Reports on Economics and Finance, Vol. 10, 2025, no. 1, 1 - 11
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/ref.2025.91421

Uncertain Autoregressive Model with External Explanatory Variables for the CNY Exchange Rate Prediction

Zhou Fang and Yuanguo Zhu

School of Mathematics and Statistics
Nanjing University of Science and Technology
Nanjing 210094, Jiangsu, China

Abstract

Uncertain time series and uncertain regression analysis are influential methods for exploring relationships between variables. As basic statistical models, these two methods cannot deal with the situation where the current observation value is simultaneously affected by past observations and external explanatory variables. This inspires us to propose an uncertain autoregressive model includes external explanatory variables. First, this paper presents an uncertain autoregressive model with external explanatory variables, and investigates its parameter estimates, predicted values, and confidence intervals. Then, the model is applied to studying CNY exchange rate data, using the gold parity exchange rate as the external explanatory variable. Finally, through comparative analysis, it is shown that the proposed model has better predictive performance than the traditional uncertain autoregressive model.

Mathematics Subject Classification: 62J86 91B84

Keywords: Uncertainty theory; Uncertain time series analysis; Uncertain autoregressive model; Gold parity exchange rate; CNY exchange rate forecasting

1 Introduction

Uncertain statistics is an important part of uncertainty theory, which is a mathematical technique for collecting, analyzing, and interpreting data using uncertainty theory[8]. Uncertainty regression analysis, as a branch of uncertain statistics, is a statistical technique that uses uncertainty theory to analyze the relationship between explanatory variable and response variable. Yao and Liu[14] established this method, assuming that the interference term is an uncertain variable rather than a random variable. In order to perform point estimation on unknown parameters in uncertain regression models, they proposed least squares estimation. Liu and Yang[9] considered minimum absolute deviation estimation, Chen[1] studied Tukey's dual weight estimation, and Lio and Liu[6] proposed maximum likelihood estimation. Lio and Liu[7] further proposed interval estimation for predicting response variables based on uncertain interference terms. To evaluate the appropriateness of fitting regression models and estimating interference terms, Ye and Liu[15] introduced uncertain hypothesis testing.

Uncertain time series analysis is another branch of uncertain statistics, which is a statistical method that predicts future values based on previous observations on the basis of uncertainty theory. Liu and Yang [13] assumed that the interference term was an uncertain variable rather than a random variable. They proposed an uncertain autoregressive (UAR) model. In addition, they applied the principle of least squares to estimate unknown autoregressions parameters. Then, Hu et al.[5] considered minimum absolute deviation estimation, while Chen and Yang[2] studied maximum likelihood estimation. To determine the optimal order of the UAR model, Liu and Yang[10] conducted cross validation. In addition, some researchers have also studied other uncertain time series models, such as the first-order uncertain moving average model[11], the uncertain vector autoregression model[12], and the uncertain autoregressive moving average (UARMA) model[4].

Currently, uncertain regression models are mostly used to analyze the relationship between explanatory variable and response variable, while autoregressive models analyze the relationship between future and past values of observed values. However, there are also some special data whose predicted values need to be analyzed not only by considering the influence of their past values, but also by taking into account the influence of closely related exogenous variables. Therefore, this paper establishes an uncertain autoregressive model with external explanatory variables. The main method is to combine the uncertain autoregressive model with external explanatory variables to propose a new regression model and apply it to CNY exchange rate prediction. This paper is arranged as follows: Section 2 presents an uncertain autoregressive model with external explanatory variables(UARX model), and presents its parameter estimates, predicted values, and confidence intervals. In Section 3 , the model is applied to actual CNY exchange rate data, using the gold parity exchange rate as the external explanatory variable. Through comparative analysis, it is proven that the UARX model has better predictive performance than the UAR model. The last section makes a conclusion.

2 Uncertain autoregressive model with exogenous variable

Assume x_t is the observed value at time t $(t = 1, 2, \cdots, n)$. Then the sequence of observed values $x_1, x_2, \cdots, x_n$ is a time series. A basic problem of uncertain time series analysis is to predict the value of x_{n+1} based on the previously observed values x_1, x_2, $\cdots$, x_n.

