Modern Management based on Big Data II and Machine Learning and Intelligent Systems III A.J. Tallón-Ballesteros (Ed.) © 2021 The authors and IOS Press. This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).

doi:10.3233/FAIA210245

Role of Big Data Analytics in Belt and Road Initiative (BRI): Multivariate Analysis with Gaussian Distribution of Data

Valliappan Raju ^{a1}, Wang Juan ^b, Sandeep Shrestha ^b, Arrunkumar Kalathinathan ^c, KK. Ramachandran ^d

^a Limkokwing University of Creative Technology, Malaysia
^b Qiqihar Medical University, China
^c Research Associate, World Research Union, India
^d GRD College of Arts & Science, India

Abstract. This manuscript focuses on the Belt and Road Initiative (BRI) of China, whereby the focus is on the engagement of big data analytics to comprehend logistics exertion. China is the trendsetter for revolutionary practices in trade, logistics, and technology. The recent progress the nation is thriving is on 'One Belt One Road' project whereby 65 countries are involved. It aims to connect continents and circulate smooth trade between them. This paper addresses the role of the database to identify the inter-model logistics in BRI. The merits of this project in the perspective of economic growth are measured through a quantitative study with 112 samples. Goal-setting theory is used to construct a conceptual framework for the research. Multivariate analysis is executed with SmartPLS 3.3.3 followed by an in-depth structural equation modeling. Normal distribution of data was given importance as in statistics the real-value of random variables whose distributions are not known, thus Gaussian distribution of data was used. Out of 6 Hypotheses, it is noted that five are significantly positive. Hypothesis testing is concluded based on p-value and t-statistics. The outcome of research suggests that big-data analytics is a major contributor in determining the significant model on logistics in Belt and Road Initiative.

Keywords. Big Data Analytics, Belt and Road Initiative, China, Economic Growth, Silk Road

1. Introduction

The Belt and Road Initiative (BRI) of China has emerged as one of the spectacular leaps in infrastructural development. Although several nations hold bilateral agreements to uplift their connectivity, China has gone far beyond expectation through its BRI project. The owners of the silk route prove to be sustainable in innovation and development [1]. It is to be noted that a study by CBER in 2019 forecasted that project BRI may boost global GDP (Gross Domestic Product) by \$7.1 trillion per annum by 2040 [2]. In the past

¹ Corresponding Author, Valliappan Raju, Limkokwing University of Creative Technology, Malaysia; Email: valliappan.raju@limkokwing.edu.my

two decades, the economy in China has undergone mammoth shifts positively. China has agreed to disseminate its funding up to USD 1.25 trillion by 2025, which shall be the highest ever FDI that has happened from the USA. The main objective was to build a singular and competent market to set up a trading place where both domestic and international commerce takes place through unity, integrating technology and cultural exchange [3]. To achieve this, it is important to understand the nuances of this project by comprehending the role of big data [4]. After completion of this project, one can evaluate the cost-opportunity of shipping, consigning, and traveling [5]. If this logistics is done effectively, then it adds more value. Through this research paper, the authors enumerate that by using Big data analytics, prominent logistics can adhere.

1.1 Overview of Big Data Analytics

Big Data Analytics comprises two distinct verbs, namely, Big Data and Analytics. Big data represents a lot of data. However, big data is not defined only in terms of volume (storage capacity) but also variety (different types of data) and velocity (the speed at which data is created). Big data can be acquired from multiple sources – web log files from the web, text data from surveys, geospatial data from mobile phones, social media are some examples [6]. Hence, data can sometimes be collected specifically for a certain form of analytics [7]. Put together, Big Data Analytics is an extremely powerful tool to understand and predict.

1.2 Big Data Analytics in Logistics

Gone are the days when Logistics used to be substandard due to data imprecision or usage of EPR systems for maintaining warehouses or unorganized inventories or limited communication with consumers or unpredictable traffic patterns and weather conditions or unpredictable market behavior [8]. Big Data Analytics has made significant benefits in Logistics. Analytics of this data leads to a better understanding of purchase patterns of various consumers, the latest demands in the market, and maintenance cycles [9].

