Hybrid Model for Unemployment Impact on Social Life

Popîrlan, Claudiu-Ionuţ; Tudor, Irina-Valentina; Dinu, Constantin-Cristian; Stoian, Gabriel; Popîrlan, Cristina; Dănciulescu, Daniela

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-84628-1

Hybrid Model for Unemployment Impact on Social Life

Claudiu-Ionuţ Popîrlan *,†, **Irina-Valentina Tudor** †, **Constantin-Cristian Dinu** †, **Gabriel Stoian** †, **Cristina Popîrlan** † and **Daniela Dănciulescu** †

Computer Science Department, Faculty of Sciences, University of Craiova, A.I. Cuza Street, No. 13, 200585 Craiova, Romania; irina.tudor@inf.ucv.ro (I.-V.T.); constantin.dinu@edu.ucv.ro (C.-C.D.); gstoian@ucv.ro (G.S.); popirlan.cristina@ucv.ro (C.P.); daniela.danciulescu@edu.ucv.ro (D.D.)
* Correspondence: popirlan@inf.ucv.ro; Tel.: +40-744-259-593
† These authors contributed equally to this work.

Abstract: In this paper, we want to examine how unemployment impacts social life, and, by using datasets from six European countries, we analyze the effect of unemployment on two of the main aspects of social life: social exclusion and life satisfaction. First, we predict unemployment rates using the Auto Regressive Integrated Moving Average (ARIMA) model and the results are further used in a linear regression model alongside social exclusion and life satisfaction data, thus obtaining the hybrid model. With the help of the point prediction method, we use the hybrid model to predict new values for the two aspects of social life for the upcoming three years and we analyze the results obtained in order to better understand their interconnection. The results suggest that unemployment has particularly adverse effects on the subjective perception of life satisfaction, furthermore increasing the social exclusion percentage.

Keywords: hybrid model; ARIMA model; linear regression model; unemployment; social life; life satisfaction; social exclusion

1. Introduction

Unemployment is one of the most important and complex phenomena in the economic system and the unemployment rate is frequently used as an indicator for analyzing other indicators. Social life is also very important for each individual and its study is useful for understanding many aspects of human life itself. Many authors studied the correlation between unemployment rate and other indicators in different countries or regions of one country ([1–3]), but there are not many studies regarding the connection between unemployment and social life.

Undoubtedly, unemployment has a significant impact on the economy, but it may also have an impact on many other aspects of life in general, such as the health and social life of individuals in society.

As an employee along with significant material benefits, one may also have non-material benefits such as social integration and relationships, whereas unemployment involves the loss of these benefits and thus can lead to poorer subjective life satisfaction [4].

The literature suggests that research done on unemployment's effects on mental health show an associated stigma and a feeling of shame for the individuals involved and, furthermore, involuntary unemployment has a universal effect on affected persons' mental health [5]. The traditional research on unemployment pays great attention to numerous intervention factors such as community support in order to minimize the negative effects unemployment has on society. Work done in [6] shows the use of online media content for unemployment prediction, while [7] studies the relationship between unemployment and psychological well-being by analyzing Twitter posts.

From another point of view, research shows that there are gender differences in the connection between unemployment and mental health, specifically women tend to be less

affected by unemployment than men in the labour market and in the family, mainly due to women's roles in the family [8].

Using data from six European countries: Bulgaria, the Czech Republic (also known as Czechia), Hungary, Poland, Romania and Slovakia, this study analyzes the relationship between unemployment and social life. We use a hybrid model consisting of the ARIMA model that computes and forecasts unemployment rates and a linear regression model to evaluate its connection to social life. We want to see if unemployment has a significant negative effect on self-perceived life satisfaction and if the effect is different across the countries as well as the impact unemployment has on the social exclusion phenomenon. The ARIMA model has been previously used and integrated in hybrid models for various time series forecasting and efficient unemployment rate forecasts with encouraging results [9–11].

In [12], Ho and Xie used ARIMA to study the reliability forecasting of a mechanical system failure, concluding that this model produces satisfactory results concerning its performance in prediction. This time series model is very flexible and makes very few assumptions [13].

The ARIMA model was used for prediction in different fields, obtaining reliable results. For example, in [14] the authors compared ARIMA and GM(1, 1) models and concluded that the first one performed better and can be a real support for controlling hepatitis B in China. Keeping in the health domain, recently, given the pandemic situation, ARIMA models were used in [15,16] to predict the evolution of Covid-19 cases. In the economic field, the ARIMA model is used in macroeconomy (such as in stock price prediction [17]) and in microeconomy (for example, in [18] the model is used to predict the demand for a food company).

The results we obtain prove a strong relationship between the examined data in most of the six countries included in the study. Our findings show that social life is strongly impacted by employment status. We take a look at two of the most important aspects of social life and observe that there is a positive slope between the unemployment rate and the percentage of people at risk of social exclusion and a negative slope between the unemployment rate and self-perceived life satisfaction.

2. Social Life

Social life and social connections have been the a point of interest for social scientists, therefore several features have been distinguished, mainly based on the existence of relationships or the lack of them [19].

The concept of social life is basically the opposite of private life, and contains all the activities that happen in the community or public zone while interacting with other individuals. Social life can be measured by the duration and quality of the social relations an individual has on a regular basis, both in person and online.

Two of the main important features of social relations are: social integration, which includes all types of informal relationships (e.g., spouse) or formal relationships (with organizations and/or institutions); and social isolation, which means the lack of social relationships. A specific type of social isolation is social exclusion which concerns those individuals that are forced beyond their control to separate themselves from society in general due to poverty, social inequality, or disadvantages in living conditions, because the lack of money makes it harder for people to relate on equal terms [20].

All of this implies that, to a certain degree, every person has a social life that is conducted following certain codes and rules of society in general and it has an important influence on their mental and physical health.

The groups and relationships of an individual led to the term network that describes all the ties a person has in the social area and has, of course, a limited set of participants and relations between them [21]. From this we can extract social capital, which is the total amount of benefits and resources that a member can extract from his social network or other social structures and can be compared to economic capital in people's bank accounts [22].

The quality of social life from a relationships perspective includes both negative aspects of relationships, such as conflict, fights and tense situations, and positive aspects such as emotional support provided by partners, spouses or friends [23]. In general, an individual with an active social life and a large network is identified as having a healthier and happier life with a high degree of self-perceived life satisfaction and well-being.

Social exclusion was first referred to in the 1970s in France and it described a person that was not protected by social insurance and was very close to being socially detached. Later, in the 1980s, the unemployment rates were high and social exclusion was a risk for social integration models [24]. The European Union declared 2010 the European Year for Combating Poverty and Social Exclusion.

Being socially excluded from society is perceived as having bad effects on living conditions, income, education and overall well-being leading to the impossibility of keeping a decent living standard and having an appropriate social life [25].

Social exclusion is not only induced by low financial income, but also by self-perceived well-being and connection to society. Social integration is partly subjective and is related to the individual's capacity for integration and acting in their own best interest, the other part being the actual behavior the individual manifests. The subjective part is highly influenced by life satisfaction in general—their mental health and their position in the social community as extrinsic factors—but also depends on intrinsic factors such as the individual's character and personality and their ability to cope with deprivation factors [26].

For this study we considered six countries from Central and South-Eastern Europe (Bulgaria, Czech Republic, Hungary, Poland, Romania and Slovakia) and we draw the data from the open-access repository Eurostat, the statistical office of the European Union (EU). The sample we considered is restricted to the working-age population with ages between 16 and 64, and the overall data coverage runs from 2005 to 2020.

For the social inclusion part of social life, we took into consideration the subjective perception of life satisfaction, summarized in Table 1 as the self-perceived life satisfaction (SPLS) for the six countries mentioned, using the data in the Health variables of The European Statistics of Income and Living Condition (EU-SILC), Quality of Life survey. The main dependent variable is subjective life satisfaction perception, represented on a scale from 0 to 4, where 0 means *very bad* and 4 means *very good* life satisfaction. We standardized the SPLS scores with a mean of 0 and a standard deviation of 1 for each dataset (1), using the corresponding weights from the survey.

$$x_{std} = \frac{x - \mu}{\sigma}. \tag{1}$$

We notice that the considered population is fairly satisfied with their lives, Romania has one of the highest self-perceived life satisfaction scores with a total of 81% between good and very good SPLS, followed by Bulgaria with 78%, Czech Republic with 72% and Slovakia with 71%. We can observe the standardization in the second row for each country, allowing us to measure the scores on the same metric. We notice that individuals with the highest SPLS score are 0.634 standard deviations above the mean in Romania, ranking the highest score once again, followed by Slovakia with 0.327 and the Czech Republic with 0.269.

