Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

North American Simulation and Gaming Association
NASAGA

Mihail.Motzev.com

ABSEL
Association for Business Simulation and Experiential Learning

ISAGA

Summer school on modelling and complex systems 2021:

“STATISTICAL LEARNING NETWORKS”

Mihail Motzev
Ph.D,M.Sc,P.D.D
(M.Motzev@wallawalla.edu)

Model Selection

Prediction (simulation) error:

$$e_t = y_t - F_t$$

where ***e**$_t$* is the error at period ***t** (t={1, 2, 3...N})*;

N is the prediction interval (or the size of the dataset);

y$_t$ is the actual value at period ***t*** and

F$_t$ is the forecast for period ***t*.**

Model Selection

Measures of Trueness (Systematic error, Statistical Bias):

- **Mean Percentage Error (MPE)**

$$\text{MPE}\ (\%) = \frac{1}{\text{N}} \sum_{t=1}^{N} (e_t / y_t) \text{x} 100$$

- **Root Mean Squared Error (RMSE)**

$$RMSE = \sqrt{\text{MSE}} \qquad \text{MSE} = \sum \left(e_t\right)^2 / (n-1)$$

Model Selection

Measures of Precision (Random Error):

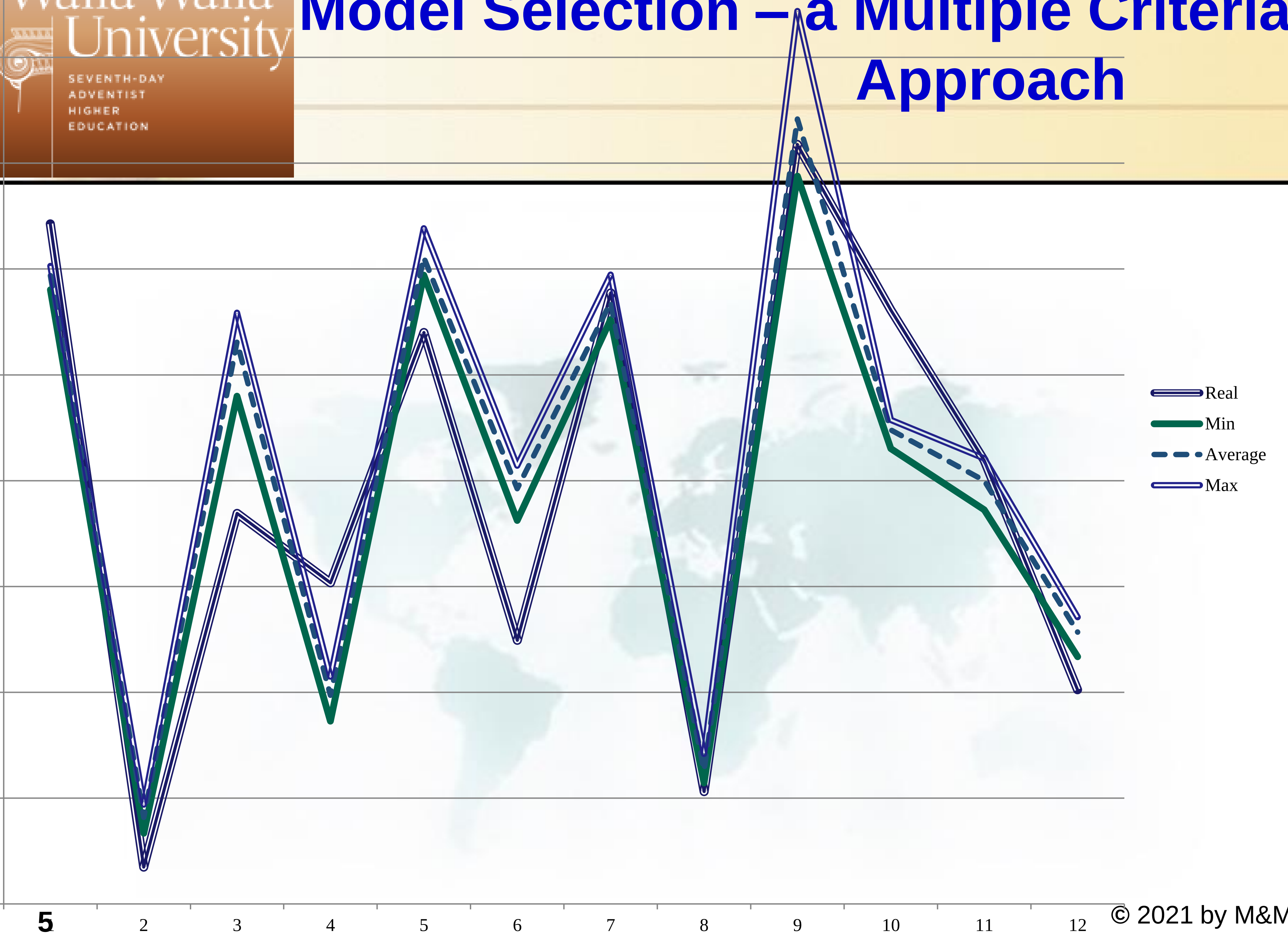
- **Mean Absolute Percentage Error (MAPE)**

