Air passengers forecasting for Australian airline based on hybrid rough set approach

H. KUMAR SHARMA, K. KUMARI AND S. KAR

Abstract

Accurate and reliable air passenger demand is very important for policy-making and planning by tourism management as well as by airline authorities. Therefore, this article proposed a novel hybrid method based on rough set theory (RST) to construct decision rules for long-term forecasting of air passengers. Level (mean) and trend components are first estimated from the air passengers time series data using DES model in the formulation of the proposed hybrid method. Then the rough set theory is employed to combine the output of DES model and generated decision rules is used to forecasting air passengers. We compare the proposed approach with other time series models using a corrected classified accuracy (CCA) criterion. For the empirical analysis, yearly air transport passenger from 1992 to 2004 is used. Empirical results show that the proposed method is highly accurate with the higher corrected classified accuracy. Also, forecasting accuracy of the proposed method is better than the other time series approaches.

Mathematics Subject Classification 2010:
Keywords: Air passengers, Forecasting, Hybrid approach, Rough set, DES model, Grey model

1. INTRODUCTION

Over the past decade, international tourism demand has become one of the most rapidly growing industries in the world. International air travel demand is often calculated regarding the number of air passenger arrivals (Lim and McAleer, 1999). For the past many years, the Australian foreign tourism demand has been growing rapidly. Moreover, total tourist arrivals have been calculated approximately 2.04 to 2.85 million from 2006 and 2012 by 39.82% growth rate during this year (Australian Bureau of Statistics, 2013a). Also, Brisbane, Melbourne, Perthand Sydney are the four key international airports in Australia by 92.73% of the monthly arrivals between 2006 and 2012, (Australian Bureau of Statistics, and 2013b). Australia has exhibited a 5.8% growth in international tourist arrivals in 2012 (United Nations World Tourism Organisation, 2014).

10.2478/jamsi-2018-0001

Consequently, highly accurate and consistent air passenger forecasts play a significant role in policy-making and planning by the private airport and airline management as well as by government tourism authorities and operators. Long-term forecasting of air passenger supplies a key input into decisions of daily operation management, including aircraft scheduling decisions, maintenance planning, advertising and sales campaigns, the opening of new sales offices, etc. More highly reliable and accurate forecasting tools are required to plan effectively. Consequently, this study focuses on the long-term forecasting of air passengers in Australian airline.

Over the past many years different univariate and multivariate time series techniques have been developed in air passenger forecasting (Samagaio and Wolters, 2010), Nguyen et al. (2013), Apergis et al. (2017). They include autoregressive integrated moving average (ARIMA) (Lim and McAleer (1999, 2001, 2002), Cho (2003), Coshall (2006) and Tsui and Balli (2016), naive and smoothing models (Martin and Witt (1989)), Law and Au (1999), Law (2000)) to forecast non-stationary data along with trend component. In particular, Xie et al. (2014) employed a hybrid approach based on decomposition and least squares support vector regression approaches. Pal and Kar (2017) applying forecast technique in the stock exchange case study. The empirical results showed that the forecasting performance was improved when the hybrid approaches of decomposition were implemented. Although, previous commonly used time series models are based on the statistical assumptions and have huge limitations such as large sample, normal distribution, and non-stationary time series data. To overcome the conditions of statistical techniques, modern soft computing method has been introduced such as rough set analysis by Pawlak (1982). Rough set theory has been consistently applied in tourism demand modeling and forecasting for the extraction of decision rules, particularly for the data on the Hong Kong tourism scenario by Law and Au (1998, 2000), Au and Law (2000) and Goh and Law (2003). Also, several variants of the rough set approach have also emerged in the tourism studies by Goh et al. (2008), Lin (2010), Xiaoya and Zhiben (2011) and Celotto et al. (2012). Moreover, Li et al. (2011) analyzed and predicted tourism in Tangshan city of China through the rough set model.

In recent times, various studies on forecasting have introduced a comparative study of rough set and other quantitative time series models. For example, Faustino et al. (2011) present a rough set analysis of electrical charge demand in the United States and the level of the Sapucal river in Brazil. Liou (2016) used the rough set theory to study the airline service quality to Taiwan. Sharma and Kar (2018) applied rough set theory in hotel industry. These authors found that the rough set is more consistent with other forecasting techniques. Most published article on airline industry is primarily based on the statistical models. However, in previous literature the application of hybrid rough set and DES techniques for air passengers forecasting has not been adopted, to the best our knowledge. Consequently, the purpose of this paper is to propose a new approach which applied rough set theory to generate decision rules for the air passengers forecasting. Based on the yearly time series of transportation data for all passengers in Australian airline from 1992 to 2004, empirical analysis is employed to compare the proposed hybrid approaches with other forecasting methods using a corrected classified accuracy.

The primary objective of the paper is to provide accurate and highly reliable forecasting of air transport passengers. The forecasting performance has been evaluated under different criteria. The contribution of the study is that the rough set is firstly applied in exponential smoothing time series modeling approach. The present article can be employed to provide the essential recommendation on tourism policy for the developing countries.

The remaining of the article is structured as follows. Section hybrid approach based on rough set proposed the hybrid method for long-term forecasting of air transport passengers. Section data describe the data. The next section illustrates the empirical analysis related to the study. Section comparison of different models compares the accuracy of different models. Finally, the last section discusses the summary, conclusions and future work.

2. HYBRID APPROACH BASED ON ROUGH SET

In this section, the proposed methodology is discussed. Firstly, DES and rough set techniques are briefly presented. Then the hybrid approach is set up for investigating the decision rules and several steps involved in their implementation are illustrated in details.

2.1. Double exponential smoothing model

Double exponential smoothing (DES) model has been proposed for level and trend estimation process (Holt, 1957). A significant advantage of smoothing model is that the future prediction is obtained using previously estimated the level and trend in time series data with some appropriate smoothing parameters, μ, and ρ, respectively which lie in the range 0 and 1. The estimated level and trend of DES model for time series X_t *(t=1, 2....T)* is given by:

$$L_t = \mu(X_t) + (1-\mu)(L_{t-1} + T_{t-1})$$

$$T_t = \rho(L_t - L_{t-1}) + (1-\rho)T_{t-1}$$

The future value of the series X_t for h period is obtained as

$$\hat{X}_{t+h} = L_t + hT_t$$

2.2. Rough set theory

The rough set is a new mathematical approach to capturing imprecision, vagueness, and uncertainty (Pawlak, 1982). For the evaluation of a vague description of the member RST is the excellent mathematical tool. The adjective vague express the information quality that is uncertainty or ambiguity that chase from information granulation. The indiscernibility relation developed is this manner is a mathematical foundation of the RST; it induces a separation of the universe into pieces of indiscernible (similar) objects, named elementary set. The main concept of the rough set model is based on indiscernible (IND) relation, two approximations set, lower and upper approximation of a set.

Let Y be the non-empty finite set of objects referred to as universe and A be a nonempty finite set of attributes, then S = (Y, A) is called an information system where C, D are two subsets of A, where C and D are condition and decision attribute, respectively. For $S = (Y, A)$ and $P \subseteq A$, $W \subseteq Y$ can be approximated based on the knowledge having in P by assembling the P-lower and P-upper approximation of W, represents by $\underline{\mathrm{P}}(\mathrm{W})$ and $\overline{\mathrm{P}}(\mathrm{W})$ respectively; where

$$\underline{\mathrm{P}}(\mathrm{W}) = \{x | [x]_P \subseteq \mathrm{W}\}$$

$$\overline{\mathrm{P}}(\mathrm{W}) = \{x | [x]_P \cap \mathrm{W} \neq \emptyset\}$$

A set P is called an elementary set if it having set of all indiscernible(similar) object with respect to particular attributes. $[x]_P$ is an equivalence class that contain x with respect to indiscernibility relation P. The objects in $\underline{\mathrm{P}}(\mathrm{W})$ is known as the set of all members of Y which can be surely classified as a member of W in the knowledge P whereas objects in $\overline{\mathrm{P}}(\mathrm{W})$ is the set of all elements of Y that can be probably classified as a member of W involving knowledge P. The boundary region of W is expressed as: $\mathrm{BN_P}(\mathrm{W}) = \overline{\mathrm{P}}(\mathrm{W}) - \underline{\mathrm{P}}(\mathrm{W})$ is the set of a member which cannot decisively classify into W consisting knowledge P. If the boundary region of an exact set is the empty set if lower approximation and upper approximation set are similar. In the adverse case, if the boundary region contains some objects than the set W is referred as rough set concerning P.

The fundamental idea of RST is the creation of decision rules which is based on the IF and THEN logical statement. Decision rules are used to preserve the core semantics of the feature set from the provided information of particular problem which is an additional significant aspect of RST.

2.3. Proposed hybrid model

Since the combination method yields, better results than a single method, the modeling, and forecasting approach with high accuracy is adopted in this study. There are three main steps involved in the proposed hybrid method, i.e., estimation, single forecast, and combination. After the parameters are estimated by DES model as X_t, L_t and T_t respectively, they are utilized for the making of prediction ($\widehat{X_t}$). Then the estimated and predicted output is combining using the rough set method. The overall procedures of deriving a hybrid model are as described in Figure 1 as the several steps:

STEP 1: Suppose the original time series is X_t (actual) at time t, and then the level (L_t) and trend (T_t) has been estimated using DES model to establish the forecast of air passengers.

STEP 2: According to the estimated and forecast ($\widehat{X_t}$) output of DES model, four attributes are formulated to ascertain an information system (data set). Then, the data set are normalized (Mahapatra and Sreekumar, 2010). All attributes have been categorized into three classes (Low, Average and High). The information system (IS) for rough set is as follows:

IS = (X_t, L_t, T_t, $\widehat{X_t}$), where, X_t, L_t, T_t and $\widehat{X_t}$ are condition and decision attributes, respectively.

STEP 3: In the next stage, an information system is used to derive the decision rules for all passengers.

To create the IF and THEN decision rules, information system has been analyzed by applying RST technique (Cheng et al., 2010). The decision rules are as follows:

IF X_t = [Low, Average, High] AND L_t = [High, Average, Low] AND T_t = [Average, High, Low] THEN $\widehat{X_t}$ = [Low, Average, High].

3. DATA

In this study, time series data is used from tourism demand according to the objective and the availability of the data. Our empirical study uses yearly air transportation data for all passengers (millions) in Australian airline from the period of 1992 to 2004, with a total of 13 observations, as shown in Figure. 2. Data are obtained from https://www.otexts.org/fpp/7/2. The R-3.0.3 software is used for the overall empirical analysis of DES and Grey models. The rough set analysis is implemented through Rough Set Data Explorer (ROSE2) software (Predki, 1998). In our data analysis, training dataset (in-sample) is used to estimate the parameters using the time series models. Testing sample is used to evaluate the forecasting performance. In time series modeling, pre-samples are used for making out-of-sample forecasting by using different models.

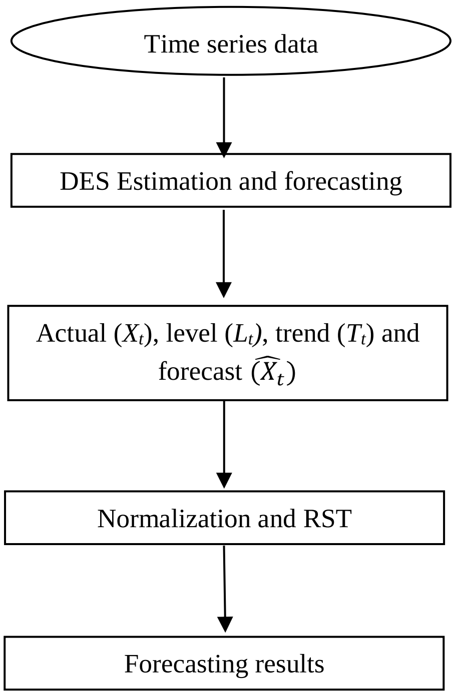


Fig. 1. The framework of DES and hybrid rough set approaches.

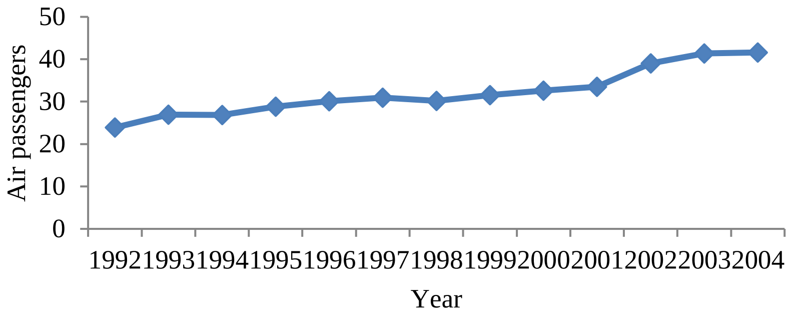


Fig.2. Yearly air passengers in Australia (in millions).

4. EMPIRICAL ANALYSIS

In the hybrid model, DES is first estimated into parameters level (L_t) and trend (T_t) from the air passenger's data (X_t). Then these parameters are used to predict the air passengers ($\widehat{X_t}$) and these actual, estimated and predicted results of the DES model are combined via rough set. In addition, Grey model (GM) is applied for the prediction of air passengers.

In the DES forecasting, the values of μ and ρ parameters are first determined using 10-fold cross-validation grid search method within the range of [0.8, 0.2] and then level (mean) and trend is estimated using these smoothing parameters. The calculated initial level and trend are 41.93, 1.87.

Apart from single DES model, other single time series model, GM is used for forecasting comparison purpose. In the GM model, forecasts are generated by using the solution of linear differential equation (LDE) of order one. The LDE for GM model is given below

$$\frac{dX_t}{dt} + \alpha\, Y_t \;= \beta \qquad (1)$$

Where, X_t *(t=1,2....T)* is a time series data and α and β are parameters. Also, the parameters α and β are calculated by applying the ordinary least squares (OLS) estimation method (Hsu and Wen, 1998; Zhou and Hu, 2008). The final solution of LDE in terms of forecasting equation for GM model for h periods is written as

$$\widehat{X_t}(h) = (1 - e^{\alpha})\left(X_1 - \frac{\beta}{\alpha}\right)e^{-\alpha(h-1)} \quad , h=1, 2, \ldots$$

Now, the results of estimated parameters, α and β for GM model are -0.04447, 21.82630.

After the DES estimation and forecasting of air passengers, we create an information system via actual (X_t), level (L_t), trend (T_t) and ($\widehat{X_t}$) as a condition and decision attributes. Then the normalization process is used for the data transformation. Now, the rough set model is used to develop the IF and THEN decision rules, expressed as a rule model (RM) for forecasting of passengers time series data. Four decision rules with their support (S) are described in Table 1. These 4 rules are used to predict air passengers to Australia. R_2 has the highest support. It indicates that R_2 is the strongest rule for the prediction.

Table 1: Decision rules for a hybrid approach

No.	Rules	S
R_1	IF [0 ≤ actual ≤ 0.3] then forecast = low	4
R_2	IF [0.3 < actual ≤ 0.7] then forecast = average	6
R_3	IF [0.3< level ≤ 0.7] and [0 ≤ trend ≤ 0.3] then forecast = average	5
R_4	IF [level > 0.7] then forecast = high	2

5. COMPARISON OF DIFFERENT MODELS

To compare the forecasting performance of time series models, we select corrected classified accuracy (CCA) criterion of measuring directional prediction accuracy. CCA is the ratio of corrected classified and the total number of observations (Nassiri and Rezaei, 2012). In general, accuracy is expressed as a percentage of the corrected direction, as follows:

$$\text{CCA} = \frac{\sum_{i=1}^{n} D_i}{n} * 100$$

$$\text{where } D_i = \begin{cases} 1, & (X_t - X_{t-1})(\widehat{X_t} - \widehat{X_{t-1}}) \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

where, D_i is the corrected direction, n is sample size, X_t is the actual value and $\widehat{X_t}$ is the forecasted value of time series data in a year. Since CCA is the measures of corrected direction between actual and forecasted values. Therefore, the higher CCA denotes highly accurate forecasting performance of time series models.

For forecasting performance, corrected classified accuracy (CCA) for three models are 92.30%, 92.30%, 100%, respectively. According to the CCA results, DES and GM model has equal forecasting performance. Also, the forecasting performance of the hybrid method is highly accurate than DES and GM models.

Empirical analysis of the study reveals that the forecasting accuracy of the hybrid model is superior to that of the individual models. The forecasting reports indicate that the prediction accuracy of the proposed approach is highly accurate to that of the single DES and Grey models. Therefore, our empirical findings reveal that the combine hybrid technique has a better prediction capacity compared with that of the individual time series models. The comparative analysis of actual and predicted values and out-of-sample forecasts from 2005 to 2017 are shown in Figure 3.

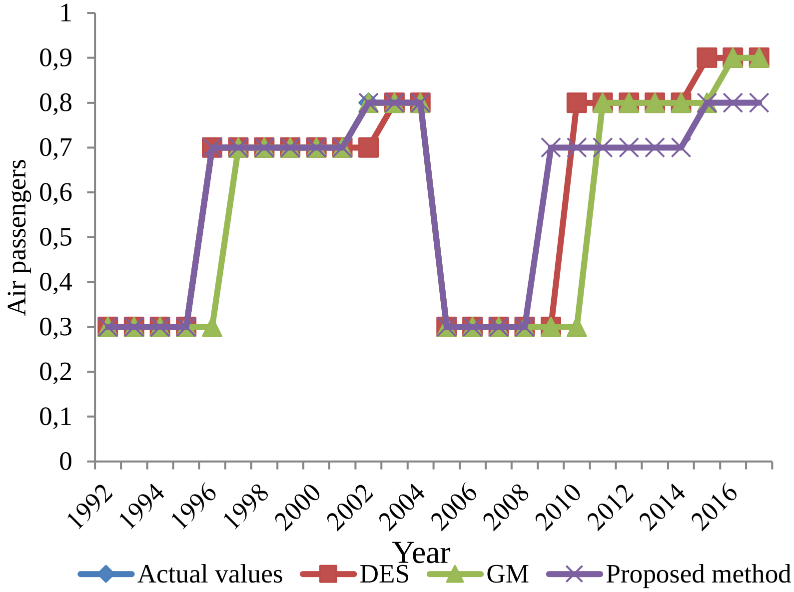


Fig. 3. Out-of-sample forecasts for different models.

6. CONCLUSIONS AND FUTURE WORK

In this study, we proposed a hybrid approach for long-term forecasting of air passenger on the basis of double exponential smoothing (DES) and rough set model. To illustrate the proposed hybrid technique, an empirical analysis has been performed and the proposed technique was compared with other time series methods using the air passengers in Australian airline. It uses corrected classified accuracy (CCA) criterion to compare the performance of hybrid method with other forecasting models.

Accurate and consistent air transport passenger is essential for policy making and planning by tourism management. Generally, tourism management is a significant source to enhance the foreign tourism demand as well as international travel to Australia and its various individual states. Therefore, the principal aim of this work is to evaluate the impact of tourism demand on Australian airline yearly air transport passengers. This work is also contributing that the combination of DES and rough set model is firstly applied for long-term forecasting of air passenger. The forecasting reports indicate that the prediction accuracy of the proposed approach is superior to that of the single DES and Grey models. Our analysis suggests that the combined hybrid technique is an efficient method for air passengers forecasting. It is important to utilize the quality and nature of air passenger for better forecasting performance. The present article can be employed to provide the essential recommendation on tourism policy for the Australian airline. Furthermore, experimental analysis of air passengers would allow airport management to assign resources effectively for the unique requirements of various visitor classes, such as holiday visitors, visiting family and relatives, business tourist, and traveler guests etc. It is expected to employ other time series and soft computing models for air passengers forecasting for the future research using the data of airline industry.

REFERENCES

Au, N. and Law, R. (2000). The application of rough sets to sightseeing expenditures, *Journal of Travel Research*, 39(1), 70-77.

Australian Bureau of Statistics (2013a). Overseas arrivals and departures, Australia. Available at: http://www. abs.gov.au (accessed 10 November 2014).

Australian Bureau of Statistics (2013b). Airport traffic statistics. Available at: http://www.abs.gov.au (accessed 10 November 2014).

Australian Bureau of Statistics (2013c). Gross state product, chain volume measures and current prices. Available at: http://www.abs.gov.au (accessed 10 November 2014).

Apergis N A, Mervar A and Payne JE (2017). Forecasting disaggregated tourist arrivals in Croatia: Evidence from seasonal univariate time series models. *Tourism Economics* 23(1): 78-98.

Chen, M,-Y., Chen, B,-T. (2014). Online fuzzy time series analysis based on entropy discretization and a Fast Fourier Transform, Applied Soft Computing. 14(B), 156-166

Cho V (2003). A comparison of three different approaches to tourist arrival forecasting. *Tourism Management* 24(3): 323-330.

Coshall J (2006). Time series analysis of UK outbound travel by air. *Journal of Travel Research* 44(3): 335-347.

Cheng, C. H., Chen, T. L. and Wei, L. Y. (2010). A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting. *Information Sciences*. 180, 1610-1629.

Celotto, E., Ellero, A. and Ferretti, P. (2012). Short-medium term tourist services demand Forecasting with rough ret theory, Procedia Economics, and Finance, 3, 62-67.

Deng, J. (1982). Control problems of Grey systems. *Systems and Control Letters*. 1(1), 288-294.

Faustino, P.C., Pinheiro, A.C., Carpinteiro, A.O. and Lima, I., (2011). Time series forecasting through rule-based models obtained via rough sets. *Artif Intell Rev*. 36, 299-310.

Goh, C., Law, R. (2003). Incorporating the rough sets theory into travel demand analysis. *Tourism Management*. 24, 511-517.

Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted moving averages. Pittsburgh PA: Carnegie Institute of Technology.

Hsu, C. I., Wen, Y. H. (1998). Improved Grey prediction models for the trans-pacific air passenger market. *Transportation Planning and Technology*. 22, 87-107.

Law, R. and Au, N. (1998). A rough set approach to hotel expenditure decision rules induction, *Journal of Hospitality and Tourism Research*, 22(4), 359-375.

Law R and Au N (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management* 20(1): 89-97.

Lee, C,-C., Wan, T.-J., Kuo, C.-Y., Chung, C.-Y. (2007). Modified Grey Model for Estimating Traffic Tunnel Air Quality, *Environmental Monitoring and Assessment*, 132(1-3): 351-364.

Lim C and McAleer M (1999). A seasonal analysis of Malaysian tourist arrivals to Australia. *Mathematics and Computers in Simulation* 48(6): 573-583.

Law, R. and Au, N. (2000). Relationship modeling in tourism shopping: a decision rules induction approach, *Tourism Management*, 21(3), 241-249.

Law R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tourism Management* 21(4): 331-340.

Lim C and McAleer M (2001). Monthly seasonal variations: Asian tourism to Australia. Annals of Tourism Research 28(1): 68–82.

Lim C and McAleer M (2002) Time series forecasts of international travel demand for Australia. *Tourism Management* 23(4): 389-396.

Lin, L. (2010). Study on the evaluation of tourist satisfaction in tourism destination based on variable precision rough set, *International Conference On Computer Design and Applications*, 1, 124-128.

Li, J., Feng, L., and Zhou, G. (2011). Travel demand prediction in Tangshan City of China based on rough set, *Springer-Verlag Berlin Heidelberg*, 440-446.

Liou, J. J. H., Chuang, Y. C., Hsu, C. C. (2016). Improving airline service quality based on rough set theory and flow graphs. *Journal of Industrial and Production Engineering*. 33(2), 123-133.

Li, J., Feng, L., Zhou, G. (2011). Travel demand prediction in Tangshan City of China based on rough set, *Springer-Verlag Berlin Heidelberg*, 440-446.

Martin A and Witt F (1989). Forecasting tourism demand: A comparison of the accuracy of several quantitative methods. *International Journal of Forecasting* 5. 1-13.

Mahapatra, S., Sreekumar, S. S. (2010). Attribute selection in marketing: A rough set approach. *IIMB Management Review*. 22, 16-24.

Nassiri, H., Rezaei, A. (2012). Air itineracy choice in a low-frequency market: A decision rule approach. *Journal of Air Transport Management*. 18, 34-37.

Nguyen L T, Shu, HM, Huang, FY and Hsu M. B. (2013). Accurate forecasting models in predicting the inbound tourism demand in Vietnam. *Journal of Statistics and Management Systems* 16(1): 25-43.

Pal, S. S., & Kar, S. (2017).Time series forecasting using fuzzy transformation and neural network with back propagation learning, Journal of Intelligent & Fuzzy Systems, 33 (1): 467 - 477.

Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Science*. 11, 341- 356.

Predki, B., Wong, S. K. M., Stefanowski, J., Susmaga, R., Wilk, Sz., (1998). ROSE-software implementation of the rough set theory. In L. Pollkowski, A. Skowron (Eds.). Rough Sets and Current Trends in Computing. *Lecture Notes in Artificial Intelligence. Berlin: Springer*. 605-608.

Software R, version R-3.0.3. https://cran.r-project.org/bin/windows/base/old/3.0.3/

Samagaio, A., Wolters, M. (2010). Comparative analysis of government forecasts for the Lisbon Airport. *Journal of Air Transport Management*. 16, 213-217.

Sharma, H. K., & Kar, S. (2018). Decision Making for Hotel Selection using Rough Set Theory: A case study of Indian Hotels. International Journal of Applied Engineering Research, 13(6), 3988-3998.

Tsui W. H. K. and Balli F. (2016). International arrivals forecasting for Australian airports and the impact of tourism marketing expenditure *Tourism Economics* 1-26.

Xie, G., Wang, S., Lai, K. K. (2014). Short-term forecasting of air passenger by using hybrid seasonal decomposition and least squares support vector regression approaches. *Journal of Air Transport Management*. 37, 20-26

Xiaoya, H. and Zhiben, J. (2011). Research on econometric model for domestic tourism income based on rough set, *Springer*, 259-266.

Zhou, Z, J., Hu, C. H. (2008). An effective hybrid approach based on Grey and ARMA for forecasting gyro drift. *Chaos, Solitons & Fractals*. 35, 525-529.

Haresh Kumar Sharma
Department of Mathematics, National Institute of Technology Durgapur,
West Bengal 713 209, India
Email: hks.14ma1106@phd.nitdgp.ac.in

Kriti Kumari
Department of Mathematics, National Institute of Technology Durgapur,
West Bengal 713 209, India
Email: kriti.kri89@gmail.com

Samarjit Kar
Department of Mathematics, National Institute of Technology Durgapur,
West Bengal 713 209, India
Email: samarjit.kar@maths.nitdgp.ac.in

On Unstability of Certain Group of ODEs with Quasiderivatives

O. PALUMBÍNY

Abstract

The article is focused on the unstability of some group of 4-th ODEs with quasi-derivatives. A sufficient condition of the unstability is followed by proper example.

Mathematics Subject Classification 2000: 34D20
Additional Key Words and Phrases: Fourth-order differential equation, Liapunov, unstability, quasiderivative

1. INTRODUCTION

The article is a continuation of [2], where some kind of the stability criterion has been derived for the equations (E) with quasiderivatives (the equations (E) as well as the quasiderivatives see below).

In [2] it was also shown an impossibility of replacement (in general) of the quasi-derivatives by ordinary derivatives in differential equations. Therefore new criteria of stability (as well as unstability) are needed.

The aim of the paper is to investigate the unstability (in Liapunov sense) of arbitrary solutions of equations of the form

$$Q_4 y + R_3(t)Q_3 y + R_2(t)Q_2 y + R_1(t)Q_1 y + R_0(t)Q_0 y = F(t, y), \tag{E}$$

where (a prime means a derivative by t)

$$Q_0 y(t) = y(t), \quad Q_i y(t) = r_i(t)(Q_{i-1}y(t))', \; i = 1,2,3, \quad Q_4 y(t) = (Q_3 y(t))',$$

$$r_k(t), \; k = 1,2,3, \quad R_j(t), \; j = 0,1,2,3$$

are continuous functions on $I_b = [b,\infty)$, $b \in R$, $F(t, y)$ is continuous on $I_b \times R$, where $R = (-\infty,\infty)$. The expressions $Q_n y(t)$, $n = 0,1,2,3,4$ are called *quasiderivatives.*

10.2478/jamsi-2018-0002

Let us consider ODE-system (“ODE” means “Ordinary Differential Equation”)

$$\dot{y}_i = f_i(t, y_1, y_2, y_3, y_4), \quad i = 1,2,3,4. \tag{S}$$

MAIN ASSUMPTION**.** Let (S) be rewritten by vectors as $\dot{\mathbf{y}} = \mathbf{f}(t,\mathbf{y})$. Let there exists c (real or $-\infty$) and an area $A \subset R^4$, $\mathbf{o} \in A$, $\mathbf{o} = (0, 0, 0, 0)$, such that $\mathbf{f}$ is continuous on an area $B = (c,\infty) \times A$ and for any couple $(s,\mathbf{m}) \in B$ the initial-value problem

$$\dot{\mathbf{y}} = \mathbf{f}(t,\mathbf{y}), \quad \mathbf{y}(s) = \mathbf{m}, \tag{1}$$

has the only one solution. Let $\mathbf{f}(t, \mathbf{o}) = \mathbf{o}$ for all $t > c$.

We shall define (by usual way) the concept of Liapunov stability of the trivial solution of (S) (for more details see Definition 1 in [2]) as well as Liapunov instability of the trivial solution of (S) (Definition 2 in [2]).

Let us consider the differential equation

$$\begin{aligned} &Q_4 w(x) + R_3(x)Q_3 w(x) + R_2(x)Q_2 w(x) + R_1(x)Q_1 w(x) + \\ &\qquad + R_0(x)Q_0 w(x) = F(x,u(x)+w(x)) - F(x,u(x)) \end{aligned} \tag{N}$$

as well as the differential system

$$\begin{aligned} &\dot{w}_i = w_{i+1}/r_i(x), \ i = 1,2,3, \\ &\dot{w}_4 = F(x,u(x)+w_1) - F(x,u(x)) - R_0(x)w_1 - R_1(x)w_2 - R_2(x)w_3 - R_3(x)w_4 \end{aligned} \tag{T}$$

One can easily see that $w_1(t)$ solves (N) iff $(z_1(t), L_1 z_1(t), L_2 z_1(t), L_3 z_1(t))$ solves (T)

(”iff ” means the usual abbreviation of “if and only if ”).

We can also state the stability of the trivial solution of (N) by means of the stability of the trivial solution of (T) (see Definition 5 in [2]). Similarly, we can also define the unstability of the trivial solution of (N) by means of the unstability of the trivial solution of (T).

2. AUXILIARY ASSERTIONS

We determine a matrix norm as a sum of absolute values of the elements of the considered matrix.

AUXILIARY THEOREM. Let us take into account a differential system such that

(2) $\dot{\mathbf{x}} = \mathbf{Cx} + \mathbf{D}(t)\mathbf{x} + \mathbf{h}(t,\mathbf{x}),\ \mathbf{h}(t,\mathbf{o}) = \mathbf{o},$ where $\mathbf{C}$ is a constant matrix, (3) $\mathbf{D}(t)$ converges to zero matrix as t tends to infinity and $\mathbf{h}$ is continuous on $(b,\infty)\times A,$ where $b \in R,\ \mathbf{o} \in A \subset R^4$ such that (4) $\|\mathbf{h}(t,\mathbf{x})\|/\|\mathbf{x}\| \to 0$ uniformly for all $t \geq b$ as $\|\mathbf{x}\| \to 0.$ If there exists an eigenvalue of $\mathbf{C}$ having a positive real part, then the null solution of (2) is unstable in Liapunov sense.

PROOF. See [1], Head 13. □

Let $u(t)$ solves (E) on I_b. Let $y(t)$ be an arbitrary function defined on I_b. Let $z(t) = y(t) - u(t),\ t \in I_b$. Then an arbitrary $z(t)$ solves (N) on I_b iff $y(t)$ solves (E) on I_b. This assertion can be proved in [2, Lemma 1].

The function $u(t)$ solving (E) is stable in Liapunov sense iff 0 solving (N) is stable in Liapunov sense. It is proved in [2, Lemma 1].

3. RESULTS

AUXILIARY LEMMA. The function $u(t)$ solving (E) is unstable in Liapunov sense iff 0 solving (N) is unstable in Liapunov sense.

PROOF. The assertion simply follows from the last assertion of the foregoing paragraph. □

Now will be proved the unstability criterion of (E).

MAIN THEOREM. Consider (E) such that Main assumption is valid. Let $u(t)$ solve (E) on I_b and

$$\lim_{t\to\infty} r_i(t) = r_i \in (0,\infty),\ i = 1,2,3, \qquad \lim_{t\to\infty} R_j(t) = R_j \in R,\ j = 0,1,2,3, \tag{a}$$

$$\frac{|F(t,u(t)+z) - F(t,u(t))|}{|z|} \to 0 \text{ uniformly for all } t \geq b \text{ as } z \to 0. \tag{c}$$

Let us denote $p(z) = z^4 + R_3 z^3 + \frac{R_2}{r_3} z^2 + \frac{R_1}{r_2 r_3} z + \frac{R_0}{r_1 r_2 r_3}$. If it holds that

$$\text{at least one real part of zeros of } p(z) \text{ is positive,} \tag{b'}$$

then $u(t)$ is unstable (in Liapunov sense).

In Main theorem we use the symbols (a), (b') and (c) which are competent to the symbols (a), (b) and (c) in [2].

PROOF OF MAIN THEOREM. From the paragraph "Auxiliary assertions" it is clear that it suffices to check up the unstability of 0-solution of (N), which correspond to (E); 0-solution of (N) is unstable iff (0,0,0,0)-solution of (T) is unstable. (T) can be expressed as (V), where

$$\dot{\mathbf{z}} = \mathbf{C}\mathbf{z} + \mathbf{D}(t)\mathbf{z} + \mathbf{h}(t,\mathbf{z}), \qquad \mathbf{h}(t,\mathbf{o}) = \mathbf{o}, \tag{V}$$

and

$$\dot{\mathbf{z}} = \begin{bmatrix} \dot z_1 \\ \dot z_2 \\ \dot z_3 \\ \dot z_4 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 1/r_1 & 0 & 0 \\ 0 & 0 & 1/r_2 & 0 \\ 0 & 0 & 0 & 1/r_3 \\ -R_0 & -R_1 & -R_2 & -R_3 \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix},$$

$$\mathbf{D}(t) = \begin{bmatrix} 0 & 1/r_1(t) - 1/r_1 & 0 & 0 \\ 0 & 0 & 1/r_2(t) - 1/r_2 & 0 \\ 0 & 0 & 0 & 1/r_3(t) - 1/r_3 \\ R_0 - R_0(t) & R_1 - R_1(t) & R_2 - R_2(t) & R_3 - R_3(t) \end{bmatrix},$$

$$\mathbf{h}(t,\mathbf{z}) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ F(t,u(t)+z_1(t)) - F(t,u(t)) \end{bmatrix}, \qquad \mathbf{o} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

If $\|\mathbf{z}\| \neq 0,\ z_1 \neq 0$, then

$$0 \le \frac{\|\mathbf{h}(t,\mathbf{z})\|}{\|\mathbf{z}\|} = \frac{|F(t,u(t)+z_1) - F(t,u(t))|}{|z_1|+|z_2|+|z_3|+|z_4|} =$$

$$= \frac{|F(t,u(t)+z_1) - F(t,u(t))|}{|z_1|} \cdot \frac{|z_1|}{|z_1|+|z_2|+|z_3|+|z_4|} \le$$

$$\le \frac{|F(t,u(t)+z_1) - F(t,u(t))|}{|z_1|}.$$

If $\|\mathbf{z}\| \ne 0,\ z_1 = 0,$ then $\dfrac{\|\mathbf{h}(t,\mathbf{z})\|}{\|\mathbf{z}\|} = 0$.

