# ITS Analysis

7/5/23

## About this script

This script was crafted using Ellen's example codes. The script has three sections:

1. Data info
2. Exploration
3. Model fit

Install and/or load packages if required

## Standard ITS

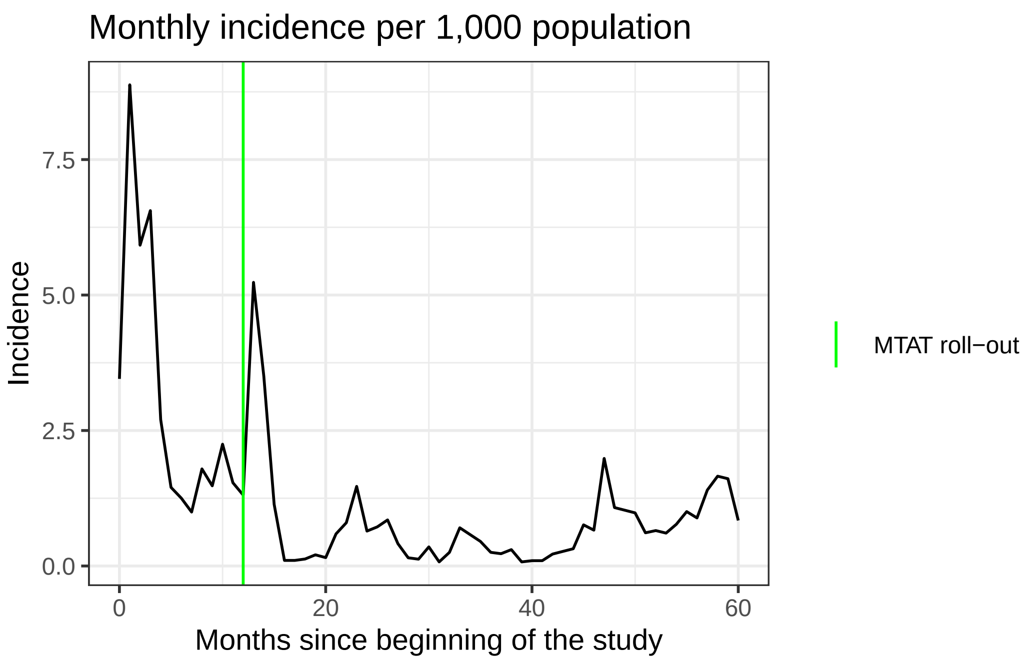
The data was aggregated monthly from 2013 to 2018 at the village level in Amhara Region, Ethiopia. The below code will read the data from the `xlsx` file. The intervention took a 2014 mass test-and-treat campaign for malaria control, followed by focal test-and-treat through 2017. Intervention impact is estimated using the INLA package.

**Case data**: There are 6 villages, all intervention groups. Villages are observed for 61 months and incidence cases were captured each month.

**Intervention (MTAT)**: the intervention rolled out on month 13 (Sept 2014) in all intervention villages. Intervention villages are filtered for standard ITS.

Inspect the data structure and variable lists

## Standard ITS - Exploratory




## Standard ITS - Modelling

Monthly incidence rate estimates assuming a negative binomial distribution, using the log population as an offset, with stratum as a fixed effect, random intercepts for each village (IID), and the month of observation (AR-1). To incorporate lagged environmental factors, a lagged 0 - 3 was created from the original **`evi, rainfall, lst_nigth, and lst_day`** variables, using the **`lag()`** function. We fitted four models: covariate-only model, lag0. lag1 and lag3. The table below showed that a model with lag 2 has a low dic value.

| model | dic |
|---|---:|
| covariate_only | 1714.670 |
| lag0 | 1489.779 |
| lag1 | 1493.649 |
| lag2 | 1488.981 |
| lag3 | 1492.203 |
| nocov | 1491.168 |

Standard ITS model - startum as random effect

| model | dic |
|---|---|
| lag0 | 1490.373 |
| lag1 | 1493.671 |
| lag2 | 1489.182 |
| lag3 | 1492.391 |
| nocov | 1491.078 |

