

EPIDEMIOLOGICAL FACTORS OF TUBERCULOSIS IN SOUTH ASIAN COUNTRIES : TIME SERIES ANALYSIS

Guided by,

Dr K Sudhakar Prasad

Associate Professor

Yenepoya research centre &

Yenepoya (Deemed to be University).

Presented by,
Sahana

MSc Biostatistics
20MSCB05

Department of Statistics
YENEPOYA
(Deemed to be University)
Mangalore

Content

Chapter I

- Introduction
- Aim
- Objectives
- Targeted population
- Data source and terminology
- Software used

Chapter II

Exploratory analysis

Chapter III

Time series analysis

Chapter IV

- Summary and conclusion
- Scope for further study
- References

Introduction

- Health science statistics are the form of evidence, or facts that can support a conclusion.
- According to Health statistics measure four types of information.
 1. **Correlates:** See how to measure the risk factors and protective factors that impact our health.
 2. **Conditions:** Learn to assess how often and how badly diseases impact a community.
 3. **Care:** Dig into how health care is delivered to the communities that need it, to treat disease and illness.
 4. **Costs:** Get more information on what health care costs, and why.

Introduction...

- Acute respiratory infections (ARI) constitute one of the principal causes of morbidity and mortality in many countries.
- Tuberculosis (TB) is a contagious bacterial infection that usually attacks the lungs. But it can also attack other parts of the body, including the kidneys, spine, and brain. It is one of lower respiratory infection.

There are two types of TB conditions:

- ✓ **Latent TB infection**, where the TB germs live in your body but don't make sick.
- ✓ **TB disease (active TB)** where you get sick from the TB germs.



Introduction...

- HIV is the major confounding for TB mortality.
- **Epidemiology** is a study of distribution determinant of health-related states or event in a specified population and application of the study to control the health problems. In a simple way it can be stated that it is a study about disease and its betterment.
- It helps to predict disease trend and Identify the health needs of community.

Aim

- To develop appropriate statistical models for Tuberculosis prediction.



Objective

- To study the pattern in annual TB incidence number and annual TB and HIV incidence number of south Asian countries.
- To study the pattern in annual TB mortality number and with HIV confounding mortality number of south Asian countries.
- To predict the annual incidence rate of TB .
- To predict the annual mortality rate of TB .

Countries considered in the study



Afghanistan



Bhutan



Bangladesh



Pakistan



India



Nepal



Sri Lanka

Data source and terminologies

- The secondary data is available at the official website of World Health Organization (WHO). and can be accessed through the link <https://www.who.int/teams/global-tuberculosis-programme/data> .
- Incidence rate and mortality rates of tuberculosis are the important terminologies in this study. They are defined as follows:

$$\text{Incidence rate} = \frac{\text{number of new cases of specific disease during the given time period}}{\text{population at risk during time period}} \times 100000$$

$$\text{Mortality rate} = \frac{\text{number of death cases of specific disease during the given time period}}{\text{population at risk during time period}} \times 100000$$

Software Used

- The R studio (version 4.1.1) is used for model building, MS excel and SPSS is used for exploratory analysis.





Exploratory
analysis

Afghanistan

- There are on an average of 29763795 population in Afghanistan from 2000 to 2020

Estimates	Mean	median	Standard deviation	Quartiles		
				25th	50th	75th
Estimated Incident number	56333.33	55000.00	10716.032	47500.00	55000.00	66000.00
Estimated TB and HIV incident number	12.71	13.00	4.941	7.50	13.00	17.50
Estimated mortality number of TB cases excluding HIV	12371.43	13000.00	1413.203	11000.00	13000.00	14000.00
Estimated mortality number of TB and HIV cases	111.19	120.00	38.505	75.00	120.00	140.00
Estimated mortality number	12523.81	13000.0	1364.516	11000.00	13000.00	14000.00

Bangladesh

- There are on an average of 147286329 population in Afghanistan from 2000 to 2020

Estimates	Mean	median	Standard deviation	Quartiles		
				25	50	75
Estimated Incident number	325619.05	326000	24430.465	305500	326000	347500
Estimated TB and HIV incident number	474.19	550	222.578	250	550	665
Estimated mortality number of TB cases excluding HIV	73095.24	76000	16364.305	63500	76000	87000
Estimated mortality number of TB and HIV cases	161.43	180	99.302	55.50	180	260
Estimated mortality number	73285.71	76000	16248.516	63500	76000	87000

Bhutan

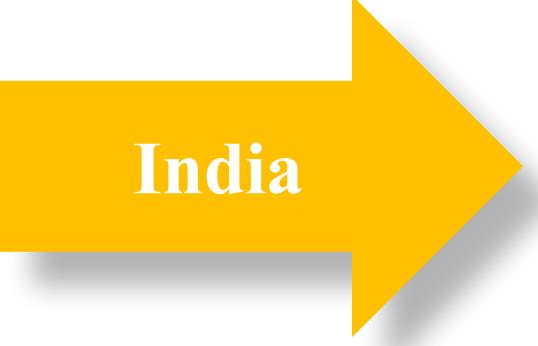
- There are on an average of 686058 population in Afghanistan from 2000 to 2020

Estimates	Mean	median	Standard deviation	Quartiles		
				25th	50th	75th
Estimated Incident number	1304.76	13	132.198	1200	13	1400
Estimated TB and HIV incident number	41.38	46	26.607	9.50	46	62
Estimated mortality number of TB cases excluding HIV	140.95	140	15.781	130	140	150
Estimated mortality number of TB and HIV cases	10.95	12	7.440	2	12	17.50
Estimated mortality number	151.43	150	18.244	140	150	160

Pakistan

- There are on an average of 180200014 population in Afghanistan from 2000 to 2020

Estimates	Mean	median	Standard deviation	Quartiles		
				25	50	75
Estimated Incident number	489809.52	495000	58298.044	437000	495000	542500
Estimated TB and HIV incident number	3721.90	2700	3598.423	745	2700	5950
Estimated mortality number of TB cases excluding HIV	47952.38	47000	3680.709	45000	47000	52000
Estimated mortality number of TB and HIV cases	821.19	610	911.025	23.50	610	1300
Estimated mortality number	48857.14	49000	3021.353	46000	49000	52000



India

- There are on an average of 1227616390 population in Afghanistan from 2000 to 2020

Estimates	Mean	median	Standard deviation	Quartiles		
				25	50	75
Estimated Incident number	2980000	3050000	197027.917	2815000	3050000	3155000
Estimated TB and HIV incident number	193428.57	195000	97043.068	103500	195000	245000
Estimated mortality number of TB cases excluding HIV	512714.29	494000	61263.482	453500	494000	566500
Estimated mortality number of TB and HIV cases	51595.24	55000	29076.459	17000	55000	74500
Estimated mortality number	564285.71	549000	88569.827	477500	549000	639000

Nepal

- There are on an average of 26565556 population in Afghanistan from 2000 to 2020

Estimates	Mean	median	Standard deviation			
				25	50	75
Estimated Incident number	83238.10	84000	11126.117	71500	84000	93500
Estimated TB and HIV incident number	2188.57	2600	920.572	1400	2600	2900
Estimated mortality number of TB cases excluding HIV	22095.24	22000	4815.649	17000	22000	26500
Estimated mortality number of TB and HIV cases	529.5	520	238.493	330	520	710
Estimated mortality number	22619.05	22000	4873.153	18000	22000	27000

Sri Lanka

- There are on an average of 20200994 population in Afghanistan from 2000 to 2020

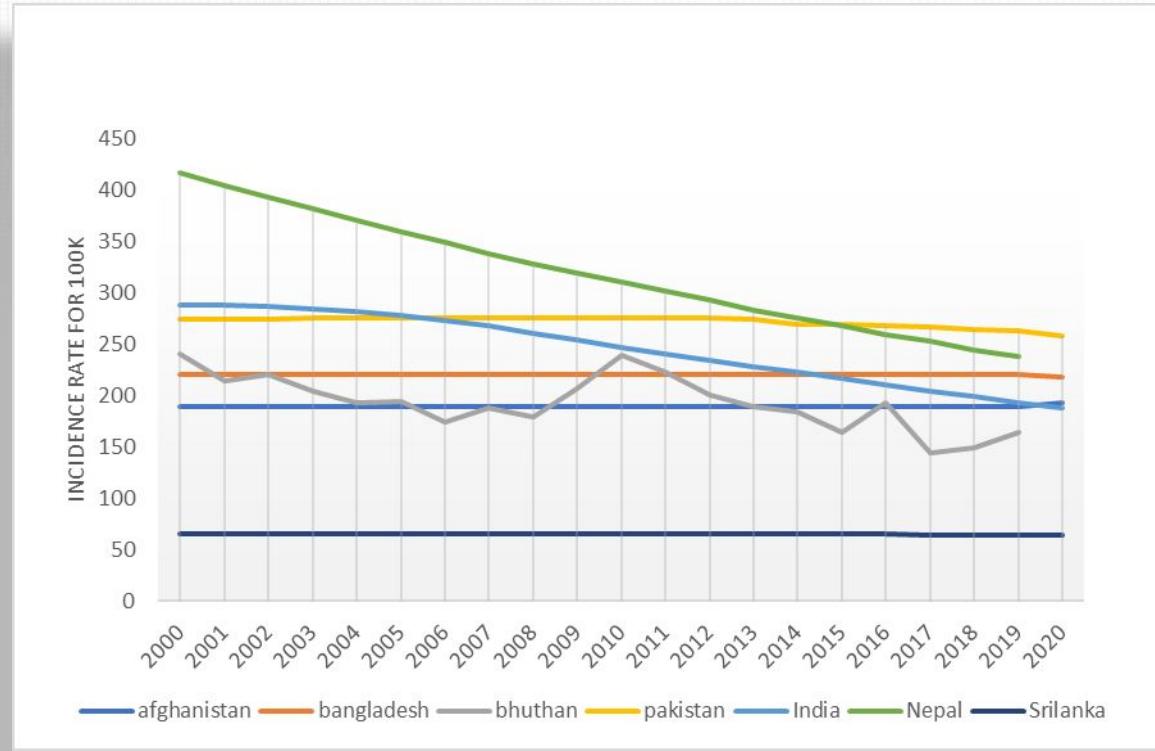
Estimates	Mean	median	Standard deviation	Quartiles		
				25	50	75
Estimated Incident number	13428.57	13000	597.614	13000	13000	14000
Estimated TB and HIV incident number	61.71	62	24.286	46.50	62	78
Estimated mortality number of TB cases excluding HIV	994.76	950	298.523	740	950	1200
Estimated mortality number of TB and HIV cases	19.19	19	7.068	14.50	19	23
Estimated mortality number	1010.48	980	296.723	760	980	1700

