) was obtained from the scanned maps which are the result of interpolation and classified into several classes. The map needed to be rectified and digitised to get a digital rainfall map.

Data processing and modelling

The response variable "distribution of malaria cases" and explanatory variables "altitude/aspect", "distance from river", "distance from lake and pond", "distance from forest" and "rainfall" were tested for multicollinearity. Therefore, the values of all explanatory variables were extracted for each case location. An index based on predictive modelling variance, the variance inflation factor (VIF) was used [29]. Multicollinearity could occur when one independent variable was a linear function of another independent variable and previously observed in GWR modelling [30].The pattern of connection between confirmed malaria cases and environmental factors was expressed by the OLS method. Here, OLS model is called global regression model because the existence of local variation had not taken into account in regression so that the estimate of the regression remained constant. Thus, the regression parameters had the same value for each point within the study area. If spatial heterogeneity occurred in regression parameters, then the information that could not be processed by the global regression model was seen as an error. In such cases, the global regression model was less able to explain the actual data phenomenon [31]. A global regression coefficient value close to zero indicated that the explanatory variables had a small effect on the response variable.

As alternative, the GWR model was used to investigate the relationships between response and explanatory variables since the study area was characterized by spatial heterogeneity [32]. A semiparametric GWR4.09 for Windows (provided by Nakaya *et al.* [32]*)* was carried out which is a new release of the windows application software tool for modelling spatially varying relationships among variables by calibrating GWR.

The estimated parameter of the GWR model uses the least squares given the location coordinates as a weighting factor. The influence of the points in this neighbourhood varies according to the distance to the central point [33]. The optimum distance threshold (also known as the bandwidth) or the optimum number of neighbours determined in two ways: by minimising the square of the residuals cross-validation (CV) or by minimising the Akaike Information Criterion (AIC) [34]. At this stage, the type of weighing (kernel type) and optimum bandwidth selection method based were selected on AIC selection criteria. Classic AIC chooses smaller bandwidths in geographically varying coefficients are possible to be under smoothed [32]. In a GWR context, the measurement of utility is the AIC to know whether a global regression model or GWR is most useful [33].

The local GWR model as earlier described is as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \tag{1}$$

Based on the model, y_i, x_{ik}, (u_i, v_i), $\beta_k(u_i, v_i)$, and ε_i were sequentially the response and explanatory variables k to location i, location coordinates to i, realization of the continuous function $\beta_k(u_i, v_i)$ at point i, and Gaussian error to location i. It is noteworthy that the kernel Fixed Gaussian function was used which highlights the optimal bandwidth found by using the Golden section search with the AIC selection criteria. Also, the Gaussian kernel supported the constant weight, and the value became less from the centre of the kernel but never touched zero. The kernel was suitable for fixed kernel because it could prevent the risk of the absence of data in the kernel. The Fixed Gaussian kernel earlier described [33] is as follows:

$$w_{ij} = exp\left[-\left(d_{ij}/b\right)^2\right] \quad (2)$$

Also, w_{ij} was the weight value observed at the location *j* to approximate the calculation of the coefficients on area i, d_{ij} was the Euclidean distance between i and j, and *b* was the size of fixed bandwidth given by the size of metric. The Golden section automatically searched the optimal frequency range value by comparing indicators of the model with the bandwidth size. A positive R^2 indicates a positive correlation. A positive coefficient means X and Y changed in the same direction and if the environmental risk factor increased, then number of confirmed malaria cases increased. Conversely, a negative coefficient means X (explanatory variable) and Y (the response variable) changed in opposite directions. Student's t distribution that had values outside the range of -1.97 and 1.97 formed a critical region with a 0.05 (95% CI) level of significance, whereas values outside the range of -2.59 and 2.59 formed critical regions with a 0.01 (99% CI) level of significance. Step-wise computation performed with these data is shown in the flowchart Fig. 5.

The locally weighed R^2 between the observed and fitted values has been calculated to measure how well the model replicates the local malaria incident values around each observation. A variable is correctly clarified for each location by the model if $R^2 = 1$ with values ranging from 0 to 1.

To compare the performance between global OLS and local GWR, GWR4 software was also used. We performed an ANOVA testing the null hypothesis that the GWR model represents no improvement over a global model. For local GWR, the sufficient number of degrees of freedom was a function of the bandwidth.

Results

Data pre-processing

Multicollinearity does not occur, because the VIF value is less than 10 and the tolerance value is higher than 0.1.

Environmental factors influencing confirmed malaria cases **at global level: OLS model**

The global OLS model reveals that altitude and distance to the forest (negative coefficients) and rainfall (positive coefficient) significantly influence the number of malaria cases. Confirmed malaria cases are more common in regions with high rainfall, lowland and areas adjacent to forest. On the other hand, environmental factors such as aspect or direction towards the slope, distance from the river, and the distance from lakes and pond do not have any significant association with malaria cases. Based on OLS model, each the variables used to assess dependent variable where each factor has a different predictor of malaria incident preferences in GWR model stage.

Environmental factors influencing confirmed malaria cases at local level: GWR model

The results of GWR using Fixed Gaussian are shown in Table 1. The best bandwidth generates 9,184 neighbours and a significant spatial relationship with a specific region has been found. The GWR model provides evidence for a locally different influence of environmental factors on malaria cases as shown by varying parameter estimate value (Fig. 6). “Altitude” and “distance from lake and pond" show a positive association and "aspect" a negative association with malaria incidence in the Northern study area (Musi Banyuasin). "Rainfall" and "distance from river" show a positive association with malaria cases in the Eastern part of Musi Rawas and Lahat. The variables "aspect", distance from lake and pond" and "distance from forest" are positively associated with confirmed malaria cases in large parts of the study area. The significance thresholds of explanatory variables according to Student's t-test in the GWR model are shown in Fig. 7. The local coefficient of determination (local R^2) for confirmed malaria cases at the local level range between 0.18 - 1 (Fig. 8).

Comparison between the two methods OLS and GWR

Like OLS, GWR is a statistical model that provides insights into the relationship between the dependent variable confirmed malaria cases and six independent explanatory variables. GWR is selected as best model based on the residual sum of square, and classic AIC, and the R^2 as stated in Table 2.

The global regression model indicates that the variables have some influence on the study area (Table 3). The global OLS model explains 6.2% variation of malaria incidences by environmental factors ($R^2 = 0.06$). This implies that 93.8% of the malaria incidence is caused by unknown environmental factors related to local variation which are not taken into account in the OLS model [33]. The local GWR explained 68.7% variation in malaria incidences (Y) by environmental factors ($R^2 = 0.69$). The DIFF criterion indicates that the spatial distribution of malaria incidence is associated with the independent variables "altitude", "distance from lakes and pond", "distance from forest", and "rainfall" with local spatial heterogeneity (Table 3). Though the testing of local coefficients for "aspect" and "distance from river" suggests no spatial variability (Table 3).

The GWR model explains the relationship between the response variable ‘confirmed malaria case’ and six explanatory variables significantly better than the global regression model OLS ($F = 2.12$, $P < 0.05$). The best model weights are automatically determined for each location and are mapped in Figure 7.

Discussion

Climate data are frequently used to predict for the spatial, seasonal and interannual variation for malaria transmission, for example the dynamic malaria model forecasting malaria prevalence with seasonal climate published by Hoshen and Morse [35]. The global OLS model revealed here that altitude, distance to forest, and rainfall significantly influence malaria incidence in South Sumatra. Similarly, land use, humidity, altitude and rainfall have been identified by GWR to determine the regional vulnerability to malaria in Purworejo, Indonesia [36]. However, the GWR model considering spatial heterogeneity explains better the association of malaria case with environmental factors in South Sumatra. Likewise in Venezuela, GWR analysis revealed that ecological interactions that act on different scales play a role in malaria transmission and that modelling enhances the understanding of relevant spatiotemporal variability [10]. The environmental factors shown to be significantly associated with malaria cases vary strongly at the village level. This finding is consistent with those obtained in studies in Ethiopia (Addis Ababa), the Amazon region of Brazil (Rondôia), and Cambodia [11, 37, 38]. A validated OLS can lead to a global policy and a validated relationship with GWR is more appropriate to drive to the local system. A geostatistical model based on analysis of residuals and using climatic, population and topographic variables has also been shown to be an important tool for local malaria prediction in Mali. [39].In the highlands of western Kenya, topographic parameters could be used to identify the risk of malaria and thereby helped to improve malaria monitoring or targeted malaria control activities [9].

