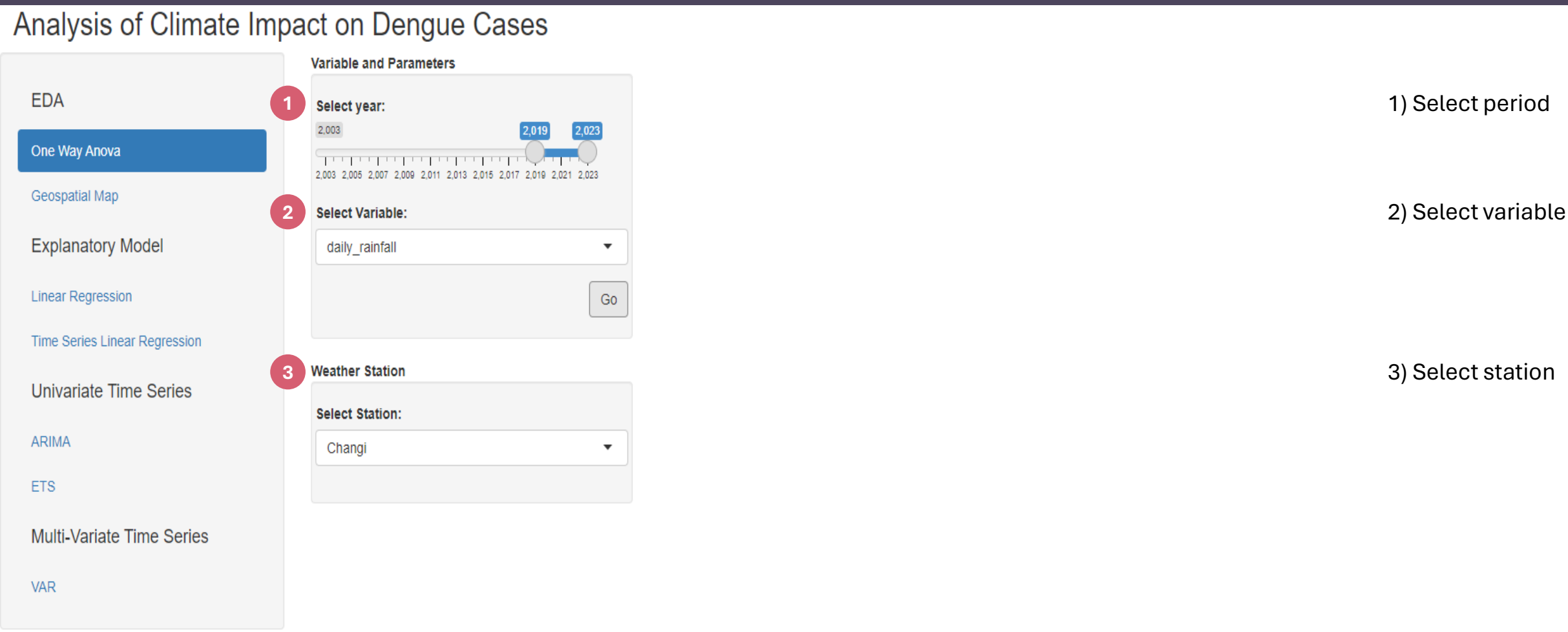
Analysis of Climate Impact on Dengue Cases

ShinyApp User Manual

Group 3: CI Hui, Colin JIANG Kelin, SUN Yiping

EDA – One Way Anova

This module allows users to analyze the difference in distribution of climate variables across time for a selected station in Singapore.



1) Select period

2) Select variable

3) Select station

EDA – One Way Anova

This module allows users to analyze the difference in distribution of climate variables across time for a selected station in Singapore.

Analysis of Climate Impact on Dengue Cases

- EDA
- One Way Anova
- Geospatial Map
- Explanatory Model
- Linear Regression
- Time Series Linear Regression
- Univariate Time Series
- ARIMA
- ETS
- Multi-Variate Time Series
- VAR

Variable and Parameters

Select year:

2,003 — 2,019 — 2,023

2,003 2,005 2,007 2,009 2,011 2,013 2,015 2,017 2,019 2,021 2,023

Select Variable:

daily_rainfall

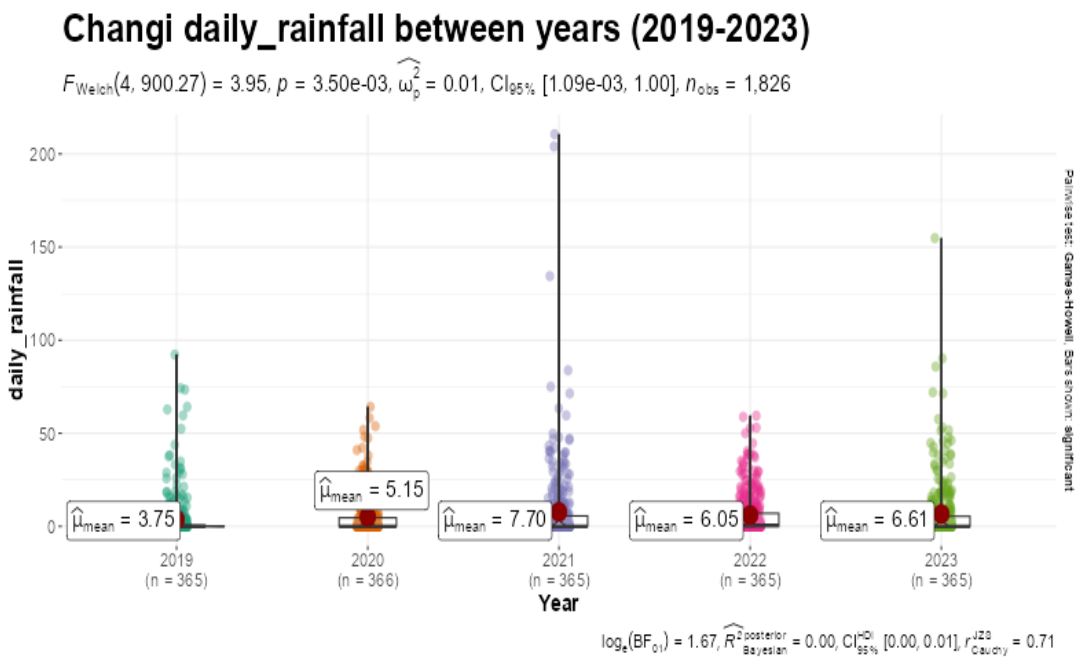
4 Go

Weather Station

Select Station:

Changi

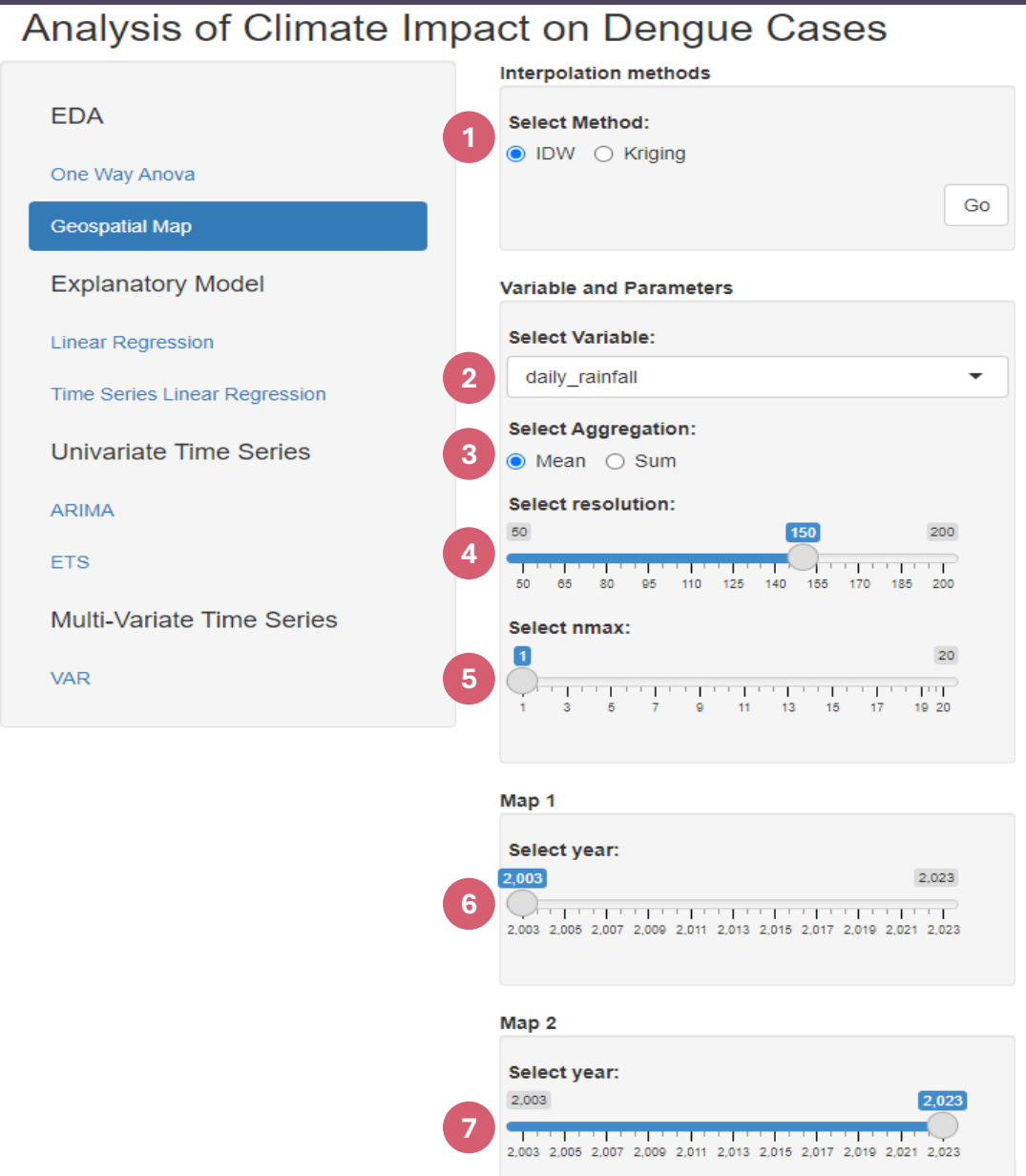
Anova



4) Click “Go” to plot

EDA – Geospatial - IDW

This module allows users to analyze the difference in geospatial distribution of climate variables between two years across Singapore.



Aggregation: calculating statistics, either average or sum of the selected variable for the selected year. Sum method is only suggested for rainfall. Mean method should be used for temperature and wind.

nmax: the maximum number of neighbouring data points considered for spatial interpolation or variogram calculation

1) Select method of interpolation. This will change the interface and parameters accordingly

2) Select variable

3) Select aggregation

4) Select resolution

5) Select nmax

6) Select year for map 1

7) Select year for map 2

EDA – Geospatial - IDW

This module allows users to analyze the difference in geospatial distribution of climate variables between two years across Singapore.

Analysis of Climate Impact on Dengue Cases

- EDA
- One Way Anova
- Geospatial Map
- Explanatory Model
- Linear Regression
- Time Series Linear Regression
- Univariate Time Series
- ARIMA
- ETS
- Multi-Variate Time Series
- VAR

Interpolation methods

Select Method:

(•) IDW ( ) Kriging

7 Go

Variable and Parameters

Select Variable:

daily_rainfall

Select Aggregation:

(•) Mean ( ) Sum

Select resolution:

50 150 200

50 65 80 95 110 125 140 155 170 185 200

Select nmax:

1 20

1 3 5 7 9 11 13 15 17 19 20

Map 1

Select year:

2,003 2,023

2,003 2,005 2,007 2,009 2,011 2,013 2,015 2,017 2,019 2,021 2,023

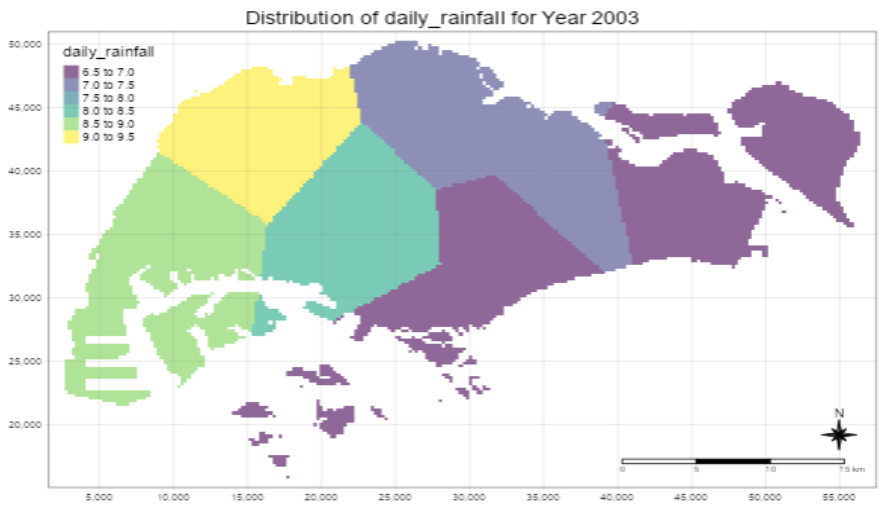
Map 2

Select year:

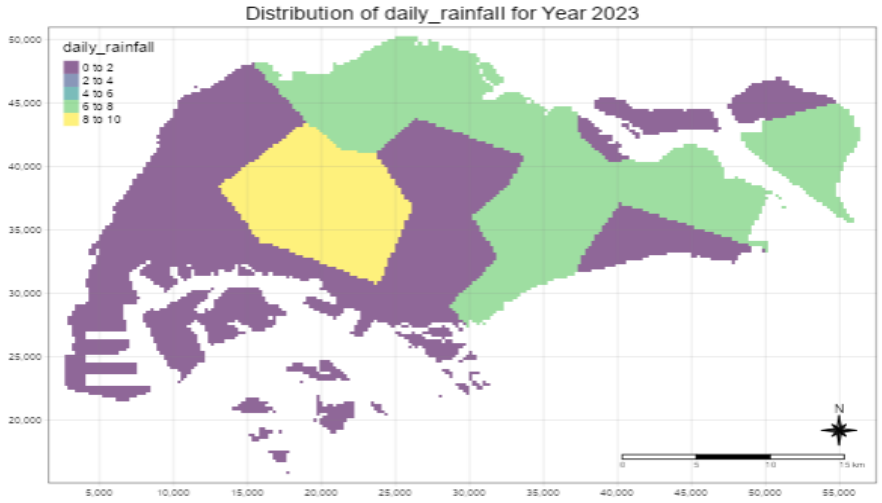
2,003 2,023

2,003 2,005 2,007 2,009 2,011 2,013 2,015 2,017 2,019 2,021 2,023

Map 1



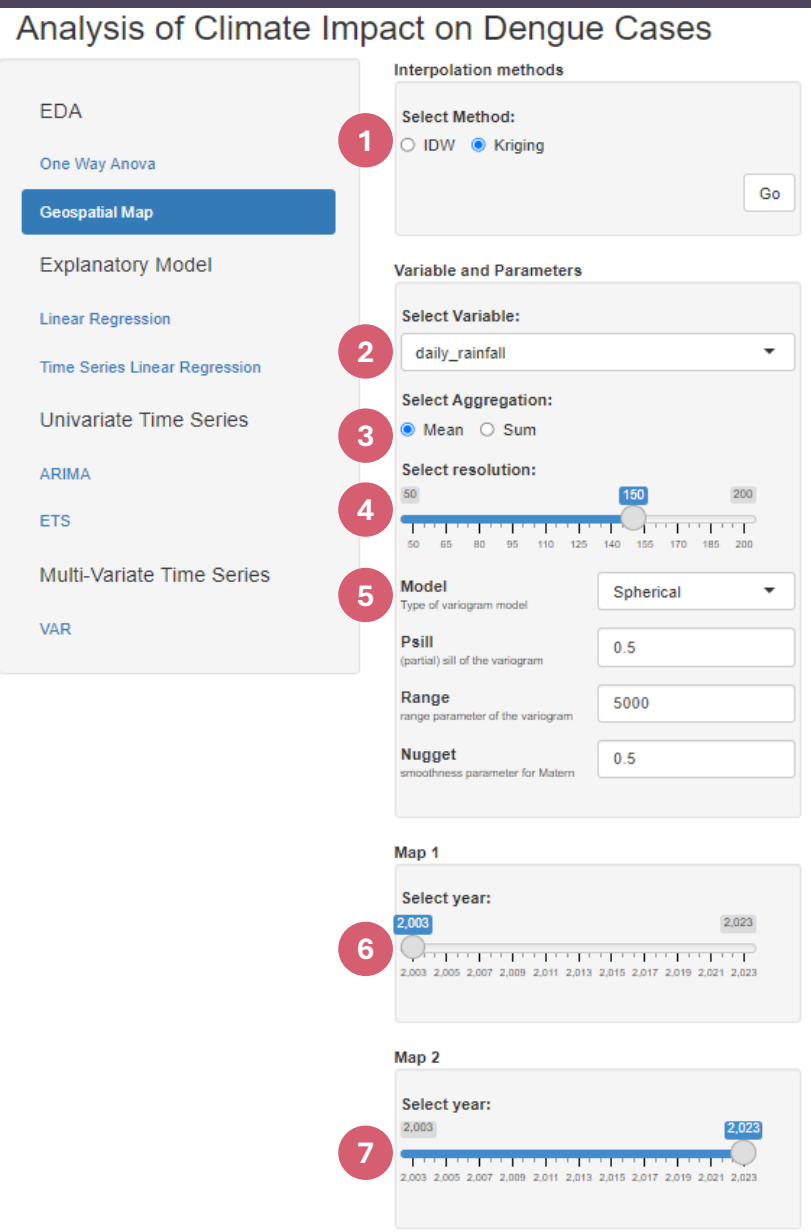
Map 2



7) Click “Go” to plot the 2 maps

EDA – Geospatial - Kriging

This module allows users to analyze the difference in geospatial distribution of climate variables between two years across Singapore.



Model: the theoretical variogram model that you want to fit to your empirical variogram

Psill: the variance contributed by spatial dependence or correlation up to the range

Range: the distance at which spatial correlation between data points reaches a plateau or levels off

Nugget: the variance at distances smaller than the smallest separation distance between data points

1) Select method of interpolation. This will change the interface and parameters accordingly

2) Select variable

3) Select aggregation

4) Select resolution

5) Fill in other parameters

6) Select year for map 1

7) Select year for map 2

EDA – Geospatial - Kriging

This module allows users to analyze the difference in geospatial distribution of climate variables between two years across Singapore.

Analysis of Climate Impact on Dengue Cases

- EDA
- One Way Anova
- **Geospatial Map**
- Explanatory Model
- Linear Regression
- Time Series Linear Regression
- Univariate Time Series
- ARIMA
- ETS
- Multi-Variate Time Series
- VAR

Interpolation methods

Select Method: ○ IDW ◉ Kriging

8 Go

Variable and Parameters

Select Variable: daily_rainfall

Select Aggregation: ◉ Mean ○ Sum

Select resolution: 50 | 150 | 200

Model (Type of variogram model): Spherical

Psill ((partial) sill of the variogram): 0.5

Range (range parameter of the variogram): 5000

Nugget (smoothness parameter for Matern): 0.5

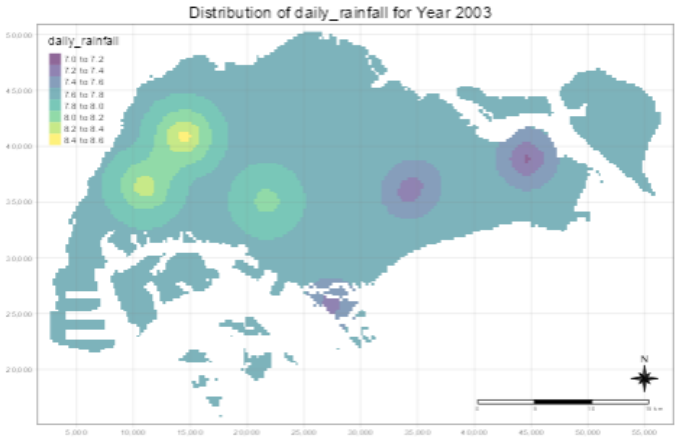
Map 1

Select year: 2,003 | 2,023

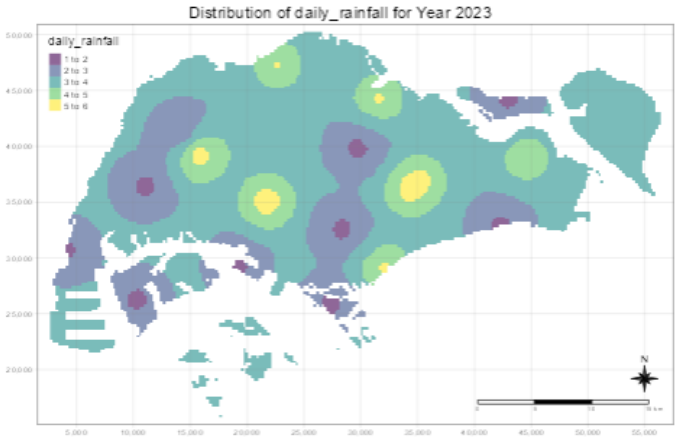
Map 2

Select year: 2,003 | 2,023

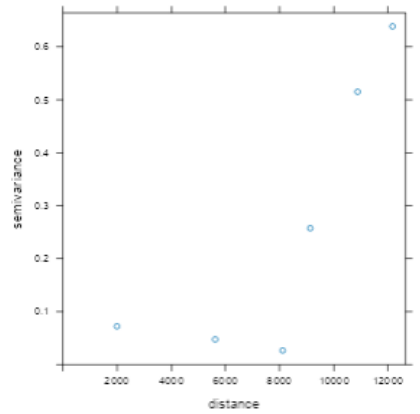
Map 1



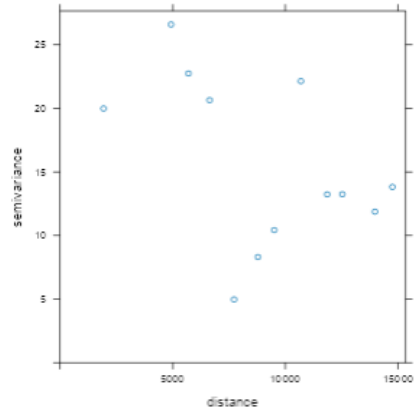
Map 2



Variogram 1



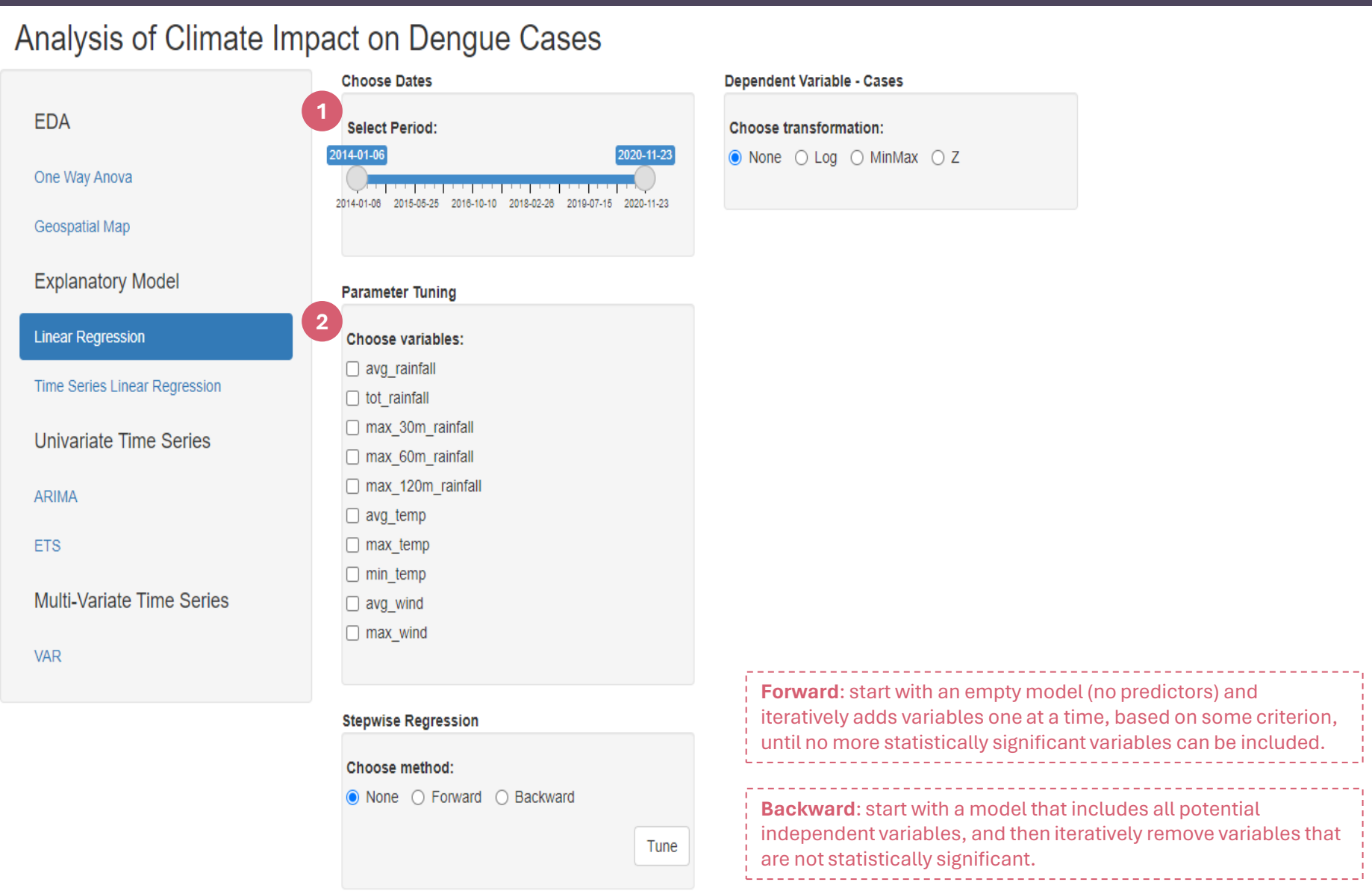
Variogram 2



8) Click “Go” to plot the 2 maps and variograms

Explanatory Model – Linear Regression

This module allows users to fit a linear regression model on Dengue Cases and analyze the variable importance via coefficients



1) Select period of analysis

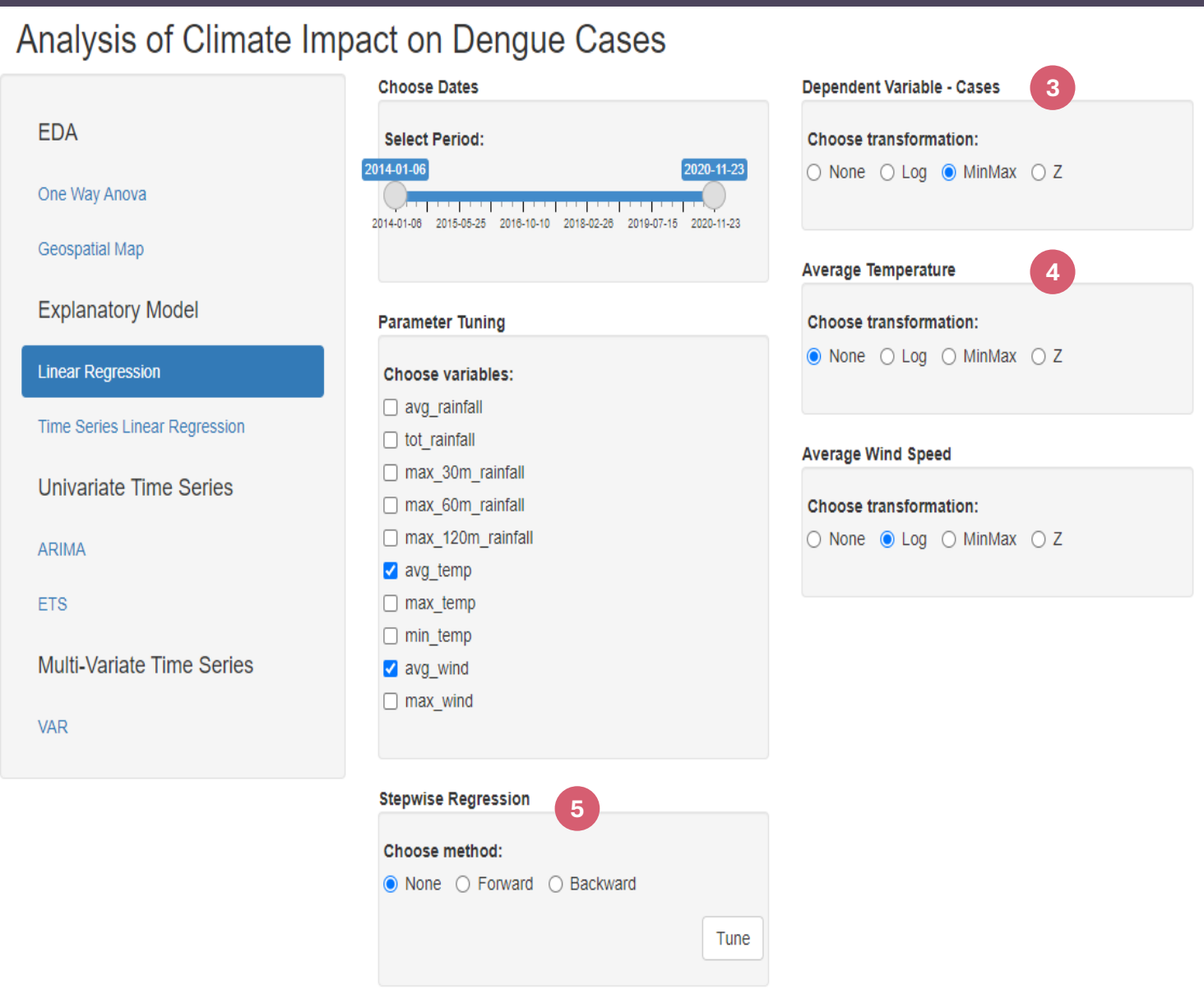
2) Choose variables to include in model

Forward: start with an empty model (no predictors) and iteratively adds variables one at a time, based on some criterion, until no more statistically significant variables can be included.