On the other side we take a look at the social exclusion part of social life, so we considered the percentage of people at risk of poverty or social exclusion (PRPSE) from the same Eurostat survey [27], examining once again only the working-age population within the same time frame from 2005 to 2020. This time the data are no longer subjective and we can compare the statistics for the six European countries in Figure 1.

We can observe a decreasing trend for all six countries, as the EU is constantly expanding their legal support to stimulate economic activity and to reduce the risks of poverty and social exclusion for the population [28].

Table 1. The distribution of self-perceived life satisfaction scores.

Country	UM	Very Bad	Bad	Fair	Good	Very Good	All
Bulgaria	percent	1.30	5.18	14.93	54.68	23.92	100
	standardized SPLS score	−0.878	−0.696	−0.238	1.629	0.184	0.00
Czechia	percent	1.26	6.04	20.12	47.68	24.92	100
	standardized SPLS score	−1.025	−0.764	0.006	1.514	0.269	0.00
Hungary	percent	2.05	8.25	24.06	45.54	20.11	100
	standardized SPLS score	−1.069	−0.699	0.242	1.520	0.006	0.00
Poland	percent	1.28	7.59	23.43	47.54	20.17	100
	standardized SPLS score	−1.049	−0.695	0.192	1.543	0.009	0.00
Romania	percent	0.90	3.71	14.11	48.44	32.83	100
	standardized SPLS score	−0.944	−0.805	−0.291	1.405	0.634	0.00
Slovakia	percent	1.51	6.32	20.34	46.09	25.76	100
	standardized SPLS score	−1.049	−0.776	0.019	1.480	0.327	0.00

Note: UM: unit of measurement. Percent: weighted percentages. Percentages may not add up to 100.00% because of rounding.
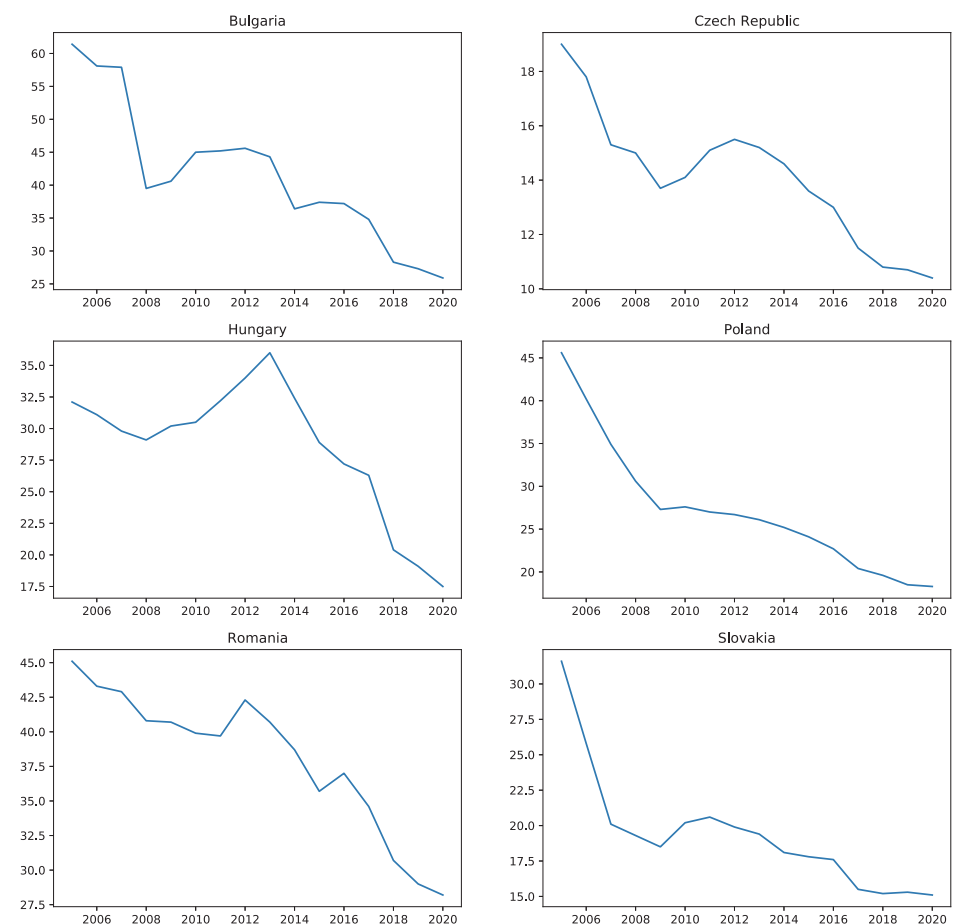
Figure 1. People at risk of poverty or social exclusion (percentage).

3. Unemployment

Unemployment has a major impact on a country's economy and predicting the unemployment rate is an essential factor for economic and financial growth due to its correlation with the country's business cycle and its influence on monetary policy.

From the perspective of social life, unemployment is one of the most important circumstances that leads to social exclusion, because being without a job means less income and the separation from society, which could affect not only one generation, but possibly the next one, too [29]. From the individual's perspective, being jobless and disposing of diminished income leads to a lowering of living standards and thus not being able to take part in cultural and social activities in the same way as before. All of these issues supply furthermore the feeling of being socially excluded [30].

Unemployment and its effects on society have been studied on both small and large scales, restricted to certain geographical areas or compared between different countries and the system of intervention they implement [31].

For the six countries analyzed in this study (Bulgaria, Czech Republic, Hungary, Poland, Romania and Slovakia), we used Eurostat data [32]. Once again, the sample we considered was restricted to the working-age population with ages between 16 to 64 and the overall data coverage runs from 2000 to 2020; monthly values were considered. From the total 252 months, we split the data into a training set of 216 months and a test set of 36 months. In Table 2, we can observe an outline of the dataset used for the unemployment rate (UR).

Table 2. Summary of the unemployment rate dataset.

Data Country	Training Set Size (2000–2017) Months	Test Set Size (2018–2020) Months	Minimum Value	Maximum Value
Bulgaria	216	36	3.6	20.7
Czechia	216	36	1.7	9.6
Hungary	216	36	2.9	12.1
Poland	216	36	2.8	21
Romania	216	36	3.6	9.5
Slovakia	216	36	5.6	20.2

Starting with the training dataset, we generated Figures 2–7 that show a visual review of training datasets for the unemployment rate for each country and its corresponding Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) plots.
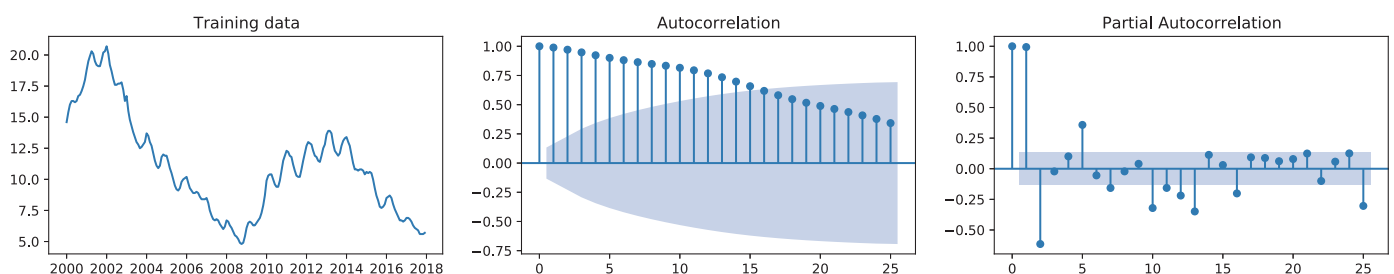
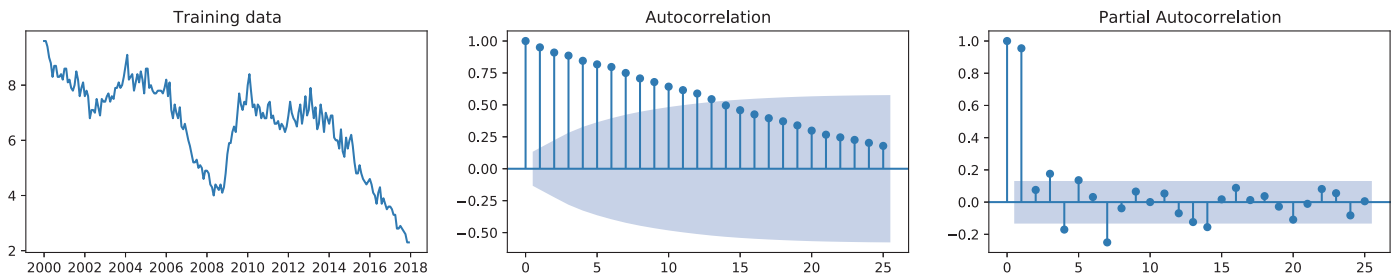**Figure 2.** Training data, ACF and PACF plots for Bulgaria.