The relationship of altitude and malaria cases has been shown in present study as well and may relate to the biology of malaria vectors. Globally, *Anopheline* species diversity and density decline from the lowlands to highlands [40]. Accordingly, poor villagers living in forested lowland areas in Papua, Indonesia, were found to be at higher risk of malaria infection than those in the highlands [41]. In contrast, a positive correlation between altitude and the abundance of *Anopheles* mosquitoes has observed in the highlands of Ethiopia, Colombia and Ecuador, particularly in warmer years [42-44]. This observation may be related to the direction towards the slopes as the distribution and density of mosquito populations may be affected by wind direction [45].In an Ethiopian study, minimum temperatures were significantly associated with malaria cases in cold areas, while precipitation was associated with transmission in hot areas [46]. In accordance to many studies, malaria case was significantly associated with rainfall in villages of South Sumatra. Rainfall showed correlation with the incidence of clinical malaria cases in Tubu village, Botswana [47]. Variations in monthly rainfall in rural Tanzania were largely associated with malaria [48]. Rainfall creates oviposition sites for female mosquitoes, whereas humidity is a key parameter for adult mosquito daily survival [49]. *Anopheline* mosquitoes require stagnant water to complete their larval and pupal development. Thus, rainfall affects the transmission of malaria by providing water to create aquatic habitats. The number of malaria cases was significantly positively connected with higher winter rainfall, but also with a higher average maximum temperature and significantly negatively associated with increasing distance from water bodies in South Africa [50]. Southern Africa Development Community estimates the positive correlation between increasing rainfall and the number of cases in Botswana during 2013 and 2014 [51].

Next to climatic and environmental factors, distance of houses to a forest are interrelated through anthropogenic activities influencing the local and regional climate [52, 53]. These observations can be confirmed for the relationship of malaria case with distance to lake, pond and forest for South Sumatra. A cross-sectional view in Brazil revealed for example that malaria incidence across health districts is positively correlated with the percentage of aggregated deforestation [54]. Indonesia contributes indeed significantly to deforestation in Southeast Asia. *Anopheles* was reported from eight sources at 47 independent sites. The first record of *Anopheles parangensis* from Sumatra was reported by O'Connor and Sopa (1981), but with no details on location [55]. *Anopheles (Cellia) leucosphyrus* is considered to be of epidemiological importance for malaria transmission in forested areas of Sumatra [55]. In current research, the main *Anopheles* vector diversity in each study area was however not investigated.

Present study have identified Lahat as the South Sumatran district in which environmental factors were of greatest relevance for malaria incidence. Lahat District has both lowland and mountain regions and is home to diverse ethnic groups, such as the Gumai who live along the rivers of the highland areas [56].

One of the key activities for malaria elimination should be the establishment of systems and tools to reduce disease burden where local transmission is high. By comparing the local GWR model with the global OLS model, it became apparent that GWR yielded new information about the spatial variation of malaria incidence and thereby better explains local phenomena. The variability of predicted malaria rates in our study was due to environmental and geographical local differences [8]. GWR should be used as a diagnostic model discovering spatially varying relationships between confirmed malaria cases and environmental factors. The use of GWR allows the uncovering of significant environmental variation for malaria incidence, which has previously been unobservable in a specific location [57].

Limitations of research

Due to practical constraints, this study was unable to encompass the entirety of environmental factors, particularly climate parameters, temperature and humidity, for which only limited data were available and hence not-representative data could not be included. Also the factor land use was eliminated. Malaria location information was plotted using a village centre approach which ignored all other locations where actual infections may have occurred (e.g., forests, plantations, lakes, rivers). The number of positive malaria per village, did not include the specific coordinates of each positive malaria case and thus, each positive case was placed in the centre of the settlement. Therefore, if land use variables would be involved, there will very likely be a strong bias. However, these eliminated or uninvestigated variables may be correlated with existing variables, for example, the temperature connected with altitude and with aspect or direction of the slope. In the same way, land use may be associated with the distance from the river and the distance from lakes and ponds. Thus, although these parameters (temperature, humidity, land use) had eliminated before analysis, these environmental factors were represented by our chosen set of variables. In the future, additional explanatory variables should be addressed to provide a comprehensive review of malaria in the study area. It should comprise, for example, the behavior of mosquito vectors and that of community members, the access to and the delivery of health services, and other eco-bio-social factors that affect the incidence of malaria. Despite these limitations, our study sheds light on relevant, not only in regional but also local realities regarding environmental variation and sociocultural practice which might interplay with vector-host relationships and provide a suitable environment for malaria mosquitoes.

Conclusion

In the present study, the importance of different environmental and geographic parameters for malaria disease was shown at global and village-level in South Sumatra, Indonesia. The independent variables altitude, distance from forest, and rainfall in global OLS were significantly associated with malaria cases. As shown by GWR model and in line with recent reviews, the relationship between malaria and environmental factors in South Sumatra was found to vary spatially in different regions. A more in-depth understanding of local ecological factors influencing confirmed malaria casecannot only be used for developing sustainable regional malaria control programs but can also benefit malaria elimination efforts at village level.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We express our gratitude to the Ministry of Research, Technology and Higher Education of the Republic of Indonesia for supporting the first author (HH) with a PhD scholarship in the context of the cooperation between Sriwijaya University and Goethe University. The authors wish to thank the Head of the Geospatial Information Agency (BIG) Indonesia for access to the digitised map, and the Head of the Indonesian Agency for Meteorological, Climatological and Geophysics (BMKG) climatology station Class I, in Kenten Palembang, for providing interpolated data of annual rainfall averages. We are also grateful to the Head and staff of the Health Office of South Sumatra Province who kindly permitted us to do further analyses on the malaria data from the study area.

Ethics approval and consent to participate

Not applicable.

Authors' contributions

HH was responsible for the management of this study, design and collection of data. Under the supervision of HH, NA performed the data analysis and was responsible for data acquisition, pre-processing, and processing. HH, UH, DM, MD, DAG, UK and RM contributed to the interpretation and visualisation of the results. HH, UH, DM, MD, NA, UK and RM wrote the paper. All authors read and approved the final version of the manuscript.

References

1. Ministry of Health Republic of Indonesia. Decree of the Minister of Health Republic of Indonesia Number 293 / MENKES / SK / IV / 2009 on Elimination of Malaria in Indonesia, Jakarta. Ministry of Health (MoH) of Indonesia, 2009. Retrieved from http://perpustakaan.depkes.go.id/.
2. Ministry of Health Republic of Indonesia. South Sumatera Province Health Profile 2013. Palembang: South Sumatra Provincial Health Office (Indonesia), 2014. Retrieved from http://www.depkes.go.id/.
3. Ministry of Health Republic of Indonesia. Malaria management: guideline. Jakarta: Directorate of Vector Borne Disease and Zoonosis Control, Directorate General of Disease Prevention and Control, Ministry of Health (MoH) of Indonesia, 2014. Retrieved from http://eliminasimalaria.blogspot.de/p/download.html.
4. WHO and UNICEF. Achieving the malaria MDG target: reversing the incidence of malaria 2000-2015. Geneva: World Health Organization, 2015. Retrieved from http://www.who.int/malaria/, isbn:978 92 4 150944 2
5. WHO. Regional action framework for malaria control and elimination in the Western Pacific: 2016-2020. Manila, Philippines WHO Regional Office for the Western Pacific, 2017. Retrieved from http://iris.wpro.who.int/handle/10665.1/13578, isbn:9789290618157
6. Elyazar IR, Hay SI, Baird JK. Malaria distribution, prevalence, drug resistance and control in Indonesia. Adv Parasitol. 2011;74:41-175.
7. WHO Country Office for Indonesia and Ministry of Health Indonesia. National malaria control programme review: Republic of Indonesia. Jakarta: WHO Country Office for

Indonesia and Ministry of Health, Indonesia, 2013. Retrieved from http://www.who.int/iris/handle/10665/253960, isbn:978 979 19477 4 9

8. Loha E, Lindtjorn B. Model variations in predicting incidence of *Plasmodium falciparum* malaria using 1998-2007 morbidity and meteorological data from south Ethiopia. Malar J. 2010;9:166.

