Article

Short Time Series Forecasting: Recommended Methods and Techniques

Mariel Abigail Cruz-Nájera [ID], **Mayra Guadalupe Treviño-Berrones** *[ID], **Mirna Patricia Ponce-Flores** [ID], **Jesús David Terán-Villanueva** [ID], **José Antonio Castán-Rocha** *[ID], **Salvador Ibarra-Martínez** [ID], **Alejandro Santiago** [ID] **and Julio Laria-Menchaca** *[ID]

Departamento de Posgrado e Investigación, Facultad de Ingeniería, Universidad Autónoma de Tamaulipas, Tampico P.C. 89109, Mexico; marielcruznajera@gmail.com (M.A.C.-N.); mirna_poncef@hotmail.com (M.P.P.-F.); david_teran01@yahoo.com.mx or jdteran@docentes.uat.edu.mx (J.D.T.-V.); sibarram@docentes.uat.edu.mx (S.I.-M.); alx.santiago@gmail.com (A.S.)
* Correspondence: mgtrevino@docentes.uat.edu.mx (M.G.T.-B.); jacastan@docentes.uat.edu.mx (J.A.C.-R.); jlaria@docentes.uat.edu.mx (J.L.-M.)

Abstract: This paper tackles the problem of forecasting real-life crime. However, the recollected data only produced thirty-five short-sized crime time series for three urban areas. We present a comparative analysis of four simple and four machine-learning-based ensemble forecasting methods. Additionally, we propose five forecasting techniques that manage the seasonal component of the time series. Furthermore, we used the symmetric mean average percentage error and a Friedman test to compare the performance of the forecasting methods and proposed techniques. The results showed that simple moving average with seasonal removal techniques produce the best performance for these series. It is important to highlight that a high percentage of the time series has no auto-correlation and a high level of symmetry, which is deemed as white noise and, therefore, difficult to forecast.

Keywords: short-sized time-series; forecasting methods; seasonal extraction; forecasting techniques

1. Introduction

Keeping cities safe is a challenge for the police, especially when crime increases in all its forms. People live in fear and feel insecure. Factors such as the environment, social network gossip, interpersonal communication with neighbors, relatives, and lack of policies in an area can alter the insecurity perception [1].

People who have not been victims of crime can fear crime just as much as people who have suffered it, the latter people being affected psychologically by the perception of insecurity [2]. According to the statistics of the National Institute of Statistics and Geography (INEGI) [3] in 2020, women had a higher perception of insecurity with 72.7%, while men had scored 62% in insecurity perception. Additionally, from INEGI 2016, the most common crimes in México are extortion, vehicle theft, assault, burglary, and robbery on public roads and transportation. These crimes' frequencies decreased compared to 2020, caused by the SARS-CoV-2 pandemic (COVID-19), because people stayed in their homes, which caused an increase in extortion and domestic violence.

Delinquency is a problem prevalent in countries in Latin America because of their high level of violence. To measure the delinquency level, some countries gather information from police databases, and others such as Norway, Denmark, Sweden, and Finland carry out surveys to know the safety perception of their people when going out on the street, to implement strategies [4].

Because of delinquency, governments lose credibility, and societies suffering from this problem tend to refrain from voting, as they do not see government intervention to stop crime rising [5].

Education and poverty are some factors that influence delinquency, from a study conducted by Millán-Valenzuela et al. [6] in Italy and México. This study showed that

the areas with greater economic prosperity have the highest crime rates compared to some states where poverty is extreme. On the other hand, it is possible to assume that delinquency and education are closely related; commonly, young people who drop out or with low levels of education engage in crime. Hence, a country with a more educated society is less prone to experiencing problems such as poverty and delinquency. Therefore, education would be a good crime prevention measure.

Latin America has the world's highest crime rate, according to Díaz et al. [7]. Some helpful prevention strategies are: monitoring crime, diseases, and new criminal organizations; updating criminal sanctions; prison expansion; and making neighborhood meetings so that neighbors get to know each other and are more united.

To collect criminal information not kept in police databases, countries from central and eastern Europe, Africa, Latin America, and Asia conducted a series of surveys, asking people about their experiences with crime. These surveys were carried out in 2000 by the international study on criminal victims, supported by the united nations interregional institute.

As a result, more information was obtained than is available in police databases; this is because sometimes crimes are not reported, out of fear, shame, or because people know there is police mismanagement or corruption, which causes more crime. Furthermore, they found that criminal activity predominates in densely populated cities, i.e., because the city's services are not enough for their number of inhabitants. Additionally, the lack of work, education, and housing, among other factors, lead to crimes [8].

The prevention of criminal activity has been tackled through several predictions and forecasting methods, including regression models and machine learning.

This paper proposes a study forecasting crime with a short series of four crimes with eight forecasting methods applied to thirty-five small-sized real crime time series. Furthermore, we propose five forecasting techniques that use the seasonal component of the time series. Additionally, we compare the performance of classical and machine-learning-based ensemble methods and analyze the impact of the seasonal component of the time series in the forecasting techniques. Finally, we compared the rankings of the Average sMAPE and the Mean Rank of a Friedman test and found slight variations in the rankings of the forecasting methods and techniques. However, both maintain the same behavior.

The remainder of this paper is structured as follows: Section 2 shows the state of the art regarding prediction and crime forecasting. In Section 3, we describe the models and methods used to produce and process the short-sized time series, which includes data extraction and preparation. Section 4 shows the computational results of the experimentation configurations. Finally, Sections 5 and 6 show the discussions of the findings and the conclusions, respectively.

