EXPERIENCE IN MODELING AND PREDICTING THE INCIDENCE OF COMMUNITY-ACQUIRED PNEUMONIA DURING THE COVID-19 PANDEMIC

E.N. Filatova, M.I. Popkova, N.N. Zaitseva, O.V. Utkin*

Nizhniy Novgorod Research Institute of Epidemiology and Microbiology named after Academician I.N. Blokhina, 71 Malaya Yamskaya St., Nizhny Novgorod, 603950, Russia.

* Corresponding author: filatova@nniiem.ru

Abstract. During the COVID-19 pandemic, the number of cases of community-acquired pneumonia (CAP) increased dramatically, which significantly changed the dynamics of its incidence time series (TS). Such changes overestimate the predicted values of the incidence of CAP and increase the forecast error. The purpose of this work was to evaluate methods for predicting the dynamics of CAP incidence during the COVID-19 pandemic. The CAP incidence data, registered within the time period from 2011 to 2022 was used. Two TS data were compiled, which did not include and included cases of CAP caused by COVID-19 in 2021-2022. TS data transformation was performed using outliers' deletion, seasonal decomposition, or X-13-ARIMA-SEATS techniques. Typical monthly dynamics calculation method and several adaptive regression models (ETS, SARIMA, decSARIMA) were used for CAP incidence modeling and forecasting. For CAP incidence TS data that excluded cases of COVID-19 pneumonia, all analyzed transformation techniques effectively smoothed out the outlier period making the TS data suitable for modeling using adaptive regression models. For CAP incidence TS data that included cases of COVID-19 pneumonia, the methods of TS decomposition turned out to be ineffective. An acceptable forecast error was obtained when using typical monthly dynamics model based on the TS data with deleted outliers.

Keywords: CAP, COVID-19, time series, forecast, STL, X-13-ARIMA-SEATS, ETS, SARIMA.

List of Abbreviations

ARIMA – autoregressive integrated moving average

CAP – community-acquired pneumonia

CAP TS – time series of CAP indices data excluding cases of pneumonia caused by COVID-19

CAP+COVID TS – time series of CAP indices data including cases of pneumonia caused by COVID-19

COVID-19 – coronavirus disease

decSARIMA – mixed model of incidence rate based on time series decomposition and SARIMA

ETS – error-trend-seasonality

SARIMA – seasonal autoregressive integrated moving average

SEATS – signal extraction and ARIMA time-series

STL – seasonal-trend decomposition using LOESS

TMD – typical monthly dynamics

TS – time-series

Introduction

Community-acquired pneumonia (CAP) is a widespread infectious disease, the causative agents of which include a wide range of microorganisms of bacterial, viral and fungal nature. CAP make a significant contribution to the structure of morbidity and mortality from infectious diseases throughout the world.

In 2020, the epidemiological situation regarding the incidence of CAP sharply worsened due to the emergence of the new coronavirus infection (COVID-19) pandemic. For example, in Russia in the period preceding the COVID-19 pandemic, the incidence rates of CAP varied from 315.1 $^{0}/_{0000}$ in 2011 to 518.9 $^{0}/_{0000}$ in 2019. In 2020, the incidence of CAP increased to 1856.2 $^{0}/_{0000}$ (3.6 times), mainly due to cases of disease caused by COVID-19 (COVID-19 pneumonia). At the same time, the number of deaths from CAP increased almost 12 times.

The seasonal factors influence the CAP incidence rate. There are certain differences in the seasonality of CAP depending on the etiology of the disease (Bobyleva *et al.*, 2018). Despite the fact that the increase of the CAP incidence generally occurs in the cold period of the year, the CAP seasonality, for example, caused by *Streptococcus pneumoniae* is limited to the autumn months (Botvinkin *et al.*, 2022) and *Mycoplasma pneumoniae* extends to the autumn-winter period (Rachina & Bobylev, 2016). The annual dynamics of CAP incidence rate caused by viral agents is closely related to the incidence of influenza and acute respiratory viral infections, so its maximum occurs at the end of winter – beginning of spring (Botvinkin *et al.*, 2022; Onishchenko *et al.*, 2013; Saltykova *et al.*, 2020). To date, most cases of CAP remain etiologically undeciphered (Bobyleva *et al.*, 2018; Yakovenko & Kravchenko, 2014). In general, the CAP incidence time series (TS) data, obtained from the federal statistical surveillance database, reflects the total effect of several etiological factors.

Despite the fact that when analyzing the total CAP incidence TS, it is often difficult to determine the true reasons underlying changes in the series (for example, a local rise or decline in incidence, changes in seasonality), identifying general patterns of the TS dynamics is a useful tool for medium- and long-term forecasting of CAP morbidity and planning sanitary and anti-epidemic measures.