Time series

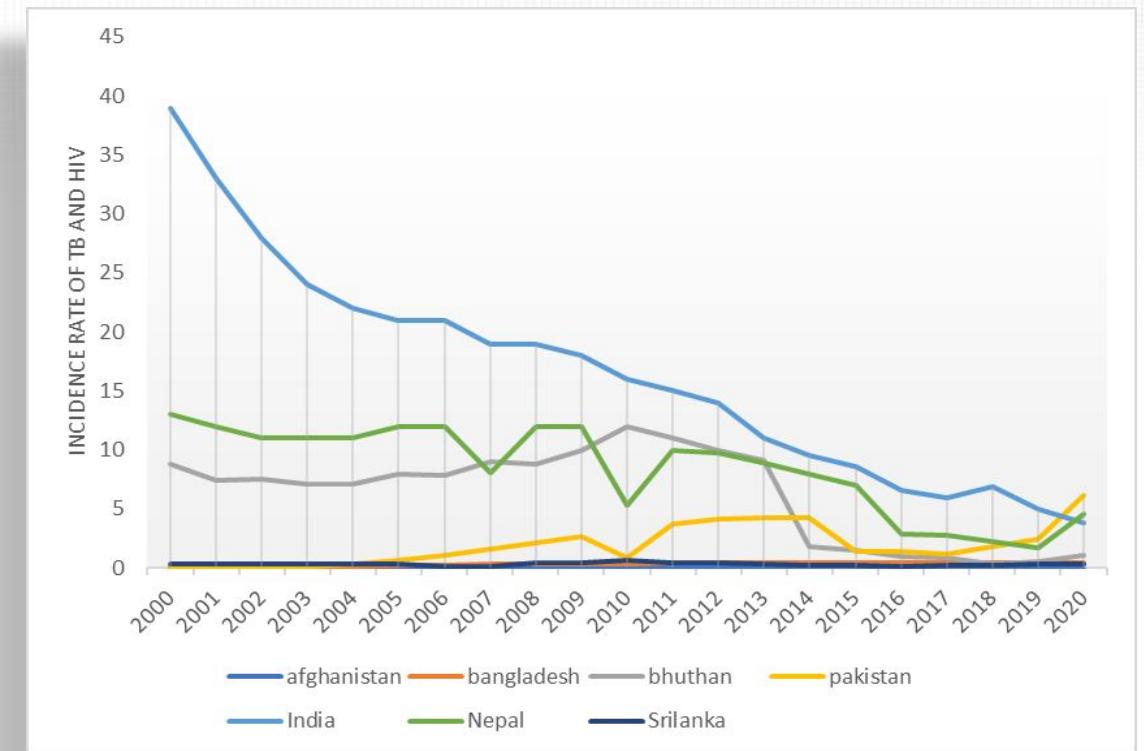
- TB incidence and mortality rate is assumed to be time dependent sequence hence time series methods are used for the analysis.
- A Time Series is a sequence of observations ordered in equally spaced, discrete time intervals.
- A basic assumption in any time series analysis/modelling is that some aspects of past pattern will continue to remain the same in the future.



The following time profiles depict the nature of Estimated Incidence rate of TB & confounding with HIV per 100k in south Asian countries

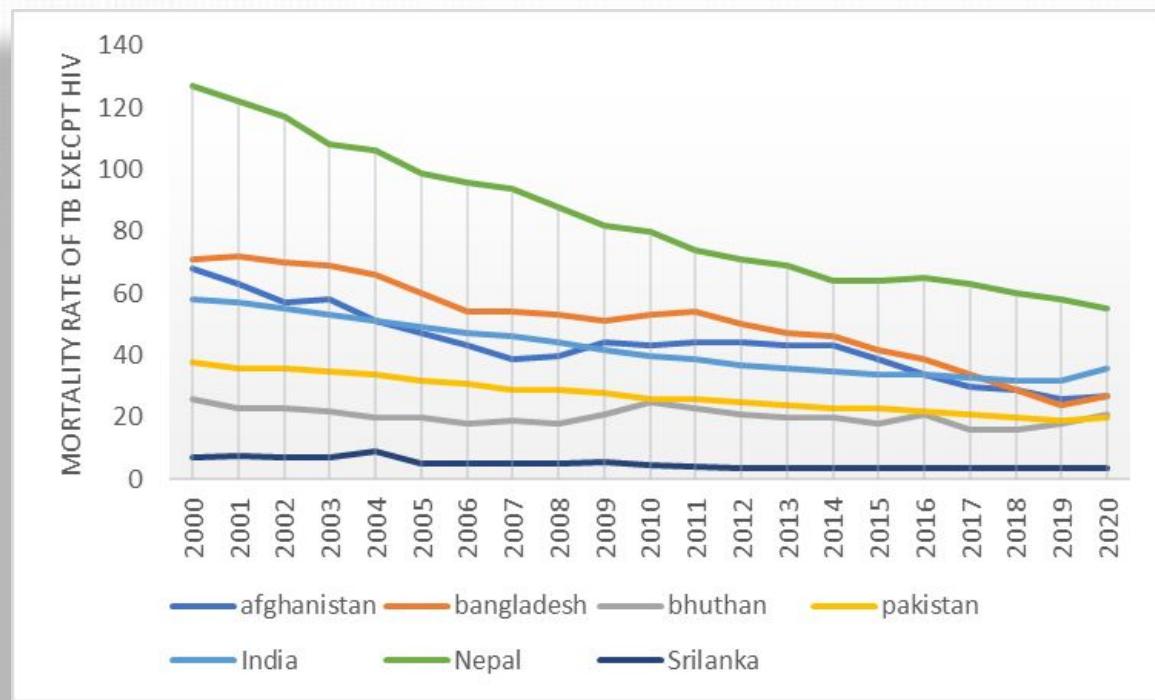


Country wise time profile of yearly incidence rate of TB

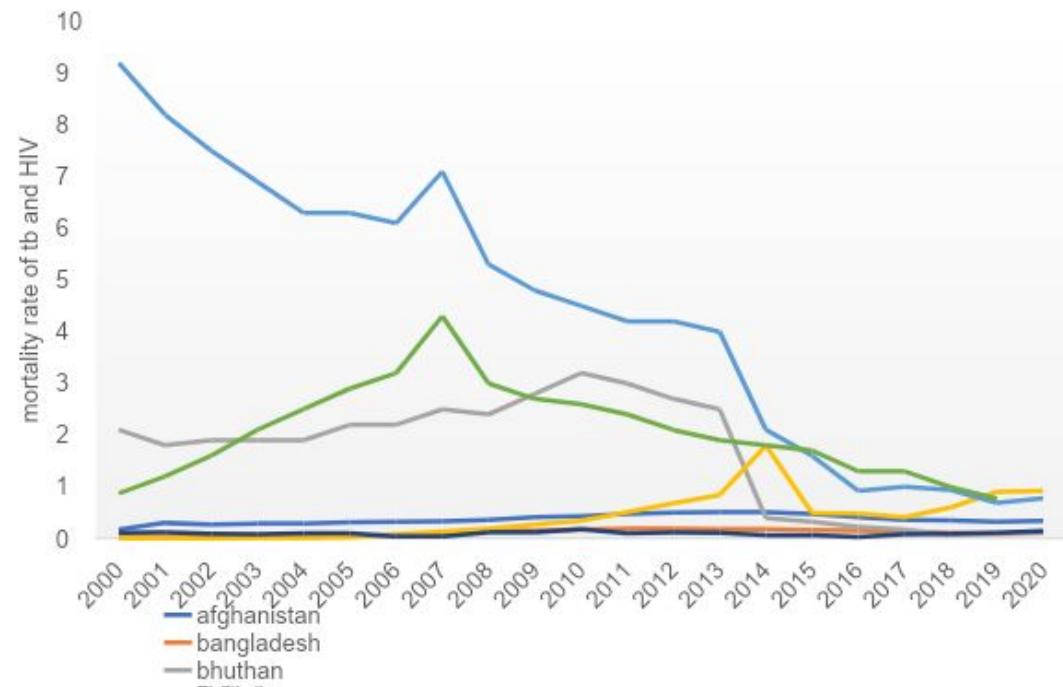


Country wise time profile of yearly incidence rate of TB and HIV

The following time profiles depict the nature of Mortality rate of TB and confounding with HIV for 100k in south Asian countries