From this and (c) it follows that (4) hold. It can be easily found out that (2), (3) hold. It is not difficult to show that $p(z)$ is the characteristic polynomial of **C**. From (b’) it implies that there exists an eigenvalue $a+bi$ of **C** such that $a>0.$ Then from Auxiliary theorem it implies that (0,0,0,0)-solution of (T) is unstable. □

EXAMPLE. Consider (E) such that

$$r_1(t) = 1 + t^{-1}, \quad r_2(t) = 2 + t^{-1}, \quad r_3(t) = 3 + t^{-1},$$

$$R_0(t) = 6 + t^{-1}, \quad R_1(t) = -24 - t^{-1}, \quad R_2(t) = 18 + t^{-1}, \quad R_3(t) = -4 - t^{-1},$$

$$F(t,y) = 6t^{-1} + 25t^{-2} + 97t^{-3} + 293t^{-4} + 674t^{-5} + 959t^{-6} + 599t^{-7} + 120t^{-8} + (y - t^{-1})^2.$$

It can be taken $b = 2.$ One can easily seen t^{-1} solves (E) on I_2 as well as r_1 = 1, r_2 = 2, r_3 = 3, R_0 = 6, R_1 = – 24, R_2 = 18, R_3 = – 4. It implies that (a) hold.

If $z \ne 0,$ then

$$\frac{|F(t,u(t)+z) - F(t,u(t))|}{|z|} = \frac{z^2}{|z|} = |z|.$$

From this it implies that $|z|$ uniformly converges to zero for $t \in [2,\infty)$ for z approaching to zero, i.e. (c) hold, too. The characteristic polynomial of **C** is $(z-1)^4$. One of its zeros is $1+0i$, i.e. (b’) hold. Then, owing to Main theorem, t^{-1} is an unstable solution of considered equation.

REFERENCES

[1] CODDINGTON, E.A. AND LEVINSON, N. 1955. *Theory of ODEs*, McGraw-Hill, NY, USA.

[2] PALUMBÍNY, O. 2016. On stability criterion of 4-th order quasilinear differential equations with quasiderivatives, *Research papers of Faculty of Materials Science and Technology of Slovak University of Technology in Bratislava, Vol. 24, No. 39*, 63–70.

Oleg Palumbíny
UIAM MTF STU
J. Bottu 25, 917 24 Trnava, Slovakia
E-mail: oleg.palumbiny@stuba.sk

Applying Fractional Calculus to Analyze Economic Growth Modelling

D. LUO, J. R. WANG AND M. FEČKAN

Abstract

In this work, we apply fractional calculus to analyze a class of economic growth modelling (EGM) of the Spanish economy. More precisely, the Grünwald-Letnnikov and Caputo derivatives are used to simulate GDP by replacing the previous integer order derivatives with the help of Matlab, SPSS and R software. As a result, we find that the data raised from the Caputo derivative are better than the data raised from the Grünwald-Letnnikov derivative. We improve the previous result in [12].

Mathematics Subject Classification 2010: 26A33, 26A51, 26D15
Keywords: Grünwald-Letnnikov derivative, Caputo derivative, Economic growth modelling, Genetic algorithm, Method of least squares.

1. INTRODUCTION

It is well known that EGM is one of the most important models in studying the dynamics of finance behaviour. After reviewing the classical EGM in the literature, one can see that the integer order derivatives and integrals are always used to characterize such procedure in the development of economics. However, there exist some gaps by using the classical calculus to simulate the data from the real models. Recently, the basic theory including existence theory, stability and control theory for all kinds of fractional differential equations and inclusions [1; 2; 3; 4; 5; 6; 7; 8; 9] is studied extensively. In addition, one can see that fractional calculus [10] is also widely used to construct economic models involving the memory effect in the evolutionary process. It has been proved that fractional models [11] are better than integer models and provide an excellent tool for the description of memory of EGM, which has been taken into account in [12; 13; 14; 15; 16; 17; 18; 19].

In [12], the authors study GDP growth for the Spanish and Portuguese cases by applying Grünwald-Letnnikov fractional EGM via data between 1960 and 2012. By

This work was supported by the National Natural Science Foundation of China (grant number 11661016), Training Object of High Level and Innovative Talents of Guizhou Province (grant number (2016)4006), Unite Foundation of Guizhou Province (grant number [2015]7640), the Slovak Research and Development Agency (grant number APVV-14-0378) and the Slovak Grant Agency VEGA (grant numbers 2/0153/16 and 1/0078/17).

10.2478/jamsi-2018-0003

setting the mean absolute deviation as performance index and using Nelder-Mead's simplex search method, the coefficients and orders proposed in the fractional EGM are obtained. By comparing the coefficients of fractional EGM and integer EGM, a new hybrid model involving integer calculus and fractional calculus is established to remove low influence variables in the models. It is shown that fractional models have a better performance than the classical models.

In the present paper, we go on the study of GDP growth for the Spanish case to improve fractional EGM in [12] by using different computational methods. More precisely, we use four different EGMs, namely Grünwald-Letnnikov integer/fractional type and Caputo integer/fractional type models. Moreover, Nelder-Mead's simplex search method is replaced by genetic algorithm to give orders in the current work. The method of least squares is used to give the estimation of the coefficients. In spite of software of Matlab, SPSS and R are also used in linear regression analysis. We note that the Spanish case is used in this paper only for possibility to compare our achievement and proposed models with previous ones.

2. INTEGER AND FRACTIONAL EGMS

Throughout of this paper, we denote land area by LA (km^2), arable land by AL (km^2), population by P, school attendance by SA, gross capital formation by GCF, exports of goods and services by EGS, general government final consumption expenditure by GGFCE, money and quasi money by MQM, number of variables of the model by NVM and number of parameters of the model by NPM. We remark that all the data used here are taken from 1960 to 2012. We also denote the mean square error by MSE, the mean absolute deviation by MAD, the coefficient of determination by R^2, Akaike Information Criterion by AIC and the weight of AIC for the i-th model by ω_i.

Consider the following general formulation of EGM: $z = f(x_1, x_2, \cdots)$ where f is a given function. For simplify, we introduce the following notations:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
LA	AL	P	SA	GCF	EGS	GGFCE	MQM

and

z	n	k	t
GDP	NVM	NPM	Year

Define

$$\begin{aligned}
\mathrm{MSE} &= \frac{\sum\limits_{i=1}^{n}(z_i-\tilde{z}_i)^2}{n},\\
\mathrm{MAD} &= \frac{\sum\limits_{i=1}^{n}|z_i-\tilde{z}_i|}{n},\\
\mathrm{R}^2 &= 1-\frac{\sum\limits_{i=1}^{n}(z_i-\tilde{z}_i)^2}{\sum\limits_{i=1}^{n}(z_i-\bar{z})^2},\\
\mathrm{AIC} &= n\log\frac{\sum\limits_{i=1}^{n}(z_i-\tilde{z}_i)^2}{n}+2k+\frac{2k(k+1)}{n-k-1},
\end{aligned}$$

if we have m model, then

$$\omega_i = \frac{\exp\left(-\frac{\mathrm{AIC}_i-\min\limits_m \mathrm{AIC}}{2}\right)}{\sum_{j=1}^{m}\exp\left(-\frac{\mathrm{AIC}_j-\min\limits_m \mathrm{AIC}}{2}\right)}.$$

Next, we recall the following standard integer order model (IOM)

$$z(t)=\sum_{i=1,2,3,4,6,7} c_i x_i(t)+c_5(I^1_{t_0,t}x_5)(t)+\sum_{i=8,9} c_i x_i'(t).$$

We also need the following modified models:

- IOM1 (Grünwald-Letnnikov integer type)

$$z(t)=\sum_{i=1,2,3,4,6,7} c_i(\,{}^{GL}D^0_{t_0,t}x_i)(t)+c_5(\,{}^{GL}D^{-1}_{t_0,t}x_5)(t)+\sum_{i=8,9} c_i(\,{}^{GL}D^1_{t_0,t}x_i)(t),$$

- IOM2 (Caputo integer type)

$$z(t)=\sum_{i=1,2,3,4,6,7} c_i(\,{}^{C}D^0_{t_0,t}x_i)(t)+c_5(I^1_{t_0,t}x_5)(t)+\sum_{i=8,9} c_i(\,{}^{C}D^1_{t_0,t}x_i)(t),$$

- FOM1 (Grünwald-Letnnikov fractional type)

$$z(t)=\sum_{i=1}^{9} c_i(\,{}^{GL}D^{\alpha_k}_{t_0,t}x_i)(t),$$

- FOM2 (Caputo fractional type)

$$z(t)=\sum_{i=1}^{9} c_i(\,{}^{C}D^{\alpha_i}_{t_0,t}x_i)(t),$$

where t_0 denotes the first year and the fractional calculus [10] is given by

$$(I_{a,t}^{\alpha}u)(t) := \frac{1}{\Gamma(\alpha)}\int_a^t \frac{u(\tau)}{(t-\tau)^{1-\alpha}}d\tau,\ 0<\alpha\leq 1,$$

and the Grünwald-Letnnikov (GL) derivative

$${}^{GL}D_{a,t}^{\alpha}u(t) = \lim_{h\to 0}\frac{\sum_{j=0}^{[(t-a)/h]}}{h^{\alpha}}(-1)^j C_{\alpha}^{j}u(t-jh),$$

$$C_{\alpha}^{j} = \frac{(-1)^j\Gamma(\alpha-j)}{\Gamma(j+1)\Gamma(-\alpha-j+1)},\ 0<\alpha\leq 1,$$

$$C_{\alpha}^{j} = \frac{\Gamma(\alpha+1)}{\Gamma(j+1)\Gamma(\alpha-j+1)},\ -1\leq\alpha<0,$$

$$C_{\alpha}^{j} = 1,\ \alpha = 0,$$

and the Caputo derivative

$${}^{C}D_{a,t}^{\alpha}u(t) = \frac{1}{\Gamma(1-\alpha)}\int_a^t \frac{u'(s)}{(t-s)^{\alpha}}ds,\ t>a,\ 0<\alpha\leq 1.$$

3. MAIN RESULTS

3.1. Economic data for Spanish economy

By using the Spanish data from 1960 to 2012 in [12, Table 5], we apply Matlab to obtain the following figures (see Figure 1).

3.2. The coefficients and orders

By using genetic algorithm in the Matlab, we obtain the following data (see Tables I and II). Here we remark that the coefficients are estimated by using the method of least squares.

Fig. 1. Data for Spanish

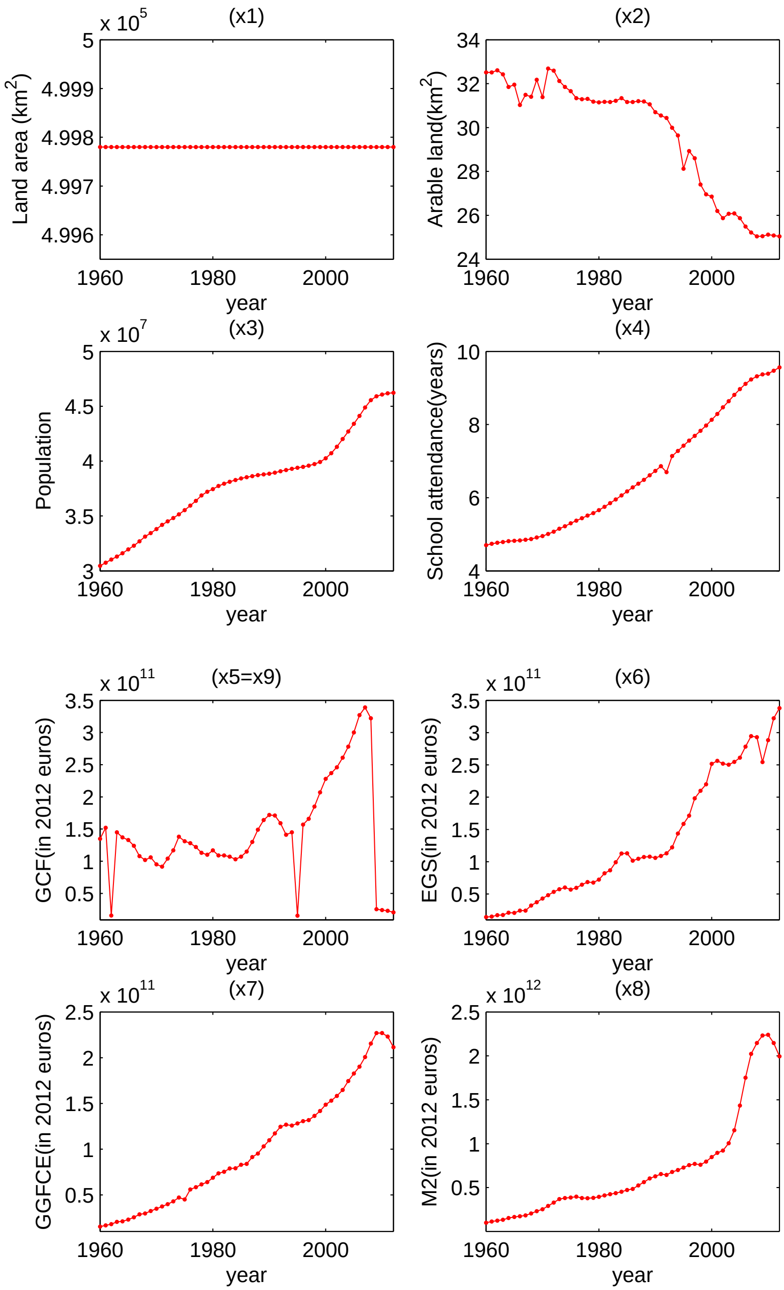


Table I. The orders of the fractional operators

	IOM1	FOM1	IOM2	FOM2
α_1	0	0.31072	0	0.26735
α_2	0	-0.75424	0	-0.69281
α_3	0	-0.73633	0	-0.70304
α_4	0	-0.99999	0	-0.67233
α_5	-1	-1	-1	-0.60068
α_6	0	-0.83616	0	-0.95969
α_7	0	0.31073	0	0.43180
α_8	1	-0.13985	1	0.31072
α_9	1	-0.34465	1	-0.94727

Table II. The coefficients of the fractional operators

	IOM1	FOM1	IOM2	FOM2
$c_1(\times 10^{05})$	9.903	1.954	10.393	227.211
$c_2(\times 10^{09})$	7.531	11.371	-0.872	9.732
$c_3(\times 10^{04})$	-1.416	-1.150	-0.851	-1.009
$c_4(\times 10^{10})$	-2.455	0.141	-0.207	1.576
$c_5(\times 10^{-1})$	2.887	0.296	1.658	1.963
$c_6(\times 10^{-1})$	-2.123	4.707	-0.269	2.464
$c_7(\times 10^{00})$	-3.845	1.513	-1.16	1.072
$c_8(\times 10^{-2})$	9.582	4.236	16.556	5.358
$c_9(\times 10^{-2})$	1.759	9.542	12.459	-6.060

Now we are ready to give analysis by virtue of estimated value from Matlab via true value.

Table III: Performance indices for the Spanish economy

index	Variable	IOM1	FOM1	IOM2	FOM2
MSE($\times 10^{20}$)		4.042	0.638	3.075	0.479
R^2		0.9947	0.9992	0.9960	0.9994
MAD($\times 10^{10}$)		1.666	0.642	1.540	0.569
AIC		2537.0	2439.1	2522.5	2423.9
AIC without	x_1	2536.8	2477.2	2523.6	2463.5
one variable	x_2	2535.3	2521.2	2519.6	2496.1
	x_3	2537.4	2508.4	2521.1	2496.3
	x_4	2534.7	2439.7	2519.5	2446.4
	x_5	2565.1	2437.0	2528.9	2440.3
	x_6	2534.6	2514.6	2519.6	2485.1
	x_7	2550.9	2445.4	2520.8	2450.6
	x_8	2538.2	2445.9	2537.4	2428.8
	x_9	2534.1	2441.8	2536.7	2422.7
ω found from the	x_1	7%	0%	3%	0%
AIC without one	x_2	15%	0%	24%	0%
variable	x_3	5%	0%	11%	0%
	x_4	20%	18%	24%	0%
	x_5	0%	73%	0%	0%
	x_6	21%	0%	24%	0%
	x_7	0%	1%	13%	0%
	x_8	4%	1%	0%	5%
	x_9	27%	7%	0%	95%

REMARK 3.1. The value 0% implies that the i-th variable for the corresponding models cannot be removed in the simulation from Table III.

3.3. Analysis of significance level

Now we apply Matlab, SPSS and R software to give the analysis of significance level (see Tables IV and V)

Table IV. Significance level of GL model

		IOM1			FOM1		
	Variable	Matlab	SPSS	R	Matlab	SPSS	R
t value	x_1	-0.765	0.557	1.637	3.326	5.378	6.841
	x_2	0.898	0.078	1.058	11.35	6.250	12.436
	x_3	-1.451	-1.711	-1.760	-10.479	-11.511	-10.623
	x_4	-0.656	0.368	-0.901	1.663	-3.215	1.736
	x_5	4.869	8.353	6.001	0.787	-1.048	0.631
	x_6	-0.530	-0.791	-0.584	9.760	3.949	11.473
	x_7	-3.156	-3.637	-4.035	2.855	.531	2.546
	x_8	2.006	2.149	2.049	2.755	.343	3.143
	x_9	0.512	0.510	0.391	2.172	7.868	2.541
P value	x_1	0.449	0.580	0.108699	0.001	0.000	1.96e-08
	x_2	0.374	0.938	0.295763	1.597e-14	0.000	5.33e-16
	x_3	0.154	0.094	0.085435	2.047e-13	0.000	1.00e-13
	x_4	0.515	0.715	0.372647	0.104	0.002	0.08954
	x_5	1.550e-5	0.000	3.37e-7	0.436	0.300	0.53117
	x_6	0.599	0.433	0.562351	1.793e-12	0.000	8.16e-15
	x_7	0.003	0.001	0.000215	0.007	0.598	0.01448
	x_8	0.051	0.037	0.046428	0.009	0.733	0.00299
	x_9	0.612	0.613	0.698030	0.035	0.000	0.01467

REMARK 3.2. The red data in Table IV denote 5% significance level.

Table V. Significance level of Caputo model

	Variable	IOM2			FOM2		
		Matlab	SPSS	R	Matlab	SPSS	R
t value	x_1	0.257	1.463	2.012	3.039	-6.802	7.020
	x_2	-0.130	-1.103	-0.126	4.478	7.638	11.312
	x_3	-1.161	-1.159	-1.241	-5.981	5.878	-11.178
	x_4	-0.043	.890	-0.248	4.338	8.019	4.798
	x_5	2.879	3.159	3.021	3.722	2.982	4.632
	x_6	-0.111	-.094	0.001	4.812	-5.245	10.100
	x_7	-1.013	-.638	-1.032	5.552	-4.452	5.638
	x_8	4.038	4.622	4.376	2.575	-2.736	2.783
	x_9	3.873	4.489	4.282	-1.311	7.173	-1.478
P value	x_1	0.798	0.150	0.05041	0.004	0.000	1.38e-08
	x_2	0.897	0.276	0.90010	5.479e-05	0.000	2.340e-14
	x_3	0.252	0.253	0.22103	3.901e-07	0.000	2.78e-14
	x_4	0.966	0.378	0.80550	8.554e-05	0.000	1.83e-05
	x_5	0.006	0.003	0.00418	0.001	0.005	4.39e-05
	x_6	0.912	0.925	0.99934	1.867e-05	0.000	8.69e-13
	x_7	0.317	0.527	0.30772	1.634e-06	0.000	1.35e-06
	x_8	2.181e-4	0.000	7.33e-05	0.014	0.009	0.0103
	x_9	3.616e-4	0.000	9.88e-05	0.197	0.000	0.1778

REMARK 3.3. The red data in Table V denote 5% significance level.

3.4. Fitting results

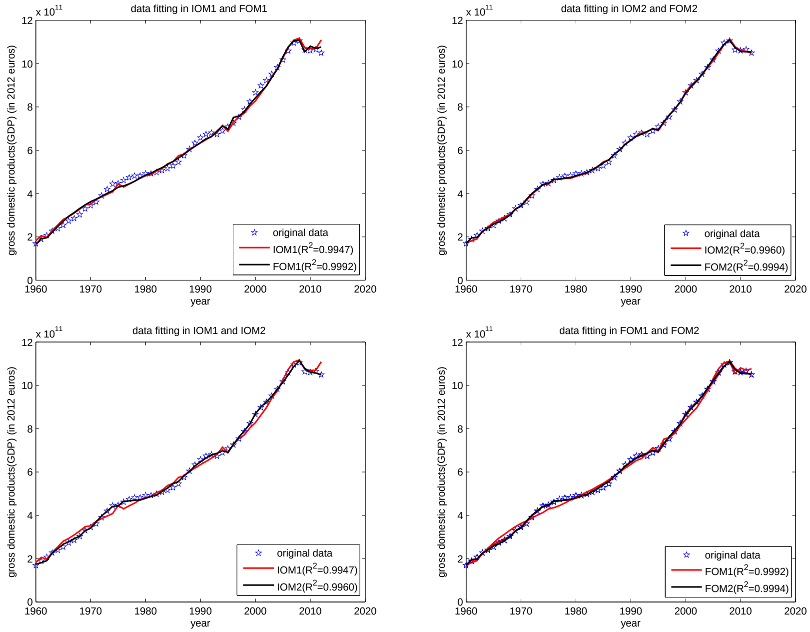
Now we are ready to give the fitting results for IOM, FOM1 and FOM2 (see Fig. 2).

REMARK 3.4. (i) From the figure of data fitting in IOM1 and FOM1, one can see that the simulation result of FOM1 is better than the simulation result of IOM1.

(ii) From the figure of data fitting in IOM2 and FOM2, one can see that the simulation results of IOM2 and FOM2 are very close to original data. However, R^2 of FOM2 is closer to 1 than R^2 of IOM2. Thus, FOM2 is better than IOM2.

(iii) From the figure of data fitting in IOM1 and IOM2 , one can see that the simulation result of IOM2 is better than the simulation result of IOM1.

Fig. 2. Data fitting



(iv) From the figure of data fitting in FOM1 and FOM2, one can see that FOM2 is closer to original data than FOM1 although the value of R^2 for both FOM1 and FOM2 tend to 1.

From above, one can deduce that FOM2 is the most suitable model for this case.

3.5. Comparison of models

Finally, we present the following tables (see Table VI and Table VII) to compare the current results with the previous ones in [12], which show that our results derived by genetic algorithm are much better than the results derived by Nelder-Mead's simplex search method [12].

Table VI. The model of [12]

	Integer (5)	Fractional (6)	Fractional (12)	Integer (13)	Fractional (14)
AIC	2554.3	2473.8	2474.4	2552.9	2472
ω_i	0%	0%	0%	0%	0%

Table VII. Our results

	IOM1	FOM1	IOM2	FOM2
AIC	2537	2439.1	2522.5	2423.9
ω_i	0%	5%	0%	95%

3.6. Conclusions

This paper studies a class of economic growth modelling for the Spanish case. Based on our results, four models of fractional calculus (IMO1, IMO2, FOM1 and FOM2) are proposed. It is shown that the date of GDP raised from the Caputo derivative are better than the Grünwald-Letnnikov derivative. They are not identical in the significance level of models (t value and P value) via Matlab, SPSS and R software. In addition, the data of FOM2 are derived via genetic algorithm and the method of least squares, which is better than IOM1, IOM2, FOM1 and the reference [12].

REFERENCES

[1] K. Diethelm, The Analysis of Fractional Differential Equations, Lecture Notes in Mathematics, Springer, New York, 2010.

[2] Y. Zhou, J. Wang, L. Zhang, Basic Theory of Fractional Differential Equations, 2nd Edn, World Scientifc, Singapore, 2016.

[3] Y. Zhou, Fractional Evolution Equations and Inclusions: Analysis and control, Academic Press, 2016.

[4] R. P. Agarwal, S. Hristova, D. O'Regan, A survey of Lyapunov functions, stability and impulsive Caputo fractional differential equations, Fract. Calc. Appl. Anal., 19(2016), 290-318.

[5] J. Wang, M. Fečkan, Y. Zhou, A survey on impulsive fractional differential equations, Fract. Calc. Appl. Anal., 19(2016), 806-831.

[6] J. Wang, X. Li, Ulam-Hyers stability of fractional Langevin equations, Appl. Math. Comput., 258(2015), 72-83.

[7] J. Wang, X. Li, A uniformed method to Ulam-Hyers stability for some linear fractional equations, Mediterr. J. Math., 13(2016), 625-635.

[8] M. Li, J. Wang, Finite time stability of fractional delay differential equations, Appl. Math. Lett., 64(2017), 170-176.

[9] J. Wang, M. Fečkan, Y. Zhou, Center stable manifold for planar fractional damped equations, Appl. Math. Comput., 296(2017), 257-269.

[10] A. A. Kilbas, H. M. Srivastava, J. J. Trujillo, Theory and Applications of Fractional Differential Equations, Elsevier, Amsterdam, 2006.

[11] R. Hilfer, Applications of Fractional Calculus in Physics, World Scientific, Singapore, 1999.

[12] I. Tejado, D. Valério, E. Pérez, N. Valério, Fractional calculus in economic growth modelling: the Spanish and Portuguese cases, Int. J. Dyn. Control, 5(2017), 208-222.

[13] J. A. T. Machado, M. E. Mata, A fractional perspective to the bond graph modelling of world economies, Nonlinear Dyn., 80(2015), 1839-1852.

[14] J. A. T. Machado, M. E. Mata, A. M. Lopes, Fractional state space analysis of economic systems, Entropy, 17(2015), 5402-5421.

[15] J. A. T. Machado, M. E. Mata, Pseudo phase plane and fractional calculus modeling of western global economic downturn, Commun. Nonlinear Sci. Numer. Simulat., 22(2015), 396-406.

[16] V. V. Tarasova, V. E. Tarasov, Elasticity for economic processes with memory: Fractional differential calculus approach, Fract. Diff. Calc., 6(2016), 219-232.

[17] S. A. David, J. A. T. Machado, D. D. Quintino, J. M. Balthazar, Partial chaos suppression in a fractional order macroeconomic model, Math. Comput. Simul., 122(2016), 55-68.

[18] T. Škovránek, I. Podlubny, I. Petráš, Modeling of the national economies in state-space: A fractional calculus approach, Economic Modelling, 29(2012), 1322-1327.

[19] I. Petras, I. Podlubny, State space description of national economies: The V4 countries, Computational Statistics & Data Analysis, 52(2007), 1223-1233.

Dahui Luo
Department of Mathematics
Guizhou University
Guiyang, Guizhou 550025, P.R. China
e-mail: dhluomath@126.com

JinRong Wang
Department of Mathematics
Guizhou University
Guiyang, Guizhou 550025, P.R. China
e-mail: wjr9668@126.com; sci.jrwang@gzu.edu.cn

Michal Fečkan
Department of Mathematical Analysis and Numerical Mathematics
Faculty of Mathematics, Physics and Informatics
Comenius University in Bratislava, Mlynská dolina, 842 48 Bratislava, Slovakia
Mathematical Institute of Slovak Academy of Sciences
Štefánikova 49, 814 73 Bratislava, Slovakia
e-mail: Michal.Feckan@fmph.uniba.sk

Received June 1, 2017

Application of Motion Capture Attributes to Individual Identification under Corridor Surveillance

A. MIHÁLIK, R. ĎURIKOVIČ AND M. SEJČ

Abstract

Accurate and fast identification of a person from a security point of view is a key procedure. The most common technique of person identification uses identity cards. In contrary to the common approach we focus our research on identification based on the body movement such as the gait in this paper. The gait and the posture belong to the unique characteristics of the person that helps us to facilitate the identification. The proposed methodology allows us to incorporate personal characteristics into the access control systems using the color depth camera (RGBD). For the sake of gait analysis, the important task is to recognize the figure and extract the skeleton data from a video recording. Besides the usage of the mathematical statistics methods, we are opting to use computer animation and computer vision methods, which makes the research interdisciplinary. The main novelty of the paper is the definition and extraction of the feature vector from motion capture data, the analysis methodology and finally the selection of few statistically dominant motion attributes for the identification purposes. Besides the development of new approaches in this field, we validate proposed approaches from the perspective of accuracy.

Mathematics Subject Classification 2000: 94A08, 68U10, 68P01, 68U05
Keywords: Motion identification, Security, Feature vector, Mocap data

1. INTRODUCTION

The right and fast identification of people from a security point of view is a key procedure, in today's world. Each person is unique. The most common way of identifying a person is based on identity cards. In a worse case, only the password or pin use is sufficient. Such identity keys can be easily stolen or falsified. It is, therefore, necessary to focus on the unique qualities of every person that cannot be imitated, lost or forgotten.

Biometrics authentication refers to the identification based on the unique features of a particular person. We can categorize biometric identifiers into two separate groups; either the physiological or the behavioral characteristics. Physiological characteristics are related to the body and its shape or structure such as fingerprint [Wang et al. 2010] or DNA [Tautz 1989]. Behavioral characteristics are related to the pattern of behavior of a person such as voice, handwriting or leg motion [Yam et al. 2002].

A person's walk is unique, as confirmed by an experiment [Barclay et al. 1978]

10.2478/jamsi-2018-0004

DE G DE GRUYTER OPEN

that was carried out in the 1970s. However, as has been shown later, a person can be identified with much greater accuracy based on the use of selected symptoms. The survey [Gianaria et al. 2014] showed that it is possible to achieve up to 96% success. In particular, they used fixed human skeleton parameters such as the height of the person or the length of the limbs. In our method, we will use some of their verified attributes. However, our project focuses mainly on walking. To analyze a walking cycle statistical methods are commonly utilized. Some approaches are based on feature learning methods for gait recognition [Feng et al. 2016], others use probabilistic methods [Bazin and Nixon 2005]. On the other hand, descriptors describing actual behavior of individuals have already been studied. In particular, there are descriptors aimed at detecting aggressive behavior [Chen et al. 2008] or detecting human action [Acar et al. 2012]. These methods share a common key feature. They try to describe all people as a group and are therefore not suitable for distinguishing individuals.

A crucial step in analyzing the walking sequence is data acquisition. In our case, it will be the captured animation of a human skeleton. Our approach to quickly extract skeleton and human movement based on depth data rely on freely available software development kit (SDK) for Kinect device [Rahman and Gavrilova 2017; Peterkova and Stremy 2015]. However, the SDK is well known for its inaccuracy. Another choice is extracting the skeleton from depth data using an anatomical human model [Zhu et al. 2015]. The skeleton thus obtained is very precise compared to the previous option, but it is paid off by a calculation time that is too large for the practical real-time application, particularly for several people at the same time. An alternative is the use of color video information [Andriluka et al. 2009]. Color-based methods are difficult to calibrate and apply in practice under unstable light conditions. Therefore, we propose a combination of color and depth information [Dubois and Bresciani 2015] that will enable us to take advantage of both approaches.

The process of the identification comprises of particular steps devoted to the feature extraction and the database of subjects search procedure. First, the multiple subjects during the gait had to be recorded. Then the skeletal representation of the data is extracted using the particular procedure [Riečický et al. 2018; Ďurikovič and Madaras 2015]. The feature vector is created from the data of each subject and subsequently stored in the database. The main focus of our work is to find the relevant parameters to construct the feature vector using in the database.

2. SKELETON MODEL

To capture the behavioral characteristics of gait we based our data structure on physiological characteristics of human skeleton depicted in Fig. 1. Our designed hierarchical data structure is meant to be compatible with Biovision Hierarchy (BVH) data format. BVH format is commonly used in many animation systems and software platforms. The proposed skeletal structure is extracted from the sequences capturing human actors walking data [Kim and Kim 2014]. We extract the 3D position of joints, representing nodes in the hierarchical structure, from default posture. The structure contains 5 leaf nodes and the root of the structure is the node in 3D space denoted as P_1. We can represent walking animation as a sequence of poses attribute to subsequent time frames. Each joint can rotate about preceding node, therefore we can create an arbitrary pose from default posture using a particular set of rotations. If we rotate root, we rotate the whole skeleton. If we rotate shoulder joint, the whole hand is moved. Therefore, we can represent skeleton in each animation frame using our hierarchical representation. The optimal number of frames per second (fps) capturing walking cycle of regular gait is 30, see Fig. 1. To capture the running person we need at least 60 fps to achieve smooth motion.

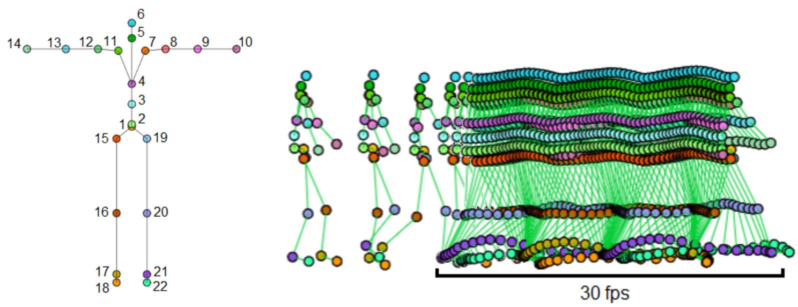


Fig. 1. Model of skeleton. Left: Joints P_1-loins (center), P_2-rump, P_3-lower back, P_4-upper back, P_5-neck, P_6-head, P_7-left arm, P_8-left hand, P_9-left elbow, P_{10}-left palm, P_{11}-right arm, P_{12}-right hand, P_{13}-right elbow, P_{14}-right palm, P_{15}-right hip joint, P_{16}-right knee, P_{17}-right heel, P_{18}-right tiptoe, P_{19}-left hip joint, P_{20}-left knee, P_{21}-left heel, P_{22}-left tiptoe. Right: Walking frame sequence.

3. GAIT ATTRIBUTES

We consider three particular parameter categories in order to analyze gait. Physiological, kinematic and time-dependent parameters. We have analyzed each group of parameters in order to find appropriate parameters suitable to form a feature vector. As an example of physiological parameter we can use the length of a limb or the height of a person which is a scalar value. The kinematic parameter is, for example, the pace, which could change during the walk. The time-dependent parameter is, for example, arm rotation, which could be different in each frame.