2. State of the Art of Crime Forecasting

Commonly, crimes are committed in groups because doing this makes them easier, making it necessary to identify the criminal gang based on criminal history, location, dates, and other variables considered essential for this study. Implementing techniques such as regressive models, neural networks, time series methods, and hybrid models has improved the forecasting or behavior prediction of these gangs or places that could be affected by crime.

Some studies using artificial intelligence have made successful predictions such as Ordóñez et al. [9], who used support vector machines. They applied a linear regression to predict the proximity of the values and obtain the proximity to the highest point of committed robberies in a specific year. They produced two models and compared their prediction accuracies using the squared error.

Cichosz [10] proposed a crime risk prediction focused on urban areas using logic regression, support vector machine, decision trees, and Random Forest. For this study, the author took information from a London database, focused on geographic data from Manchester, Liverpool, Bournemouth, and Wakefield, considering antisocial behavior,

the rate of sexual crimes, robbery with violence, shoplifting, and other types of robbery and theft.

Another tool used in crime prediction is deep neural networks. In [11], Chun et al. used this technique to predict the number of crimes using data from 1997 to 2017, considering the type of crime, severity level, and the number of faced charges. They used a criminal code, gender, age, race, and record date. This information successfully predicted the criminality level based on the record of each person for five years, labeling the crime from 0 to 3, depending on the severity. They identify that when the prediction finds a high-level crime, the probability of this crime is taken for granted for the next few years.

Wang et al. [12] used ARIMA with information from a London police database to carry out prediction research based on historical data from 2016 to 2018 for analysis and predictions using ARIMA with fourteen different types of crimes. Liu and Lu [13] used a hybrid model and STARMA model with LSTM for San Francisco's crime prediction using the United States historic data from 2003 to 2017. From this database, the authors used the number of incidents, description, day of the week, date, time, resolution, district, and location, *X*,*Y*. Their best model was a hybrid model which incorporates improvements to adapt to the data, adding a linear function and improving the accuracy to forecast the numbers of crimes committed for a specific hour, day, week, month, or year.

With criminal data from India, Jha et al. [14] made a comparison of ARIMA, neural networks, and machine learning. For this study, ARIMA models produce better results than the other methods.

Yadav and Kumari [15] used autoregressive models and fuzzy membership functions for multivariate crime predictions. They used a dataset from Delhi city in India for the following crimes: vehicular crime, arson, assaults, kidnappings, robberies, theft, and violent assault. They found that the fuzzy membership simplified multivariate crime prediction.

Shi et al. [16] conducted a study of an ARMAX model applied to US crimes related to age, the period of the committed crime, and similar characteristics regarding the offender. These characteristics help identify the trend of crime and the relationship with other planned crimes, among others presented in each case.

Melgarejo et al. [17] used two different databases, the first from San Francisco regarding burglary crime and the second from Colombia regarding cellphone theft. The authors used clustering algorithms to study the dynamics of criminal groups to produce time series. Additionally, they identified criminal gangs, minimizing the fuzzy partition index and generating fuzzy evolutionary predictors. Finally, they found that this clustering improved the time series to enhance crime predictions with non-linear analysis.

On the other hand, Izonin et al. [18] and Tkachenko et al. [19] tackled the problem of processing short datasets and missing values from a regression perspective. They used techniques such as support vector regression with kernels, adaptive boosting and stochastic gradient descent, neural networks, multilayer perceptron, and random forest. The authors have a medical dataset to predict calcium concentration in human urine and air pollution monitoring data, respectively.

3. Models and Methods

This section includes the data extraction, preparation and the experimentation methodology. All these processes were carried out using R.

3.1. Data Extraction

First, we built the time series using surveys of the frequency of each crime in several sectors of each city. These sectors respond to each city's socioeconomic, demographic, and geographic aspects. Then, we condensed this information producing thirty-five series ranging in size from fourteen to twenty-one observations.

3.2. Data Preparation

Additionally, we noticed that every series had an outlier value in the second element; therefore, we changed that value with the average value of the series.

Finally, we conducted some preliminary experiments and noticed that several series produce a symmetric mean average percentage error (sMAPE) of 200; this happened because of zero values in the series. Therefore, we increase every value of all the time series by one to avoid this issue. Figure 1 shows a time series example for each one of the four types of crime.
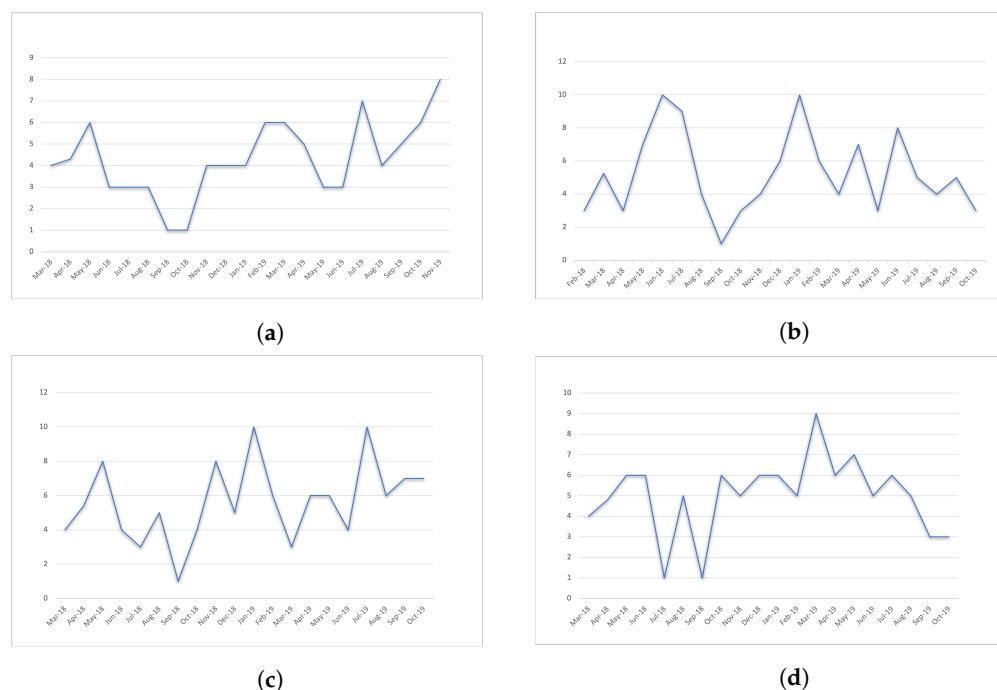