9. Atieli HE, Zhou G, Lee M-C, Kweka EJ, Afrane Y, Mwanzo I, et al. Topography as a modifier of breeding habitats and concurrent vulnerability to malaria risk in the western Kenya highlands. Parasit Vectors. 2011;4:241.

10. Grillet ME, Barrera R, Martinez JE, Berti J, Fortin MJ. Disentangling the effect of local and global spatial variation on a mosquito-borne infection in a neotropical heterogeneous environment. Am J Trop Med Hyg. 2010;82:194-201.

11. Alemu A, Abebe G, Tsegaye W, Golassa L. Climatic variables and malaria transmission dynamics in Jimma town, South West Ethiopia. Parasit Vectors. 2011;4:30.

12. Yamana TK, Eltahir EA. Incorporating the effects of humidity in a mechanistic model of Anopheles gambiae mosquito population dynamics in the Sahel region of Africa. Parasit Vectors. 2013; 6:235.

13. Abiodun GJ, Maharaj R, Witbooi P, Okosun KO. Modelling the influence of temperature and rainfall on the population dynamics of Anopheles arabiensis. Malar J. 2016;15:364.

14. Weaver HJ. Climate change and human parasitic disease. In: Climate change and global health. CABI Nosworthy Way Wallingford UK: ©CAB International; 2014:95. Retrieved from https://www.cabi.org/cabebooks/ebook/20143328432.

15. Mordecai EA, Paaijmans KP, Johnson LR, Balzer C, Ben-Horin T, Moor E, et al. Optimal temperature for malaria transmission is dramatically lower than previously predicted. Ecol Lett. 2013;16:22-30.

16. Villarreal-Treviño C, Penilla-Navarro RP, Vázquez-Martínez MG, Moo-Llanes DA, Ríos-Delgado JC, Fernández-Salas I, et al. Larval habitat characterization of *Anopheles darlingi* from its northernmost geographical distribution in Chiapas, Mexico. Malar J. 2015;14:517.

17. Ernst KC, Lindblade KA, Koech D, Sumba PO, Kuwuor DO, John CC, et al. Environmental, socio-demographic and behavioural determinants of malaria risk in the western Kenyan highlands: a case-control study. Trop Med Int Health. 2009;14:1258-65.

18. Haque U, Scott LM, Hashizume M, Fisher E, Haque R, Yamamoto T, et al. Modelling malaria treatment practices in Bangladesh using spatial statistics. Malar J. 2012; 11:63.

19. Dhimal M, O'Hara RB, Karki R, Thakur GD, Kuch U, Ahrens B. Spatio-temporal distribution of malaria and its association with climatic factors and vector-control interventions in two high-risk districts of Nepal. Malar J. 2014;13:457.

20. Haque U, Glass GE, Bomblies A, Hashizume M, Mitra D, Noman N, et al. Risk factors associated with clinical malaria episodes in Bangladesh: a longitudinal study. Am J Trop Med Hyg. 2013; 88:727-32.

21. Brunsdon C, Fotheringham A, Charlton M. Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr Anal. 1996;28:281-98.

22. Badan Pusat Statistik. Sumatera Selatan dalam angka (Sumatera Selatan in figures) 2013. Sumatera Selatan: BPS – Statistics Sumatera Selatan, 2014. Retrieved from http://sumsel.bps.go.id.

23. BMKG. Annual average rainfall 2007 - 2013 Indonesian Agency for Meteorological, Climatological and Geophysics (BMKG) Climatology Station Class I Kenten Palembang, 2013. Retrieved from www.bmkg.go.id

24. Lee JSH, Abood S, Ghazoul J, Barus B, Obidzinski K, Koh LP. Environmental impacts of large-scale oil palm enterprises exceed that of smallholdings in Indonesia. Conserv Lett. 2014;7:25-33.

25. Abood SA, Lee JSH, Burivalova Z, Garcia-Ulloa J, Koh LP. Relative contributions of the logging, fiber, oil palm, and mining industries to forest loss in Indonesia. Conserv Lett. 2015;8:58-67.

26. Olson SH, Gangnon R, Silveira GA, Patz JA. Deforestation and Malaria in Mâncio Lima County, Brazil. Emerg Infect Dis. 2010;16:1108-15.

27. Huisman O, De By R. Principles of geographic information systems. The Netherlands: The International Institute for Geo-Information Science and Earth Observation, 2009. Retrieved from http://www.gdmc.nl/oosterom/PoGISHyperlinked, isbn:978–90–6164–269–5

28. Perdana AP, Juniati E, Hidayat J. Transformation of topographic data visualization from freehand drawing to cartographic representation. Jurnal Ilmiah Geomatika. 2013;19:2. Retrieved from http://jurnal.big.go.id/.

29. Halimi M, Farajzadeh M, Delavari M, Takhtardeshir A, Moradi A. Modelling spatial relationship between climatic conditions and annual parasite incidence of malaria in southern part of Sistan & Balouchistan Province of Iran using spatial statistic models. Asia Pac J Public Health. 2014;4:S167-S72.

30. Wheeler D, Tiefelsdorf M. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. J Geogr Syst. 2005;7:161-87.

31. Yasin H. Pemilihan variabel pada model geographically weighted regression. Media Statistika. 2011;4:63-72.

32. Nakaya T. GWR4 Windows Application for Geographically Weighted Regression Modelling. Kyoto: Ritsumeikan University, Department of Geography. 24 March 2016. Retrieved from https://raw.githubusercontent.com/gwrtools/gwr4/master/GWR4manual_409.pdf,

33. Fotheringham A, Brunsdon C, Charlton M. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. University of Newcastle, UK: John Wiley & Sons, 2002. Retrieved from https://www.wiley.com/, isbn:0-471-49616-2

34. Rodrigues M, de la Riva J, Fotheringham S. Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. Applied Geogr. 2014;48:52-63.

35. Hoshen MB, Morse AP. A weather-driven model of malaria transmission. Malar J. 2004;3:32.

36. Widayani P, Danoedoro P, Mardihusodo SJ. Pemodelan spasial kerentanan wilayah terhadap penyakit menular terkait lingkungan berbasis penginderaan jauh (kasus malaria, leptospirosis dan tuberkulosis di sebagian wilayah Provinsi Jawa Tengah dan DIY). Universitas Gadjah Mada, Fakultas Geografi; 2016. Retrieved from http://etd.repository.ugm.ac.id/.

37. de Castro MC, Monte-Mór RL, Sawyer DO, Singer BH. Malaria risk on the Amazon frontier. Proc Natl Acad Sci USA. 2006;103:2452-57.

38. Dysoley L, Kaneko A, Eto H, Mita T, Socheat D, Börkman A, et al. Changing patterns of forest malaria among the mobile adult male population in Chumkiri District, Cambodia. Acta Trop. 2008;106:207-12.

39. Kleinschmidt I, Bagayoko M, Clarke G, Craig M, Le Sueur D. A spatial statistical approach to malaria mapping. Int J Epidemiol. 2000;29:355-61.

40. Ovadje L, Nriagu J. Malaria as an Environmental Disease. In: Encyclopedia of Environmental Health. Burlington: Elsevier; 2011:558–67 Retrieved from https://www.researchgate.net/publication/268515612_Ovadje_L_Nriagu_J_2011_Malaria_as_an_Environmental_Disease_In_Nriagu_JO_ed_Encyclopedia_of_Environmental_Health_volume_3_pp_558-567_Burlington_Elsevier.

41. Hanandita W, Tampubolon G. Geography and social distribution of malaria in Indonesian Papua: a cross-sectional study. Int J Health Geogr. 2016;15:13.

42. Siraj AS, Santos-Vega M, Bouma MJ, Yadeta D, Ruiz Carrascal D, Pascual M. Altitudinal changes in malaria incidence in highlands of Ethiopia and Colombia. Science. 2014;343:1154-8.

43. Pinault LL, Hunter FF. New highland distribution records of multiple Anopheles species in the Ecuadorian Andes. Malar J. 2011;10:236.