Backward: start with a model that includes all potential independent variables, and then iteratively remove variables that are not statistically significant.

Explanatory Model – Linear Regression

This module allows users to fit a linear regression model on Dengue Cases and analyze the variable importance via coefficients



3) Once a variable is added into the model, an additional panel will appear in the middle column

4) Select log, min-max transformation or standardization for each variable, if needed

5) Select forward or backward stepwise method for regression, if needed

Explanatory Model – Linear Regression

This module allows users to fit a linear regression model on Dengue Cases and analyze the variable importance via coefficients

Analysis of Climate Impact on Dengue Cases

- EDA
- One Way Anova
- Geospatial Map
- Explanatory Model
- Linear Regression
- Time Series Linear Regression
- Univariate Time Series
- ARIMA
- ETS
- Multi-Variate Time Series
- VAR

Choose Dates

Select Period: 2014-01-06 – 2020-11-23

2014-01-06 2015-05-25 2016-10-10 2018-02-26 2019-07-15 2020-11-23

Parameter Tuning

Choose variables:

- [ ] avg_rainfall
- [ ] tot_rainfall
- [ ] max_30m_rainfall
- [ ] max_60m_rainfall
- [ ] max_120m_rainfall
- [x] avg_temp
- [ ] max_temp
- [x] min_temp
- [x] avg_wind
- [ ] max_wind

Stepwise Regression

Choose method:

( ) None ( ) Forward (•) Backward

6 Tune

Dependent Variable - Cases

Choose transformation:

(•) None ( ) Log ( ) MinMax ( ) Z

Average Temperature

Choose transformation:

(•) None ( ) Log ( ) MinMax ( ) Z

Min Temperature

Choose transformation:

(•) None ( ) Log ( ) MinMax ( ) Z

Average Wind Speed

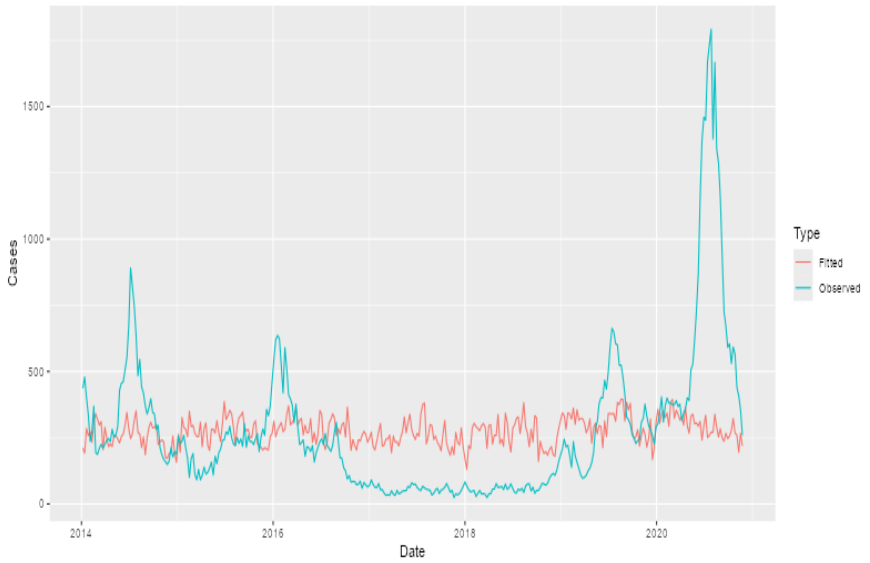
Choose transformation:

(•) None ( ) Log ( ) MinMax ( ) Z

7

Observed vs Fitted | Coefficients | Diagnostics | Raw Variables

Observed vs Fitted



Model Metrics

Adjusted.R.2	F.Statistics	MAPE
0.03	5.84	1.52

Variables in Model (after stepwise)

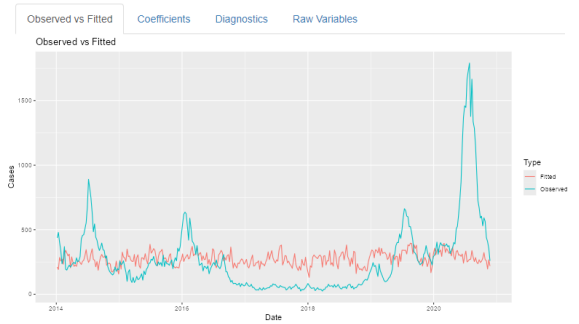
term	estimate	std.error	statistic	p.value
(Intercept)	-1122.84	513.04	-2.19	0.03
avg_temp	43.38	18.01	2.41	0.02
avg_wind	21.12	8.51	2.48	0.01

7) Toggle through tab panels to view various model diagnostics

6) Click “Tune” to generate results

Explanatory Model – Linear Regression

This module allows users to fit a linear regression model on Dengue Cases and analyze the variable importance via coefficients

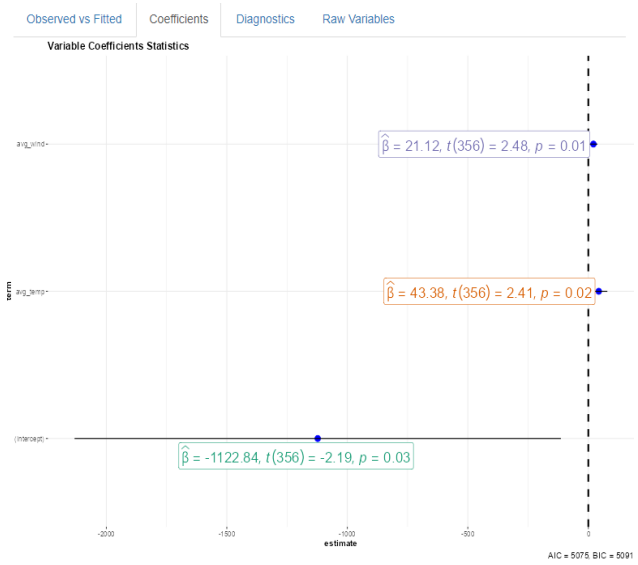


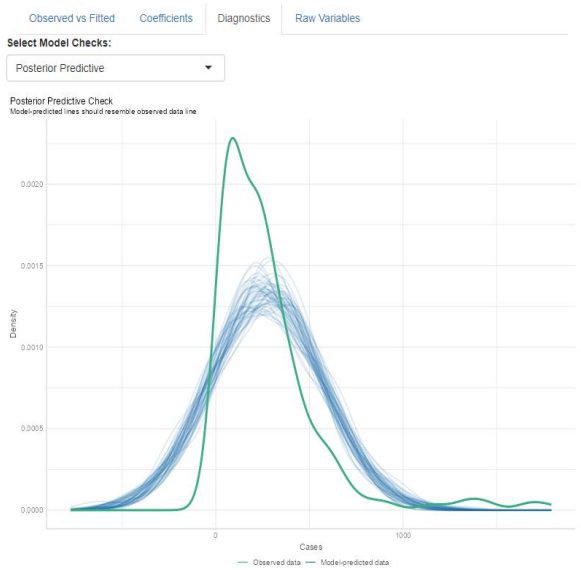
Model Metrics

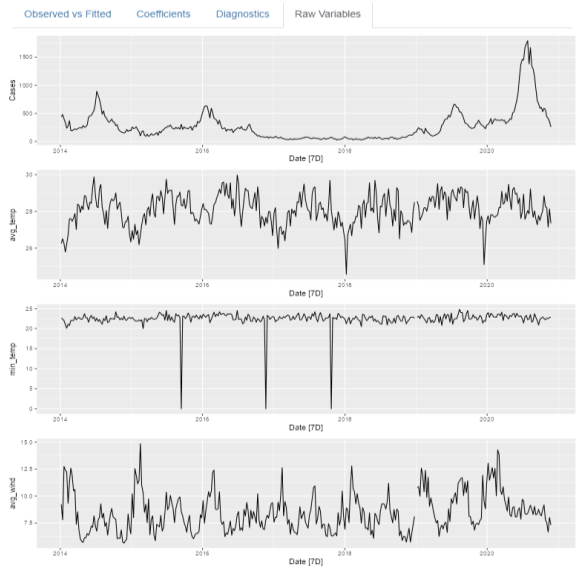
Adjusted.R.2	F.Statistics	MAPE
0.03	5.84	1.52

Variables in Model (after stepwise)

term	estimate	std.error	statistic	p.value
(Intercept)	-1122.84	513.04	-2.19	0.03
avg_temp	43.38	18.01	2.41	0.02
avg_wind	21.12	8.51	2.48	0.01







Actual vs Fit
Check the fit of the LM model against the actual values

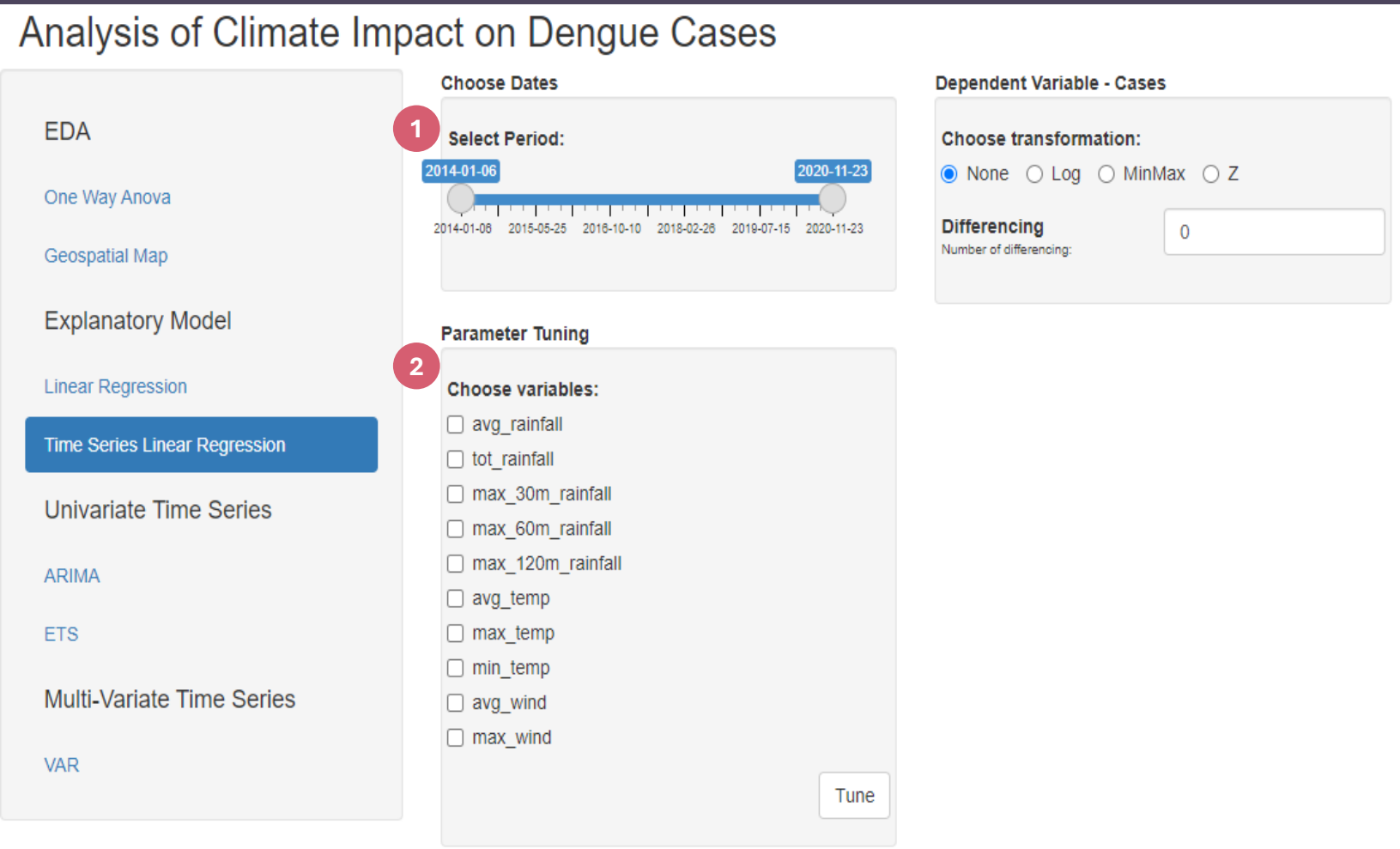
Coefficients
View coefficients and error of variables in model

Diagnostics
Check the various diagnostics of the model such as normality of residuals and multi-collinearity

Raw Variables
Check the plots of the individual variables in the model to determine if transformations required

Explanatory Model – Time Series Linear Model

This module allows users to fit a Time Series Linear model on Dengue Cases and analyze the variable importance via coefficients