$$\text{MAPE}\,(\%) = \frac{1}{\text{N}}\sum_{t=1}^{N}\left(\left|e_t\right| / \, y_t\right)\text{x}100$$

- **Coefficient of Variation of the RMSE, CV(RMSE)**

$$\text{CV(RMSE)} = \text{RMSE}/\overline{\text{y}}$$

Model Selection – a Multiple Criteria Approach



	GMDH algorithms	
Variables	**Parametric**	**Non-parametric**
Continuous	- Combinatorial (COMBI) - Multilayered Iterative (MIA) - Objective System Analysis (OSA) - Harmonical - Two-level (ARIMAD) - Multiplicative-Additive (MAA)	- Objective Computer Clusterization (OCC); - "Pointing Finger" (PF) clusterization algorithm; - Analogues Complexing (AC)
Discrete or binary	- Harmonical Rediscretization	- Algorithm on the base of Multilayered Theory of Statistical Decisions (MTSD)

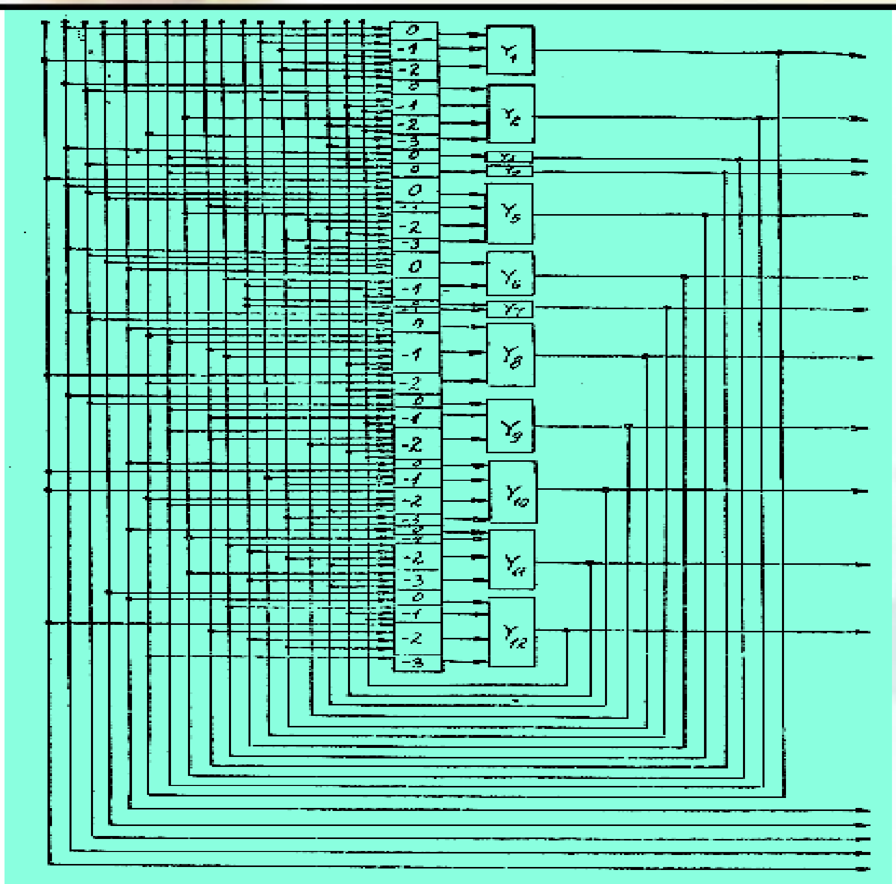
Basic GMDH algorithms

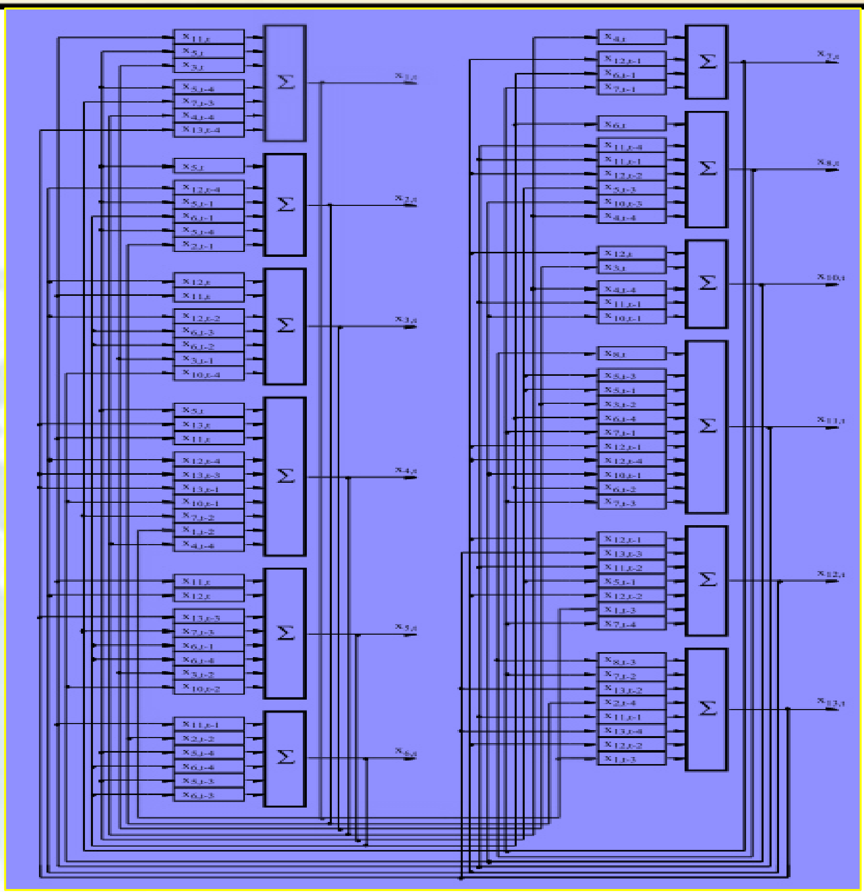
Source: http://www.gmdh.net/GMDH_alg.htm

Example 1: A Family of Models "SIMUR"

Year of design	Main purpose & Accuracy	Model characteristics
1981	Analysis of the possibilities for automated model building in the form of SE using MSSP. Accuracy = 2.7%.	A one-product macroeconomic model in the form of 5 SE. Contains 5 endogenous, 5 lag and 1 exogenous variables
1985	Software development for simulation experiments with SE. Analysis of different criteria for model evaluation & selection. Accuracy = 2.0%.	Aggregated macroeconomic model in the form of 12 SE. Contains 12 endogenous, 5 exogenous and 26 lag variables with lag of up to 3 years.
1987	Improving MSSP for synthesis of a big number of SE. Simulation and prediction of the main macroeconomic indexes. Accuracy < 1%.	Complex macro-economic model of 39 SE, with 39 endogenous, 7 exogenous and 82 lag variables with a time lag of up to 5 years.