Save the model results in data frame

| | Incidence rate ratio | 2.5th quantile, IRR | 97.5th quantile, IRR |
|---|---|---|---|
| Intercept | 0.000 | 0.000 | 0.003 |
| Months since study beginning | 0.851 | 0.686 | 1.055 |
| Months since MTAT | 1.146 | 0.910 | 1.442 |
| Post_MTAT_period | 0.628 | 0.112 | 3.522 |
| 2-month lagged EVI | 0.060 | 0.001 | 2.718 |
| 2-month lagged rainfall | 1.003 | 0.999 | 1.007 |
| 2-month lagged night temp | 0.950 | 0.825 | 1.094 |
| 2-month lagged day temp | 1.123 | 1.036 | 1.218 |
| Feb | 0.532 | 0.195 | 1.467 |
| Mar | 0.617 | 0.213 | 1.813 |
| Apr | 0.361 | 0.109 | 1.217 |
| May | 0.448 | 0.124 | 1.655 |
| Jun | 0.857 | 0.232 | 3.245 |
| Jul | 1.560 | 0.444 | 5.607 |
| Aug | 3.699 | 1.143 | 12.219 |
| Sep | 4.059 | 0.945 | 17.796 |
| Oct | 12.013 | 2.921 | 49.650 |
| Nov | 7.645 | 2.164 | 27.358 |
| Dec | 3.547 | 1.276 | 9.875 |

**Model 2 result interpretation**.

The intercept, which represents the estimated incidence rate ratio when all other variables are at their reference levels, does not show a significant association with the outcome. The uncertainty intervals for the intercept include the value 1, indicating a lack of evidence for a significant effect.

Regarding the months since the study began and months since MTAT, there is some evidence of a slight increase in the incidence rate with an increase in these variables. However, the uncertainty intervals include the value 1, suggesting that the associations may not be statistically significant.

During the post-MTAT period, there is evidence to suggest a lower estimated incidence rate compared to the pre-MTAT period. The uncertainty intervals for the post-MTAT period exclude the value 1, indicating a significant association.

When examining the variables Stratum2b and Stratum3, the estimated incidence rate ratios suggest different levels of association with the outcome. Stratum 2 b has a relatively smaller association, while Stratum 3 shows a significantly higher estimated incidence rate ratio. The uncertainty intervals for Stratum 3 indicate high variability, indicating the small sample size though it's statistically significant association.

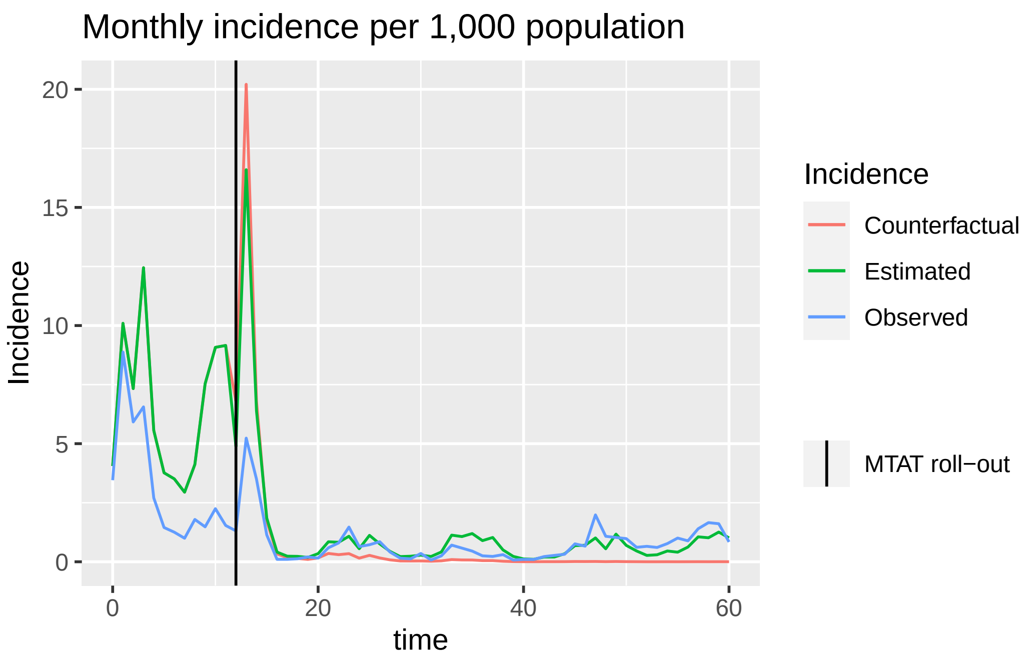
The environmental factors, such as 2-month lagged EVI, rainfall, night temperature, and day temperature, show varying levels of association with the outcome. The estimated incidence rate ratios suggest their impact on the outcome, but the uncertainty intervals indicate some uncertainty in the estimates.