The COVID-19 pandemic and the peculiarities of recording of CAP cases in federal statistical forms during this period have made significant changes to the CAP incidence TS. From a statistical point of view, the CAP incidence TS data for 2020–2021 contain numerous outliers – data points that stand out from the general sample. Such outliers overestimate the average values of the TS which affects the results of analysis and forecasting of morbidity. On the other hand, the spread of COVID-19 and the complex of sanitary and anti-epidemic measures carried out during a pandemic had an impact on the CAP incidence in general, and these changes must be taken into account when making medium- and long-term forecasts.

Thus, analysis and forecasting of the CAP incidence requires the usage of flexible statistical approaches that make it possible to reduce the influence of periods of epidemic without losing sensitivity to possible TS level changes. Widely used methods of TS analysis and modeling, as a rule, solve only one of two problems. Thus, simple averaging methods make it possible to identify and eliminate outliers in a time series, but have low sensitivity to levels change. Adaptive regression models reproduce well the dynamics of the incidence TS, but are susceptible to the influence of outliers. A good result can be shown by a combination of several methods using preliminary transformation of a TS data (Kondratyev, 2013).

The purpose of this work was to evaluate methods for predicting the dynamics of CAP incidence during the COVID-19 pandemic.

Materials and Methods

Data source

The study was performed using data obtained in the Nizhny Novgorod region – a large region in central Russia with a population of 3,144,245 people (data as of 2022). Population characteristics of the region, such as birth rate, mortality and morbidity, as well as income levels are comparable to federal ones. The total number of registered CAP incidences for the period of 2011-2022 and the size of the permanent population of the region for the specified period were obtained from the database of the Federal State Statistics Service. Two monthly CAP incidence (0/0000) TS data were created for the period from 2011 to 2022 inclusive. The first TS did not include cases of COVID-19 pneumonia (CAP TS), and the second TS included cases of COVID-19 pneumonia (CAP+COVID TS) for 2021 and 2022. Both TS data included cases of COVID-19 pneumonia for 2020, since at that time they were recorded as CAP cases in statistical forms.

Software

Transformation of TS data, construction of models, forecasting and calculation of error rates was performed in the R software environment (version 4.2.1 (R Core Team, 2022)) using “stats” (R Core Team, 2022), “forecast” (Hyndman & Khandakar, 2008) and “seasonal” (Sax & Eddelbuettel, 2018) packages.

Time-series modeling

Each of the two TS data was used to build several types of models with preliminary data transformation. The following approaches were used for TS data decomposition:

1. Outliers deletion from the set of monthly incidence values. Outliers were defined as values above Q3 + 1.5*IQR or below Q1 - 1.5*IQR, where Q1, Q3 are the values of the first and third quartiles, IQR is the interquartile range. To preserve the integrity of TS deleted points were replaced by the values Q3 + 1.5*IQR or Q1 -1.5*IQR, respectively.
2. Seasonal-Trend decomposition using LOESS (STL) (Cleveland, 1990; Sanchez-Vazquez, 2012). After decomposition, the resulting trend and seasonal components or their sum were used for modeling.
3. X-13-ARIMA-SEATS1 decomposition (Sax & Eddelbuettel, 2018). After decomposition, the resulting trend and seasonal components or their sum were used for modeling.

Transformed TS dats was modeled using the following models and techniques:

1. Typical (averaged) monthly dynamics (TMD) (Slobodenyuk, 2015).
2. Exponential smoothing model with additive seasonality (additive Holt-Winters model, ETSad, ETS (A, A, A)) (Ke *et al.*, 2016; Kuan *et al.*, 2022).
3. Exponential smoothing model with multiplicative seasonality (Holt-Winters multiplicative model, ETSmult, ETS (A, A, M)) (Ke *et al.*, 2016; Liu *et al.*, 2020).
4. Seasonal multiplicative autoregressive model - integrated moving average (seasonal ARIMA, SARIMA) (Chen *et al.*, 2022; Kuan, 2022; Liu *et al.*, 2020; Tan *et al.*, 2022).
5. Mixed model of incidence rate based on time series decomposition and SARIMA methods (decSARIMA) (Filatova & Solntsev, 2019).

For building models 1-4 in cases of data transformation with STL and X-13-ARIMA-SEATS techniques, the sum of the trend and the seasonal components was used. For building model 5, the trend and the seasonal component were modeled separately and the resulting models were summed up. The selection of model coefficients was carried out by enumeration; the best model was selected based on the values of the Akaike criterion. The autocorrelation of the residuals was tested using the Ljung-Box Q test (Q test, did not perform with TMD model).

Forecasting quality

Transformed TS data for 2011-2021 inclusive were used as a training sample. The values of the obtained models were extrapolated for a lead period of 12 months for forecasting. The limits of the predictive interval for the TMD model were calculated as lower and upper confidence limits (Rachina & Bobylev, 2016), for other models - in accordance with the formula for calculating the forecast variance. To test the predictive ability of the constructed models, actual data on CAP incidence in 2022 was used. The predictive ability of the models was assessed using mean absolute error (MAE), mean absolute percentage error (MAPE) and mean absolute scaled error (MASE) (Hyndman & Koehler, 2006). In the latter case, the TMD model forecast based on the original not transformed TS values was used as a naive forecast.