44. Alimi TO, Fuller DO, Qualls WA, Herrera SV, Arevalo-Herrera M, Quinones ML, et al: Predicting potential ranges of primary malaria vectors and malaria in northern South America based on projected changes in climate, land cover and human population. Parasit Vectors. 2015;8:431.

45. Messina JP, Taylor SM, Meshnick SR, Linke AM, Tshefu AK, Atua B, et al. Population, behavioural and environmental drivers of malaria prevalence in the Democratic Republic of Congo. Malar J. 2011;10:161.

46. Midekisa A, Beyene B, Mihretie A, Bayabil E, Wimberly MC. Seasonal associations of climatic drivers and malaria in the highlands of Ethiopia. Parasit Vectors. 2015;8:339.

47. Chirebvu E, Chimbari MJ, Ngwenya BN, Sartorius B. Clinical malaria transmission trends and its association with climatic variables in Tubu Village, Botswana: a retrospective analysis. PLoS One. 2016;11:e0139843.

48. Thomson MC, Ukawuba I, Hershey CL, Bennett A, Ceccato P, Lyon B, et al. Using rainfall and temperature data in the evaluation of national malaria control programs in Africa. Am J Trop Med Hyg. 2017;97:32-45.

49. Day JF. Mosquito oviposition behavior and vector control. Insects 2016;7:65.

50. Kleinschmidt I, Sharp BL, Clarke GPY, Curtis B, Fraser C. Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in KwaZulu Natal, South Africa. Am J Epidemiol. 2001;153:1213-21.

51. Chihanga S, Haque U, Chanda E, Mosweunyane T, Moakofhi K, Jibril HB, et al. Malaria elimination in Botswana, 2012–2014: achievements and challenges. Parasit Vectors. 2016;9:99.

52. Steffen W, Persson Å, Deutsch L, Zalasiewicz J, Williams M, Richardson K, et al. The anthropocene: from global change to planetary stewardship. Ambio 2011;40:739-61.

53. Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li W, Nobre CA. Climate change, deforestation, and the fate of the Amazon. Science. 2008;319:169-72.

54. Olson SH, Gangnon R, Silveira GA, Patz JA. Deforestation and malaria in Mancio Lima county, Brazil. Emerg Infect Dis. 2010;16:1108.

55. Elyazar IR, Sinka ME, Gething PW, Tarmidzi SN, Surya A, Kusriastuti R, et al. The distribution and bionomics of Anopheles malaria vector mosquitoes in Indonesia. Adv Parasitol. 2013;83:173-266.

56. Sakai M. Remembering origins ancestors and places in the Gumai society of South Sumatra. In: The Poetic Power of Place comparative perspectives on Austronesian ideas of locality Australia: ANU Press; 2006:42-55. Retrieved from http://www.jstor.org/stable/j.ctt2jbjrm.2.

57. Barati M, Keshavarz-Valian H, Habibi-Nokhandan M, Raeisi A, Faraji L, Salahi-Moghaddam A. Spatial outline of malaria transmission in Iran. Asian Pac J Trop Med. 2012;5:789-95.

Figure legends

Fig. 1 Map of the study area covering one city and seven districts of South Sumatra Province, Indonesia.

Fig. 2 Each explanatory variable mapped in the study area.

Fig. 3 Malaria cases and their geographical locations in the study area.

Fig. 4 Malaria cases at village level.

Fig. 5 Flow chart of the research strategy.

Fig. 6 Predicted value from GWR for parameter estimates of explanatory variables of malaria cases in the study area.

Fig. 7:Student’s test significance (95% and 99% confidence interval) for each explanatory variable and village location.

Fig. 8 Goodness-of-fit of GWR model (local Coefficient of determination R^2) for malaria cases associated with environmental factors in South Sumatra, Indonesia.

Table 1: GWR result based on Fixed Gaussian (distance) kernel function for geographical weighting.

Bandwidth and Geographic Ranges	Value
Bandwidth size:	9,184.47
Diagnostic information	
Residual sum of squares:	33,549.28
Classic AIC:	3,482.17
BIC/MDL:	4,198.30
CV:	178.92
R square:	0.69
Adjusted R square:	0.41

Table 2: Comparison between Global OLS and Local GWR models

Value	OLS	GWR
Residual sum of square	100,625.26	33,549.28
Classic AIC	3,625.82	3,482.17
R^2	0.06	0.69
Adjusted R^2	0.05	0.41

Table 3: The result of global regression model and geographical variability test of local coefficients for six environmental factors.

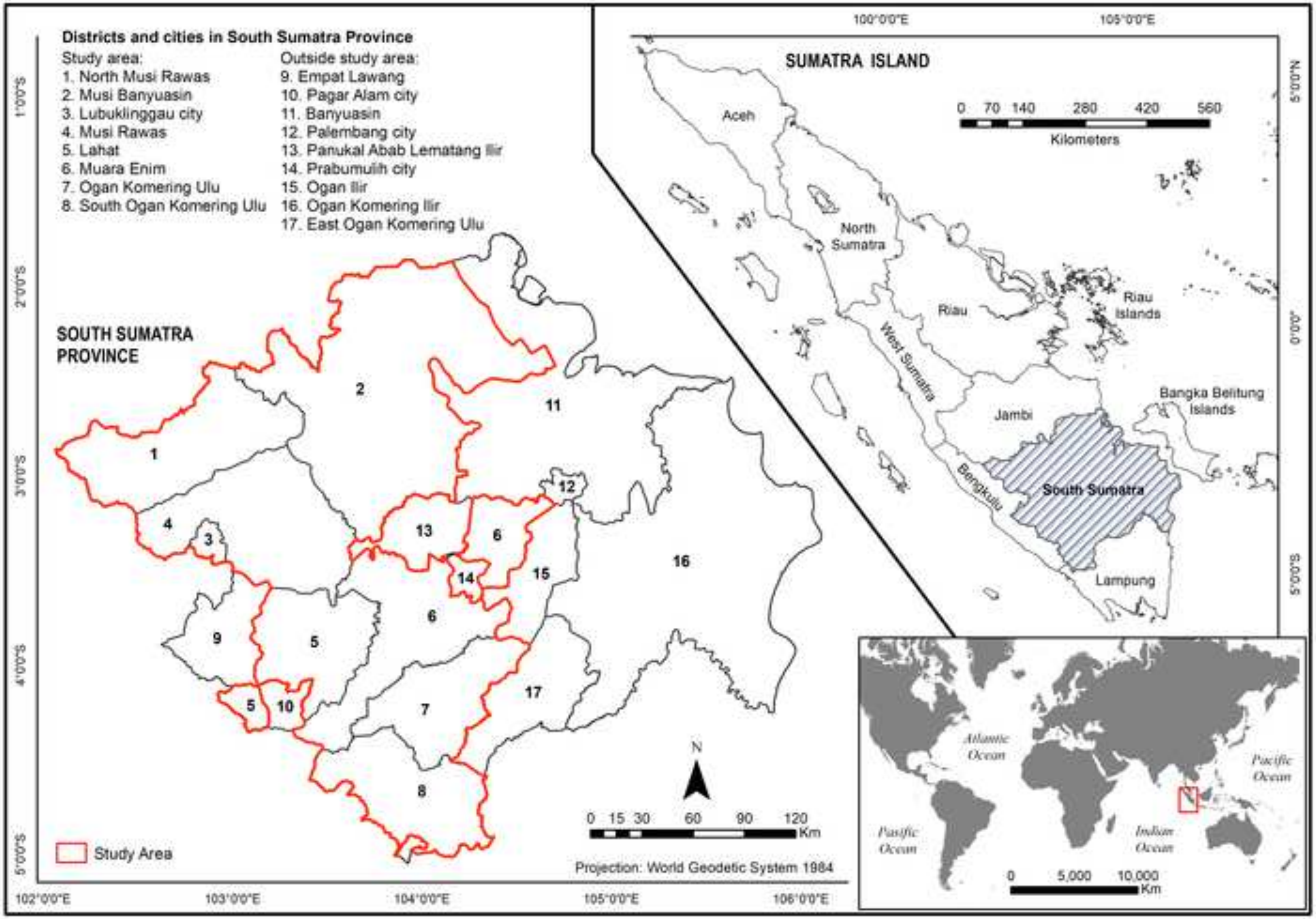
Variables	Global regression model output				Geographical variability test			
	Estimate	SE	T value	P value	F	DOF for F test		DIFF of Criterion
Intercept	7.98	4.63	1.72	0.04	33.20	10.48	261.38	-347.99
"Altitude (X1)"	-0.02	0.00	-4.03	0.00	0.24	12.02	261.38	19.19
"Aspect (X2)"	-0.01	0.01	-1.60	0.05	0.55	22.68	261.38	24.91
"Distance from the river (X3)"	0.00	0.00	-0.84	0.24	1.84	18.15	261.38	-16.03
"Distance from lakes and pond (X4)"	0.00	0.00	0.39	0.71	0.90	15.04	261.38	7.99
"Distance from forest (X5)"	0.00	0.00	-3.69	0.00	2.99	14.61	261.38	-38.12
"Rainfall (X6)"	0.00	0.00	2.38	0.02	13.07	10.17	261.38	-158.91

Table 4: ANOVA testing the null hypothesis that the GWR model represents no improvement over a global model.