Assume y_t is an external explanatory variable $(t = 1, 2, \cdots, n)$. The functional relationship between y_t and $(x_1, x_2, \cdots, x_n)$ is assumed to be expressed by the regression model

$$x_t = a_0 + \sum_{i=1}^{k} a_i x_{t-i} + \beta y_t + \varepsilon_t, \tag{1}$$

where k is the order of the uncertain autoregressive model, a_0, a_1, $\cdots$, a_k, β are unknown parameters, and ε_t is an uncertain disturbance term (uncertain variable). When β=0, the model is the traditional uncertain autoregressive model.

Parameter Estimation

Based on the observed values x_1, x_2, $\cdots$, x_n, the least squares estimate of a_0, a_1, $\cdots$, a_k, β in the uncertain autoregressive model (1) is the solution of the minimization problem

$$\min_{a_0, a_1, \cdots, a_k, \beta} \sum_{t=k+1}^{n} (x_t - a_0 - \sum_{i=1}^{k} a_i x_{t-i} - \beta y_t)^2. \tag{2}$$

If the minimization solution of (2) is $(\hat{a}_0,\ \hat{a}_1,\ \cdots,\ \hat{a}_k,\ \hat{\beta})$, then the fitted autoregressive model is

$$x_t = \hat{a}_0 + \sum_{i=1}^{k} \hat{a}_i x_{t-i} + \hat{\beta} y_t. \tag{3}$$

Residual Analysis

For each index t $(t = k+1, k+2, \cdots, n)$, the difference between the actual observed value and the value predicted by the model (3),

$$\varepsilon_t^* = x_t - \hat{a}_0 - \sum_{i=1}^{k} \hat{a}_i x_{t-i} - \hat{\beta} y_t \tag{4}$$

is called the t-th residual. If the uncertain disturbance terms ε_{k+1}, ε_{k+2}, $\cdots$ are assumed to be independent and identically distributed uncertain variables (hereafter called iid hypothesis), then the expected value of uncertain disturbance terms can be estimated as the average of residuals, i.e.,

$$\hat{e} = \frac{1}{n-k} \sum_{t=k+1}^{n} \varepsilon_t^*, \tag{5}$$

and the variance can be estimated as

$$\hat{\sigma}^2 = \frac{1}{n-k} \sum_{t=k+1}^{n} (\varepsilon_t^* - \hat{e})^2, \tag{6}$$

where ε_t^* are the t-th residuals, $t = k+1, k+2, \cdots, n$. Therefore, for each t, the estimated disturbance term $\hat{\varepsilon}_t$ is an uncertain variable with expected value $\hat{e}$ and variance $\hat{\sigma}^2$.

If ε_t is assumed to obey the uncertain normal distribution, an uncertain autoregressive model is obtained

$$\hat{x}_{n+1} = \hat{a}_0 + \sum_{i=1}^{k} \hat{a}_i x_{n+1-i} + \hat{\beta} y_{n+1} + \hat{\varepsilon}_{n+1}, \;\; \hat{\varepsilon}_{n+1} \sim \mathcal{N}(\hat{e}, \hat{\sigma}). \tag{7}$$

Uncertain Hypothesis Test

According to the uncertain hypothesis test proposed by Ye and Liu[15], in order to test whether the uncertain autoregressive model (7) fits the observed data well, we should test whether the normal uncertainty distribution $\mathcal{N}(\hat{e}, \hat{\sigma})$ fits the $n-k$ residuals, i.e.,

$$\varepsilon_{k+1}^*, \varepsilon_{k+2}^*, \cdots, \varepsilon_n^* \sim \mathcal{N}(\hat{e}, \hat{\sigma}). \tag{8}$$

Consider the following two hypotheses:

$$H_0 : e = \hat{e} \text{ and } \sigma = \hat{\sigma} \text{ versus } H_1 : e \neq \hat{e} \text{ or } \sigma \neq \hat{\sigma}. \tag{9}$$