System’s graph of the generated macroeconomic models for Bulgarian and German Economy





Results show that SLNs (MLNAN) provide opportunities in both shortening the design time and reducing the cost and efforts in model building, as well as developing reliably even complex models with high level of accuracy and low overall error rates

SODM with GMDH - Applications

Prediction errors for forecasts calculated using SIMUR II model

Variable Equation	MAPE%				Average for the model:			
	1977	1978	1979	1980	MAPE%	CV(RMSE)%	R	U
1	8.28%	2.58%	4.52%	3.85%	4.81%	5.10%	0.999	0.0260
2	4.26%	6.45%	10.16%	0.61%	5.37%	6.40%	0.975	0.0040
3	3.71%	6.54%	8.31%	0.62%	4.80%	5.84%	0.977	0.0310
4	5.81%	6.15%	15.67%	0.56%	7.05%	8.93%	0.993	0.0078
5	6.17%	0.35%	6.32%	3.81%	4.16%	4.47%	0.998	0.0022
6	0.48%	0.40%	0.04%	1.36%	0.57%	1.82%	0.996	0.0003
7	1.43%	1.15%	1.44%	1.64%	1.42%	1.77%	0.986	0.0035
8	0.96%	0.83%	0.22%	0.62%	0.66%	0.72%	0.995	0.0001
9	5.24%	6.15%	5.60%	8.23%	6.31%	6.68%	0.995	0.0044
10	0.77%	2.22%	1.20%	3.34%	1.88%	2.31%	0.998	0.0005
11	13.89%	6.26%	6.39%	3.86%	7.60%	7.72%	0.996	0.0056
12	1.03%	3.16%	3.16%	1.18%	2.13%	3.28%	0.995	0.0010
Average	4.34%	3.52%	5.25%	2.47%	3.90%	4.59%	0.992	0.0098

Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

SODM with GMDH - Applications

Prediction errors for German economy forecasts using *Knowledge-Miner*

Variable	% Differences Between Predictions & Real Data			MAPE	MSE (%)
	1988	1989	Average %		
Y_{1t}	2.58%	3.85%	3.22%	3.48%	5.10%
Y_{2t}	6.45%	0.61%	3.53%	3.73%	6.40%
Y_{3t}	6.54%	0.62%	3.58%	3.84%	5.85%
Y_{4t}	6.15%	0.56%	3.36%	3.59%	8.93%
Y_{5t}	0.35%	3.81%	2.08%	2.30%	4.47%
Y_{6t}	0.04%	1.36%	0.70%	1.10%	1.82%
Y_{7t}	16.51%	17.60%	17.06%	17.80%	10.12%
Y_{8t}	0.83%	0.62%	0.73%	1.15%	0.72%
Y_{9t}	6.15%	8.23%	7.19%	7.43%	6.68%
Y_{10t}	2.22%	3.34%	2.78%	2.98%	2.31%
Y_{11t}	6.26%	3.86%	5.06%	5.45%	7.72%
Y_{12t}	3.16%	1.18%	2.17%	2.48%	3.28%
Mean	4.77%	3.80%	4.29%	4.61%	5.28%

RESEARCH PROJECTS :

The family of models "S I M U R"

Model No . Year of design	Objectives – Improving Accuracy	Characteristics
SIMUR 0 (1977—1978)	First step of the SIMUR project. Traditional methods used. Accuracy = 14%.	**A one—product macroeconomic model in the form of SSE with 5 equations. Contains 5 endogenous, 5 lag and 1 exogenous variables.**
SIMUR I (1978—1980)	Analysis of possibilities for automatic synthesis of SSE during the run of multi-stage selection procedure. Accuracy = 2.7%.	**A one—product macroeconomic model in the form of SSE with 5 equations. Contains the same set of variables.**
SIMUR II (1981—1982)	Design and experimenting of a programming system for automatic holding of simulation experiments with SSE. Analysis of different criteria for the precision's estimating of SSE. Accuracy = 2.0%.	**Aggregated macroeconomic model in the form of 12 interdepending simultaneous equations. Contains 12 endogenous, 5 exogenous and 26 lag variables with lag of up to 3 years.**
SIMUR III (1983—1985)	Improving the MSSP for synthesis of SSE with many equations. Simulating and forecasting of main macroeconomic indexes. Accuracy < 1%.	**Macroeconomic simulation model. Contains 39 equations and 39 endogenous, 7 exogenous and 82 lag variables with a time lag of up to 5 years.**
SIMUR IV (1989 -)	**Next step of the SIMUR project.**	**Multisector macroeconomic model. Contains more than 100 equations.**