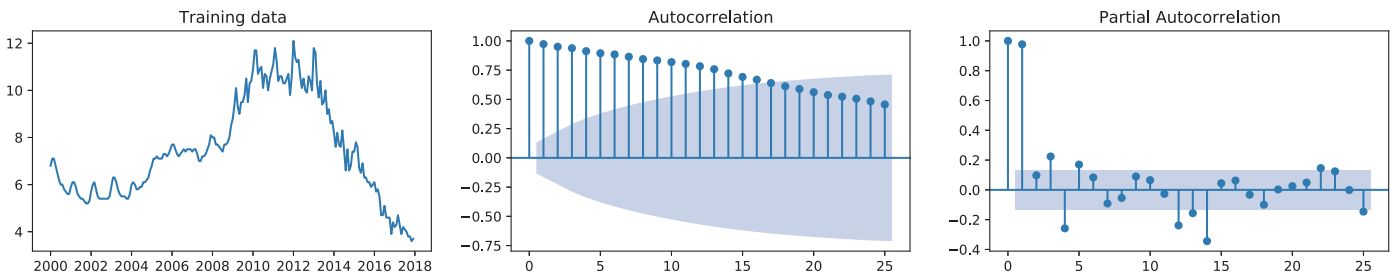**Figure 3.** Training data, ACF and PACF plots for Czechia.

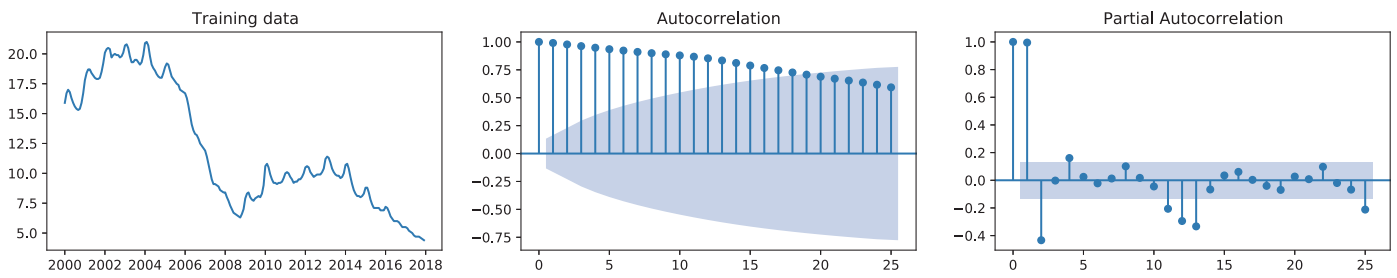**Figure 4.** Training data, ACF and PACF plots for Hungary.

Figure 5. Training data, ACF and PACF plots for Poland.
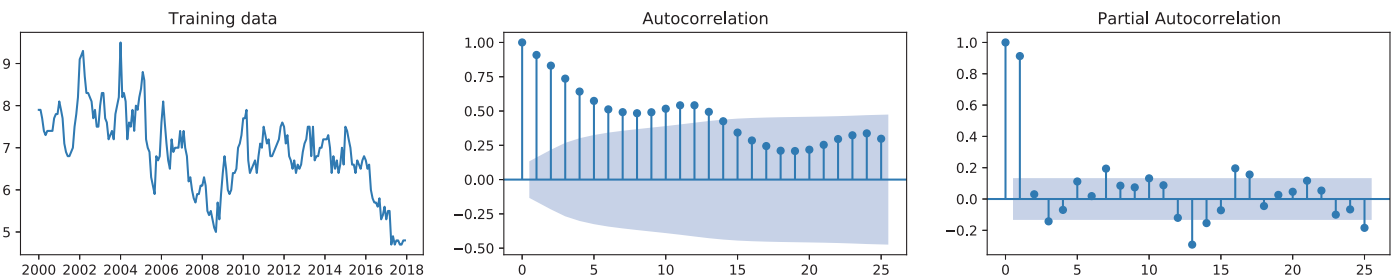

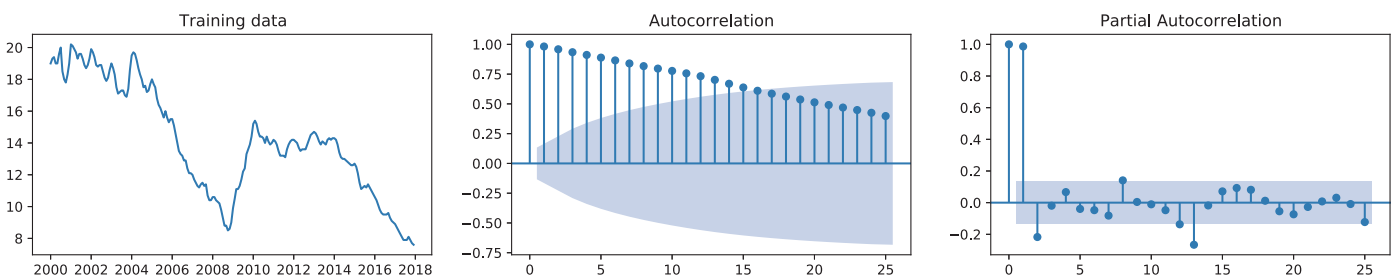**Figure 6.** Training data, ACF and PACF plots for Romania.

Figure 7. Training data, ACF and PACF plots for Slovakia.

We used a Shapiro–Wilk's test ($p > 0.05$) [33,34] and we concluded that the UR values are approximately normally distributed for each country considered, with a skewness of 0.640 and a kurtosis of -0.232 for Bulgaria, a skewness of -0.664 and a kurtosis of -0.869 for Czechia, a skewness of 0.309 and a kurtosis of -0.875 for Hungary, a skewness of 0.410 and a kurtosis of -1.134 for Poland, a skewness of -0.850 and a kurtosis of 0.009 for Romania and a skewness of -0.122 and a kurtosis of -0.972 for Slovakia [35–37]. We computed that the skewness standard error (SE) is 0.501 and the kurtosis SE is 0.972, for all six countries considered.

4. Hybrid Model for Unemployment Rate over Social Exclusion and Life Satisfaction

The proposed hybrid model for unemployment rate, people at risk of poverty or social exclusion and self-perceived life satisfaction approach followed the next steps:

- ARIMA model training the UR data for 2000–2017;
- ARIMA model testing the UR data for 2018–2020;
- Linear regression model for correlation between UR, PRPSE and SPLS data for 2005–2020;
- ARIMA model forecasting UR for 2021–2023;
- Linear regression model used UR forecasts for point predictions of PRPSE and SPLS for 2021–2023.

In the next section, we will briefly describe the ARIMA model and the results we obtained by applying it to the unemployment rate data for the six mentioned countries.

4.1. ARIMA (Autoregressive Integrated Moving Average) Model

Autoregressive models are used to make predictions based on past values and historical observations that can affect the current values and are used to analyze different time-varying systems.

ARIMA is an autoregression model that uses linear stationary time series in an attempt to forecast a value of a dependent variable at the present moment depending on the values it had in previous moments in the past. Any non-seasonal time series that displays patterns and is not just a random series can be modeled using ARIMA models.

