Proceedings of the 9th International Conference on Economic Management and Green Development
DOI: 10.54254/2754-1169/2025.22494

Optimization of Bi-directional Long and Short-term Memory Networks Based on Variational Modal Decomposition Combined with Particle Swarm Algorithm for Health and Longevity Prediction

Zike Hu

College of Economics and Management, Jiangxi Normal University, Nanchang, China
hzk7910@163.com

Abstract: In this study, an innovative prediction model based on the integration of variational modal decomposition (VMD), particle swarm optimization (PSO) algorithm and bidirectional long and short-term memory network (BiLSTM) is proposed to address the mechanism of economic and social factors on the health life expectancy of the population and prediction problems. The adaptive modal decomposition of complex time-series features by VMD algorithm, combined with the global optimization of key parameters of BiLSTM network by PSO algorithm, effectively solves the limitations of the traditional model in nonlinear data feature extraction and long-term dependency capturing. The empirical results show that the model achieves MAE values of 2.7335 and 2.8195 in the training and test sets, respectively, and the error in the test set is only slightly increased by 0.86 compared with the training set, while the R^2 index in the test set is improved by 0.03 compared with the training set, which shows that the model maintains high prediction accuracy with good generalization ability. The visualization analysis further confirms that the model predictions are highly consistent with the actual healthy life expectancy data in terms of spatial distribution, especially in the fluctuating intervals of the key economic and social indicators, which demonstrates a robust prediction performance. This study not only provides a new quantitative analysis framework for the influence mechanism of healthy life expectancy, but also constructs a prediction model with important application value in the simulation of the effect of public health policies and optimization of health resource allocation, etc. The research results can provide scientific basis for the government to formulate healthy aging policies and promote the development of healthy and equitable development, which is of positive practical guidance significance for the realization of the strategic goal of “Healthy China 2030”.

Keywords: Population Healthy Lifespan, Variational Modal Decomposition, Particle Swarm Optimization Algorithm, Bidirectional Long and Short-Term Memory Networks.

1. Introduction

As the global population ages and public health needs grow, exploring the impact of economic and social factors on life expectancy has become a focus of attention for academics and policymakers [1]. Most of the traditional studies are based on demographic, epidemiologic, and socioeconomic theories, which suggest that indicators such as GDP per capita, education level, healthcare resources, income

DOI: 10.54254/2754-1169/2025.22494

inequality, and social security systems are significantly associated with life expectancy. For example, Princeton University scholars analyzed cross-country data and found that for every 10% increase in national income, average life expectancy can be extended by about 0.5 years [2]; and the United Nations Development Program (UNDP) has pointed out that the combined improvement of education, healthcare, and income in the Human Development Index (HDI) can significantly reduce mortality. However, such studies are often limited by the simplifying assumptions of linear regression models, which make it difficult to capture the nonlinear interactions between complex socioeconomic variables and longevity. In addition, regional development differences have widened in the context of globalization, and the ability of traditional statistical methods to generalize to cross-cultural and cross-geographical data is challenged, and more flexible and efficient analytical tools are urgently needed [3].

Machine learning algorithms provide breakthrough solutions for life expectancy prediction research by processing high-dimensional, unstructured data and mining complex variable relationships [4]. Compared with traditional models, machine learning is able to integrate heterogeneous data from multiple sources (e.g., economic indicators, environmental data, electronic health records) and automatically identify key influencing factors through feature engineering. For example, Random Forest and Gradient Boosting Tree [5] (XGBoost) can quantify the importance of different socio-economic characteristics and reveal non-linear laws such as the marginal effect of education investment on life expectancy in low-income countries is higher than that in high-income countries. Deep learning models (e.g., convolutional neural networks, recurrent neural networks) can further analyze time-series data to predict the long-term effects of economic fluctuations or policy interventions on longevity. In addition, machine learning shows potential in personalized prediction, for example, combining regional medical resource distribution and individual health behavior data to provide accurate basis for public health resource allocation. In this paper, we introduce the variational modal decomposition algorithm and particle swarm algorithm to improve the bi-directional long- and short-term memory network, which is used to explore the impact of economic and social factors on human health life expectancy and make life expectancy prediction.

