www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277- 7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2022; SP-11(4): 816-821

www.thepharmajournal.com
Received: 04-02-2022
Accepted: 06-03-2022

An artificial neural network approach for predicting area, production and productivity of Banana in Gujarat

Prity Kumari, DJ Parmar, Sathish Kumar M, YA Lad and AB Mahera

Prity Kumari
Assistant Professor, Department of Basic Science, College of Horticulture, Anand Agricultural University, Anand, Gujarat, India

DJ Parmar
Associate Professor, Department of Agricultural Statistics, B. A. College of Agriculture, Anand Agricultural University, Anand, Gujarat, India

Sathish Kumar M
Ph.D. Research Scholar (Agribusiness Management), International Agri-Business Management Institute, Anand Agricultural University, Anand, Gujarat, India

YA Lad
Associate Professor & Head, Department of HRM & PM, International Agribusiness Management Institute, Anand Agricultural University, Anand, Gujarat, India

AB Mahera
Assistant Professor & Head, Department of Marketing Management, International Agribusiness Management Institute, Anand Agricultural University, Anand, Gujarat, India

Corresponding Author
Prity Kumari
Assistant Professor, Department of Basic Science, College of Horticulture, Anand Agricultural University, Anand, Gujarat, India

Abstract
The present study steered to predicting area, production and productivity of banana in Gujarat by using different models. The secondary data on area, production and productivity of banana in Gujarat (1991-92 to 2017-18) were collected from Directorate of Horticulture, Gujarat. Time series secondary data on area, production and productivity of Banana were collected for the period 1958-59 to 2017-18. The collected data were analyzed in R Studio (version 3.5.2) software. Different Artificial Neural Network models employed to forecast area, production and productivity of fruits crops and also find out best models through comparison of all models 2:2s:1l, 2:3s:1l & 2:2s:1l ANN architectures models were best predicted area, production and productivity of banana in Gujarat with predicted value for 2018-19, 78.59 thousand hectares area, 6286.89 metric tonnes production and 55.29 metric tonnes per hectare productivity where area and production are likely to increase while productivity will go down in upcoming years.

Keywords: Forecasting, area, production, productivity, banana and artificial neural network model

Introduction
Horticulture plays important role in economy of all developing countries especially in India. India is the second largest producer of fruits and vegetables after china. Horticulture contributes 28 percent Gross Domestic Product (GDP) of India. Overall 9,49, 115 metric tonnes of fruits produced from 56, 761 hectares of land during the period 2019-20 in the world. In India, 32,597 thousand tonnes of banana produced from 897 thousand hectares of land with average productivity of 36.3 metric tonnes per hectare during the period 2019-20. In Gujarat, 4472.32 thousand tonnes of banana produced from 68.15 thousand hectares with average productivity of 65.63 metric tonnes per hectare during the period 2017-18 (Ministry of Agriculture & Farmers Welfare). Artificial Neural Network (ANN) model holds different features which attract the researchers at same time it is contrast to many traditional techniques. In Bangladesh, production of major fruit crops such as mango, banana and guava had forecasted by using Box- Jerkins Arima Model. The study found that ARIMA (2,1,3), ARIMA (3,1,2) and ARIMA (1,1,2) were the best model to forecasted the mango, banana and guava. Comparison between original series of data and forecasted series which shown same manner indicating fitted model were statically behaved well to forecast the fruits production in Bangladesh (Hamja, 2014) [4]. In Bangladesh, banana production was forecasted by using Auto- Regressive Integrated Moving Average (ARIMA) model with help of secondary data which has collected over the period 1972 to 2013. ARIMA (0,2,1) selected best model to forecasting the production of banana crop. Observed and forecasted data comparison indicated fitted model behaved statistically well during and beyond the estimation period (Hossain *et al*., 2016). Mango and banana yield forecasted in Karnataka, India by using linear, nonlinear and non-parametric statistical models. Auto-Regressive Integrated Moving Average (ARIMA) model found as better to forecasting of banana yield rather than other models (Rathod and Mishra, 2018). In Varanasi region, pigeon pea yield was forecasted by using statistical models. Different linear and non-linear models like multiple linear regression (MLR), autoregressive integrated moving average (ARIMA) and artificial neural network (ANN) models were used to analyze 27 years data from 1985-86 to 2011-12. The best suited model was identified based on root mean squared error (RMSE). The study revealed that ANN model was best model with lowest RSME which forecasted pigeon pea yield well during 2012-13 for Varanasi region (Kumari Prity *et al*., 2016) [6].