3.1. Physiological parameters

Physiological parameters of the same person are constant at each time of movement. They can be calculated from one arbitrary frame. In our method, we used 5 physiological parameters. These include the height of the torso, left and right-hand lengths, left and right leg lengths. When evaluating these parameters, the distance between the joints is calculated. Each of the joints is represented as a point P in a three-dimensional space. So the distance between the two joints is calculated as the Euclidean distance of two points. The resulting value will be the sum of the distances between every two adjacent joints of the sequence. Parameters representing the lengths of particular body part L_w are calculated using the formula:

$$L_w = \sum_{i \in w} \| P_{w_{i+1}} - P_{w_i} \|, \tag{1}$$

where w represents a set of joints from a given limb (see Fig. 1). The indexes of the joints from these sets are found in the Tab. I. We have calculated these distances for each frame separately. The resulting value was the average of the values from all frames to minimize minor inaccuracies.

Table I. Joints sequences. Symbol of the parameter its explanation and the corresponding sequence of joints for calculation.

Symbol	Parameter	w
γ	Torso height	$\{6, 5, 4, 3, 2, 1\}$
α_P	Right hand length	$\{12, 13, 14\}$
α_L	Length of left hand	$\{8, 9, 10\}$
β_P	Right leg length	$\{15, 16, 17\}$
β_L	Length of left foot	$\{19, 20, 21\}$

3.2. Kinematic parameters

Kinematic parameters include attributes that can be changed at each step of the walk. These include, for example, walking speed, step length, etc. Altogether, we investigated up to 30 parameters that could be relevant to the identification. However, it turned out that most of them depend on the spatial position of the skeleton. Therefore, we have excluded such parameters. Finally, we have 10 parameters that we have further analyzed. The parameter list is listed in the Tab. II.

Table II. List of kinematic parameters with parameter notation and the explanation.

Parameter	Description
υ	Speed
δ_L	Step length of left leg
δ_P	Step length of right leg
κ	Step width
$\sigma^2_{P_6,x}$	Variance of head on x axis
$\sigma^2_{P_6,y}$	Variance of head on y axis
$\sigma^2_{P_{20},y}$	Variance of left knee on y axis
$\sigma^2_{P_{16},y}$	Variance of right knee on y axis
$\sigma^2_{P_7,x}$	Variance of left arm on x axis
$\sigma^2_{P_{11},x}$	Variance of right arm on x axis

Kinematic parameters are not constant as physiological parameters. For one person, they can change at each step. To calculate them, we will need several frames capturing at least 2 steps gait cycle from the whole walk. The gait cycle means that the left and right legs touch the ground twice so that the length of the step can be calculated. The number of frames captured during the gait cycle cannot be uniquely determined. It depends on the number of frames per second and the speed of the recorded walk.

3.3. Walking speed

Walking speed is a numeric value indicating the length of the path per time interval. Specifically, the distance that the point P_1 (see Fig. 1) passes. However, we need to choose a suitable time step. Because the input data contains an integer number of fps (30, 60, ...), one second was the most appropriate choice.

Input data we have contains 30 fps, but it does not have to be the rule. Thus we select every thirty-one frame from the whole walk sequence. From each of the two selected consecutive frames, we calculated what distance point P_1 passes. These distances indicate the length of the path the person underwent for each second of the walk. Even though the walking speed should be theoretically constant, not all of these values are the same. In the same second, the observed person can slightly accelerate or slow down. We have attempted to use the arithmetic mean of this data, which has proven to be an incorrect choice. In some cases, the tracking person started to slow down at the end of the walk, causing the results to be distorted. Finally, we determine the resulting velocity as the median of the obtained data.

3.4. Step length

In our experiments, we use two parameters for step length. One for the left and the other for the right leg. It turns out that one does not have to do the same long steps for each of the lower limbs. The length of the step was counted as the distance of two points in which the foot of the same leg touched the ground. These points are shown in the Fig. 2. The image shows the touch of the left foot to the floor.

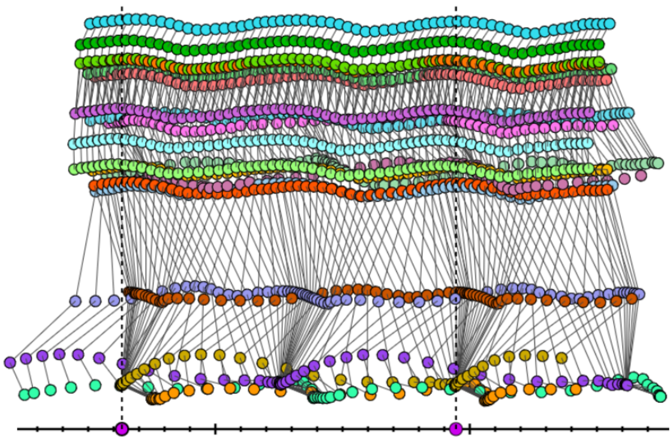


Fig. 2. Step length of left leg. X-axis shows the step start and the end.

To calculate the distance we first need to detect the positions where the foot touched the ground. However, the floor was not defined. It could be at any height relative to the coordinate system. Using height coordinates to detect contact with the floor was not possible.

During walking legs are alternating. For example, the right foot moves while the left is standing in place and does not change its position. The left foot will move only when we are touching the floor with the right foot. So at the moment of touching the floor, the foot position does not change. This means that the position of the foot is the same in several consecutive frames. From the Fig. 2 we can see that the density of the points representing the heel is higher in this particular location.

The positions where the density was the highest are exactly the points where the foot touched the floor. In this way, we received several consecutive points representing the touch with the floor. Finally, we calculate the distances between consecutive touch points representing the particular step length. However, we always needed a sequence of walks in which the person would have done at least two steps. For the same person, the length of the steps during the whole walk could be slightly different. We obtain the resulting value as the median of the calculated data. Along the whole walking sequence, we captured variable step lengths and create the list of step lengths. As a resulting length, we choose median computed from the list of captured lengths. However, it is true that the longer the record of the walk, the more accurate the results are.

3.5. Step width

This parameter determines what the distance between each heel is. To calculate this value, we need to detect where the feet touch the ground. Acquired points are projected onto the floor represented by the XZ plane. Therefore, we can consider these points as 2D, shown in the Fig. 3 and indexed from 1 to 5.

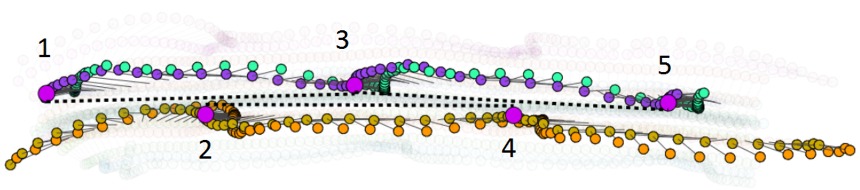


Fig. 3. Step width - top view. Dots show the left and the right leg movements, the distance between parallel dot lines can proximate the width of a step in that location.

We have found a linear regression for three consecutive points corresponding to left and right leg, respectively. Two lines we found are sketched in the Fig. 3 with a dashed line. We then calculate the distances of the line $(1,3,5)$ from the center of the line $(2,4)$. In this way, we also obtained variance data caused by the person being occasionally directed to the sides during the movement. These occasional extremes could cause a significant amount of distortion when averaging the values, so we again used the median of the data.

3.6. Variance of head, knee and arm

Variance gives us the size of the range in which the observed values are found. It is denoted as σ^2 and can be calculated using the formula:

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)^2, \tag{2}$$

where N is the number of samples and μ is the mean value. In Fig. 4 we can see the top view of the walking person. Particular points represent the movement of the head along the Z axis. The solid line, denoted as μ, represents the mean value. Variable σ^2 is represented by a dotted line and represents the range at which the points are distributed.

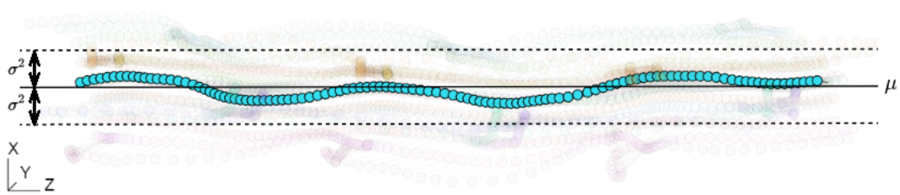


Fig. 4. Variance of head movement. Middle line is the mean value μ of head position, the dots show the head movement and σ^2 is the variance.

This approach is valid only if the monitored person always walks in the direction of the Z axis. For arbitrary direction, we use the linear regression instead of the mean value (see Fig. 5). Let us consider a line $\hat{y}$ that approximates particular points much better than average μ. Using this line we are able to calculate the size of the mean quadratic deviation MSE:

$$MSE = \frac{\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}{N-2}, \tag{3}$$

where N is the number of points, y_i is y coordinate of the i-th point and $\hat{y}_i$ is the mean value at point i calculated as line $\hat{y}_i = \beta_0 + \beta_1 x_i$. β_0 and β_1 are coefficients obtained by linear regression. The variance of knee and arm is estimated same way tracking particular joints instead of the head.

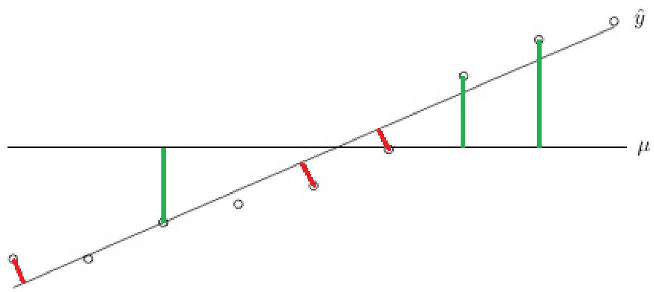


Fig. 5. Linear regression $\hat{y}$ and mean value μ.

4. ANALYSIS OF PHYSIOLOGICAL AND KINEMATIC PARAMETERS

Firstly, we have focused to find the similarities in the multiple records of the walking sequence recording same person. We have searched for the scalar values similar to multiple recordings of the same person. PCA can be utilized to find relevant axes in multidimensional space. We have multidimensional space of multiple parameters. PCA is used to find the direction of maximal variance in physiological and kinematic parameters multidimensional space. The most relevant parameters have its axes oriented in the direction of maximal variance.

5. TIME DEPENDENT PARAMETERS

In this section, we describe parameters that are time dependent. It will no longer be a one-dimensional value, as it was with physiological or kinematic parameters, but a sequence of values depending on time. We may find many parameters that vary with time. However, we decided to investigate joint acceleration, arm rotation, bending and leg distances. A complete overview of the parameters is described in Tab. III.

5.1. Acceleration

We have examined the acceleration of the various joints over time. The result was a signal that shows the magnitude of the acceleration as a function of time. The resulting acceleration values a are calculated using the formula:

$$a = \frac{v_2 - v_1}{t_2 - t_1}. \tag{4}$$

Table III. List of time dependent parameters. Parameter corresponding to relevant joints and their explanation.

Parameter	Description	Parameter	Description
a_{P_6}	head acceleration	ω_v	vertical rotation of arm
a_{P_1}	hip acceleration	ω_h	horizontal rotation of arm
a_{P_9}	left knee acceleration	$d_{P_{10,14}}$	palms distance
$a_{P_{13}}$	left elbow acceleration	$d_{P_{9,13}}$	elbows distance
$a_{P_{10}}$	left palm acceleration	$d_{P_{16,20}}$	knees distance
$a_{P_{14}}$	right palm acceleration	$d_{P_{17,21}}$	heels distance
$a_{P_{20}}$	left knee acceleration	$d_{P_{18,22}}$	tiptoes distance
$a_{P_{16}}$	right knee acceleration	φ_n	tilt of the head
$a_{P_{21}}$	left heel acceleration	φ_γ	tilt of the torso
$a_{P_{17}}$	right heel acceleration	φ_{α_L}	left arm bending
$a_{P_{22}}$	left tiptoe acceleration	φ_{α_P}	right arm bending
$a_{P_{18}}$	right tiptoe acceleration	φ_{β_L}	left leg bending
		φ_{β_P}	right leg bending

To calculate acceleration, we need to know the speeds of v_1 and v_2, that can be calculated using the formula:

$$v = \frac{\|P_2 - P_1\|}{t_2 - t_1}, \tag{5}$$

where P_1 and P_2 are points representing joint positions at t_1 and t_2, v_1 and v_2 are speeds in time t_1 and t_2, respectively.

This way we calculated the magnitudes of acceleration over time. We plotted the resulting values in a graph where axis x represents time and axis y acceleration, depicted in Fig. 6.

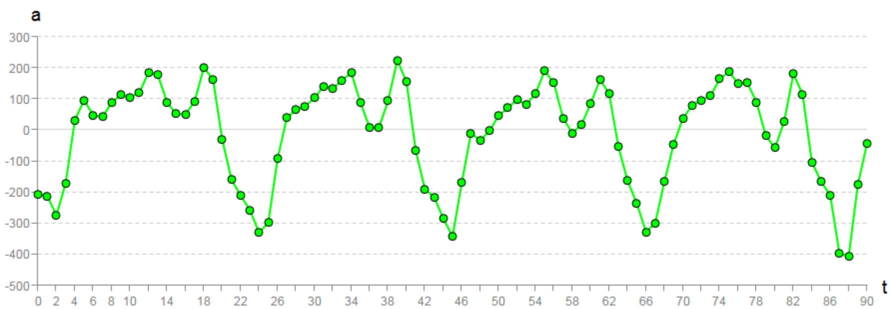


Fig. 6. Hip joint acceleration.

As we can see from the figure, the signal is not smooth but it is slightly shaken.

This is because footage consists of only 30 frames per second. If we had a record of walking with more robust sampling, the signal would be smoother. We used the triangle smoothing function [O'haver 2016], that uses the weighted average of the ambient values for signal smoothing. These values were calculated using the formula in Eq. 6.

$$X_i = \frac{Y_{i-2} + 2Y_{i-1} + 3Y_i + 2Y_{i+1} + Y_{i+2}}{9}, \tag{6}$$

where Y represents the input and X output signal. After using the smoothing function, we received a signal that can be seen in the Fig. 7.

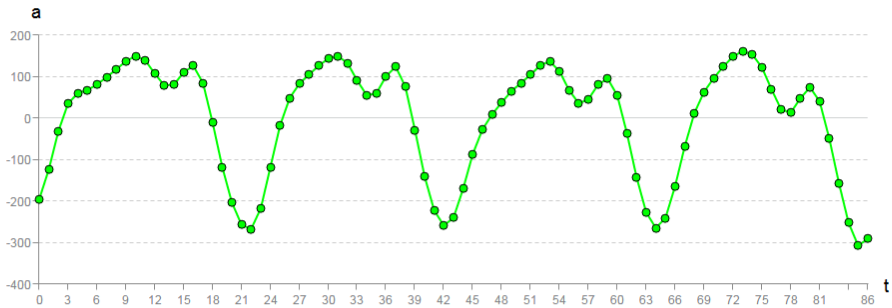


Fig. 7. Hip joint acceleration - smoothed signal.

5.2. Arms rotation

Another parameter we studied was the rotation of the arms. Walking person usually moves his shoulders. One shoulder moves upwards, while the other downwards. Concurrently, one moves forward and the other moves backwards. These are the two basic moves we perform with our shoulders as we walk. If we imagine a straight line passing through both arms, it would change its angle with respect to the floor and also to any vertical plane. That is why we call these movements as horizontal and vertical rotation of the arms, respectively. The illustration of the moves can be seen in the Fig. 8.

We use P_8 and P_{12}, see Fig. 1 and calculate the directional vector:

$$\vec{v} = \left(P_{12,x} - P_{8,x},\ P_{12,y} - P_{8,y},\ P_{12,z} - P_{8,z}\right), \tag{7}$$

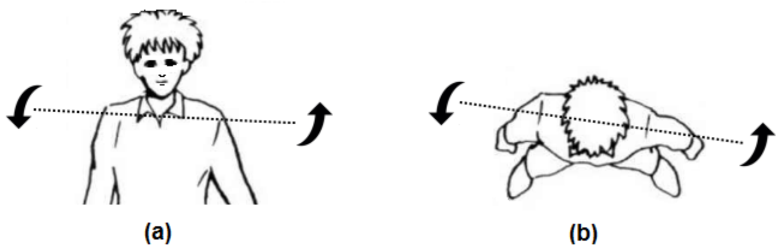


Fig. 8. Horizontal (a) a vertical (b) arms rotation.

where $P_{i,x}$, $P_{i,y}$ and $P_{i,z}$ represent the coordinates x, y, z of P_i, $i \in \{12, 8\}$. Subsequently, we calculated cosine of angle between vector $\vec{v}$ and the horizontal plane (Eq. 8) and the cosine angle between vector $\vec{v}$ and the vertical plane (Eq. 9).

$$\omega_h = |\vec{n_h}| \cdot |\vec{v}|, \quad (8)$$

$$\omega_v = |\vec{n_v}| \cdot |\vec{v}|, \quad (9)$$

where vector $\vec{n_h}$ was a normal vector of floor (0, 1, 0). For vertical rotation, vector $\vec{n_v}$ represented the normal vector of the plane YX, i.e. (0, 0, 1). Change of the cosine angles between individual frames plot to a graph, which can be seen in the Fig.9. The signal is also smoothed by the triangle smoothing function.

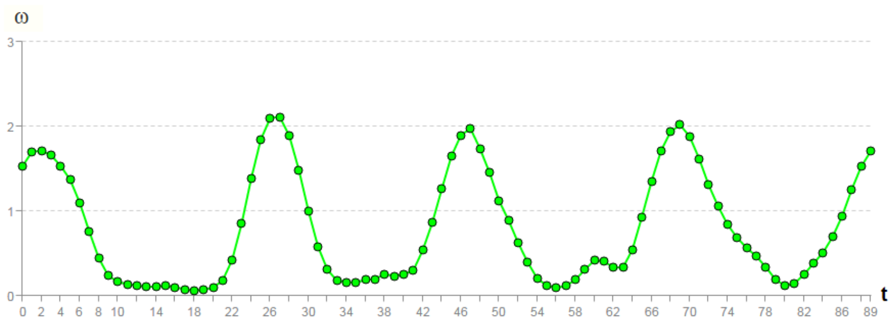


Fig. 9. Magnitude of vertical rotation of arms.

5.3. Distances

If we observe walking person, we may notice that the distance from the left palm to the right one is changing. We have decided to examine whether this change in the distance could be a characteristic feature for someone. During the investigation, we also focused on changing the elbows, knees and foot's distance. In addition, we studied the heel and tiptoe of the foot. So, overall, we have got up to 5 parameters.

Individual joints are defined as points in space and the Euclidean distance is calculated for the given joints in each frame. We plot the resulting values in the graph with the corresponding frame time, see Fig. 10. Where the axis y represents the distance value and the axis x is time expressed by the number of frames. There was almost no noise in this signal and it was not necessary to filter it.

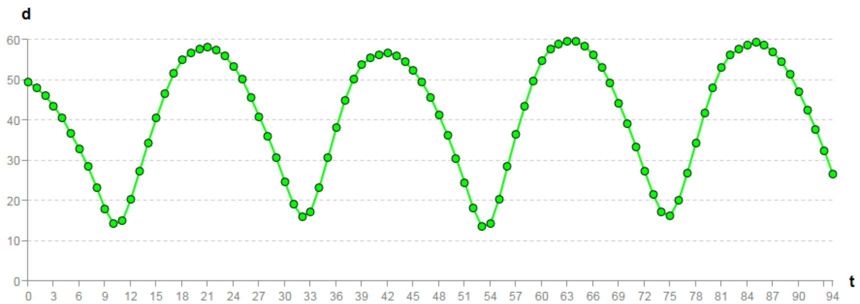


Fig. 10. Heels distance

5.4. Tilt and bending

Next we analyze the attributes such as a bending of an arm, a leg, a tilting of the neck and the whole body.

The hand and the leg have three joints in our case. Let us denote these joins as A, B and C, see Fig. 11. The resulting bend was calculated as the ratio of the length of the limb to the distance between the points A and C. For example, the right leg bending was calculated using Eq. 10.

$$\varphi_{\beta_P} = \frac{\beta_P}{\|P_{15} - P_{17}\|}, \tag{10}$$

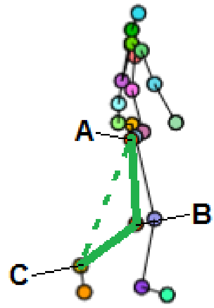


Fig. 11. Leg bending.

where β_P is the length of the right leg. We applied this calculation to each slide. The result is a graph that can be seen in Fig. 12, where axis y represents the ratio of the distances φ_{β_P} and axis x time by the number of frames.

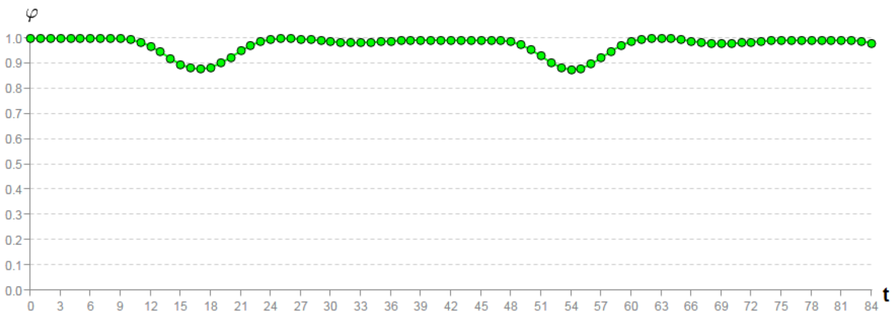


Fig. 12. Right leg bending.

The tilt of the whole body was calculated as the distance of the vertical line passing through point P_1 from point P_6, see Fig. 1. This way we found the movement of the head in relation to the center of body. The resulting graph, however, rather resembled a mixture of random values than the periodic signal. For this reason, we consider the signal irrelevant.

6. ANALYSIS OF TIME DEPENDENT PARAMETER

When analyzing time-dependent parameters, we have 10 records of the same person. From each record, we calculate all 25 parameters listed on the Tab. III. The purpose of the analysis was to find out which of these parameters could be used to uniquely identify a person.

We compared each parameter using the cross correlation of two signals [Telford et al. 1990]. We were able to calculate the match between the two signals. The match can be expressed by a numerical value in the range $\langle$-1,1$\rangle$, where 1 represents a 100% match.

Consequently, we were able to create a 10×10 matrix for each of the parameters, where we compared every record with each other based on a specific parameter. Each of these matrices was symmetric because the match between record 1 and record 2 must be the same as the match between record 2 and record 1. The diagonal was always 1 because it was a comparison of two identical records. Overall, we have received 25 matrices and our goal was to find those where the diagonal elements were as close as possible to 1. For this purpose, we used multidimensional scaling [Borg and Groenen 2013] method (MDS) to help us select these matrices.

From our input matrices, however, we needed to choose those with the lowest scatter of data. This meant that the size of the resulting dimension should be small. Using the MDS we can compute a matrix containing the coordinates of the points in the new space. The smaller the size, the parameter was more relevant. Based on the calculated eigenvalues, we can sort input parameters. The first eigenvalue was always the largest and it was the dominant axis of the new space. The Tab. IV contains the parameters sorted according to the eigenvalues from the smallest to the largest.

Based on this analysis, we were able to select parameters that are the most similar to records of one person. We chose to select those whose value is $e_1 < 0.1$. Consequently, we needed to find out which of these parameters are different for different people. The procedure was the same as in the previous case, with the difference that we used 8 records from 8 different people. In this case, however, we were looking for parameters whose value e_1 would be as large as possible to determine which parameters differentiate people. The results of the above analysis are written in the Tab. V.

Table IV. Time dependent parameters. Parameter notation, the parameter explanation and the relevant eigenvalue.

	Parameter	Description	e_1
1	$a_{P_{20}}$	left knee acceleration	0.0586
2	φ_{β_P}	bending of right leg	0.0600
3	$d_{P_{18,22}}$	distance of tiptoes	0.0617
4	$a_{P_{22}}$	acceleration of left tiptoe	0.0651
5	$d_{P_{17,21}}$	heels distance	0.0658
6	$a_{P_{17}}$	right heel acceleration	0.0706
7	$a_{P_{18}}$	right tiptoe acceleration	0.0733
8	φ_{β_L}	left leg bending	0.0805
9	$d_{P_{16,20}}$	knees distance	0.0831
10	$a_{P_{21}}$	left heel acceleration	0.0836
11	a_{P_1}	hips acceleration	0.1077
12	a_{P_6}	head acceleration	0.1114
13	$a_{P_{13}}$	right elbow acceleration	0.1147
14	$a_{P_{16}}$	right knee acceleration	0.1156
15	a_{P_9}	left elbow acceleration	0.1385
16	$a_{P_{14}}$	right palm acceleration	0.1399
17	$a_{P_{10}}$	left palm acceleration	0.1426
18	$d_{P_{9,13}}$	elbows distance	0.1809
19	φ_γ	tilt of torso	0.2010
20	$d_{P_{10,14}}$	palms distance	0.2036
21	φ_{α_P}	bending of right arm	0.3648
22	ω_h	horizontal arm rotation	0.3923
23	φ_{α_L}	bending of left arm	0.3935
24	ω_v	arm rotation vertical	0.4794
25	φ_n	tilt of the head	0.5272

Table V. Sorted time-dependent parameters that are different for different people. Parameter notation, the parameter explanation and the relevant eigenvalue.

	Parameter	Description	e_1
1	$d_{P_{17,21}}$	heels distance	0.4325
2	$d_{P_{16,20}}$	knees distance	0.4000
3	$a_{P_{17}}$	right heel acceleration	0.2251
4	$a_{P_{21}}$	left heel acceleration	0.2124
5	φ_{β_L}	bending of left leg	0.2119
6	φ_{β_P}	bending of right leg	0.1597
7	$a_{P_{20}}$	left knee acceleration	0.1273
8	$d_{P_{18,22}}$	tiptoes distance	-
9	$a_{P_{22}}$	left tiptoe acceleration	-
10	$a_{P_{18}}$	right tiptoe acceleration	-

7. CONCLUSIONS

Novelty of our work is in the determination of parameter that most relevantly describes the individual in the context of gait. We have 32 records capturing gait of 8 different people. From the original 32 records, 11 of them were stored in the database. Each person had at least one record. Our goal was to assign the remaining 21 records to the right person.

During testing, we relied on data obtained from the analysis. We tested whether we can uniquely assign the gait to the right person. We tested various combinations of parameters. The results of the comparison are listed the Tab. VI. In addition to the selection of parameters, we also changed their weights in the test. We tried to give more weight to the more relevant parameter than the less relevant. Weight was a real number that multiplied the strength of each parameter. Parameters that were not listed in the table had a weight set to 0. It has turned out, that even for the parameters that have been identified as least relevant (e.g. arm rotation), we have managed to correctly assign up to 62% of walking records. This meant that each parameter, regardless of its relevance, uniquely describes the gait of person. In addition, further research has shown that different group of parameters is able to describe each group of people differently. For example, for individuals 1-4, some parameters have greater significance than for people 5-8. For another set of parameters, this was the opposite. This means that each person performs one of the movements of the limbs in a unique way that is specific to that person only. For each person, however, this limb or joint is different. This makes it very difficult to find a universal parameter that would identify the person with a 100% success rate. Success rate indicates how two particular feature vectors are similar to each other. For example, 100% success rate says, that the feature vector obtained during the gait analysis of investigated person is exactly same as a particular vector in the database of recorded subjects and therefore it is expected, that the person is identical. However, the key to clear identification may lie in the possible combinations of parameters, which would be perfectly balanced.

Table VI. Result validation. List of nonzero parameter weights used in identification process and the corresponding success rate.

Parameter weights	Success rate
$V(\gamma) = 1, V(\alpha_P) = 1, V(\alpha_L) = 1, V(\beta_P) = 1, V(\beta_L) = 1$	100%
$V(\gamma) = 3, V(\alpha_P) = 3, V(\alpha_L) = 3, V(\beta_P) = 3, V(\beta_L) = 3, V(d_{P_{16,20}}) = 1, V(d_{P_{17,21}}) = 1$	95%
$V(\gamma) = 2, V(\alpha_P) = 2, V(\alpha_L) = 2, V(\beta_P) = 2, V(\beta_L) = 2, V(d_{P_{16,20}}) = 1, V(d_{P_{17,21}}) = 1, V(\sigma^2_{P_6,y}) = 1, V(\sigma^2_{P_{20},y}) = 1, V(\sigma^2_{P_{16},y}) = 1$	66%
$V(\gamma) = 5, V(\alpha_P) = 5, V(\alpha_L) = 5, V(\beta_P) = 5, V(\beta_L) = 5, V(d_{P_{16,20}}) = 1, V(d_{P_{17,21}}) = 1, V(\sigma^2_{P_6,y}) = 1, V(\sigma^2_{P_{20},y}) = 1, V(\sigma^2_{P_{16},y}) = 1$	80%
$V(d_{P_{16,20}}) = 1, V(d_{P_{17,21}}) = 1$	33%
$V(\sigma^2_{P_6,y}) = 1, V(\sigma^2_{P_{20},y}) = 1, V(\sigma^2_{P_{16},y}) = 1$	57%
$V(d_{P_{16,20}}) = 1, V(d_{P_{17,21}}) = 1, V(\sigma^2_{P_6,y}) = 1, V(\sigma^2_{P_{20},y}) = 1, V(\sigma^2_{P_{16},y}) = 1$	57%
$V(\omega_h) = 1, V(\omega_v) = 1$	62%
$V(a_{P_1}) = 1, V(a_{P_{10}}) = 1, V(a_{P_{14}}) = 1, V(a_{P_{17}}) = 1, V(a_{P_{21}}) = 1,$	57%
$V(a_{P_1}) = 1$	47%
$V(a_{P_1}) = 1, V(a_{P_6}) = 1, V(a_{P_9}) = 1, V(a_{P_{13}}) = 1, V(a_{P_{21}}) = 1, V(a_{P_{17}}) = 1$	71%
$V(\varphi_{\beta_L}) = 1, V(\varphi_{\beta_P}) = 1, V(d_{P_{10,14}}) = 1, V(d_{P_{9,13}}) = 1, V(d_{P_{16,20}}) = 1, V(d_{P_{17,21}}) = 1, V(a_{P_1}) = 1, V(a_{P_9}) = 1, V(a_{P_{13}}) = 1, V(a_{P_1}0) = 1, V(a_{P_{14}}) = 1, V(a_{P_{20}}) = 1, V(a_{P_{16}}) = 1, V(a_{P_{21}}) = 1, V(a_{P_{17}}) = 1$	62%

In this work, we have described the parameters of the gait, the ways we calculated them and the methods we examined their relevance. Finally, we used these results to validate their accuracy. We came to the conclusion that the unique parameters are ones that are constant in time. They are body proportions such as height, length of arms and legs. All other parameters were behavioral, and it could be somewhat imitated. We found out that the gait hides a unique biometric signature that could help to identify a person.

REFERENCES

ACAR, E., SENST, T., KUHN, A., KELLER, I., THEISEL, H., ALBAYRAK, S., AND SIKORA, T. 2012. Human action recognition using lagrangian descriptors. In *2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP)*. 360–365.

ANDRILUKA, M., ROTH, S., AND SCHIELE, B. 2009. Pictorial structures revisited: People detection and articulated pose estimation. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 1014–1021.

BARCLAY, C. D., CUTTING, J. E., AND KOZLOWSKI, L. T. 1978. Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics 23,* 2, 145–152.

BAZIN, A. I. AND NIXON, M. S. 2005. Gait verification using probabilistic methods. In *Application of Computer Vision, 2005. WACV/MOTIONS '05 Volume 1. Seventh IEEE Workshops on*. Vol. 1. 60–65.

BORG, I. AND GROENEN, P. 2013. *Modern Multidimensional Scaling: Theory and Applications*. Springer Series in Statistics. Springer New York.

CHEN, D., WACTLAR, H., Y. CHEN, M., GAO, C., BHARUCHA, A., AND HAUPTMANN, A. 2008. Recognition of aggressive human behavior using binary local motion descriptors. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 5238–5241.

DUBOIS, A. AND BRESCIANI, J. P. 2015. Person identification from gait analysis with a depth camera at home. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 4999–5002.

ĎURIKOVIČ, R. AND MADARAS, M. 2015. Controllable skeleton-sheets representation via shape diameter function. In *Mathematical Progress in Expressive Image Synthesis II*, H. Ochiai and K. Anjyo, Eds. Springer Japan, Tokyo, 79–90.

FENG, Y., LI, Y., AND LUO, J. 2016. Learning effective gait features using lstm. In *2016 23rd International Conference on Pattern Recognition (ICPR)*. 325–330.

GIANARIA, E., GRANGETTO, M., LUCENTEFORTE, M., AND BALOSSINO, N. 2014. *Human Classification Using Gait Features*. Springer International Publishing, Cham, 16–27.

KIM, J. AND KIM, M. 2014. Motion capture with high-speed rgb-d cameras. In *2014 International Conference on Information and Communication Technology Convergence (ICTC)*. 394–395.

O'HAVER, T. 2016. *A Pragmatic Introduction to Signal Processing: With Applications in Scientific Measurement*. CreateSpace Independent Publishing Platform.

PETERKOVA, A. AND STREMY, M. 2015. Obtaining the gait parameters from kinect sensor for the person identification. In *2015 IEEE 19th International Conference on Intelligent Engineering Systems (INES)*. 337–340.

RAHMAN, M. W. AND GAVRILOVA, M. L. 2017. Kinect gait skeletal joint feature-based person identification. In *2017 IEEE 16th International Conference on Cognitive Informatics Cognitive Computing (ICCI*CC)*. 423–430.

RIEČICKÝ, A., MADARAS, M., PIOVARCI, M., AND DURIKOVIC, R. 2018. Optical-inertial synchronization of mocap suit with single camera setup for reliable position tracking. In *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 1: GRAPP,*. INSTICC, SciTePress, 40–47.

TAUTZ, D. 1989. Hypervariability of simple sequences as a general source for polymorphic dna markers. *Nucleic Acids Research 17,* 16, 6463.

TELFORD, W., TELFORD, W., GELDART, L., AND SHERIFF, R. 1990. *Applied Geophysics*. Monograph series. Cambridge University Press.

WANG, Y., LAU, D. L., AND HASSEBROOK, L. G. 2010. Fit-sphere unwrapping and performance analysis of 3d fingerprints. *Appl. Opt. 49,* 4 (Feb), 592–600.

YAM, C., NIXON, M. S., AND CARTER, J. N. 2002. On the relationship of human walking and running: automatic person identification by gait. In *Object recognition supported by user interaction for service robots*. Vol. 1. 287–290 vol.1.

ZHU, L., HU, X., AND KAVAN, L. 2015. Adaptable anatomical models for realistic bone motion reconstruction. *Comput. Graph. Forum 34,* 2 (May), 459–471.