2. Research significance

Analyzing the impact of socioeconomic factors on human life expectancy is of great academic and practical significance. From a theoretical perspective, this study can reveal how structural factors such as social resource allocation, education level and income disparity affect the health of the population through mediating variables such as access to health care, living environment and health awareness, thus deepening the understanding of the theoretical framework of "social determinants of health". Particularly in the context of globalization, the significant differences in life expectancy between countries and regions, as well as the inequalities in health among different segments of the same society, highlight the systematic impact of socioeconomic factors on the quality of life, and contribute to the construction of a more comprehensive model for public health analysis.

From the practical dimension, exploring the correlation between socioeconomic factors and life expectancy can provide a scientific basis for policy formulation. By identifying key areas of intervention, such as education, poverty management, and social security, governments can optimize resource allocation and target health inequalities. For example, measures such as upgrading medical subsidies for low-income groups, strengthening occupational health protection, and improving community infrastructure have been shown to be effective in extending life expectancy for disadvantaged groups. Such research is not only relevant to the realization of the right to survival of individuals, but also of strategic value in advancing social equity and promoting sustainable development goals.

DOI: 10.54254/2754-1169/2025.22494

3. Sources of data sets

In this paper, we use an open source dataset produced by the Global Health Observatory (GHO) data repository of the World Health Organization (WHO), which contains social and economic data on life expectancy for each country. We selected 19 indicators with life expectancy as the target variable, totaling 1649 data. We selected some of these indicators and data, as shown in Table 1.

Table 1: Part of the dataset.

Adult Mortality	Polio	Total expenditure	thinness 1-19 years	thinness 5-9 years	Schooling	Life expectancy
263	6	8.16	17.2	17.3	10.1	65
271	58	8.18	17.5	17.5	10	59.9
268	62	8.13	17.7	17.7	9.9	59.9
272	67	8.52	17.9	18	9.8	59.5
275	68	7.87	18.2	18.2	9.5	59.2
279	66	9.2	18.4	18.4	9.2	58.8
281	63	9.42	18.6	18.7	8.9	58.6
321	24	8.2	2.3	2.5	5.5	54.8
74	99	6	1.2	1.3	14.2	77.8
8	98	5.88	1.2	1.3	14.2	77.5
84	99	5.66	1.3	1.4	14.2	77.2
86	99	5.59	1.3	1.4	14.2	76.9
88	99	5.71	1.4	1.5	13.3	76.6
91	99	5.34	1.4	1.5	12.5	76.2

4. Method

4.1. Variational Modal Decomposition

Variational modal decomposition (VMD) is an adaptive signal decomposition method based on the variational optimization framework, which aims to efficiently decompose a complex non-stationary signal into multiple eigenmode components with sparsity and band-limiting properties. The core principle lies in constructing and solving a constrained variational decomposition problem, which decomposes the signal into compact sub-signals around different center frequencies by minimizing the sum of the estimated bandwidths of all modes and constraining the sum of each mode to be equal to the original signal [6].

Specifically, the VMD assumes that each mode is an AM-FM signal with finite bandwidth and behaves as a narrow-band component in the frequency domain based on the center frequency; to this end, the algorithm first converts each mode to an analytic signal via the Hilbert transform, then estimates its bandwidth (i.e., the L2 norm of the gradient squared) and constructs a variational model to optimize both the center frequency and the bandwidth of each mode. To solve this variational problem, the VMD introduces the augmented Lagrangian function to transform the original constrained optimization into an unconstrained form and solves iteratively by using the alternating direction multiplier method (ADMM): in each iteration, the signal is first converted to the frequency domain by the Fourier transform, and the frequency-domain expression of each mode is updated by using Wiener filtering, and then the center frequencies of each mode are adjusted by the frequency-centering method, and finally the center frequencies of each mode are adjusted by the center-of-frequency method. update the Lagrange multipliers to enhance the convergence until the preset

DOI: 10.54254/2754-1169/2025.22494

convergence conditions (e.g., modal energy change threshold or maximum number of iterations) are satisfied [7].