Kumari Prity *et al*. (2017) [7] studied forecasting models for predicting pod damage of pigeon pea in Varanasi region. Autoregressive integrated moving average (ARIMA) and artificial neural network (ANN) with multiple linear regression models were used to predicting perc cent pod damage in pigeon pea by pod borer in Varanasi region, Uttar Pradesh by using 27 years of data (1985-86 to 2011-12). The best suited model was assessed by root mean squared error (RSME). The study revealed that ANN was found best model with lowest RSME having forecasted per cent of pod damage in pigeon pea 2012-13. Kumari Prity and Sathish Kumar M (2021) [8] forecasted area, production and productivity of Citrus in Gujarat – An application of artificial neural network. Times series data were collected for this study from 1991-92 to 2017-18 and different artificial neural network models were used. This study found that area, production and productivity of citrus was best explained by 4:1s:1l, 2:2s:1l & 3:2s:1l ANN architectures, with forecasted value for 2018-19, 41.43 ('000' Ha.), 4143.00 ('000' MT) & 10.20 (MT/Ha.) respectively, where area, production & productivity are likely to go down for the next year. Sathish Kumar M and Kumari Prity (2021) [8] used different artificial neural network models for predicting area, production and productivity of sapota in Gujarat. They found that area, production and productivity of sapota was best explained by 4:1s:1l, 4:1s:1l and 2:2s:1l ANN architectures, with forecasted value for 2018-19, 28.48 ('000' Ha.), 320.89 ('000' MT) and 10.51 (MT/Ha.) respectively, where area, production and productivity are likely to go decrease for the next year.

Objectives

To evaluate appropriate Artificial Neural Network (ANN) for forecasting area, production and productivity of Banana.

To compare the performance of developed models and estimate the forecast for area, production and productivity of Banana.

Methodology

Source of data

Secondary data on area, production and productivity of Banana in Gujarat were collected from Directorate of Horticulture, Govt. of Gujarat from 1991-92 to 2017-18. Time series secondary data on area, production and productivity of Banana were collected for the period 1958-59 to 2017-18.

Analytical framework

In the present study, different neural network architectures were used to compare their ability for predicting area, production and productivity of four major fruit crops in Gujarat. Analysis was done by RStudio (version 3.5.2) software.

Artificial neural network (ANN)