Finally, for the different months or time periods, some months show lower estimated incidence rates (e.g., February, April, May, and June), while others exhibit higher estimated incidence rates (e.g., July, August, September, October, November, and December). The uncertainty intervals can provide further information on the significance of these associations.

| | Incidence rate ratio | 2.5th quantile, IRR | 97.5th quantile, IRR |
|---|---|---|---|
| Intercept | 0.000 | 0.000 | 0.001 |
| Months since study beginning | 0.857 | 0.690 | 1.062 |
| Months since MTAT | 1.139 | 0.904 | 1.433 |
| Post_MTAT_period | 0.609 | 0.110 | 3.435 |
| Stratum2b | 0.846 | 0.119 | 6.284 |
| Stratum3 | 14.498 | 2.014 | 118.748 |
| 2-month lagged EVI | 0.080 | 0.002 | 3.487 |
| 2-month lagged rainfall | 1.003 | 0.999 | 1.006 |
| 2-month lagged night temp | 0.943 | 0.820 | 1.083 |
| 2-month lagged day temp | 1.122 | 1.035 | 1.217 |
| Feb | 0.540 | 0.198 | 1.490 |
| Mar | 0.636 | 0.219 | 1.874 |
| Apr | 0.378 | 0.114 | 1.275 |
| May | 0.470 | 0.130 | 1.738 |
| Jun | 0.899 | 0.244 | 3.411 |
| Jul | 1.674 | 0.478 | 6.013 |
| Aug | 3.834 | 1.182 | 12.730 |
| Sep | 4.175 | 0.969 | 18.357 |
| Oct | 12.110 | 2.942 | 50.149 |
| Nov | 7.691 | 2.168 | 27.577 |
| Dec | 3.547 | 1.275 | 9.895 |

A counterfactual analysis is below to estimate how many additional cases would have been observed if not for the intervention. To accomplish this, we convert the predicted case counts, which are labelled `mean` in `models$summary.fitted.values`, back to the log scale so that we can remove the effects associated with the intervention. Specifically, the regression coefficient for the post-intervention period (level change) and time since intervention roll-out (trend change), multiplied by time since roll-out, are subtracted. Now, we have counterfactual case counts in log-scale, which we can exponentiate to get counterfactual cases. In our data, cases declined by 769, or a 35% decrease, during the MTAT/FTAT period.

Counter-factual plot




## Controlled ITS

The data was aggregated monthly from 2013 to 2018 at the village level in Amhara Region, Ethiopia. The code provided below will read the data from the specified `.xlsx file` (e.g., "filename.xlsx") and the corresponding sheet name (if applicable). The purpose of this analysis is to conduct a controlled interrupted time series analysis in R, focusing on a mass test-and-treat campaign in 2014 (MTAT), followed by focal test-and-treat interventions until 2017 for malaria control. The intervention impact will be estimated using the Integrated Nested Laplace Approximation (INLA) package.