Given a significance level α (e.g. 0.05), the rejection domain can be provided by

$$W = \Big\{(z_{k+1}, z_{k+2}, \cdots, z_n): \text{ there are at least } \alpha \text{ of indexes } i's \text{ with } k+1 \le i \le n \text{ such that } z_i < \Phi^{-1}(\frac{\alpha}{2}) \text{ or } z_i > \Phi^{-1}(1-\frac{\alpha}{2})\Big\}, \tag{10}$$

where $\Phi^{-1}(\alpha)$ is the inverse uncertainty distribution of $\mathcal{N}(\hat{e}, \hat{\sigma})$, i.e.,

$$\Phi^{-1}(\alpha) = \hat{e} + \frac{\hat{\sigma}\sqrt{3}}{\pi} \ln \frac{\alpha}{1-\alpha}.$$

For the $n-k$ residuals ε^*_{k+1}, ε^*_{k+2}, $\cdots$, ε^*_n, if

$$(\varepsilon^*_{k+1}, \varepsilon^*_{k+2}, \cdots, \varepsilon^*_n) \in W,$$

then we reject H_0. That is, the uncertain autoregressive model (7) is not a good fit to the observed data. In this case, we have to re-choose an uncertain autoregressive model. If

$$(\varepsilon^*_{k+1}, \varepsilon^*_{k+2}, ..., \varepsilon^*_n) \notin W,$$

then we accept H_0. That is, the uncertain autoregressive model (7) is a good fit to the observed data.

Forecast Uncertain Variable

The forecast value of x_{n+1} is defined as the expected value of the uncertain variable $\hat{x}_{n+1}$, i.e.,

$$\hat{\mu} = \hat{a}^*_0 + \sum_{i=1}^{k} \hat{a}^*_i x_{n+1-i} + \hat{\beta}^* y_{n+1} + \hat{e}. \tag{11}$$

And the α confidence interval of x_{n+1} is

$$\hat{\mu} \pm \frac{\hat{\sigma}\sqrt{3}}{\pi} \ln \frac{1+\alpha}{1-\alpha}. \tag{12}$$

3 Empirical analysis

In this section, we use exchange rate data as an example to demonstrate that the proposed model in the previous section is better than traditional autoregressive models in prediction.

Yang and Fang [3] proposed the gold parity CNY exchange rate, the construction of the underlying choice of spot gold, using the same point in time of the international spot price of gold and the domestic spot price of gold to define the ratio, which can be expressed as:

$$CNY^{GP} = \frac{G^D}{G^I}, \tag{13}$$

where CNY^{GP} represents the gold parity CNY exchange rate, and G^D and G^I represent the domestic gold price and international gold price, respectively. In accordance with the theory of purchasing power parity, after deducting the relevant costs the immediate exchange rate adjustment of the domestic gold price should be equal to the foreign gold price, but in the actual situation there is a certain deviation, and the deviation causes a certain arbitrage space.

In the paper we selects the morning fixing price of the London gold market as the international gold spot price, because the morning fixing price can be matched with the closing time of the domestic gold market. "Shanghai Gold" Au9995 from the Shanghai Gold Exchange, its gold purity in 99.95 or more, is similar to the quality of London standard gold spot, so we select it as the domestic gold spot price. London gold is priced in "dollars/oz", Shanghai gold is priced in "yuan/g", before calculating CNY^{GP}, the London gold price unit is converted to "dollars/g".

In Fig. 1, we show the weekly frequency gold price data and the weekly closing price of CNY exchange rate in 2020 ∼ 2023 (excluding holidays). It is