An ARIMA model features three parameters, $ARIMA(p, d, q)$, where p is the order of the AR term, q is the order of the MA term and d is the level of differencing used to transform the time series into a stationary one. Autoregression models work best when the predictors are not correlated and are independent of each other.

The easiest and most used approach to make the series stationary is to difference it by subtracting the previous value from the current value. Sometimes, more than one differencing may be needed depending on the complexity of the series.

p is the order of the AR, that is, a pure Auto Regressive model where Y_t depends only on its own lags:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \epsilon_1, \tag{2}$$

where Y_{t-1} is the lag1 of the series, β_1 is the coefficient of lag1 that the model estimates and α is the intercept term, also estimated by the model.

q is the order of the MA, that is a Moving Average model where Y_t depends only on the lagged forecast errors.

$$Y_t = \alpha + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}, \tag{3}$$

where $\theta_1, \ldots, \theta_q$ are the parameters of the model and the $\epsilon_t, \ldots, \epsilon_{t-q}$ are white noise error terms.

The ARIMA model can be expressed as follows:

$$Y_t = \alpha + \sum_{i=1}^{p} \beta_i Y_{t-i} + \epsilon_t - \sum_{j=1}^{q} \theta_j \epsilon_{t-j}, \tag{4}$$

where Y_t is the value of the variable considered at time t, ϵ_t is the random error at time t, β_i and θ_j are the coefficients of the model [11].

Basically, the model can be summarized in words as follows:

$$Predicted\ Y_t = Constant + Linear\ Combination\ Lags\ of\ Y + Linear\ Combination\ of\ Lagged\ Forecast\ Errors. \qquad (5)$$

In order to build our ARIMA model, we have to estimate the parameters of the model using ACF and PACF plots that help us determine the p and q parameters for the AR and MA, respectively. A value of 0 can be used for any of the parameters, which means we will not be using that element of the model.

If we look at the graphical representation of the monthly unemployment data (Figures 2–7), we can observe that the datasets are nonstationary and nonlinear, thus we considered the ARIMA model.

We used the Augmented Dickey–Fuller (ADF) Test [38,39] to check whether the data were stationary or not and we observed that the test rejects the null hypothesis of stationarity in most of the data considered. All data failed to reject the null hypothesis and became stationary after applying first-order differencing and testing again with the ADF Test gave us the differencing parameter $d = 1$ for the ARIMA model.

In order to build the appropriate ARIMA model we used the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) in order to find the best forecasting parameters for the model [40]. After fitting the model, we tested the training data predicted values and compared them to the actual values; we calculated the residual errors and we generated monthly predictions for 36 months ahead (2021–2023). The predictions for the six countries' test data, along with actual test values and future forecasts, are plotted in Figure 8.

From Figure 8, we observe that the constructed ARIMA models achieve quite good prediction results and fit the observed time series. Some deviations can be noticed between the time series values and the model predictions, and this can be explained by the fact that the model has a relatively simple structure.

The findings will be further used in the next section to evaluate how unemployment relates to social life, considering its two aspects: the subjective life satisfaction score and the risk of social exclusion.

4.2. Linear Regression of PRPSE and SPLS on UR

Starting from the desire to analyze the connection between unemployment on one hand and two of the most important parts of social life on the other hand, we studied the correlation between the unemployment rate (UR), the percentage of people at risk of poverty or social exclusion (PRPSE) and the self-perceived life satisfaction (SPLS), considering them two-by-two, pairwise.

Since the data about life satisfaction are subjective opinions, we computed a new variable as a weighted mean using the following weights: 0 for *very bad*, 1 for *bad*, 2 for *fair*, 3 for *good* and 4 for *very good*. Thus, a higher score means that more people have a better self-perception about their life, while a lower score is translated into a poor life-perception for most people. These variables were analyzed for the six countries for the last 16 years (2005–2020).

First, we want to observe the relationship between the unemployment rate and PRPSE and between the unemployment rate and SPLS.
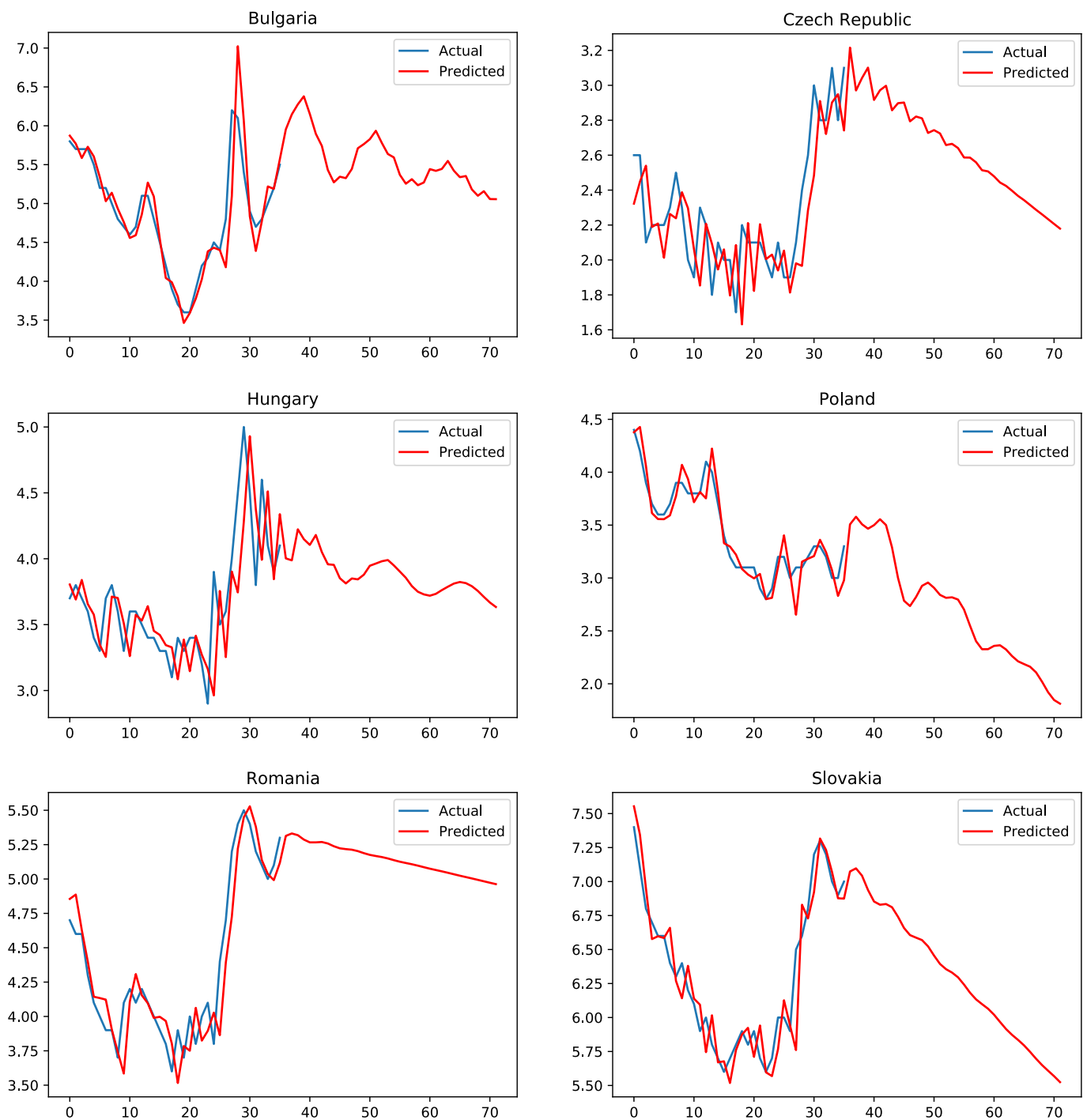
Figure 8. Test data predictions and 3 years ahead forecasts (unemployment rates in months).