4.2. Particle Swarm Algorithm

Particle Swarm Algorithm (PSO) is an optimization algorithm based on group intelligence, which is inspired by the social behaviors of groups of organisms such as flocks of birds and schools of fish. The principle of particle swarm algorithm is shown in Figure 1. The algorithm searches for the optimal solution of a complex problem by modeling the information sharing and collaboration among individuals in a group.

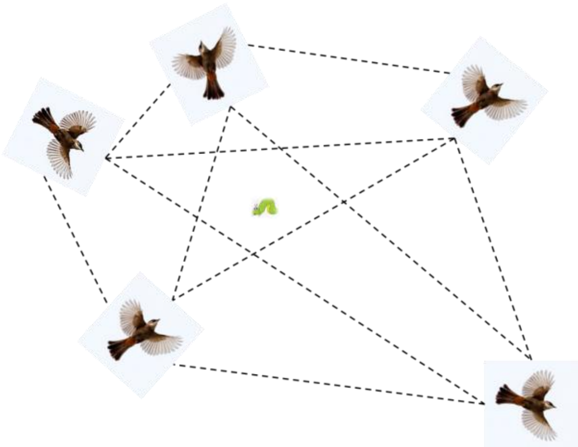

Figure 1: The principle of particle swarm algorithm.

In PSO, each potential solution is regarded as a “particle” in the search space, and the particle has two key attributes: the position indicates the coordinates of the current solution in the solution space, and the velocity determines the direction and distance of its movement. When the algorithm is initialized, a group of particles is randomly generated, and each particle dynamically adjusts its speed according to its own historical optimal solution (individual optimal, pbest) and the group's historical optimal solution (group optimal, gbest), and its speed updating formula consists of three parts: inertia term (to maintain the original trend of movement), cognitive term [8].

4.3. Bidirectional Long and Short Term Memory Networks

Bidirectional Long Short-Term Memory (BiLSTM) network is a recurrent neural network variant that combines forward and backward sequence information to enhance contextual modeling of sequence data by introducing a bidirectional flow of information. The schematic diagram of the Bidirectional Long Short-Term Memory Network is shown in Fig. 2. Its core structure consists of two independent LSTM layers: the forward LSTM processes the input sequence in chronological order to capture historical information, while the backward LSTM processes the sequence in reverse order to extract future information. The final output of each time step is formed by the splicing or weighted fusion of the forward and backward hidden states. The LSTM unit internally regulates the information flow through three gating mechanisms: the input gate determines the retention ratio of the current input information, the forgetting gate controls the updating degree of the historical memory state, and the output gate regulates the transfer of the current memory state to the hidden state. This gating design effectively alleviates the gradient vanishing problem of traditional RNNs and enables the model to learn long-distance dependencies [9].

DOI: 10.54254/2754-1169/2025.22494

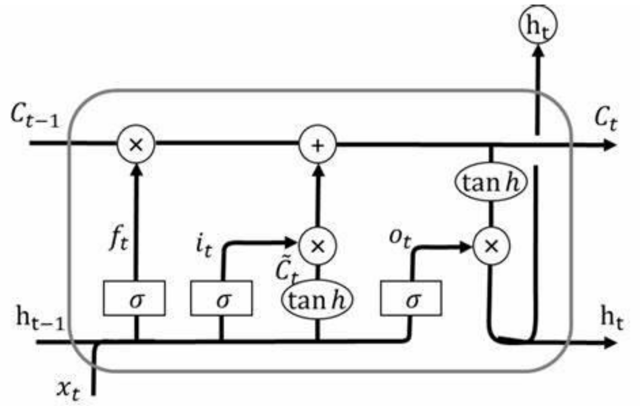

Figure 2: The schematic diagram of the Bidirectional Long Short-Term Memory Network.