Figure 1. Time series example of the four types of crime considered in this paper. (**a**) shoplifting, (**b**) vehicular, (**c**) theft, (**d**) burglary.

Once processed, we split the time series into train and test. Here, we set a horizon equal to three ($h = 3$) and set it as test, leaving the rest as train (see Figure 2). The time series are available at https://mega.nz/folder/osFwxIia#vrnxaIFAJmSQ6_pPaFvjfQ accessed on 28 May 2022.
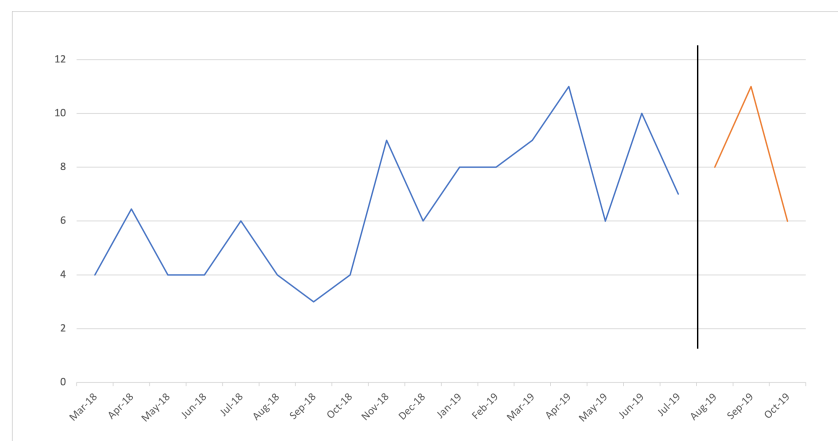

Figure 2. —: train. —: test.

3.3. Forecasting Methods

In this section, we show the methods used to forecast.

We use four simple forecasting methods: ARIMA, Holt–Winters (HW), Artificial Neural Networks (ANN), and Simple Moving Averages (SMA).

Additionally, we used four state-of-the-art forecasting methods: Jaganathan, which achieves the fourth place in the M4 competitions (Makridakis) [20]; FFORMA, which is at second place in M4; Hybrid, which is a simple hybrid forecasting method; and LightGBM, whose performance achieved first place in M5 competition [21].

These last four forecasting methods use the following components:

[Jaganathan] combines numerous statistical and machine learning methods [22]: naive/snaive; ExponenTial Smoothing (ETS); dampened ETS; bagged ETS; exponential smoothing, complex exponential smoothing, general exponential smoothing; multi-aggregation prediction algorithm (MAPA); temporal hierarchical forecasting; autoregressive integrated moving average (ARIMA); ThetaH; hybrid Theta; forecast pro; seasonal and trend decomposition using loess forecast; trigonometric Box–Cox transform, ARMA errors, trend and seasonal components (TBATS); double seasonal Holt-Winters; and multilayer perceptron and extreme learning machines.
[FFORMA] uses the following forecasting methods [23]: naive, random walk with drift, seasonal naive, theta method, automated ARIMA algorithm, ETS, TBATS, STLM-AR seasonal and trend decomposition, and neural network time series forecasts (NNETAR).
[Hybrid] ensembles [24]: auto ARIMA, ETS, Theta, NNETAR, seasonal and trend decomposition using loess, TBATS, snaive.
[LightGBM] [21] applies decision trees.

3.4. Proposed Forecasting Techniques

For every forecasting method, we propose five different forecasting techniques.

Re is the regular forecasting of the training section of the series, ARIMA, HW, ANN, SMA, Jaganathan, FFORMA, Hybrid, or LightGBM (see Figure 3); see Equation (1).

$$Re = Forecasting\_Method(series) \tag{1}$$

where the Forecasting_Method is any of the previously presented forecasting methods, and *series* is the training section of the time series.
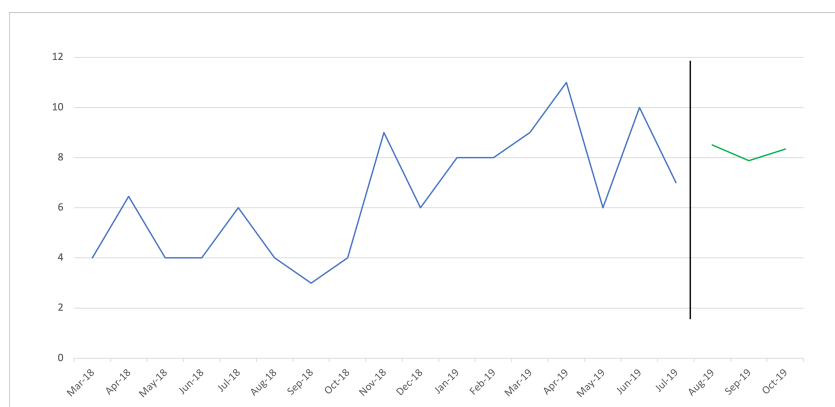

Figure 3. Regular Forecasting (Re): —: train. —: forecasting.