Andrej Mihálik
Mathematics, Physics and Informatics,
Comenius University,
842 48 Bratislava, Slovak Republic
email: andrej.mihalik@fmph.uniba.sk

Roman Ďurikovič
Faculty of Mathematics, Physics and Informatics,
Comenius University,
842 48 Bratislava, Slovak Republic
email: roman.durikovic@fmph.uniba.sk

Michal Sejč
Faculty of Mathematics, Physics and Informatics,
Comenius University,
842 48 Bratislava, Slovak Republic

Estimation Methods for a Flexible INAR(1) COM-Poisson Time Series Model

Y. SUNECHER, N. MAMODE KHAN AND V. JOWAHEER

Abstract

Time series of counts occur in many real-life situations where they exhibit various forms of dispersion. To facilitate the modeling of such time series, this paper introduces a flexible first-order integer-valued non-stationary autoregressive (INAR(1)) process where the innovation terms follow a Conway-Maxwell Poisson distribution (COM-Poisson). To estimate the unknown parameters in this model, different estimation approaches based on likelihood and quasi-likelihood formulations are considered. From simulation experiments and a real-life data application, the Generalized Quasi-Likelihood (GQL) approach yields estimates with lower bias than the other estimation approaches.

Mathematics Subject Classification 2010:
Keywords:

1. INTRODUCTION

Counts observations are often exposed to over- and under-dispersion relative to their means. In the existing literature, the modeling of such data is handled mainly by some Generalized forms of the Poisson distribution such as the Generalized Poisson Distribution (GPD) [Consul and Shoukri 1984], the family of the Weighted Poisson (WP) distributions [Castillo and Perez-Cassany 2005], the Negative Binomial (NB) model specifically for the over-dispersion case and lately by the COM-Poisson model, among others. For some of the Generalized forms of the Poisson Distribution (PD) such as GPD and WP, the drawback is that such models do not form part of the class of exponential densities and Generalized Linear Models (GLMs) whilst for the renovated COM-Poisson distribution, the properties of GLMs such as the availability of the link function predictor and the forms of its mean and variance functions are compliant with the theory of exponential densities. These features in turn facilitate the use of standard estimation methods such as Maximum Likelihood Estimation (MLE) and Fisher-Scoring (FS) algorithm. As at date, many researchers such as Shmueli et al. [2005], Guikema and Coffelt [2008], Nadarajah [2009], Lord et al. [2010], Sellers et al. [2012] and Dey et al. [2016] have extensively applied the COM-Poisson to analyze many real-life data sets in linguistics, in modelling traffic accidents and number of sales in supermarkets and lately due to an upsurge of

10.2478/jamsi-2018-0005

interest in this model, Sellers et al. [2016] have developed the bivariate COM-Poisson. Shmueli et al. [2005] have also compared the consistency of the fitted values obtained from applying COM-Poisson, NB and GPD to real-life data and concluded that COM-Poisson is equally able to provide reliable values under any situation of over-, equi- and under-dispersion. The COM-Poisson structure is made up of a normalizing constant based on an infinite summation, which has been approximated to an asymptotic closed form expression to facilitate the computation of the moments and derivation of the moment generating function. However, an important note here is that the COM-Poisson under this approximated formula is especially suitable for over- and equi-dispersion but may still be used for the under-dispersion case as the means outweigh the variances under this approximation. Based on these elegant advantages, this paper proposes to extend the univariate COM-Poisson to an integer-valued autoregressive (INAR(1)) COM-Poisson time series model.

Time series of counts are usually modeled using parameter-driven (PD) or observation-driven (OD) approaches [Cox 1981]. In the parameter-driven models, the serial-correlation is induced by correlated randomly-distributed latent effects while under observation-driven, the observation at the current time point is related to previous lagged observation through the binomial thinning operation [Steutel and Van Harn 1979] which makes the serial-correlation and an innovation term taken at the current time point. Various studies have shown that it is more appropriate to model the time series using the OD model since parameter estimation becomes less cumbersome (See Ravishanker et al. [2014], Jung et al. [2006], Jung and Tremayne [2003], Jung and Liesenfeld [2001], Durbin and Koopman [2000] and Chan and Ledolter [1995]). One such class of OD model constitutes of the classical INAR(1) process introduced by McKenzie [1986] and Alzaid and Al Osh [1990] in the modeling of discrete time series processes where the innovation terms are Poisson or Negative Binomial marginals. In this paper, we explore the INAR(1) process where the innovation terms are COM-Poisson distributed. Since under such set-up, the counts are also mostly influenced by time-varying covariates as illustrated in many real-life studies, the proposed INAR(1) process in this paper is developed based on the error terms following COM-Poisson under non-stationary moments specified by time-dependent covariates. Thus, it is of worth interest to explore the different estimation methods to estimate the regression effects and the serial-correlations. Firstly, the Conditional Maximum Likelihood Estimation (CMLE) method is appropriate in this context since conditioned on the previous observation, the likelihood function may be derived using convolution and thereafter, estimation of parameters is carried out by maximizing the conditional likelihood and solve for the unknown parameters using iterative approach such as Newton-Raphson method. The Conditional Least Squares (CLS) is also considered in this paper. In many time series studies conducted by Quoreshi [2008], CLS was used for estimation purpose and yielded consistent parameter estimates. Another popular technique is the Generalized Method of Moments (GMM) which combines a set of moment scores into an objective function and this is minimized and solved iteratively using the Newton-Raphson technique. However, in this paper, we use a modified GMM that is based on only two moment adaptive estimating equations developed by Qu and Lindsay [2003]. Note that the modified GMM has been tested and yields equally consistent and efficient estimates as the traditional GMM with more than two score functions [Qu and Lindsay 2003]. In the last ten years, the Generalized Quasi-Likelihood (GQL) method has also gained a lot of popularity due to its flexibility in requiring only the population means and variances of the model rather than the full likelihood function. The GQL technique is originated from the Generalized Estimating Equations (GEE) [Liang and Zeger 1986] and quasi-likelihood function [Wedderburn 1974]. In fact, in GEE it was difficult to specify the true correlation structure and Liang and Zeger [1986] suggested to approximate this structure by some working correlation matrices. However, Crowder [1995] and Sutradhar and Das [1999] noted that GEE estimates under working structure are comparatively inefficient as compared to the GEE based on independence structure approach and this is a clear contradiction since the observations are serially correlated. Then Sutradhar and Das [1999] developed a robust autocorrelation structure where the correlation coefficients are obtained through the method of moments and this robust correlation structure was incorporated into the GEE to yield the GQL estimating function. The GQL approach has been explored in many longitudinal studies, where consistent and efficient parameter estimates were generated and even in the late findings by Sutradhar et al. [2014], the GQL method was shown to provide asymptotically efficient estimates as the maximum likelihood based method.

This paper underpins all these approaches and compares the efficiency of the estimates under the different methods on the basis of the standard errors. The organization of the paper is as follows: In the next section, the INAR(1)

COM-Poisson time series model is developed and its autocorrelation structure is derived taking into account that the series of observations is influenced by some time-varying covariates. Section 3 of the paper focuses on the estimation methods. In section 4, a simulation study is presented and the different approaches are compared. In section 5, the INAR(1) COM-Poisson model is fitted to a real-life dataset in Mauritius and the paper is concluded in section 6.

2. THE INAR(1) COM-POISSON TIME SERIES PROCESS

The COM-Poisson probability function is expressed as

$$P(Y=y_j)=\frac{(\lambda_j)^{y_j}}{(y_j!)^{\nu}}\frac{1}{Z(\lambda_j,\nu)}, \quad y_j=0,1,2,3,\ldots, \quad for \quad \lambda_j\geq 0, \nu\geq 0 \tag{1}$$

where $Z(\lambda_j,\nu)=\sum_{k=0}^{\infty}\frac{(\lambda_j)^k}{(k!)^{\nu}}$.
The moments of the COM-Poisson are calculated iteratively using

$$E(Y^{r+1})=\begin{cases}\lambda_j E(Y+1)^{1-\nu} & r=0,\\ \\ \lambda_j\frac{d}{d\lambda_j}E(Y^r)+E(Y)E(Y^r) & r>0.\end{cases} \tag{2}$$

From Shmueli et al. [2005], Minka et al. [2003] and Sellers et al. [2012], the first two moments can be approximated by

$$E(Y_j)=(\lambda_j)^{1/\nu}-(\frac{\nu-1}{2\nu}) \quad and \quad V(Y_j)=\frac{(\lambda_j)^{1/\nu}}{\nu} \tag{3}$$

where the normalizing constant $Z(\lambda_j,\nu)$ is approximated by

$$Z(\lambda_j,\nu)\simeq\frac{\exp(\nu\lambda_j^{\frac{1}{\nu}})}{\lambda_j^{\frac{\nu-1}{2\nu}}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}} \tag{4}$$

From numerical studies, Shmueli et al. [2005] showed that these approximations work suitably well for all $\nu>0$ but in particular for $\nu<1$ or $\lambda_j>10^{\nu}$. Under these expectations and variances, it can easily be verified that for $\nu<1$, $V(Y_j)>E(Y_j)$ (over-dispersion) and for $\nu>1, E(Y_j)>V(Y_j)$ (under-dispersion) since $\lambda_j>0$. Under a regression set-up, these assumptions about over- and under-dispersion hold since $\lambda_j=exp(x_j^T\beta)$ for $x_j=[x_{j1},x_{j2},\ldots,x_{jp}]^T$ and $\underset{\sim}{\beta}=[\beta_1,\beta_2,\ldots,\beta_p]^T$. Using McKenzie [1986] and Al-Osh and Alzaid [1988] approach, the observation-driven equation for the first-order AR(1) process is given by

$$y_t = \rho * y_{t-1} + d_t \tag{5}$$

where d_t corresponds to the innovation term, $\rho * y_{t-1}|y_{t-1} \sim \text{Binomial}(y_{t-1},\rho)$. d_t and y_{t-k} are independent, for $k = 1,2,\ldots,T$.

Assuming that Y_t is a COM-Poisson observation taken at the t^{th} time point, then the mean and variance at the t^{th} time point is given by

$$E(Y_t) = \theta_t = \lambda_t^{1/\nu} - \frac{\nu-1}{2\nu} \quad and \quad Var(Y_t) = \sigma_t^2 = \frac{\lambda_t^{1/\nu}}{\nu}, \quad for \quad t = 1,2,3,\ldots \tag{6}$$

Under a re-parametrization, $Var(Y_t) = \frac{\theta_t}{\nu} + \frac{\nu-1}{2\nu^2}$. Note that, under non-stationary assumption which is due to the presence of time-dependent covariates;

$$\lambda_1 \neq \lambda_2 \neq \ldots \neq \lambda_t \neq \ldots \neq \lambda_T \Rightarrow \theta_1 \neq \theta_2 \neq \ldots \neq \theta_t \neq \ldots \neq \theta_T \Rightarrow \sigma_1^2 \neq \sigma_2^2 \neq \ldots \neq \sigma_t^2 \neq \ldots \neq \sigma_T^2 \tag{7}$$

From equation (5) and equation (6), the mean and variance of the innovation term d_t are determined using the same procedure as defined by Mamode khan and Jowaheer [2013] for the stationary case. Under the moment expression obtained for d_t, d_t is shown to be COM-Poisson with

$$E(d_t) = \theta_t - \rho\theta_{t-1} \tag{8}$$

and the dispersion

$$\tilde{\nu} = \frac{(2\theta_t - 2\rho\theta_{t-1} + 1) + \sqrt{(2\theta_t - 2\rho\theta_{t-1} + 1)^2 - 8[(1-\rho^2)(\frac{\nu-1}{2\nu^2}) + \frac{\theta_t}{\nu} - \frac{\rho\theta_t}{\nu}(\rho + \nu(1-\rho))]}}{4[(1-\rho^2)(\frac{\nu-1}{2\nu^2}) + \frac{\theta_t}{\nu} - \frac{\rho\theta_{t-1}}{\nu}(\rho + \nu(1-\rho))]} \tag{9}$$

and under these distributional assumptions for d_t, it can easily be re-verified that Y_t has same mean and variance as in equation (6). The serial-correlation is shown to be

$$\begin{aligned} Corr(Y_t, Y_{t-k}) &= \frac{Cov(Y_t, Y_{t-k})}{\sqrt{\frac{\theta_t}{\nu} + \frac{\nu-1}{2\nu^2}}\sqrt{\frac{\theta_{t-k}}{\nu} + \frac{\nu-1}{2\nu^2}}} \\ &= \rho^k \frac{\sqrt{(\frac{\theta_{t-k}}{\nu} + \frac{\nu-1}{2\nu^2})}}{\frac{\theta_t}{\nu} + \frac{\nu-1}{2\nu^2}} \end{aligned} \tag{10}$$

for $k = 1,\ldots,T-1$ and $0 < \rho < \min(\frac{\theta_2}{\theta_1},\ldots,\frac{\theta_t}{\theta_{t-1}},1)$.

3. ESTIMATION METHODS

3.1. Conditional Maximum Likelihood Estimation (CMLE)

In this section, the different estimation approaches are presented. The CML approach consists of using a conditional likelihood function of the form where

$$L(\rho,\beta,\nu|y) = \prod_{t=1}^{T} Pr(y_t|y_{t-1},\rho,\beta,\nu)$$
$$= \sum_{t=1}^{T}\sum_{i=0}^{m}\left(\frac{\lambda_t^{y_t-i}}{(y_t-i)!^{\nu}}\right)\left(\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})}{\lambda_t^{\frac{\nu-1}{2\nu}}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}}\right)\binom{y_{t-1}}{i}\rho^i(1-\rho)^{y_{t-1}-i} \quad (11)$$

where m=min(y_t, y_{t-1}) and by maximizing this function with respect to the regression vector and the serial coefficient using the R optim function or the MatLab software, the likelihood scores are obtained (See Appendix). The objective function in the CLS approach is expressed as

$$Q(\psi) = \sum_{t=2}^{T}[y_t - E(y_t|y_{t-1})]^2 = \sum_{t=2}^{T}[y_t - \rho y_{t-1} - (\lambda_t^{1/\nu} - \frac{\nu-1}{2\nu}) + \rho(\lambda_{t-1}^{1/\nu} - \frac{\nu-1}{2\nu})]^2$$

and again here, the derivatives are obtained by minimizing the above (See Appendix) and use Newton-Raphson iterative procedure to solve.

3.2. The GQL Approach

The GQL estimating function depends only on the specification of the mean and variance function and does not require the expression of the likelihood function. However, one important component in the GQL formulation is the modeling of the auto-covariance structure. In previous estimation approaches such as GEE, the auto-covariance was constructed using the working correlation structures but in the case of GQL, the auto-covariance is built under the assumption of a robust autocorrelation function [Sutradhar 2003; Sutradhar and Das 1999].
The GQL for the INAR(1) COM-Poisson time series model is given by

$$D^T\Sigma^{-1}(\underset{\sim}{y} - \underset{\sim}{\theta}) = 0 \quad (12)$$

where $\underset{\sim}{y} = [y_1, y_2, \ldots, y_T]^T$ and $\underset{\sim}{\theta} = [\theta_1, \theta_2, \ldots, \theta_T]^T$. Assuming a set of p explanatory variables $\underset{\sim}{\beta} = [\beta_1, \beta_2, \ldots, \beta_p]^T$, the derivative matrix is a block diagonal $T \times (p+1)$

matrix and

$$D^T = \begin{pmatrix} \frac{\partial\theta_1}{\partial\beta_1} & \frac{\partial\theta_2}{\partial\beta_1} & \cdots & \frac{\partial\theta_T}{\partial\beta_1} \\ \frac{\partial\theta_1}{\partial\beta_2} & \frac{\partial\theta_2}{\partial\beta_2} & \cdots & \frac{\partial\theta_T}{\partial\beta_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\theta_1}{\partial\beta_p} & \frac{\partial\theta_2}{\partial\beta_p} & \vdots & \frac{\partial\theta_T}{\partial\beta_p} \\ \frac{\partial\theta_1}{\partial\nu} & \frac{\partial\theta_2}{\partial\nu} & \cdots & \frac{\partial\theta_T}{\partial\nu} \end{pmatrix}_{(p+1)\times T} \tag{13}$$

where $\frac{\partial\theta_t}{\partial\beta_k} = \frac{(\lambda_t)^{\frac{1}{\nu}} x_j^T}{\nu}$ and $\frac{\partial\theta_t}{\partial\nu} = -\frac{(\lambda_t)^{\frac{1}{\nu}} ln(\lambda_t)}{\nu^2} - \frac{1}{2\nu} + \frac{\nu-1}{2\nu^2}$.
The covariance structure Σ is computed using

$$\frac{\sum_{t=1}^{T-1} \tilde{Y}_t \tilde{Y}_{t+1}/(T-1)}{\sum_{t=1}^{T} \tilde{Y}_t^2/T} = \rho \frac{\sum_{t=2}^{T} \sqrt{\frac{\frac{\lambda_{t-1}^{\frac{1}{\nu}}}{\nu}}{\frac{\lambda_t^{\frac{1}{\nu}}}{\nu}}}}{(T-1)} \tag{14}$$

Thus

$$\rho = \frac{T\sum_{t=1}^{T-1} \tilde{Y}_t \tilde{Y}_{t+1}}{[\sum_{t=1}^{T} \tilde{Y}_t^2][\sum_{t=2}^{T} \sqrt{\frac{\frac{\lambda_{t-1}^{\frac{1}{\nu}}}{\nu}}{\frac{\lambda_t^{\frac{1}{\nu}}}{\nu}}}]} \tag{15}$$

where $\tilde{y}_t = \frac{y_t - \theta_t}{\sqrt{\theta_t}}$ and this value of ρ is substituted in equation (10) to calculate the serial-correlation. The Newton-Raphson iterative technique is used to solve the GQL in equation (12) which yields

$$\begin{pmatrix} \hat{\beta}_{r+1} \\ \hat{\nu}_{r+1} \end{pmatrix} = \begin{pmatrix} \hat{\beta}_r \\ \hat{\nu}_r \end{pmatrix} + [D^T\Sigma^{-1}D]_r^{-1}[D^T\Sigma^{-1}(\underset{\sim}{y} - \underset{\sim}{\theta})]_r \tag{16}$$

where $\hat{\beta}_r$ is the value of $\hat{\beta}$ at the r^{th} iteration. $[.]_r$ is the value of the expression at the r^{th} iteration. The algorithm works by assuming an initial value of $\hat{\beta}$ and $\hat{\nu}$, we calculate $\hat{\rho}$ using equation (15) and then use this parameter to update the values of $\hat{\beta}$ and $\hat{\nu}$. Then the new set of parameters is used to calculate $\hat{\rho}$ and the iteration continues in this way until convergence. The estimators are consistent and under mild regularity conditions, it may be shown that $((\hat{\beta},\hat{\nu}) - (\beta,\nu))^T$ has an asymptotic normal distribution with mean 0 and covariance matrix

$[D^T\widetilde{\Sigma}^{-1}D]^{-1}[D^T\widetilde{\Sigma}^{-1}(\underset{\sim}{y}-\underset{\sim}{\theta})(\underset{\sim}{y}-\underset{\sim}{\theta})^T\widetilde{\Sigma}^{-1}D][D^T\widetilde{\Sigma}^{-1}D]^{-1}$ [Sutradhar 2003; Sutradhar et al. 2014; Mamode khan et al. 2016].

3.3. Generalized Method of Moments (GMM)

For the GMM methodology, we follow the adaptive GMM equation developed by Qu and Lindsay [2003] where only 2 moment conditions are taken into account,

$$g = \begin{bmatrix} D^T(y-\theta) \\ \hat{\alpha}^T D^T V(y-\theta) \end{bmatrix} \tag{17}$$

where $V = \frac{1}{T}(y-\theta)(y-\theta)^T$ and to solve the over-determined system of equations, the moment vector (17) is combined to form an objective function

$$Q(\beta) = g^T C^{-1} g \tag{18}$$

where C is the sample variance of g:

$$\begin{pmatrix} D^T VD & (D^T V^2 D)\hat{\alpha} \\ \hat{\alpha}^T(D^T V^2 D) & \hat{\alpha}^T(D^T V^3 D)\hat{\alpha} \end{pmatrix} \tag{19}$$

where $\hat{\alpha}$ is the product of the inverse of the upper triangular matrix,G, obtained from the Cholesky decomposition of the matrix $(D^T V - LD^T)V(D^T V - LD^T)$ where $L = (D^T V^2 D)(D^T VD)^{-1}$ and the eigenvector corresponding to the largest eigenvalue of the matrix $(G^T)^{-1}\widetilde{D}\widetilde{D}^T G^{-1}$ where $\widetilde{D} = D^T VD - LD^T D$.
By minimizing the function (18) and applying the Newton-Raphson algorithm, we obtain the iterative

$$\hat{\beta}_{r+1} = \hat{\beta}_r - [\ddot{Q}(\hat{\beta}_r)]^{-1}\dot{Q}(\hat{\beta}_r) \tag{20}$$

where asymptotically, $\dot{Q}(\beta) = 2\dot{g}^T C^{-1} g$ and the hessian matrix $\ddot{Q}(\beta) = 2\dot{g}^T C^{-1}\dot{g}$ with $\dot{g} = \frac{\partial g}{\partial \beta^T}$ where $\dot{g}$ is of dimension $(p+1)\times p$. The iterative equation (20) works by assuming an initial value of $\hat{\beta}$ which is used to calculate $\dot{Q}(\beta)$ and $\ddot{Q}(\beta)$. After obtaining an updated value of $\hat{\beta}$, these 2 quantities are re-computed until the previous and current elements of $\hat{\beta}$ converges such that $||\hat{\beta_{r+1}} - \hat{\beta}|| < 10^{-5}$. Qu and Lindsay [2003] showed that the estimator of β obtained by minimizing the objective function (18) is consistent and asymptotically normal with $\hat{\beta} - \beta \sim \mathrm{N}\{0, (E[\dot{g}^T]E[C^{-1}]E[\dot{g}])^{-1}\}$.
Using the converged estimates from these methods, the forecasting equations are given by:

Using the model (5), for $t = T$ and $j = 1,2,3,\ldots$

$$y_{T+j} = \hat{\rho}^j * y_T + \sum_{i=0}^{j-1} \hat{\rho}^i * d_{T+j-i} \tag{21}$$

where

$$d_{T+j} \sim \text{CMP}\left(\left(\frac{(\theta_{T+j} - \rho\theta_{T+j-1})}{\tilde{\nu}}, \tilde{\nu}\right)\right) \tag{22}$$

The conditional expectation and variance of the future observation $Y_{T+j}^{[k]}$ given $y_T^{[k]}$ are as follows:

$$E(Y_{T+j}|y_T) = \theta_{T+j} + \hat{\rho}^j(y_T - \theta_T) \tag{23}$$

$$Var(Y_{T+j}|y_T) = \left(\frac{\theta_{T+j}}{\nu} + \frac{\nu-1}{2\nu^2}\right) + \hat{\rho}^j(1-\hat{\rho}^j)(y_T - \theta_T) - (\hat{\rho}^2)^j\left(\frac{\theta_T}{\nu} + \frac{\nu-1}{2\nu^2}\right) \tag{24}$$

Using equations (21), (23) and (24), the variance of the j-step ahead forecasting error, $e_{T+j} = y_{T+j} - E(Y_{T+j}|y_T)$, is given by

$$\begin{aligned} Var(e_{T+j}) &= Var[E(e_{T+j}|y_T)] + E[Var(e_{T+j}|y_T)] \\ &= \left(\frac{\theta_{T+j}}{\nu} + \frac{\nu-1}{2\nu^2}\right) - (\hat{\rho}^2)^j\left(\frac{\theta_{T+j-1}}{\nu} + \frac{\nu-1}{2\nu^2}\right) \end{aligned} \tag{25}$$

4. SIMULATIONS AND RESULTS

Time series data of size $T = 60, 100$, and 500 are generated using equation (5) under $\nu = 0.3, 0.5, 0.9, 1, 1.5, 2$ and 3. A 2×1 regression vector is assumed where for the first covariate

$$x_{t1} = \begin{cases} cos(2\pi t/12) & (t = 1,\ldots,T/4), \\ rnorm(1,0,1) & (t = (T/4)+1,\ldots,3T/4), \\ sin(2\pi t/12) & (T = (3T/4)+1,\ldots,T), \end{cases}$$

and x_{t2}, the second covariate, follows standard normal distribution with true values of $\beta_1, \beta_2 = 1$. Note that for this simulation study, $\theta_0 = \theta_1$ and for each combination of ν, ρ and T, 5000 simulations are run.

			$\rho = 0.5$				$\rho = 0.9$			
ν	T	Method	$\hat{\beta_0}$	$\hat{\beta_1}$	$\hat{\nu}$	$\hat{\rho}$	$\hat{\beta_0}$	$\hat{\beta_1}$	$\hat{\nu}$	$\hat{\rho}$
3	60	CMLE	1.0002	0.9956	2.9901	0.4921	0.9912	1.0036	3.0096	0.9012
			(0.1030)	(0.0954)	(0.1384)		(0.1162)	(0.1291)	(0.1515)	
		CLS	0.9848	0.9802	3.1531	0.4800	0.9840	0.9885	2.9828	0.9109
			(0.3114)	(0.3295)	(0.3789)		(0.3822)	(0.3360)	(0.3237)	
		GMM	1.1003	0.9850	2.9882	0.5175	0.9843	1.1371	3.1387	0.8896
			(0.4162)	(0.4702)	(0.4380)		(0.4131)	(0.4674)	(0.4720)	
		GQL	0.9864	0.9992	2.9951	0.4958	0.9985	1.0033	2.9932	0.8964
			(0.1029)	(0.0952)	(0.1383)		(0.1160)	(0.1290)	(0.1515)	
	100	CMLE	1.0120	0.9906	3.0119	0.4971	1.0122	0.9912	3.0028	0.8912
			(0.0971)	(0.0940)	(0.1019)		(0.0949)	(0.1087)	(0.1007)	
		CLS	0.9833	0.9843	2.9820	0.4841	0.9871	0.9829	2.9877	0.9115
			(0.2609)	(0.2518)	(0.2011)		(0.2215)	(0.2473)	(0.2670)	
		GMM	1.1644	0.9853	2.9882	0.5131	0.9819	1.1811	2.9821	0.8871
			(0.3111)	(0.3271)	(0.3202)		(0.3297)	(0.3712)	(0.3110)	
		GQL	0.9971	0.9906	2.9960	0.5047	1.0111	1.0209	3.0189	0.8911
			(0.0970)	(0.0938)	(0.1018)		(0.0949)	(0.1086)	(0.1005)	
	500	CMLE	0.9984	0.9902	2.9904	0.5068	0.9961	1.0005	3.0281	0.9088
			(0.0311)	(0.0403)	(0.0790)		(0.0738)	(0.0969)	(0.0740)	
		CLS	0.9864	0.9808	2.9878	0.4833	0.9806	0.9895	2.9818	0.9113
			(0.1682)	(0.1167)	(0.1111)		(0.1304)	(0.1378)	(0.1620)	
		GMM	1.1669	0.9827	2.9853	0.5103	0.9852	1.1678	2.9861	0.8920
			(0.2433)	(0.2540)	(0.2796)		(0.2490)	(0.2533)	(0.2332)	
		GQL	1.0066	0.9987	3.0108	0.5013	1.0195	0.9978	3.0052	0.9069
			(0.0310)	(0.0401)	(0.0788)		(0.0736)	(0.0967)	(0.0738)	
2.5	60	CMLE	1.0250	1.0120	2.5003	0.4928	1.0182	1.0245	2.4945	0.9002
			(0.1141)	(0.1205)	(0.1740)		(0.1050)	(0.1326)	(0.1651)	
		CLS	0.9857	0.9875	2.4801	0.4803	0.9899	0.9872	2.5171	0.8822
			(0.3282)	(0.3774)	(0.3317)		(0.3677)	(0.3824)	(0.3570)	
		GMM	1.1173	0.9821	2.4832	0.5103	0.9823	1.1293	2.5152	0.8869
			(0.4122)	(0.4361)	(0.4573)		(0.4813)	(0.4052)	(0.4153)	
		GQL	1.0980	1.0166	2.4922	0.4904	1.0153	0.9924	2.5015	0.8925
			(0.1141)	(0.1204)	(0.1739)		(0.1048)	(0.1325)	(0.1650)	
	100	CMLE	1.0203	1.0178	2.5025	0.4996	1.0190	1.0130	2.4920	0.9076
			(0.1022)	(0.1116)	(0.1280)		(0.0902)	(0.1167)	(0.1339)	
		CLS	0.9809	0.9858	2.5170	0.4845	0.9832	0.9807	2.4836	0.8819
			(0.2386)	(0.2507)	(0.2394)		(0.2033)	(0.2744)	(0.2642)	
		GMM	1.1212	0.9850	2.4859	0.5179	0.9840	1.1042	2.5121	0.8866
			(0.3194)	(0.3207)	(0.3667)		(0.3543)	(0.3318)	(0.3201)	
		GQL	1.0248	1.0128	2.5085	0.4983	1.0206	1.0187	2.5047	0.8966
			(0.1021)	(0.1115)	(0.1279)		(0.0900)	(0.1166)	(0.1339)	
	500	CMLE	1.0106	1.0112	2.4971	0.4903	1.0195	1.0190	2.5018	0.8930
			(0.0861)	(0.0725)	(0.0796)		(0.0792)	(0.0810)	(0.0930)	
		CLS	0.9889	0.9826	2.4891	0.4848	0.9888	0.9807	2.5159	0.8819
			(0.2082)	(0.2403)	(0.2273)		(0.2181)	(0.2339)	(0.2698)	
		GMM	1.1231	0.9808	2.4883	0.5192	0.9856	1.1058	2.5189	0.8812
			(0.3120)	(0.3243)	(0.3415)		(0.3316)	(0.3169)	(0.3253)	
		GQL	1.0159	1.0171	2.5013	0.5062	1.0185	1.0134	2.5008	0.8937
			(0.0858)	(0.0723)	(0.0795)		(0.0790)	(0.0810)	(0.0929)	

Table I. Estimates of the parameters and standard errors under different estimation methods for non-stationary Com-Poisson model

			$\rho=0.5$				$\rho=0.9$			
ν	T	Method	$\hat{\beta_0}$	$\hat{\beta_1}$	$\hat{\nu}$	$\hat{\rho}$	$\hat{\beta_0}$	$\hat{\beta_1}$	$\hat{\nu}$	$\hat{\rho}$
2	60	CMLE	1.0042	0.9991	2.0040	0.4910	0.9914	1.0127	2.0108	0.8907
			(0.1110)	(0.1105)	(0.1278)		(0.1382)	(0.1283)	(0.1100)	
		CLS	0.9855	0.9808	1.8991	0.4815	0.9836	0.9839	2.0174	0.9120
			(0.3323)	(0.3069)	(0.3160)		(0.3153)	(0.3749)	(0.3560)	
		GMM	1.1141	0.9824	2.1157	0.5189	0.9890	1.1137	2.1032	0.8834
			(0.4581)	(0.4999)	(0.4860)		(0.4281)	(0.4442)	(0.4668)	
		GQL	0.9913	0.9901	2.0175	0.4941	0.9970	1.0097	1.9568	0.8955
			(0.1110)	(0.1104)	(0.1277)		(0.1380)	(0.1281)	(0.1100)	
	100	CMLE	1.0131	0.9996	2.0141	0.4924	1.0027	0.9981	2.0160	0.8974
			(0.0761)	(0.0860)	(0.0903)		(0.0978)	(0.0717)	(0.0913)	
		CLS	0.9851	0.9828	2.1291	0.4819	0.9858	0.9821	2.1386	0.9114
			(0.2481)	(0.2906)	(0.2419)		(0.2727)	(0.2586)	(0.2440)	
		GMM	1.1454	0.9834	2.1020	0.5105	0.9867	1.1478	2.1411	0.8889
			(0.3317)	(0.3147)	(0.3022)		(0.3895)	(0.3098)	(0.3252)	
		GQL	0.9913	0.9934	2.0139	0.5022	1.0071	1.0007	2.0111	0.8970
			(0.0760)	(0.0859)	(0.0901)		(0.0975)	(0.0715)	(0.0913)	
	500	CMLE	0.9966	0.9999	2.0130	0.5085	0.9901	1.0017	2.0140	0.9080
			(0.0580)	(0.0578)	(0.0539)		(0.0491)	(0.0503)	(0.0527)	
		CLS	0.9899	0.9873	2.1804	0.4856	0.9863	0.9880	2.1721	0.9102
			(0.1357)	(0.1191)	(0.1695)		(0.1837)	(0.1680)	(0.1487)	
		GMM	1.1550	0.9856	1.9117	0.5133	0.9840	1.1064	2.1128	0.8822
			(0.2605)	(0.2256)	(0.2925)		(0.2108)	(0.2027)	(0.2968)	
		GQL	1.0085	0.9961	2.0121	0.5048	1.0091	0.9958	2.0102	0.9090
			(0.0578)	(0.0575)	(0.0539)		(0.0490)	(0.0502)	(0.0527)	
1.5	60	CMLE	1.0044	1.0156	1.4988	0.4922	1.0102	1.0129	1.5054	0.9007
			(0.1092)	(0.1231)	(0.1042)		(0.1204)	(0.1263)	(0.1142)	
		CLS	0.9809	0.9898	1.5156	0.4890	0.9872	0.9892	1.5117	0.8809
			(0.3432)	(0.3209)	(0.3449)		(0.3821)	(0.3585)	(0.3937)	
		GMM	1.1131	0.9808	1.5182	0.5133	0.9899	1.1117	1.5108	0.8891
			(0.4362)	(0.4218)	(0.4195)		(0.4185)	(0.4033)	(0.4093)	
		GQL	1.0088	1.0111	1.5001	0.4942	1.0115	1.0166	1.5054	0.8903
			(0.1090)	(0.1230)	(0.1041)		(0.1204)	(0.1263)	(0.1141)	
	100	CMLE	1.0101	1.0189	1.5014	0.4903	1.0141	1.0193	1.5007	0.9085
			(0.1045)	(0.1023)	(0.1015)		(0.0997)	(0.1091)	(0.1010)	
		CLS	0.9807	0.9839	1.5136	0.4814	0.9827	0.9804	1.5152	0.8893
			(0.2314)	(0.2913)	(0.2513)		(0.2916)	(0.2328)	(0.2755)	
		GMM	1.1504	0.9839	1.5179	0.5139	0.9880	1.1330	1.5186	0.8898
			(0.3218)	(0.3129)	(0.3412)		(0.3335)	(0.3529)	(0.3405)	
		GQL	1.0171	1.0169	1.5050	0.4993	1.0187	1.0103	1.5077	0.8937
			(0.1044)	(0.1022)	(0.1015)		(0.0996)	(0.1090)	(0.1010)	
	500	CMLE	1.0055	1.0031	1.5008	0.4948	1.0014	1.0071	1.5009	0.8931
			(0.0882)	(0.0940)	(0.0801)		(0.0852)	(0.0936)	(0.0961)	
		CLS	0.9811	0.9855	1.5102	0.4871	0.9819	0.9841	1.5150	0.8899
			(0.2547)	(0.2777)	(0.2141)		(0.2256)	(0.2380)	(0.2177)	
		GMM	1.1674	0.9810	1.5140	0.5183	0.9829	1.1466	1.5190	0.8886
			(0.2768)	(0.2879)	(0.2822)		(0.2889)	(0.2835)	(0.2736)	
		GQL	1.0018	1.0021	1.5050	0.5042	1.0044	1.0092	1.5012	0.8960
			(0.0880)	(0.0940)	(0.0800)		(0.0851)	(0.0935)	(0.0960)	
1	60	CMLE	1.0122	0.9979	0.9991	0.4918	0.9972	1.0101	0.9916	0.8951
			(0.1002)	(0.0917)	(0.1702)		(0.1014)	(0.1121)	(0.1678)	
		CLS	0.9729	0.9761	0.9742	0.4728	0.9761	0.9701	0.9783	0.9311
			(0.3517)	(0.3718)	(0.3850)		(0.3900)	(0.3104)	(0.3061)	
		GMM	1.3325	0.9746	0.9738	0.5340	0.9764	1.3109	0.9790	0.8720
			(0.4072)	(0.4680)	(0.4414)		(0.4259)	(0.4767)	(0.4993)	
		GQL	0.9881	0.9991	0.9989	0.4911	0.9995	1.0097	0.9988	0.8977
			(0.1001)	(0.0915)	(0.1701)		(0.1012)	(0.1120)	(0.1677)	
	100	CMLE	1.0056	0.9991	1.0010	0.4943	1.0035	0.9999	1.0001	0.8984
			(0.0850)	(0.0888)	(0.1112)		(0.0975)	(0.1181)	(0.1101)	
		CLS	0.9721	0.9715	0.9760	0.4726	0.9748	0.9722	0.9788	0.9322
			(0.2887)	(0.2633)	(0.2120)		(0.2060)	(0.2528)	(0.2718)	
		GMM	1.3406	0.9747	0.9700	0.5328	0.9703	1.3355	0.9756	0.8702
			(0.3684)	(0.3034)	(0.3136)		(0.3338)	(0.3960)	(0.3022)	
		GQL	0.9998	0.9995	0.9999	0.5008	1.0020	1.0010	1.0015	0.8912
			(0.0842)	(0.0880)	(0.1110)		(0.0974)	(0.1180)	(0.1100)	
	500	CMLE	0.9995	0.9999	0.9997	0.5002	0.9999	1.0002	1.0125	0.9015
			(0.0438)	(0.0501)	(0.0919)		(0.0870)	(0.1000)	(0.0868)	
		CLS	0.9815	0.9829	0.9844	0.4852	0.9860	0.9832	0.9824	0.9269
			(0.1787)	(0.1019)	(0.1230)		(0.1073)	(0.1202)	(0.1906)	
		GMM	1.2130	0.9871	0.9829	0.5230	0.9832	1.2154	0.9832	0.8815
			(0.2621)	(0.2764)	(0.2900)		(0.2770)	(0.2763)	(0.2511)	
		GQL	1.0025	0.9936	1.0010	0.5014	1.0025	0.9999	1.0010	0.9015
			(0.0438)	(0.0500)	(0.0915)		(0.0869)	(0.0990)	(0.0867)	