From Figure 9, one can notice that there is a positive slope between PRPSE and UR, meaning that there is a positive linear correspondence for all six countries, while in Figure 10 one can notice a negative slope between SPLS and UR, which means there is a negative linear relationship for Czechia, Poland, Romania and Slovakia but it is less clear for Bulgaria and Hungary. For those countries where the relation is significant, as the UR increases, the PRPSE also increases and the SPLS decreases.
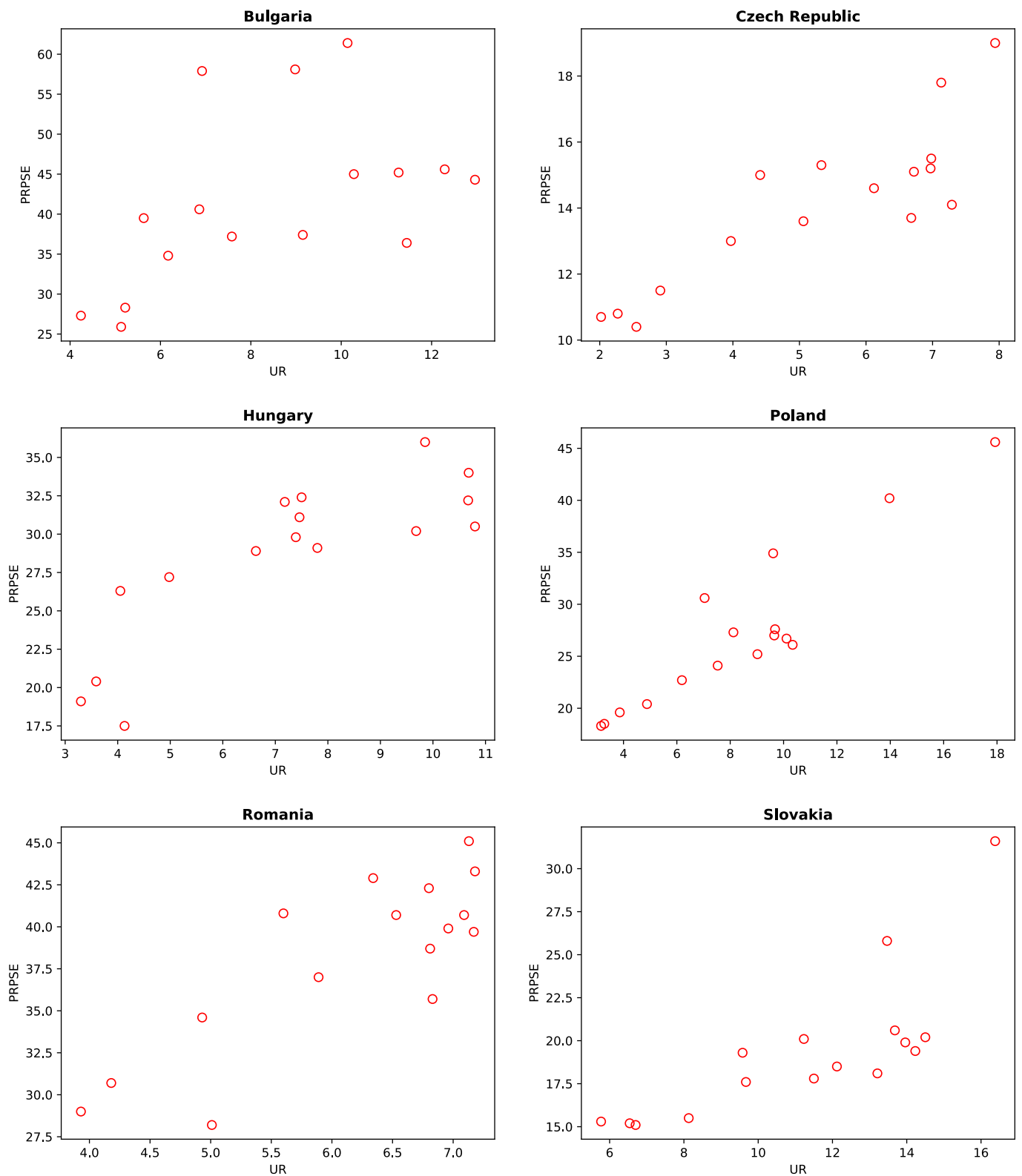
Figure 9. Scatter plot of PRPSE (percentage) on unemployment rate (percentage).

Computing the Pearson correlation coefficient on the three variables (UR, PRPSE and SPLS) for every studied country (detailed in Table 3), we noticed a very strong positive correlation between UR and PRPSE (the correlation coefficient takes values from +0.501 to +0.921, except for Bulgaria, the rest of the values are above +0.765 at a significance level

lower than 0.01), while the correlation between UR and SPLS is a strong negative and one, in general, with values from -0.548 to -0.851, except for Bulgaria and Hungary. One can also notice a very strong negative correlation between PRPSE and SPLS, suggesting the two opposite aspects of what we denote social life are likewise highly connected. The values of Pearson correlation coefficients range from -0.801 to -0.970 except for Hungary, suggesting an almost perfect correlation between these indicators.
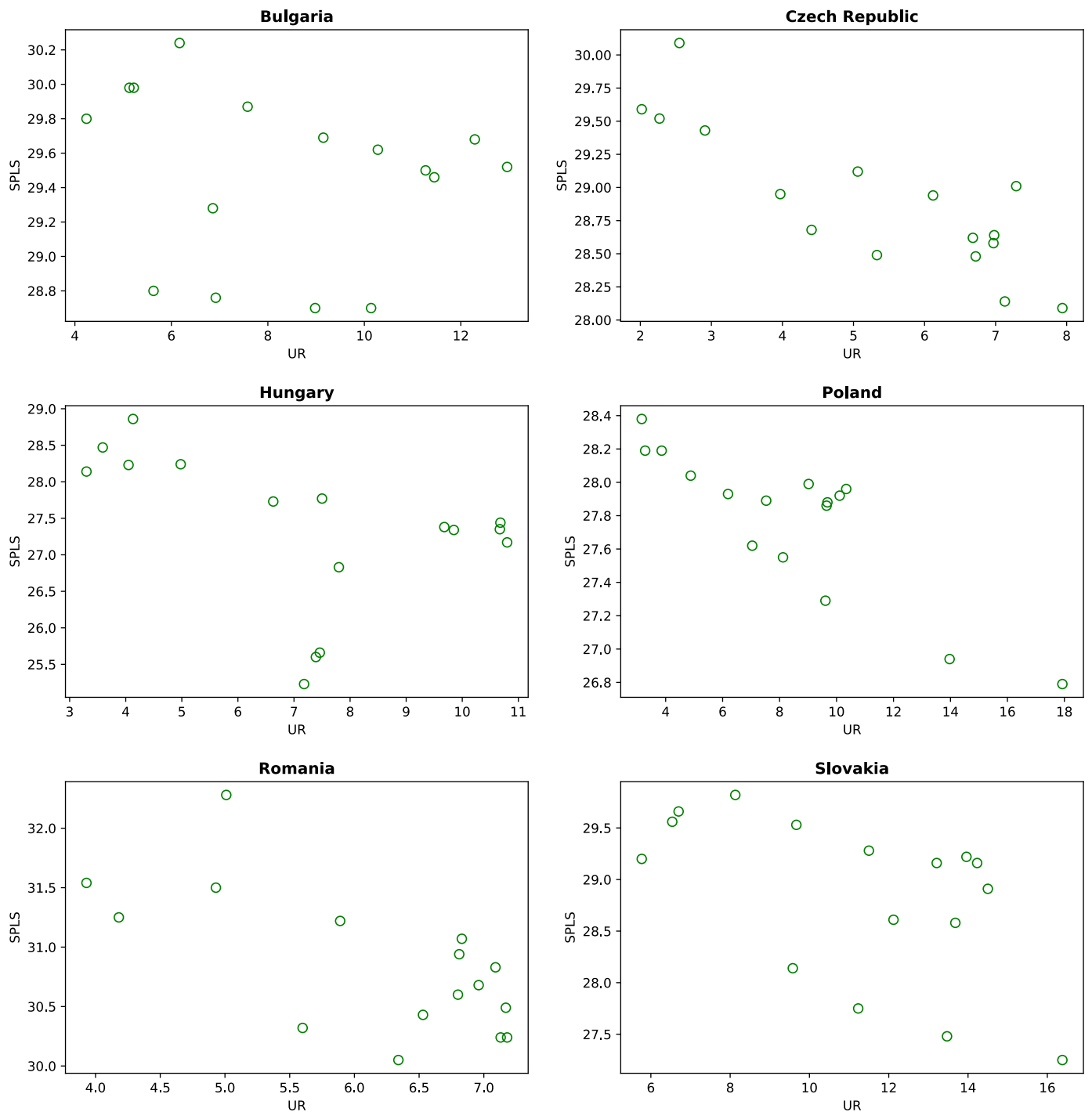

Figure 10. Scatter plot of SPLS (percentage) on unemployment rate (percentage).