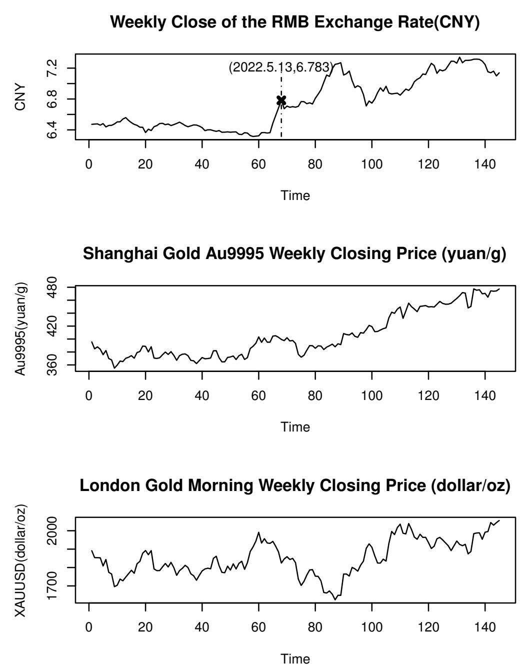

Figure 1: The price plot for 2020∼2023

obvious that the CNY has been fluctuating up and down at 6.4 before May 13, 2022 due to the influence of the international situation and COVID-19. After May 13, 2022, the CNY rose to a new level of $6.8 \sim 7.4$, and the overall level was higher than before, and the price of Shanghai gold also gradually increased. Considering that the CNY exchange rate is greatly affected by short-term events, we selects some data from May 13, 2022 to December 29, 2023 (excluding holidays and non-trading days) as the model fitting objects.

According to the formula (13), we calculate the weekly frequency CNY^{GP} and plot the trend shown in Fig. 2. The long-term trend of them is basically same and the fluctuation of CNY^{GP} is slightly larger than the spot CNY exchange rate. In addition, it can also be seen that most of the time the deviation of CNY from CNY^{GP} is negative, that is to say, there is a tendency for the CNY to depreciate from the exchange rate in the gold parity.

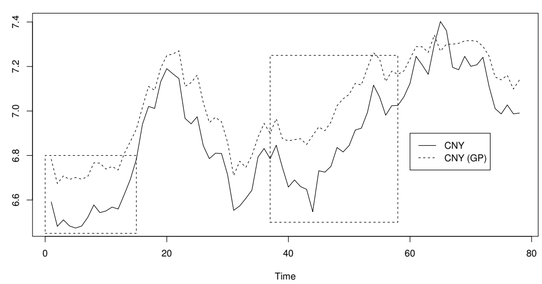

Figure 2: Deviation between CNY and CNY^{GP}

The weekly CNY closing prices from May 13, 2022 to December 29, 2023 (excluding holidays and non-trading days) in Table 1 are chosen as observations, denoted by X_t with $t = 1, 2, \cdots, 78$, and the following uncertain autoregressive model (UARX) is constructed:

$$CNY_t = \alpha_0 + \sum_{i=1}^{k} \alpha_i CNY_{t-i} + \varepsilon_t + \beta CNY_t^{GP},\ \varepsilon_t \sim \mathcal{N}(\hat{e}, \hat{\sigma}), \tag{14}$$

where k is the order, ε_t is an uncertain variable. The parameters to be estimated are β, α_i $(i = 1, 2, \cdots, k)$. In addition, when $\beta = 0$, it is the uncertain model without exogenous variables, i.e., the UAR model, which is constructed from a single series of CNY. Let $k = 2$, it is the 2nd order uncertain autoregressive model. The parameters are estimated by (2) as follows in Table 2.

For the UARX model, residuals $\varepsilon_1, \cdots, \varepsilon_{78}$ are obtained from the observations based on (4). The expectation and the variance of the uncertain distur-

Table 1: The weekly CNY closing prices from May 13, 2022 to December 29, 2023(excluding holidays and non-trading days)

6.7830	6.6740	6.7081	6.6927	6.7010	6.6936	6.7051	6.7670
6.7657	6.7390	6.7507	6.7349	6.8088	6.8621	6.9192	7.0166
7.1104	7.0931	7.1931	7.2494	7.2565	7.2703	7.1106	7.1275
7.1615	7.0380	6.9480	6.9716	6.9514	6.8588	6.7099	6.7740
6.7465	6.7999	6.8818	6.9442	6.9002	6.9653	6.8765	6.8655
6.8713	6.8754	6.8495	6.8898	6.9284	6.9114	6.9506	7.0235
7.0547	7.0750	7.1241	7.1168	7.1938	7.2620	7.2336	7.1325
7.1790	7.1645	7.1802	7.2340	7.2896	7.2884	7.2633	7.3415
7.2691	7.3002	7.3002	7.3040	7.3150	7.3166	7.3133	7.2906
7.2465	7.1529	7.1400	7.1607	7.0987	7.1393		