Table II. Estimates of the parameters and standard errors under different estimation methods for non-stationary Com-Poisson model

			$\rho = 0.5$				$\rho = 0.9$			
ν	T	Method	$\hat{\beta_0}$	$\hat{\beta_1}$	$\hat{\nu}$	$\hat{\rho}$	$\hat{\beta_0}$	$\hat{\beta_1}$	$\hat{\nu}$	$\hat{\rho}$
0.9	60	CMLE	1.0110	1.0107	0.8989	0.4909	1.0110	1.0105	0.8990	0.9028
			(0.1201)	(0.1175)	(0.1881)		(0.1111)	(0.1230)	(0.1776)	
		CLS	0.9778	0.9794	0.9112	0.4782	0.9779	0.9734	0.9132	0.8714
			(0.3463)	(0.3871)	(0.3417)		(0.3718)	(0.3961)	(0.3850)	
		GMM	1.3583	0.9744	0.9166	0.5330	0.9750	1.3340	0.9187	0.8710
			(0.4259)	(0.4254)	(0.4311)		(0.4926)	(0.4180)	(0.4041)	
		GQL	1.0112	1.0113	0.9001	0.4905	1.0108	1.0106	0.9008	0.8945
			(0.1200)	(0.1173)	(0.1880)		(0.1110)	(0.1227)	(0.1775)	
	100	CMLE	1.0105	1.0118	0.9112	0.4930	1.0104	1.0103	0.8883	0.9073
			(0.1001)	(0.1097)	(0.1331)		(0.0909)	(0.1021)	(0.1512)	
		CLS	0.9736	0.9715	0.9199	0.4740	0.9742	0.9763	0.9112	0.8714
			(0.2228)	(0.2721)	(0.2530)		(0.2156)	(0.2895)	(0.2923)	
		GMM	1.3390	0.9713	0.9134	0.5329	0.9783	1.3350	0.9181	0.8709
			(0.3528)	(0.3680)	(0.3824)		(0.3728)	(0.3613)	(0.3472)	
		GQL	1.0102	1.0106	0.8991	0.4947	1.0103	1.0104	0.8980	0.8919
			(0.1000)	(0.1095)	(0.1330)		(0.0907)	(0.1020)	(0.1512)	
	500	CMLE	1.0106	1.0101	0.8999	0.4918	1.0104	1.0102	0.9001	0.8962
			(0.0985)	(0.0997)	(0.0882)		(0.0825)	(0.0922)	(0.0901)	
		CLS	0.9895	0.9805	0.9160	0.4826	0.9805	0.9837	0.9109	0.8816
			(0.2135)	(0.2529)	(0.2119)		(0.2011)	(0.2267)	(0.2821)	
		GMM	1.2595	0.9810	0.9107	0.5212	0.9840	1.2246	0.9141	0.8829
			(0.3027)	(0.3546)	(0.3306)		(0.3299)	(0.3026)	(0.3313)	
		GQL	1.0104	1.0157	0.9033	0.5010	1.0103	1.0102	0.9083	0.8991
			(0.0980)	(0.0995)	(0.0882)		(0.0824)	(0.0921)	(0.0900)	
0.5	60	CMLE	1.0039	0.9930	0.4914	0.4966	0.9959	1.0113	0.4944	0.8917
			(0.1095)	(0.1057)	(0.1105)		(0.1260)	(0.1131)	(0.1244)	
		CLS	0.9710	0.9750	0.4753	0.4762	0.9711	0.9710	0.4755	0.9350
			(0.3993)	(0.3128)	(0.3057)		(0.3293)	(0.3203)	(0.3263)	
		GMM	1.3376	0.9712	0.4765	0.5334	0.9777	1.3145	0.4727	0.8787
			(0.4691)	(0.4482)	(0.4214)		(0.4513)	(0.4632)	(0.4130)	
		GQL	0.9950	0.9967	0.4910	0.4909	0.9935	1.0045	0.4960	0.8943
			(0.1092)	(0.1055)	(0.1100)		(0.1260)	(0.1130)	(0.1243)	
	100	CMLE	1.0050	0.9936	0.4901	0.4921	1.0085	0.9938	0.4917	0.8990
			(0.0882)	(0.0881)	(0.0861)		(0.0971)	(0.0808)	(0.0885)	
		CLS	0.9723	0.9742	0.4758	0.4703	0.9791	0.9731	0.4714	0.9305
			(0.2396)	(0.2092)	(0.2151)		(0.2371)	(0.2110)	(0.2450)	
		GMM	1.3960	0.9752	0.4725	0.5380	0.9749	1.3355	0.4771	0.8786
			(0.3540)	(0.3533)	(0.3008)		(0.3586)	(0.3921)	(0.3052)	
		GQL	0.9954	0.9932	0.4906	0.5011	1.0015	1.0050	0.5094	0.8942
			(0.0880)	(0.0880)	(0.0861)		(0.0970)	(0.0807)	(0.0884)	
	500	CMLE	0.9936	0.9983	0.4930	0.5064	0.9905	1.0061	0.5096	0.9035
			(0.0431)	(0.0476)	(0.0412)		(0.0400)	(0.0560)	(0.0497)	
		CLS	0.9872	0.9840	0.4892	0.4817	0.9859	0.9835	0.4836	0.9200
			(0.1453)	(0.1115)	(0.1302)		(0.1484)	(0.1817)	(0.1813)	
		GMM	1.2871	0.9891	0.4858	0.5213	0.9862	1.2066	0.4864	0.8802
			(0.2027)	(0.2948)	(0.2960)		(0.2198)	(0.2740)	(0.2422)	
		GQL	1.0079	0.9954	0.5068	0.5090	1.0049	0.9956	0.5022	0.9035
			(0.0430)	(0.0475)	(0.0412)		(0.0400)	(0.0559)	(0.0496)	
0.3	60	CMLE	1.0087	1.0248	0.2905	0.4980	1.0261	1.0223	0.3097	0.9018
			(0.1188)	(0.1162)	(0.1182)		(0.1118)	(0.1115)	(0.1106)	
		CLS	0.9787	0.9709	0.2741	0.4739	0.9759	0.9720	0.2715	0.8701
			(0.3884)	(0.3816)	(0.3620)		(0.3512)	(0.3048)	(0.3810)	
		GMM	1.3642	0.9781	0.2702	0.5321	0.9796	1.3327	0.2729	0.8700
			(0.4453)	(0.4126)	(0.4313)		(0.4275)	(0.4330)	(0.4730)	
		GQL	1.0126	1.0160	0.3008	0.4903	1.0107	1.0183	0.3040	0.8971
			(0.1186)	(0.1160)	(0.1182)		(0.1118)	(0.1114)	(0.1105)	
	100	CMLE	1.0112	1.0197	0.3056	0.4933	1.0130	1.0115	0.3045	0.9015
			(0.1037)	(0.1006)	(0.1022)		(0.0985)	(0.1052)	(0.1022)	
		CLS	0.9735	0.9752	0.2730	0.4729	0.9714	0.9774	0.2734	0.8700
			(0.2809)	(0.2484)	(0.2054)		(0.2400)	(0.2307)	(0.2464)	
		GMM	1.3301	0.9761	0.2712	0.5378	0.9714	1.3340	0.2703	0.8796
			(0.3804)	(0.3205)	(0.3608)		(0.3421)	(0.3204)	(0.3735)	
		GQL	1.0164	1.0120	0.2904	0.4962	1.0148	1.0175	0.2991	0.8930
			(0.1036)	(0.1005)	(0.1022)		(0.0984)	(0.1052)	(0.1020)	
	500	CMLE	1.0020	1.0083	0.2961	0.4998	1.0050	1.0014	0.3064	0.8924
			(0.0955)	(0.0919)	(0.0820)		(0.0838)	(0.0969)	(0.0931)	
		CLS	0.9869	0.9808	0.2818	0.4851	0.9804	0.9865	0.2880	0.8862
			(0.2303)	(0.2287)	(0.2000)		(0.2206)	(0.2085)	(0.2231)	
		GMM	1.2903	0.9857	0.2897	0.5229	0.9880	1.2178	0.2898	0.8832
			(0.2540)	(0.2475)	(0.2820)		(0.2760)	(0.2806)	(0.2406)	
		GQL	1.0031	1.0088	0.3091	0.5027	1.0090	1.0048	0.3087	0.8988
			(0.0954)	(0.0919)	(0.0820)		(0.0837)	(0.0967)	(0.0930)	

Table III. Estimates of the parameters and standard errors under different estimation methods for non-stationary Com-Poisson model

The tables display the simulated mean estimates of the regression, dispersion and correlation parameters along with the standard errors in brackets. These converged estimates are obtained by assuming small initial values in the iterative processes described in equations (16) and (20). As the number of time points increases, the standard errors across the different methods decreases, but the standard errors in the GQL approach is lower than those based on the GMM and CLS approaches while comparatively same as in the CMLE approach. This pattern is noted irrespective of the correlation parameter. In term of non-convergent simulations, for $T = 100$ and $\rho = 0.9$, CMLE failed in 2500 simulations, CLS in 1200, GMM in 1000 and GQL in 550, for $\nu = 0.5$. Almost the same number of failures is noted when $T = 500$, for the same level of dispersion. On the other hand, for $T = 500$, $\rho = 0.5$ and $\nu = 0.3$, CMLE flopped in 2700 simulations, CLS in 1350, GMM in 1100 and GQL in 600. When $T = 100$, $\nu = 1.5$ and $\rho = 0.5$, CMLE failed in 1650 simulations, CLS failed in 950 simulations, GMM in 830 and GQL in 425. For $T = 500$, CMLE flopped in 2525 simulations, CLS flopped in 1350, GMM in 1275 and GQL in 525. When $\nu = 2$, T=100 and $\rho = 0.5$, the non-convergent simulations are 1775 for CMLE, 975 for CLS, 825 for GMM and 320 for GQL. When T=500, the non-convergent simulations are 1550 for CMLE, 875 for CLS, 675 for GMM and 325 for GQL. In addition, when $T = 100$, $\nu = 2.5$ and $\rho = 0.9$, CMLE failed in 1950 simulations, CLS in 1020, GMM in 880 and GQL in 250. When $T = 500$, the number of non-convergent simulations are 2025 for CMLE, 1075 for CLS, 930 for GMM and 315 for GQL for similar values of serial and dispersion parameter. However, the non-convergent simulations are slightly less when $\rho = 0.5$. For $\nu = 3$, $T = 100$ and $\rho = 0.5$, CMLE failed in 2075 simulations, CLS in 1220, GMM in 980 and GQL in 350. As T increases to 500, for similar values of ν and ρ, the non-convergent simulations are 2040 for CMLE, 1180 for CLS, 1050 for GMM and 390 for GQL. Moreover, there were also some non-convergent simulations when $\nu = 1$. For $T = 100$ and $\rho = 0.5$, CMLE flopped in 1525 simulations, CLS in 810, GMM in 720 and GQL in 280. When T=500, the non-convergent simulations increases and are 1650 for CMLE, 880 for CLS, 775 for GMM and 320 for GQL. It is noted that the non-convergent simulations increase slightly when $\rho = 0.9$. To summarize, we conclude here that the GQL approach yields far better estimates than CLS and GMM under both high and low serial-correlation and dispersion parameter and is almost equally efficient as CMLE as confirmed by Sutradhar et al. [2014], but computationally more stable.

5. APPLICATION OF THE COM-POISSON INAR(1) PROCESS TO THE MCB DATA

The Mauritius Commercial Bank is one of the most long lasting and proficient banking institution in Mauritius and is also an active institution in the stock exchange market in the country. We have collected some of its daily number of transactions (TRANS) that took place on the Stock Exchange Mauritius (SEM) from the period 01.03.2015 to 01.06.2015 and that makes a total of 68 days excluding Saturdays, Sundays and public holidays. Note that SEM starts its operation at 09.00 a.m and closes at 13.30 p.m local time. The summary statistics show that TRANS has a mean of 7.36 with a variance of 10.36 thus clearly indicating over-dispersion and with an empirical lag 1 correlation of 0.69. The factors influencing TRANS were the news or rumour effect, the Monday effect and the time of the day. The news variable is a categorical variable with 0 indicating that no news were in circulation in day t while 1 indicates either positive or negative news were rumoured in the local or international market that influences investor's decision. The Monday effect is another binary variable with 1 indicating Monday and 0 for the other days in the week. The purpose of this variable is that it is perceived in the global market that Monday is the day where the highest number of investments may occur. The time of day variable is categorized into the morning session coded as 0 and the afternoon session coded as 1. Thus, along with the number of transactions, data related to these three factors were also recorded. The INAR(1) CMP model was fitted to the data and the parameters were estimated by CMLE, GMM, GQL and CLS. The results are displayed below:

Method	INTC	Monday	Time of Day	News	$\hat{v}$	$\hat{\rho}$
CMLE	0.2582	0.1480	-0.0327	0.3210	0.8111	0.4532
	(0.4321)	(0.0057)	(0.0040)	(0.0109)	(0.0048)	
CLS	0.2601	0.2512	-0.0418	0.2133	0.9120	0.5529
	(0.5521)	(0.0227)	(0.0190)	(0.0311)	(0.0178)	
GMM	0.2799	0.2410	-0.0428	0.2158	0.9185	0.5147
	(0.6101)	(0.0360)	(0.0198)	(0.0382)	(0.0195)	
GQL	0.2601	0.1485	-0.0330	0.3222	0.8164	0.4529
	(0.4318)	(0.0056)	(0.0040)	(0.0107)	(0.0047)	

Table IV. Estimates of the regression and over-dispersion parameters using CMLE, CLS, GMM and GQL approaches

From the above results, Monday and the news effects have positive coefficients indicating that there is a significant number of movements in the MCB transactions at SEM on Monday and during the day while the negative sign in the time of day variable shows that more of the transactions is taking place in the morning session rather than in the afternoon. The estimates under the different methods are almost consistent, but GQL yields lower standard errors than GMM and CLS, however, comparatively equally efficient estimates as CMLE. However, CMLE yielded some computational problems in estimating the hessian entries due to the singularity of the covariance matrix. To avoid these complications, the initial values were refined. The over-dispersion estimates are all significant and the correlation parameter is reliable. Using the correlation coefficient, the non-stationary correlations are easily obtained through equations (10).

6. CONCLUSION

This paper has considered a novel non-stationary INAR(1) discrete time series observation-driven process based on the COM-Poisson model. The INAR(1) generating process was developed along with the derivation of the moments and covariance expressions. As for the estimation of the unknown parameters, the CMLE, CLS, GMM and GQL estimating functions were developed and it was remarked that the CMLE approach was rather computationally cumbersome and yielded a significant number of non-convergent simulations while GQL was computationally more appealing and provides consistent and equally efficient estimates than the other approaches.

REFERENCES

AL-OSH, M. A. AND ALZAID, A. A. 1988. Integer-valued moving average (inma) process. *Statistical Papers 29*, 281–300.

ALZAID, A. AND AL OSH, M. 1990. An integer-valued pth order autoregressive structure (INAR(p)) process. *Journal of Applied 27*, 314–324.

CASTILLO, J. AND PEREZ-CASSANY, M. 2005. Overdispersed and underdispersed Poisson generalizations. *Journal of Statistical Planning and Inference 134*, 486–500.

CHAN, K. AND LEDOLTER, J. 1995. Monte carlo EM estimation for time series models involving counts. *Journal of the American Statistical Association 90(429)*, 242–252.

CONSUL, P. AND SHOUKRI, M. 1984. Maximum likelihood estimation for the generalized Poisson distribution. *Communication in Statistics, Theory and Methodology 17*, 1533–1547.

COX, D. 1981. Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics 8(2)*, 93–115.

CROWDER, M. 1995. On the use of a working correlation matrix in using generalized linear models for repeated measures. *Biometrika 82*, 407–10.

DEY, S., DEY, T., ALI, S., AND MULEKAR, M. 2016. Two-parameter Maxwell distribution: Properties and different methods of estimation. *Journal of Statistical Theory and Practice 10,* 2, 291–310.

DURBIN, J. AND KOOPMAN, S. 2000. Time series analysis of non-gaussian observations based on state space models from both classical and bayesian perspectives. *Journal of the Royal Statistical Society B62*, 3–56.

GUIKEMA, S. AND COFFELT, J. 2008. A flexible count data regression model for risk analysis. *Risk Analysis 28*, 213–223.

JUNG, R., KUKUK, M., AND TREMAYNE, A. 2006. Time series of count data:,modelling, estimation and diagnostics. *Computational Statistics and Data Analysis 51(4)*, 2350–2364.

JUNG, R. AND LIESENFELD, R. 2001. Estimating time series models for count data using efficient importance sampling. *Allgemeines Statistical Archives 85(4)*, 387–407.

JUNG, R. AND TREMAYNE, A. 2003. Testing for serial dependence in time series models of counts. *Journal of Time Series Analysis 24,* 1, 65–84.

LIANG, K. AND ZEGER, S. 1986. Longitudinal data analysis using generalized linear models. *Biometrika 73*, 13–22.

LORD, D., GEEDIPALLY, S., AND GUIKEMA, S. 2010. Extension of the application of Conway-Maxwell Poisson models:analysing traffic crash data exhibiting underdispersion. *Risk Analysis 30*, 1268–1276.

MAMODE KHAN, N. AND JOWAHEER, V. 2013. Comparing joint GQL estimation and GMM adaptive estimation in COM-Poisson longitudinal regression model. *Commun Stat-Simul C. 42(4)*, 755–770.

MAMODE KHAN, N., SUNECHER, Y., AND JOWAHEER, V. 2016. Modelling a non-stationary BINAR(1) Poisson process. *Journal of Statistical Computation and Simulation 86*, 3106–3126.

MCKENZIE, E. 1986. Autoregressive moving-average processes with Negative Binomial and geometric marginal distributions. *Advanced Applied Probability 18*, 679–705.

MINKA, T., SHMUELI, G., KADANE, J., BORLE, S., AND BOATWRIGHT, P. 2003. Computing with the COM-Poisson distribution. Tech. rep., Carnegie Mellon University.

NADARAJAH, S. 2009. Useful moment and CDF formulations for the com-poisson distribution. *Statistical papers 50*, 617–622.

QU, A. AND LINDSAY, B. 2003. Building adaptive estimating equations when inverse of covariance estimation is difficult. *Journal of Royal Statistical Society 65*, 127–142.

QUORESHI, A. 2008. A vector integer-valued moving average model for high frequency financial count data. *Economics Letters 101*, 258–261.

RAVISHANKER, N., SERHIYENKO, V., AND WILLIG, M. 2014. Hierarchical dynamic models for multivariate time series of counts. *Statistics and Its Interface 7*, 559–570.

SELLERS, K., MORRIS, D., AND BALAKRISHNAN, N. 2016. Bivariate Conway-Maxwell Poisson distribution: Formulation, properties and inference. *Journal of Multivariate Analysis 150*, 152–168.

SELLERS, K., S.BORLE, AND SHMUELI, G. 2012. The COM-Poisson model for count data: A survey of methods and applications. *Applied Stochastic Models in Business and Industry, 22*, 104–116.

SHMUELI, G., MINKA, T., BORLE, J., AND BOATWRIGHT, P. 2005. A useful distribution for fitting discrete data. *Applied Statistics,Journal of Royal Statistical Society C*.

STEUTEL, F. AND VAN HARN, K. 1979. Discrete analogues of self-decomposability and stability. *The Annals of Probability 7*, 3893–899.

SUTRADHAR, B. 2003. An overview on regression models for discrete longitudinal responses. *Statistical Science 18,* 3, 377–393.

SUTRADHAR, B. AND DAS, K. 1999. On the efficiency of regression estimators in generalised linear models for longitudinal data. *Biometrika 86*, 459–65.

SUTRADHAR, B., JOWAHEER, V., AND RAO, P. 2014. Remarks on asymptotic efficient estimation for regression effects in stationary and non-stationary models for panel count data. *Brazilian Journal of Probability and Statistics 28(2)*, 241–254.

WEDDERBURN, R. 1974. Quasi-likelihood functions,generalized linear models,and the Gauss-Newton method. *Biometrika 61,* 3 (December), 439–47.

The derivatives of the conditional likelihood function for the CMLE are:

Differentiating with respect to ρ, we have

$$\sum_{t=1}^{T}\frac{\partial[log \quad L(\rho,\beta,\nu|y)]}{\partial\rho}=\sum_{t=1}^{T}[\sum_{i=0}^{m}(\frac{\lambda_t^{y_t-i}}{(y_t-i)!^{\nu}})(\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})}{\lambda_t^{\frac{\nu-1}{2\nu}}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}})\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}[-(y_{t-1}-i)\rho^i(1-\rho)^{y_{t-1}-i-1}$$
$$+i\rho^{i-1}(1-\rho)^{y_{t-1}-i}]/[Pr(y_t|y_{t-1})] \quad (26)$$

Differentiating with respect to ν, we have

$$\sum_{t=1}^{T}\frac{\partial[log \quad L(\rho,\beta,\nu|y)]}{\partial\nu}=\sum_{t=1}^{T}[\sum_{i=0}^{m}\lambda_t^{y_t-i}\rho^i(1-\rho)^{y_{t-1}-i}\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}$$
$$[\frac{(\lambda_t^{\frac{1}{\nu}}-\frac{\lambda_t^{\frac{1}{\nu}}ln(\lambda_t)}{\nu})\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}}$$
$$-\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}ln((y_t-i)!)\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}}$$
$$+\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})ln(\lambda_t)(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}}$$
$$-\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}ln(2\pi)}{2\sqrt{\nu}}$$
$$-\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{2\nu^{\frac{3}{2}}}]]/[Pr(y_t|y_{t-1})] \quad (27)$$

Differentiating with respect to β_j, we have

$$\sum_{t=1}^{T}\frac{\partial[log \quad L(\rho,\beta,\nu|y)]}{\partial\beta_j}=\sum_{t=1}^{T}[[\sum_{i=0}^{m}\frac{1}{(y_t-i)!^{\nu}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}}\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}\rho^i(1-\rho)^{y_{t-1}-i}$$
$$[\frac{\lambda_t^{y_t-i}\lambda_t^{\frac{1}{\nu}}x_t^T exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}+\frac{(x_t^T\beta_j)^{y_t-i}(y_t-i)\lambda_t^{y_t-i}exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\beta_j\lambda_t^{\frac{1-\nu}{2\nu}})}$$
$$-\frac{\lambda_t^{y_t-i}x_t^T exp(\nu\lambda_t^{\frac{1}{\nu}})(\frac{1-\nu}{2\nu})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}]]/[Pr(y_t|y_{t-1})]] \quad (28)$$

The derivation of the double derivatives in the CMLE approach are shown below:

$$
\begin{aligned}
\sum_{t=1}^{T} \frac{\partial^2 [log \quad L(\rho,\beta,\nu|y)]}{\partial \nu^2} = & \sum_{t=1}^{T} [\sum_{i=0}^{m} \lambda_t^{y_t-i} \rho^i (1-\rho)^{y_{t-1}-i} \begin{pmatrix} y_{t-1} \\ i \end{pmatrix} \\
& [-\frac{2\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu} ln((y_t-i)!)\lambda_t^{\frac{1-\nu}{2\nu}}(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})ln(\lambda_t)(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
& -\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})ln(\lambda_t)(2\pi)^{\frac{1-\nu}{2}}}{\nu^{\frac{3}{2}}} \\
& -\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})ln(\lambda_t)(2\pi)^{\frac{1-\nu}{2}} ln(2\pi)}{\sqrt{\nu}} \\
& +\frac{\lambda_t^{\frac{1}{\nu}} ln(\lambda_t)^2 \exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\nu^{\frac{7}{2}}} \\
& +\frac{2(\lambda_t^{\frac{1}{\nu}}-\frac{\lambda_t^{\frac{1}{\nu}} ln(\lambda_t)}{\nu}) ln(\lambda_t)\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
& +\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}} ln((y_t-i)!)^2(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
& +\frac{1}{4}\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}} ln(2\pi)^2(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
& +\frac{3}{4}\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\nu^{\frac{5}{2}}} \\
& +\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(\frac{1}{\nu^2}+\frac{1-\nu}{\nu^3})ln(\lambda_t)(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
& +\frac{1}{2}\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}} ln(2\pi)(2\pi)^{\frac{1-\nu}{2}}}{\nu^{\frac{3}{2}}} \\
& -\frac{(\lambda_t^{\frac{1}{\nu}}-\frac{\lambda_t^{\frac{1}{\nu}} ln(\lambda_t)}{\nu})\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\nu^{\frac{3}{2}}} \\
& +\frac{(\lambda_t^{\frac{1}{\nu}}-\frac{\lambda_t^{\frac{1}{\nu}} ln(\lambda_t)}{\nu})^2\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
& -2\frac{(\lambda_t^{\frac{1}{\nu}}-\frac{\lambda_t^{\frac{1}{\nu}} ln(\lambda_t)}{\nu})\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu} ln((y_t-i)!)\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}}
\end{aligned}
$$

$$
\begin{aligned}
&- \frac{(\lambda_t^{\frac{1}{\nu}} - \frac{\lambda_t^{\frac{1}{\nu}} ln(\lambda_t)}{\nu}) \exp(\nu \lambda_t^{\frac{1}{\nu}})(y_t - i)!^{-\nu} ln(2\pi) \lambda_t^{\frac{1-\nu}{2\nu}} (2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&+ \frac{\exp(\nu \lambda_t^{\frac{1}{\nu}})(y_t - i)!^{-\nu} ln((y_t - i)!) \lambda_t^{\frac{1-\nu}{2\nu}} (2\pi)^{\frac{1-\nu}{2}}}{\nu^{\frac{3}{2}}} \\
&+ \frac{ln(\lambda_t)^2 \exp(\nu \lambda_t^{\frac{1}{\nu}})(y_t - i)!^{-\nu} \lambda_t^{\frac{1-\nu}{2\nu}} (-\frac{1}{2\nu} - \frac{1-\nu}{2\nu^2})^2 (2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&+ \frac{\exp(\nu \lambda_t^{\frac{1}{\nu}})(y_t - i)!^{-\nu} ln((y_t - i)!) \lambda_t^{\frac{1-\nu}{2\nu}} (2\pi)^{\frac{1-\nu}{2}} ln(2\pi)}{\sqrt{\nu}}]/[Pr(y_t|y_{t-1})]
\end{aligned}
$$

$$
\begin{aligned}
&-[\sum_{i=0}^{m} \lambda_t^{y_t-i} \rho^i (1-\rho)^{y_{t-1}-i} \begin{pmatrix} y_{t-1} \\ i \end{pmatrix} [\frac{(\lambda_t^{\frac{1}{\nu}} - \frac{\lambda_t^{\frac{1}{\nu}} ln(\lambda_t)}{\nu}) \exp(\nu \lambda_t^{\frac{1}{\nu}})(y_t - i)!^{-\nu} \lambda_t^{\frac{1-\nu}{2\nu}} (2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&- \frac{\exp(\nu \lambda_t^{\frac{1}{\nu}})(y_t - i)!^{-\nu} ln((y_t - i)!) \lambda_t^{\frac{1-\nu}{2\nu}} (2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&+ \frac{\exp(\nu \lambda_t^{\frac{1}{\nu}})(y_t - i)!^{-\nu} \lambda_t^{\frac{1-\nu}{2\nu}} (-\frac{1}{2\nu} - \frac{1-\nu}{2\nu^2}) ln(\lambda_t) (2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&- \frac{\exp(\nu \lambda_t^{\frac{1}{\nu}})(y_t - i)!^{-\nu} \lambda_t^{\frac{1-\nu}{2\nu}} (2\pi)^{\frac{1-\nu}{2}} ln(2\pi)}{2\sqrt{\nu}} \\
&- \frac{\exp(\nu \lambda_t^{\frac{1}{\nu}})(y_t - i)!^{-\nu} \lambda_t^{\frac{1-\nu}{2\nu}} (2\pi)^{\frac{1-\nu}{2}}}{2\nu^{\frac{3}{2}}}]^2 / [Pr(y_t|y_{t-1})]^2]
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
\sum_{t=1}^{T} \frac{\partial^2 [log \quad L(\rho, \beta, \nu|y)]}{\partial \rho^2} = &\sum_{t=1}^{T} [[\sum_{i=0}^{m} (\frac{\lambda_t^{y_t-i}}{(y_t - i)!^{\nu}})(\frac{\exp(\nu \lambda_t^{\frac{1}{\nu}})}{\lambda_t^{\frac{\nu-1}{2\nu}} (2\pi)^{\frac{\nu-1}{2}} \sqrt{\nu}}) \begin{pmatrix} y_{t-1} \\ i \end{pmatrix} [-i\rho^{i-1}(1-\rho)^{y_{t-1}-i-1} \\
&+ (y_{t-1} - i)\rho^i (1-\rho)^{y_{t-1}-i-2} + i(i-1)\rho^{i-2}(1-\rho)^{y_{t-1}-i} \\
&- i(y_{t-1} - i)\rho^{i-1}(1-\rho)^{y_{t-1}-i-1}]] / \sum_{i=0}^{m} Pr(y_t|y_{t-1}) \\
&-[\sum_{i=0}^{m} (\frac{\lambda_t^{y_t-i}}{(y_t - i)!^{\nu}})(\frac{\exp(\nu \lambda_t^{\frac{1}{\nu}})}{\lambda_t^{\frac{\nu-1}{2\nu}} (2\pi)^{\frac{\nu-1}{2}} \sqrt{\nu}}) \begin{pmatrix} y_{t-1} \\ i \end{pmatrix} \\
&\times [-(y_{t-1} - i)\rho^i (1-\rho)^{y_{t-1}-i-1} + i\rho^{i-1}(1-\rho)^{y_{t-1}-i}]]^2 / Pr(y_t|y_{t-1})^2]
\end{aligned} \tag{30}
$$

$$\sum_{t=1}^{T}\frac{\partial^2[log \quad L(\rho,\beta,\nu|y)]}{\partial\beta_j^2} = \sum_{t=1}^{T}[[[\sum_{i=0}^{m}\frac{1}{(y_t-i)!^{\nu}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}}\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}\rho^i(1-\rho)^{y_{t-1}-i}$$

$$[\frac{(x_t^T\beta_j)^{y_t-i}(y_t-i)^2\lambda_t^{y_t-i}exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\beta_j^2\lambda_t^{\frac{1-\nu}{2\nu}})} - \frac{(x_t^T\beta_j)^{y_t-i}(y_t-i)\lambda_t^{y_t-i}exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\beta_j^2\lambda_t^{\frac{1-\nu}{2\nu}})}$$

$$+\frac{[(x_t^T\beta_j)^{y_t-i}]^2(y_t-i)^2\lambda_t^{y_t-i}exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\beta_j^2\lambda_t^{\frac{1-\nu}{2\nu}})}$$

$$+\frac{2(x_t^T\beta_j)^{y_t-i}(y_t-i)\lambda_t^{y_t-i}\lambda_t^{\frac{1}{\nu}}x_t^Texp(\nu\lambda_t^{\frac{1}{\nu}})}{(\beta_j\lambda_t^{\frac{1-\nu}{2\nu}})}$$

$$-\frac{2(x_t^T\beta_j)^{y_t-i}(y_t-i)\lambda_t^{y_t-i}x_t^Texp(\nu\lambda_t^{\frac{1}{\nu}})(\frac{1-\nu}{2\nu})}{(\beta_j\lambda_t^{\frac{1-\nu}{2\nu}})} + \frac{\lambda_t^{y_t-i}\lambda_t^{\frac{1}{\nu}}(\frac{1}{\nu})(x_t^T)^2exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}$$

$$+\frac{\lambda_t^{y_t-i}(\lambda_t^{\frac{1}{\nu}})^2(x_t^T)^2exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\lambda_t^{\frac{1-\nu}{2\nu}})} - \frac{2\lambda_t^{y_t-i}\lambda_t^{\frac{1}{\nu}}(x_t^T)^2exp(\nu\lambda_t^{\frac{1}{\nu}})(\frac{1-\nu}{2\nu})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}$$

$$+\frac{\lambda_t^{y_t-i}(x_t^T)^2exp(\nu\lambda_t^{\frac{1}{\nu}})(\frac{1-\nu}{2\nu})^2}{(\lambda_t^{\frac{1-\nu}{2\nu}})}]]/[Pr(y_t|y_{t-1})]]$$

$$-[\sum_{i=0}^{m}\frac{1}{(y_t-i)!^{\nu}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}}\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}\rho^i(1-\rho)^{y_{t-1}-i}$$