Example 2: Model-Based Business Games

About WWU | Academics | Life at WWU | Resources | Attend WWU | Email | MyWWU | Login

Walla Walla University

NAVIGATION
- News
  - Archive
  - Name Change

RELATED LINKS
- Inside WWU Newsletter

WHAT'S HAPPENING
APIC Chinese New Year Party ...

Quicklinks

BEST COLLEGES U.S.News REGIONAL UNIVERSITIES WEST 2012

HOME > ABOUT WWU > GENERAL INFORMATION > NEWS >

Business Professor Designs Game to Help Industry Professionals

Motzev Has Shared Research Results at Worldwide Conferences

By: Becky St. Clair

Mihail Motzev, School of Business professor

Who says professionals can't have fun? Mihail Motzev, professor in the School of Business at Walla Walla University, spent three years developing what is essentially a game for businesspeople. His latest project, titled "Intelligent Techniques in Simulations and Management Games–A Hybrid Approach: Multi-Stage Selection Procedures for Complex Systems Model Building" was nominated and approved by WWU faculty for a faculty research grant three consecutive years.

"It's one of my favorite areas of research," says Motzev. "I enjoy the work and the presentation as much as I enjoy the end result."

As a member of the International Simulation and Gaming Association (ISAGA), Motzev has presented results from his research at many conventions, most recently in Romania, Poland, and Spokane, Wash. He has also been invited to present at the ISAGA/IFIP (International Federation for Information Processing) world conference in Sweden this summer.

STATISTICAL LEARNING NETWORKS

Example 2: Model-Based Business Games

Original Version	New Version
A one-product macro-economic model developed as a system of five SE. Contains five endogenous, one exogenous, and five lag variables.	A one—product macroeconomic model with the same structure. Contains same set of variables.
Indirect OLS used to estimate unknown coefficients in equations.	Model synthesized using the hybrid algorithm.
Model accuracy - mean squared error (MSE) = 14%	Model accuracy MSE = 2.7%

Example 2: *Model-Based Business Games*

Model characteristics and comparisons

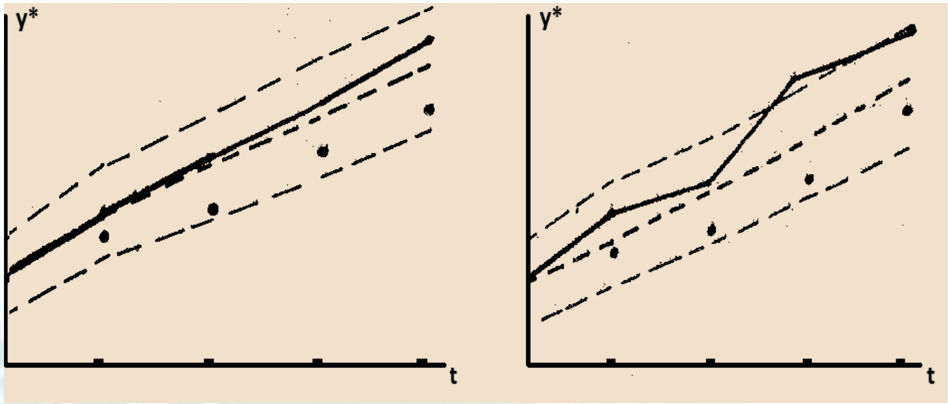
Characteristics	Old Version	Improved Version
Model description	A one-product macro-economic model developed as a system of five SE. Contains five endogenous, one exogenous, and five lag variables.	A one-product macroeconomic model with the same structure. Contains same set of variables.
Model-building technique	Indirect OLS used to estimate unknown coefficients in equations.	Model synthesized using the GMDH procedure.
Model accuracy	Mean squared error relative to the mean (MSE%) = 14%	MSE% = 2.7%