Ethical approval

Anonymized data that is freely available was used in this work.

Results

We analyzed two CAP incidence TS data for the period from 2011 to 2022. For each TS data, 14 morbidity predictive models were built.

The CAP TS in 2021-2022, was characterized by a pronounced increase of the incidence rate from May 2020 to January 2021 inclusive (Fig. 1, supplementary materials). Subsequently, the CAP incidence rate decreased to values comparable to the period preceding the COVID-19 pandemic. The increase in CAP incidence rates within CAP+COVID TS data was more extended and covered the period from May 2020 to February 2022 inclusive.

As shown in the Figure 2 and supplementary materials, the data transformation eliminated the overestimation of CAP incidence rates during the COVID-19 pandemic for both CAP TS and CAP+COVID TS. Outliers deletion technique reduced the CAP incidence rates for

2020-2021 to Q3 + 1.5*IQR. TS decomposition methods included increased CAP incidence values into the random component, so those were excluded from further analysis. However, for CAP+COVID TS data the lasting outliers period was partially regarded by decomposition algorithms as a level change, so the increased CAP incidence values constituted both trend and random component. Thus, the decomposition methods turned out to be effective only for CAP TS data.

It should be noted that TS decomposition methods allowed to identify the contribution of the seasonal factor into the CAP incidence rate. For both CAP TS and CAP+COVID TS data decomposition methods demonstrated a positive contribution of the seasonal factor into the CAP incidence within the periods from January to April and from October to December, as well as a negative contribution within the period from May to September (Fig. 2, supplementary materials).

Based on the analysis of residuals autocorrelation and forecast errors of CAP TS data models we selected three best forecasting approaches: outliers deletion + multiplicative ETS, X-13_ARIMA-SEATS decomposition + multiplicative ETS and outliers deletion + SARIMA (Table 1, Fig. 3, supplementary materials). Despite the autocorrelation of the residuals, all models of CAP TS data demonstrated a smaller forecast error rates compared to the naive forecast. In all cases, the forecasting error was associated with an overprediction of the CAP incidence, while periods of increase and decrease of the morbidity were predicted adequately (supplementary materials).

Forecast errors for CAP+COVID TS data, on the contrary, had extremely high values (Table 2, supplementary materials) and predicted morbidity values were overestimated tenfold. The only approach that made it possible to obtain a forecast with an acceptable error level was outliers deletion + TMD. Note that this approach, among all those tested, obtained a smaller error rate compared to the naive forecast (Table 2, Fig. 3).

Table 1

Characteristics of predictive models of CAP indices excluding COVID-19 cases

Data transformation	**Model**	**Model residual independence (Q*, p)**	**Forecast error**		
			MAE, 0/0000	MAPE, %	MASE
Outliers deletion	TMD	NT	25.21	135.88	0.76
Outliers deletion	ETSad	Q* = 13.15 p = 0.358	20.00	107.38	0.60
Outliers deletion	ETSmult	Q* = 9.07 p = 0.697	10.44	48.73	0.31
Outliers deletion	ARIMA (3, 1, 0) (1, 1, 2) (Liu *et al.*, 2020)	Q* = 9.58 p = 0.653	16.26	82.98	0.49
STL	TMD	NT	25.36	143.04	0.76
STL	ETSad	Q* = 228.24 p < 0.001	15.36	78.54	0.46
STL	ETSmult	Q* = 130.03 p < 0.001	20.96	116.95	0.63
STL	ARIMA (1, 1, 4) (1, 1, 2) (Liu *et al.*, 2020)	Q* = 5.7 p = 0.930	18.42	99.27	0.55

End of table 1

Data transformation	**Model**	**Model residual independence (Q*, p)**	**Forecast error**		
			MAE, 0/0000	MAPE, %	MASE
X-13-ARIMA-SEATS	TMD	NT	27.23	148.40	0.82
X-13-ARIMA-SEATS	ETSad	Q* = 55.97 p < 0.001	15.33	75.79	0.46
X-13-ARIMA-SEATS	ETSmult	Q* = 16.00 p = 0.191	15.46	83.62	0.46
X-13-ARIMA-SEATS	ARIMA (1, 1, 0) (1, 0, 0) (Liu *et al.*, 2020)	Q* = 39.08 p < 0.001	16.27	83.21	0.49
STL	decSARIMA T: ARIMA (1, 1, 0) S: ARIMA (0, 1, 0)	Q* = 24.48 p = 0.017	17.21	91.27	0.52
X-13-ARIMA-SEATS	decSARIMA T: ARIMA (2, 1, 0) (2, 0, 0) (Liu *et al.*, 2020) S: ARIMA (1, 1, 0) (2, 1, 0) (Liu *et al.*, 2020)	Q* = 37.03 p < 0.001	18.87	102.80	0.57