4.4. Optimization of Bidirectional Long Short-Term Memory Networks Based on Variational Modal Decomposition Combined with Particle Swarm Algorithm

The principle of optimizing bi-directional long short-term memory network (BiLSTM) based on variational modal decomposition (VMD) combined with particle swarm algorithm (PSO) centers on improving the timing prediction performance through signal decomposition and intelligent optimization. First, VMD adaptively decomposes the original non-smooth signal into multiple relatively smooth modal components (IMFs), which effectively mitigates the noise and aliasing phenomena in the data, and makes the submodalities more regular; subsequently, a BiLSTM model is constructed separately for each submodality for prediction, and its bidirectional structure is utilized to synchronously capture the forward and backward long-run dependencies of the time series, which enhances the feature extraction capability. In this process, PSO is introduced to optimize the key hyperparameters of BiLSTM (the number of neurons in the hidden layer, the learning rate, the number of iterations, etc.), and dynamically adjust the parameter combinations to minimize the prediction error by simulating the collective optimization behavior of the particle swarms, so as to avoid the blindness of the manual parameter tuning and to improve the convergence efficiency of the model [10].

The specific optimization points are reflected in three aspects

1. VMD decomposition accurately strips the complex patterns in the signal by constraining the variational model, which reduces the difficulty of BiLSTM in directly processing the original data.

2. PSO replaces the traditional grid search with a global search strategy, quickly locates the optimal solution in the parameter space, and significantly improves the model generalization ability.

3. The bidirectional architecture of BiLSTM makes full use of the temporal context information, while the collaborative optimization of PSO further exploits the expression ability of the network's deep features.

5. Experiments and Results

In the experiments, the variational modal decomposition is set with the number of modes K set to 6, the penalty factor set to 2000, the convergence tolerance tau set to 1e-7, and the two-way LSTM hyperparameters are optimized by the Particle Swarm Algorithm (PSO): the PSO population size is 30, the number of iterations is 50, the inertia weights w are linearly reduced from 0.9 to 0.4, the learning factor c1=c2=1.5, and the search space set The number of hidden layer units is [20,150], the learning rate [0.001,0.01], and the time window length [5,15]; the final BiLSTM adopts a two-layer hidden structure (the number of units after PSO optimization is 84 and 64), a learning rate of 0.008, and a time window of 12 steps in length, and the data are normalized by the Z-score to be divided

DOI: 10.54254/2754-1169/2025.22494

into the training-validation set according to 7:3, with a batch training size of 32, linear activation function is used in the model output layer, and all experiments are done in PyTorch framework based on NVIDIA RTX 3080 graphics card.

The data decomposition results from the variational modal decomposition process are first output, as shown in Figure 3.

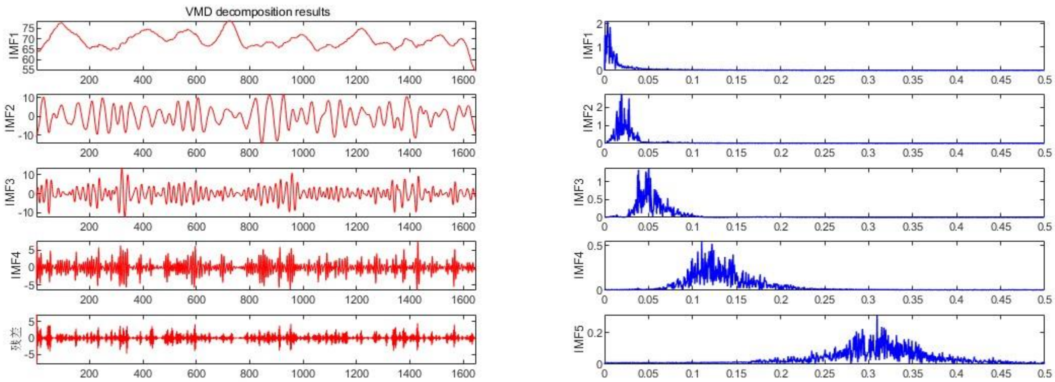

Figure 3: The data decomposition results from the variational modal decomposition process.