$$[\frac{\lambda_t^{y_t-i}\lambda_t^{\frac{1}{\nu}}x_t^Texp(\nu\lambda_t^{\frac{1}{\nu}})}{(\lambda_t^{\frac{1-\nu}{2\nu}})} + \frac{(x_t^T\beta_j)^{y_t-i}(y_t-i)\lambda_t^{y_t-i}exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\beta_j\lambda_t^{\frac{1-\nu}{2\nu}})}$$

$$-\frac{\lambda_t^{y_t-i}x_t^Texp(\nu\lambda_t^{\frac{1}{\nu}})(\frac{1-\nu}{2\nu})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}]]^2/[Pr(y_t|y_{t-1})]^2] \tag{31}$$

$$
\begin{aligned}
\sum_{t=1}^{T}\frac{\partial^2[log\quad L(\rho,\beta,\nu|y)]}{\partial\rho\partial\nu} &= \sum_{t=1}^{T}[\sum_{i=0}^{m}\lambda_t^{y_t-i}\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}[-(y_{t-1}-i)\rho^i(1-\rho)^{y_{t-1}-i-1} \\
&+ i\rho^{i-1}(1-\rho)^{y_{t-1}-i}]\times[\frac{(\lambda_t^{\frac{1}{\nu}}-\frac{\lambda_t^{\frac{1}{\nu}}ln(\lambda_t)}{\nu})\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&- \frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}ln((y_t-i)!)\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&+ \frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})ln(\lambda_t)(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&- \frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}ln(2\pi)}{2\sqrt{\nu}} \\
&- \frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{2\nu^{\frac{3}{2}}}]]/[Pr(y_t|y_{t-1})] \\
&- [(\frac{\lambda_t^{y_t-i}}{(y_t-i)!^{\nu}})(\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})}{\lambda_t^{\frac{\nu-1}{2\nu}}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}})\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}[-(y_{t-1}-i)\rho^i(1-\rho)^{y_{t-1}-i-1} \\
&+ i\rho^{i-1}(1-\rho)^{y_{t-1}-i}] \\
&\times[\lambda_t^{y_t-i}\rho^i(1-\rho)^{y_{t-1}-i}\begin{pmatrix} y_{t-1} \\ i \end{pmatrix} \\
&[\frac{(\lambda_t^{\frac{1}{\nu}}-\frac{\lambda_t^{\frac{1}{\nu}}ln(\lambda_t)}{\nu})\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&- \frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}ln((y_t-i)!)\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&+ \frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})ln(\lambda_t)(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&- \frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}ln(2\pi)}{2\sqrt{\nu}} \\
&- \frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{2\nu^{\frac{3}{2}}}]]/[Pr(y_t|y_{t-1})]^2]] \qquad (32)
\end{aligned}
$$

$$\sum_{t=1}^{T}\frac{\partial^2[log \quad L(\rho,\beta,\nu|y)]}{\partial\beta_j\partial\rho} = \sum_{t=1}^{T}[[\sum_{i=0}^{m}\frac{1}{(y_t-i)!^{\nu}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}}\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}[-(y_{t-1}-i)\rho^i(1-\rho)^{y_{t-1}-i-1}$$
$$+i\rho^{i-1}(1-\rho)^{y_{t-1}-i}]\times[\frac{\lambda_t^{y_t-i}\lambda_t^{\frac{1}{\nu}}x_t^T exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}$$
$$+\frac{(x_t^T\beta_j)^{y_t-i}(y_t-i)\lambda_t^{y_t-i}exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\beta_j\lambda_t^{\frac{1-\nu}{2\nu}})}$$
$$-\frac{\lambda_t^{y_t-i}x_t^T exp(\nu\lambda_t^{\frac{1}{\nu}})(\frac{1-\nu}{2\nu})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}]]/[Pr(y_t|y_{t-1})]$$
$$-(\frac{\lambda_t^{y_t-i}}{(y_t-i)!^{\nu}})(\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})}{\lambda_t^{\frac{\nu-1}{2\nu}}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}})\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}[-(y_{t-1}-i)\rho^i(1-\rho)^{y_{t-1}-i-1}$$
$$+i\rho^{i-1}(1-\rho)^{y_{t-1}-i}]\times[\frac{1}{(y_t-i)!^{\nu}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}}\begin{pmatrix} y_{t-1} \\ i \end{pmatrix}\rho^i(1-\rho)^{y_{t-1}-i}$$
$$[\frac{\lambda_t^{y_t-i}\lambda_t^{\frac{1}{\nu}}x_t^T exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}+\frac{(x_t^T\beta_j)^{y_t-i}(y_t-i)\lambda_t^{y_t-i}exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\beta_j\lambda_t^{\frac{1-\nu}{2\nu}})}$$
$$-\frac{\lambda_t^{y_t-i}x_t^T exp(\nu\lambda_t^{\frac{1}{\nu}})(\frac{1-\nu}{2\nu})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}]]/[Pr(y_t|y_{t-1})]^2] \quad (33)$$

$$
\begin{aligned}
\sum_{t=1}^{T}\frac{\partial^2[log \quad L(\rho,\beta,\nu|y)]}{\partial\beta_j\partial\nu} &= \sum_{t=1}^{T}[[\sum_{i=0}^{m}\rho^i(1-\rho)^{y_{t-1}-i}\left(\begin{array}{c} y_{t-1} \\ i \end{array}\right) \\
&[\frac{\lambda_t^{y_t-i}(y_t-i)x_t^T(\lambda_t^{\frac{1}{\nu}}-\frac{\lambda_t^{\frac{1}{\nu}}ln(\lambda_t)}{\nu})\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&-\frac{\lambda_t^{y_t-i}\lambda_t^{\frac{1}{\nu}}x_t^T ln(\lambda_t)\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\nu^{\frac{5}{2}}} \\
&-\frac{\lambda_t^{y_t-i}(y_t-i)x_t^T\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}ln((y_t-i)!)\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&+\frac{\lambda_t^{y_t-i}(y_t-i)x_t^T\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})ln(\lambda_t)\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&+\frac{\lambda_t^{y_t-i}\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}x_t^T(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&-\frac{1}{2}\frac{\lambda_t^{y_t-i}(y_t-i)x_t^T\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}ln(2\pi)\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&-\frac{1}{2}\frac{\lambda_t^{y_t-i}(y_t-i)x_t^T\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\nu^{\frac{3}{2}}}]]/[Pr(y_t|y_{t-1})] \\
&-\sum_{i=0}^{m}\lambda_t^{y_t-i}\rho^i(1-\rho)^{y_{t-1}-i}\left(\begin{array}{c} y_{t-1} \\ i \end{array}\right) \\
&[\frac{(\lambda_t^{\frac{1}{\nu}}-\frac{\lambda_t^{\frac{1}{\nu}}ln(\lambda_t)}{\nu})\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&-\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}ln((y_t-i)!)\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&+\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(-\frac{1}{2\nu}-\frac{1-\nu}{2\nu^2})ln(\lambda_t)(2\pi)^{\frac{1-\nu}{2}}}{\sqrt{\nu}} \\
&-\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}ln(2\pi)}{2\sqrt{\nu}} \\
&-\frac{\exp(\nu\lambda_t^{\frac{1}{\nu}})(y_t-i)!^{-\nu}\lambda_t^{\frac{1-\nu}{2\nu}}(2\pi)^{\frac{1-\nu}{2}}}{2\nu^{\frac{3}{2}}}] \\
&\times[\frac{1}{(y_t-i)!^{\nu}(2\pi)^{\frac{\nu-1}{2}}\sqrt{\nu}}\left(\begin{array}{c} y_{t-1} \\ i \end{array}\right)\rho^i(1-\rho)^{y_{t-1}-i} \\
&[\frac{\lambda_t^{y_t-i}\lambda_t^{\frac{1}{\nu}}x_t^T exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}+\frac{(x_t^T\beta_j)^{y_t-i}(y_t-i)\lambda_t^{y_t-i}exp(\nu\lambda_t^{\frac{1}{\nu}})}{(\beta_j\lambda_t^{\frac{1-\nu}{2\nu}})} \\
&-\frac{\lambda_t^{y_t-i}x_t^T exp(\nu\lambda_t^{\frac{1}{\nu}})(\frac{1-\nu}{2\nu})}{(\lambda_t^{\frac{1-\nu}{2\nu}})}]/[Pr(y_t|y_{t-1})]^2]
\end{aligned} \tag{34}
$$

The derivatives of the criterion function for the CLS estimator are:

$$\frac{\partial Q(\psi)}{\partial \rho}=2\sum_{t=2}^{T}[y_t-\rho y_{t-1}-(\lambda_t^{1/\nu}-\frac{\nu-1}{2\nu})+\rho(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})](-y_{t-1}+(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})) \tag{35}$$

$$\frac{\partial^2 Q(\psi)}{\partial \rho^2}=2\sum_{t=2}^{T}[-y_{t-1}+(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})](-y_{t-1}+(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})) \tag{36}$$

$$\begin{aligned}\frac{\partial Q(\psi)}{\partial \nu}=&2\sum_{t=2}^{T}[y_t-\rho y_{t-1}-(\lambda_t^{1/\nu}-\frac{\nu-1}{2\nu})+\rho(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})](\frac{\lambda_t^{1/\nu}ln(\lambda_t)}{\nu^2}-\frac{\nu-1}{2\nu^2}+\frac{1}{2\nu}\\&-\rho\frac{(\lambda_{t-1}^{1/\nu}ln(\lambda_{t-1}))}{\nu^2}+\rho\frac{(\nu-1)}{2\nu^2}-\rho\frac{1}{2\nu})\end{aligned} \tag{37}$$

$$\begin{aligned}\frac{\partial^2 Q(\psi)}{\partial \nu^2}=&2\sum_{t=2}^{T}[\frac{\lambda_t^{1/\nu}ln(\lambda_t)}{\nu^2}+\frac{1}{2\nu}-\frac{1}{2}\frac{\nu-1}{\nu^2}+\rho(\frac{-\lambda_{t-1}^{1/\nu}ln(\lambda_{t-1})}{\nu^2}+\frac{(\nu-1)}{2\nu^2}-\frac{1}{2\nu})]^2\\&+[y_t-\rho y_{t-1}-\lambda_t^{1/\nu}+\frac{\nu-1}{2\nu}+\rho(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})]\times[-\frac{\lambda_t^{1/\nu}ln(\lambda_t)^2}{\nu^4}-2\frac{\lambda_t^{1/\nu}ln(\lambda_t)}{\nu^3}\\&-\frac{1}{\nu^2}+\frac{(\nu-1)}{\nu^3}+\rho(\frac{\lambda_{t-1}^{1/\nu}ln(\lambda_{t-1})^2}{\nu^4}+2\frac{\lambda_{t-1}^{1/\nu}ln(\lambda_{t-1})}{\nu^3}+\frac{1}{\nu^2}-\frac{\nu-1}{\nu^3})]\end{aligned} \tag{38}$$

$$\frac{\partial Q(\psi)}{\partial \beta_j}=2\sum_{t=2}^{T}[[y_t-\rho y_{t-1}-(\lambda_t^{1/\nu}-\frac{\nu-1}{2\nu})+\rho(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})](-\frac{\lambda_t^{(1/\nu)}x_t^T}{\nu}+\frac{\rho}{\nu}\lambda_{t-1}^{(1/\nu)}x_{t-1}^T)] \tag{39}$$

$$\begin{aligned}\frac{\partial^2 Q(\psi)}{\partial \beta_j^2}=&2\sum_{t=2}^{T}[[-\frac{\lambda_t^{(1/\nu)}x_t^T}{\nu}+\frac{\rho}{\nu}\lambda_{t-1}^{(1/\nu)}x_{t-1}^T]^2+[y_t-\rho y_{t-1}-\lambda_t^{1/\nu}+\frac{\nu-1}{2\nu}+\rho(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})]\\&\times[-\frac{\lambda_t^{(1/\nu)}(x_t^T)^2}{\nu^2}+\frac{\rho}{\nu^2}\lambda_{t-1}^{(1/\nu)}(x_{t-1}^T)^2]]\end{aligned} \tag{40}$$

$$\begin{aligned}\frac{\partial^2 Q(\psi)}{\partial \beta_j\partial \rho}=&2\sum_{t=2}^{T}[[-y_{t-1}+\lambda_t^{1/\nu}-\frac{\nu-1}{2\nu}]\times[-\frac{\lambda_t^{(1/\nu)}(x_t^T)}{\nu}+\frac{\rho}{\nu}\lambda_{t-1}^{(1/\nu)}(x_{t-1}^T)]\\&+\frac{[y_t-\rho y_{t-1}-\lambda_t^{1/\nu}+\frac{\nu-1}{2\nu}+\rho(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})]\times(\lambda_{t-1}^{(1/\nu)}x_{t-1}^T)}{\nu}]\end{aligned} \tag{41}$$

$$\begin{aligned}\frac{\partial^2 Q(\psi)}{\partial \beta_j\partial \nu}=&2\sum_{t=2}^{T}[[\frac{\lambda_t^{1/\nu}ln(\lambda_t)}{\nu^2}+\frac{1}{2\nu}-\frac{\nu-1}{2\nu^2}+\rho(-\frac{\lambda_{t-1}^{1/\nu}ln(\lambda_{t-1})}{\nu^2}-\frac{1}{2\nu}+\frac{\nu-1}{2\nu^2})]\\&\times(-\frac{\lambda_t^{1/\nu}x_t^T}{\nu}+\frac{\rho\lambda_{t-1}^{1/\nu}x_{t-1}^T}{\nu})+[y_t-\rho y_{t-1}-\lambda_t^{1/\nu}+\frac{\nu-1}{2\nu}+\rho(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})]\\&\times[\frac{\lambda_t^{(1/\nu)}ln(\lambda_t)(x_t^T)}{\nu^3}+\frac{\lambda_t^{(1/\nu)}(x_t^T)}{\nu^2}-\frac{\rho}{\nu^3}\lambda_{t-1}^{(1/\nu)}ln(\lambda_{t-1})(x_{t-1}^T)-\frac{\rho}{\nu^2}\lambda_{t-1}^{(1/\nu)}(x_{t-1}^T)]]\end{aligned} \tag{42}$$

$$\begin{aligned}\frac{\partial^2 Q(\psi)}{\partial \nu\partial \rho}=&2\sum_{t=2}^{T}[[\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu}]\times[\frac{\lambda_t^{1/\nu}ln(\lambda_t)}{\nu^2}-\frac{\nu-1}{2\nu^2}+\frac{1}{2\nu}-\rho\frac{(\lambda_{t-1}^{1/\nu}ln(\lambda_{t-1}))}{\nu^2}+\rho\frac{(\nu-1)}{2\nu^2}-\rho\frac{1}{2\nu}]\\&+[y_t-\rho y_{t-1}-\lambda_t^{1/\nu}+\frac{\nu-1}{2\nu}+\rho(\lambda_{t-1}^{1/\nu}-\frac{\nu-1}{2\nu})]\times[-\frac{\lambda_{t-1}^{1/\nu}ln(\lambda_{t-1})}{\nu^2}+\frac{(\nu-1)}{2\nu^2}-\frac{1}{2\nu}]]\end{aligned} \tag{43}$$

Yuvraj Sunecher,
University of Technology Mauritius,
Pointe-Aux-Sables, Mauritius.
email: yuvisun@yahoo.com

Naushad Ali Mamode Khan,
University of Mauritius,
Reduit, Mauritius.

Vandna Jowaheer,
University of Mauritius,
Reduit, Mauritius.

The relationship between internet addiction and personality traits in Slovak secondary schools students

M. HOSŤOVECKÝ AND P. PROKOP

Abstract

A new phenomenon, which increasingly rises among the young generation all over the world, is Internet addiction. Information and communication technology (ICT) has a major influence not only on individual sectors of the national economy (services, industry, agriculture, etc.), but also on free time activities and interests of the young generation. ICT-related free time activities of the young generation have notably changed over the last decade (Slovakia was connected to the Internet in November 1992). The aim of the research was to determine the symptoms of Internet addiction in Slovak secondary school students, i.e. to detect significant differences between the following: a personality trait – addiction, addiction – age, addiction – a place of residence, etc. The standardized questionnaire was used in all regions of Slovakia. We collected data by means of an online questionnaire over the period of more than six months. The research results show that female have more potential to become addicted than male, which could be caused by feelings and sadness (Ha & Hwang., 2014). Some other significant results show that more extroverted students are less Internet-addicted than their less extroverted counterparts. According to Zamani, extroverts prefer social and face-to-face interaction with others (Zamani et al., 2011). Other questions focused on neuroticism - emotional stability, where more neurotic students were more addicted than less neurotic students. Finally, significant results were related to residence. Findings showed that students who come from cities (towns) had higher symptoms of Internet addiction than villagers. Other details of this research are shown in the results and discussion section.

Mathematics Subject Classification 2000: 97C30 [**Psychology of and research in mathematics education**]
General Terms: addiction, personality, students, internet.
Additional Key Words and Phrases: Big Five test, Internet addiction, online questionnare.

1. INTRODUCTION

Nowadays, computers and other information technologies play a crucial role in the learning and teaching process. Computers are also an integral part of student's everyday life outside of school and they are used for writing homework or essays, as a part of every school subject, not only at informatics classes; this pattern has been recently documented at almost every Slovak school (Demkanin, 2008). The Internet has long been not only a medium allowing sharing information in text or graphic form or a medium providing "chat" and e-mail based on a text. Today, in connection

10.2478/jamsi-2018-0006

DE GRUYTER OPEN

with the Internet constantly increasing development of technical means (e.g. parameters and data), networks offer a wider range of options for developing interactive communication (Hosťovecký & Zaťková, 2010; Horrigan & Rainie, 2006). The strong integration between social networking and the Internet represents a new Copernican revolution for net users. Most of them choose to use social media platforms for regulating their personal and interpersonal relationships, communications and sociability with peers (Hawi, 2012; Kalkan, 2012; Odaci & Çelik, 2013).

Today's youth spend a lot of time on the Internet, whether it's gaming, watching movies or social networking. As for these activities, there are a number of disadvantages related to using the Internet by high school students. The explosive growth of the Internet in the last decade has had a huge impact on communication and interpersonal behaviour. The Internet was originally designed to facilitate communication and research activities. However, the dramatic increase in the use of the Internet in recent years has led to pathological use - Internet addiction (Takeshi, 2006). Many students even feel out of control and helpless and report serious impairments in their lives as a result of their Internet use (Orzack, 1999; Young, 1996). Internet addiction (also referred to as Internet dependence) occurs when excessive Internet use begins to impact other areas of a person's life. For example, many people have reported experiencing marital problems because of compulsive Internet use (Bocij, 2006). Addictive use of the Internet is a new phenomenon.

Addiction is defined as a primary, chronic disease affecting brain centres, motivation, memory and related circuits. Dysfunction of these circuits leads to characteristic biological, psychological, social and spiritual manifestations. Addiction is characterized by damage to behavioral control and desire, a failure to recognize significant issues and interpersonal relationships, and dysfunctional emotional reactions. Like other chronic diseases, addiction often involves relapse cycles. Without treatment, addiction is progressive and may result in disability or premature death (American Society of Addiction Medicine, 2011). According to the definition in the Mosby's Medical Dictionary, it is a compulsive, uncontrollable addiction to chemical substances, a habit or a practice to such an extent that stopping the activity may cause severe emotional, mental or physiological reactions. The use of the term uncontrollable excludes cases in which an individual successfully tries to control their behaviour (West, 2013). We were examining differences in addiction between sexes, age groups, places of residence (city – the country), regions (Western, Middle, Eastern Slovakia), classes (lower grades 1,2 and upper grades 3,4), extroverts and introverts, more and less agreeable students, more and less conscientious students, more nervous and calmer students.

Before 1997, the number of computer users was very small and this technology was used only by military or large financial centres. Nowadays, however, the number of computer and Internet users increases on a daily basis (Young, 1996). Slovakia was connected to the Internet in November 1992. After 2000, the situation changed. More and more universities and secondary schools were connected to the Internet and the quality of connection improved year after year. Nowadays, after almost 25 years, more than 81% households have an access to the Internet in Slovakia (over 4.4 million inhabitants, Eurostat 2016). Computers have thus also become a sort of a companion of today's youth. The young generation uses computer in their everyday life. One of the areas affected by the influence of computers is free time. Much research (e.g., Vrabec, 2009; Hirzalla & Zoonen, 2008; Uyenco & Kingdon, 2010) has confirmed that children at a younger and older school age, such as adolescents, spend an average of 1.5 – 3 hours on the computer.

The past decade has seen a tremendous increase in internet use and computer-mediated communication (Fox et al., 2001). Especially the young generation carries out a lot of work with the help of a computer. The society expects pupils already in the first grade of primary school to have at least partially developed skills in using some of the "new" media, e.g. computer, smart phone, tablet. Computer and the Internet have become an indispensable passion for them. The young generation uses computers and the Internet more successfully and faster than previous generations (Tosun & Baris, 2011; Mišút & Pribilová, 2013). It creates new ways for citizens to communicate, congregate and share information of a social nature (Kedem, 1999). The Internet provides an array of tools for people to use for information retrieval and communication in individual, group and mass contexts, but can current notions of media be used to define communication on the Internet? They can no longer imagine their lives without a computer and internet connection. Computers help students in their studies and in communication with their classmates, teachers, friends and family, but have also become one of the main free time activities. Nowadays, television, radio, but also books and magazines, sport, personal contact and conversations are to a large extent substituted by computers. The well-known phrase “Everything’s on the Web” is slowly becoming reality.

Another very important factor of using the Internet is user’s personality. Personality has a psychological impact on how subjects interact with information technologies. According to Amichai-Hamburger, these studies are decisive because personality traits are relevant factors in determining subjects’ behaviour when using Internet technologies (Amichai -Hamburger, 2002).

Personality not only defines the subjects’ behavioural style but also represents relatively enduring characteristics of subjects, and it refers to all aspects of individuality. Human activities and types of behaviour are consistent with specific traits of personality.

There are three ingredients required for the initiation of scientific research on traits: systematic data collection, statistical techniques for data analysis, and development of testable theories. These prerequisites became available around the beginning of the 20th century. Of key importance were the new techniques of correlation and, somewhat later, factor analysis (Matthews et al., 2009; Gorsuch, 1983).

One way to study the relationship between subjects’ personality traits and the use of the Internet is to apply specific conceptual frameworks, such as the Big Five factor model (Montag, Jurkiewicz, & Reutera, 2010). Following the advances in research, previous investigations established a strong connection between personality traits and Internet addiction (Chang & Law, 2008; Correa, Hinsley, & de Zúñiga, 2010; Landers & Lounsbury, 2006; Rice & Markey, 2009). Moreover, personality traits, parenting and familial influence, alcohol use, and social anxiety are considered to be predictive factors of Internet addiction disorder (Ko et al., 2012). The exponential use of the Internet has stimulated debate on examining how personality traits impact on the use of technologies, particularly the Internet, social networking, virtual environments, online and offline games, and so on. Thus, rather than investigate only the relationships between Internet overuse and subjects' behaviour, current research work is interested in exploring the influence of personality traits on Internet addiction risks (Buckner et al., 2012; Landers & Lounsbury, 2006; Ryan & Xenos, 2011).

When the effect of gender was taken into consideration, males had significantly higher PIU scores than females (Ozcan & Buzlu, 2007; Tsai et al., 2009; Young, 1996). Internet use also has an impact on the life of school-age young people. Young and Rogers state that difficulties concerning studying, for instance a drop in grades and changing sleep patterns can occur when school-age youngsters spend too much time on the web (Young & Roger, 1998).

Further research using various samples, such as elderly, adults, and adolescents, is necessary to evaluate the divergent validity of the findings. In addition, the effects of some other variables on PIU should be taken into consideration in future studies, such as personality variables neuroticism (Tsai et al., 2009), self-esteem (Yang & Tung, 2007; Douglas et al., 2008; Ko et al., 2007), and conscientiousness (Durak & Senol-Durak, 2010), life satisfaction (Stepanikova, Nie, & He, 2010) and subjective well-being (Kraut et al., 1998). The results show that the five dimensions of personality add significantly to the variance explained by gender and Facebook experience concerning time spent on Facebook (De Cock et al., 2014).

Besides developed countries, one of the countries where addiction is starting to show is China. According to Ferraro et al.: "Internet addiction is a problem not only for adolescents in China; it also presents headaches for parents and teachers all over the world" (Ferraro et al., 2007). Other research realized in Greece, Siomos and his colleagues (Siomos et al., 2008) surveyed 2,200 Greek adolescent students and showed that 8.2% of them were addicted to the Internet, mainly male students who play online games and visit Internet cafés. In Iran, a study conducted in Iranian high schools revealed that, among 1,968 high school students, 977 students were Internet users: 37 were classified as Internet addicts, 304 as possible Internet addicts, and 636 as moderate users. Researchers found that Internet addicts were lonelier than moderate users and had lower self-esteem and poorer social skills (Ghassemzadeh et al., 2008). Because accessibility to computers in Slovakia was, at least according to past report, much lower compared with western countries (Fančovičová & Prokop, 2008) and no reports concerning IAD in Slovakia exist, we investigated the prevalence of IAD among Slovak high school students.

Griffiths proposed the concept of „technological addiction“, which is nonchemical but behavioral bin nature and involves excessive human-machine interaction. Technological addiction can be either passive, such as viewing television, or active, such as gaming on the computer or chatting online usually comprises inducing and reinforcing features that may contribute to the promotion of addictive tendencies (Griffiths 1996). Internet addiction disorder (IAD) was first proposed by Ivan Goldberg in 1995. Derived from the substance-dependence criteria of the Diagnostic and Statistical Manual of Mental Disorders (4th Ed. DSM-IV American Psychiatric Association, 1994), IAD is the first listed Internet-related disorded. It is defined as a behavioral addiction consisting of six core components: salience, mood modification, tolerance, withdrawal symptoms, conflict, and relapse (Griffiths, 1998). Internet Addiction Disorder (IAD) ruins lives by causing neurological complications, psychological disturbances, and social problems.

2. METHODS

- ***Purpose of study***

The aim of our research was to determine Internet addiction in students of secondary professional schools and, at the same time, to find out how student’s character and personality affects the time spent on the computer in relation to the risk of development of addiction.

These hypotheses were evaluated:

- Males will show a higher rate of Internet addiction than females.
- Students from lower grades will be more addicted than students from upper grades.

- Students from cities will show a statistically significant higher rate of Internet addiction than students coming from the village.
- Extroverted students will be less addicted to Internet than introverted students.
- More agreeable students will be less addicted to Internet than less agreeable students.
- More conscientious students will show a lower rate of Internet addiction than less conscientious ones.
- More neurotic students will be more addicted than less neurotic students.
- Students who are more open in their behaviour will be less addicted to Internet than students who are more closed.

- ***Research instrument***

An anonymous questionnaire was used in the research that we carried out to determine Internet addiction in secondary school students. It was divided into three sections.

In the first section of the questionnaire, a short introductory text about the intent of the questionnaire was presented followed by demographic data about a student: age, sex, place of residence (the village/city), region, secondary school's specialisation, student's current grade.

In the second section, the Internet Addiction Test (hereinafter referred to as IAT) by Kimberly Young (Year) was used. It is the first ratified and reliable method of measurement of addictive Internet use. This questionnaire was translated from English into Slovak and contained 18 items investigating questions regarding Internet addiction. Questions were translated freely; however, the emphasis was on maintaining the meaning and structure of the questionnaire. Each item in the questionnaire was evaluated by a 5-point Likert scale using the following description: "strongly agree" (5 points), "slightly agree" (4 points), "neutral" (3 points), "slightly disagree" (2 points), "strongly disagree" (1 point). For determining Internet addiction, we used a differentiation scale of addicted and non-addicted users by K. Young.

In the third section, a questionnaire for assessing personality traits of students (10-item short version of the Big Five Inventory) was used in comparison with their access to the Internet by Rammstedt & John. The Big Five Inventory is one of the most widespread personality models. Thus, basic factors or personality traits according to the Big Five Model are neuroticism, openness to experience, extraversion, agreeableness and conscientiousness. As a part of BFI, a 5-item scale of responses from “strongly disagree” (1) to “strongly agree” (5) was designed. Individual’s personality may be found at a various level of the scale. The Big Five Model is commonly used in various cultural areas as well as some areas of psychological research (Rammstedt & John, 2007).

The present study was conducted online, which affords a high degree of anonymity, which elicits more candid responses to questions about socially undesirable behaviour and emotions than do paper and pencil methods or interview methods (Locke and Gilbert, 1995; Musch et al., 2001). In addition, internet surveys are a convenient, user-friendly, comfortable, and secure data gathering method (e.g., Campos et al., 2011; Lewis, et al., 2009; van Gelder et al., 2010).

- ***Participants***

Participants were 707 high school students (340 males and 367 females) with the mean age of 16.9 years (SE = 0.06, range: 15 – 21 years). The research sample consisted of secondary school students from all regions of Slovakia, i.e. the Banská Bystrica, Bratislava, Košice, Nitra, Prešov, Trenčín, Trnava and Žilina region. Students participated in the research on a voluntary basis. Respondents were addressed by the school’s management which received an accompanying letter clarifying the research and its purpose as well as the internet link to an electronic questionnaire. The alternative way of addressing them was a social network Facebook. We contacted all secondary professional schools in all regions of Slovakia, i.e. 448 secondary schools, by means of sending an e-mail and via Facebook. Response rate was 5.5% which represents 25 secondary schools in Slovakia.

Secondary professional schools were divided by their specialisation following the International Standard Classification of Education (hereinafter referred to as ISCED) by UNESCO. The reason was the fact that this instrument serves as a means of comparing, compiling and presenting statistical data on education. According to this categorization, secondary professional schools belong to category ISCED 3 – Upper secondary education (secondary schools with a school-leaving exam). Representation of the number of respondents according to ISCED 3 specialisation who participated in our research is shown below:

- Social sciences and humanities (music, dramatic arts, history...) 10 % (71 students);
- Social sciences, business, law (economics, journalism, tourism...) 15 % (108 students);
- Natural sciences, mathematics, computer sciences (geodesy, biology, chemistry...) 7% (47 students);
- Engineering, production, construction (mechanics, electronics, architecture...) 32% (228 students);
- Agriculture and veterinary medicine (food, horticulture...) 2% (15 students);
- Health and health care (social work, pharmaceutics...) 12% (84 students);
- Services (hotel, security, hairdressing and beauty treatment services...) 22% (154 students).

About half of participants (56%) reported to live in villages and remaining 46% reported to live in cities. Pursuant to the Act of the National Council of the Slovak Republic No. 369/1990 on Municipal Establishment, Slovakia is divided into 8 large regions and 79 smaller counties. Each of these 8 so-called Higher-Tier or Upper-Tier Territorial Units (VÚC – Vyššie Územné Celky) are named after the largest city in that region. Each of the counties has some cities and villages. There is not a classification in Slovakia which determines the difference between a city and a town. We can define a city and a village as follows.

A Village is a clustered human settlement which is smaller than a city. Today, there are 2,933 villages in all regions of Slovakia. About 45% of Slovaks live in villages with less than 5,000 inhabitants, and 14% in villages with less than 1,000.

A city is a human settlement larger than a village. Nowadays, there are 140 cities in Slovakia (2017). These requirements are determined by the Act of the National Council of the Slovak Republic called "the Act on Municipal Establishment". A city can be considered every community or a place of living which meets the requirements listed below:

- an economic, administrative and cultural center or a tourist center, or a spa,
- provides services for inhabitants of surrounding villages,
- has a secured transport connection with surrounding villages,
- has an urban building character at least in some parts of its area,
- has at least 5,000 inhabitants (§ 22, Act No. 369/1990).

- ***Statistical analyses***

Multiple regression was applied to investigate associations between Big Five personality traits (agreeableness, extraversion, conscientiousness, neuroticism and openness), residence (village/city), the region of student's origin – the part of Slovakia (western, central or eastern region), secondary school focus and grade (independent variables) and internet dependence (dependent variable). Continuous predictors were initially checked for normality with Kolmogorov-Smirnov test. Only mean scores of the Internet Dependence scale differed from normal distribution, thus Box-Cox (x+1) transformation was applied in order to achieve normality. All VIF (variance inflation factor) values were < 2 suggesting that there was no collinearity between variables in our sample (Allison, 1999).

3. RESULTS

Multiple regression resulted in significant model (R2 = 0.12, F (10,696) = 9.28, p < 0.001) that explained 12% of total variance of the results. Detailed report is shown in Table 1. Participants who scored high in Extraversion and Conscientiousness domains showed lower Internet Dependence score. In contrast, Neuroticism showed positive associations with Internet Dependence score. Agreeableness and Openness were not associated with the Internet Dependence score. Perhaps surprisingly, females showed higher Internet Dependence score than males and participants from cities scored higher than participants from villages. Other correlations were not statistically significant (Table 1).

Table 1. Multiple regression on Internet Dependence score. Significant associations are marked **bold**.

	β	SE of β	B	Se of B	t(696)	P
Intercept			5.556	1.245	4.462	**<0.001**
Extraversion	-0.117	0.037	-0.007	0.002	-3.143	**0.001**
Agreeableness	-0.050	0.036	-0.003	0.002	-1.398	0.163
Conscientiousness	-0.152	0.038	-0.010	0.002	-4.030	**<0.001**
Neuroticism	0.132	0.037	0.008	0.002	3.554	**<0.001**
Openness	0.004	0.036	0.000	0.003	0.113	0.910
Gender	-0.119	0.037	-0.026	0.008	-3.187	**0.001**
Residence	-0.091	0.036	-0.020	0.008	-2.499	**0.01**
Region	0.052	0.036	0.003	0.002	1.434	0.152
Type of secondary school	-0.053	0.036	-0.004	0.003	-1.446	0.149
Grade	-0.049	0.036	-0.005	0.004	-1.335	0.182

4. DISCUSSION

The present study offers findings that the age of Internet-addicted Slovak adolescents ranges from 15 to 21. This is the first study which compares the Internet addiction and Big Five personality factors of secondary school students in Slovakia. It gives us understanding of behavioural addiction and its personality correlates. Therefore, the findings support some of our above mentioned hypotheses. A representative web-based online survey was used to examine the relationship between personal traits and Internet addiction among 707 Slovak adolescents. Our findings are in accordance with previous researches carried out in some other regions of Europe, Asia and the United States.

The exponential use of the Internet has stimulated debate on examining how personality traits impact the use of technologies, particularly the Internet, social networking, virtual environments, online and offline games, and so on. Thus, rather than investigate only the relationships between Internet overuse and subjects' behaviour, current research work is interested in exploring the influence of personality traits on Internet addiction risks (Buckner et al., 2012; Landers & Lounsbury, 2006; Ryan & Xenos, 2011). Internet use also has an impact on the life of school-age young people. Young and Rogers state that difficulties concerning studying, for instance a drop in grades and changing sleep patterns can occur when school-age youngsters spend too much time on the web (Young and Rogers, 1998).

Gender

The present study investigated whether males are more Internet addicted than females. It was surprising to find that females obtained higher score than males in web self-assessment questionnaire focused on Internet addiction. In line with Tsai et al. (Tsai & Lin, 2004), we confirm that findings may contradict the gender stereotype in the use of technology. We can see similar results in the increase of female scores compared to males in some technological areas, e.g. mobile phone addiction. There are some findings that mobile phone addiction was associated with female gender. This finding is in line with previous studies (Andreassen et al., 2013; Augner & Hacker, 2012).