Source	SS	DF	MS	F Count	F Table
Global Residuals	100,625.26	429.00			
GWR Improvement	67,075.98	197.74	339.22		
GWR Residuals	33,549.28	231.26	145.07	2.34	2.12

Figure 1

Click here to download Figure Figure 1.png



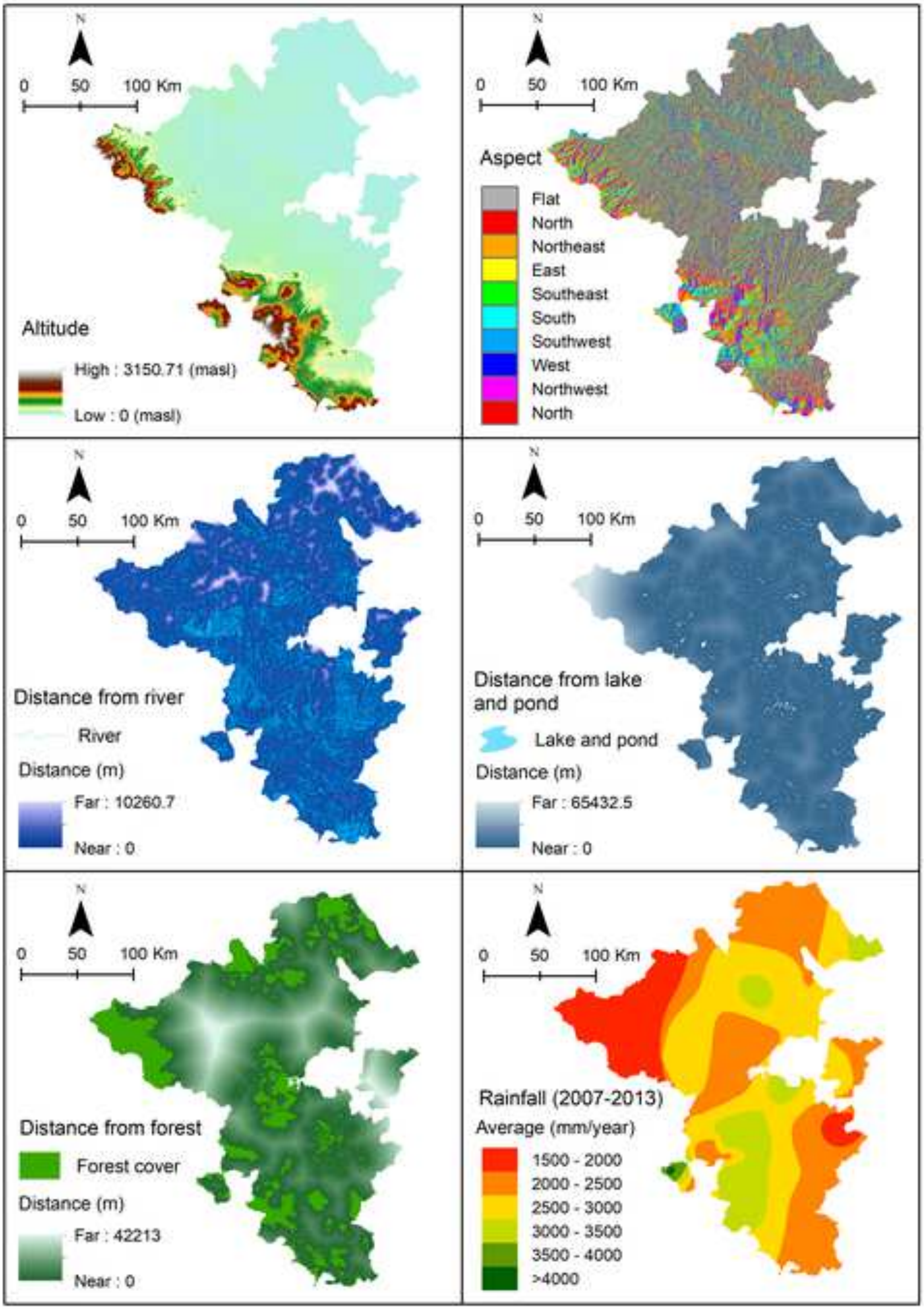
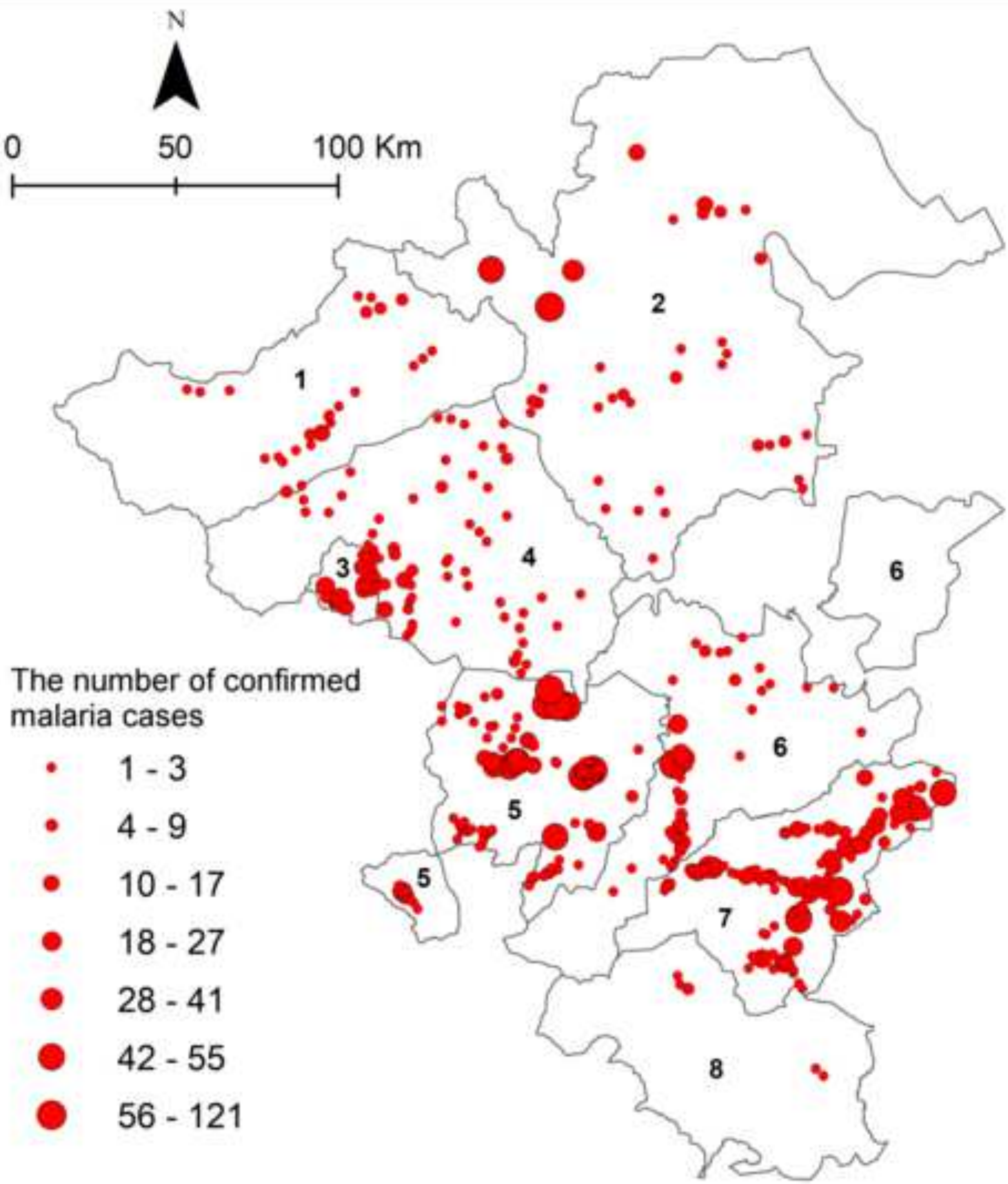