2. Literature Review

The core purpose of a literature review is to identify scholarly inputs in the respective area of research and to provide immense knowledge to readers about the respective research [10]. While reviewing, strengths and weaknesses can be comprehended. A meticulous review of the literature was done on Belt and Road Initiative (BRI) and Big Data Analytics (BDA). Coherently both these concepts are new and limited literature was found [9]. However, several scholars have aligned their views that new technologies can build economies with rapid growth. It has to be understood that Machine Learning is part of Big Data, thus Machine Learning plays a dominant role [4]. In the context of BRI, the codes for machine learning are to be used during the implementation phase. With this seed of thought, literature from 2014 was considered, the reason being, project BRI was officially initiated in the year 2013 [1].

There are 347 research papers about Belt and Road Initiative (BRI) as of October 2018 particularly which are indexed with Web of Science core collection journals [3]



Figure 1- BRI on WoS Database (Source: Shahriar, 2018)

Figure 1 illustrates the bibliographic analysis between Web of Science core collections and research papers. Furthermore, below Table 1 represents the details of 32 research manuscripts that dealt with about Belt and Road Initiative along with their findings

The key five points from the CEBR study pointed out that,

- 1) By this project, BRI China may continue to lead the world economy in terms of production, infrastructure, employment rate etcetera. With this, China has a high chance to lead the World sustainably for the next few decades [11], [12]
- The advantages of project BRI are meritorious because 70 countries have agreed to be part of it, and around 56 countries forecast that their GDP will be escalated to USD 10 billion and beyond by 2040.
- The impact of this project can trigger economical shifts in Russia, India, Japan, Korea, Indonesia, Netherlands, and United Kingdom [13].
- 4) The United States of America could be a potential beneficiary from this project BRI though it has no direct involvement. Thus by 2040, the USA also has the potential to get benefitted from project BRI [14]
- Planet Earth is supposed to be transformed by BRI. This makes the clear notion that project BRI has the meritocracy to change the landscape of business in the World [15]

2.1 Underpinning Theory

Edwin A. Locke constructed Goal Setting theory in the year 1986. Goal Setting Theory has been considered in this research as underpinning theory as one of the main components to distinguish between Belt and Road Initiative and Big Data Analytics [16].

2.1.1 Rationale of Underpinning Theory

The by-product of this research is on project performance, that is, by identifying the role of big data analytics in BRI project, the ultimate aim is to successfully complete the project. Big Data is a means [17]. Therefore, the it would be right to presume that 'project



Figure 2- Goal Setting Theory Model

performance' is the dependent variable for this research. To measure and successfully complete a project, the short-term and long-term goals has to be set. Thus, this theory has been adopted in this respective research.

2.2 Conceptual Framework

Based on Goal Setting Theory, a conceptual framework has been developed for respective research whereby the core purpose is to relate big data analytics in Belt and Road initiative (BRI). Here, BRI is considered as the principal (main goal), and big data is considered as an agent (processor). In this connotation, the variables were retrieved from literature and the above-mentioned theories.



Figure 3- Conceptual Framework

3. Research Methodology

The independent variables in the above conceptual framework were randomly reviewed and enforced to have a normal distribution of data in data collection [18]. Gaussian distribution of data means to be normally distributed, that is, to normally deviate without hitting extremes [1]. The standard deviation to be within +1 to -1 as the curve to be in bell shape. This can be known only after data is collected.

The sample size can be arrived based on following formula,

$$n = N * \frac{X}{X + N - 1}$$

Where N = *population size;* X = *target population*

However, for this research, data were collected from 112 samples which were based on Krejcie & Morgan's (1970) sample-size calculation method. These 112 respondents are from IT, Supply chain management, Chinese diplomats, and economists. Thus, it is a stratified random sampling that consists of four different strata namely IT industry, personnel from the logistics sector, diplomats from the Chinese embassy in ASEAN, and lastly from economists who possess technical understanding towards research analysis [19].