This result could be caused by feelings such as sadness, which is in line with Ha (Ha & Hwang, 2014) who states that when males and females are stressed, females are encouraged to express feelings such as sadness and to use more emotional strategies, whereas males are discouraged from expressing their feelings, and they tend to adopt distracting or aggressive strategies. Gender differences in preferred online activities may contribute to a higher risk of Internet addiction in females who are depressed (Ha & Hwang, 2014).

Extroversion & Neuroticism

Results showed that more *extroverted participants* are less Internet-addicted than less extroverted participants. It appears that students with high extraversion scores prefer social and face-to-face interaction with others to interaction with the virtual world. On the contrary, more introverted students avoid contact with other people because of their shyness, so they communicate with the virtual world more (Zamani et al., 2011). Introverted (or less extraverted) individuals fail to cope with emotional and physical stress effectively, have difficulty in establishing social relationships and, thus, are relatively more susceptible to addiction. Some previous researches, including the study of Batıgün (Batıgün & Kilic, 2011) and Floros (Floros & Siomos, 2014), indicate that people with low levels of extraversion use the Internet more frequently and may be at a higher risk of internet addiction. Our findings confirm some previous studies which investigate extroversion (Andreassen et al., 2013; Chou et al., 2005).

Neuroticism or emotional stability is one of the five components of Big Five personality model. In our study, we predicted that more neurotic students will be more addicted than less neurotic students. Our findings have confirmed our hypothesis: Slovak secondary school students who are more neurotic are more Internet-addictive than their less neurotic counterparts. Our findings are in line with and confirm some other research in this field from different regions from all over the world. Kunimura and Thomas assessed 113 students at Loyola Marymount University and found a significant positive relation between neurosis and Internet dependency as well as a significant negative relation between extroversion and

Internet dependency (Kunimura & Thomas, 2000; Zamani et al., 2011). Results of another study confirm that neuroticism was positively associated with internet addiction (Saini et al., 2016)

Residence

The main findings showed that students coming from cities (towns) had higher symptoms of Internet addiction than villagers. It could be caused by the fact that villagers spend more time in the garden (gardening, field work, work in orchards, etc.), in nature (walking, tourism) or sporting (cycling, natural swimming, mountaineering, etc.). It is, in particular, very typical of some parts of our country (the northern, southern and eastern part of Slovakia) that young people are helpful to parents. Our findings are in accordance with previous studies (Lićwinko, Krajewska-Kułak & Łukaszuk, 2011) showing similar results: students living in the city had more problems with Internet addiction than students from the villages.

Limitations

The present study has some limitations, which must be addressed. Our sample size was 707 participants (340 males and 367 females). We wanted to obtain a more representative sample size from schools in Slovakia. Also, even though the present study sample size provided acceptable statistical power (Cohen, 1988), it was still quite small. Thus, cross-validation of the current study with larger samples should be carried out. The instruments used in the present study were standardized. Other studies in this field should therefore use longitudinal designs in order to better assess the directionality between the concepts. In addition, future studies should use larger and more representative samples in terms of gender, age, etc. (Andreassen, Griffiths, Gjertsen, Krossbakken,,Kvam, Pallesen, 2013).

Further research using various samples, such as elderly, adults, and adolescents, is necessary to evaluate the divergent validity of the findings. In addition, the effects of some other variables on PIU should be taken into consideration in future studies, such as personality variables (neuroticism (Tsai et al., 2009), self-esteem (Yang & Tung, 2007; Douglas et al., 2008; Ko et al., 2007), and conscientiousness (Durak &

Senol-Durak, 2010), life satisfaction (Stepanikova, Nie, & He, 2010) and subjective well-being (Kraut et al., 1998).

In addition, the validity of some of the participants' answers may be questionable and motivational factors of participants responding to this study should be discussed as well. An online questionnaire without personal involvement and without experts in the topic of Internet addiction was used. Therefore, some answers might reflect different level of scores. Also, all data in the present study was based on self-report, so the results may have been influenced by the common method bias (Podsakoff et al., 2003).

5. CONCLUSION

Nowadays, Internet addiction disorder (AID) belongs to disorders that not only students, but anyone else can have. The aim of this research was to find out whether secondary school students from secondary schools in Slovakia show potential signs of Internet addiction. We were examining addiction between sexes, older and younger students, types of specialization of secondary professional schools, and personality characters (introvert-extrovert). Our results confirm that:

- female scored higher than male in the self-assessment questionnaire;
- more extroverted participants had higher symptoms of Internet addiction than less extroverted participants;
- more neurotic participants had higher symptoms of Internet addiction than less neurotic participants;
- less conscious students had higher symptoms of Internet addiction than more conscious students;
- respondents living in the cities were more Internet-addicted than respondents living in villages.

In another prospective future study, it could be interesting to see the relationship and the comparison between Internet addiction and personality traits of Slovak university students in different fields of study: science, medicine, engineering, agriculture, economy, law and the relationship between private and public university students.

ACKNOWLEDGMENTS

Many thanks to students who participated in the research. Many thanks to Mr. Halás for improve the English language.

REFERENCES

ALLISON, P. D. 1999. Multiple Regression: a primer. Pine Forge Press: Thousand Oaks.

AMERICAN PSYCHIATRIC ASSOCIATION. 1994. DSM-IV. 4th Ed.

AMERICAN SOCIETY OF ADDICTION MEDICINE. 2011. Definition of Addiction, http://www.asam.org for-the-public/definition-of-addiction

AMICHAI–HAMBURGER, Y. 2002. Internet and personality. Computers in Human Behavior, 18, 1–10.

ANDREASSEN, C.S., GRIFFITHS, M.D., GJERTSEN, S.R., KROSSBAKKEN, E., KVAM, S., PALLESEN, S. (2013). The relationships between behavioral addictions and the five-factor model of personality. *Journal of Behavioral Addictions,* 2(2), 90 – 99.

AUGNER, C., HACKER, G. W. 2012. Associations between problematic mobile phone use and psychological parameters in young adults. In: *International Journal of Public Health*, 57, 437 – 441.

BATIGÜN, A.D., KILIC, N. 2011. The relationships between internet addiction, social support, psychological symptoms and some socio-demographical variables. *Türk Psikol Derg,* 26 (67), 1 – 10.

BOCIJ, P. 2006. The dark side of the internet: protecting yourself and your family from online criminals. ISBN: 0-275-98757-X

CAMPOS, J. A., ZUCOLOTO, M. L., BONAFE, F. S., JORDANI, P. C., MAROCO, J. 2011. Reliability and validity of self-reported burnout in college students: A cross randomized comparison of paper-and-pencil vs. online administration. In: *Computers in Human Behavior,* 27, 1875–1883.

COHEN, J. 1988. Statistical power analysis for the behavioral sciences (2nd ed.). New Jersey: Lawrence Erlbaum.

CORREA, T., HINSLEY, A. W., DE ZÚÑIGA, H. G. 2010. Who interacts on the Web? The intersection of users' personality and social media use. In: *Computers in Human Behavior*, 26(2), 247-253. DOI: 10.1016/j.chb.2009.09.003

CHANG, M.K., LAW, S.P.M. 2008. Factor Structure for Young's Internet Addiction Test: A confimatory study. *Computers in Human Behavior*. 24(6), 2597–2619.

ODACI, H., ÇELIK, Ç. B. 2013. Who are problematic internet users? An investigation of the correlations between problematic internet use and shyness, loneliness, narcissism, aggression and self-perception. Computers in Human Behavior 29(6), 2382–2387.

CHOU C., CONDRON L., BELLAND J.C.. 2005. A review of the research on Internet addiction. Educational Psychology Review, 17(4), 363–88.

De COCK, R., VANGEEL, J., KLEIN, A., MINOTTE, P., ROSAS, O., MEERKERK, G.J. 2014. Compulsive use of social networking sites in Belgium: prevalence, profile, and the role of attitude toward work and school. *Cyberpsychology, Behavior and Social Networking*, 17(3), 166–171. http://dx.doi.org/10.1089/cyber.2013.0029.

DEMKANIN, P. 2008. One small, but significant step in science education in Slovakia. In P. Demkanin, B. Kibble, J. Lavonen, J. G. Mas, & J. Turlo Eds., Effective use of ICT in science education. 120–140. Edinburgh: University of Edinburgh, School of Education.

DOUGLAS A.C., MILLS J.E., NIANG M., STEPCHENKOVA S., BYUN S., RUFF;INI C., et al. 2008. Internet addiction: meta-synthesis of qualitative research for the decade 1996-2006. *Computers in Human Behavior*. 24(6), 302–744.

DURAK, M., SENOL-DURAK, E. 2010. Which personality traits are associated with cognitions related to problematic Internet use? Manuscript under evaluation.

FANČOVIČOVÁ, J., PROKOP, P. 2008. Students' attitudes toward computer use in Slovakia. *Eurasia Journal of Mathematics, Science & Technology Education*, 4(3), 255 – 262.

FERRARO, G., CACI, B., D'AMICO, A., BLASI, M. D. 2007. Internet addiction disorder: An Italian study. In: *CyberPsychology & Behavior,* 10(2), 170–175.

FLOROS, G., SIOMOS, K. 2014. Excessive internet use and personality traits. In: *Curr Behav Neurosci Reports*, 1, 19–26.

FOX, S., RAINIE, L., LARSEN, E., HORRIGAN, J., LENHART, A., SPOONER, T., CARTER, C. 2001. Wired Seniors. *The Pew Internet and American Life Project. Retrieved May 4th 2005 from:* http://www.pewinternet.org/pdfs/PIP_Wired_Seniors_Report.pdf

GHASSEMZADEH, L., SHAHRARAY, M., MORADI, A. 2008. Prevalence of Internet addiction and comparison of Internet addicts and non-addicts in Iranian high schools. In: *CyberPsychology* & *Behavior*, 11(6), 731–733.

GORSUCH, R. L., 1983. Factor Analysis, second edition, Hillsdale: Lawrence Erlbaum Associates.

GRIFFITHS, M. D. 1996. Gambling on the Internet: A brief note. In: *Journal of Gambling Studies*, 12, 471–473.

GRIFFITHS, M.D., HUNT, N. 1998. Dependence on computer games by adolescents. In: *Psychology Reports*, 82, 475–480.

HA, Y. M., HWANG, W. J. 2014. Gender Differences in Internet Addiction Associated with Psychological Health Indicators Among Adolescents Using a National Web-based Survey. In: *International Journal of Mental Health and Addiction,* 12 (5), 660. https://doi.org/10.1007/s11469-014-9500-7

HAWI, N. S. 2012. Internet addiction among adolescents in Lebanon. In: *Computers in Human Behavior* 28(3): 1044–1053.

HIRZALLA, F., VAN ZOONEN, L. 2008. Usesofthe Web for Civic Participation: výskumná správa. London: Institute of Education. 140 pages. http://www.civicweb.eu/images/stories/reports/wp9civicwebdeliverable8.pdf

HORRIGAN, J., RAINIE, L. 2006. The internet's growing in life's major moments. Retrieved August 30, http://www/pewinternet.org/PPF/r/181/repo rtdisplay.asp

HOSŤOVECKÝ, M. ZAŤKOVÁ, T. 2010. Options for implementing videoconferencing system in a combined form of education. In: *Acta Universitatis Facultatis Paedagogicae Tyrnaviensis.* Ser. C – Trnava, 157-164.

KALKAN, M. 2012. Predictiveness of interpersonal cognitive distortions on university students' problematic internet use. In: *Children and Youth Services Review* 3(7), 1305–1308.

KEDEM, C. 1999. The social impact of the internet on our society. Retrieved August 30, 2006, http://www.math.umd.edu/~bnk/CAR/ project.htm.

KO, CH., YEN, J.Y., YEN, C.F., LIN, H.C., YANG, M.J. 2007. Factors predictive for incidence and remission of Internet addiction in young adolescents: A prospective study. In: *CyberPsychology & Behavior*. 10, 545-551.

KO, CH., YEN, J.Y., YEN, C.F., CHEN, C.S., CHEN, C.C. 2012. The association between Internet addiction and psychiatric disorder: a review of the literature. In: *Eur Psychiatry*. 27(1), 1-8. https://doi.org/10.1016/j.eurpsy.2010.04.011

KRAUT, R., PATTERSON, M., LUNDMARK, V., KIESLER, S., MUKOPHADHYAY, T., SCHERLIS, W. 1998. Internet paradox: A social technology that reduces social involvement and psychological well-being? In: *American Psychologist*, 53(9), 1017-1031.

KUNIMURA, M., THOMAS, V. 2000. Summary and review of the NEO-PI-R personality test. *Journal of LoyolaMary mount University*. 2000;38(4):1–13

LANDERS, R. N., LOUNSBURY, J. W. 2006. An investigation of Big Five and narrow personality traits in relation to Internet usage. In: *Computers in Human Behavior*, 22, 283–293. doi:10.1016/j.chb.2004.06.001.

LEWIS, I., WATSON, B., WHITE, K. M. 2009. Internet versus paper-and-pencil survey methods in psychological experiments: Equivalence testing of participant responses to health-related messages. In: *Australian Journal of Psychology*, 61, 107–116.

LIĆWINKO, J., KRAJEWSKA-KUŁAK, E., & ŁUKASZUK, C. 2011. Internet addiction among academic youth in Białystok. Progress in Health Sciences, 1(1), 124-130.

LOCKE, S. D., GILBERT, B. O. 1995. Method of psychological assessment, self-disclosure, and experiential differences: A study of computer, questionnaire, and interview assessment formats. In: *Journal of Social Behavior and Personality*, 10, 255–263.

MATTHEWS, G., DEARY, I. J., WHITEMAN, M. C. 2009. Personality traits (3rd ed.) Cambridge: Cambridge University Press. In Internet addiction in adolescents: Prevalence and risk factors. In: *Computer in Human Behavior.*

MIŠÚT, M., PRIBILOVÁ, K. 2013: Communication impact on project oriented teaching in technology supported education. In *Lecture Notes in Electrical Engineering*, 152 LNEE, 559-567.

MONTAG, C., JURKIEWICZ, M., REUTER, M. 2010. Low self-directedness is a better predictor for problematic internet use than high neuroticism. In: *Computers in Human Behavior*, 26(6), 1531–1535. doi:10.1016/j.chb.2010.05.021

MUSCH, J., BRODER, A., KLAUER, K. C. 2001. Improving survey research on the world-wide web using the randomized response technique. In U-D. Reips and M. Bosnjak (Ed.), Dimensions of internet science (179–192). Lengerich: Pabst Science Publishers.

OZCAN, N.K. BUZLU, S. 2007. Internet use and its relation with the psychosocial situation for a sample of university students. In: *Cyberpsychology and Behavior*, 10(6), 767–772.

ORZACK, M. 1999. Computer addiction: Is it real or virtual? Harvard Mental Health Letter, 15(7), 8.

PODSAKOFF, P. M., MACKENZIE, S., Lee, J.Y., Podsakoff, P. N. 2003. Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. In: *The Journal of applied psychology*, 88(5). 879–903. DOI: 10.1037/0021-9010.88.5.879

VAN GELDER, M. , BRETVELD, R. W., ROELEVELD, N. 2010. Web-based questionnaires: The future in epidemiology? Internet versus paper-and-pencil survey methods in psychological experiments: Equivalence testing of participant responses to health-related messages. In: *American Journal of Epidemiology*, 172, 1292–1298.

RAMMSTEDT, B., JOHN, O. P. 2007. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. In: *Journal of Research in Personality*. 41, 203-212. ISSN: 0092-6566. http://www.sciencedirect.com/science/article/pii/S0092656606000195

RICE, L., & MARKEY, P. M. 2009. The role of extraversion and neuroticism in influencing anxiety following computer-mediated interactions. In: *Personality and Individual Differences, 46* (1), 35-39. http://dx.doi.org/10.1016/j.paid.2008. 08.022

RYAN, T., XENOS, S. 2011. 'Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage', In: *Computers in Human Behavior*, 27 (5), 1658-1664.

SAINI, V. K., BANIYA, G. CH., VERMA, K. K., SONI, A., KESHAWANI, S. 2016. A study on relationship of internet addictive behavior with personality traits among medical students. In: *Journal of Mental Health and Human Behaviour*, 21(2), 108–111.

SIOMOS, K. E., DAFOULI, E. D., BRAIMIOTIS, D. A., MOUZAS, O. D., & ANGELOPOULOS, N. V. 2008. Internet addiction among Greek adolescent students. In: *CyberPsychology & Behavior*, 11 (6), 653–657.

SLOVAK REPUBLIC. 2017. Law no. 369/1990, § 22 "The Law about General Settlement" http://www.zakonypreludi.sk/zz/1990-369

STEPANIKOVA, I., Nie, N. H., & He, X. 2010. Time on the Internet at home, loneliness, and life satisfaction: Evidence from panel time-diary data. *Computers in Human Behavior*, 26(3), 329-338.

TAKESHI, S. 2006. Internet Addiction among Students: Prelevance and psychological problems in Japan. In *JMAJ* 49 (7-8). 279-283. http://www.med.or.jp/english/pdf/2006_07%2B/279_283.pdf

VRABEC, N. 2009. On-line generácia: informácie, komunikácia a digitálna participácia mládeže v informačnej spoločnosti: výskumná správa. Bratislava: MŠ SR a IUVENTA. 59 s. http://www.iuventa.sk/files/documents/7_vyskummladeze/spravy/davm027/on_line_generacia_publikacia.pdf

WEST, R. 2013. Models of addiction. In EMCDDA Insights [online]. No. 14, 161p. ISSN 2314-9264. http://www.emcdda.europa.eu/ publications/insights/models-addiction

UYENCO, B., KINGDON, A. 2010. Young Adults Revealed: The lives and motivations of 21st century youth, advertising.microsoft.com/ wwdocs/.../Young-Adults-Revealed.pdf,

YOUNG, K. S. 1996. Internet addiction: The emergence of a new clinical disorder. In: *Cyber Psychology and Behavior*. 237-244. ISSN 1094-9313. http://www.chabad4israel.org/tznius4israel/newdisorder.pdf

YOUNG, K. 1998. Internet addiction: The emergence of a new clinical disorder. In: *CyberPsychology and Behavior*, 1, 237-244.

YOUNG, K., ROGERS, R. C. 1998. The Relationship between depression and Internet Addiction. In: *CyberPsychology & Behavior*, 1(1): 25-28. https://doi.org/10.1089/cpb.1998. 1.25

YANG, S., & TUNG, C. 2007. Comparison of Internet addicts and non-addicts in Taiwanese high school. In: *Computers in Human Behavior*, 23, 79–96.

TOSUN, N. – BARIS, M. F. 2011. The Place and Importance of Computer and Internet's In Secondary School Students' Life. In: *Procedia Social and Behavioral Sciences* 28, 530 – 535.

TSAI, C. C., & LIN, C.C. 2004. Taiwanese adolescents' perceptions and attitudes regarding the Internet: Exploring gender differences. In: *Adolescence*, 39, 725-734.

TSAI, H. F., CHENG, S. H., YEH, T. L., SHIH, C. C., CHEN, K. C., YANG, Y. C., YANG, Y. K. 2009. The risk factors of internet addiction—A survey of university freshmen. In: *Psychiatry Research*, 167, 294-299.

ZAMANI, B. E., ABEDINI, Y., KHERADMAND, A. 2011. Internet Addiction Based on Personality Characteristics of High School Students in Kerman, Iran. In: *Addict Health*. Summer-Autumn; 3(3-4), 85–91.

Marián Hosťovecký,
Department of Applied Informatics and Mathematics,
University of Ss. Cyril and Methodius, Slovakia, 917 01
Email: marian.hostovecky@ucm.sk

Pavol Prokop,
Department of Biology, Faculty of Education,
Trnava University in Trnava, Slovakia, 917 01
Email: pavol.prokop@truni.sk

Received April 2018

Performance improvement of using lambda expressions with new features of Java 8 vs. other possible variants of iterating over ArrayList in Java

J. JURINOVÁ

Abstract

A fluent programming is the programming technique where operations return a value that allows the invocation of another operation. With the fluent programming, it is perfectly natural to end up with one huge statement that is the concatenation of as many operations as you like. The Java Development Kit (JDK) streams (added in Java 8) are designed to support fluent programming. Instead of looping over all elements in the sequence repeatedly (once for filter, then again for map, and eventually for toArray), the chain of filtermapper-collector can be applied to each element in just one pass over the sequence. In this context, we often encounter lambda expressions used to create locally defined anonymous functions. They provide a clear and concise way to represent one method interface using an expression. Oracle claims that use of lambda expressions also improve the collection libraries making it easier to iterate through, filter, and extract data from a collection. In addition, new concurrency features improve performance in multicore environments.
There are multiple ways to traverse, iterate, or loop collection in Java. Therefore, to solve one problem, we have several options for solutions that differ by undeniably increasing of the code readability. Searching for answers to the question of whether these new features really bring performance benefits over conventional way, is the subject of this paper.

1. INTRODUCTION

Iteration is the process of traversing a sequence. It can be performed in two ways: as an internal or external iteration. The external iteration uses an iterator for access to all sequence elements. The internal iteration is performed by the sequence itself; the user just supplies an operation to be applied to all sequence elements. Traditionally, Java collections offer external iteration. However, Java 8 provides excellent features to support iterating, filtering and extracting of elements in Java collections. Prior to Java 8, better (concise) way to iterate elements is by using *for each loop* (added in Java 5) or iterating over collection using *iterator* and selecting the required object. Though that approach works, it was very difficult to run them in parallel and take advantage of all the available cores in the processor. Many of the techniques that

10.2478/jamsi-2018-0007

programmers have relied on in the past are now being replaced by better, more powerful constructs.
Lambda expressions are one of the most important features added in Java 8. They are such important since they add functional programming features to Java. In addition, a lambda expression addresses the bulkiness of anonymous inner classes by converting five or even more lines of code into a single statement. According Schildt (2014), the use of lambda expressions in Java “can simplify and reduce the amount of source code needed to create certain constructs, such as some types of anonymous classes.” The conclusions of this paper demonstrate compliance with this statement. For example, this is particularly helpful when implementing many commonly used event handlers. Lambdas also make it easy to pass what is a piece of executable code as an argument to a method. Before introducing lambda expressions, programmers were forced into writing more words in code than needed. Now, we can pass in a function as a parameter.

2. DEFINITION OF OUR PROBLEM

The most common operation with collection classes are iterating over them and applying business logic on each element. In other words, anytime you have a collection of things you will need some mechanism to systematically step through the items in that collection. Suppose that you are creating an application and you want to create a feature that enables to perform any kind of action on members of the application that satisfy certain criteria. Frequency of the occurrence of this action is many times during the day. Therefore, the performance of implemented approaches would be very important for us. Our motivation for this work came from the diverse argument that the use of lambda expressions would result in more efficient code. According Angelika Langer (2015) the motivation of the designers of the Java programming language “for inventing streams for Java was performance or — more precisely — making parallelism more accessible to software developers.” Another reason for our interest in this issue is an ongoing conflict between the younger and older generation of programmers, when many of the older programmers often have a distrust of new approaches or a strong belief in the unnecessary change in validated and routine procedures and approaches. The question we are trying to answer is the following: *"Are lambda expressions with connection to the new features in Java 8 really faster than non-lambda expressions at accomplishing the same tasks?"* Therefore, our main goal is to determine whether the introduced lambda expressions have a speed advantage over non-lambda expressions for an identical task by getting a quantitative speed difference in nanoseconds. At the same time, we want to look at the presented approaches from a more complex point of view and together with an illustrative demonstration of legibility and the length of the code necessary for solving the identical task about the possible implemented variants.

3. MATERIALS AND METHODS

For implementation, we use open source framework Spring (http://spring.io/). Spring Tool Suite™ (STS) is an eclipse-based development environment that is customized for developing Spring applications. It provides a ready-to-use environment to implement, debug, run, and deploy Spring applications. We use Oracle 11g database, layered with Hibernate as the object mapper and then Spring Data JPA for abstraction and some nice boilerplate crud (create, read, update, delete) operations. As it is mentioned on the official web page of Spring (2017) "the goal of Spring Data repository abstraction is to significantly reduce the amount of boilerplate code required to implement data access layers for various persistence stores."
Users in applications are represented by the simple class with four fields (identification number, user name, age and e-mail addresses). Below you can see annotations that instruct hibernate how to persist the object.

```
@Entity
@Table(name = "PERF_TEST_USER")
public class User implements Serializable {
        private static final long serialVersionUID = 1L;

        @Id
```

```
        @SequenceGenerator(name = "PERF_TEST_USER_ID_GEN", sequenceName =
        "PERF_TEST_USER_SEQ", allocationSize = 1)
        @GeneratedValue(strategy = GenerationType.SEQUENCE,
                generator = "PERF_TEST_USER_ID_GEN")
        @Column(name = "ID")
        private Long id;

        @Column(name = "USER_NAME", nullable = false)
        private String userName;

        @Column(name = "EMAIL_ADDRESS", nullable = false)
        private String emailAddress;

        @Column(name = "AGE", nullable = false)
        private int age;

        public Long getId() { return this.id; }

        public void setId(final Long id) { this.id = id; }

        public String getUserName() { return this.userName; }

        public void setUserName(final String userName) {
                this.userName = userName; }

        public String getEmailAddress() { return this.emailAddress; }

        public void setEmailAddress(final String emailAddress) {
                this.emailAddress = emailAddress; }

        public int getAge() { return this.age; }

        public void setAge(final int age) { this.age = age; }
}
```

The reason for class simplicity is also the fact that we do not work with real data (reason: protection of personal data), but we generate them, what is sufficient for our test purposes. Work with the database, i.e. generating users, storing them, or removing them from the database is solved by using the *IUserService* interface implemented by the *UserService* class. This class for the data acquisition from the database implements two methods: ***public*** *List<User> getTestUsersAsList(**final int** countOfTestUsers)* and ***public*** *Iterator<User> getTestUsersAsIterator(**final int** countOfTestUsers)*, the principle of which is the same. The goal is to return the required number of users from the database either as a collection or as an iterator. An *iterator* in Java is an instance of the *Iterator<E>* interface (Interface List<E> 2017). We can get an iterator for a collection using the *iterator()* method from the *Collection* interface. In the *getTestUsersAsList* method we use a method *List<E> subList(int fromIndex, int toIndex)*, which returns a view of the portion of the list between the specified *fromIndex*, inclusive, and *toIndex*, exclusive (Interface List<E> 2017).

```
public List<User> getTestUsersAsList(final int countOfTestUsers) {
     final List<User> testUsers = users.subList(0, countOfTestUsers - 1);
     return testUsers;
}

public Iterator<User> getTestUsersAsIterator(final int countOfTestUsers) {
     final List<User> testUsers =
             this.getTestUsersAsList(countOfTestUsers);
     return testUsers.iterator();
}
```

If there are not enough users in the database, they will be generated.

```
public User generateRandomUser() {
  final User user = new User();

  user.setAge(this.getRandomAge());
  user.setUserName(this.getRandomUserName());
  user.setEmailAddress(this.getRandomEmailAddress());

  return user;
}

// USER ATRIBUTES RANDOM GENERATOR

private int getRandomAge() {
  final int randomAge = new Random().nextInt(this.maxAgeTo -
this.maxAgeFrom + 1) + this.maxAgeFrom;
  return randomAge;
}

private String getRandomUserName() {
```

```
    final UUID randomUuid = UUID.randomUUID();
    return String.valueOf(randomUuid);
  }

  private String getRandomEmailAddress() {
    final String randomUserName = this.getRandomUserName();
    return randomUserName.concat("@user.com");
  }
```

A collection is an object that groups multiple elements into a single unit. Collections are used to store, retrieve, manipulate, and communicate aggregate data (Collections 2017). They are fundamental to many programming tasks. We use an *ArrayList* class that is implemented using resizable array, because retrieval is the key operation for us. According Paul Javin (2017), the main difference between *ArrayList* and *LinkedList* is that *ArrayList* is implemented using resizable array while *LinkedList* is implemented using doubly linked list. Since array is an index-based data-structure, searching or getting element from array with index is fast. Array provides $O(1)$ performance for *get(index)* method but remove is costly in *ArrayList* as you need to rearrange all elements. On the other hand, *LinkedList* does not provide random or index based access and you need to iterate over linked list to retrieve any element that is of order $O(n)$. Insertions are easy and fast in *LinkedList* comparing to *ArrayList* because there is no risk of resizing array and copying content to new array if array gets full what makes adding into *ArrayList* of $O(n)$ in worst case, while adding is $O(1)$ operation in *LinkedList* in Java. *ArrayList* also needs to update its index if you insert something anywhere except at the end of array. Removal is like insertions better in *LinkedList* than *ArrayList*. *LinkedList* has more memory overhead than *ArrayList* because in *ArrayList* each index only holds actual object (data) but in case of *LinkedList* each node holds both data and address of next and previous node. According Paul Javin (2012) "*ArrayList* is more popular among Java programmer than *LinkedList*, but there are few scenarios on which *LinkedList* is a suitable collection than *ArrayList*." This is not the subject of this paper. However, based on the above, it is obvious that the results achieved by using *LinkedList* would be different.

The developed Spring Boot application provides a simple Rest API interface for testing. These results are used to determine the difference, between the lambda and non-lambda expression in nanoseconds. For testing, we use a single method that can be called with the following mandatory parameters:

- *testTypes* (enum — for, iterator, stream…);
- *approach* (enum — 1, 2, 3…);
- *skipWarmUp* (boolean);
- *countOfWarmUpIterations* (number);
- *countOfAllIterations* (number);
- *countOfTestUsers* (number);
- *getDurationsList* (boolean — displaying also the results for each iteration of the test).

From the above, we can use multiple test variants (the *testTypes* attribute) for one tested approach at a time. With the *skipWarmUp* parameter, we can exclude the so-called "warm-up iterations" from the test results, the number of which is determined by the *countOfWarmUpIterations* attribute. The reason is the selection of data that could distort the results. The principle of testing could be defined in simple terms by this construction:

```
long startTime = System.nanoTime();
     // ... THE CODE BEING MEASURED ...
long estimatedTime = System.nanoTime() - startTime;
```

To communicate with our API, we use the SoapUI (5.3.0) tool, which is primarily designed for functional testing. At each call of our method we can see a clear result in the JSON format, from which we can unambiguously conclude testing for a given combination of input parameters.

4. DEFINITION OF LAMBDA EXPRESSIONS

Kishori Sharan (2014) defines a lambda expression as "an unnamed block of code (or an unnamed function) with a list of formal parameters and a body." However, this method is not executed on its own. Instead, it is used to implement a method defined by a functional interface. The lambda expression signature must be the same as the functional interface method's signature, as the target type of the lambda expression is inferred from that context. Here is a key point: a lambda expression can be used only in a context in which its target type is specified. A lambda expression is a poly expression — the type of a lambda expression is inferred from the target type thus the same lambda expression could have different types in different contexts. A functional interface is sometimes referred to as a SAM type, where SAM stands for Single Abstract Method. Lambda expressions are also commonly referred to as closures (Schildt 2014).

According Schildt (2014) "a functional interface is an interface that contains one and only one abstract method." A functional interface typically represents a single action. In the past, no method in an interface could include a body. Thus, all methods in an interface were implicitly abstract. With the release of JDK 8, this situation has changed dramatically. It is now possible for an interface method to define a default implementation. This new capability is called the default method (Default Methods 2017).

5. SYNTAX OF LAMBDA EXPRESSIONS IN JAVA (LAMBDA EXPRESSIONS 2017)

A lambda expression consists of the following: **(arguments) -> body**

- A comma-separated list of formal parameters enclosed in parentheses specifies parameters required by the lambda expression. If the parameter types of a lambda expression can be inferred, you can omit them. In addition, you can omit the parentheses if there is only one parameter. Parameters are optional, if no parameters are needed an empty parenthesis can be given.
- The arrow token -> (arrow operator, lambda operator).
- A body, which consists of a single expression or a statement block. If you specify a single expression, then the Java runtime evaluates the expression and then returns its value. Alternatively, you can use a return statement. A return statement is not an expression, in a lambda expression, you must

enclose statements in braces ({}). However, you do not have to enclose a void method invocation in braces. It is illegal for a lambda expression to return a value in some branches but not in others. You can write following kind of code using lambdas:

- (params) -> expression
- (params) -> statement
- (params) -> { statements }

For example: lambda expression to test if the given number is odd or not: *n -> n % 2 != 0*. This lambda expression returns *true* if the value of parameter *n* is odd. Compared to a method a lambda definition lacks a return type, a throws clause and a name. Return type and exceptions are inferred by the compiler from the lambda body. So, the only thing that is really missing is the name.

6. RESULTS AND DISCUSSION

The Java platform includes a variety of ways to iterate over a collection of objects, including new options based on features introduced in Java 8. One common approaches to iterate over *ArrayList* in Java is *advanced for loop* (also known as *for each loop*, introduced to Java 5). If I have to remove elements while iterating then using *Iterator* or *ListIterator* is the best solution. Although iterators in Java have taken different forms, using an active iterator was essentially the only viable option prior to Java 8. For an active iterator (also known as explicit iterator or external iterator), the client controls the iteration in the sense that the client creates the iterator, tells it when to advance to the next element, tests to see if every element has been visited, and so on. For a passive iterator (also known as an implicit iterator, internal iterator, or callback iterator), the iterator itself controls the iteration. The client essentially says to the iterator, “perform this operation on the elements in the collection.” Now we have possibility to use Java 8’s *forEach()* method and features of the stream API which helps us fine-tune and parallelize the behaviour of Java iterators. For each tested option we defined a single class, as you can see in figure 1.

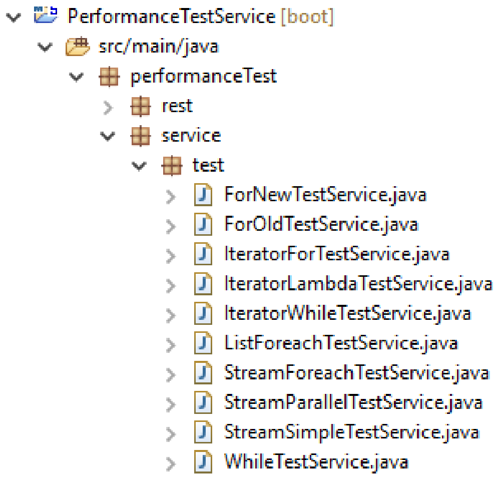


Figure. 1 List of classes for testing purposes.

In the examples below, we present these different iteration options to solve simple basic logic in the simplest way. One simplistic approach is to create several methods. Each method searches for members that match one or more characteristics. Our implemented methods return a collection of names of users who meet the simple filtering criterion. The task and even the selection criteria are not crucial, the applied approach is important. Consequently, the appropriate action can be taken over this data. There are more approaches to the problem solution that can be applied regarding the development of the Java programming language. We applied five different approaches that eliminate lack of the simplest approach presented below. It improves upon this approach with local and anonymous classes, and then finishes with an efficient and concise approach using lambda expressions. These will be presented in detail in another paper.