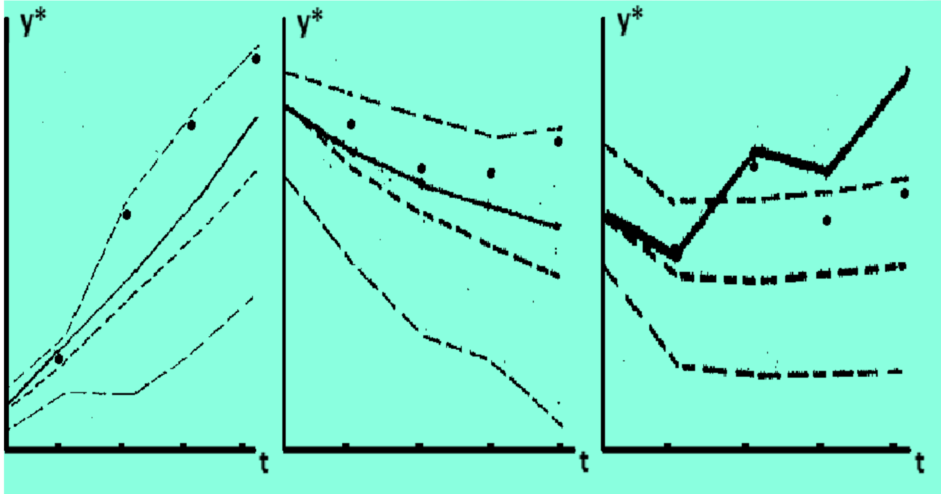
Improving Model Accuracy

Statistics and test results for ***ex-post*** forecasts

Eq. no	Coefficient of multiple correlation (R)	Coefficient of multiple determination (R^2)	F value	CV(RMSE)%	MAPE %
1	0.9949	0.9898	355.81	3.40%	2.69%
2	0.9980	0.9960	936.50	2.16%	1.77%
3	0.9991	0.9982	1848.52	1.76%	1.40%
4	0.9637	0.9287	47.69	0.66%	0.52%
5	0.9522	0.9067	24.30	11.09%	7.13%
Total	0.9816	0.9635	642.56	3.81%	2.70%

Ex-ante predictions four years ahead for variables $y_{1,t}$ and $y_{2,t}$





Ex-ante predictions 4 years ahead for variables $y_{3,t}$, $y_{4,t}$ and $y_{5,t}$

Eq. no	Coefficient of multiple correlation (R)	Coefficient of multiple determination (R^2)	F value	CV(RMSE) %	MAPE %
1	0.9928	0.9857	254.55	4.03%	3.57%
2	0.9925	0.9851	242.42	4.22%	3.78%
3	0.9993	0.9986	1664.17	3.84%	3.39%
4	0.8288	0.6869	8.05	1.39%	1.11%
5	0.9362	0.8765	17.76	12.76%	8.56%
Total	0.9499	0.9023	437.39	5.25%	4.08%