We made preliminary forecasting tests and found that several forecasting methods performed poorly. To improve the forecasting, we used a Fourier decomposition as suggested by Hyndman in [25–27] to decompose a time series in trend, seasonal, and residuals.

The following techniques use a Fourier decomposition to extract the seasonal component of the train section of a series; for simplicity, we will use the term seasonal component.

ReMS is the regular forecasting of the series minus the seasonal component. Figure 4 shows this technique's scheme, where Figure 4a shows the train section of the series, Figure 4b is the seasonal component and Figure 4c shows the forecasting of Figure 4a minus Figure 4b, see Equations (2)–(4).

$$seasonal = \text{FourierDecomposition}(series)[seasonal] \tag{2}$$

$$seriesMS = series - seasonal \tag{3}$$

$$\text{ReMS} = \text{Forecasting\_Method}(seriesMS) \tag{4}$$

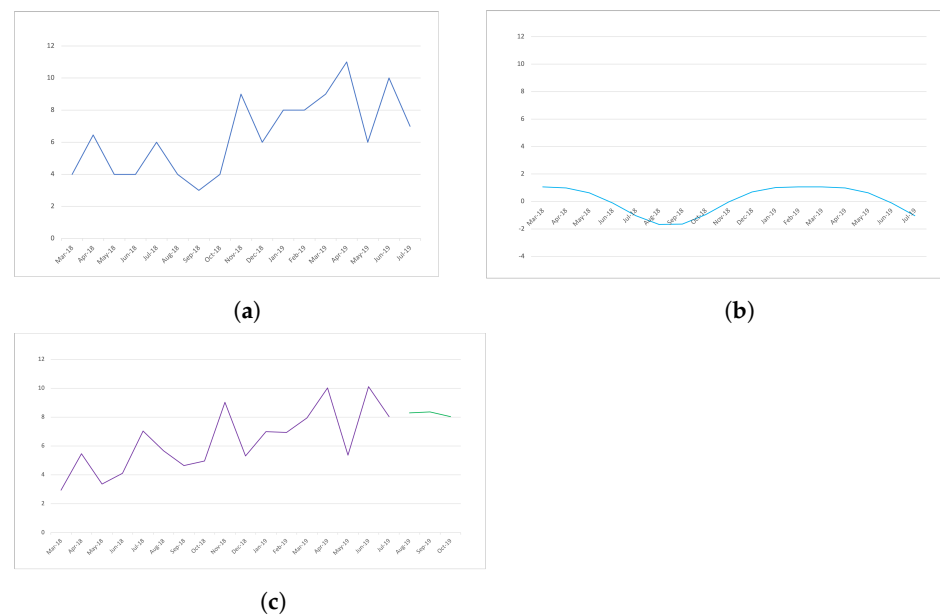where *seriesMS* is the training section of the time series minus the *seasonal* component.

(a)

(b)

(c)

Figure 4. The regular forecasting of the time series minus the seasonal component (ReMS). (**a**) — Train, (**b**) — Seasonal, (**c**) — (Train-Seasonal) — Forecasting ReMS.

The CAS technique is a composed forecasting of the ReMS plus the average of the whole seasonal component (see Figure 5). Here, we use the forecasting of ReMs (Figure 5a) and add the average of the seasonal component (Figure 5b) to produce the final forecasting (Figure 5b); see Equation (5).

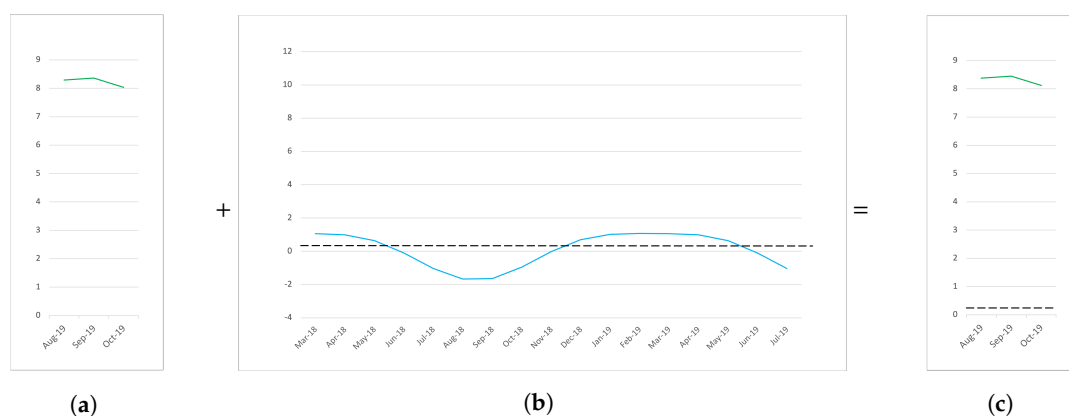$$\text{CAS} = \text{Forecasting\_Method}(seriesMS) + \text{Avg}(seasonal) \tag{5}$$

(a)

(b)

(c)

Figure 5. The ReMS plus the average of the seasonal component (CAS). (**a**) — Forecasting ReMS, (**b**) — Seasonal - - - Average Seasonal, (**c**) — Forecasting CAS.