Note: NT – not tested

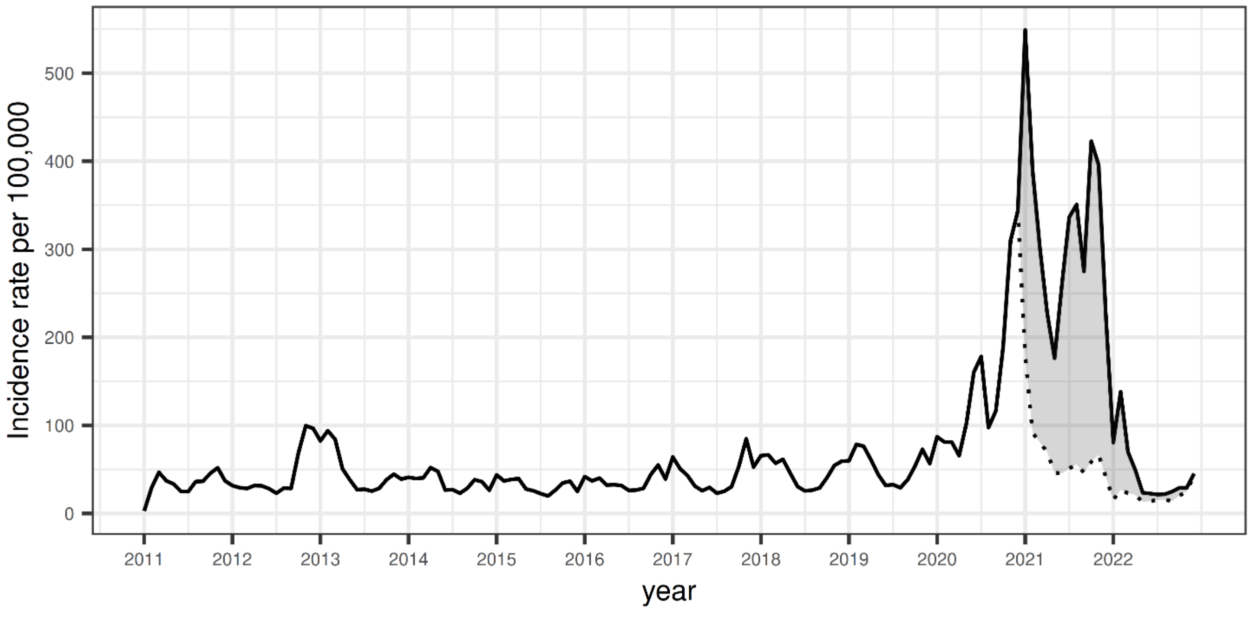

Fig. 1. Time series of CAP incidence for 2011–2022
The solid line indicates the incidence of CAP, including cases of COVID-19 pneumonia in 2021–2022. The dotted line indicates the incidence of CAP, excluding cases of COVID-19 pneumonia in 2021–2022. Gray color indicates cases of COVID-19 pneumonia in 2021–2022

Table 2

Characteristics of predictive models of CAP indices including COVID-19 cases

Data transformation	**Model**	**Model residual independence (Q*, p)**	**Forecast error**		
			MAE, 0/0000	MAPE, %	MASE
Outliers deletion	TMD	NT	21.05	51.39	0.53
Outliers deletion	ETSad	Q* = 22.07 p = 0.037	85.18	284.11	2.14
Outliers deletion	ETSmult	Q* = 2.74 p = 0.997	74.00	233.51	1.86
Outliers deletion	ARIMA (2, 1, 1) (1, 1, 0) (Liu *et al.*, 2020)	Q* = 6.10 p = 0.911	64.76	198.44	1.63
STL	TMD	NT	39.60	128.12	1.00
STL	ETSad	Q* = 280.36 p <0.001	252.49	805.30	6.36
STL	ETSmult	Q* = 184.94 p < 0.001	288.14	911.89	7.25
STL	ARIMA (1, 1, 0) (1, 1, 1) (Liu *et al.*, 2020)	Q* = 9.87 p = 0.628	235.74	748.72	5.93
X-13-ARIMA-SEATS	TMD	NT	40.93	134.98	1.03
X-13-ARIMA-SEATS	ETSad	Q* = 61.25 p < 0.001	424.22	1344.73	10.68
X-13-ARIMA-SEATS	ETSmult	Q* = 32.85 p = 0.001	386.39	1205.19	9.73
X-13-ARIMA-SEATS	ARIMA (3, 1, 0) (1, 0, 0) (Liu *et al.*, 2020)	Q* = 21.56 p = 0.043	364.53	1143.92	9.18
STL	decSARIMA T: ARIMA (1, 1, 0) (0, 0, 1) (Liu *et al.*, 2020) S: ARIMA (5, 0, 0)	Q* = 3.44 p = 0.922	232.08	737.54	5.84
X-13-ARIMA-SEATS	decSARIMA T: ARIMA (1, 1, 0) (1, 0, 0) (Liu *et al.*, 2020) S: ARIMA (2, 1, 1) (2, 1, 0) (Liu *et al.*, 2020)	Q* = 34.36 p < 0.001	357.51	1113.17	9.00