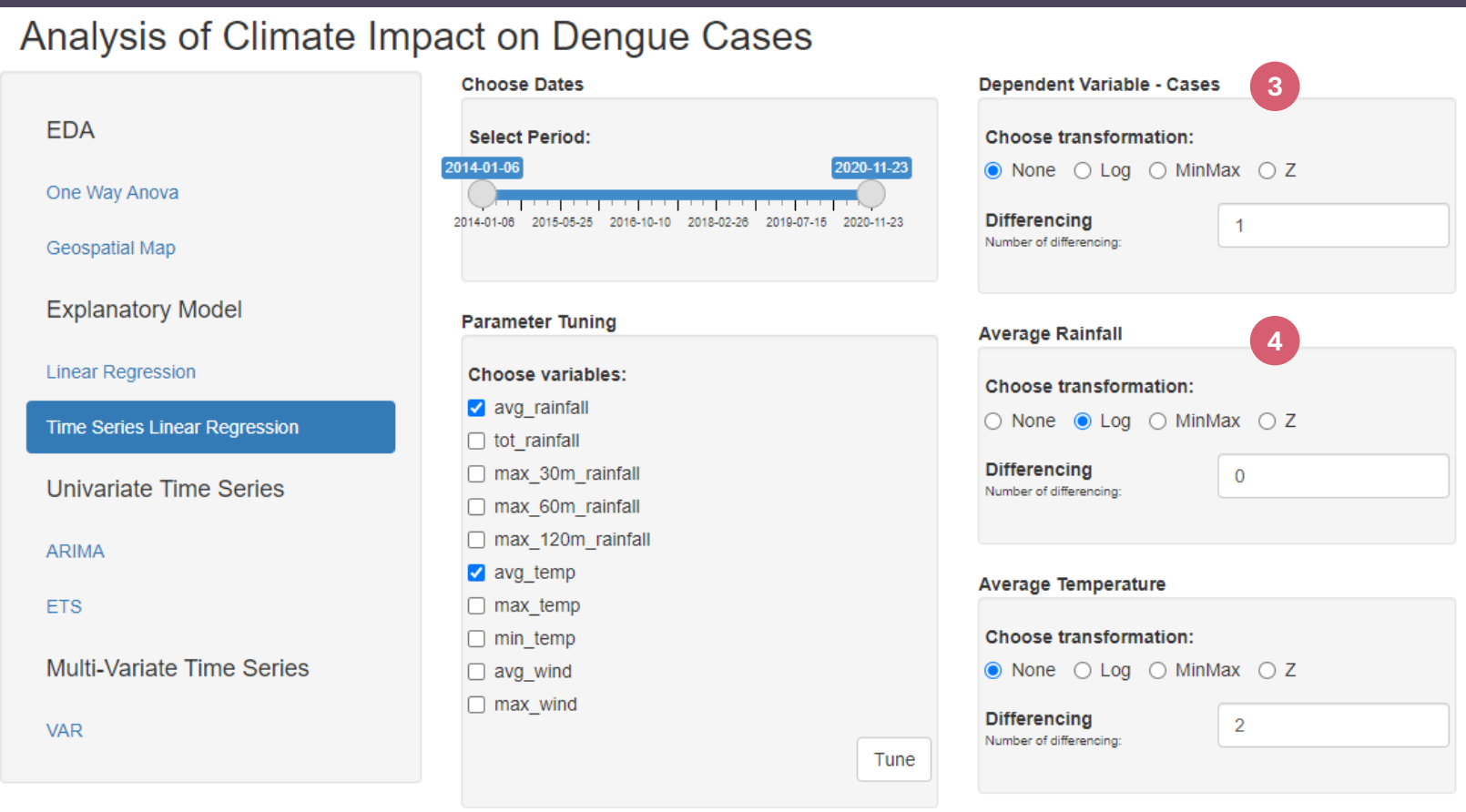


1) Select period of analysis

2) Choose variables to include in model

Explanatory Model – Time Series Linear Model

This module allows users to fit a Time Series Linear model on Dengue Cases and analyze the variable importance via coefficients



3) Once a variable is added into the model, an additional panel will appear in the middle column

4) Select log, min-max transformation or standardization for each variable, if needed.

Users can also determine if differencing is required

Explanatory Model – Time Series Linear Model

This module allows users to fit a Time Series Linear model on Dengue Cases and analyze the variable importance via coefficients

Analysis of Climate Impact on Dengue Cases

- EDA
- One Way Anova
- Geospatial Map
- Explanatory Model
- Linear Regression
- Time Series Linear Regression
- Univariate Time Series
- ARIMA
- ETS
- Multi-Variate Time Series
- VAR

Choose Dates

Select Period: 2014-01-06 – 2020-11-23

2014-01-06 2015-05-25 2016-10-10 2018-02-26 2019-07-15 2020-11-23

Parameter Tuning

Choose variables:

- [x] avg_rainfall
- [ ] tot_rainfall
- [ ] max_30m_rainfall
- [ ] max_60m_rainfall
- [ ] max_120m_rainfall
- [x] avg_temp
- [ ] max_temp
- [ ] min_temp
- [x] avg_wind
- [ ] max_wind

5 Tune

Dependent Variable - Cases

Choose transformation: ○ None ◉ Log ○ MinMax ○ Z

Differencing — Number of differencing: 0

Average Rainfall

Choose transformation: ○ None ◉ Log ○ MinMax ○ Z

Differencing — Number of differencing: 1

Average Temperature

Choose transformation: ◉ None ○ Log ○ MinMax ○ Z

Differencing — Number of differencing: 0

Average Wind Speed

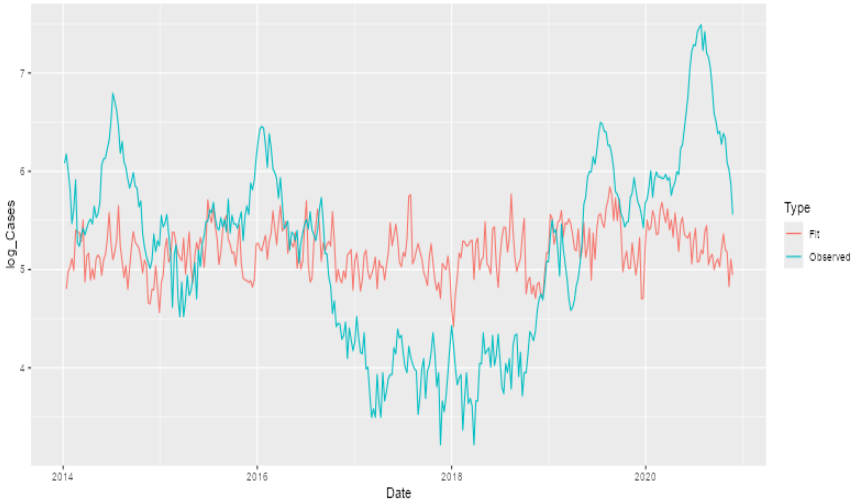
Choose transformation: ◉ None ○ Log ○ MinMax ○ Z

Differencing — Number of differencing: 0

6

Actual vs Fit | Residuals | Coefficients

Observed vs Fitted



Model Metrics

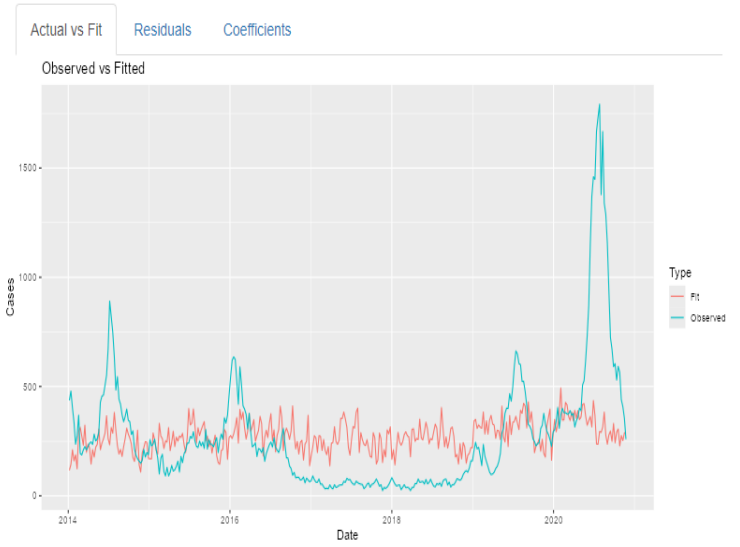
RMSE	MAE	MAPE
0.91	0.74	15.29

6) Toggle through tab panels to view various model diagnostics

5) Click “Tune” to generate results

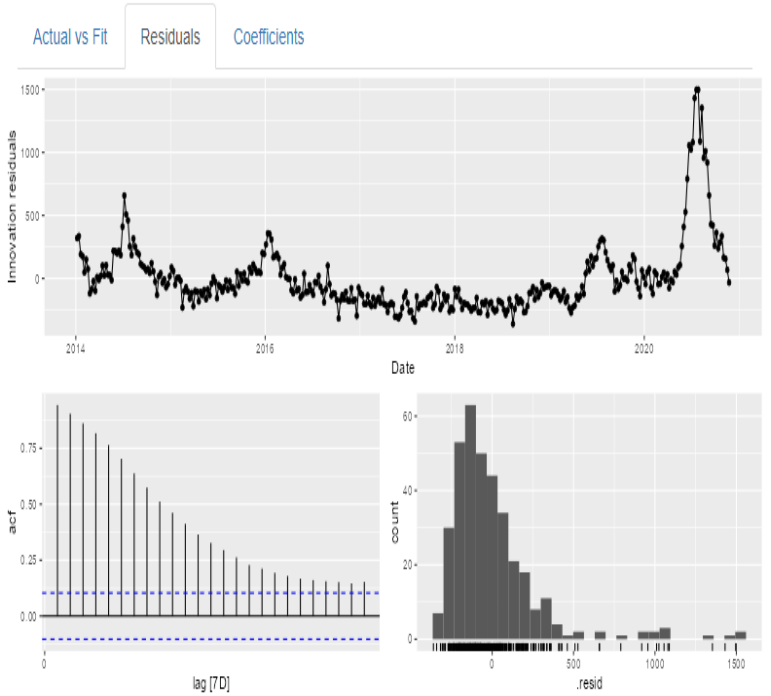
Explanatory Model – Time Series Linear Model

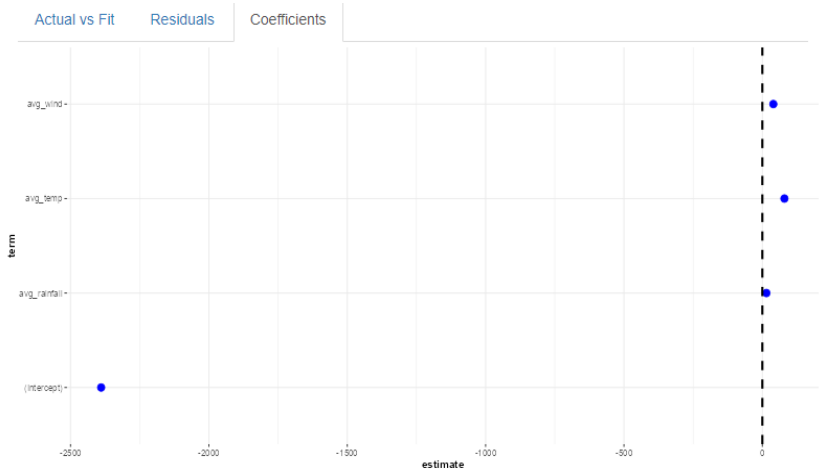
This module allows users to fit a Time Series Linear model on Dengue Cases and analyze the variable importance via coefficients



Model Metrics

RMSE	MAE	MAPE
276.89	180.25	149.44





Model Estimates

term	estimate	std.error	statistic	p.value
(Intercept)	-2389.41	645.82	-3.70	0.00
avg_rainfall	14.71	4.64	3.17	0.00
avg_temp	79.80	21.19	3.77	0.00
avg_wind	39.65	10.23	3.88	0.00

Actual vs Fit

Check the fit of the TSLM model against the actual values

Residuals

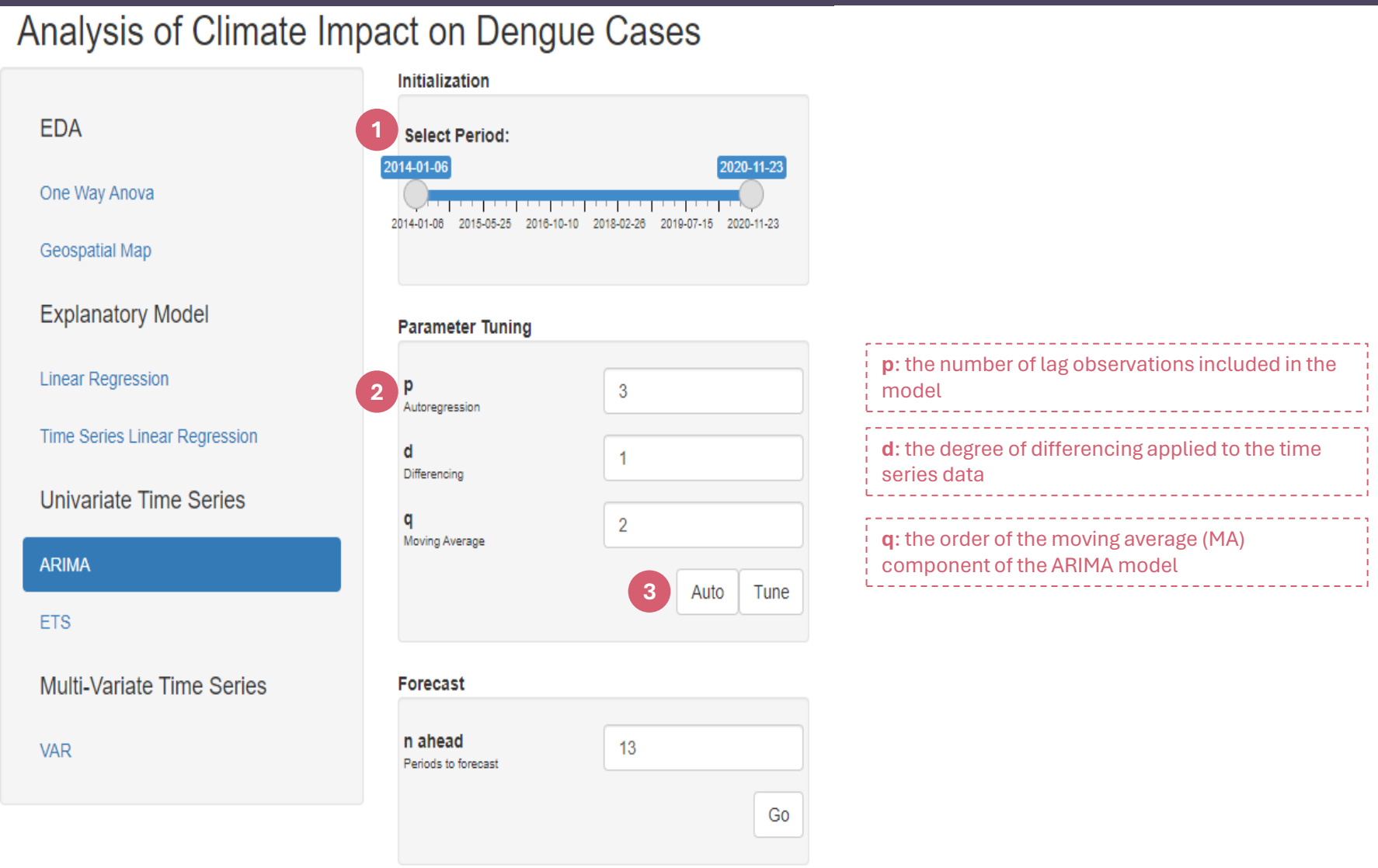
Check the normality of residuals to determine stationarity