CAHS is the same as CAS, but instead of using the average of the whole seasonal component, we use the average of the last h values of the seasonal component (see Figure 6). We propose this variation to highlight the latest tendencies of the seasonal component shown in Figure 6b, to increase or decrease the ReMS forecasting (Figure 6a) to produce the new CAHS forecast in Figure 6c; see Equation (6).

$$\text{CAHS} = \text{Forecasting\_Method}(seriesMS) + \text{Avg}(seasonal_{\text{last}h}) \tag{6}$$

where $seasonal_{\text{last}h}$ is a fragment of $seasonal$ that contains its last h elements.
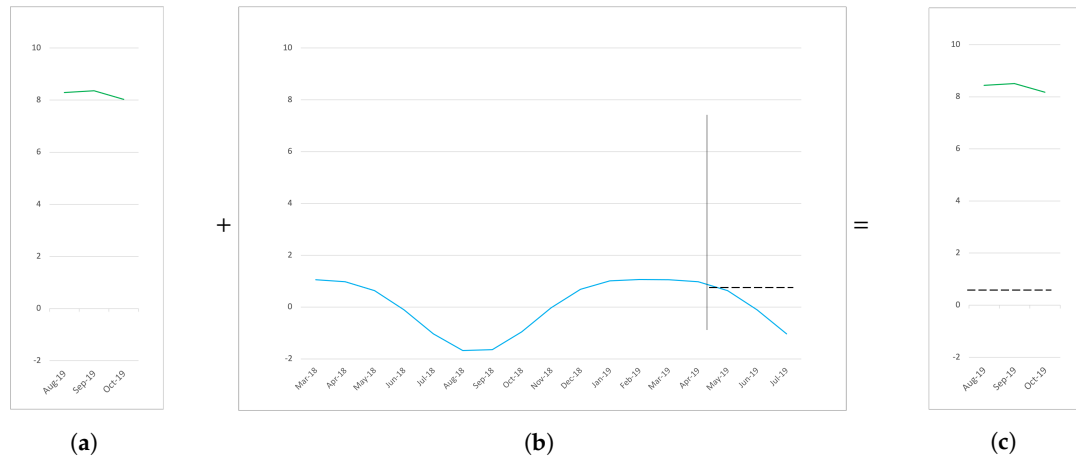

Figure 6. CAHS uses the same structure as CAS; however, it only uses the last h values instead of the whole seasonal component. (**a**) — Forecasting ReMS, (**b**) — Seasonal - - - - Average of last h = 0.1458, (**c**) — Forecasting CAHS.

Figure 7 shows the CReMSPlusFS scheme. This forecasting technique uses the ReMS (Figure 7a) and adds the independent forecasting of the seasonal component (Figure 7b), to yield a combined forecasting (Figure 7c). It is important to highlight that both forecastings are produced using the same forecasting method; see Equation (7).

$$\text{CReMSPlusFS} = \text{Forecasting\_Method}(seriesMS) + \text{Forecasting\_Method}(seasonal) \tag{7}$$

where the Forecasting_Method is the same for both components.

Therefore, we use eight forecasting methods and five forecasting techniques for each one, producing forty experimentation configurations.
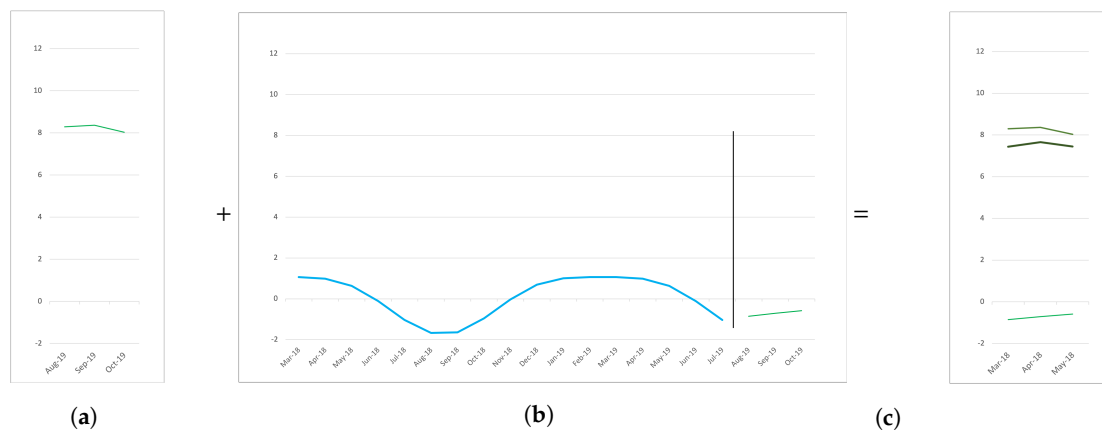

Figure 7. The ReMS plus the independent forecasting of the seasonal component (CReMSPlusFS). (**a**) — Forecasting ReMS, (**b**) — Seasonal — Seasonal Forecasting, (**c**) — Forecasting (CReMSPlusFS).

4. Experimental Result

In this section, we present the results of the forty experimentation configurations applying the symmetric mean average percentage error (sMAPE) used in M4 competitions [20] and the Mean Rank of the Friedman nonparametric test.

Table 1 shows the forecasting methods and techniques in columns one and two, which are ordered according to the Average sMAPE in column three. Here, we can see that the forecasting method that produces the best average sMAPE is ARIMA, followed by Jaganathan, SMA, and Hybrid, while the first three use Re as forecasting technique. These methods continue until the fifteenth forecasting method, which is the FFORMA. On the other hand, the last positions correspond to ANN, LightGBM, and HW with the CReMSPlusFS technique.