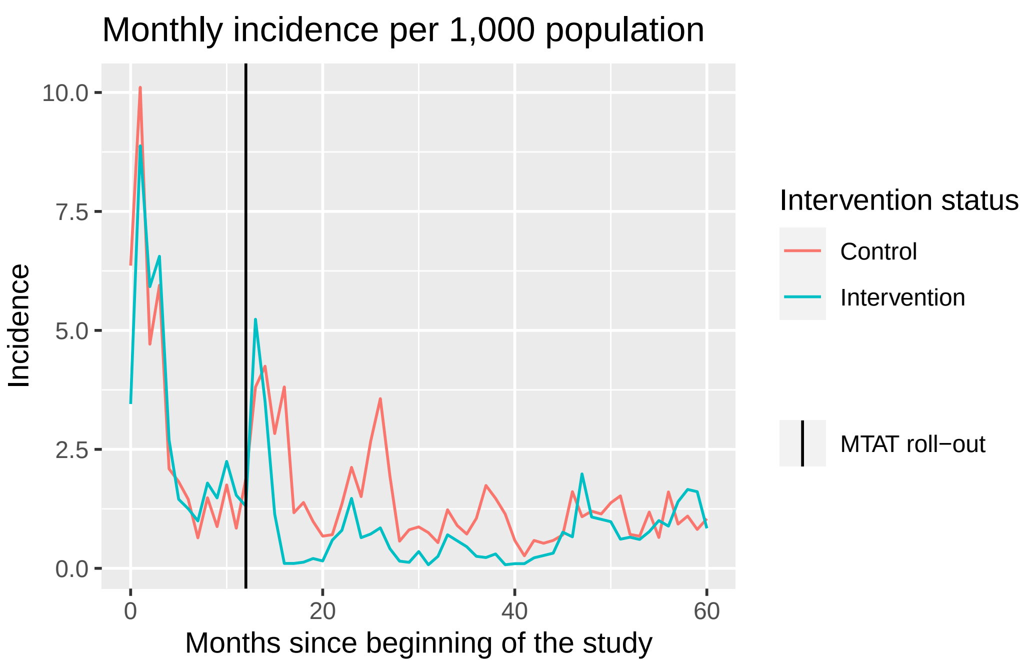
**Case data**: The study includes a total of 12 villages, with 6 villages assigned to the intervention group and 6 to the control group. These villages were observed for 61 months, and the incidence cases were captured monthly.

**Intervention (MTAT)**: The intervention was implemented in all intervention villages in September 2014 (month 13). The villages were divided into two groups: control and intervention.

**Time**: Time is defined in months. In the dataset, there are three-time variables: (1) time since the beginning of the study, a spline term representing months since the beginning of the intervention, (2) time_post spline representing the post-intervention period, and (3) a post-indicator to determine observations following the intervention roll-out. Our exposures of interest are the interactions of the group with (1) the post-intervention period and (2) the time since the intervention.

**Covariates**: The analysis includes environmental data extracted from the Malaria Atlas Project database, such as evi, daylight temp, night light temp, and rainfall. Additionally, intervention data related to vector control measures, including Long-Lasting Insecticidal Nets (LLINs) and Indoor Residual Spraying (IRS), are considered covariates.

Incidence plot: Intervention vs control




We estimate monthly incidence rates, assuming a negative binomial distribution, using log population as an offset with, `fixed effects of stratum`, random intercepts for each village (IID) and the month of observation (AR-1). Interaction terms are used to estimate the relative change in incidence rates in intervention areas compared to control areas 1) immediately following MTAT (level change) and 2) each additional month thereafter (trend change). Similar to the SITS, we fitted four different models and a model with lag 2 showed low dic.

Save the model results (stratum as fixed effect) in data frame.