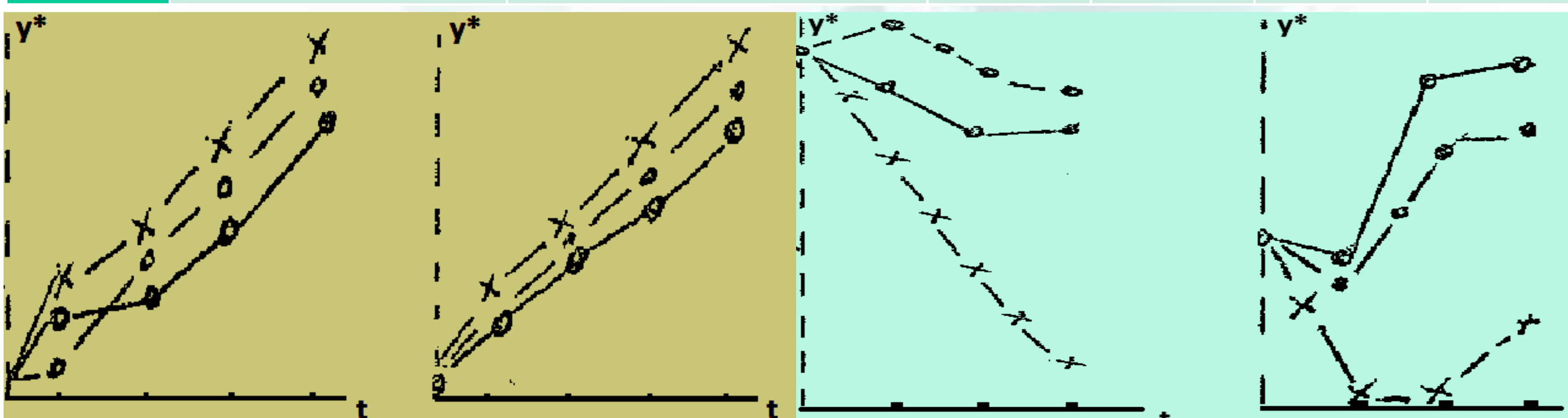
Statistics and test results for ***ex-ante*** forecasts

Statistical Learning Networks

Improving Model Accuracy

Comparisons for *ex-post* forecasts from stationary and non-stationary models

Equation No.	Coefficient of multiple determination (R^2)		von Neumann test for Autocorrelation (Q)		CV(RMSE)%		MAPE%	
	Stationary	Non-Stat.	Stationary	Non-Stat.	Stationary	Non-Stat.	Stationary	Non-Stat.
1	0.9898	0.9920	1.010	0.965	3.40%	2.78%	2.69%	2.35%
2	0.9960	0.9980	0.735	1.212	2.16%	1.44%	1.77%	1.23%
3	0.9982	0.9980	1.070	1.069	1.76%	1.63%	1.40%	1.34%
4	0.9287	0.9900	0.846	1.081	0.66%	0.81%	0.52%	0.72%
5	0.9067	0.9460	1.020	1.233	11.09%	7.61%	7.13%	6.73%
For the model	xxx		xxx		3.81%	2.85%	2.70%	2.48%



Real values

Non-stationary

Stationary

Ex-ante **prediction comparisons for three years ahead with stationary and non-stationary models**

STATISTICAL LEARNING NETWORKS

Improving Model Accuracy

Comparisons for ***ex-ante*** forecasts from stationary and non-stationary models

Equation No.	Theil's U-statistics (U)		Standard error of the forecast (S)		CV(RMSE)%		MAPE%	
	Stationary	Non-Stat.	Stationary	Non-Stat.	Stationary	Non-Stat.	Stationary	Non-Stat.
1	0.018	0.014	579.20	463.30	3.58%	2.87%	3.27%	2.80%
2	0.027	0.012	675.30	311.26	5.49%	2.53%	5.31%	2.18%
3	0.001	0.003	95.64	298.28	0.22%	0.69%	0.19%	0.54%
4	0.016	0.003	115.84	27.24	3.22%	0.76%	2.88%	0.75%
5	0.128	0.027	1013.6	238.86	22.74%	5.36%	22.01%	5.03%
Average	xxx		xxx		7.05%	2.44%	6.73%	2.26%