Coefficients

Gain insights on variable estimates and statistics

Univariate Time Series – ARIMA

This module allows users to fit a ARIMA model on Dengue Cases and generate forecast



p: the number of lag observations included in the model

d: the degree of differencing applied to the time series data

q: the order of the moving average (MA) component of the ARIMA model

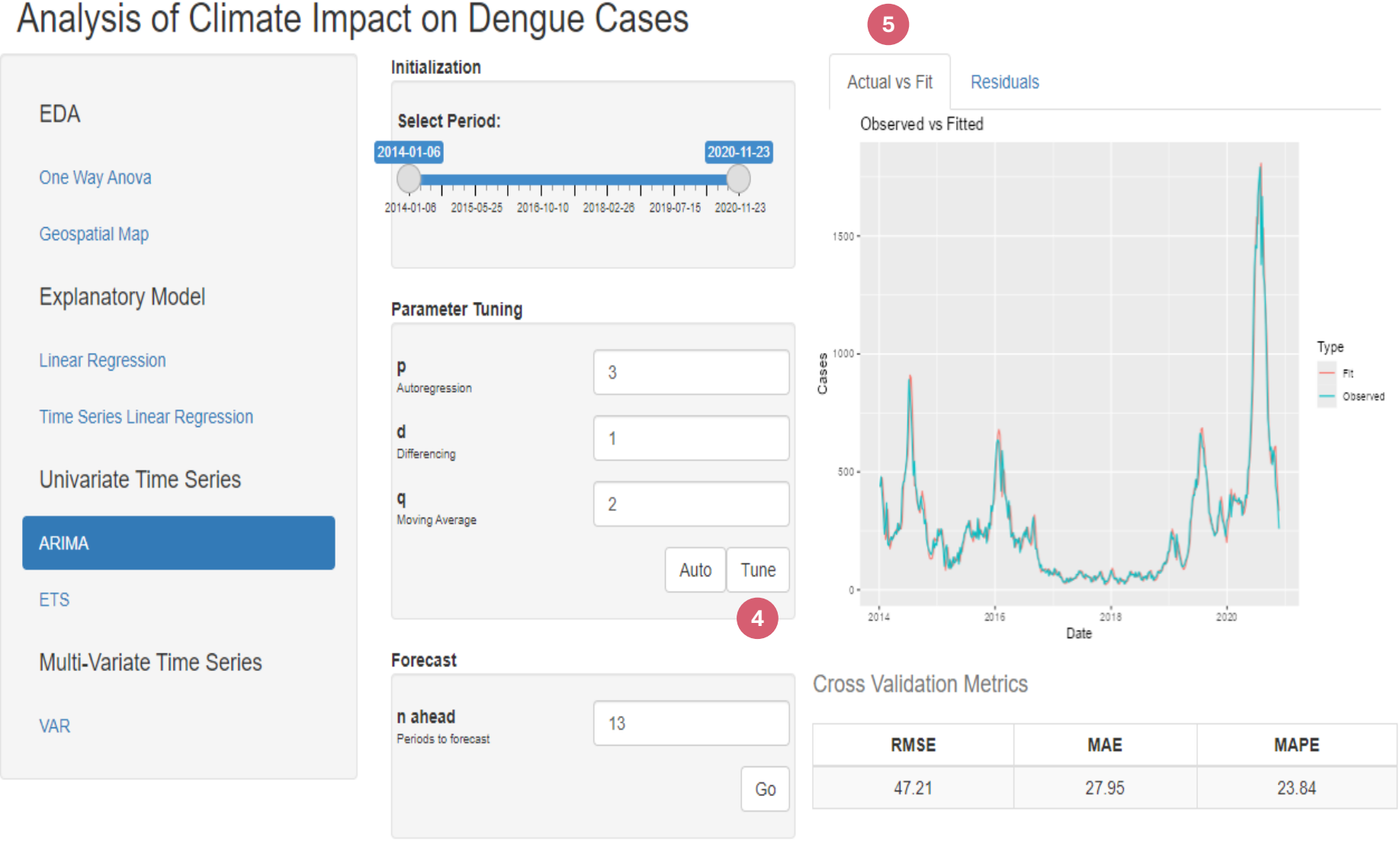
1) Select period of analysis

2) Select Input parameters manually for the model

3) Alternative to step 2, users can click “Auto” to fill parameter based on model’s AICC automatically

Univariate Time Series – ARIMA

This module allows users to fit a ARIMA model on Dengue Cases and generate forecast

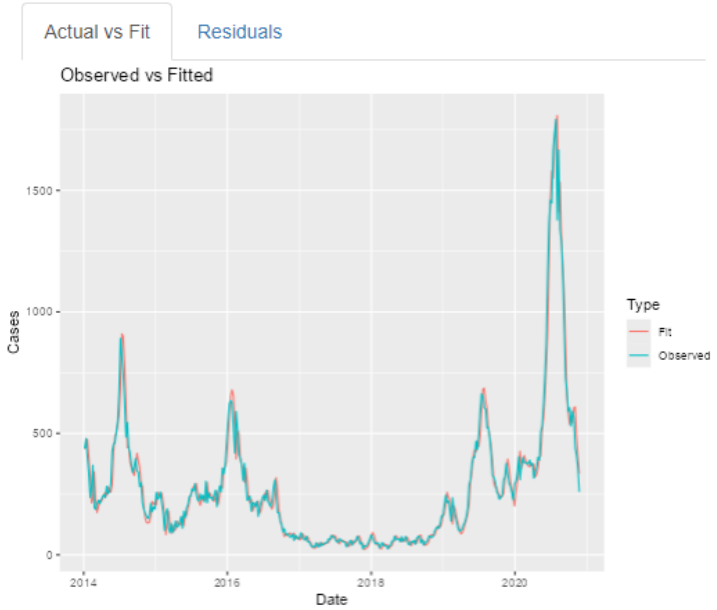


5) Toggle between tab panels for various diagnostics

4) Click “Tune” to generate model diagnostics

Univariate Time Series – ARIMA

This module allows users to fit a ARIMA model on Dengue Cases and generate forecast



Cross Validation Metrics

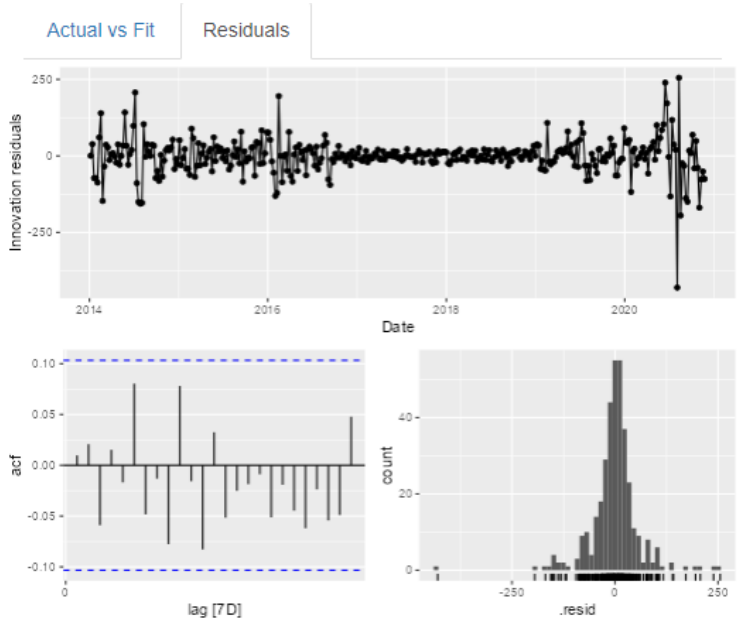
RMSE	MAE	MAPE
47.21	27.95	23.84

Actual vs Fit

Check the fit of the ARIMA model against the actual values

Cross Validation Metrics

Check and compare the performance of the model using Cross Validation metrics



Cross Validation Metrics

RMSE	MAE	MAPE
47.21	27.95	23.84

Residuals

Check the normality of residuals to determine stationarity

Model Diagnostics

Users can use these diagnostics to determine if parameters are suitable, and re-tune models accordingly

Univariate Time Series – ARIMA

This module allows users to fit a ARIMA model on Dengue Cases and generate forecast

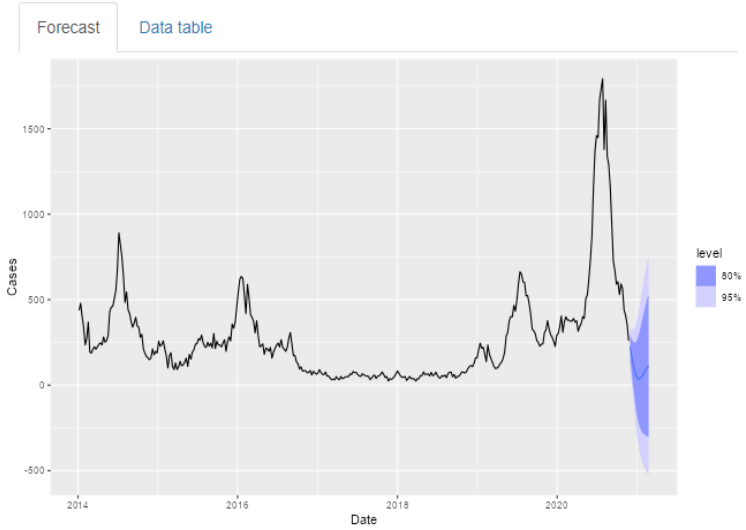
8) Toggle between tab panels for various forecast results

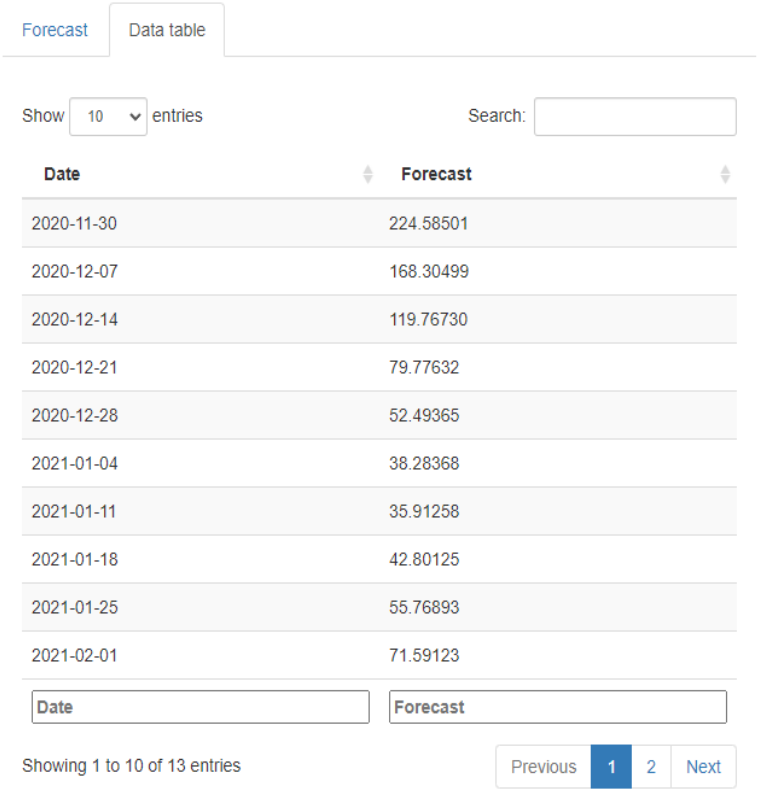
6) Key in number of forecast period

7) Click“Go” to generate forecast

Univariate Time Series – ARIMA

This module allows users to fit a ARIMA model on Dengue Cases and generate forecast