Table 1. Experimentation configuration ordered by sMAPE.

F. Method	F. Configuration	Average sMAPE
ARIMA	Re	37.5052
Jaganathan	Re	37.6426
SMA	Re	38.535
Hybrid	CAS	39.4308
Hybrid	ReMS	39.7137
SMA	CReMSPlusFS	39.7689
SMA	ReMS	40.0366
Hybrid	Re	40.2257
SMA	CAS	40.2651
Hybrid	CReMSPlusFS	40.6069
Hybrid	CAHS	40.8827
Jaganathan	CReMSPlusFS	41.1746
ARIMA	CAHS	41.793
SMA	CAHS	41.7961
FFORMA	ReMS	42.8928
FFORMA	CReMSPlusFS	42.934
FFORMA	CAS	43.0287
Jaganathan	CAS	43.0652
Jaganathan	ReMS	43.3135
FFORMA	Re	43.8414
FFORMA	CAHS	44.1245
Jaganathan	CAHS	45.7345
ARIMA	CReMSPlusFS	45.9772
ARIMA	CAS	46.1315
ARIMA	ReMS	46.6199
HW	Re	47.3814
HW	CAS	52.8533
HW	ReMS	52.9905
HW	CAHS	55.4958
LigthGBM	ReMS	57.7173
LigthGBM	CAS	58.2243
LigthGBM	CAHS	59.0801
LigthGBM	Re	64.1037
ANN	ReMS	68.258
ANN	CAS	68.289
ANN	CAHS	69.6555
HW	CReMSPlusFS	69.7481
LigthGBM	CReMSPlusFS	72.307
ANN	Re	72.5841
ANN	CReMSPlusFS	74.0957

Finally, Table 2 adds the Mean Rank of a non-parametric Friedman Test to Table 1 in column three, which ranks the experimental configuration (the pair of forecasting method and technique). Additionally, we included a fifth column that shows, for every

experimental configuration, the position changes shown in Table 1. Here, we can see that some experimental configurations change ranks by several units; however, the general performance remains the same. We consider the mean rank of the Friedman test to be more reliable than the average sMAPE, given that it was designed for statistical comparisons.

Table 2. Experimentation configurations ordered by Mean Rank. Here, $\uparrow \times r$ or $\downarrow \times k$ means that the experimentation configuration went up or down r or k rows, respectively, regarding Table 1.

F. Method	F. Technique	Mean Rank	Average sMAPE	Position Change
SMA	CAHS	14.31	41.8	↑ × 13
SMA	ReMS	14.49	40.04	↑ × 5
ARIMA	Re	14.54	37.51	↓ × 2
SMA	CAS	14.63	40.27	↑ × 5
Hybrid	CAS	14.71	39.43	↓ × 1
Jaganathan	Re	14.77	37.64	↓ × 4
SMA	Re	15.23	38.54	↓ × 4
Hybrid	ReMS	15.66	39.71	↓ × 3
SMA	CReMSPlusFS	15.91	39.77	↓ × 3
Hybrid	CAHS	16.00	40.88	↑ × 1
Hybrid	Re	17.00	40.23	↓ × 3
Jaganathan	CReMSPlusFS	17.14	41.17	- × -
Hybrid	CReMSPlusFS	17.46	40.61	↓ × 3
FFORMA	ReMS	17.60	42.89	↑ × 1
FFORMA	CAS	17.97	43.03	↑ × 2
FFORMA	CAHS	18.21	44.12	↑ × 5
FFORMA	CReMSPlusFS	18.24	42.93	↓ × 1
ARIMA	CReMSPlusFS	18.49	45.98	↑ × 5
FFORMA	Re	18.51	43.84	↑ × 1
ARIMA	CAHS	19.00	41.79	↓ × 7
Jaganathan	CAS	19.00	43.07	↓ × 3
HW	ReMS	19.07	52.99	↑ × 6
Jaganathan	CAHS	19.16	45.73	↓ × 1
Jaganathan	ReMS	19.41	43.31	↓ × 5
HW	CAS	19.46	52.85	↑ × 2
ARIMA	CAS	19.66	46.13	↓ × 2
HW	Re	20.34	47.38	↓ × 1
HW	CAHS	20.53	55.5	↑ × 1
ARIMA	ReMS	20.54	46.62	↓ × 4
HW	CReMSPlusFS	25.51	69.75	↑ × 7
ANN	ReMS	27.16	68.26	↑ × 3
LigthGBM	ReMS	27.67	57.72	↓ × 2
LigthGBM	CAHS	27.73	59.08	↓ × 1
ANN	CAS	27.83	68.29	↑ × 1
LigthGBM	CAS	27.89	58.22	↓ × 4
ANN	Re	28.20	72.58	↑ × 3
ANN	CAHS	28.24	69.66	↓ × 1
LigthGBM	Re	29.46	64.1	↓ × 5
ANN	CReMSPlusFS	30.74	74.1	↑ × 1
LigthGBM	CReMSPlusFS	32.51	72.31	↓ × 2

Now, we can see that the first forecasting method is SMA followed by ARIMA, Hybrid, and Jaganathan until the fourteenth experimentation configuration with FFORMA. The last forecasting methods are LightGBM, ANN, and HW mixed with ARIMA. We can see that SMA CAHS, SMA ReMS, and SMA CAS have increased their rankings, to positions thirteen ($\uparrow \times 13$), five ($\uparrow \times 5$), and five ($\uparrow \times 5$), respectively, highlighting that the forecasting methods with the most variations are in the top rankings.