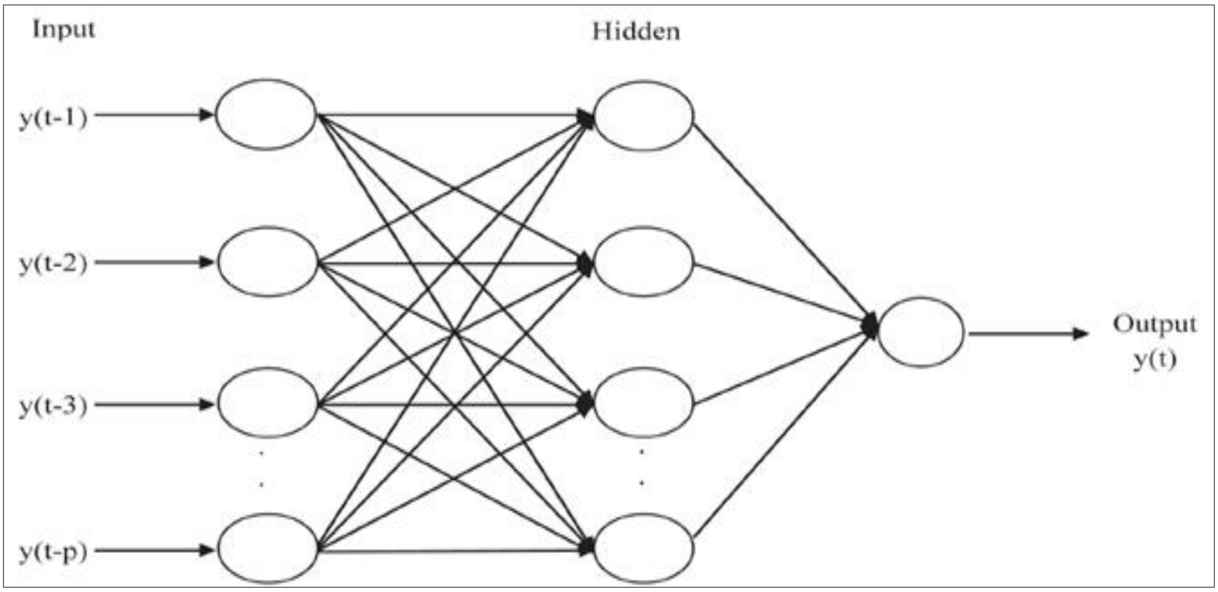

Fig 1: Time-Delay Neural Network (TDNN) with one hidden layer

ANNs are nonlinear data-driven models capable to perform modeling without a prior knowledge about the relationships between input and output variables. Its generalizing ability, after learning the data presented to structure, can often correctly infer the unseen part of a population even if the sample data contain noisy information. Time series can be modelled with the structure of a neural network by the use of time delay, which can be implemented at the input layer of the neural network. Such an ANN is termed as Time Delay Neural Network.

The structure of the neural network consists of

1. Input Layer
2. Hidden Layer
3. Output Layer

The general expression for the final output value y_{t+1} in a multilayer feed forward time delay neural network is given by equation:

$$y_{t+1} = g\left[\sum_{j=1}^{q} \alpha_j f\left(\sum_{i=0}^{p} \beta_{ij} y_{t-i}\right)\right]$$

Where

f and g denote the activation function at the hidden and output layers, respectively. p is the number of input nodes (tapped delay), q is the number of hidden nodes, β_{ij} is the weight attached to the connection between i^{th} input node to the j^{th} node of hidden layer, α_j is the weight attached to the connection from the j[th] hidden node to the output node, y_{t-i} is the ith input (lag) of the series. The main task of activation function is to map the outlying values of the obtained neural input back to a bounded interval such as [0,1] or [-1,1].

Research Results

In this study, area, production and productivity (yield) of banana were analyzed by different neural network architecture. The empirical findings of banana as follow:

Forecasting of area, production and productivity for Banana

Forecasting of area for banana

Fig. 2 illustrates chart series of area dataset for banana from 1958-59 to 2017-18. Also, the characteristics (basic statistics) of the data set used were presented in the Table 1.

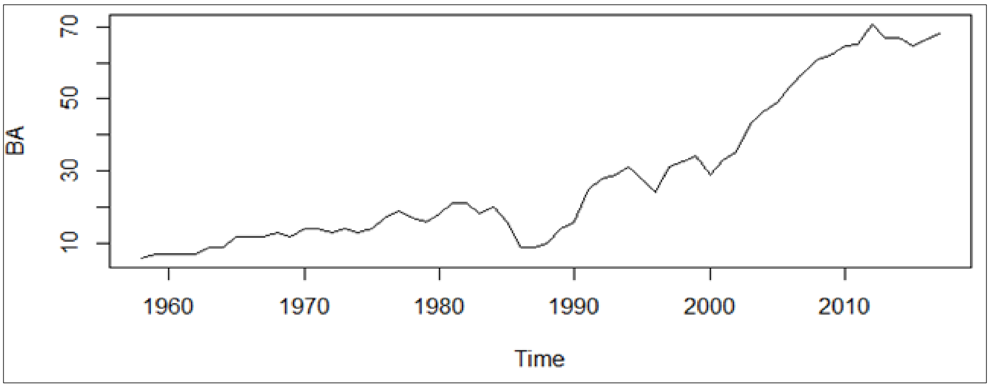

Fig 2: Area (In ' 000 Hectare) under banana in Gujarat

Table 1: Summary statistics of banana area time series

No. of observations	60
Minimum	6
Maximum	70.58
Mean	28.37
Median	19.50
Standard Deviation	20.48
Sem	2.64
Skewness	0.86
Kurtosis	-0.74

Various architectures of neural network were tried considering the availability of data. Further, the model performance in training set and testing data set is given in Table 2.