Model Forecast

Users can use these forecast results to compare with that from other models

Forecast

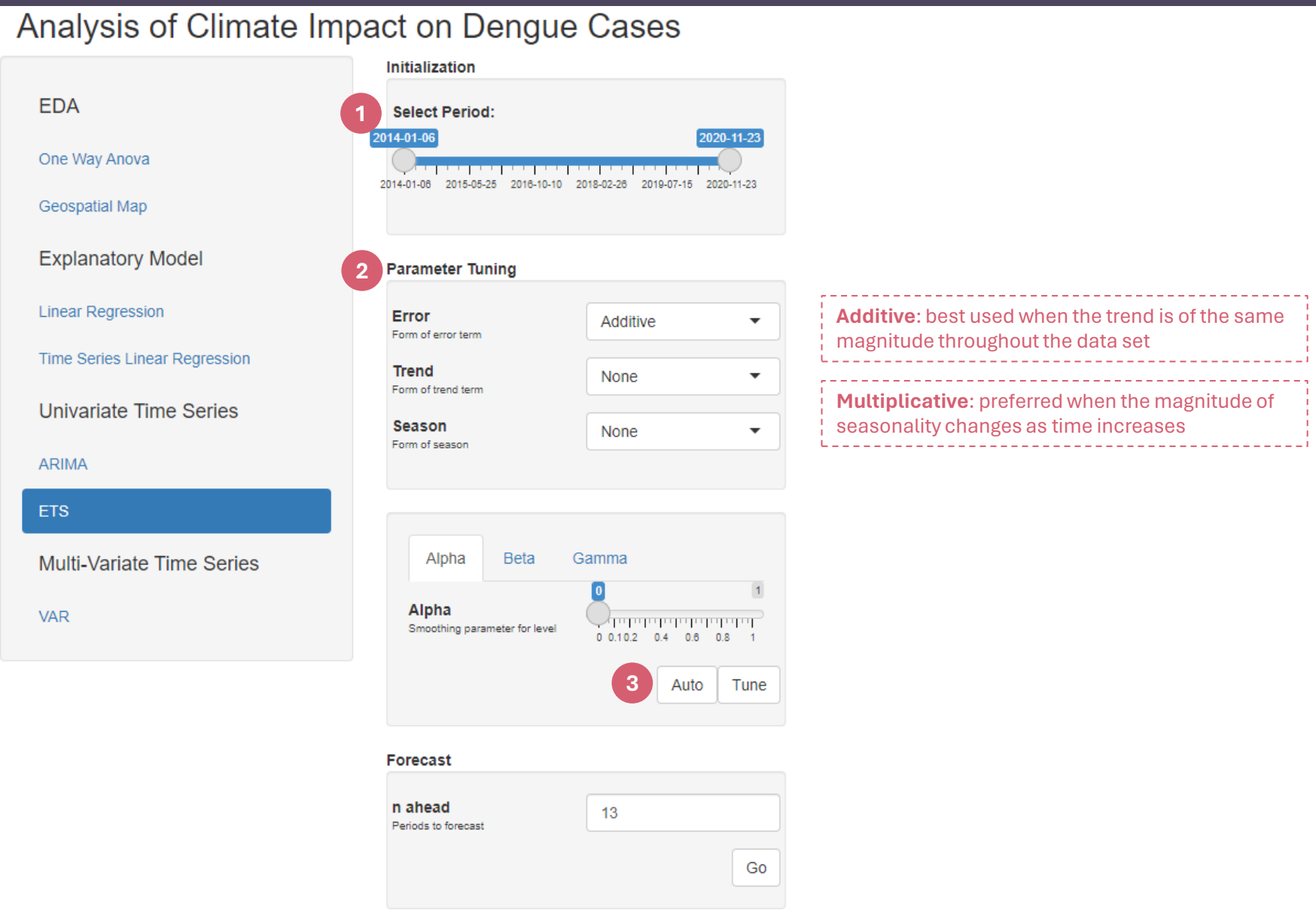
View forecast results in plot, along with 80% and 95% confidence interval

Data table

View forecast results in table format

Univariate Time Series – ETS

This module allows users to fit an Exponential Smoothing model on Dengue Cases and generate forecast



Additive: best used when the trend is of the same magnitude throughout the data set

Multiplicative: preferred when the magnitude of seasonality changes as time increases

1) Select period of analysis

2) Select Input parameters to use for the model

3) Alternative to step 2, users can click “Auto” to fill parameter based on model’s AICC automatically

Univariate Time Series – ETS

This module allows users to fit an Exponential Smoothing model on Dengue Cases and generate forecast

Analysis of Climate Impact on Dengue Cases

EDA
One Way Anova
Geospatial Map
Explanatory Model
Linear Regression
Time Series Linear Regression
Univariate Time Series
ARIMA
ETS
Multi-Variate Time Series
VAR

Initialization

Select Period:

2014-01-06 — 2020-11-23

2014-01-06 2015-05-25 2016-10-10 2018-02-26 2019-07-15 2020-11-23

Parameter Tuning

Error
Form of error term: Multiplicative

Trend
Form of trend term: Additive

Season
Form of season: None

Alpha | Beta | Gamma

Alpha
Smoothing parameter for level

0 0.83 1

0 0.1 0.2 0.4 0.6 0.8 1

Auto | Tune

4

Forecast

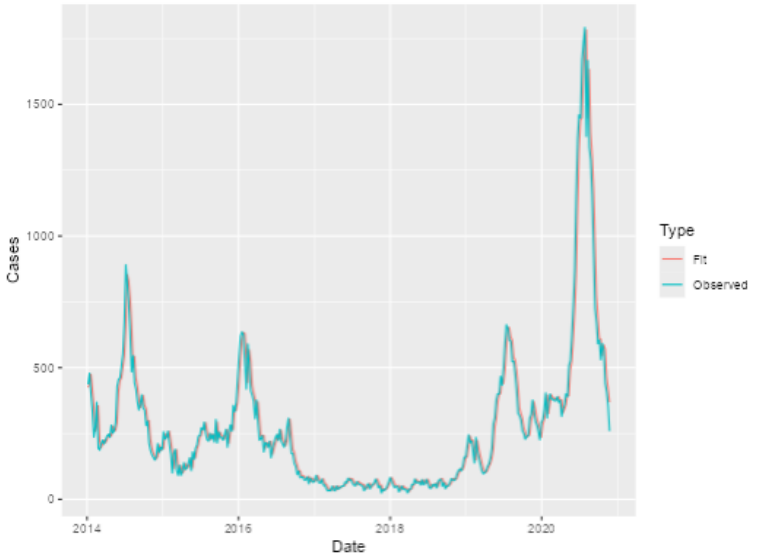
n ahead
Periods to forecast: 13

Go

5

Actual vs Fit | Components | Report

Observed vs Fitted



Cross Validation Metrics

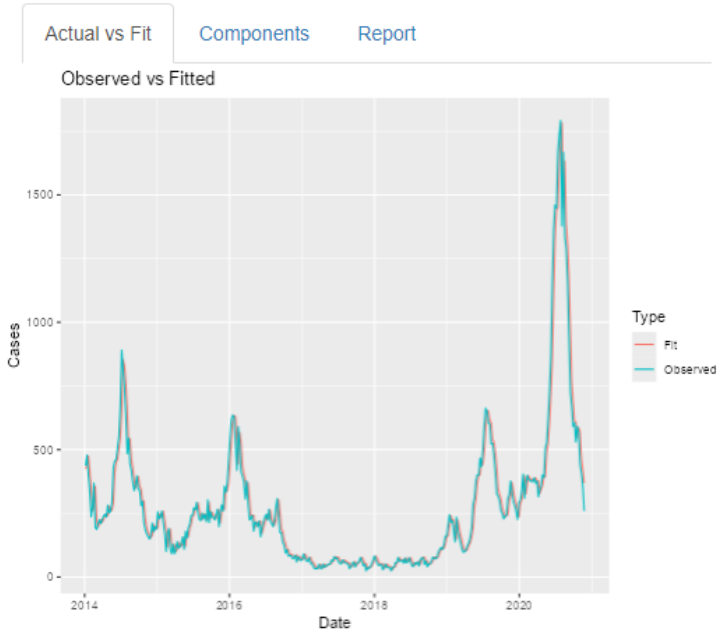
RMSE	MAE	MAPE
60.64	35.15	22.08

5) Toggle between tab panels for various diagnostics

4) Click “Tune” to generate model diagnostics

Univariate Time Series – ETS

This module allows users to fit an Exponential Smoothing model on Dengue Cases and generate forecast



Cross Validation Metrics

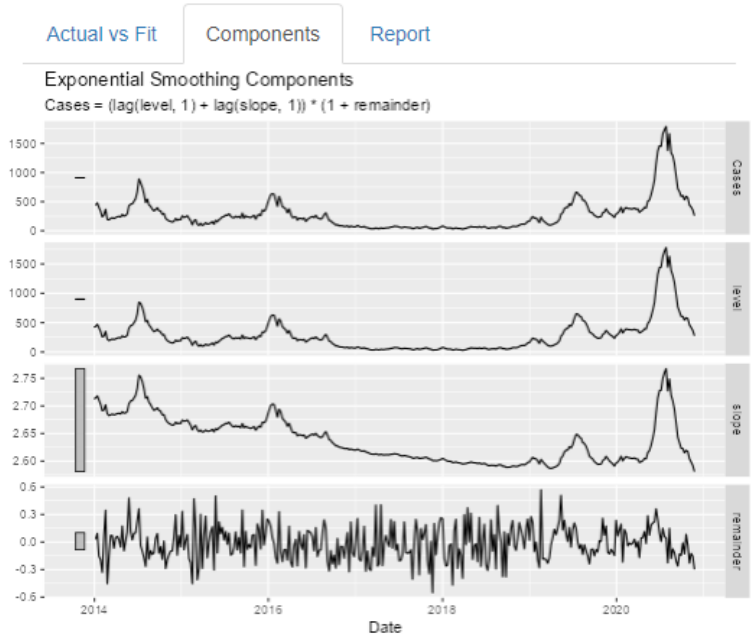
RMSE	MAE	MAPE
60.64	35.15	22.08

Actual vs Fit

Check the fit of the ETS model against the actual values

Cross Validation Metrics

Check and compare the performance of the model using Cross Validation metrics

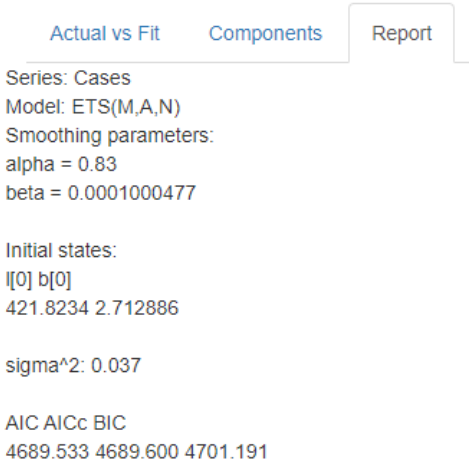


Cross Validation Metrics

RMSE	MAE	MAPE
60.64	35.15	22.08

Components

Check the decomposition of dependent variable to observe the levels of Trend, Error, and Seasonality, if any.



Report

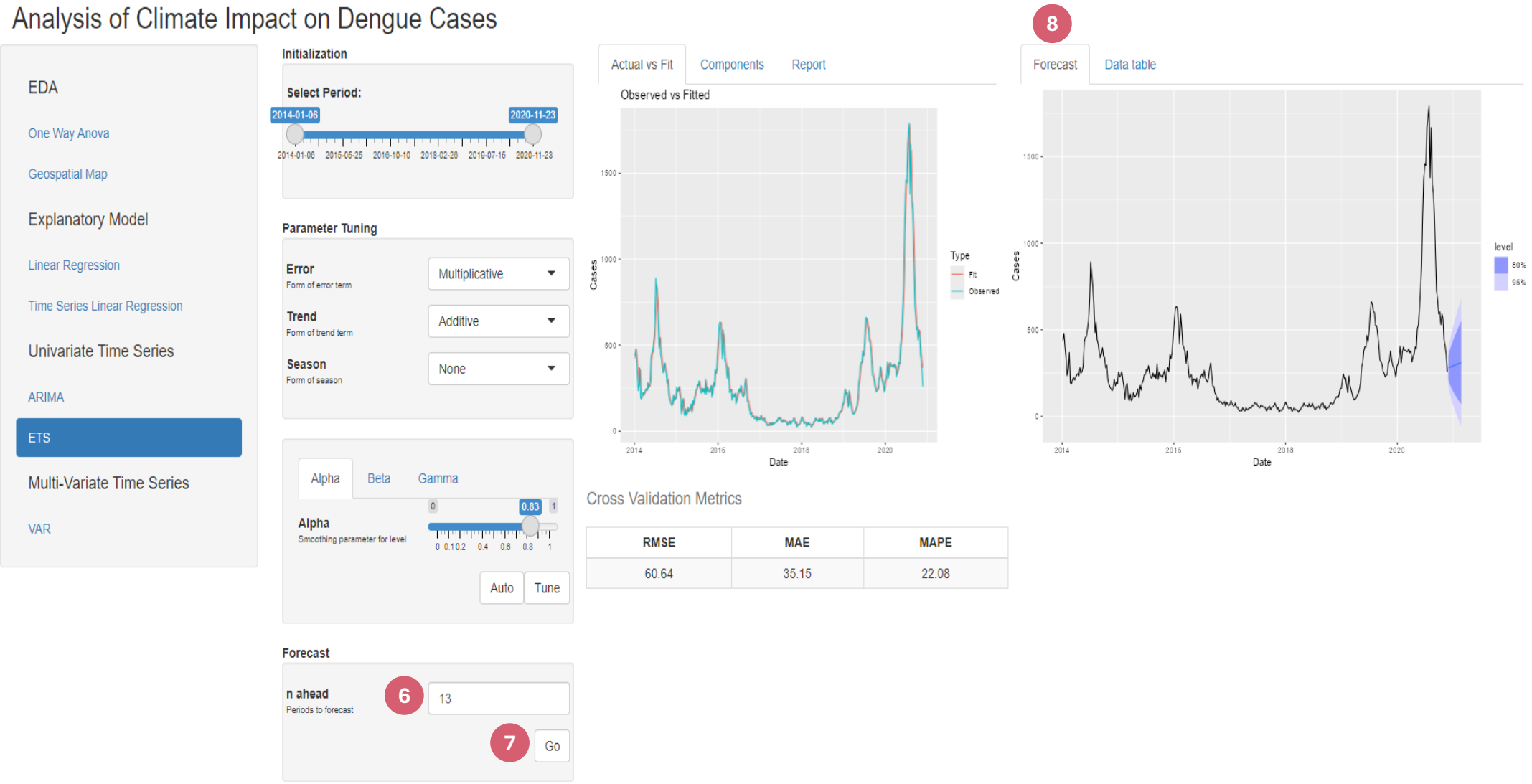
Check for other stats and parameter values in the model report

Model Diagnostics

Users can use these diagnostics to determine if parameters are suitable, and re-tune models accordingly

Univariate Time Series – ETS

This module allows users to fit an Exponential Smoothing model on Dengue Cases and generate forecast



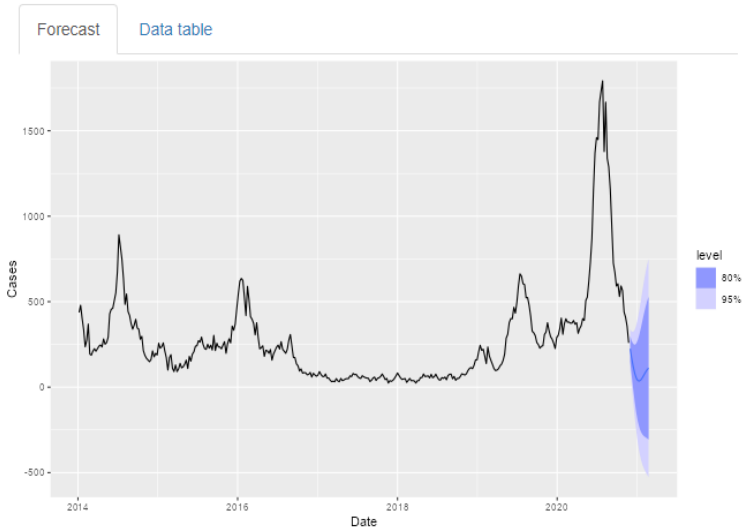
8) Toggle between tab panels for various forecast results