Note: NT – not tested

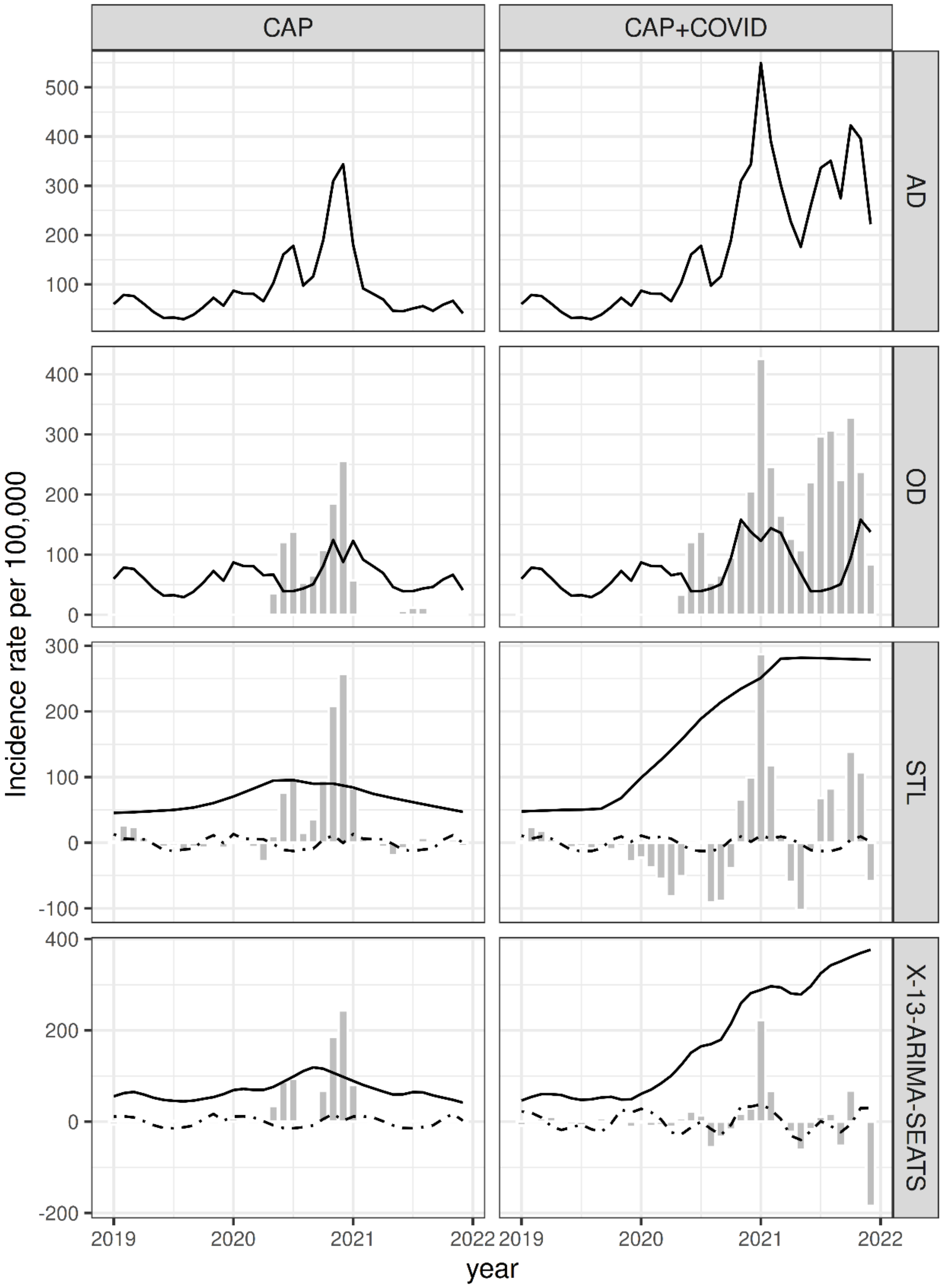

Fig. 2. Transformation of the CAP incidence time series for the period 2011–2021 The figure shows the time interval from 2019 to 2021 inclusive

CAP – the incidence of CAP, excluding cases of COVID-19 pneumonia in 2021–2022. CAP+COVID – the incidence of CAP, including cases of COVID-19 pneumonia in 2021–2022. AD – the actual data of the incidence of CAP. OD – transformation with outliers deletion. The solid line shows the original TS (AD), the result of the transformation (OD) or the trend (STL, X-13-ARIMA-SEATS). The dotted line indicates the seasonal component (STL, X-13-ARIMA-SEATS). Columns indicate transformation residuals (OD) or random component (STL, X-13-ARIMA-SEATS)