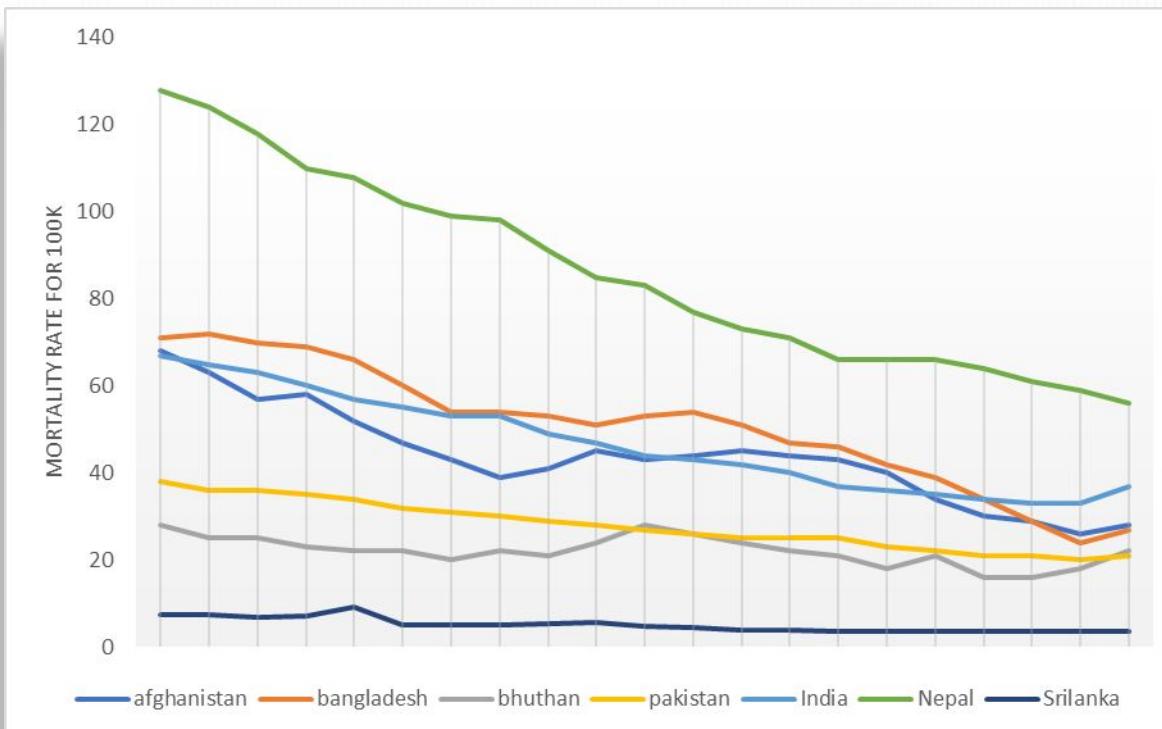


Country wise time profile of yearly mortality rate of TB except HIV



Country wise time profile of yearly incidence rate of TB and HIV

The following time profiles depict the nature of Mortality rate of TB per 100k in south Asian countries

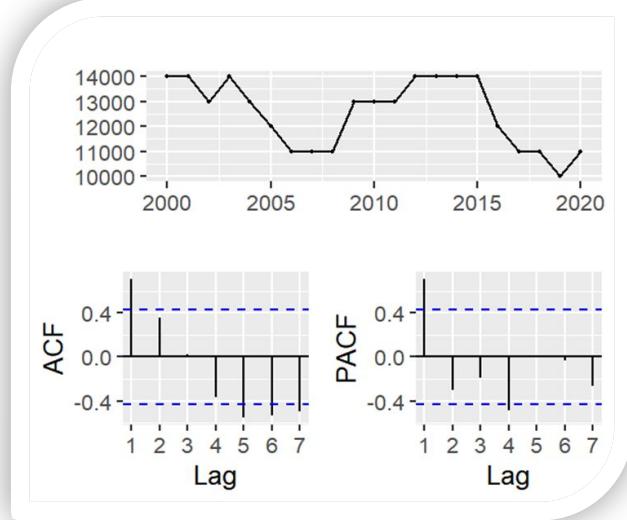
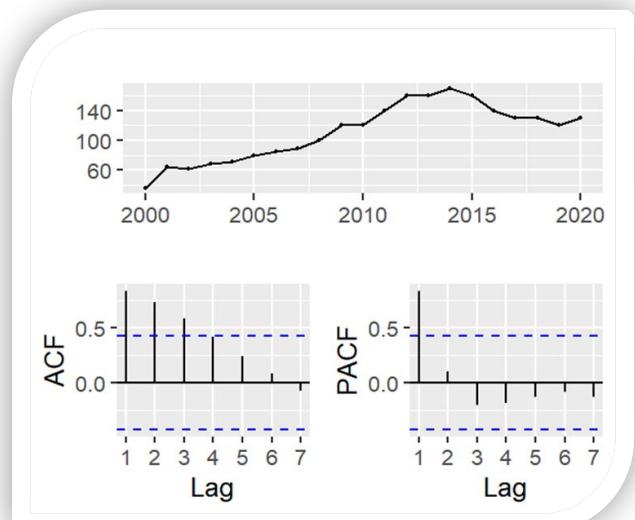
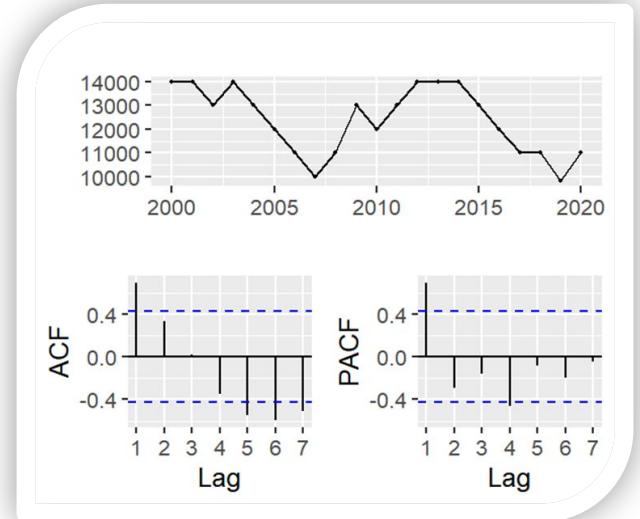
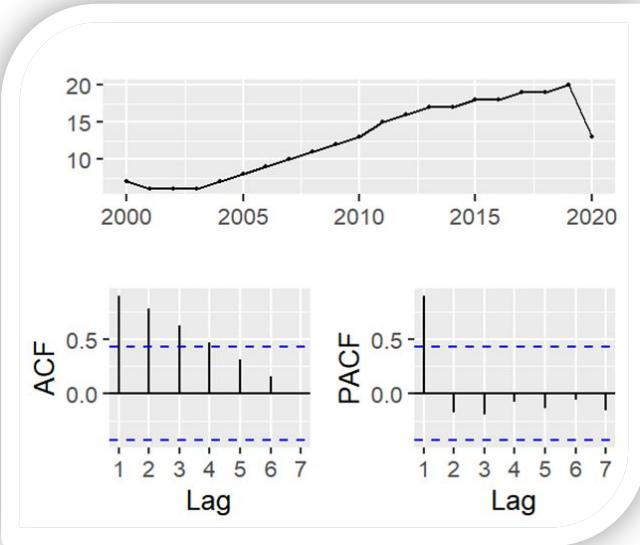
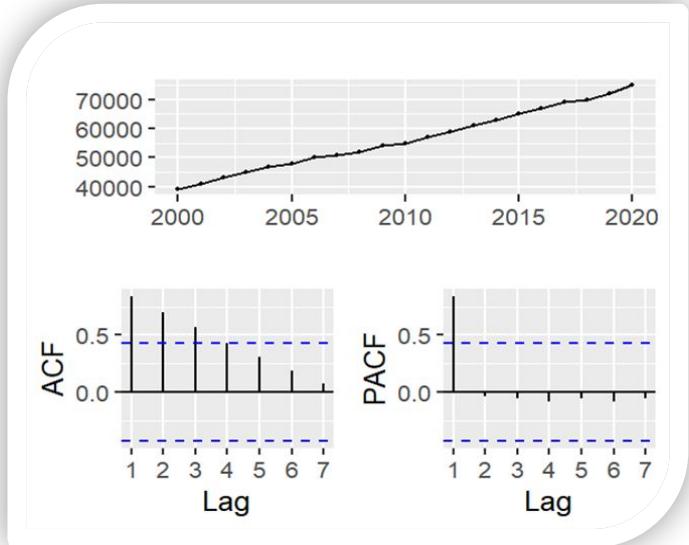


Country wise time profile of yearly mortality rate of TB

Afghanistan

- Time profile , ACF and PACF
- Time series model
- Prediction table
- Interpretation





Time series Models

Country	Estimates	Order			AIC	ME	RMSE	MAE
				q				
Afghanistan	Estimated Incident number	0	1	0	310.13	1.7714	497.6796	382.723
	Estimated TB and HIV incident number	0	1	0	82.64	0.2860	1.7728	1.0480
	Estimated mortality number of TB cases excluding HIV	1	0	0	353.6	-81.7908	932.8482	850.1889
	Estimated mortality number of TB and HIV cases	0	1	0	159.93	0.0014	11.6451	9.0014
	Estimated mortality number	1	0	0	350.72	-79.3141	869.2394	731.1792

Prediction Table

Afghanistan					
	2021	2022	2023	2024	2025
Estimated Incident number	76800	78600	80400	82200	84000
Estimated TB and HIV incident number	13	13	13	13	13
Estimated mortality number of TB cases excluding HIV	11360.10	11627.52	11826.12	11973.60	12083.12
Estimated mortality number of TB and HIV cases	134.75	139.50	144.25	149	153.75
Estimated mortality number	11354.92	11626.87	11835.25	11994.91	12117.25