Table 2: Parameters for UARX and UAR Models

	α_0	α_1	α_2	β
UARX	0.9064	0.3501	0.0306	0.5011
UAR	0.0885	0.9274	0.0612	-

bance term ε are estimated as

$$\hat{e} = \frac{1}{78}\sum_{t=1}^{78}\varepsilon_t = 0.00004,\ \hat{\sigma}^2 = \frac{1}{78}\sum_{t=1}^{78}(\varepsilon_t - \hat{e})^2 = 0.0013. \tag{15}$$

The expectation and variance of the uncertain disturbance term for the UAR model were obtained in the same way as -0.009 and 0.0559. Thus, we have UARX model

$$\begin{aligned} CNY_t = &\, 0.9064 + 0.3501 CNY_{t-1} + 0.0306 CNY_{t-2} + 0.5011 CNY_t^{GP} + \varepsilon_t, \\ &\varepsilon_t \sim \mathcal{N}(0.00004, 0.0013). \end{aligned} \tag{16}$$

and UAR model

$$\begin{aligned} CNY_t = &\, 0.0885 + 0.9274 CNY_{t-1} + 0.0612 CNY_{t-2} + \varepsilon_t, \\ &\varepsilon_t \sim \mathcal{N}(-0.009, 0.0559). \end{aligned} \tag{17}$$

Uncertainty hypothesis test is used to determine if the models fit the observed data well. The rejection domain W can be expressed as (10) with $\alpha = 0.05$ and $\Phi^{-1}(\alpha) = -0.07266$, $\Phi^{-1}(1-\alpha) = 0.07274$.

In Table 3, we show the 76 residuals for model (16). It can be seen that only 3 residuals are not in $[-0.07266, 0.07274]$, the residuals do not belong to W, so the model (16) fits the observed data. The same way tests that model (17) is also well fitted.

Table 3: The residuals for model (16)

−0.00503	−0.01524	0.00214	−0.01217	−0.01773	0.01216
0.00620	−0.02572	−0.01333	−0.02810	0.01856	0.01242
0.00299	0.00003	0.01788	−0.03086	0.01209	0.00466
0.00051	0.02093	−0.05430	0.03048	0.04734	−0.02345
−0.04191	0.00481	−0.02037	−0.05859	−0.09444	0.01467
−0.04765	−0.00507	−0.01560	−0.00286	−0.04818	0.00008
−0.06078	0.00217	−0.00160	0.01587	−0.00541	0.09410
0.02725	−0.00149	0.02951	0.04679	0.06126	0.05376
0.06009	0.03080	0.07327	0.05337	0.02551	−0.02684
0.03459	0.000625	0.00627	0.02480	0.00013	−0.00282
−0.00808	0.01132	−0.13904	−0.06428	0.00956	0.01779
−0.00256	0.01737	0.00980	−0.02872	0.00065	−0.02669
0.00658	0.01432	−0.03441	0.02553		

Based on the models (16) and (17), the forecast value and the real value error of the CNY exchange rate in the next 3 weeks are obtained by substituting the data, as shown in Table 4. Notice that the mean squared error (MSE) is defined as

$$MSE = \frac{1}{3}\sum_{t=1}^{3}(X_t - \hat{X}_t)^2$$

It can be seen that the mean squared error(MSE) by (16) of the forecast results of the UARX model is smaller, so we conclude that the UARX model has a higher forecast accuracy compared to the UAR model.