| | Incidence rate ratio | 2.5th quantile, IRR | 97.5th quantile, IRR |
|---|---|---|---|
| Intercept | 0.000 | 0.000 | 0.001 |
| Months since study beginning | 0.807 | 0.705 | 0.921 |
| Intervention | 0.971 | 0.081 | 11.752 |
| Months since MTAT | 1.209 | 1.051 | 1.392 |
| Post period | 4.572 | 1.504 | 14.311 |
| Stratum2b | 1.293 | 0.114 | 14.688 |
| Stratum3 | 9.641 | 0.849 | 112.730 |
| 2-month lagged EVI | 0.167 | 0.018 | 1.534 |
| 2-month lagged rainfall | 1.002 | 1.000 | 1.004 |
| 2-month lagged night temp | 0.925 | 0.860 | 0.996 |
| 2-month lagged day temp | 1.112 | 1.062 | 1.163 |
| Feb | 0.685 | 0.401 | 1.176 |
| Mar | 0.643 | 0.358 | 1.160 |
| Apr | 0.447 | 0.231 | 0.868 |
| May | 0.514 | 0.252 | 1.049 |
| Jun | 0.704 | 0.336 | 1.480 |
| Jul | 1.289 | 0.638 | 2.620 |
| Aug | 2.542 | 1.313 | 4.948 |
| Sep | 3.480 | 1.505 | 8.117 |
| Oct | 9.488 | 4.238 | 21.328 |
| Nov | 6.328 | 3.083 | 13.079 |
| Dec | 3.762 | 2.134 | 6.653 |
| Intervention:time | 1.125 | 1.017 | 1.245 |
| Intervention:time_post | 0.893 | 0.807 | 0.988 |
| Intervention:post | 0.072 | 0.029 | 0.175 |
| Intervention:Stratum2b | 0.623 | 0.020 | 19.317 |
| Intervention:Stratum3 | 1.510 | 0.049 | 46.109 |

A random intercept for each village (IID), the month of observation (AR-1), and stratum (iid)

| model | dic |
|---|---|
| lag0 | 3342.247 |
| lag1 | 3342.214 |
| lag2 | 3324.511 |
| lag3 | 3339.188 |

**Model comparison: DIC, CPO and/or PIT**

The precision of stratum has been fixed with the prior information.

Deviance information criterion (DIC) between two model stratum as the fixed effect and stratum as a random effect, respectively. The smallest DIC (stratum as a random effect) indicates the best-fit model.

```
[1] 3324.761 3324.511
```

Conditional predictive ordinate - the smallest CPO indicates the best-fit model.

A histogram of PIT must resemble a uniform distribution; extreme values indicate outlying observations. Both model's outputs don't show a normal distributions - skewed to the right