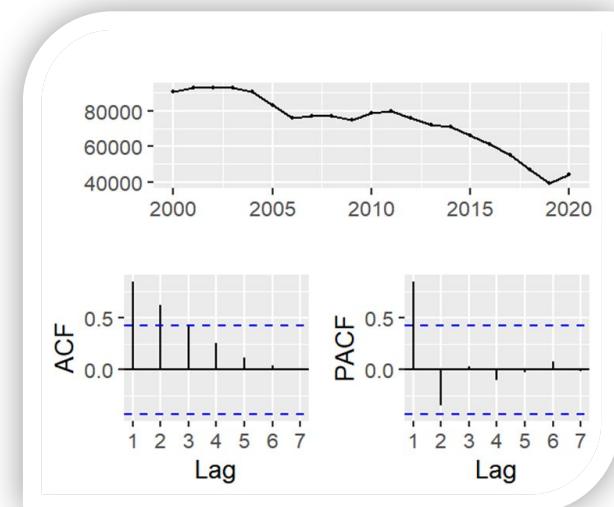
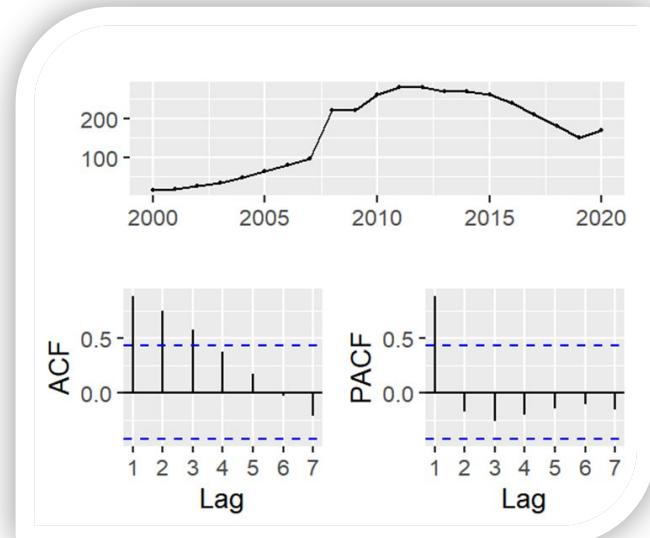
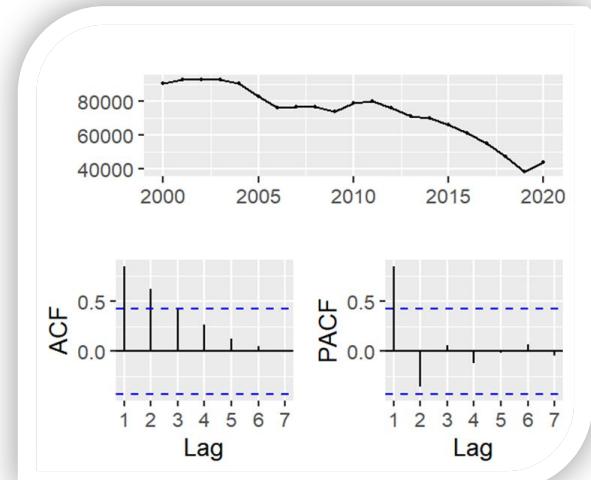
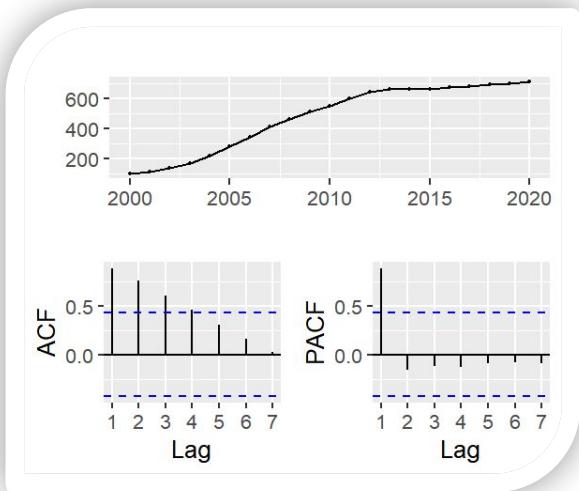
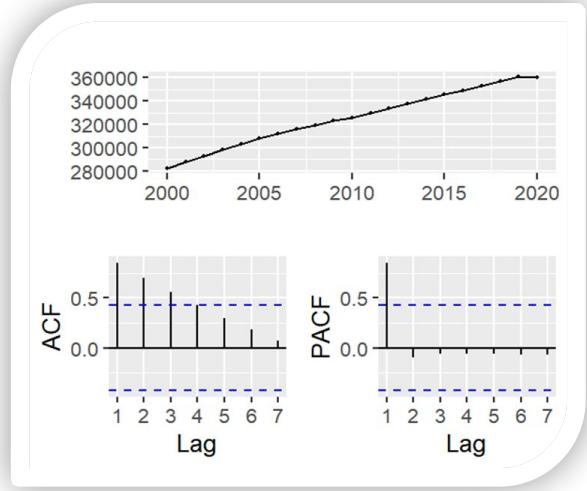
Interpretation

- Time profile shows the trend in the graph.
- ACF outside the threshold bound then it is considered as significant.
- PACF outside the threshold bound then it is considered as significant.
- For Incidence number, Incidence of TB and HIV, Mortality number of TB and HIV model is fitted after differencing once.
- For mortality number of TB excluding HIV and mortality of TB, AR(1) model is fitted.

Bangladesh

- Time profile , ACF and PACF
- Time series model
- Prediction table
- Interpretation





Time series Models

Country	Estimates	Order			AIC	ME	RMSE	MAE
				q				
Bangladesh	Estimated Incident number	0	2	0	328.9	-344.7053	1256.395	642.4305
	Estimated TB and HIV incident number	0	2	0	148.81	-0.0981	10.9631	10.96317
	Estimated mortality number of TB cases excluding HIV	0	1	0	394.11	4.4452	4062.363	3385.398
	Estimated mortality number of TB and HIV cases	0	1	0	198.79	7.4292	32.3463	19.8101
	Estimated mortality number	0	1	1	390.17	-1188.769	3628.134	2933.527

Prediction Table

Bangladesh					
	2021	2022	2023	2024	2025
Estimated Incident number	359000	358000	357000	356000	355000
Estimated TB and HIV incident number	720	730	740	750	760
Estimated mortality number of TB cases excluding HIV	41650	39300	36950	34600	32250
Estimated mortality number of TB and HIV cases	170	170	170	170	170
Estimated mortality number	49462.55	49462.55	49462.55	49462.55	49462.55

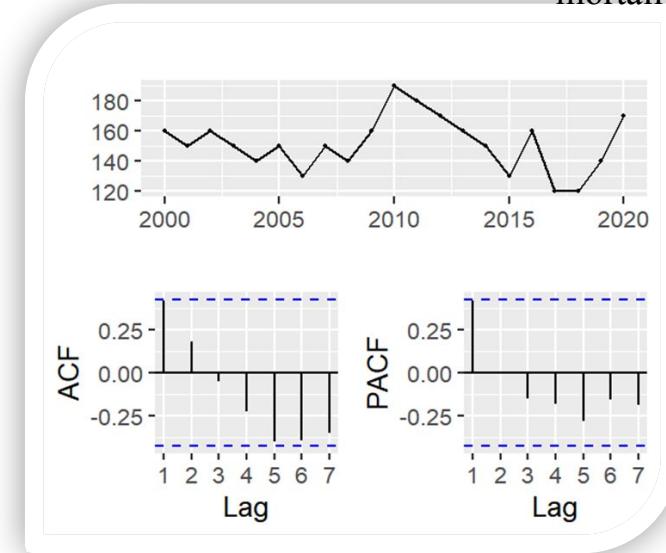
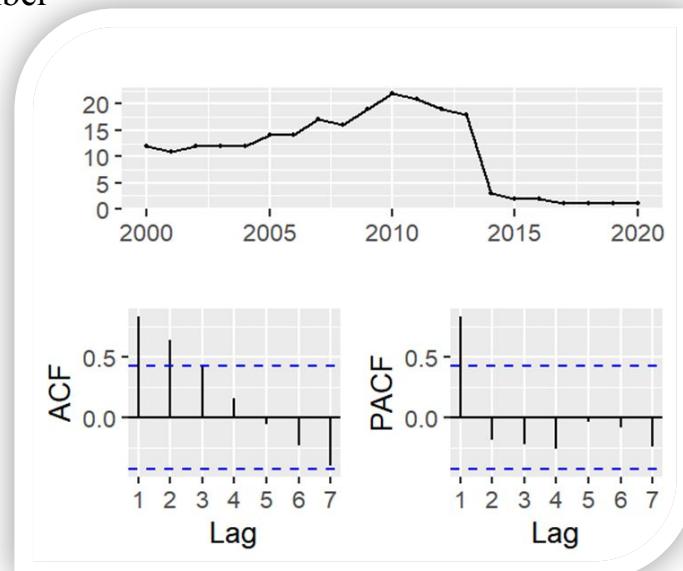
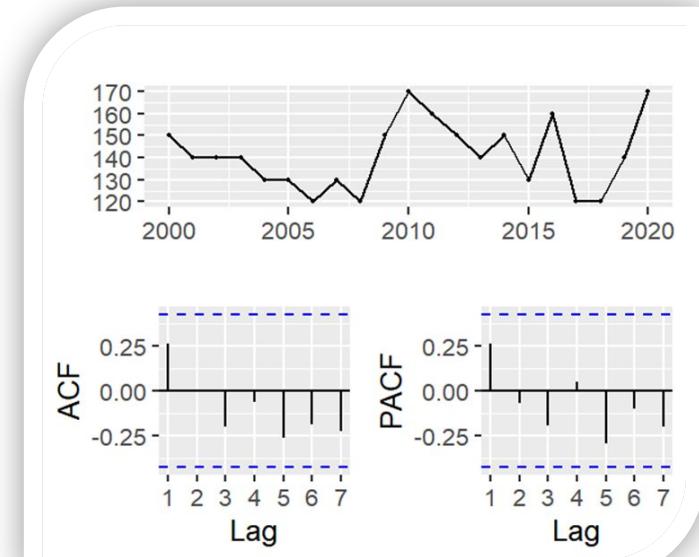
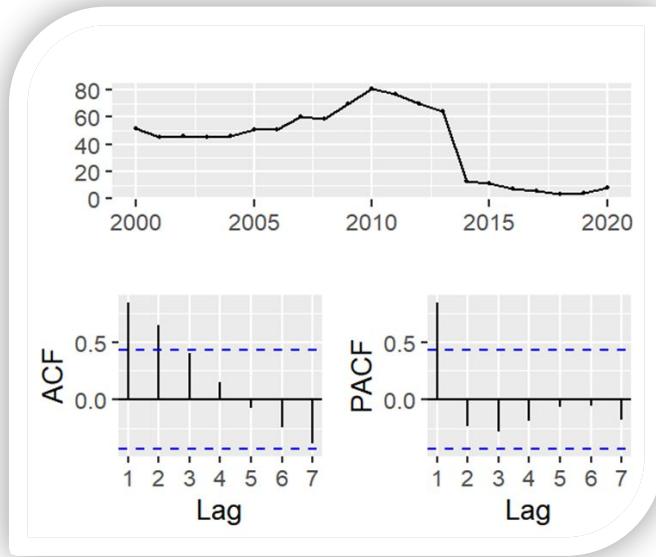
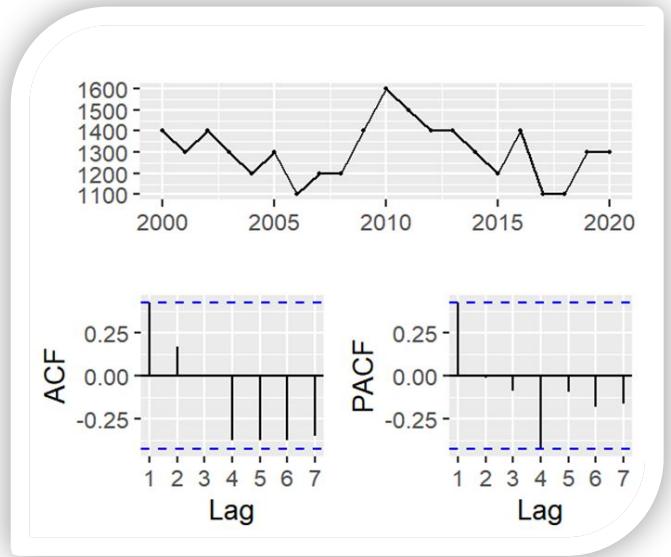
Interpretation