Table 3. Correlations for the selected countries.

Country		UR Pearson Correlation	UR Sig. (2-Tailed)	PRPSE Pearson Correlation	PRPSE Sig. (2-Tailed)	SPLS Pearson Correlation	SPLS Sig. (2-Tailed)
Bulgaria	UR			0.501 *	0.048	−0.181	0.502
	PRPSE	0.501 *	0.048			−0.801 **	0.000
	SPLS	−0.181	0.502	−0.801 **	0.000		
Czechia	UR			0.868 **	0.000	−0.841 **	0.000
	PRPSE	0.868 **	0.000			−0.941 **	0.000
	SPLS	−0.841 **	0.000	−0.941 **	0.000		
Hungary	UR			0.833 **	0.000	−0.430	0.097
	PRPSE	0.833 **	0.000			−0.579 *	0.019
	SPLS	−0.430	0.097	−0.579 *	0.019		
Poland	UR			0.921 **	0.000	−0.851 **	0.000
	PRPSE	0.921 **	0.000			−0.970 **	0.000
	SPLS	−0.851 **	0.000	−0.970 **	0.000		
Romania	UR			0.828 **	0.000	−0.667 **	0.005
	PRPSE	0.828 **	0.000			−0.916 **	0.000
	SPLS	−0.667 **	0.005	−0.916 **	0.000		
Slovakia	UR			0.765 **	0.001	−0.548 *	0.028
	PRPSE	0.765 **	0.001			−0.849 **	0.000
	SPLS	−0.548 **	0.028	−0.849 **	0.000		

Note: *. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).

All these observations show a powerful connection between unemployment and poverty or the social exclusion phenomenon. An increase of UR leads to the increase of PRPSE, that is, to poverty and social exclusion and to a decrease of self-perceived life satisfaction. Of course there are several other possible factors for a decline in self-perceived life satisfaction, but it is obvious that unemployment has a significant role among those factors. That is somehow logical, but this study shows some relevant results to support these ideas.

Since we obtained good estimations of UR using the ARIMA model, we also obtained new predicted values for PRPSE and SPLS using the linear modeling provided by the linear regression:

$$Y = a + b * X + \epsilon, \tag{6}$$

where X is the independent variable UR and Y is PRPSE or SPLS.

Thus, the predicted values $\hat{Y}$ are determined ($\hat{Y} = b_0 + b_1 * X$), leading to the forecast of the impact of unemployment rate over the social life for the following years.

In our case, we want to find a linear model for the dependence between PRPSE or SPLS and UR. That is,

$$PRPSE = a_1 + b_1 * UR \tag{7}$$

and

$$SPLS = a_2 + b_2 * UR. \tag{8}$$

We used linear regression to find the linear coefficients in (7) and (8) for each country. These coefficients are different in every country due to the specificity of the social interactions and economic environment of each country and the differences can be observed in Table 4.

Table 4. Linear coefficients for the selected countries.

Country	PRPSE	SPLS
Bulgaria	$PRPSE = 25.423 + 1.923 * UR$	-
Czechia	$PRPSE = 8.532 + 1.053 * UR$	$SPLS = 30.108 - 0.23 * UR$
Hungary	$PRPSE = 16.35 + 1.688 * UR$	-
Poland	$PRPSE = 12.031 + 1.804 * UR$	$SPLS = 28.585 - 0.096 * UR$
Romania	$PRPSE = 14.158 + 3.891 * UR$	$SPLS = 33.088 - 0.363 * UR$
Slovakia	$PRPSE = 8.056 + 1.003 * UR$	$SPLS = 30.354 - 0.135 * UR$

Table 5 presents the R square, Adjusted R square and Standard error of estimate for PRPSE and SPLS for every country.

Notice that for Bulgaria and Hungary the null hypothesis cannot be rejected for SPLS, while for their PRPSE formula and for both PRPSE and SPLS for the other countries we can reject the null hypothesis and we can conclude that there is a significant linear relationship between the unemployment rate and the two aspects of social life.

For Romania and Slovakia, SPLS R-square values are lower than 0.5 suggesting that for these countries some other variables must be included in a multi-linear model, in future research. The same applies to the Bulgaria PRPSE variable.

Table 5. R square, Adjusted R square and Standard error of estimate.

Country	Variable	R Square	Adjusted R Square	Standard Error of Estimate
Bulgaria	PRPSE	0.251	0.198	9.595
Czechia	PRPSE	0.753	0.735	1.252
	SPLS	0.707	0.686	0.306
Hungary	PRPSE	0.693	0.671	3.054
Poland	PRPSE	0.848	0.838	3.074
	SPLS	0.724	0.704	0.239
Romania	PRPSE	0.686	0.663	2.990
	SPLS	0.444	0.405	0.461
Slovakia	PRPSE	0.585	0.555	2.825
	SPLS	0.300	0.250	0.688

4.3. Forecasting the PRPSE and SPLS

In this section, we want to predict the values of PRPSE and SPLS. Our input for prediction is given by the predicted values of the unemployment rate.

Using the formulas from the linear regression and the predicted values of UR for the next three years, given by the ARIMA model, we can compute the estimated values for PRPSE and SPLS.

For every value of UR for 2021, 2022 and 2023, we get a point prediction PRPSE for every country and a point prediction SPLS for Czechia, Poland, Romania and Slovakia (see Table 6).

Table 6. The point prediction of PRPSE and SPLS.

Country	Variable	2021	2022	2023
Bulgaria	UR (predicted)	5.780	5.557	5.293
	PRPSE (computed)	36.538	36.110	35.602
Czechia	UR (predicted)	2.954	2.640	2.329
	PRPSE (computed)	11.643	11.312	10.984
	SPLS (computed)	29.428	29.500	29.572
Hungary	UR (predicted)	4.010	3.881	3.750
	PRPSE (computed)	23.119	22.902	22.681
Poland	UR (predicted)	3.263	2.635	2.086
	PRPSE (computed)	17.918	16.785	15.795
	SPLS (computed)	28.271	28.332	28.384
Romania	UR (predicted)	5.055	4.892	4.765
	PRPSE (computed)	33.829	33.196	32.699
	SPLS (computed)	31.252	31.311	31.358
Slovakia	UR (predicted)	6.858	6.305	5.753
	PRPSE (computed)	14.934	14.380	13.826
	SPLS (computed)	29.428	29.502	29.577

Now, let us consider a desired level of prediction at 90%, that means α is 10%. Since in our study we analyze a period spanning 16 years, we have $n = 16$ and $1 - \frac{\alpha}{2} = 95\%$. We obtain a *t*-value equal to 1.761 and we can determine the prediction interval using the formula:

$$\hat{Y} \pm t_{1-\frac{\alpha}{2}, n-2} \cdot SE_{\hat{Y}_{pred}}, \tag{9}$$

where

- $\hat{Y}_{pred}$ is the point prediction calculated for each country in Table 6;
- $t_{1-\frac{\alpha}{2}, n-2}$ is the *t*-value for the established confidence level and degree of freedom;
- $SE_{\hat{Y}_{pred}}$ is the standard error of estimate.

The margin of error is computed by the *t*-value multiplied by the standard error of estimate; thus, the lower bound of the predicted interval is equal to the point prediction minus the margin of error, while the upper bound of the predicted interval is equal to the point prediction plus the margin of error.

Regarding the distribution of residues, we observe that their distribution is approximately normal. We applied the Shapiro–Wilk's test ($p > 0.05$) for PRPSE and we obtained a skewness of 0.144 and a kurtosis of 0.269 for Bulgaria, a skewness of 0.234 and a kurtosis of -0.231 for Czechia, a skewness of -0.434 and a kurtosis of -0.942 for Hungary, a skewness of 0.598 and a kurtosis of 0.074 for Poland, a skewness of -0.283 and a kurtosis of -0.096 for Romania and a skewness of 0.322 and a kurtosis of 0.976 for Slovakia.