Table 2: Forecasting performance of ANN model for Banana area time series

Model	Parameters	RMSE	
		Training	Testing
2-1S-1L	5	58.169	2.822
3-1S-1L	6	58.838	2.842
4-1S-1L	7	59.669	2.822
2-2S-1L	9	58.041	2.738
3-2S-1L	11	58.635	2.605
4-2S-1L	13	58.870	2.419
2-3S-1L	13	58.309	2.629
3-3S-1L	16	59.036	2.348

Based on the lowest training RMSE, 3 ANN models *viz.* 2:2s:1l, 2:1s:1l and 2:3S:1l are selected. These models were further assessed based on their hold out sampling (testing set) forecasting performance. Out of all neural network architectures, a model with two tapped delay and two hidden nodes (2:2s:1l), was selected for forecasting area of banana in Gujarat.

Table 3: ANN model parameter Banana area time series

Weights between nodes		Biases	
I1:H1	0.443	Hidden node	
I1:H2	-1.299	H_{B1}	2.762
I2:H1	0.352	H_{B2}	-4.610
I2:H2	-3.994	Output node	
H1:O	-3.965	O_B	2.791
H2:O	5.084		
Forecasting (2018-19)		C.I.	
78.59		70.60	88.10

Table 3 reflects that the estimates of all weights associated with nodes of different layer. Input layer lag1 and lag2 are denoted by I1 & I2, Hidden layer node1 and node2 are denoted by H1 & H2 and output node is denoted by O, where biases of three nodes are given by the notation H_{B1}, H_{B2} & O_B. The forecasted value of banana area in Gujarat for the year 2018-19 by 2:2s:1l neural network architecture was obtained as 78.59('000' Hectares) with confidence interval 70.60 to 88.10.

Forecasting of production for banana

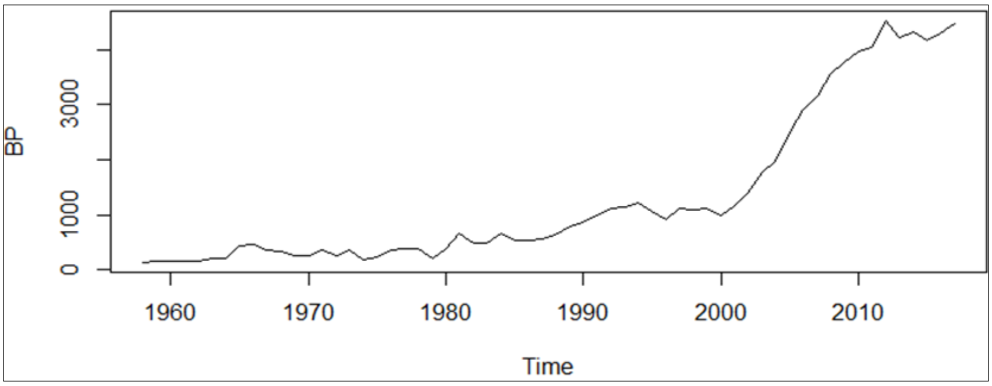

Fig 3: Production (In '000 MT) of banana in Gujarat

Fig. 3 illustrate chart series of production dataset for banana from 1958-59 to 2017-18. The characteristics (basic statistics) of the data set used presented in Table 4.

Table 4: Summary statistics of banana production time series

No. of observations	60
Minimum	140
Maximum	4523.49
Mean	1328.25
Median	666
Standard Deviation	1420.69
Sem	183.41
Skewness	1.22
Kurtosis	-0.08

Artificial neural network architecture was tried considering the availability of data. Further, the model performance in training set and testing data set is given in Table 5.

Table 5: Forecasting performance of ANN model for banana production time series

Model	Parameters	RMSE	
		Training	Testing
2-1S-1L	5	4115.380	129.20
3-1S-1L	6	4178.678	128.854
4-1S-1L	7	4264.281	129.950
2-2S-1L	9	4099.537	117.238
3-2S-1L	11	4171.632	113.215
4-2S-1L	13	4250.322	112.667
2-3S-1L	13	4098.418	114.472
3-3S-1L	16	4176.452	106.259

Based on the lowest training RMSE, 3 ANN models *viz.* 2:1s:1l, 2:2s:1l and 2:3S:1l are selected. These models were further assessed based on their hold out sampling (testing set) forecasting performance. Out of all neural network architectures, a model with two tapped delay and three hidden nodes (2:3s:1l), was selected for forecasting production of banana in Gujarat.