- Time profile shows the trend in the graph.
- ACF outside the threshold bound then it is considered as significant.
- PACF outside the threshold bound then it is considered as significant.
- For incidence number of TB and incidence number of TB confounding with HIV model is fitted after differencing twice.
- For mortality rate of TB & HIV, mortality rate of TB cases excluding HIV the model is fitted after differencing once.
- MA(1) model is fitted for mortality rate of TB after differencing once.

Bhutan

- Time profile , ACF and PACF
- Time series model
- Prediction table
- Interpretation





Time series Models

Country	Estimates	Order			AIC	ME	RMSE	MAE
				q				
Bhutan	Estimated Incident number	1	0	0	265.51	-2.2566	116.1938	93.6402
	Estimated TB and HIV incident number	0	1	0	159.95	-2.0928	12.24745	-23.56168
	Estimated mortality number of TB cases excluding HIV	0	0	0	178.44	0.000	15.4009	12.6077
	Estimated mortality number of TB and HIV cases	0	1	0	110.59	-0.5232	3.5657	1.6672
	Estimated mortality number	1	0	0	182.34	-0.19175	16.0337	12.78799

Prediction Table

Bhutan					
	2021	2022	2023	2024	2025
Estimated Incident number	1304.453	1306.327	1307.115	1307.447	1307.586
Estimated TB and HIV incident number	8	8	8	8	8
Estimated mortality number of TB cases excluding HIV	140.9524	140.9524	140.9524	140.9524	140.9524
Estimated mortality number of TB and HIV cases	1	1	1	1	1
Estimated mortality number	159.9121	155.5849	153.7287	152.9325	152.5910

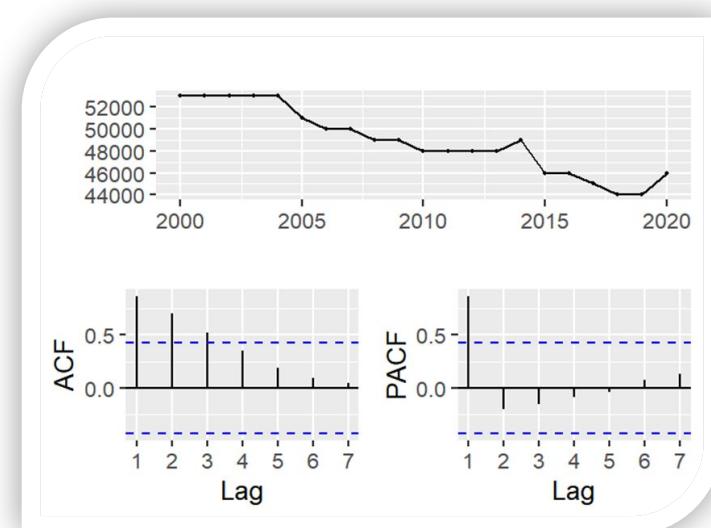
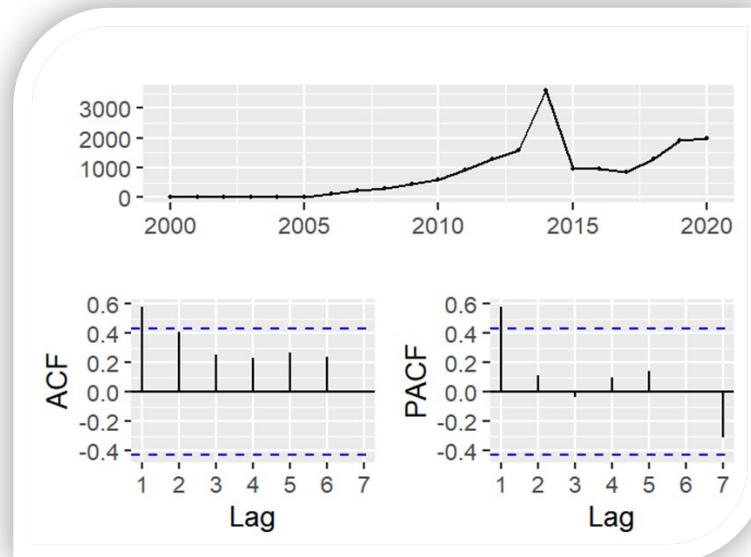
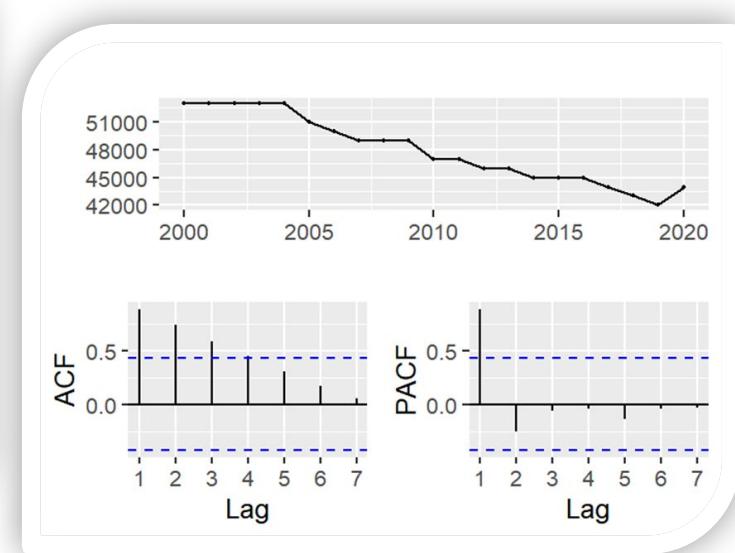
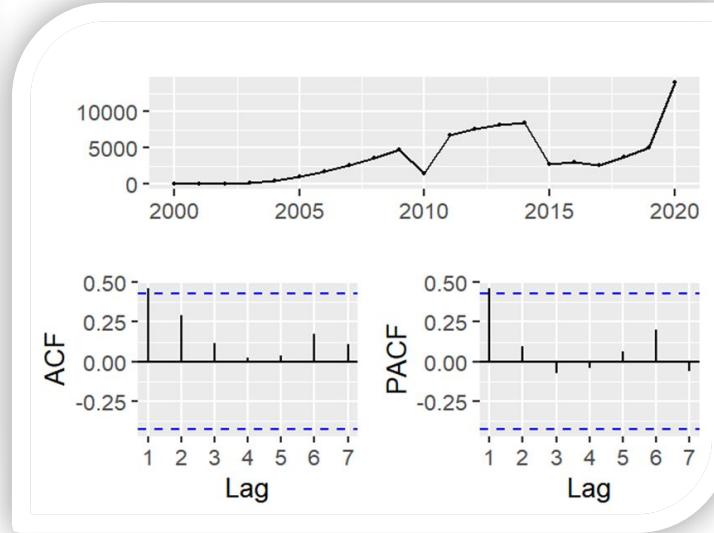
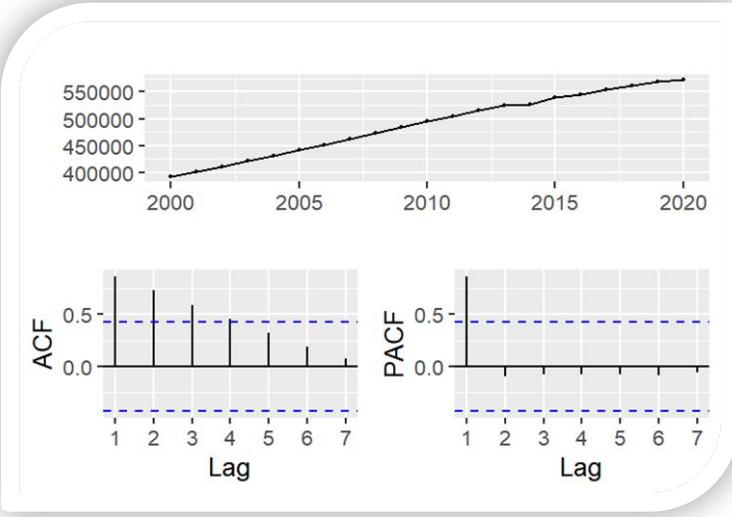
Interpretation

- Time profile shows the trend in the graph.
- ACF outside the threshold bound then it is considered as significant.
- PACF outside the threshold bound then it is considered as significant.
- AR(1) model is fitted for incidence number and mortality number of TB.
- For incidence number of TB&HIV and mortality number of TB&HIV model is fitted after differencing once.