N
0 50 100 Km
Altitude
High : 3150.71 (masl)
Low : 0 (masl)
Aspect
Flat
North
Northeast
East
Southeast
South
Southwest
West
Northwest
North
Distance from river
River
Distance (m)
Far : 10260.7
Near : 0
Distance from lake and pond
Lake and pond
Distance (m)
Far : 65432.5
Near : 0
Distance from forest
Forest cover
Distance (m)
Far : 42213
Near : 0
Rainfall (2007-2013)
Average (mm/year)
1500 - 2000
2000 - 2500
2500 - 3000
3000 - 3500
3500 - 4000
>4000

Figure 3

Click here to download Figure Figure 3 .png



The primary Anopheles malaria vectors in South Sumatra Provinces:
An. letifer, An. nigerrimus, An. maculatus, An. sinensis, An. barbirostris, An. vagus, and An. sundaicus

Source: Vector and Animal-Borne Disease Control Unit of Research and Development, National Institute of Health Research and Development (NIHRD), Ministry of Health (Indonesia) at Baruraja and some relevant references

Figure 4 Click here to download Figure Figure 4.png

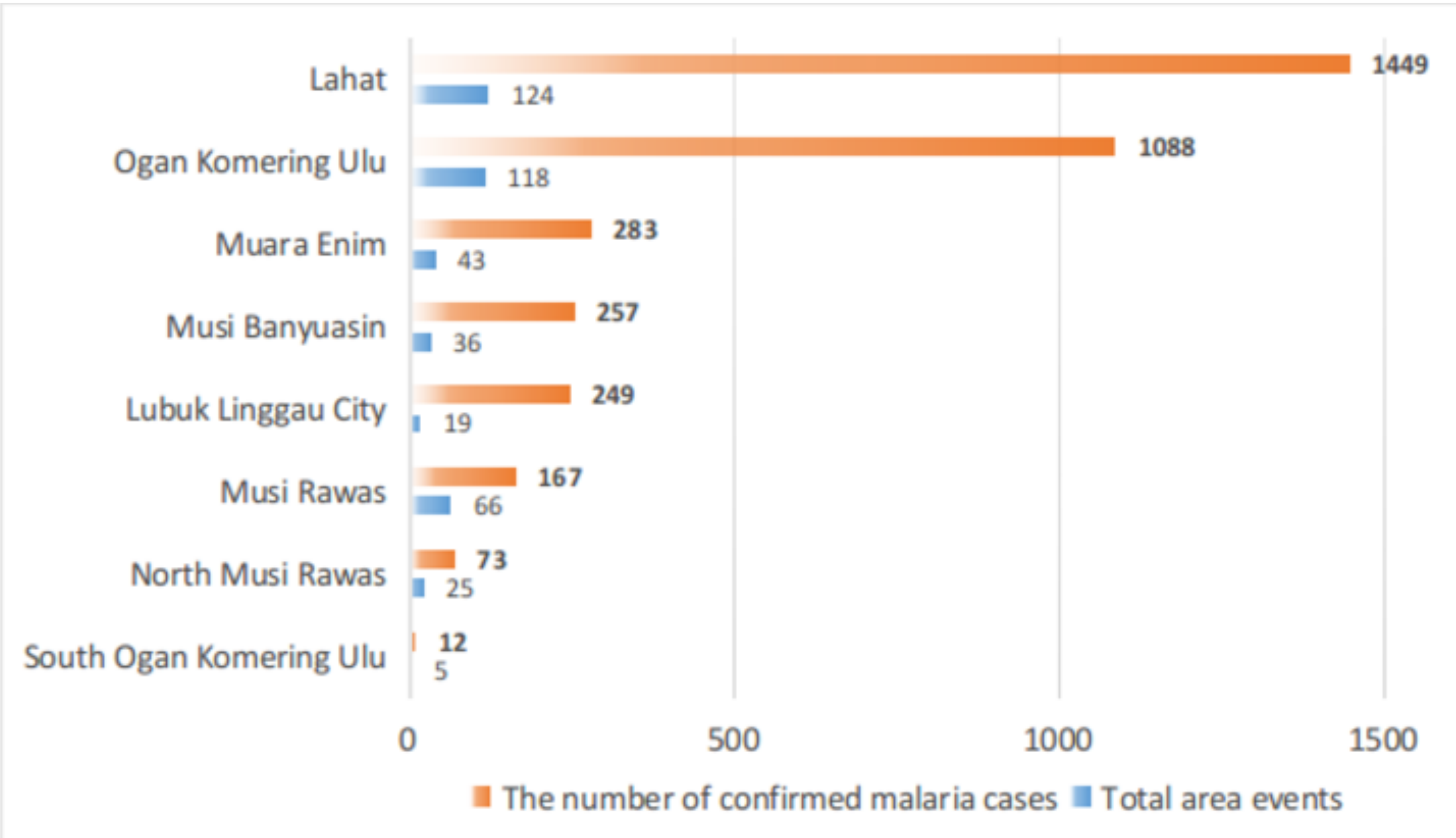


Figure 5

Click here to download Figure Figure 5.png

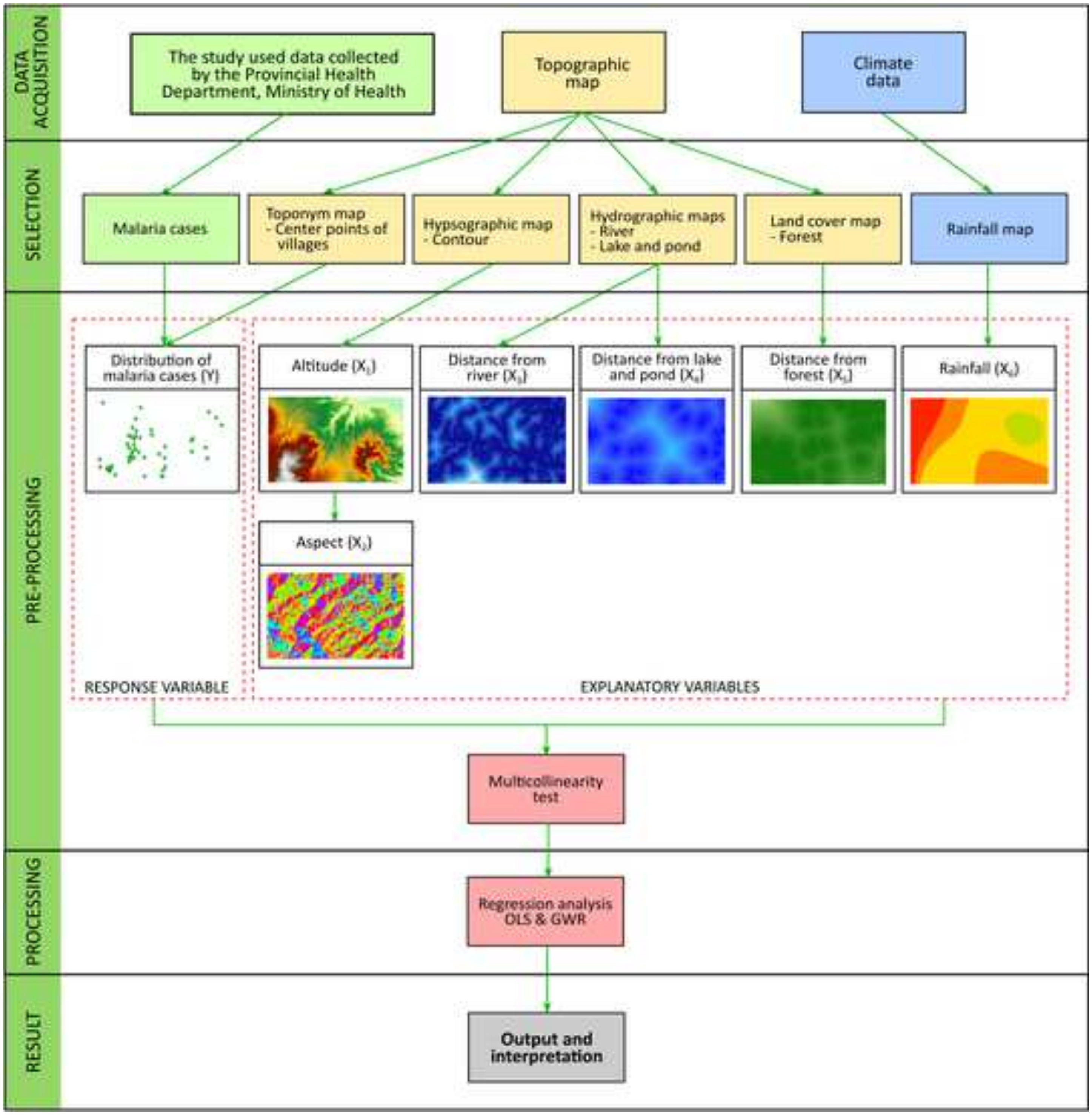


Figure 6

Click here to download Figure Figure 6.png

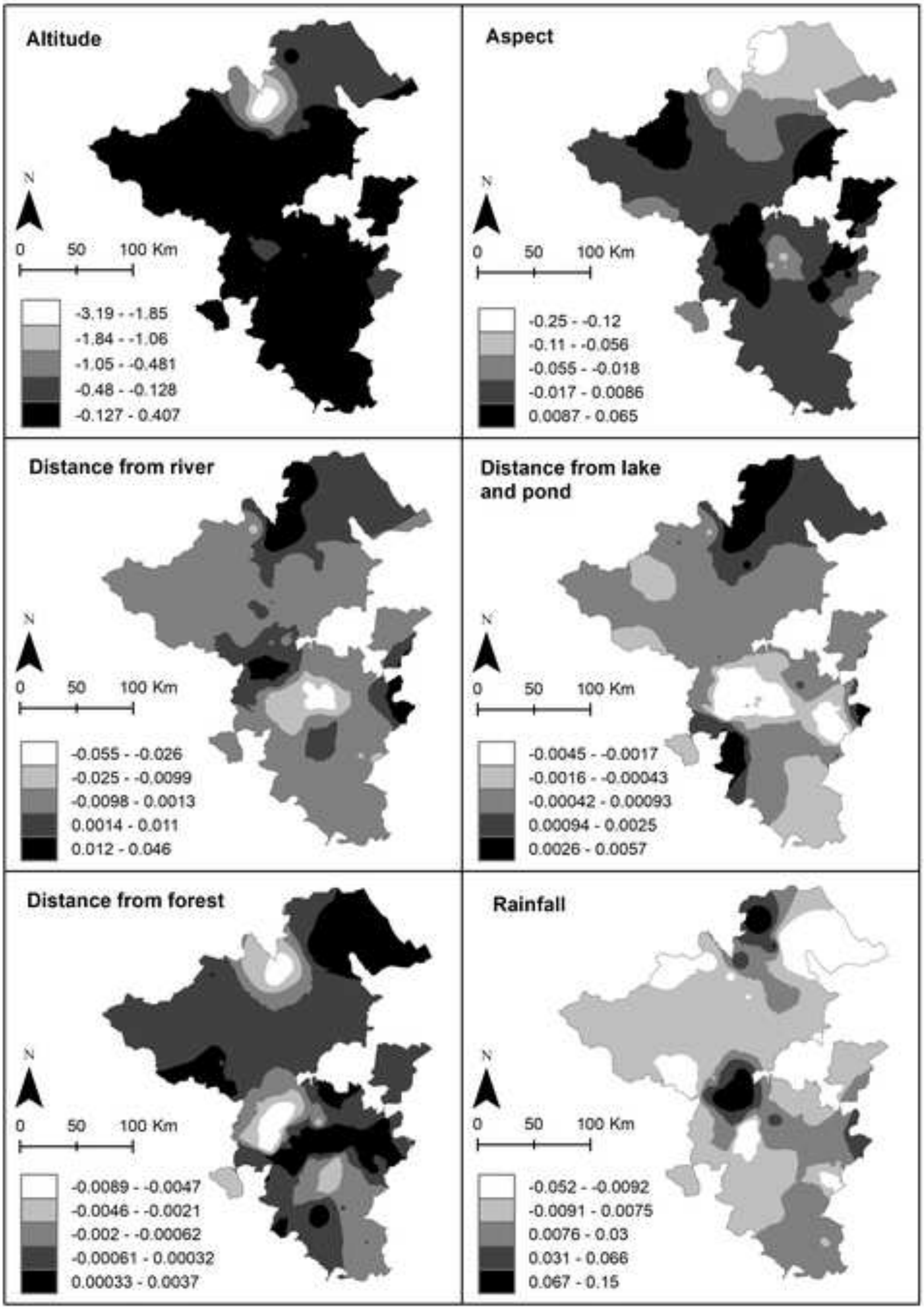


Figure 7

Click here to download Figure Figure 7.png

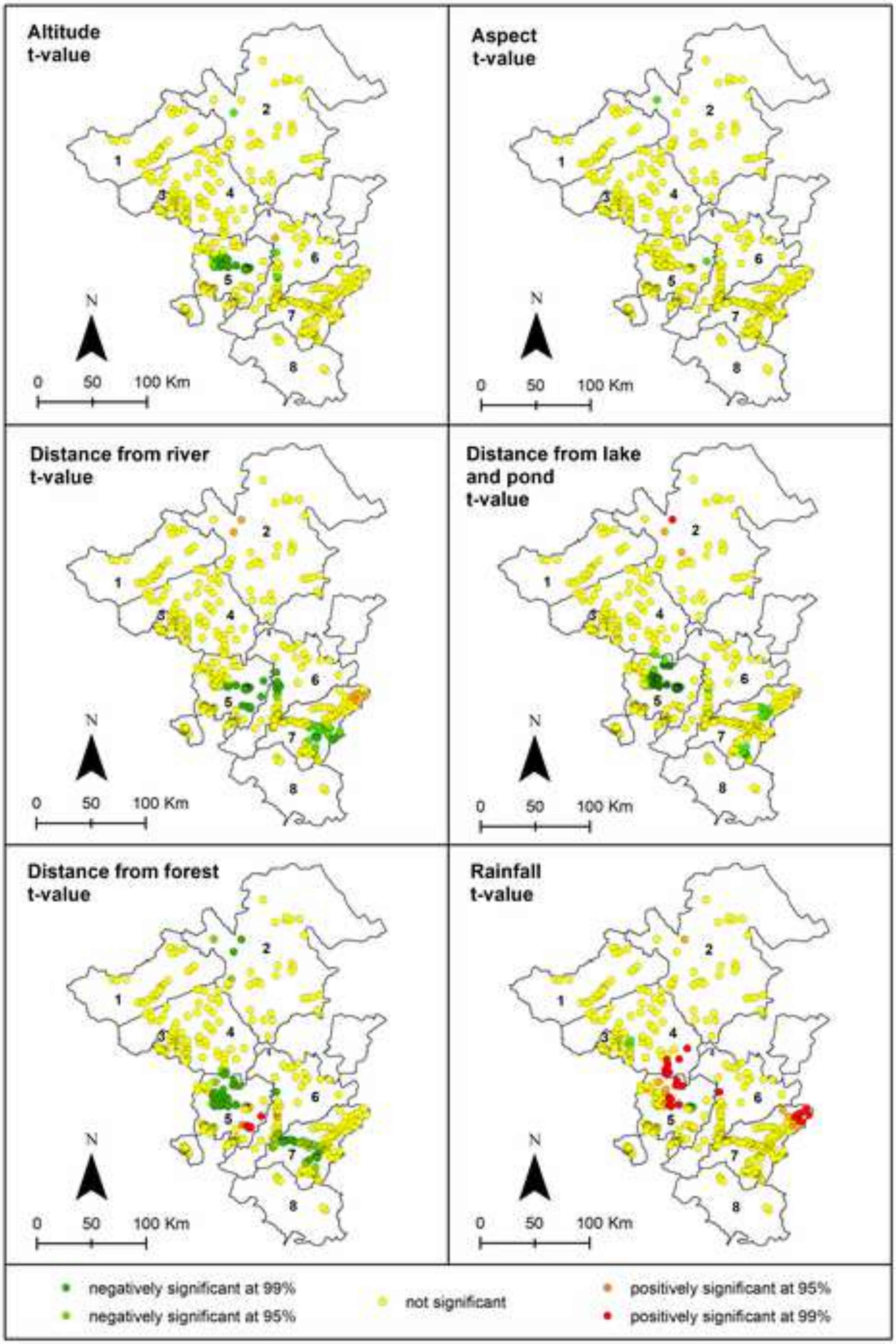
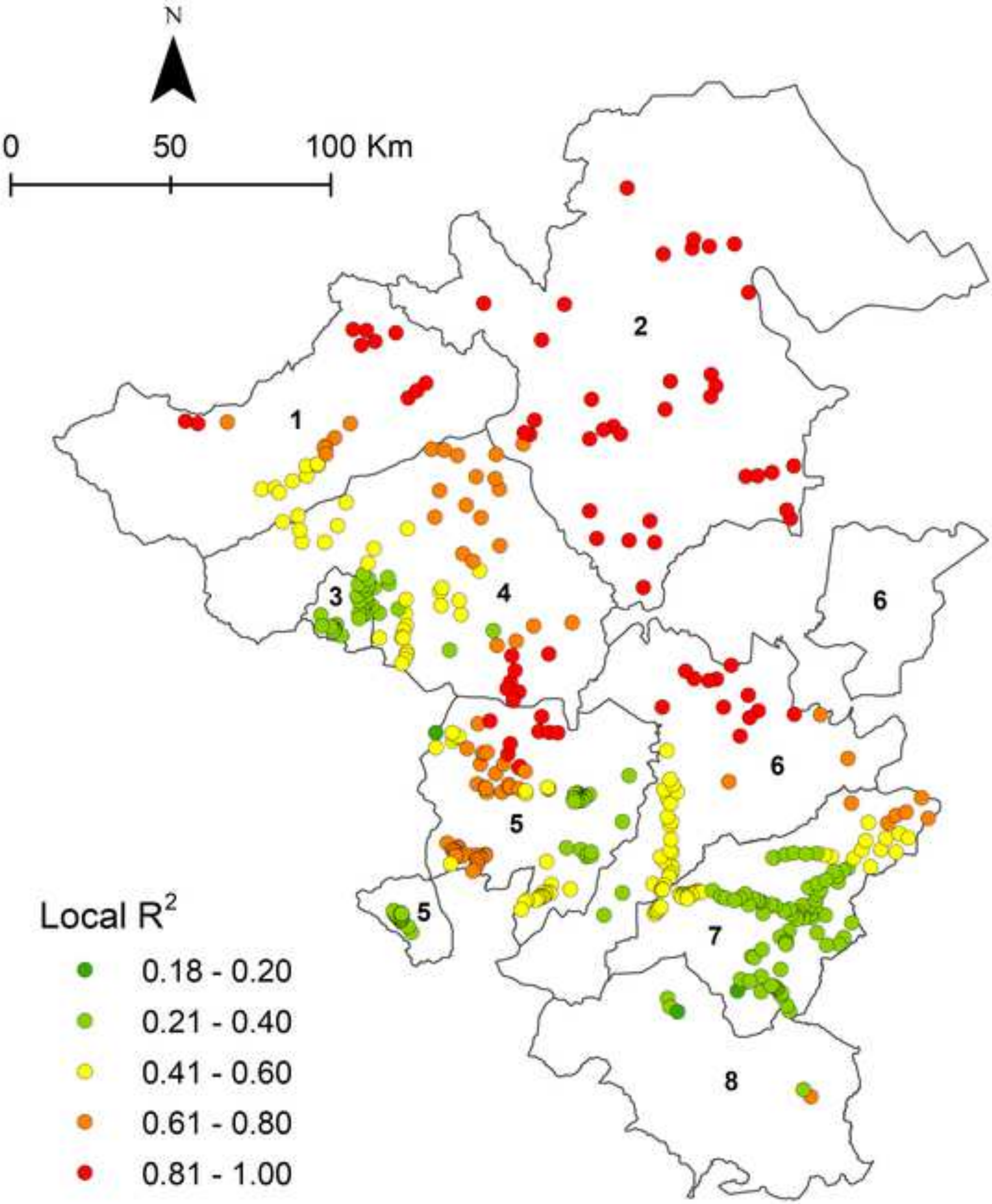


Figure 8

Click here to download Figure Figure 8.png



Date:	11 Feb 2018