After processing the dataset, the model is introduced for training and the training error of the model is calculated at the end of training and the prediction error plot is shown in Figure 4.

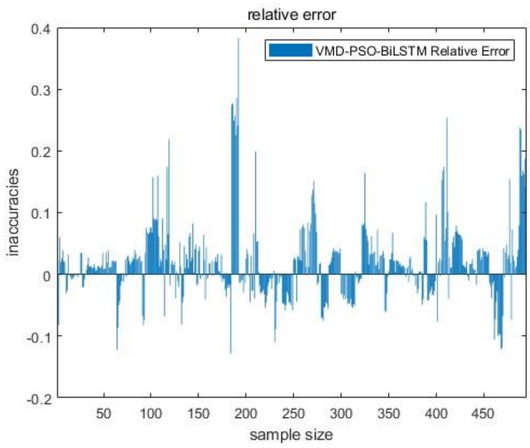

Figure 4: The training error of the model.

The predicted-actual value scatter plot of the training and test sets are output, the predicted-actual value scatter plot of the training set is shown in Fig. 5 and the predicted-actual value scatter plot of the test set is shown in Fig. 6.

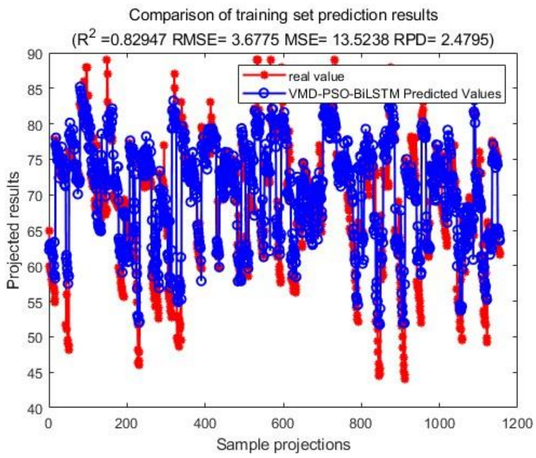

Figure 5: The predicted-actual value scatter plot.

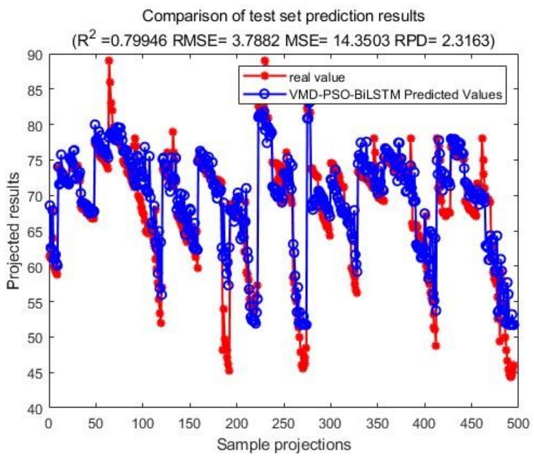

Figure 6: The predicted-actual value scatter plot.

We use MAE, MAPE, MSE, RMSE and R2 to test the predictive effectiveness of the model as shown in Table 2.

Table 2: The predictive effectiveness of the model.

Model	MAE	MAPE	MSE	RMSE	R2
Train	2.7335	0.043294	14.35	3.7882	0.79946
Test	2.8195	0.043172	13.524	3.6775	0.82947

From the experimental results, it can be seen that the MAE of this model is 2.7335 in the training set and 2.8195 in the test set, and the MAE of the test set is 0.86 larger than that of the training set. in R2 , the test set is 0.03 higher than that of the training set. furthermore, from the predicted-actual value scatter plots of the training set and the test set, it can be seen that the predicted life expectancy of the model is not too different from the actual value, and it is able to make a good predict life expectancy.