Pakistan

- Time profile , ACF and PACF
- Time series model
- Prediction table
- Interpretation





Time series Models

Country	Estimates	Order			AIC	ME	RMSE	MAE
				q				
Pakistan	Estimated Incident number	0	2	1	359.1	-604.5552	2578.372	1266.636
	Estimated TB and HIV incident number	0	1	0	376.45	660.9581	2746.015	1558.101
	Estimated mortality number of TB cases excluding HIV	0	1	0	331.25	2.545237	843.8251	664.45
	Estimated mortality number of TB and HIV cases	0	1	0	324.6	94.95267	751.2351	357.4289
	Estimated mortality number	0	1	0	337.86	-330.8095	1046.6	621.5714

Prediction Table

Pakistan					
	2021	2022	2023	2024	2025
Estimated Incident number	579926.9	586853.9	593780.8	600707.8	607634.7
Estimated TB and HIV incident number	14000	14000	14000	14000	14000
Estimated mortality number of TB cases excluding HIV	43550	43100	42650	42200	41750
Estimated mortality number of TB and HIV cases	2000	2000	2000	2000	2000
Estimated mortality number	46000	46000	46000	46000	46000

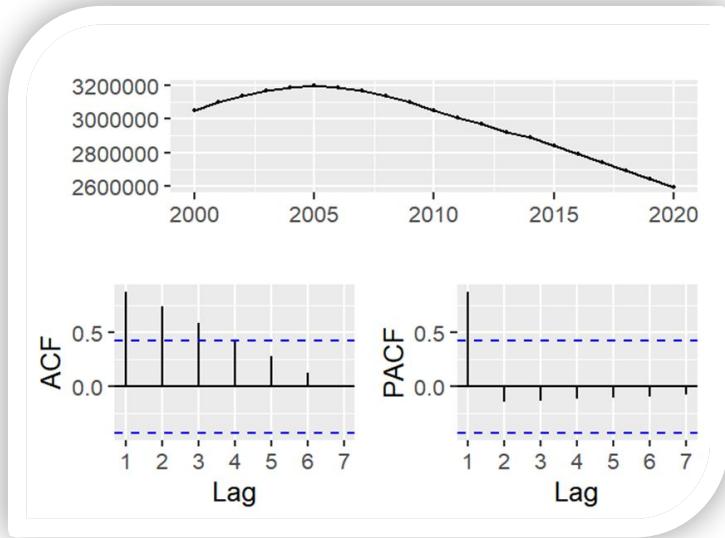
Interpretation

- Time profile shows the trend in the graph.
- ACF outside the threshold bound then it is considered as significant.
- PACF outside the threshold bound then it is considered as significant.
- MA(1) model is fitted for incidence number of TB after differencing twice.
- For Incidence of TB and HIV, mortality number of TB cases excluding HIV, mortality number of TB and HIV cases, mortality number of TB model is fitted for differencing once.

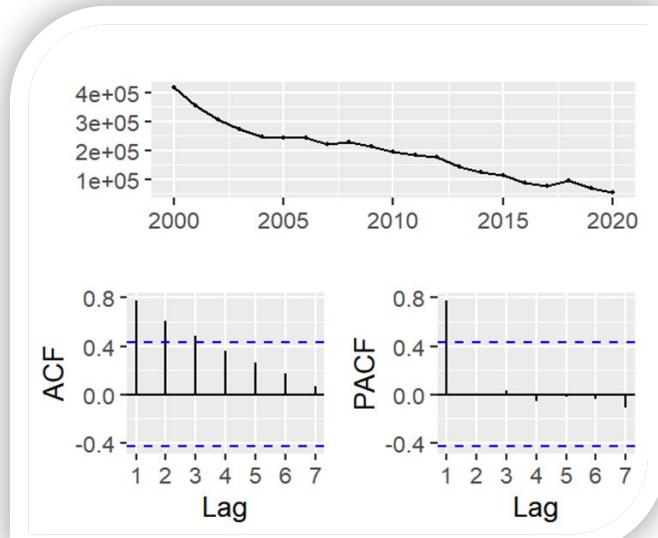
India

- Time profile , ACF and PACF
- Time series model
- Prediction table
- Interpretation

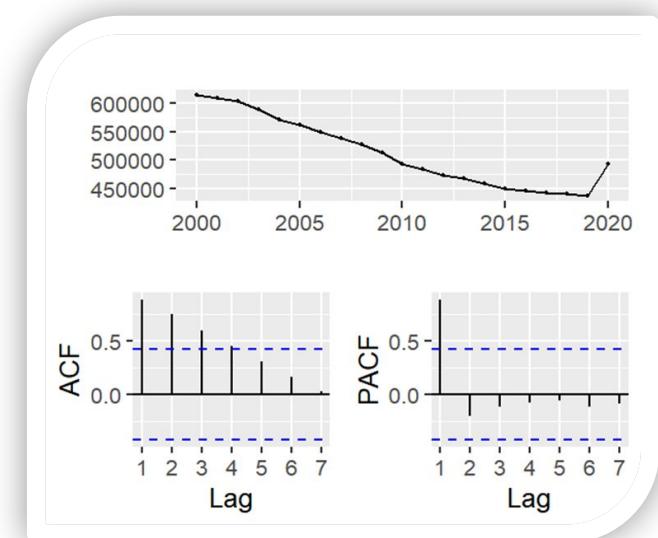




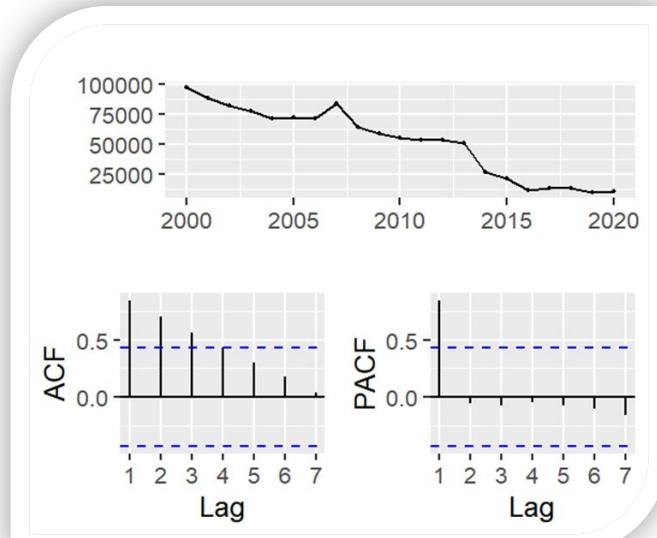
Time profile ,ACF ,PACF of
Incidence number



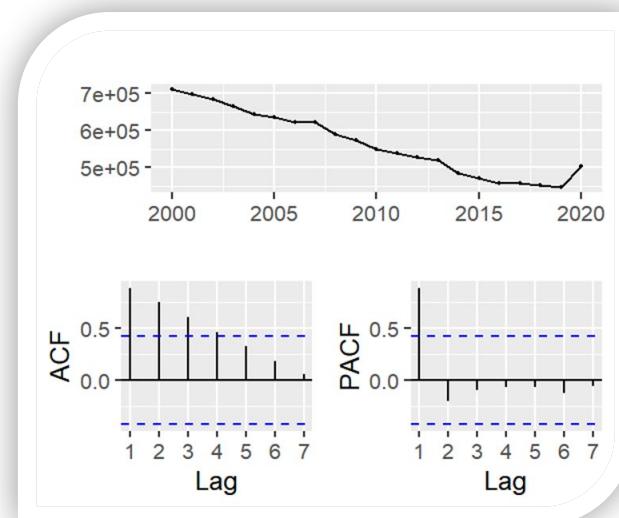
Time profile ,ACF, PACF of
Incidence with HIV number



Time profile ,ACF, PACF of
mortality with HIV number



Time profile ,ACF, PACF of
mortality without HIV number



Time profile, ACF, PACF of
mortality number

Time series Models

Country	Estimates	Order			AIC	ME	RMSE	MAE
				q				
India	Estimated Incident number	0	2	0	408.7	-4886.485	10276.42	7873.533
	Estimated TB and HIV incident number	1	1	0	450.99	1121.476	15933.26	12346.48
	Estimated mortality number of TB cases excluding HIV	0	2	0	420.6	2925.655	14001.32	5767.206
	Estimated mortality number of TB and HIV cases	0	1	0	417.87	4.823807	7357.245	4890.538
	Estimated mortality number	0	1	0	453.58	34.39998	17649.12	9767.733

Prediction Table

India					
	2021	2022	2023	2024	2025
Estimated Incident number	2540000	2490000	2440000	2390000	2340000
Estimated TB and HIV incident number	34897.218	15880.002	3534.873	23122.677	42785.683
Estimated mortality number of TB cases excluding HIV	550000	607000	664000	721000	778000
Estimated mortality number of TB and HIV cases	6700	2400	1900	6200	10500
Estimated mortality number	493600	483200	472800	462400	452000

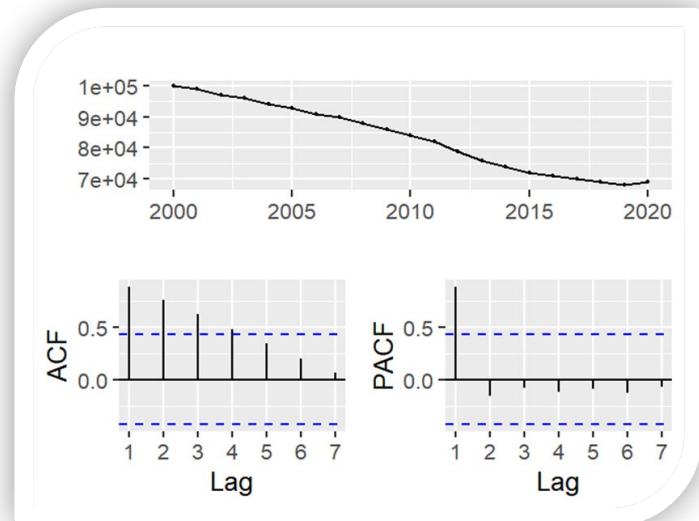
Interpretation

- Time profile shows the trend in the graph.
- ACF outside the threshold bound then it is considered as significant.
- PACF outside the threshold bound then it is considered as significant.
- For incidence number of TB , mortality number of TB excluding HIV model is fitted for differencing twice.
- For mortality number of TB & HIV and mortality number of TB model is fitting for differencing once.
- AR(1) model is fitted for incidence number of TB & HIV cases after differencing once.