For SPLS, we obtained a skewness of 0.589 and a kurtosis of 0.033 for Czechia, a skewness of 0.020 and a kurtosis of -1.086 for Poland, a skewness of 0.300 and a kurtosis of 0.656 for Romania and a skewness of -0.648 and a kurtosis of -1.144 for Slovakia.

We computed that, for PRPSE and SPLS, the skewness standard error (SE) is 0.564 and the kurtosis SE is 1.091, for all six countries considered.

So, with a 90% probability, we can state that, if the unemployment rates for the next three years for the six studied countries are those in Table 6, then PRPSE and SPLS for the corresponding period of time are somewhere between the lower bound and the upper bound of each country in every year, 2021, 2022 and 2023 (Table 7). Therefore, the best estimations are the point predictions from Table 6.

Table 7. The lower and upper bound of the confidence intervals for PRPSE and SPLS.

Country	Variable	SE of Estimate	Margin of Error	Lower Bound 2021	Upper Bound 2021	Lower Bound 2022	Upper Bound 2022	Lower Bound 2023	Upper Bound 2023
Bulgaria	PRPSE (computed)	9.595	16.896	19.642	53.435	19.213	53.007	18.705	52.499
Czechia	PRPSE (computed)	1.252	2.206	9.437	13.849	9.106	13.518	8.778	13.191
	SPLS (computed)	0.306	0.540	28.888	29.968	28.960	30.040	29.032	30.112
Hungary	PRPSE (computed)	3.054	5.378	17.741	28.497	17.524	28.280	17.303	28.059
Poland	PRPSE (computed)	3.074	5.414	12.503	23.333	11.370	22.200	10.381	21.210
	SPLS (computed)	0.239	0.422	27.849	28.693	27.909	28.754	27.962	28.806
Romania	PRPSE (computed)	2.990	5.266	28.563	39.096	27.929	38.462	27.432	37.965
	SPLS (computed)	0.461	0.812	30.440	32.065	30.499	32.124	30.546	32.170
Slovakia	PRPSE (computed)	2.825	4.975	9.959	19.910	9.404	19.356	8.850	18.802
	SPLS (computed)	0.688	1.212	28.215	30.640	28.290	30.714	28.365	30.789

5. Discussion

Unemployment causes deprivation not only from a financial point of view, but also in five important psychosocial areas according to [41]: the need for social connections outside the family, the need for personal identity, the need for a daily time structure, of having regular activity and the necessity of being part of a collective purpose. Being deprived of all this could explain why social integration is reduced and self-perceived life satisfaction declines when one becomes unemployed.

We started building the hybrid model with the evaluation of unemployment rate first, trying to fit the classical $ARIMA(p, d, q)$, using the Python programming language. For that, we generated the ACF and PACF plots and used the ADF test for stationarity thus obtaining the value 1 for the d parameter of the model. Then we used the best AIC value for each dataset for the six countries in order to obtain the p and q parameters for the model. After fitting the ARIMA model we tested, generated predictions for the following 3 years and we evaluated the performance based on the root mean square error (RMSE).

We fitted the following ARIMA models:

- ARIMA(5, 1, 4) to Bulgaria unemployment rate data with AIC = 11.474, BIC = 50.210 and obtained the RMSE = 0.310 for the predicted values, coefficient of determination = 0.833;
- ARIMA(2, 1, 2) for Czechia UR data with AIC = 133.833, BIC = 154.961 and RMSE = 0.246 for the predicted values, coefficient of determination = 0.794;
- ARIMA(4, 1, 4) for Hungary UR data with AIC = 218.825, BIC = 254.039 and RMSE = 0.337 for the predicted values, coefficient of determination = 0.812;
- ARIMA(5, 1, 4) having AIC = −86.956, BIC = −48.291 and RMSE = 0.142 for the Poland dataset, coefficient of determination = 0.904;
- ARIMA(4, 1, 2) for Romania with AIC = 174.236, BIC = 202.407 and RMSE = 0.238 for the predictions, coefficient of determination = 0.840;
- ARIMA(3,1,1) for Slovakia UR data having AIC = 77.396, BIC = 98.524 and RMSE = 0.202 for the predicted values, coefficient of determination = 0.883.

In order to check the presence of autocorrelation in the residuals we used the Ljung–Box portmanteau test [42] with a number of lags equal to five as suggested by [43]. For each model, we obtained Ljung–Box p-values greater than 0.05 that fail to reject the null hypothesis (Bulgaria p-value = 0.394, Czechia p-value = 0.359, Hungary p-value = 0.940, Poland p-value = 0.242, Romania p-value = 0.770, Slovakia p-value = 0.758).

In the second part of the study, with the help of the SPSS Statistics software package, we used the UR obtained predictions, the PRPSE and SPLS data, and with the help of point prediction method, we predict the PRPSE and SPLS for the same three future years

using the linear regression model. The statistical data proves to us that there is a powerful connection between the unemployment rate and the different aspects of what we call social life.

We can easily observe that there is a positive slope between the unemployment rate and the percentage of people at risk of poverty or social exclusion (PRPSE) and a negative slope between UR and the new defined variable called self-perceived life satisfaction (SPLS). Moreover, the relationship between these variables is quite significant. We presented tables and figures for all obtained data, in order to make the results more understandable. Available datasets span over a 21 year interval for UR (2000–2021) and a 16 year interval for PRPSE and SPLS, reflecting the period 2005–2020.

The obtained results show a strong relationship between the examined factors in most of the six studied countries. Our findings reveal that social life is strongly impacted by the employment status. However, in Bulgaria and Hungary, self-perceived life satisfaction is weakly correlated with unemployment rate. Some future research must be conducted in these countries in order to identify the factors that have a great influence on SPLS, other than UR. On the other hand, Poland had the strongest correlation coefficient for both PRPSE and SPLS, suggesting that social life is strongly connected with labor life in this country.

6. Conclusions and Future Work

The economic environment can be extremely oscillating over time because of random external factors; making accurate predictions therefore becomes quite challenging. Although social life is a complex concept that depends on many aspects, we concentrated our research on the impact of unemployment rate on two measurable parameters of this concept (people at risk of poverty or social exclusion (PRPSE), self-perceived life satisfaction (SPLS)), because unemployment is one of the most important social-economic phenomena.

In our study we implemented the ARIMA model fitting the unemployment rate data for an 18 year span in six European countries (Bulgaria, Czechia, Hungary, Poland, Romania and Slovakia) and then we tested the model for the next 3 years (2018–2020). The next step was to apply the linear regression model to obtain the mathematical formula for the connection between unemployment rate (UR) and PRPSE and SPLS, respectively.

Using the forecast provided by the ARIMA model for 2021, 2022 and 2023 in the linear regression model previously computed, we obtained point prediction forecasts of PRPSE and SPLS for the same period of time.

Analyzing these results, we concluded that there is a strong connection between UR and PRPSE and SPLS, suggesting that our social life is highly impacted by the unemployment rate. All the obtained results were presented in tables and suggestive figures. Our proposed hybrid model can be successfully applied for other countries or regions, if data were previously collected. This research can be viewed as a helpful tool in the implementation of government policies in order to minimize the negative effects of unemployment on the population. One of the expansions of our work may be the analysis of the implementation level of the Sustainable Development Goals of the UN 2030 Agenda [44].

Some new studies should be conducted involving other factors that can have a great impact on social life, such as the current situation of the Covid-19 pandemic, which can affect the economic indicators of this period, but also the social life indicators. The results thus obtained must be compared with these ones or better integrated. We can also use a multiple linear regression model for the second part of our hybrid model in order to find a better regression formula for the studied variables. Likewise, a gender-oriented study should be taken into consideration as it has been observed that, even though unemployed, women's life satisfaction scores are in general higher than men's, possibly due to women's role in the family.