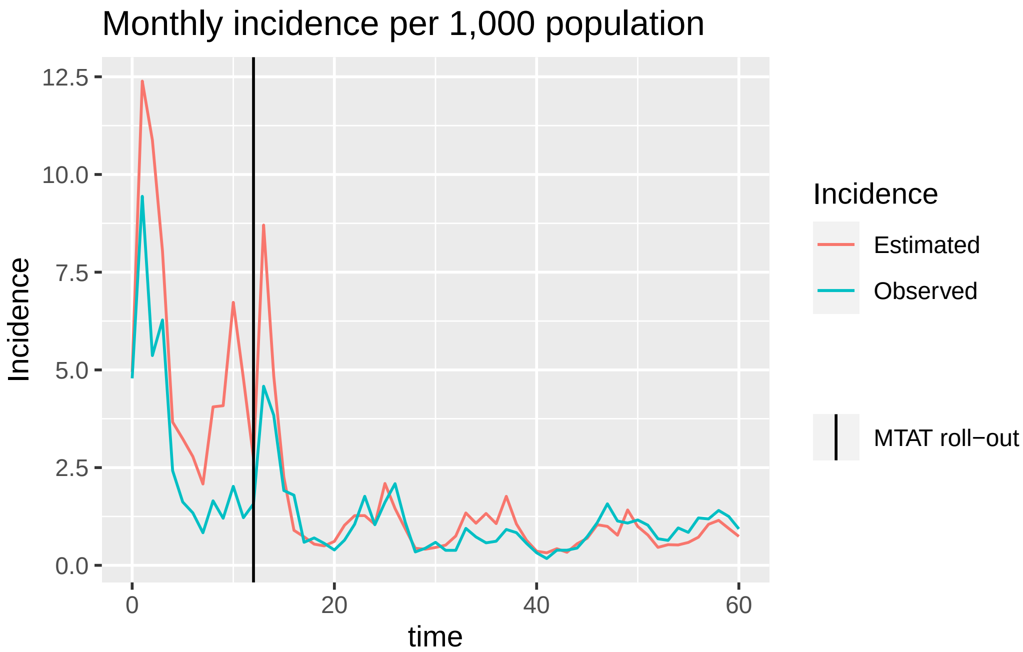
**Final Results - Modelc22**

Over time, for every one-unit increase in months, since the study began, there is an estimated decrease in the incidence rate by approximately 19.3%. Conversely, for every one-unit increase in months since the MTAT event, there is an estimated increase in the incidence rate by approximately 21.0%. During the post-period, there is a substantial estimated increase in the incidence rate by approximately 358.1%. The intervention shows a potential decrease in the incidence rate by approximately 4.3%, although there is some uncertainty in the effects. These findings provide valuable insights for understanding the dynamics of the incidence rate and can inform targeted interventions to mitigate the risk of the outcome. The "2-month lagged EVI" suggests a potential decrease in the incidence rate by approximately 84.7% to 1.2%. Conversely, the "2-month lagged rainfall" and "2-month lagged day temp" indicate a slight increase in the incidence rate by approximately 0.2% to 11.2% and 6.3% to 16.3%, respectively. However, the wide uncertainty intervals highlight the need for further investigation and consideration of other factors to better understand the associations. The analysis uncovered significant interactions between variables, revealing the dynamic nature of interventions over time. The "time: Intervention" interaction indicated that the intervention's effect varied with time, resulting in a 12.3% increase in the incidence rate for each unit increase in time. The "Intervention:time_post" interaction showed an 11.0% decrease in the incidence rate for each unit increase in time during the post period. Moreover, the "Intervention:post" interaction demonstrated a substantial reduction of approximately 92.7% in the incidence rate during the post period when the intervention was implemented. These findings underscore the importance of considering the timing of interventions for understanding their impact on the outcome.

Among the months examined, April, May, and June showed estimated incidence rate ratios below 1, indicating a lower incidence rate compared to the reference month. Specifically, April had an estimated incidence rate ratio of approximately 44.1%, May had 50.2%, and June had 68.9% compared to the reference month. However, the uncertainty intervals for these months suggest some variability and the possibility of non-significant associations.

For the remaining months (February, March, July, August, September, October, November, and December), the estimated incidence rate ratios ranged from approximately 36.0% to 267.2% compared to the reference month. However, since the uncertainty intervals include the value 1 for most of these months, it suggests that the associations may not be statistically significant.

| | Incidence rate ratio | 2.5th quantile, IRR | 97.5th quantile, IRR |
|---|---|---|---|
| Intercept | 0.000 | 0.000 | 0.006 |
| Months since study beginning | 0.807 | 0.706 | 0.920 |
| Intervention | 0.958 | 0.248 | 3.717 |
| Months since MTAT | 1.210 | 1.053 | 1.391 |
| Post period | 4.527 | 1.501 | 13.999 |
| 2-month lagged EVI | 0.152 | 0.017 | 1.383 |
| 2-month lagged rainfall | 1.002 | 1.000 | 1.004 |
| 2-month lagged night temp | 0.931 | 0.864 | 1.001 |
| 2-month lagged day temp | 1.112 | 1.063 | 1.163 |
| Feb | 0.683 | 0.400 | 1.170 |
| Mar | 0.640 | 0.357 | 1.150 |
| Apr | 0.440 | 0.228 | 0.852 |
| May | 0.501 | 0.246 | 1.021 |
| Jun | 0.687 | 0.329 | 1.441 |
| Jul | 1.264 | 0.626 | 2.557 |
| Aug | 2.527 | 1.309 | 4.904 |
| Sep | 3.525 | 1.528 | 8.199 |
| Oct | 9.641 | 4.315 | 21.607 |
| Nov | 6.404 | 3.127 | 13.197 |
| Dec | 3.773 | 2.143 | 6.659 |
| time:Intervention | 1.123 | 1.015 | 1.242 |
| Intervention:time_post | 0.894 | 0.807 | 0.990 |
| Intervention:post | 0.073 | 0.030 | 0.178 |


Monthly incidence per 1,000 population
Incidence
12.5
10.0
7.5
5.0
2.5
0.0
0
20
40
60
time
Incidence
Estimated
Observed
MTAT roll−out