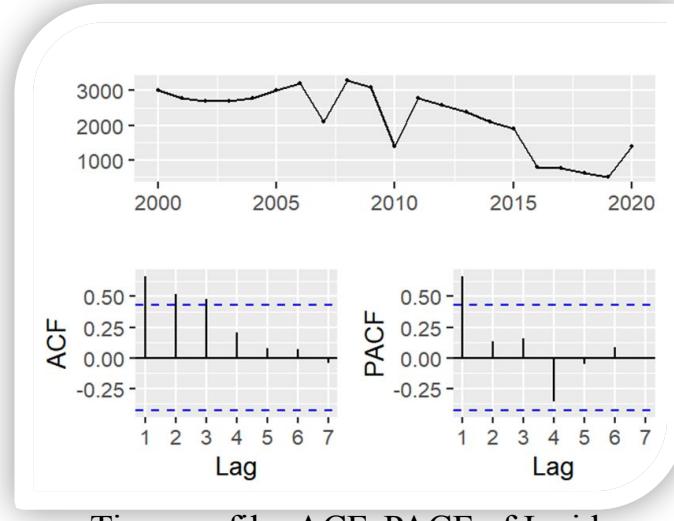
Nepal

- Time profile , ACF and PACF
- Time series model
- Prediction table
- Interpretation

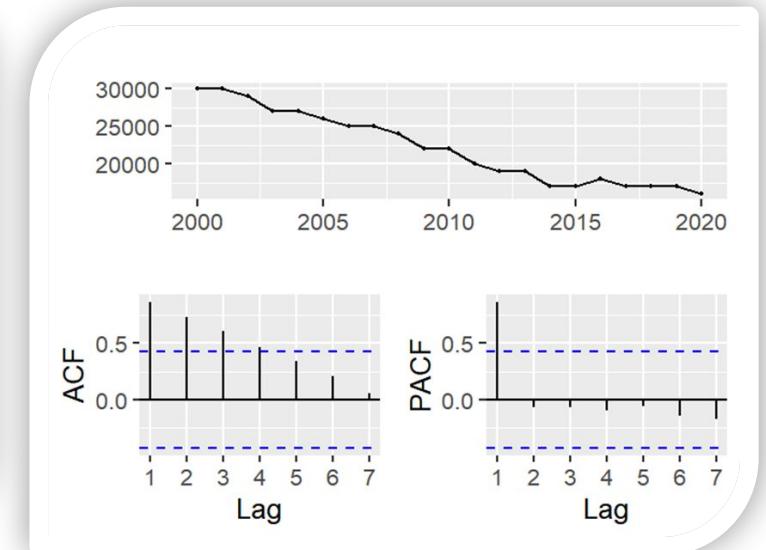




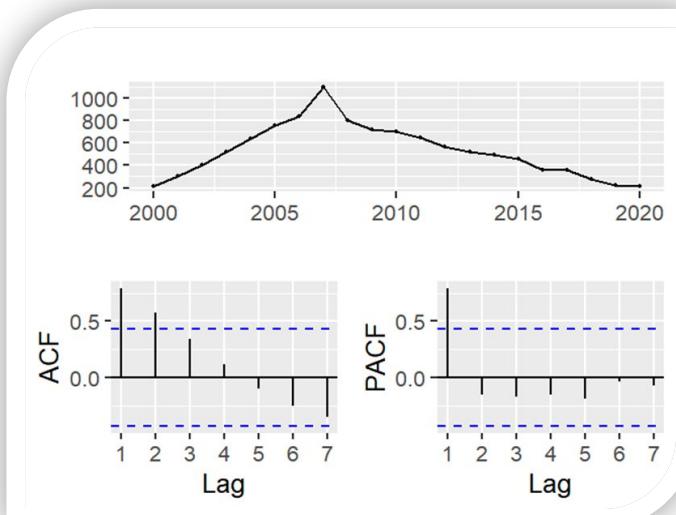
Time profile ,ACF ,PACF of
Incidence number



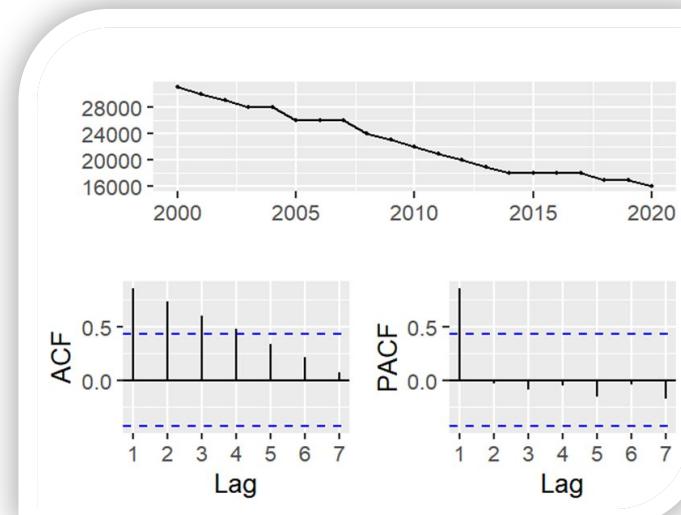
Time profile ,ACF, PACF of Incidence
with HIV number



Time profile ,ACF, PACF of
mortality with HIV number



Time profile ,ACF, PACF of
mortality without HIV number



Time profile, ACF, PACF of
mortality number

Time series Models

Country	Estimates	Order			AIC	ME	RMSE	MAE
				q				
Nepal	Estimated Incident number	2	1	0	327.3	-39.4421	695.6819	547.3757
	Estimated TB and HIV incident number	0	1	1	319.05	-168.4186	618.2626	456.324
	Estimated mortality number of TB cases excluding HIV	0	1	0	330.22	1.461904	822.3353	696.7
	Estimated mortality number of TB and HIV cases	1	0	0	264.78	21.98327	113.0481	83.08083
	Estimated mortality number	0	1	0	318.11	1.511904	607.5324	501.5119

Prediction Table

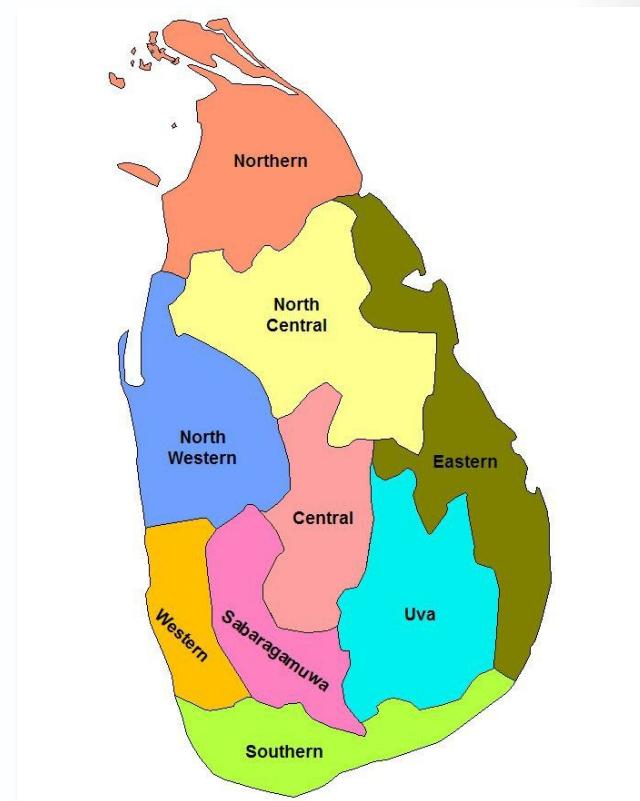
Nepal					
	2021	2022	2023	2024	2025
Estimated Incident number	68783.08	69275.33	69325.08	69622.43	69755.26
Estimated TB and HIV incident number	1051.281	1051.281	1051.281	1051.281	1051.281
Estimated mortality number of TB cases excluding HIV	15300	14600	13900	13200	12500
Estimated mortality number of TB and HIV cases	202.1984	194.6866	187.4539	180.4899	173.7846
Estimated mortality number	15250	14500	13750	13000	12250

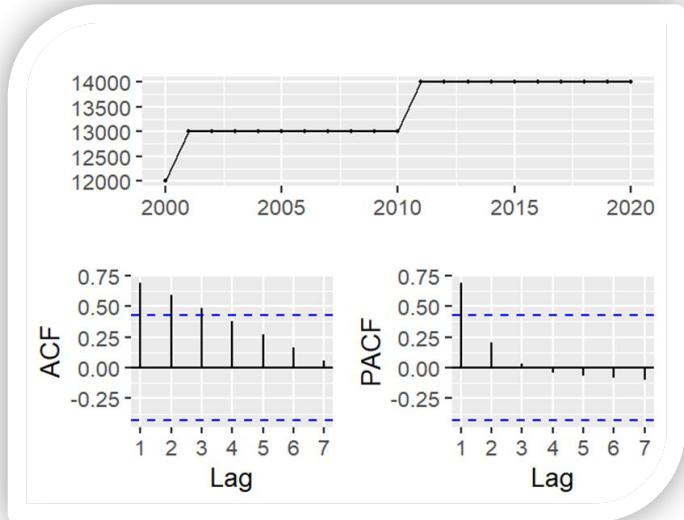
Interpretation

- Time profile shows the trend in the graph.
- ACF outside the threshold bound then it is considered as significant.
- PACF outside the threshold bound then it is considered as significant.
- AR(2) model is fitted for incidence number of TB after differencing once.
- For incidence number of TB & HIV MA(1) model is fitted after differencing once.
- AR(1) model is fitted for mortality number of TB&HIV cases.
- For mortality number of TB and incidence number of TB&HIV cases model is fitted after differencing once.