Table 6: ANN model parameter for banana production time series

Weights between nodes		Biases	
I1:H1	-6.497	Hidden node	
I1:H2	5.347	H_{B1}	-0.021
I1:H3	5.599	H_{B2}	0.283
I2:H1	0.594	H_{B3}	-2.455
I2:H2	-3.641	Output node	
I2:H3	2.252	O_B	8.426
H1:O	-1.427		
H2:O	-0.597		
H3:O	0.889		
Forecasting (2018-19)		C.I.	
6286.89 (In '000 MT)		6008.51	6565.93

Table 6 reflects that the estimates of all weights associated with nodes of different layer. Input layer lag1 and lag2 are denoted by I1 & I2, Hidden layer node1, node2 and node3 are denoted by H1, H2 & H3 and output node is denoted by O, where biases of four nodes are given by the notation H_{B1}, H_{B2}, H_{B3} & O_B. The forecasted value of banana production in Gujarat for the year 2018-19 by 2:3s:1l neural network architecture was obtained as 6286.89 ('000' MT) with confidence interval 6008.51 to 6565.93.

Forecasting of productivity for banana

Fig. 4 illustrate chart series of productivity dataset for banana from 1958-59 to 2017-18. Also, the characteristics (basic statistics) of the data set used were presented in Table 7.

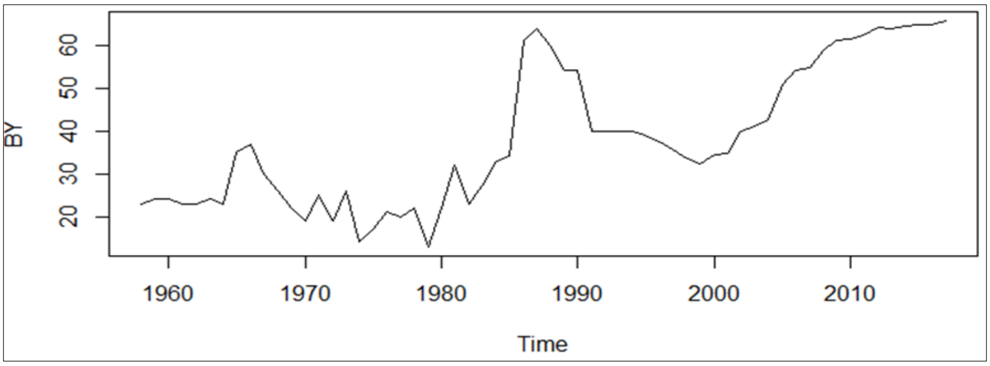

Fig 4: Productivity (In MT/Ha.) of banana in Gujarat

Table 7: Summary statistics of banana productivity time series

No. of observations	60
Minimum	13
Maximum	65.62
Mean	38.27
Median	34.91
Standard Deviation	16.25
Sem	2.10
Skewness	0.39
Kurtosis	-1.25

Different neural network architecture was tried considering the availability of data. Further, the model performance in training set and testing data set is given in Table 8.

Table 8: Forecasting performance of ANN model for Banana productivity time series

Model	Parameters	RMSE	
		Training	Testing
2-1S-1L	5	40.427	6.040
3-1S-1L	6	40.428	6.067
4-1S-1L	7	41.412	5.984
2-2S-1L	9	41.212	5.454
3-2S-1L	11	41.434	5.338
4-2S-1L	13	42.046	4.978
2-3S-1L	13	41.477	5.137
3-3S-1L	16	40.979	4.378

Based on the lowest training RMSE, 4 ANN models *viz*. 2:1s:1l, 3:1s:1l, 2:2s:1l and 3:3S:1l are selected. These models were further assessed based on their hold out sampling (testing set) forecasting performance. In this scenario, lowest training RMSE are in the case of first two and last two model (as given in Table 5) but due to underestimation and overestimation case, neural network of 2:2s:1l architecture is considered to be the best explained model for predicting banana productivity. Out of all neural network architectures, a model with two tapped delay and two hidden nodes (2:2s:1l), was selected for forecasting productivity of banana in Gujarat.