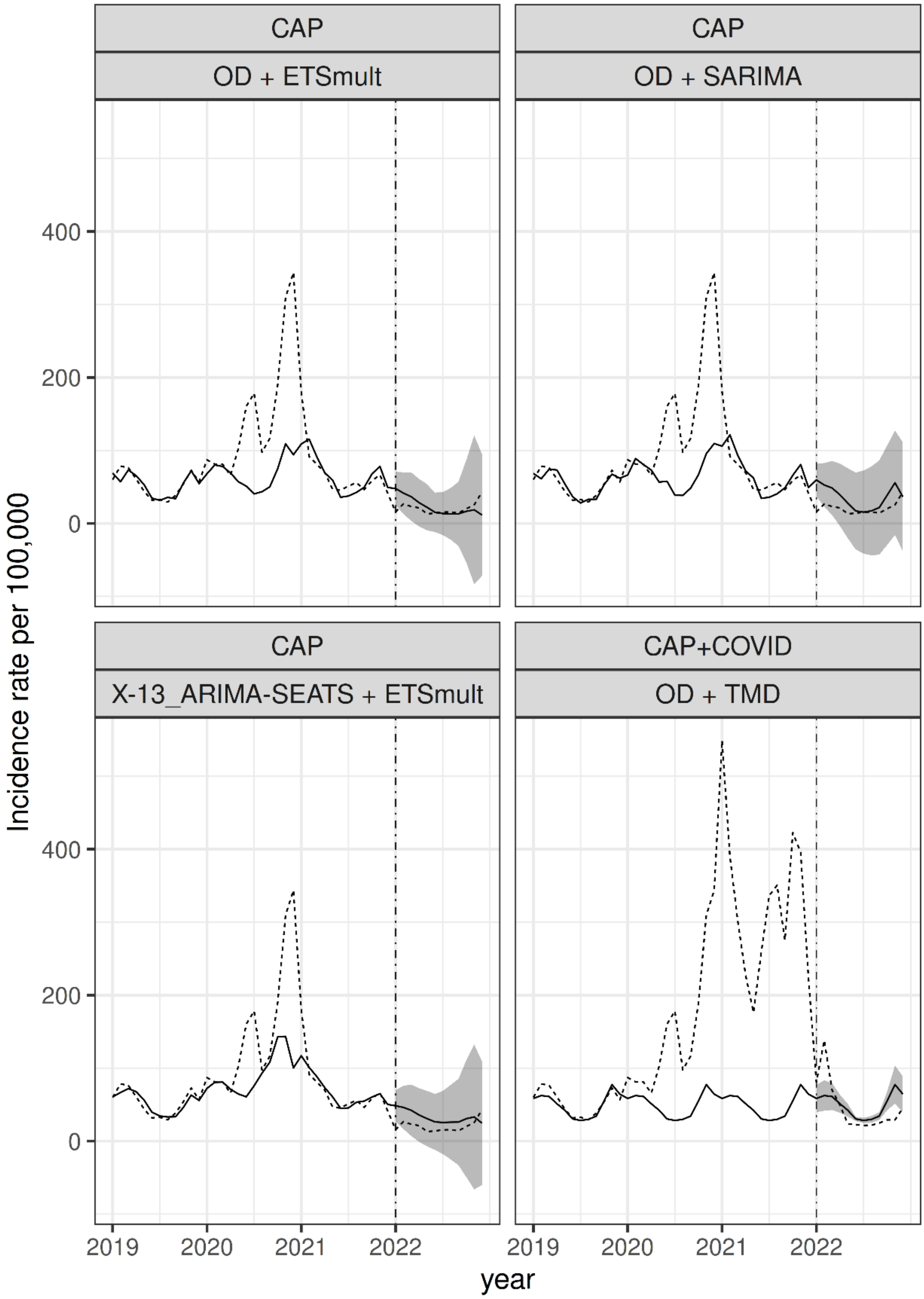

Fig. 3. Some models and forecast of CAP incidence. The figure shows the time interval from 2019 to 2021 (for actual incidence and model) and 2022 (for actual incidence and forecast). CAP – the incidence of CAP, excluding cases of COVID-19 pneumonia in 2021–2022. CAP+COVID – the incidence of CAP, including cases of COVID-19 pneumonia in 2021–2022. OD – transformation with outliers deletion. The solid curve indicates the model values and the prognosis of the incidence of CAP. The dotted curve indicates the actual incidence rates. The colored area represents the forecast 95% prediction interval. The vertical dotted line indicates the boundary point of the simulation (January 2022)

Discussion

The COVID-19 pandemic, which has been going on for more than three years, has had a negative impact on many statistical and analytical data, primarily medical, but also economic, demographic, logistics, social and others. The negative impact of the COVID-19 pandemic for statistical analysis was the formation of a set of outlier points located at the end of the TS data. The forecast for the period immediately following the outliers area often has big error rate due to the influence of latter points. Borderline outliers provide an extremely negative impact on the quality of the prognosis made with adaptive regression models (ETS, ARIMA, SARIMA), which are widely used for short-term and middle-term forecasting.