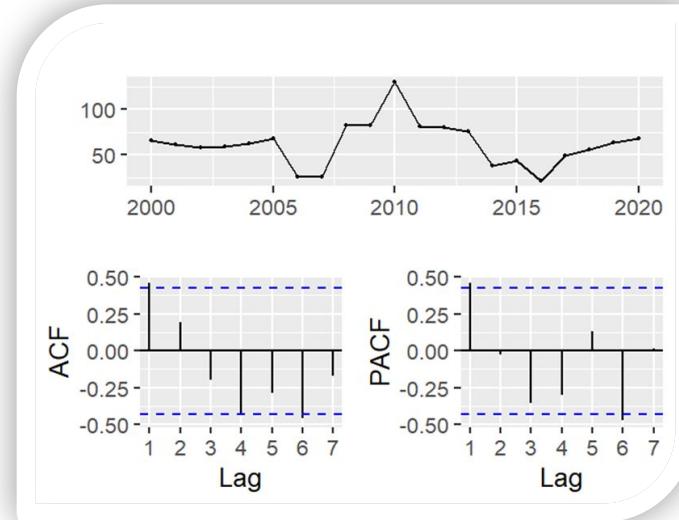
Sri Lanka

- Time profile , ACF and PACF
- Time series model
- Prediction table
- Interpretation

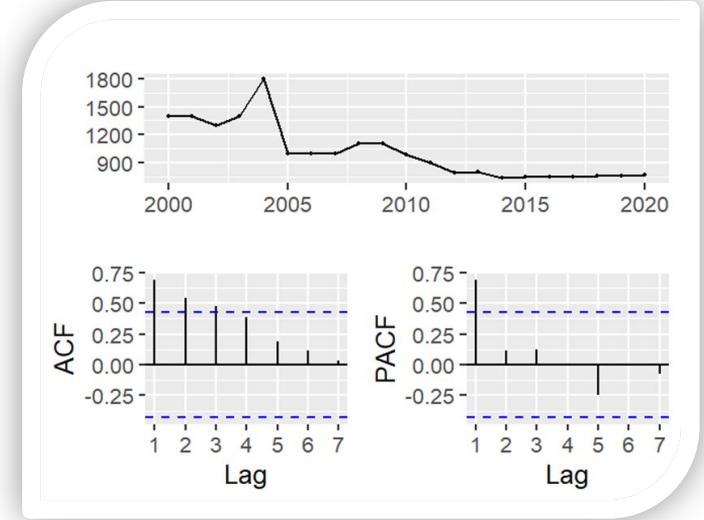




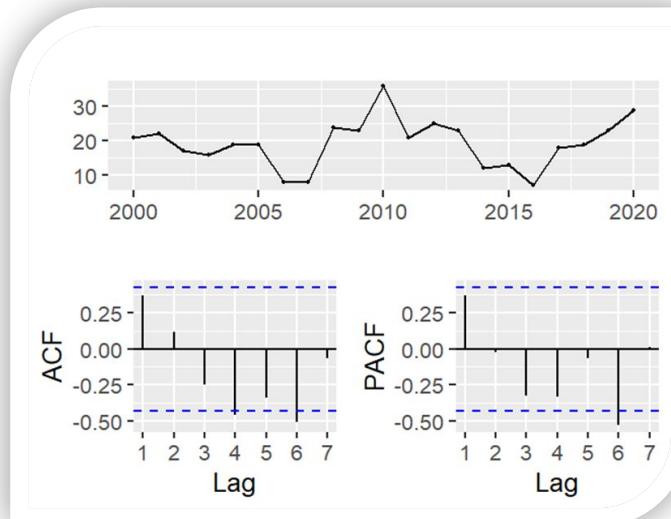
Time profile ,ACF ,PACF of
Incidence number



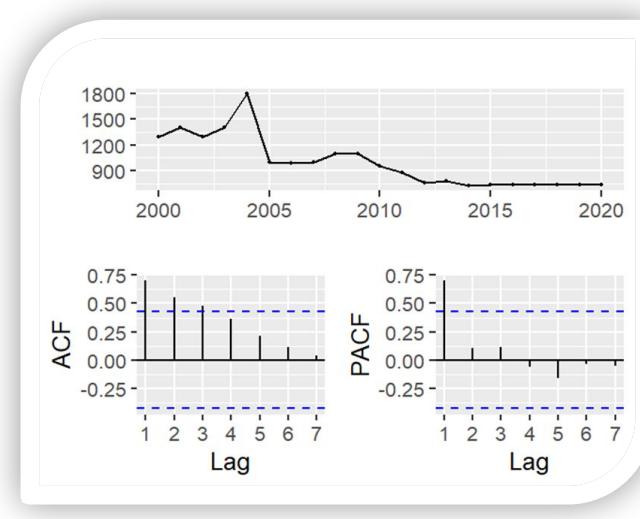
Time profile ,ACF, PACF of Incidence
with HIV number



Time profile ,ACF, PACF of
mortality with HIV number



Time profile ,ACF, PACF of
mortality without HIV number



Time profile, ACF, PACF of
mortality number

Time series Models

Country	Estimates	Order			AIC	ME	RMSE	MAE
				q				
Sri Lanka	Estimated Incident number	0	1	0	289.02	95.80952	308.6178	95.80952
	Estimated TB and HIV incident number	1	0	0	193.69	-0.10230	21.00134	14.06102
	Estimated mortality number of TB cases excluding HIV	0	1	0	272.74	-26.60476	205.4266	98.15714
	Estimated mortality number of TB and HIV cases	1	0	0	143.4	-0.033150	6.350372	4.928525
	Estimated mortality number	0	1	0	272.33	-29.93333	203.3414	91.97143

Prediction Table

Sri Lanka					
	2021	2022	2023	2024	2025
Estimated Incident number	14000	14000	14000	14000	14000
Estimated TB and HIV incident number	64.71543	63.25588	62.60730	62.31910	62.19103
Estimated mortality number of TB cases excluding HIV	740	740	740	740	740
Estimated mortality number of TB and HIV cases	23.24416	20.98563	20.09940	19.75166	19.61521
Estimated mortality number	770	770	770	770	770

Interpretation

- Time profile shows the trend in the graph.
- ACF outside the threshold bound then it is considered as significant.
- PACF outside the threshold bound then it is considered as significant.
- For incidence number of TB and mortality number of TB excluding HIV and mortality number of TB model is fitted for differencing once.
- For incidence and mortality number of TB &HIV AR(1) model is fitted.

Summary and Conclusion

- The time series is divided into test and train data and the autoregressive integrated moving average models are fitted for given data. The best model is selected using different criteria like AIC, ME, RMSE and MAE.
- The predicted annual incidence and mortality rates are the projected values under ideal conditions. These predictions can be compared with the actual incidence.
- The annual incidence number of TB, Incidence of TB and HIV, mortality number of TB cases excluding HIV, mortality number of TB and HIV cases, mortality number of TB of South Asian countries were predicted for 5 years from 2021 to 2025.

Summary and Conclusion...

- The incidence rate of TB cases in Afghanistan, Bangladesh , Pakistan, and Sri Lanka is increasing. There is a decrease in the trend in India and Nepal. There is variation in the incidence rate of Bhutan.
- The incidence rate and HIV cases in Bangladesh and Pakistan show an increasing rate , and Afghanistan, Nepal, and Sri Lanka show that there is variation in incidence and HIV rate. It is rapidly declining in India and Bhutan.
- The mortality rate of Afghanistan, Bangladesh, Bhutan, and Pakistan varies. In Nepal and Sri Lanka, the mortality rate is decreasing. But only in India is the mortality rate of TB increasing.

Summary and Conclusion...

- The mortality rate of TB and HIV varies in Afghanistan, Bangladesh, Bhutan, Pakistan, and India.
And it is decreasing in Nepal and Sri Lanka .

Scope for further study

- Study can be extended to all the countries.
- It can be conducted in State level.
- It is also useful for public health researchers to formulate the hypothesis which is specific to a particular country.

References

- Montgomery D.C, Jennings C.L,Kculachi M(2008). *Introduction Time Series Analysis and Forecasting*. John Wiley & Sons. New Jersey.
- Box G.E.P, Jenkins.G.M, Reinsel G.C, Ljung G.M(2008). *Time Series Analysis Forecasting and Control*. John Wiley & Sons. New Jersey.
- Sullivan A, Nathavitharana RR. Addressing TB-related mortality in adults living with HIV: a review of the challenges and potential solutions. Therapeutic Advances in Infectious Disease. 2022 Mar;9:20499361221084163.
- Kober L, Solovic I, Littva V, Siska V. Prediction of the Epidemiological Situation of Tuberculosis in Slovakia by 2040-Data Update. Iranian Journal of Public Health. 2021 Nov;50(11):2229.

References

- Carwile ME, Hochberg NS, Sinha P. Undernutrition is feeding the tuberculosis pandemic: a perspective. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*. 2022 Mar 17;100311.
- Aryee G, Kwarteng E, Essuman R, Nkansa Agyei A, Kudzawu S, Djagbletey R, Owusu Darkwa E, Forson A. Estimating the incidence of tuberculosis cases reported at a tertiary hospital in Ghana: a time series model approach. *BMC Public Health*. 2018 Dec;18(1):1-8.
- Zheng Y, Zhang X, Wang X, Wang K, Cui Y. Predictive study of tuberculosis incidence by time series method and Elman neural network in Kashgar, China. *BMJ open*. 2021 Jan 1;11(1):e041040.
- Liu K, Li T, Vongpradith A, Wang F, Peng Y, Wang W, Chai C, Chen S, Zhang Y, Zhou L, Chen X. Identification and prediction of tuberculosis in eastern China: analyses from 10-year population-based notification data in Zhejiang Province, China. *Scientific reports*. 2020 May 4;10(1):1-0.

THANK YOU

March 24th is world TB Day