APPROACH 1: CREATE METHODS THAT SEARCH FOR MEMBERS THAT MATCH ONE CHARACTERISTIC

1st variant: Simple for loop and an integer index — old practice

```
public List<String> approach1(final int countOfTestUsers) {
  final List<User> users =
     this.userService.getTestUsersAsList(countOfTestUsers);
  final List<String> userNames = new ArrayList<>();

  for (int index = 0; index < users.size(); index++) {
     final User user = users.get(index);

     if ((user.getId() + user.getAge()) % 2 == 1) {
             userNames.add(user.getUserName());
     }
  }
  return userNames;
}
```

2nd variant: While loop — old practice

The use of the cycle with a condition at the beginning (*while loop*) if we need to iterate the whole collection has no logical justification. In such a case, a cycle with an explicitly defined number of repeats (*for loop*) is more appropriate and readable. For the sake of completeness, we present also this variant, as it is the case with the 4th variant. There is the assumption that there will be no performance differences between the 1st and 2nd (similarly between 3rd and 4th) variants.

```
public List<String> approach1(final int countOfTestUsers) {
  final List<User> users =
     this.userService.getTestUsersAsList(countOfTestUsers);
  final List<String> userNames = new ArrayList<>();

  int index = 0;
  while (index < users.size()) {
     final User user = users.get(index++);

     if ((user.getId() + user.getAge()) % 2 == 1) {
                     userNames.add(user.getUserName());
```

```
        }
    }
    return userNames;
}
```

3rd variant: Iterator + for loop

An iterator is a mechanism that permits all elements of a collection to be accessed sequentially, with some operation being performed on each element while isolating the user from the internal structure of the container. An iterator provides a means of “looping” over an encapsulated collection of objects. It was introduced in the Java JDK 1.2. An iterator should also be non-destructive in the sense that the act of iteration should not, by itself, change the collection. Of course, the operation being performed on the elements in a collection could possibly change some of the elements. It might also be possible for an iterator to support removing an element from a collection or inserting a new element at a point in the collection, but such changes should be explicit within the program and not a by-product of the iteration. Typically, the *hasNext()* and *next()* methods are used together in a loop. The *next()* method returns the next element from the collection. We should always call the *hasNext()* method before calling the *next()* method. If not, the *next()* method throws a *NoSuchElementException*. *ListIterator* extends *Iterator* to allow bidirectional traversal of a list, and the modification of elements, what we do not need now. Only collections that implement *List*, you can obtain an iterator by calling *ListIterator()*.

```
public List<String> approach1(final int countOfTestUsers) {
  final List<String> userNames = new ArrayList<>();

for (final Iterator<User> users =
     this.userService.getTestUsersAsIterator(countOfTestUsers);
     users.hasNext();) {
     final User user = users.next();

     if ((user.getId() + user.getAge()) % 2 == 1) {
             userNames.add(user.getUserName());
     }
  }
  return userNames;
```

```
}
```

4thvariant: Iterator + while loop

The same philosophy as in the 3rd variant, but using *while loop*. We have met such a use in the articles more often than the use *for loop*. This was the reason for the inclusion of this option.

```
public List<String> approach1(final int countOfTestUsers) {
  final List<String> userNames = new ArrayList<>();

  final Iterator<User> users =
      this.userService.getTestUsersAsIterator(countOfTestUsers);

  while (users.hasNext()) {
      final User user = users.next();

      if ((user.getId() + user.getAge()) % 2 == 1) {
              userNames.add(user.getUserName());
      }
  }
  return userNames;
}
```

5th variant: Advanced (enhanced) for loop = for each loop

From Java 5, we can use the *for each loop* to iterate over any collection whose implementation class implements the *Iterable* interface. This iteration is a more convenient way than using the *traditional for loop*. The creation of the iterator and calls to its *hasNext()* and *next()* methods are not expressed explicitly in the code, but they still take place behind the scenes. Thus, even though the code is more compact.

```
public List<String> approach1(final int countOfTestUsers) {
    final List<String> userNames = new ArrayList<>();

  for (final User user :
        this.userService.getTestUsersAsList(countOfTestUsers)) {
        if ((user.getId() + user.getAge()) % 2 == 1) {
                userNames.add(user.getUserName());
        }
```

```
  }
  return userNames;
}
```

6th variant: forEach method with lambda expressions

Collection classes that implement *Iterable* (for example all list classes) from Java 8 have a *void forEach(Consumer<? super T> action)* method (Interface Iterable<T> 2017). This method accepts a *consumer type*, which is a functional interface and that is why we can pass a lambda expression to it that has been passed as argument to every single element of the stream. Using it leads to shorter code constructs. *forEach()* is a terminal operation, which means that can only once be applied on a stream. The greatest difference comparing a *for loop* is that it cannot be interrupted ahead of time — neither with break nor with return. The form of using active iterators is not wrong and it is still being in use. Java 8 simply provides additional capabilities and new ways of performing iteration. For some scenarios, the new ways can be better. In particular, the *Iterable* interface provides a passive iterator in the form of a default method called *forEach()*. In this case, the *forEach()* method is actually implemented using an active iterator in a manner similar to what you saw in *Advanced (enhanced) for loop = for each loop* example. Note the difference between the passive iterator in this example and the active iterator in the previous examples. There is no explicit loop. We simply tell the *forEach()* method what to do with the objects in the list. Control of the iteration resides within the *forEach()* method.

```
public List<String> approach1(final int countOfTestUsers) {
  final List<String> userNames = new ArrayList<>();

    this.userService
      .getTestUsersAsList(countOfTestUsers)
      .forEach(user -> {
                                if ((user.getId() + user.getAge()) % 2 == 1)
{
                                        userNames.add(user.getUserName());
                                }
                         });
  return userNames;
}
```

7th variant: Iterator + forEachRemaining method with lambda expressions

In Java 8, the *Iterator<T>* interface contains the extension method *void forEachRemaining(Consumer<? super T> action)* which performs the given action for each remaining element until all elements have been processed or the action throws an exception. Actions are performed in the order of iteration, if that order is specified. The action is specified as a *consumer* that takes a lambda. A functional *consumer* interface represents an operation that accepts a single input argument and returns no result (API Specification 2017).

```
public List<String> approach1(final int countOfTestUsers) {
  final List<String> userNames = new ArrayList<>();

  this.userService.getTestUsersAsIterator(countOfTestUsers)
      .forEachRemaining(user -> {
              if ((user.getId() + user.getAge()) % 2 == 1) {
                      userNames.add(user.getUserName());
              }
      });

  return userNames;
}
```

8th variant: The stream API

Since Java 8, internal iteration is supported via streams. Stream is a key abstraction from the JDK collection framework that supports bulk operations with internal iteration. These bulk operations include *forEach()*, *filter()*, *map()*, *reduce()*, and many more methods. Thanks to the internal iteration, streams support sequential as well as parallel execution in a very convenient way. Arguably, one of the most important new features of JDK 8 is the stream API, which is packaged in *java.util.stream*. Collections and arrays can be used to generate streams — but beware: streams are not collections that store elements. Rather, think of a stream as a mechanism for carrying a sequence of values from a source through a pipeline of operations. Underlying data structures such as arrays or lists are therefore not changed. A stream pipeline consists of the stream source, intermediate operations that transform the stream and produce a new stream, and a terminal operation that either produces a result or calls the *forEach()* method (Aggregate Operations 2017). In general, intermediate stream operations are lazy, which means that they are not immediately executed, but delayed until a terminal operation is applied. After the invocation of terminal methods, no further stream operations can be performed.

The key aspect of the stream API is its ability to perform pipeline operations that search, filter, map, or otherwise manipulate data. Often, you will use lambda expressions to specify the behaviour of these types of operations. Furthermore, in many cases, such actions can be performed in parallel, thus providing a high level of efficiency, especially when large data sets are involved. Put simply, the stream API provides a powerful means of handling data in an efficient, yet easy to use way (Schildt 2014).

```
public List<String> approach1(final int countOfTestUsers) {
  return this.userService
      .getTestUsersAsList(countOfTestUsers).stream()
      .filter(user -> (user.getId() + user.getAge()) % 2 == 1)
      .map(user -> user.getUserName())
      .collect(Collectors.toList());
}
```

The *Stream<T> filter(Predicate<? super T> predicate)* method expect a *predicate* instance which is also a functional interface as an argument and returns a new stream containing the elements that matched the conditions of predicate. Just remember that filter does not remove elements which matches the condition given in predicate, instead it selects them in output stream. Each predicate can have multiple conditions that need to be satisfied. A lambda expression can be passed into the *filter()* method instead, enabling you to customize how they behave. A lambda expression is the perfect way to express the actual parameter, which is simply a boolean-valued function. Aggregate operations process elements from a stream, not directly from a collection, which is the reason why the first method invoked in this example is *stream*.

Filter() is an intermediate operation, which means you can call any other method of stream e.g. *map()* as stream will not be closed due to filtering. The *<R> Stream<R> map(Function<? super T,? extends R> mapper)* method is an intermediate operation which returns a stream consisting of the results of applying the given function to the elements of this stream. Stream *map* method takes function as argument that is also a functional interface (Interface Stream<T>, 2017).

The *collect* operation is best suited for collections. Unlike the *reduce* method (Aggregate Operations 2017), which always creates a new value when it processes an element, the *collect* method modifies, or mutates, an existing value. If us *reduce* operation involves adding elements to a collection, then every time our accumulator function processes an element, it creates a new collection that includes the element, which is inefficient. It would be more efficient for us to update an existing collection instead. This version of the *collect* operation takes one parameter of type *Collector*. The *Collectors* class contains many useful reduction operations, such as accumulating elements into collections and summarizing elements according to various criteria. These reduction operations return instances of the class *Collector*, so you can use them as a parameter for the *collect* operation. We use the *Collectors.toList* operation, which accumulates the stream elements into a new instance of *List*. As with most operations in the *Collectors* class, the *toList* operator returns an instance of *Collector*, not a collection.

The intermediate operation *filter()* and *map()* shown in example is only two example of a transformation that can be performed on a stream. Other examples include *distinct()* and *sorted()*, which applies a function to the elements of the stream to produce a new stream. Similarly, there are a number of terminal operations other including *count()*, *average()*, *sum()*, *min()*, *max()*, and *forEach()*. For more information about stream pipelines, see “Aggregate Operations” in the Java tutorial (Aggregate Operations 2017) or “Chapter 2: The Stream API” in Cay Horstmann’s book (2014).

9th variant: The stream API with forEach method with lambda expressions

The *forEach()* method is defined at two places, on *Iterable* interface (*6th variant: forEach method with lambda expressions example*) as well as on stream, which means *list.forEach()* and *list.stream.forEach()* both are valid.

```
public List<String> approach1(finalintcountOfTestUsers) {
  final List<String> userNames = new ArrayList<>();
  this.userService
        .getTestUsersAsList(countOfTestUsers).stream()
        .forEach(user -> {
                if ((user.getId() + user.getAge()) % 2 == 1) {
                        userNames.add(user.getUserName());
                }
        });
  return userNames;
}
```

10th variant: The parallel stream API

Collection classes do not only have a *stream()* method, which returns a sequential stream, but they also have a *parallelStream()* method, which returns a parallel stream. Aggregate operations and parallel streams enable you to implement parallelism with non-thread-safe collections provided that you do not modify the collection while you are operating on it. When a stream executes in parallel, the Java runtime partitions the stream into multiple substreams. Aggregate operations iterate over and process these substreams in parallel and then combine the results. In other words: parallel streams have the potential to improve performance by allowing pipeline operations to be executed concurrently in separate Java threads; but thought that the order in which collection elements are processed can change.

Implementation of parallel bulk operations for collections requires a better separation of concerns. And this is what lambdas are for: they provide a convenient and concise notation for functionality, which can be passed as an argument to a bulk operation of a collection, which in turn applies this functionality in parallel to all its elements.

```
public List<String> approach1(final int countOfTestUsers) {
```

```
    return this.userService
          .getTestUsersAsList(countOfTestUsers).parallelStream()
          .filter(user -> (user.getId() + user.getAge()) % 2 == 1)
          .map(user -> user.getUserName())
          .collect(Collectors.toList());
}
```

Evaluation

Java 8 stream is very efficient replacement of looping both design and performance wise, because it separates what to do from how to do and supported internal iteration. In *the stream API* example, with one line of code a new list is created, containing only the desired names of users. Over traditional *for loop* we can write more succinct code, sometimes just one line long. You can pass lambda expression, which gives you the immense flexibility to change what you do in the loop. Aggregate operations do not contain a method like *next* to instruct them to process the next element of the collection. With internal delegation, your application determines what collection it iterates, the JDK determines how to iterate the collection. You did not iterate over the elements in the list when you use the stream. The stream did that for you internally. With external iteration, your application determines both what collection it iterates and how it iterates it. However, external iteration can only iterate over the elements of a collection sequentially, that is the code can be executed only by one thread. Internal iteration does not have this limitation. It can more easily take advantage of parallel computing, which involves dividing a problem into subproblems, solving those problems simultaneously, and then combining the results of the solutions to the subproblems. In *the parallel stream API* example by changing only one word we take advantage of multiple CPU. So, you can perform the looping in parallel without writing a single line of multi-threading code. Of course, it is possible only on a computer with multiple cores, and with operations that can and should be executed in parallel. Kishori (2014) mentioned that Java has supported concurrent programming since the beginning. "It added support for parallel programming in Java 7 through the *Fork/Join framework*, which is not so simple to use, especially for beginners." On the other side, *streams* come as a rescue. They are designed to process their elements in parallel without you even noticing it. When you want to process elements in parallel you use *parallelStream()* method and the stream will take care of the rest. Streams take care of the details of using the *Fork/Join framework* internally.

It is obvious that these two approaches (traditional imperative programming vs. declarative programming) look very different. The imperative style intermingles the logic of iterating (the various nested loops) and the application of the declarative style, in contrast, separates concerns. It describes which functionality is supposed to be applied and leaves the details of how the functions are executed to the various operations such as *filter*, *map*, etc., what is according us more readable. Due to new *stream API* features you can write more expressive code, which is easy to understand, easy to read, easy to optimize and super easy to run in parallel without you worrying about multi-threading nightmare. The *stream API* supports the functional programming. Operations on a stream produce a result without modifying the data source. Like in the functional programming, when you use streams, you specify "what" operations do you want to perform on its elements using the built-in methods provided by the *stream API*, typically by passing a lambda expression to those methods, customizing the behaviour of those operations.

As we can see above (*6th a 9th variant*) the *forEach()* method is defined at two places. Prefer use *forEach()* with streams because streams are lazy and not evaluated until a terminal operation is called. Also obtaining stream gives you more choices e.g. *filtering*, *mapping* etc. (*8th variant*). The *forEach()* method is a terminal operation, which means you cannot reuse the stream after calling this method. It will throw *IllegalStateException* if you try to call another method on this stream. When you run *forEach()* method on parallel stream the order on which elements are processed is not guaranteed, though you can use *forEachOrdered()* to impose ordering. Based on the above, when comparing the 8th and 9th variants, for the reasons of legibility and clarity, we are for the use 8th variant. It is illogical to combine new approaches with old ones. The if-statement can be easily replaced by the filter method.

A non-lambda expression (*variants 1–5*) to accomplish the same feet requires creating a new list, using an if-statement to check the condition of each user, and then adding the wanted names of user to the list. So instead of filtering, the non-lambda expression uses an if-statement. For inserting the specified elements into the list we use *java.util.ArrayList.add(int index, E element)* method (Class ArrayList<E> 2017). The *enhanced for loop* actually uses an iterator behind the scenes. That means the Java compiler will convert the *enhanced for loop* syntax to iterator construct when compiling. The new syntax just gives the programmers a more convenient way for iterating over collections than old fashion practices, but it is more limited. You cannot remove elements from the collection while iterating, there is no reference to the iterator to call the *remove()* method. Even though you would end up with with *ConcurrentModificationException*. The other practical difference between *for loop* and *enhanced for loop* is that, in the case of indexable objects, you do not have access to the index. Although you could manually create a separate index integer variable with *for each loop*. But there is extra operation, which is not needed. With *for each loop* we have no way to start from the middle of the collection. The *for each loop* provides no way to visit the previously visited elements. So, if you want to selective modify only certain object then your best bet is to use *traditional for loop* which keeps track of indexes. Also, streams do not offer random access on the source data via index or the like. Access to the first element is possible, however not on any following elements.

The *advanced for loop* is same as using *iterator* with *for* or *while loop*. However, both *while* and *for* in combination with *iterator* gives more flexibility and power as you can iterate collection. On the other hand, we can say that the *iterator* is error prone. Because, at the use of the *iterator* there is nothing special if the Java program goes into infinite loop, which eventually cause out of memory error. This error message (exception) may be caused by use of the *Iterator.next()* inside loop. An *iterator* is a one-time object. We cannot reset an *iterator* and it cannot be reused. To iterate over the elements of the same collection again, it is necessary to create a new *iterator* by calling the *iterator()* method of the collection. But using *forEachRemaining()* method with a lambda also lets you control the collection type or even to use a pre-existing collection instance if wanted. Lambda expressions have the great advantage of creating much more readable andclean code. I tend to use them wherever I get a real

benefit — in terms of readability — from them. Undoubtedly, there exists a lambda example that is not more or less readable then the example that does not use a lambda expression. In fact, there also exist an example, that has the similar number of code lines, but in the example above you can see that even when dealing with a simple basic logic, readability as well as verbosity are obvious.
You might still wonder why anyone would prefer the code in *Advanced (enhanced) for loop example* over *The stream API example*, which probably looks more familiar to most Java developers. It can be argued that the declarative approach in *The stream API example* is actually more readable and less error prone, at least once you understand the basic concepts of streams and stream operations. In particular, depending on the context, the logic in *Advanced (enhanced) for loop example* might not be thread-safe, while *The stream API example* is thread safe. It is also easier to parallelize the operations shown in *The stream API example* as you see in *The parallel stream API example.*

While there are several advantages to use the passive iterators and new features of Java 8, I thought that it would be interesting to test whether or not they provide a performance improvement.
The testing ran on the computer with an Intel® Core™ i7-5500U CPU @ 2.40GHz processor, 16 GB RAM, the 64-bit version of Windows 10 and a 64-bit version of Java (version 1.8.0_131). I timed the calls to each of these ten methods using *ArrayList* classes with collection sizes (1 000 users) but with different count of iterations. The data in table 1 presents the average duration of the operation for the presented approaches at a varying number of iterations. The aim was to find the answer to the research question, whether the number of iterations performed has a significant impact on the measured values. For the sake of the utmost accuracy, we measure the values in nanoseconds, but for the sake of clarity and conclusions, we also report them in milliseconds. It follows that for our other testing purposes, where we will only change the size of the data structure, it will be sufficient to perform 10 000 iterations.

Table 1. The result of testing — average duration of operation in ns and ms — different numbers of iterations.

approach	Approach1							
countOfIterations	1 000		10 000		100 000		1 000 000	
countOfTestUsers	1 000		1 000		1 000		1 000	
time unite	ns	ms	ns	ms	ns	ms	ns	ms
FOR_OLD	275 809,75	0,28	228 089,50	0,23	239 959,39	0,24	244 379,47	0,24
WHILE	292 745,69	0,29	221 698,49	0,22	241 535,65	0,24	235 898,24	0,24
ITERATOR_FOR	296 489,30	0,30	227 772,45	0,23	224 985,29	0,22	232 021,68	0,23
ITERATOR_WHILE	283 545,75	0,28	279 073,43	0,28	225 070,37	0,23	224 372,72	0,22
FOR_ADVANCED	326 107,87	0,33	237 142,25	0,24	215 781,47	0,22	214 921,40	0,21
LIST_FOREACH	226 403,02	0,23	192 308,70	0,19	179 811,89	0,18	178 805,51	0,18
ITERATOR_LAMBDA	233 293,20	0,23	188 423,71	0,19	169 199,60	0,17	173 992,42	0,17
STREAM_SIMPLE	248 457,09	0,25	205 321,19	0,21	174 555,27	0,17	178 316,26	0,18
STREAM_FOREACH	211 732,14	0,21	182 507,96	0,18	166 926,15	0,17	188 002,86	0,19
STREAM_PARALLEL	161 079,86	0,16	175 881,53	0,18	143 907,95	0,14	139 356,43	0,14

I timed the calls to each of these ten methods using *ArrayList* classes, with different collection sizes (1 000, 10 000 and 100 000 users). I ran the benchmark several times (10 000 iterations). To remove any startup or just-in-time (JIT) effects, we exclude first five iterations. The data presented in table 2 are arranged from the shortest processing of the operation to the longest one.

Table 2. The result of testing Approach 1 — average duration of operation in nanoseconds and milliseconds.

approach	Approach1							
countOfIterations	10 000		10 000			10 000		
countOfTestUsers	1 000		10 000			100 000		
time unit	ns	ms		ns	ms		ns	ms
STREAM_PARALLEL	161 159,79	0,16	STREAM_PARALLEL	913 255,42	0,91	STREAM_PARALLEL	11 974 573,92	11,97
STREAM_FOREACH	168 929,19	0,17	STREAM_SIMPLE	2 030 324,34	2,03	STREAM_FOREACH	26 971 814,24	26,97
ITERATOR_LAMBDA	169 719,38	0,17	ITERATOR_LAMBDA	2 129 308,08	2,13	ITERATOR_LAMBDA	27 185 775,28	27,19
STREAM_SIMPLE	174 435,76	0,17	STREAM_FOREACH	2 148 108,51	2,15	STREAM_SIMPLE	27 212 472,88	27,21
LIST_FOREACH	175 472,36	0,18	LIST_FOREACH	2 163 293,37	2,16	LIST_FOREACH	28 273 787,57	28,27
WHILE	210 944,88	0,21	WHILE	2 495 592,67	2,50	FOR_ADVANCED	32 324 300,40	32,32
FOR_ADVANCED	217 192,56	0,22	FOR_OLD	2 588 545,95	2,59	WHILE	32 880 768,40	32,88
ITERATOR_FOR	217 794,62	0,22	FOR_ADVANCED	2 645 102,00	2,65	ITERATOR_WHILE	33 434 215,47	33,43
FOR_OLD	217 822,30	0,22	ITERATOR_FOR	2 645 248,20	2,65	ITERATOR_FOR	33 703 254,11	33,70
ITERATOR_WHILE	221 575,12	0,22	ITERATOR_WHILE	2 652 470,27	2,65	FOR_OLD	34 245 669,42	34,25

In some cases, the benchmark results were not what we expected. For example, we expected that there would be no differences between using *for* (FOR_OLD) and *while loop*. Similarly, we did not expect any difference when implementing the *iterator* (ITERATOR_FOR and ITERATOR_WHILE). We can see that these differences are not significant, but they still exist at a certain level of detail. According Javin Paul (2016) "the *enhanced for loop* provides the cleanest way to loop through an array or collection in Java, as you don't need to keep track of counter or you don't need to call the *hasNext()* method to check whether there are more elements left." Paul (2016) also states that "an *enhanced for loop* is nothing but a syntactic sugar over *iterator* in Java." Also, based on the above facts, our expectation were that the use of an *enhanced for loop* (FOR_ADVANCED) will not bring any significant performance improvements over older approaches, which was confirmed by the investigation results. On the other hand, we expected the duration of the operation to be the same for these two approaches. However, we cannot confirm this definitively, since there are some minor differences. So, we can state that this is an improvement in terms of clarity and readability of the code, but not in terms of performance.

More extensive benchmarking should be performed using different computer configurations in a production or different simulated production scenarios before drawing any definitive conclusions. But based on this simple benchmark and the results shown in the table 2, the following statements appeared to be true: New features added to Java 8 improve performance, even though we were expecting greater differences between the individual options. As you can see the *parallel stream* is roughly 50% faster in big collection. At the processing of the data on the size of 1 000 users (small collection), these differences are not significant. We can say that using multiple threads to process them is overkill. The second shortest duration of the operation was achieved in two cases using the *list.stream.forEach()* method (STREAM_FOREACH). The use of the *forEach()* method applied directly over the collection (LIST_FOREACH) achieves slightly worse results. Very interesting results were achieved using *iterator* with the new extension method *forEachRemaining()* (ITERATOR_LAMBDA). We do not need to omit the use of streams with aggregate operations (STREAM_SIMPLE), which, beside the code clarity, also brings positive performance results.

In general, benchmark experiment results must be taken with a grain of salt. Every benchmark is an experiment. Minor deviations in measurements can be caused either by running background processes, but most of them we eliminated, or we have to consider the fact that when measuring the duration of a process, we never achieve exactly the same results when performing multiple tests. Obviously, the deviations are minimal and do not have a significant impact on the conclusions. We tried to increase the accuracy of the data by many iterations as well as the calculation of the mean values.

APPROACH 2: CREATE MORE GENERALIZED SEARCH METHODS

The philosophy of all ten variants of iteration over the collection remains the same at Approach 2 as in Approach 1, only the number of conditions that are set for selecting suitable users changes. For this reason, just one example is illustrated.

1st variant: Simple for loop and an integer index — old practice

```
public List<String> approach2(final int countOfTestUsers) {
  final List<User> users =
        this.userService.getTestUsersAsList(countOfTestUsers);
  final List<String> userNames = new ArrayList<>();

  for (int index = 0; index < users.size(); index++) {
     final User user = users.get(index);

     if ((user.getId() + user.getAge()) % 2 == 1
             && user.getId() % 2 == 1
             && user.getUserName().contains("5")
             && user.getEmailAddress().contains("1")
             && user.getEmailAddress().startsWith("a")
             && user.getAge() % 2 == 0) {
                     userNames.add(user.getUserName());
                     }
     }
  return userNames;
```

```
}
```

We also timed the calls to each of these ten methods using *ArrayList* classes, with different collection sizes (1 000, 10 000 and 100 000 users). I ran the benchmark several times (10 000 iterations). To remove any startup or just-in-time (JIT) effects, we exclude first five iterations. The data presented in table 3 are arranged from the shortest duration of the operation processing to the longest.

Table 3. The result of testing Approach 2 — average duration of operation in nanoseconds and milliseconds.

approach	Approach2							
countOfIterations	10 000		10 000			10 000		
countOfTestUsers	1 000		10 000			100 000		
time unit	ns	ms		ns	ms		ns	ms
STREAM_PARALLEL	131 073,19	0,13	STREAM_PARALLEL	775 377,32	0,78	STREAM_PARALLEL	10 352 412,13	10,35
STREAM_FOREACH	145 108,63	0,15	STREAM_SIMPLE	1 754 125,24	1,75	STREAM_SIMPLE	23 272 473,62	23,27
STREAM_SIMPLE	151 853,15	0,15	STREAM_FOREACH	1 854 730,98	1,85	STREAM_FOREACH	24 315 980,06	24,32
ITERATOR_LAMBDA	152 857,10	0,15	LIST_FOREACH	1 926 932,75	1,93	ITERATOR_LAMBDA	25 028 585,49	25,03
LIST_FOREACH	167 999,86	0,17	ITERATOR_LAMBDA	1 931 942,53	1,93	LIST_FOREACH	25 575 949,43	25,58
FOR_ADVANCED	187 141,34	0,19	WHILE	2 213 931,07	2,21	WHILE	27 889 886,24	27,89
FOR_OLD	189 976,61	0,19	FOR_ADVANCED	2 231 369,21	2,23	ITERATOR_FOR	28 978 475,68	28,98
WHILE	192 609,15	0,19	FOR_OLD	2 270 359,61	2,27	FOR_ADVANCED	29 085 196,01	29,09
ITERATOR_WHILE	194 885,71	0,19	ITERATOR_WHILE	2 276 332,35	2,28	FOR_OLD	29 437 404,21	29,44
ITERATOR_FOR	195 339,55	0,20	ITERATOR_FOR	2 323 049,71	2,32	ITERATOR_WHILE	29 482 042,72	29,48

The order of individual variants remains more or less unchanged, minor changes in the order of some variants are not significant. However, it is worth to mention that the average duration of an operation is in each case shorter comparing Approach 1, even though the applied business logic of choice is more complex. Thus, we can conclude that the difficulty of the selection criterion does not have a significant effect on the prolongation of the operation duration.

CONCLUSION

A fluent programming (the chaining of operations) is a programming technique where operations return a value that allows the invocation of another operation. With fluent programming, it is perfectly natural to end up with one huge statement that is the concatenation of as many operations as you need. The JDK streams are designed to support fluent programming: all intermediate operations return a result stream to which we can apply the next operation in the chain until we apply a terminal operation as the tail of the chain. The chaining of operations is a key feature of streams as it enables a number of optimizations. For instance, the stream implementation can arrange the execution of functions as a pipeline. Instead of looping over all elements in the sequence repeatedly (once for filter, then again for map, and eventually for *toArray*) the chain of filtermapper-collector can be applied to each element in just one pass over the sequence. Creating such a pipeline per element does not only reduce multiple passes over the sequence to a single pass, but also allows further optimizations. So, benefits of the new approach are that it can be more readable, less error prone, and more easily parallelized.

Our results can partly compare the results of research Ward and Diego (2015) who looked at the issue from a different perspective, but their main goal was also to determine if the newly introduced lambda expressions have a speed advantage over non-lambda expressions for an identical task. The conclusion of their research is that, "the addition of lambda expressions to Java 8 provide for more functional, concise, and readable coding. In addition, given enough execution time, the new lambda expressions can provide a performance advantage." We absolutely agree with them and results of our research is proof.

Of course, a readability is a matter of practice and of habit. Moreover, there is not a black and white answer to this problem. Javin Paul (2014) wrote: "When I first saw a Java code written using lambda expression, I was very disappointed with cryptic syntax and thinking they are making Java unreadable now, but I was wrong." We find similar opinions in practice, among IT students and at programming oriented blogs very often. Frequently, there is presented a belief that it is not worth paying attention to the new features of Java 8. As a result, these programmers remain in the use of old practices. However, we think that programmers who fail to adopt the new strategies will soon find themselves left behind. As far as the code length is concerned, in some simple examples, the version that uses aggregate operations is longer than the one that uses a *for each loop*, in some more complex tasks that versions that use bulk-data operations will be more concise. We agree with Angelika Langer (2015) statement: “Expecting that parallel stream operations are always faster than sequential stream operations is naive.” Using functional programming is not always better for various reasons. Some of them we try to mention briefly above and others require a deeper investigation of the problem. In fact, it is never “better”, it is just different and allows us to reason about problems differently. Never expect a miracle, also, do not guess; instead, benchmark a lot.

REFERENCES

Aggregate Operations. The Java™ Tutorials. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://docs.oracle.com/javase/tutorial/collections/streams/index.html⟩.

API Specification. Java™ Platform, Standard Ed. 8. 2017. © Oracle 1993–2017. [online]. [cit. 2017-03-10]. Available on internet: ⟨https://docs.oracle.com/javase/8/docs/api/overview-summary.html⟩.

Class ArrayList<E>. Java™ Platform Standard Ed. 8. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨https://docs.oracle.com/javase/8/docs/api/java/util/ArrayList.html⟩.

Class Colectors. Java™ Platform Standard Ed. 8. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨https://docs.oracle.com/javase/8/docs/api/java/util/stream/Collectors.html⟩.

Collections. The Java™ Tutorials. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://docs.oracle.com/javase/tutorial/collections/index.html⟩.

Default Methods. The Java™ Tutorials. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨https://docs.oracle.com/javase/tutorial/java/IandI/defaultmethods.html⟩.

HORSTMANN, S. CAY. 2014. Java SE 8 for the Really Impatient. Addison-Wesley Professional, 2014. ISBN-13: 978-0-321-92776-7, ISBN-10: 0-321-92776-1. 240 pp.

Interface Iterable<T>. Java™ Platform Standard Ed. 8. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨https://docs.oracle.com/javase/8/docs/api/java/lang/Iterable.html⟩.

Interface Iterator<E>. Java™ Platform Standard Ed. 8. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨https://docs.oracle.com/javase/8/docs/api/java/util/Iterator.html⟩.

Interface List<E>. Java™ Platform Standard Ed. 8. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨https://docs.oracle.com/javase/8/docs/api/java/util/List.html⟩.

Interface Stream<T>. Java™ Platform Standard Ed. 8. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨https://docs.oracle.com/javase/8/docs/api/java/util/stream/Stream.html⟩.

JAVIN, P., 2012. Difference between LinkedList vs ArrayList in Java. © 2012–2017. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://javarevisited.blogspot.sk/2012/02/difference-between-linkedlist-vs.html⟩.

JAVIN, P., 2014. 10 Example of Lambda Expressions and Streams in Java 8. © 2012–2017. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://javarevisited.blogspot.sk/2014/02/10-example-of-lambda-expressions-in-java8.html#axzz4oaVpthXZ⟩.

JAVIN, P., 2016. How does Enhanced for loop works in Java? © 2012–2017. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://javarevisited.blogspot.sk/2016/02/how-does-enhanced-for-loop-works-in-java.html#ixzz4jcEyMSvz⟩.

JAVIN, P., 2017. Java67: Java Programming tutorials and Interview Questions. © 2012–2017. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://www.java67.com/2012/12/difference-between-array-vs-arraylist-java.html#ixzz4i7DJtfrq⟩.

KISHORI, S. 2014. Beginning Java 8 Language Features: Lambda Expressions, Inner Classes, Threads, I/O, Collections, and Streams. Apress, 2014. ISBN 1430266597, 9781430266594. 704 pp.

Lambda Expressions. The Java™ Tutorials. 2017. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://docs.oracle.com/javase/tutorial/java/javaOO/lambdaexpressions.html#syntax⟩.

LANGER, A. and KREFT, K. 2013. Lambda Expressions in Java. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://www.angelikalanger.com/Lambdas/Lambdas.pdf⟩.

LANGER, A., 2015. Java performance tutorial — How fast are the Java 8 streams? [online]. [cit. 2017-03-10]. Available on internet: ⟨https://jaxenter.com/java-performance-tutorial-how-fast-are-the-java-8-streams-118830.html⟩.

SCHILDT, H., 2014. Java 8 eSampler. Preview exclusive excerpts from brand-new and forthcoming Oracle Press Java JDK 8 books. Oracle Press. © 2014 McGraw-Hill Education. [online]. [cit. 2017-03-10]. © 2010. Available on internet: ⟨http://www.oracle.com/technetwork/java/newtojava/java8book-2172125.pdf⟩.

Spring. 2017. Working with Spring Data Repositories. © 2011–2017. [online]. [cit. 2017-03-10]. Available on internet: ⟨https://docs.spring.io/spring-data/data-commons/docs/1.6.1.RELEASE/reference/html/repositories.html⟩.

WARD, A., DEUGO, D. 2015. Performance of Lambda Expressions in Java 8. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://worldcomp-proceedings.com/proc/p2015/SER2509.pdf⟩.

WILLIAMS, M. and J. Q. NUNEZ, 2015. Java SE 8: Lambda Quick Start. © Oracle. [online]. [cit. 2017-03-10]. Available on internet: ⟨http://www.oracle.com/webfolder/technetwork/tutorials/obe/java/Lambda-QuickStart/index.html⟩.

Jana Jurinová
Department of Applied Informatics and Mathematics,
University of SS. Cyril and Methodius, 917 01 Trnava,
Slovak Republic
Email: jana.jurinova@ucm.sk