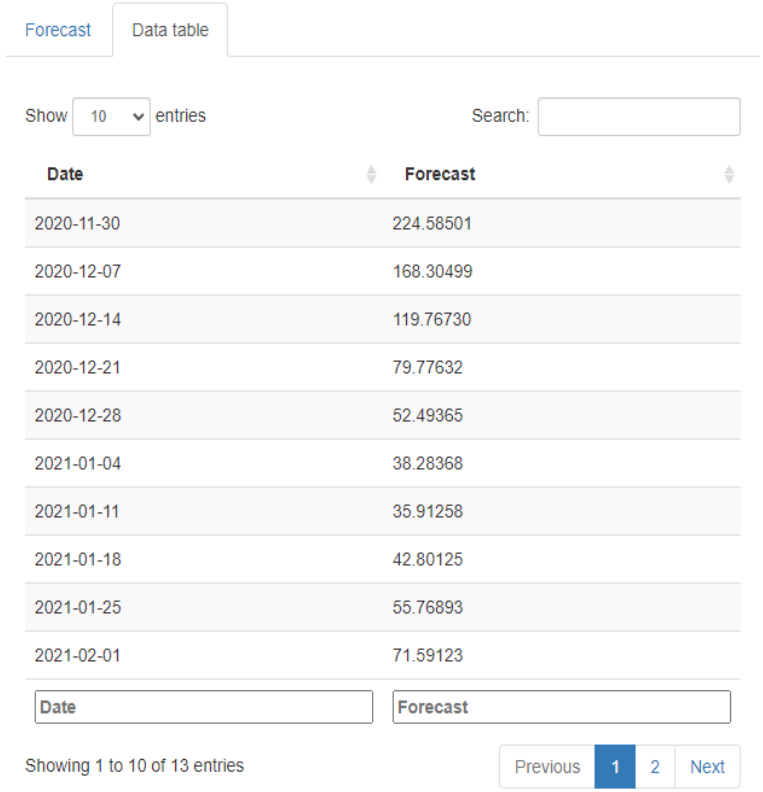
6) Key in number of forecast period

7) Click“Go” to generate forecast

Univariate Time Series – ETS

This module allows users to fit an Exponential Smoothing model on Dengue Cases and generate forecast





Model Forecast

Users can use these forecast results to compare with that from other models

<u>Forecast</u>

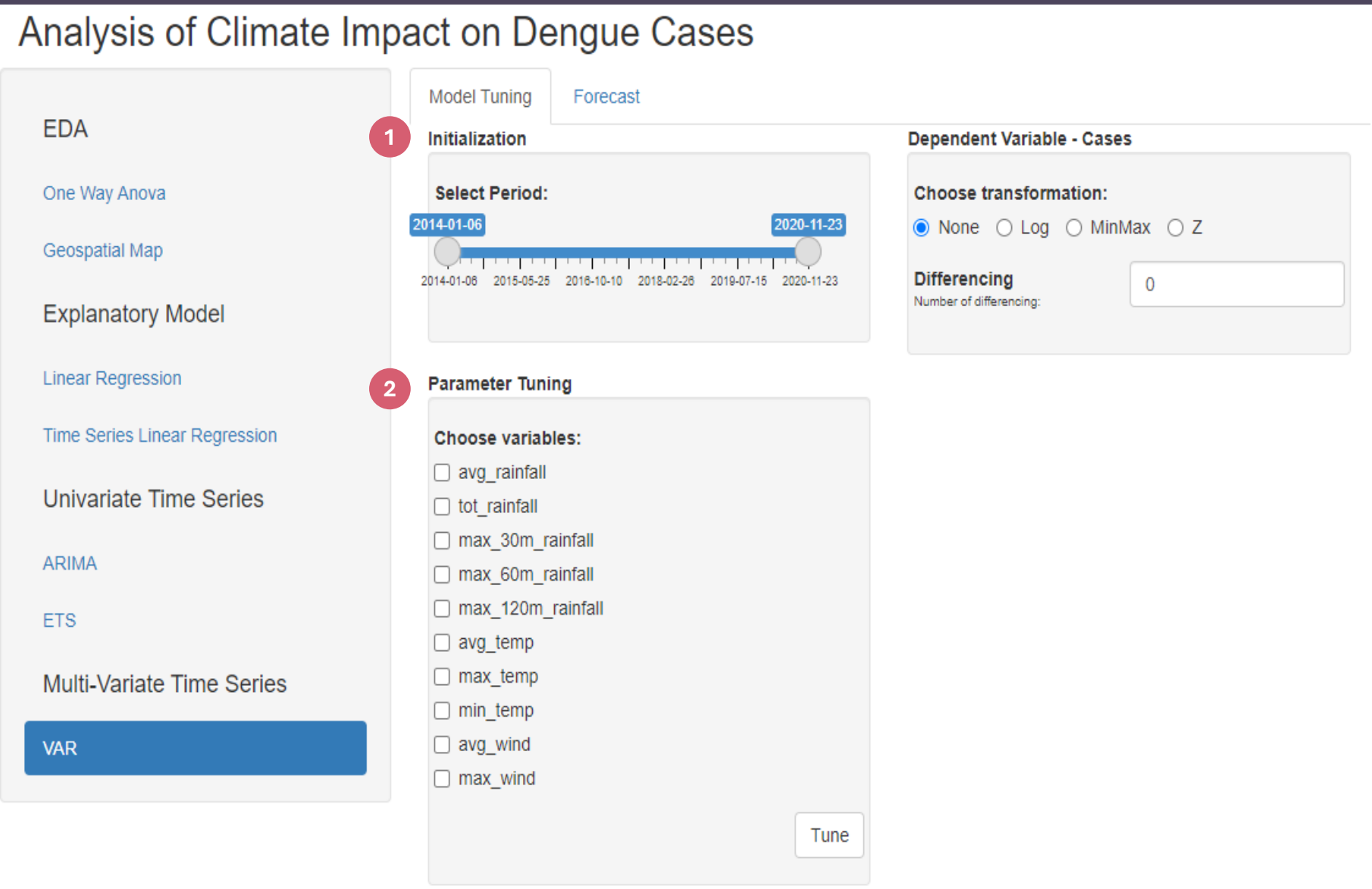
View forecast results in plot, along with 80% and 95% confidence interval

<u>Data table</u>

View forecast results in table format

Multivariate Time Series – VAR

This module allows users to fit a Vector AutoRegressions model on Dengue Cases and generate forecast

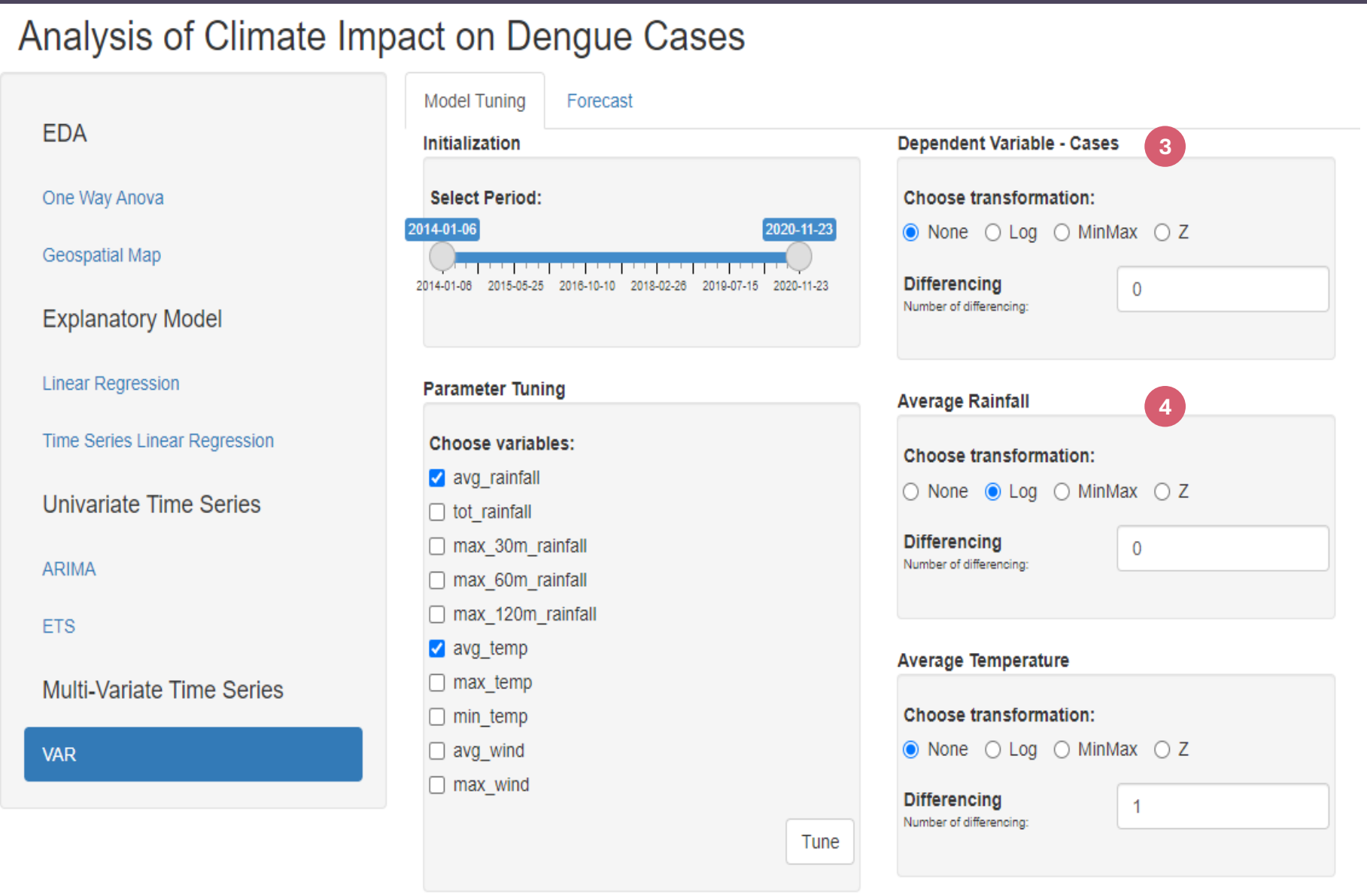


1) Select period of analysis

2) Select Input parameters to use for the model

Multivariate Time Series – VAR

This module allows users to fit a Vector AutoRegressions model on Dengue Cases and generate forecast



3) Once a variable is added into the model, an additional panel will appear in the middle column

4) Select log, min-max transformation or standardization for each variable, if needed

Users can also determine the differencing required

Multivariate Time Series – VAR

This module allows users to fit a Vector AutoRegressions model on Dengue Cases and generate forecast

Analysis of Climate Impact on Dengue Cases

- EDA
- One Way Anova
- Geospatial Map
- Explanatory Model
- Linear Regression
- Time Series Linear Regression
- Univariate Time Series
- ARIMA
- ETS
- Multi-Variate Time Series
- VAR

Model Tuning | Forecast

Initialization

Select Period: 2014-01-06 – 2020-11-23

2014-01-06 2015-05-25 2016-10-10 2018-02-26 2019-07-15 2020-11-23

Parameter Tuning

Choose variables:

- [x] avg_rainfall
- [ ] tot_rainfall
- [ ] max_30m_rainfall
- [ ] max_60m_rainfall
- [ ] max_120m_rainfall
- [x] avg_temp
- [ ] max_temp
- [ ] min_temp
- [ ] avg_wind
- [ ] max_wind

5 Tune

Dependent Variable - Cases

Choose transformation: (•) None ( ) Log ( ) MinMax ( ) Z

Differencing Number of differencing: 0

Average Rainfall

Choose transformation: ( ) None (•) Log ( ) MinMax ( ) Z

Differencing Number of differencing: 0

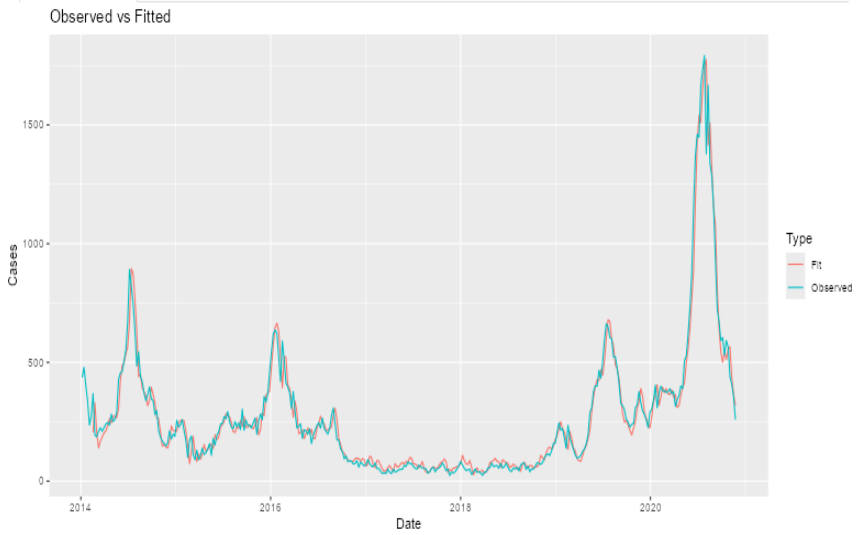
Average Temperature

Choose transformation: (•) None ( ) Log ( ) MinMax ( ) Z

Differencing Number of differencing: 1

6

Actual vs Fit | ACF



Cross Validation Metrics

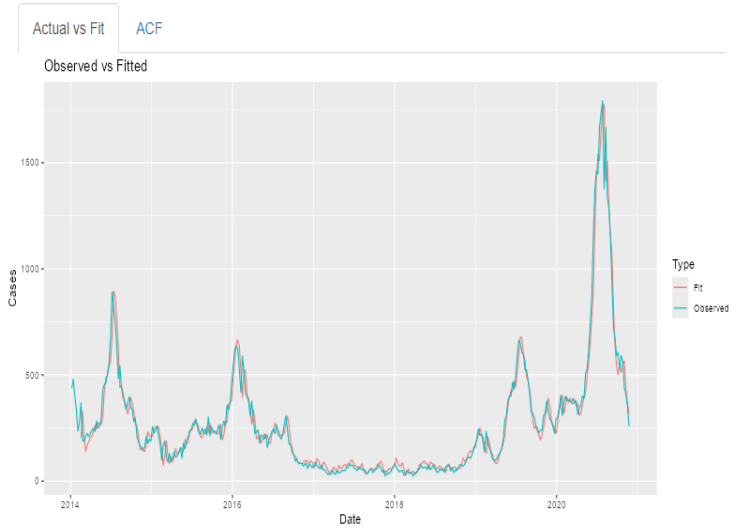
.response	RMSE	MAE	MAPE
Cases	61.29	37.26	36.52
diff1_avg_temp	0.80	0.74	111.69
log_avg_rainfall	0.60	0.49	41.27