To:	"Hamzah Hasyim" hamzah.hasyim@gmail.com;hamzah.hasyim@stud.uni-frankfurt.de;hamzah@fkm.unsri.ac.id
From:	"Malaria Journal Editorial Office" magesh.murugappan@springer.com
Subject:	Decision has been reached on your submission to Malaria Journal - MALJ-D-17-00578R3

MALJ-D-17-00578R3
Spatial modelling of malaria case associated with environmental factors in South Sumatra, Indonesia
Hamzah Hasyim, PhD candidate; Afi Nursafingi, M.Sc; Ubydul Haque, PhD; Doreen Montag, DPhil; David Groneberg, Prof,Dr,PhD; Meghnath Dhimal, PhD; Ulrich Kuch, Dr; Ruth Müller, Dr
Malaria Journal

Dear Mr Hasyim,

I am pleased to inform you that your manuscript "Spatial modelling of malaria case associated with environmental factors in South Sumatra, Indonesia" (MALJ-D-17-00578R3) has been accepted for publication in Malaria Journal.

Before publication, our production team will check the format of your manuscript to ensure that it conforms to the standards of the journal. They will be in touch shortly to request any necessary changes, or to confirm that none are needed.

Any final comments from our reviewers or editors can be found, below. Please quote your manuscript number, MALJ-D-17-00578R3, when inquiring about this submission.

We look forward to publishing your manuscript and I do hope you will consider Malaria Journal again in the future.

Best wishes,

Marcel Hommel, MD, PhD
Malaria Journal
https://malariajournal.biomedcentral.com/

Comments:

****

(ID MALJ-D-17-00578) - Spatial Modelling of Malaria Incidence in South Suma...

You have requested to pay offline for the article-processing charge. Your order number is **0007001606**.

An invoice will be sent as a PDF attachment to the email address you supplied within 72 working hours.

Order Confirmation

Biomed Central Ltd
The Campus
4 Crinan Street
London N1 9XW
UK

Telephone: +44 (0) 20 3192 2000
Facsimile: +44 (0) 20 3192 2009/10
Email: ORSupport@springernature.com

Biomed Central is part of Springer Science+Business Media, Tax No. GB 823826326 Registered in England and Wales No. 03680030

Invoicing information

Manuscript details:

MAN_ID	MALJ-D-17-00578
Title	Spatial Modelling of Malaria Incidence in South Sumatra, Indonesia
Journal	Malaria Journal
Authors	Hamzah Hasyim,Afi Nursafingi,Ubydul Haque,Meredian Alam,Doreen Montag,David G...
Submitted by	Hamzah Hasyim (hamzah.hasyim@gmail.com)
Submitted	06 Sep 2017

Invoice address:

Email	hamzah.hasyim@gmail.com
Phone number	0049 69 6301 6650
Department	Occupational Health, Safety, and Environmental Health Department
Organization	Goethe Universit\u00E4t Frankfurt
Address line 1	Haus 9
Address line 2	Theodor-Stern-Kai 7
City	Frankfurt am Main
State/Province	Hessen
Zip/Postcode	60590
Country	Germany

Order details:

OrderNumber	0007001606
Order date	13 Feb 2018
Purchase order no	
Currency	EUR
VAT number	DE 212137461
Product amount	EUR 1274.00
Admin charge	EUR 55.00
Total amount	EUR 1329.00