Regardless of whether COVID-19 pneumonia cases were excluded or included, the outlier periods in both analyzed TS data had three similar characteristics. Firstly, both periods had distinct start and end points, before and after which the CAP incidence values were comparable. Thus, the relation of two consecutive points was 1.64 at the beginning of the period of rising morbidity, 0.52 or 0.51 at the end of the period (for CAP TS or CAP+COVID TS data, respectively). Secondly, both TS data reflected natural CAP seasonality with an increase of the incidence rate in the autumn-winter period and a decline in the summer months, including the pandemic period. The third feature was the relatively high length of the outlier period. Thus, for CAP TS data the outlier period lasted for 7% (9/132) of the training period data points. For the CAP+COVID TS data the outlier period lasted for 15% (20/132) of the training data points and 17% (2/12) of the test data points.

The above indicates the necessity of treating the outlier period as a period of temporary level change and, therefore, of excluding this period from the analysis. There are three major ways to eliminate inflated values from a TS data: to eliminate the whole outlier period, to normalize the TS values (substitute outlier points only) or to decompose the TS with assigning outlier points values to the random component. Since the first method provide data loss, we only analyzed two remaining. Those demonstrated different results for CAP TS and CAT+COVID TS, which was due to TS features.

A major difference between the outliers period within CAP TS and CAP+COVID TS data was the proximity to the training and test data boundary point. In the case of the CAP TS data, the outliers period ended at the beginning of 2021, that is 11 points before training and test boundary point. Those data amount was sufficient for the outlier values to be classified by data transformation algorithms as “random” and excluded from further analysis. In our study, data transformation with outliers deletion technique gave predictive models with the least error rate, although all analyzed transformation techniques could be successfully applied in the case. Thus, the choice of data transformation method should be determined by the goals of the analysis: to make an accurate forecast (outliers deletion method) or to obtain maximum information about morbidity dynamics or its seasonality (TS decomposition methods). The transformed TS data is further suitable for building adaptive regression models, while the use of more complex mixed modeling techniques (such as decSARIMA) did not provide any advantages in forecast accuracy.

In the case of the CAP+COVID TS, the training and test border point was withing the outliers period, so all of the tested transformation algorithms based on TS data decomposition (STL and X-13-ARIMA-SEATS) defined the outliers period as the period of level change. This led to the inclusion of outlier points in the trend component, thereby retaining overestimated CAP incidence values in the data structure for further analysis, and subsequently provided a significant forecast error. In this situation, point-wise averaging techniques for both TS data transformation (outliers deletion) and TS modeling and forecasting (TPM) procedures had a significant advantage over adaptive regression models due to high resistance to TS level changes.

It should be noted that at the time of the study, we already had data on the CAP incidence rate in 2022 and knew about the coming period of steady decline of the intensity of the

COVID-19 epidemic process. In a real situation, if it is necessary to predict the level of morbidity at the time of epidemic challenge, it is advisable to analyze two situations. The first forecast with point-wise averaging techniques should model the normalization of the incidence rate during the lead-time period. The second forecast should model the situation when the changed incidence rate does not tend to normalization within the forecast period. In the latter case, the use of time series decomposition methods and adaptive regression models, on the contrary, will take into account the change in the incidence trend and will make it possible to create a more accurate prediction.

Conclusion

Using the CAP incidence data for 2011-2022, including the COVID-19 pandemic period, we examined the data transformation, modeling and forecasting techniques. Withing the time of transient epidemic challenge it is advisable to transform a data by replacing statistical outlier points by averaged values or using time series decomposition methods. In case of CAP TS data, that is when challenge period is fully withing training data, the most accurate forecast of CAP incidence was obtained by usage of the outlier deletion transformation technique and adaptive regression models (ETS, SARIMA). In case of CAP+COVID TS data, that is when challenge period is partially within training data, the most accurate forecast was obtained by usage of the outlier deletion transformation technique and TMD model. Results of this work would be useful for analyzing and forecasting morbidity based on time series with periods of atypical rise or decline in morbidity.

Acknowledgments

The study was carried out at the expense of funds received for financial support for the implementation of the Industry Research Program for 2021-2025. “Scientific support of epidemiological surveillance and sanitary protection of the territory of the Russian Federation. Creation of new technologies, means and methods for the control and prevention of infectious and parasitic diseases” (approved by the order of Rospotrebnadzor dated December 24, 2020 No. 869, clause 1.3.4.1, reg. No. 121091400195-7).

Supplementary materials: supporting information can be downloaded at: http://opera-medphys.org/sites/default/files/manuscripts/files/supplementary%20materials.xls.

References

BOBYLEVA E.S., GORBUNOV A.YU., STARODUBTSEVA O.I. & VAKHRUSHEV YA.M. (2018): The medic-statistical characteristics incidence of pneumonia in the Udmurt Republoc. *The Russian Archives of Internal Medicine* **8**(6), 438–443. (in Russ.)