5. Discussion

We suspect that ANN and LightGBM produce the highest sMAPE and worst Mean Rank because they did not have enough training data to identify consistent patterns for the forecasting. Furthermore, HW uses exponential smoothing and works with the seasonal component. It is important to highlight that given the size of the time series, we consider that it is hard to identify the seasonal component correctly, which might lead to a poor performance.

On the other hand, SMA does not use complex parameters nor the seasonal component; it only requires the last values to conduct the forecasting. Furthermore, ARIMA models try to describe the autocorrelations in the data, and it is composed of a combination of differencing with autoregression and a moving average model; these parts are described by the p, d and q parameters, respectively, in ARMIA (p, d, q). It is important to highlight that 85% of the total time series were classified as white noise by ARIMA [28], meaning that they lack autocorrelation and have high levels of symmetry by having a uniform distribution of data, similar to a random walk process [26]. Therefore, it is difficult to forecast these series because they have the same probability for the series to go up or down.

These time series used the model ARIMA (0,0,0), which forecasts a straight line with the value of the average of the whole time series, similar to SMA. Additionally, ARIMA can produce different models, leading to variations in its performance.

Additionally, Hybrid and Jaganathan also exhibited good performances but not enough to be called the best, and according to the parsimony principle, if several forecasting methods exhibit similar performances, we should use the simplest among them.

Regarding the forecasting techniques, most of the best five methods tend to disregard the seasonal component, whether ignoring it completely (ReMS) or using only its average (CAS) or a fragment of it (CAHS).

6. Conclusions

In this paper, we tackle the problem of forecasting short-sized crime time series for common crimes (theft, shoplifting, vehicular, and burglary) in several sectors of three cities of Tamaulipas, México. For this problem, we propose the use of simple and machine-learning-based ensemble methods. Additionally, for each one of the forecasting methods, we propose five different forecasting techniques that imply the extraction of the seasonal component of the series, producing a total of forty experimentation configurations.

We rank the experimentation configurations using two criteria, the average sMAPE and the mean rank of the Friedman test. As expected, both rankings produce different orders but with the same pattern, i.e., SMA, ARIMA, Hybrid, and Jaganathan as the best forecasting methods and LightGBM, ANN, and HW as the methods with the lowest performance, leaving FFORMA in the middle.

The best five experimentation configurations tend to dismiss the seasonal component CAHS, ReMS, and CAS. As stated before in Section 5, most of the internal configurations of ARIMA tend to forecast using a simple average of the whole series, ignoring the seasonal component even if it uses the Re forecasting technique. Finally, the worst experimentation configurations use the series with the seasonal (Re) or forecast the seasonal and add it to the forecasting of the series minus the seasonal (CReMSPlusFS).

Therefore, for this study, we can conclude that the best-performing forecasting methods, starting with the best, are SMA, ARIMA, Hybrid, and Jaganathan. In contrast, the worst forecasting methods are LightGBM and ANN. The best forecasting techniques tend to disregard the seasonal component. Furthermore, SMA, Hybrid, FFORMA, HW, ANN, and LightGBM tend to exhibit their best performances by ignoring the seasonal component and their worst otherwise, whereas Jaganathan and ARIMA have an inverse relationship regarding the use of the seasonal component. However, the internal components of ARIMA identify most of the time series as white noise (no autocorrelation and high symmetry), ignoring the seasonal by generating models ARIMA (0,0,0). Moreover, Jaganathan also uses ARIMA as part of its components; hence, we consider it to work similarly.

It is important to remember that these results hold for small-sized time series, and might be specific to these kind of time series. In future work, we recommend using other techniques that involve the tendency and residuals, as well as the development of hyper-heuristics to find better forecasting techniques and the use of a larger set of time series.