Experimental Test Results - 2018

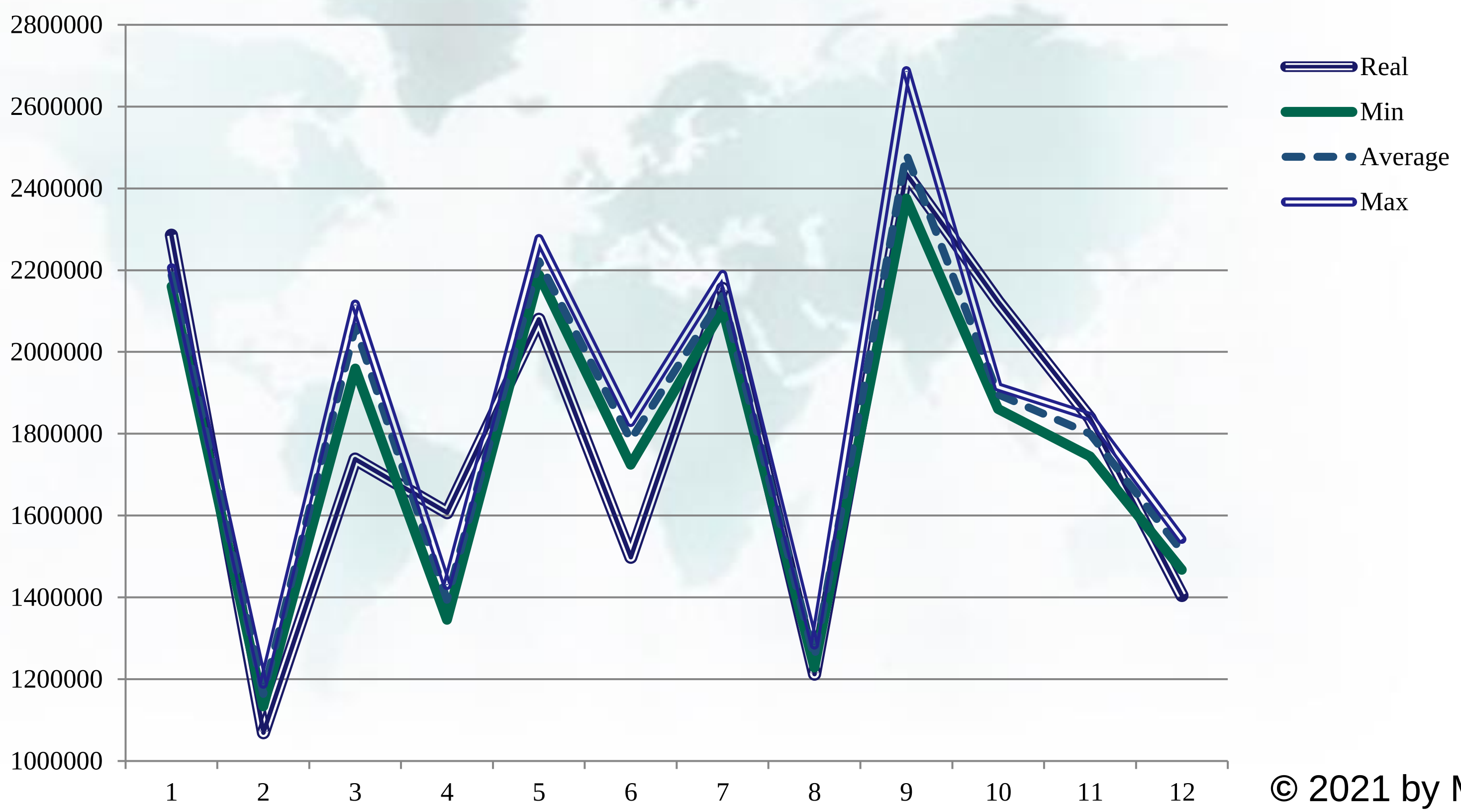
Best Model	Second Best	Third Best
MLNAN: MASE: 0.0414 MPE = 1.42% MAPE = 1.42% CV(RMSE) = 1.56%	Triple Exponential MASE = 0.0627 MPE = -0.57% MAPE = 1.76% CV(RMSE) = 2.45%	Multiple Autoregression MASE = 0.0908 MPE = 2.03% MAPE = 2.58% CV(RMSE) = 3.17%

Experimental Test Results - 2019

Best Model	Second Best	Third Best
MLNAN: MASE: 0.0446 MPE = 1.55% MAPE = 1.55% CV(RMSE) = 1.56%	Multiple Regression with Time and Dummy Seasonal MASE = 0.0508 MPE = -1.09% MAPE = 1.59% CV(RMSE) = 1.56%	Triple Exponential MASE = 0.0627 MPE = -0.57% MAPE = 1.76% CV(RMSE) = 2.45%

Model Selection – a Multiple Criteria Approach

Experimental Test Results – 2021 copy rights protected



Thank You!

Questions?

Thank You!

and I’ll

See You again…