6) Toggle through tab panels to view various model diagnostics

5) Click “Tune” to generate results

Multivariate Time Series – VAR

This module allows users to fit a Vector AutoRegressions model on Dengue Cases and generate forecast

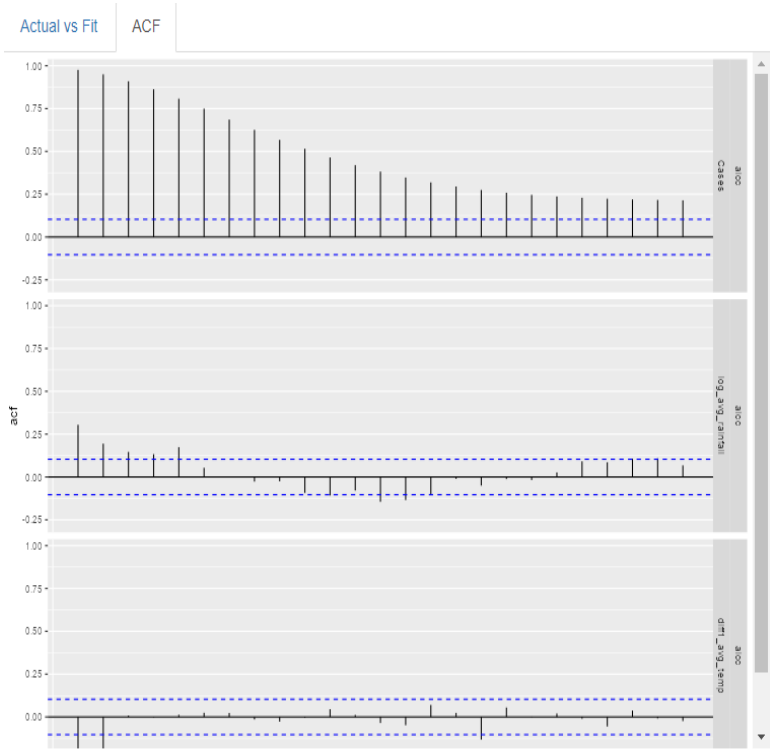


Cross Validation Metrics

.response	RMSE	MAE	MAPE
Cases	61.29	37.26	36.52
diff1_avg_temp	0.80	0.74	111.69
log_avg_rainfall	0.60	0.49	41.27

Actual vs Fit

Check model fit and metrics of individual responses

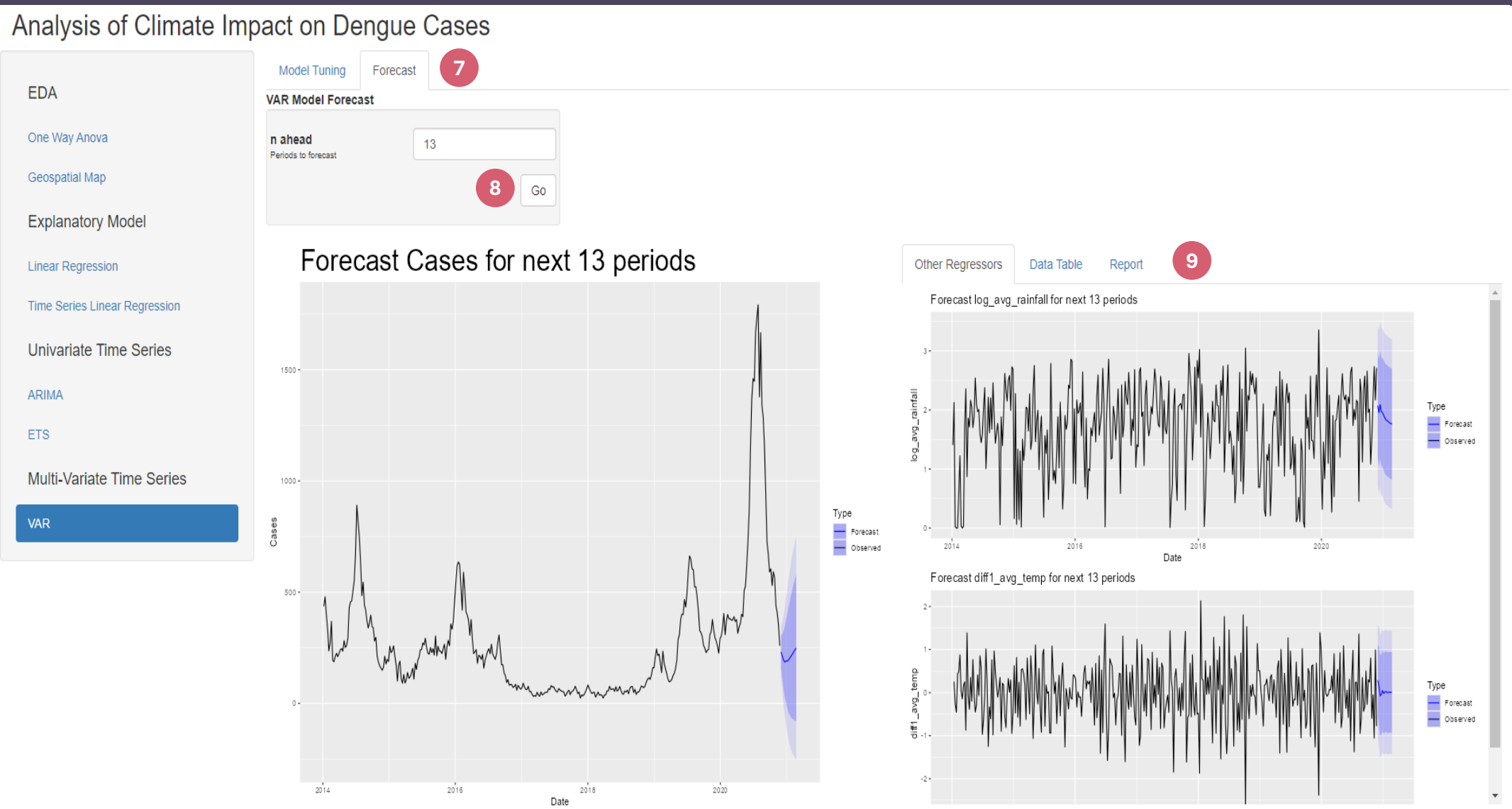


ACF

Check ACF plots for all variables to determine stationarity

Multivariate Time Series – VAR

This module allows users to fit a Vector AutoRegressions model on Dengue Cases and generate forecast



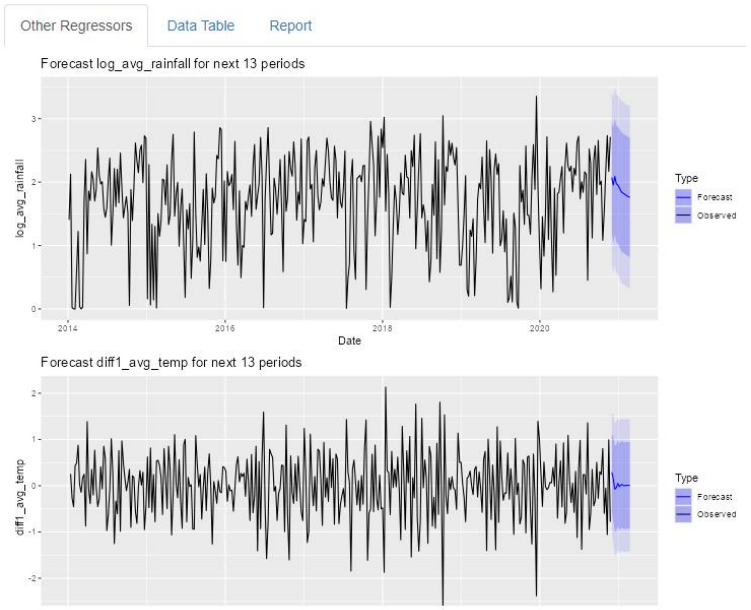
7) Toggle to “Forecast” tab panel for forecasting

8) Click“Go” to generate forecast

9) Toggle between tab panels for various forecast results

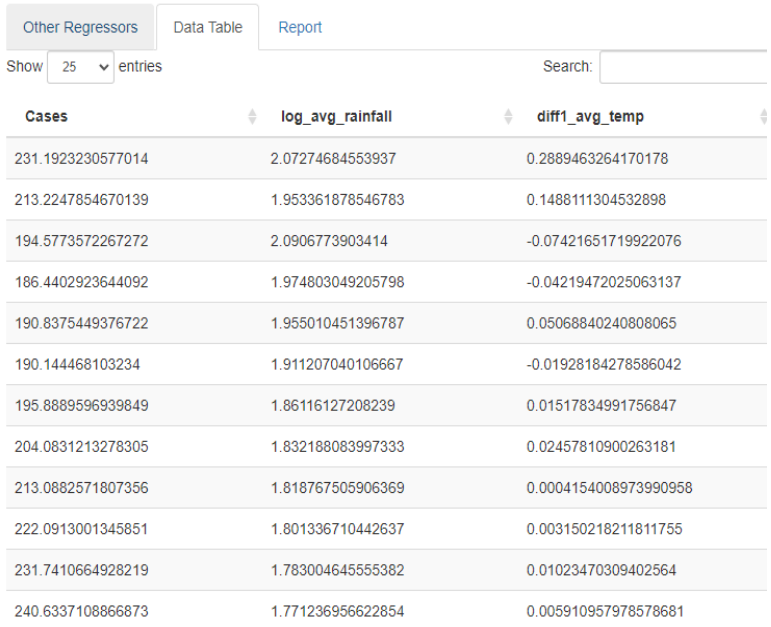
Multivariate Time Series – VAR

This module allows users to fit a Vector AutoRegressions model on Dengue Cases and generate forecast



Other Regressors

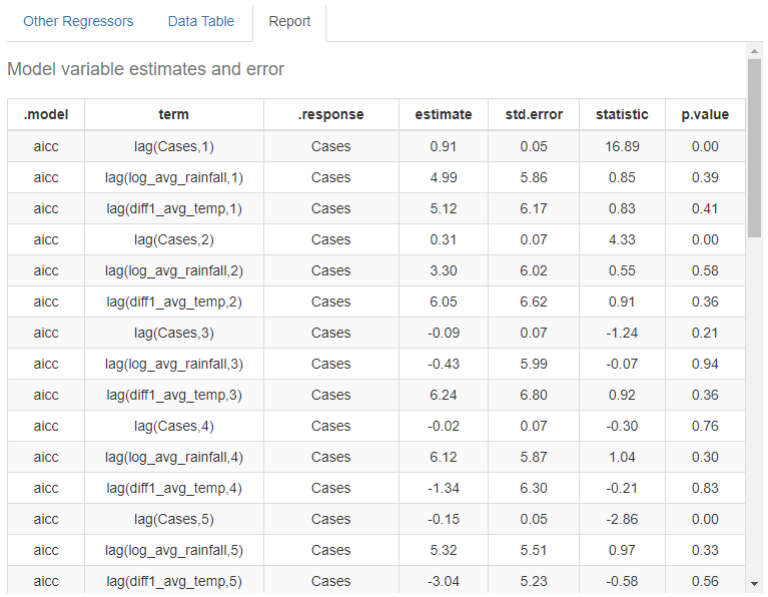
View the time series forecast of each individual regressors in the VAR model



Cases	log_avg_rainfall	diff1_avg_temp
231.1923230577014	2.07274684553937	0.2889463264170178
213.2247854670139	1.953361878546783	0.1488111304532898
194.5773572267272	2.0906773903414	-0.07421651719922076
186.4402923644092	1.974803049205798	-0.04219472025063137
190.8375449376722	1.955010451396787	0.05068840240808065
190.144468103234	1.911207040106667	-0.01928184278586042
195.8889596939849	1.86116127208239	0.01517834991756847
204.0831213278305	1.832188083997333	0.02457810900263181
213.0882571807356	1.818767505906369	0.0004154008973990958
222.0913001345851	1.801336710442637	0.003150218211811755
231.7410664928219	1.783004645555382	0.01023470309402564
240.6337108866873	1.771236956622854	0.005910957978578681

Data Table

Obtain the forecast results in table format



.model	term	.response	estimate	std.error	statistic	p.value
aicc	lag(Cases,1)	Cases	0.91	0.05	16.89	0.00
aicc	lag(log_avg_rainfall,1)	Cases	4.99	5.86	0.85	0.39
aicc	lag(diff1_avg_temp,1)	Cases	5.12	6.17	0.83	0.41
aicc	lag(Cases,2)	Cases	0.31	0.07	4.33	0.00
aicc	lag(log_avg_rainfall,2)	Cases	3.30	6.02	0.55	0.58
aicc	lag(diff1_avg_temp,2)	Cases	6.05	6.62	0.91	0.36
aicc	lag(Cases,3)	Cases	-0.09	0.07	-1.24	0.21
aicc	lag(log_avg_rainfall,3)	Cases	-0.43	5.99	-0.07	0.94
aicc	lag(diff1_avg_temp,3)	Cases	6.24	6.80	0.92	0.36
aicc	lag(Cases,4)	Cases	-0.02	0.07	-0.30	0.76
aicc	lag(log_avg_rainfall,4)	Cases	6.12	5.87	1.04	0.30
aicc	lag(diff1_avg_temp,4)	Cases	-1.34	6.30	-0.21	0.83
aicc	lag(Cases,5)	Cases	-0.15	0.05	-2.86	0.00
aicc	lag(log_avg_rainfall,5)	Cases	5.32	5.51	0.97	0.33
aicc	lag(diff1_avg_temp,5)	Cases	-3.04	5.23	-0.58	0.56

Report

Gain insights on variable estimates and statistics