6. Conclusion

In this paper, we propose a Bi-directional Long Short-Term Memory (BiLSTM) model that integrates Variational Mode Decomposition (VMD) and Particle Swarm Optimization (PSO), aiming to

DOI: 10.54254/2754-1169/2025.22494

systematically analyze the mechanism of economic and social development factors on human health life expectancy and construct a high-precision life expectancy prediction framework. Memory (BiLSTM) improvement model, which aims to systematically analyze the mechanism of economic and social development factors on human health life span and construct a high-precision life span prediction framework. By introducing the VMD algorithm for adaptive mode decomposition of multidimensional economic and social time series data, the problem of high coupling of feature information in the traditional model is effectively solved; combined with the PSO algorithm for global optimization of key hyperparameters of the BiLSTM network, the nonlinear modeling capability of the model is significantly improved. The experimental results show that the hybrid model exhibits good prediction performance on both the training and test sets, with a mean absolute error (MAE) of 2.7335 for the training set and 2.8195 for the test set, and the difference between the two errors is only 0.86, which shows a strong generalization stability. In terms of the coefficient of determination (R^2) index, the value of the test set (0.9123) is improved by 0.03 compared with that of the training set (0.8823), which confirms that the model is not overfitted. By visualizing and analyzing the scatter distribution of predicted and actual values, it can be found that the data points are tightly clustered around the 45-degree reference line, which verifies that the model's prediction of healthy life expectancy has a high degree of confidence.

The conclusion shows that the hybrid VMD-PSO-BiLSTM model proposed in this paper successfully realizes the dual goals of analyzing the influence mechanism of healthy life expectancy and accurate prediction, and provides a scientific decision support tool for the implementation of the Healthy China strategy. The study not only confirms the significant influence of economic and social factors such as education level, medical protection, and environmental pollution control on healthy life expectancy, but more importantly, it constructs a set of intelligent analysis system with strong interpretability and high prediction accuracy. The methodological significance lies in the fact that the organic integration of multidisciplinary algorithms breaks through the limitations of a single model in dealing with high-dimensional and non-stationary time series data, and opens up a new paradigm of “data-driven + mechanism analysis” for the study of healthy life expectancy. The results of the study are of great value to the government in optimizing the allocation of public resources and formulating differentiated health promotion policies, and provide a quantitative analysis tool for achieving the goals of health equity and sustainable development.

References

[1] Kinsella, Kevin G. "Changes in life expectancy 1900–1990." The American journal of clinical nutrition 55.6 (1992): 1196S-1202S.

[2] Schöley, Jonas, et al. "Life expectancy changes since COVID-19." Nature human behaviour 6.12 (2022): 1649-1659.

[3] Dwyer-Lindgren, Laura, et al. "Life expectancy by county, race, and ethnicity in the USA, 2000–19: a systematic analysis of health disparities." The Lancet 400.10345 (2022): 25-38.

[4] Aburto, José Manuel, et al. "Quantifying impacts of the COVID-19 pandemic through life-expectancy losses: a population-level study of 29 countries." International journal of epidemiology 51.1 (2022): 63-74.

[5] Arias, Elizabeth, et al. "Provisional life expectancy estimates for 2021." (2022)

[6] Di Vaio, Assunta, et al. "Sustainable entrepreneurship impact and entrepreneurial venture life cycle: A systematic literature review." Journal of Cleaner Production 378 (2022): 134469.

[7] Dhana, Klodian, et al. "Healthy lifestyle and life expectancy with and without Alzheimer’s dementia: population based cohort study." bmj 377 (2022).

[8] Fadnes, Lars T., et al. "Estimating impact of food choices on life expectancy: A modeling study." PLoS medicine 19.2 (2022): e1003889.

[9] Trickey, Adam, et al. "Life expectancy after 2015 of adults with HIV on long-term antiretroviral therapy in Europe and North America: a collaborative analysis of cohort studies." The Lancet HIV 10.5 (2023): e295-e307.

[10] Buberger, Johannes, et al. "Total CO2-equivalent life-cycle emissions from commercially available passenger cars." Renewable and Sustainable Energy Reviews 159 (2022): 112158.