Table 9: ANN model parameter Banana productivity time series

Weights between nodes		Biases	
I1:H1	2.901	Hidden node	
I1:H2	-3.460	H_{B1}	1.054
I2:H1	0.979	H_{B2}	-1.937
I2:H2	8.985	Output node	
H1:O	27.941	O_B	0.983
H2:O	5.131		
Forecasting (2018-19)	C.I.		
55.29	30.39	69.93	

Table 9 reflects that the estimates of all weights associated with nodes of different layer. Input layer lag1 and lag2 are denoted by I1 & I2, Hidden layer node1 and node2 are denoted by H1 & H2 and output node is denoted by O, where biases of three nodes are given by the notation H_{B1}, H_{B2} & O_B. The forecasted value of banana productivity in Gujarat for the year 2018-19 by 2:2s:1l neural network architecture was obtained as 55.29(MT/ha.) with confidence interval 30.39 to 69.69.

Table 10: Performance of different models for banana

Model for crops		Area (In '000' Ha.)	Production (In '000 MT)	Productivity (MT/Ha.)
Banana	Model	2:2s:1l	2:3s:1l	2:2s:1l
	RMSE	58.04	4098.418	41.212
	Forecast	78.59 (68.75)	6286.89 (4472.32)	55.29 (65.62)
	C.I.	70.60 to 88.10	6008.51 to 6565.93	30.39 to 69.93

*values within in parenthesis are previous year APY of all crops

Table 10 shows area, production and productivity of banana were best explained by 2:2s:1l, 2:3s:1l & 2:2s:1l ANN architectures, with forecasted value for 2018-19, 78.59 thousand hectares area, 6286.89 thousand metric tonnes production and 55.29 metric tonnes per hectare productivity respectively, where area and production are likely to increase while productivity will go down.

Conclusion

The study found artificial neural network model as better for predicting area, production and productivity of crops than classical models. Based on result of the study farmers and policy makers in production of banana can make plan in advance. This model result also helps to increase the productivity by using suitable management of inputs and available weather information. The artificial neural network approach can be further extended using some other machine learning techniques so that practical validity of the model can be well established. This artificial neural model can generalize by applying this approach to all other agricultural and horticultural crops.

Acknowledgement

Authors are thankful to Anand Agricultural University, Anand -388110, Gujarat, India.

References

1. Aguilar KL, Mendoza AB, Morales SG, Maldonado AJ. Artificial neural network modelling of greenhouse tomato yield and aerial dry matter. Agriculture. 2020;10(2):2-14.
2. Dhaikar SS, Rode SV. Agricultural crop yield prediction

using artificial neural network approach. International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering. 2014;2(1):683-686.
3. Guo WW, Xue H. Crop yield forecasting using artificial neural networks. A comparison between spatial and temporal models. Mathematical problems in Engineering. 2014;8(3):4-10.
4. Hamjah MA. Forecasting major fruit crops productions in Bangladesh using Box-Jenkins ARIMA model. Journal of Economics and Sustainable Development. 2014;5(7):96-107.
5. Hossain MM, Abdulla F, Majumder AK. Forecasting of Banana Production in Bangladesh. American Journal of Agricultural and Biological Sciences. 2016;11(2):93-99.
6. Kumari Prity, Mishra GC, Srivastava CP. Statistical models for forecasting pigeon pea yield in Varanasi region. Journal of Agrometeorology. 2016;18(2):306-310.
7. KumariPrity, Mishra GC, Srivastava CP. Forecasting models for predicting damage of pigeon pea in Varanasi region. Journal of Agrometeorology. 2017;19(3):265-269.
8. Kumari Prity, Sathish Kumar M. Forecasting area, production and productivity of Citrus in Gujarat- An application of artificial neural network. International Journal of Agricultural Sciences. 2021;13(10):10913-10916.
9. Rathod S. Modelling and forecasting of oil seed production of India through artificial intelligence techniques. Journal of Agricultural Sciences. 2018;88(1):22-27.
10. Rathod S, Mishra CG. Statistical models for forecasting Mango and Banana yield of Karnataka, India. Journal of Agricultural Science and Technology. 2018;20(3):803-816.
11. Sathish Kumar M, Kumari Prity. Artifical neural network model for predicting area, production and productivity of sapota in Gujarat. International Journal of Agricultural Sciences. 2021;13(10):10909-10912.