BOTVINKIN A.D., KRAVCHENKO N.A., BAYANOVA T.A., KHAKIMOVA M.I., GAVRILOVA T.A. & LIKHANOVA N.A. (2022): Calculation of epidemic thresholds for incidence of community-acquired pneumonia. *Fundamental and clinical medicine* **7**(2), 45–55. (in Russ.)

CHEN Q., ZHAO H., QIU H., WANG QI., ZENG D. & YE M. (2022): Time series analysis of rubella incidence in Chongqing, China using SARIMA and BPNN mathematical models. *J. Infect. Dev. Ctries* **16**(8), 1343–1350.

CLEVELAND R.B., CLEVELAND W.S., MCRAE J.E. & TERPENNING I. (1990): STL: A seasonal-trend decomposotion procedure based on Loess. *J. Off. Stat.* **6**(1), 3–33.

FILATOVA E.N. & SOLNTSEV L.A. (2019): Mixed infectious disease forecasting technique, based upon seasonal decomposition and SARIMA. *Zh. Microbiol. (Moscow)* **3,** 76–83. (in Russ.)

GRIGOROVA E.V., NEMCHENKO U.M., IVANOVA E.I., TUNIK T.V., KUNGURTZEVA E.A., KOZLOVA L.S. & RYCHKOVA L.V. (2019): The seasonality of the bacteria Streptococcus pneumoniae in children with respiratory diseases. *Dal'nevostochnyy zhurnal infektsionnoy patologii* **36**, 69–73. (in Russ.)

HYNDMAN R.J. & KHANDAKAR Y. (2008): Automatic Time Series Forecasting: The forecast Package for R. *J. Stat. Softw.* **27**(3), 1–22.

HYNDMAN R.J. & KOEHLER A.B. (2006): Another look at measures of forecast accuracy. I*nt. J. Forecast* **22**(4), 679–688.

KE G., HU YA., HUANG X., PENG X., LEI M., HUANG CH., GU L., XIAN P. & YANG D. (2016): Epidemiological analysis of hemorrhagic fever with renal syndrome in China with the seasonal-trend decomposition method and the exponential smoothing model. *Sci. Rep.* **6**, 39350.

KONDRATYEV M. (2013): Forecasting methods and models of disease spread. *Computer Research and Modeling* **5**(5), 863–882. (in Russian)

KUAN M.-M. (2022): Applying SARIMA, ETS, and hybrid models for prediction of tuberculosis incidence rate in Taiwan. *PeerJ.* **10,** e13117.

LIU H., LI CH., SHAO YI., ZHANG X., ZHAI ZH., WANG X., HAO Y., WU Q. & JIAO M. (2020): Forecast of the trend in incidence of acute hemorrhagic conjunctivitis in China from 2011-2019 using the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ETS) models. *J. Infect. Public Health* **13**(2), 287–294.

ONISHCHENKO G.G., EZHLOVA E.B. & DEMINA YU. V. (2013): Epidemiological surveillance over community-acquired pneumonia as one of means for biological safety provision. *Problems of particularly dangerous infections* **4**, 24–27. (in Russ.)

RACHINA S.A. & BOBYLEV A.A. (2016): Atypical pathogens of community-acquired pneumonia: epidemiology, diagnosis and treatment. *Prakticheskaya pul'monologiya* **2**, 20–27. (in Russ.)

R CORE TEAM (2022): *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

SALTYKOVA T.S., ZHIGARLOVSKIY B.A., BRIKO N.I. & VYAZOVICHENKO YU.V. (2020): Epidemiological parallels of community-acquired pneumonia, influenza and ARVI in Moscow. *Tuberculosis and Lung Diseases* **98**(3), 6–12. (in Russ.)

SANCHEZ-VAZQUEZ M.J., NIELEN M., GUNN G.J. & LEWIS F.I. (2012): Using seasonal-trend decomposition based on loess (STL) to explore temporal patterns of pneumonic lesions in finishing pigs slaughtered in England, 2005-2011. *Prev. Vet. Med.* **104**(1-2), 65–73.

SAX C. & EDDELBUETTEL D. (2018): Seasonal Adjustment by X-13ARIMA-SEATS in *R. J. Stat. Softw.* **87**(11), 1–18.

SLOBODENYUK A.V., KOSOVA A.A. & AN R.N. (2015): Epidemiologicheskiy analiz: uchebnoye posobiye. Yekaterinburg: izd. GBOU VPO UGMU Minzdrava Rossii, 36 pp. (in Russ.)

TAN C.V., SINGH S., LAI CH.H., SYAHMI A., ZAMRI S.MD, DASS S.CH., ARIS T.B., IBRAHIM H.M., GILL B.S. (2022): Forecasting COVID-19 Case Trends Using SARIMA Models during the Third Wave of COVID-19 in Malaysia. *Int. J. Environ. Res. Public. Health* **19**(3), 1504.

YAKOVENKO O.N. & KRAVCHENKO N.A. (2014): Epidemiology features of community-acquired pneumonia. *Siberian medical journal (Irkutsk)* **125**(2), 8–11. (in Russ.)