Table 4: Prediction Results

	Forecast Period	$\hat{X}_t$	X_t	Bias	MSE
	1	7.1792	7.1671	0.0121	
UARX	2	7.1944	7.1777	0.0167	0.000142
	3	7.1992	7.1999	−0.0007	
	1	7.1349	7.1671	−0.0322	
UAR	2	7.1332	7.1777	−0.0445	0.002574
	3	7.1313	7.1999	−0.0686	

Remark 1. *The Shapiro-Wilk test on the residuals for Models (16) and (17) shows that the p-values are* 0.0063 *and* 0.0033*, respectively, and significantly less than* 0.05*, indicating that the disturbance term is not randomly normally distributed and cannot be considered as a random variable. So it is more appropriate to use the uncertain time series analysis than the probabilistic time series analysis in studying the corresponding problem.*

4 Conclusion

This paper proposes an uncertain regression model that combines uncertain autoregression model with external variables and gives its parameter estimates, predicted values, and confidence intervals. By fitting the actual data, we obtained the predicted value of the RMB exchange rate for the next three weeks, and compared the MSE of the predicted value with the real value. It turned out that that the UARX model has better predictive performance for CNY exchange rate than the UAR model. In the future, we will try to explore the combination of uncertain time series analysis with other methods, and choose methods with different characteristics for different economic data.

References

[1] Chen D, Tukeys biweight estimation for uncertain regression model with imprecise observations, *Soft Computing*, **24** (2020), 16803-16809. https://doi.org/10.1007/s00500-020-04973-x

[2] Chen D, Yang X, Maximum likelihood estimation for uncertain autoregressive model with application to carbon dioxide emissions, *Journal of Intelligent & Fuzzy Systerm*, **40** (2021), 1391-1399. https://doi.org/10.3233/jifs-201724

[3] Fang Q, Yang Nan, Research on ARMAX-GARCHX exchange rate model based on gold parity, *Journal of Jinan University* (Social Sciences Edition), **33** (2023), 76-87+179.

[4] Gao J, Xin Y, Yang X, Yang J, Maximum likelihood estimation for uncertain autoregressive moving average model with application in financial market, *Journal of Computational and Applied Mathematics*, **417** (2023), 114604. https://doi.org/10.1016/j.cam.2022.114604

[5] Hu Y, Park G, Yang X, Least absolute deviations estimation for uncertain autoregressive model, *Soft Computing*, **24** (2020), 18211-18217. https://doi.org/10.1007/s00500-020-05079-0

[6] Lio W, Liu B, Uncertain maximum likelihood estimation with application to uncertain regression analysis, *Soft Computing*, **24** (2020), 9351-9360. https://doi.org/10.1007/s00500-020-04951-3

[7] Lio W, Liu B, Initial value estimation of uncertain differential equations and zero-day of COVID-19 spread in China, *Fuzzy Optimization and Decision Making*, **20** (2021), 177-188. https://doi.org/10.1007/s10700-020-09337-6

[8] Liu B, *Uncertainty Theory*, 2nd edn, Springer-Verlag, Berlin, 2007.

[9] Liu Z, Yang Y, Least absolute deviations estimation for uncertain regression with imprecise observations, *Fuzzy Optimization and Decision Making*, **19** (2020), 33-52. https://doi.org/10.1007/s10700-019-09312-w

[10] Liu Z, Yang X, Cross validation for uncertain autoregressive model, *Communications in Statistics-Simulation and Computation*, **51** (2020), 4715-4726. https://doi.org/10.1080/03610918.2020.1747077

[11] Ni Y, Yang X, Least-squares estimation for uncertain moving average model, *Communication in Statistics-Theory and Methods*, **50** (2021),4134-4143. https://doi.org/10.1080/03610926.2020.1713373

[12] Tang H, Uncertain vector autoregressive model with imprecise observations, *Soft Computing*, **24** (2020),17001-17007. https://doi.org/10.1007/s00500-020-04991-9

[13] Yang X, Liu B, Uncertain time series analysis with imprecise observations, *Fuzzy Optimization and Decision Making*, **18** (2019), 263-278. https://doi.org/10.1007/s10700-018-9298-z

[14] Yao K, Liu B, Uncertain regression analysis: An approach for imprecise observations, *Soft Computing*, **22** (2018), 5579-5582. https://doi.org/10.1007/s00500-017-2521-y

[15] Ye T, Liu B, Uncertain hypothesis test with application to uncertain regression analysis, *Fuzzy Optimization and Decision Making*, **21** (2021), 157-174. https://doi.org/10.1007/s10700-021-09365-w

Received: January 7, 2025; Published: January 27, 2025