Author Contributions: Conceptualization, M.A.C.-N. and M.G.T.-B.; Data curation, M.A.C.-N., M.G.T.-B. and M.P.P.-F.; Formal analysis, M.P.P.-F. and A.S.; Investigation, M.A.C.-N., S.I.-M. and J.L.-M.; Methodology, M.A.C.-N., M.P.P.-F. and J.D.T.-V.; Project administration, S.I.-M. and A.S.; Software, M.P.P.-F. and J.A.C.-R.; Supervision, S.I.-M. and A.S.; Validation, M.G.T.-B., J.D.T.-V. and S.I.-M.; Visualization, J.L.-M.; Writing—original draft, M.G.T.-B., M.P.P.-F. and J.A.C.-R.; Writing—review and editing, J.D.T.-V., S.I.-M. and J.L.-M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The time series used can be found here: https://mega.nz/folder/osFwxIia#vrnxaIFAJmSQ6_pPaFvjfQ accessed on 28 May 2022.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Arnoso Martínez, A.; Vozmediano Sanz, L.; Martínez de Taboada Kutz, C. Inseguridad subjetiva y representaciones sociales de la delincuencia. *Univ. Psychol.* **2018**, *17*, 1–14. [CrossRef]
2. Livier, M.; Martínez, G. Confianza, victimización y desorden en la percepción de inseguridad en una población mexicana Trust, victimization and disorder in a Mexican population's perception of insecurity Resumen. *Psicumex* **2019**, *9*, 1–17.
3. Envipe, S.P. Encuesta Nacional De Victimización Y Percepción Sobre Seguridad Pública (Envipe) 2020. *Inst. Nac. Estadística Geogr.* **2020**, *10*, 1–58.
4. Santos, T.; Jiménez, M.A. El miedo de las víctimas: Diseccionando la Criminología del Control. *Utopía Prax. Latinoam.* **2019**, *24*, 133–153.
5. Armesto, A. Quality of government, crime victimization and particularistic political participation in Latin America. *Perfiles Latinoam.* **2019**, *27*, 1–27. [CrossRef]
6. Millán-Valenzuela, H.; Pérez-Archundia, E. Education, poverty and crime: Links of violence in Mexico? *Convergencia* **2019**, *80*, 1–26. [CrossRef]
7. Pavel, M.; Román, D.; Cecilia, M.; Minchel, J.; Lara, O.P. Reflexiones alternas en torno al tratamiento de las violencias y la delincuencia desde América Latina: La prevención del delito como estrategia. *Med. Soc.* **2019**, *12*, 110–117.
8. Guilmartin, C.E.K. No hay "delitos comunes" Un Planteamiento Alternativo Para Asegurar Puntos Calientes Globales y áreas Urbanas Densamente Pobladas. 2019. Available online: https://www.armyupress.army.mil/Journals/Edicion-Hispanoamericana/Archivos/Segundo-Trimestre-2019/No-hay-delitos-comunes/ (accessed on 28 May 2022).
9. Ordóñez, H.; Cobos, C.; Bucheli, V. Machine learning model for predicting theft trends in Colombia | Modelo de machine learning para la predicción de las tendencias de hurto en Colombia. *RISTI -Rev. Iber. Sist. Tecnol. Inf.* **2020**, *2020*, 494–506.
10. Cichosz, P. Urban crime risk prediction using point of interest data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 459. [CrossRef]
11. Chun, S.A.; Pathak, R.; Paturu, V.A.; Atluri, V.; Yuan, S.; Adam, N.R. Crime Prediction Model using Deep Neural Networks. In Proceedings of the 20th Annual International Conference on Digital Government, Dubai, United Arab Emirates, 18–20 June 2019; pp. 512–514. [CrossRef]
12. Wang, K.; Li, W. Application of Electrical Automation Technology in Power System. *J. Power Energy Eng.* **2019**, *7*, 8–13. [CrossRef]
13. Liu, M.; Lu, T. A Hybrid Model of Crime Prediction. *J. Phys. Conf. Ser.* **2019**, *1168*, 032031. [CrossRef]
14. Jha, S.; Yang, E.; Almagrabi, A.O.; Bashir, A.K.; Joshi, G.P. Comparative analysis of time series model and machine testing systems for crime forecasting. *Neural Comput. Appl.* **2021**, *33*, 10621–10636. [CrossRef]
15. Yadav, R.; Kumari, S.; Savita. *Autoregressive Model for Multivariate Crime Prediction*; Springer: Singapore, 2020; pp. 301–307.
16. Shi, L.; Lu, Y.; Pickett, J.T. The public salience of crime, 1960–2014: Age–period–cohort and time–series analyses. *Criminology* **2020**, *58*, 568–593. [CrossRef]
17. Melgarejo, M.; Rodriguez, C.; Mayorga, D.; Obregón, N. Time Series from Clustering: An Approach to Forecast Crime Patterns. In *Recent Trends in Artificial Neural Networks: from Training to Prediction*; IntechOpen: London, UK, 2020; pp. 1–20. [CrossRef]
18. Izonin, I.; Tkachenko, R.; Shakhovska, N.; Lotoshynska, N. The additive input-doubling method based on the svr with nonlinear kernels: Small data approach. *Symmetry* **2021**, *13*, 612. [CrossRef]

19. Tkachenko, R.; Mishchuk, O.; Izonin, I.; Kryvinska, N.; Stoliarchuk, R. A non-iterative neural-like framework for missing data imputation. *Procedia Comput. Sci.* **2019**, *155*, 319–326. [CrossRef]

20. Makridakis, S.; Spiliotis, E.; Assimakopoulos, V. The M4 Competition: 100,000 time series and 61 forecasting methods. *Int. J. Forecast.* **2020**, *36*, 54–74. [CrossRef]

21. Spiliotis, E.; Assimakopoulos, V.; Makridakis, S.; Assimakopoulos, V. The M5 Accuracy competition: Results, findings and conclusions. *Int. J. Forecast.* 2022, *in press*. [CrossRef]

22. Jaganathan, S.; Prakash, P.K. A combination-based forecasting method for the M4-competition. *Int. J. Forecast.* **2020**, *36*, 98–104. [CrossRef]

23. Montero-Manso, P.; Athanasopoulos, G.; Hyndman, R.J.; Talagala, T.S. FFORMA: Feature-based forecast model averaging. *Int. J. Forecast.* **2020**, *36*, 86–92. [CrossRef]

24. Atıcı, R.; Pala, Z. Prediction of the Ionospheric foF2 Parameter Using R Language Forecasthybrid Model Library Convenient Time Series Functions. *Wirel. Pers. Commun.* **2022**, *122*, 3293–3312. [CrossRef]

25. Hyndman, R.J. Seasonal Decomposition of Short Time Series. 2018. Available online: https://robjhyndman.com/hyndsight/tslm-decomposition/ (accessed on 2 June 2022).

26. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*, 2nd ed.; OTexts: Melbourne, Australia, 2018.

27. Hyndman, R.J.; Athanasopoulos, G. 12.7 Very long and very short time series. In *Forecasting: Principles and Practice*, 2nd ed.; OTexts: Melbourne, Australia, 2018; Chapter 12.

28. Moffat, I.U.; Akpan, E.A. White Noise Analysis: A Measure of Time Series Model Adequacy. *Appl. Math.* **2019**, *10*, 989–1003. [CrossRef]