Author Contributions: Writing–original draft, C.-I.P., I.-V.T., C.-C.D., G.S., C.P. and D.D. These authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

References

1. Ladislav, M.; Zsigmond, T.; Kovacs, A.; Baloghova, E. Unemployment and GDP relationship analysis in the Visegrad four countries. *J. Model. New Eur.* **2020**, *12/13*, 118–134.
2. Ciucu, S. A Quantitative Analysis of the Correlation between Unemployment and GDP in Transition Countries. In Proceedings of the International Conference New Trends in Modelling and Economic Forecast, Romanian Academy—Institute For Economic Forecasting & Nicolae Titulescu University of Bucharest—Faculty of Economic Sciences, Bucharest, Romania, 5–7 March 2014.
3. Chiripuci, B.C.; Constantin, M.; Popescu, M.F.; Scrieciu, A. The Socio-Economic Impact of Migration on the Labor Market in the Romanian Danube Region. *Sustainability* **2020**, *12*, 8654. [CrossRef]
4. Frey, B.S.; Stutzer, A. *Happiness and Economics: How the Economy and Institutions Affect Well-Being*; Princeton University Press: Princeton, NJ, USA, 2002.
5. Ahmed, S.; Rajput, A.E.; Sarirete, A.; Aljaberi, A.; Alghanem, O.; Alsheraigi, A. Studying Unemployment Effects on Mental Health: Social Media versus the Traditional Approach. *Sustainability* **2020**, *12*, 8130. [CrossRef]
6. Ryu, P.-M. Predicting the Unemployment Rate Using Social Media Analysis. *J. Inf. Process. Syst.* **2018**, *14*, 4904–4915.
7. Proserpio, D.; Counts, S.; Jain, A. The Psychology of Job Loss: Using social Media Data to Characterize and Predict Unemployment. In Proceedings of the 8th ACM Conference on Web Science (WebSci '16), Hannover, Germany, 22–25 May 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 223–232.
8. Strandh, M.; Hammarstrom, A.; Nilsson, K.; Nordenmark, M.; Russel, H. Unemployment, gender and mental health: The role of the gender regime. *Sociol. Health Illn.* **2013**, *35*, 649–665. [CrossRef] [PubMed]
9. Zhang, G.P. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing* **2003**, *50*, 159–175. [CrossRef]
10. Pai, P.F.; Lin, C.S. A hybrid arima and support vector machines model in stock price forecasting. *Omega* **2005**, *33*, 497–505. [CrossRef]
11. Chakraborty, T.; Chakraborty, A.K.; Biswas, M.; Banerjee, S.; Bhattacharya, S. Unemployment Rate Forecasting: A Hybrid Approach. *Comput. Econ.* **2021**, *57*, 183–201. [CrossRef]
12. Ho, S.L.; Xie, M. The use of ARIMA models for reliability forecasting and analysis. *Comput. Ind. Eng.* **1998**, *35*, 213–216. [CrossRef]
13. Brockwell, P.J.; Davis, R.A. *Introduction to Time Series and Forecasting*; Springer: New York, NY, USA, 2002.
14. Wang, Y.; Shen, Z.; Jiang, Y. Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China. *PLoS ONE* **2018**, *13*, e0201987. [CrossRef] [PubMed]
15. Benvenuto, D.; Giovanetti, M.; Vassallo, L.; Angeletti, S.; Ciccozzi, M. Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data Brief* **2020**, *29*, 105340. [CrossRef] [PubMed]
16. Khan, F.M.; Gupta, R. ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India. *J. Saf. Sci. Resil.* **2020**, *1*, 12–18. [CrossRef]
17. Ariyo, A.A.; Adewumi, A.O.; Ayo, C.K. Stock Price Prediction Using the ARIMA Model. In Proceedings of the UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, Cambridge, UK, 26–28 March 2014; pp. 106–112.
18. Fattah, J.; Ezzine, L.; Aman, Z.; El Moussami, H.; Lachhab, A. Forecasting of demand using ARIMA model. *Int. J. Eng. Bus. Manag.* **2018**, *10*, 1847979018808673. [CrossRef]
19. Smith, K.P.; Christakis, N.A. Social Networks and Health. *Annu. Rev. Sociol.* **2008**, *34*, 405–429. [CrossRef]
20. Barry, B. *Social Exclusion, Social Isolation and the Distribution of Income*; Centre for Analysis of Social Exclusion, London School of Economics: London, UK, 1998.
21. Amati, V.; Meggiolaro, S.; Rivellini, G.; Zaccarin, S. Social relations and life satisfaction: The role of friends. *Genus* **2018**, *74*, 7. [CrossRef]
22. Portes, A. Social capital: Its origins and applications in modern sociology. *Annu. Rev. Sociol.* **1998**, *24*, 1–24. [CrossRef]
23. Umberson, D.; Montez, J.K. Social Relationships and Health: A Flashpoint for Health Policy. *J. Health Soc. Behav.* **2010**, *51* (Suppl. 1), 54–66. [CrossRef] [PubMed]
24. Kronauer, M. Social exclusion and underclass—New concepts for the analysis of poverty. In *Empirical Poverty Research in a Comparative Perspective*; Andrej, H.-J., Ed.; Ashgate Publishing: Aldershot, UK, 1998; pp. 51–75.
25. Furaker, B. *On the Social Consequences of Unemployment*; European Parliament's Special Committee on the Financial, Economic and Social Crisis (CRIS): Brusseles, Belgium, 2010.

26. Sen, A.K. *Social Exclusion: Concept, Application and Scrutiny*; Social Development Papers No. 1; Asian Development Bank: Manila, Philippines, 2000.
27. Eurostat. People at Risk of Poverty or Social Exclusion by Age and Sex. 2021. Available online: https://ec.europa.eu/eurostat/databrowser/view/ILC|_PEPS01_custom_1170455/default/table?lang=en (accessed on 20 April 2021).
28. Kwilinski, A.; Vyshnevskyi, O.; Dzwigol, H. Digitalization of the EU Economies and People at Risk of Poverty or Social Exclusion. *J. Risk Financ. Manag.* **2020**, *13*, 142. [CrossRef]
29. Jenkins, S.P.; Cappellari, L. Summarizing multiple deprivation indicators. In *Inequality and Poverty Re-Examined*; Micklewright, J., Ed.; Oxford University Press: Oxford, UK, 2007; pp. 166–184.
30. Pohlan, L. Unemployment and social exclusion. *J. Econ. Behav. Organ.* **2019**, *164*, 273–299. [CrossRef]
31. Chen, W.H.; Hou, F. The Effect of Unemployment on Life Satisfaction: A Cross-National Comparison Between Canada, Germany, the United Kingdom and the United States. *Appl. Res. Qual. Life* **2019**, *14*, 1035–1058. [CrossRef]
32. Eurostat. Unemployment by Sex and Age—Monthly Data. 2021. Available online: https://ec.europa.eu/eurostat/databrowser/view/UNE_RT_M_custom_1166829/default/table?lang=en (accessed on 20 April 2021).
33. Shapiro, S.S.; Wilk, M.B. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* **1965**, *52*, 591–611. [CrossRef]
34. Razali, N.M.; Wah, Y.B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Liliefors and Anderson-Darling test. *J. Stat. Model. Anal.* **2011**, *2*, 21–33.
35. Cramer, D. *Fundamental Statistics for Social Research*; Routledge: London, UK, 1998.
36. Cramer, D.; Howitt, D. *The SAGE Dictionary of Statistics*; SAGE: London, UK, 2004.
37. Doane, D.P.; Seward, L.E. Measuring Skewness. *J. Stat. Educ.* **2011**, *19*, 1–18. [CrossRef]
38. Dickey, D.A.; Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **1979**, *74*, 427–431.
39. Cheung, Y.W.; Lai, K. Lag Order and Critical Values of the Augmented Dickey-Fuller Test. *J. Bus. Econ. Stat.* **1995**, *13*, 277–280.
40. Vrieze, S.I. Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol. Methods* **2012**, *17*, 228–243. [CrossRef]
41. Jahoda, M. Work, employment, and unemployment: Values, theories, and approaches in social research. *Am. Psychol.* **1981**, *36*, 184–191. [CrossRef]
42. Ljung, G.M.; Box, G.E.P. On a measure of a lack of fit in time series models. *Biometrika* **1987**, *65*, 297–303. [CrossRef]
43. Ljung, G.M. Diagnostic testing of univariate time series models. *Biometrika* **1986**, *73*, 725–732. [CrossRef]
44. Sustainable Development Goals of the UN 2030 Agenda. Available online: https://sdgs.un.org/goals (accessed on 23 August 2021).