4. Data Analysis

The analysis is done to determine the project performance of BRI using Big Data Analytics as moderator and data inputs as mediator [20]. While the motive is to comprehend a prominent model to spur trade across the globe, inter-model transportation was one familiar strategy [21]. Adding to this, if big data analytics are done then the outcome is larger than inter-model transportation. To testify this, 112 respondents were surveyed, and data were collected.

4.1 Measurement Model

The measurement model is also called an outer model, whereby the indicators or items are analyzed with their significance with constructs. Reliability, validity, R square adjusted, multicollinearity is analyzed. Reliability is about consistency and validity is about the accuracy of items. Table 1 illustrates the reliability and validity

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
BigData_Analytics	0.850	0.864	0.912	0.777
Challenge	0.915	0.938	0.941	0.766
Data_Inputs	0.830	0.841	0.899	0.748
MD-DI*BD*PP	1.000	1.000	1.000	1.000
Project_Clarity	0.858	0.914	0.908	0.683
Project_Performance	0.813	0.842	0.872	0.584
Task_complexity	0.886	0.911	0.914	0.682

Table 1 - Construct Reliability and Convergent Validity

(Source: Generated from SmartPLS 3.3.3)

4.2. Parameters for Measurement Model

The threshold for reliability is <0.7 and the threshold for convergent validity is <0.5. In Table 3, the reliability is identified by both Cronbach's Alpha and Composite Reliability. Cronbach's Alpha is calculated for vetting Confirmatory[22]. Factor Analysis (CFA) whereas Composite Reliability is calculated for vetting Exploratory Factor Analysis (EFA). Convergent Validity is calculated through AVE. According to [23] the threshold for reliability (represented by Cronbach's alpha) is it should be more than 0.7. By this, all variables are reliable, as the least value is 0.813 (for Project Performance). Similarly, composite reliability needs to be greater than 0.7, which is also complied with. Average Variance Extracted (AVE) is the convergent validity of items or indicators. The threshold for AVE has to be greater than 0.5 [24]Above table reflects a similar outcome, so the validity of the instrument is high. The square root of the AVE value will be the same as the respective construct's value in discriminant validity. Discriminant validity represents the accuracy of items in the research instrument. Following Table 2 highlights of Discriminant Validity

Fornell-Larcker Criterion							
	BigData Analytics	Challenge	Data_ Inputs	MD- DI*BD *PP	Project_ Clarity	Project Performance	Task_ complexity
BigData_Analytics	0.881						
Challenge	0.801	0.875					
Data_Inputs	0.879	0.928	0.865				
MD-DI*BD*PP	-0.147	-0.447	-0.423	1.000			
Project_Clarity	0.905	0.940	0.971	-0.338	0.826		
Project_Performance	0.945	0.877	0.931	-0.219	0.958	0.764	
Task_complexity	0.913	0.846	0.930	-0.183	0.945	0.954	0.826

Table 2 - Discriminant Validity

(Source: Generated from SmartPLS 3.3.3)

The value of construction in each column has to be more than the value of the rest for discriminant validity. There are certain values lesser, namely, the project performance has got the value of 0.764 whereas the other value is 0.954. This proves that the validity is weak, however, for other constructs, it's within the limit. VIF (Variance Inflation Factor) measures the multicollinearity of indicators of constructs. According to [24]the VIF in the outer range should not be more than 10 whereas according to [25] 5 is the threshold.

Table 3 –	Multicollinearity	(VIF)
-----------	-------------------	-------

	VIF
BD1	1.414
DI1	1.558
Data_Inputs * BigData_Analytics	1.000
PC1	9.127
PP1	2.571
TC1	7.542

(Source: Generated from SmartPLS 3.3.3)

In this analysis, only one item in each construct is considered. Coherently convincing both these scholarly thresholds, the respective research has less than 5 for all constructs except for Project Complexity. This proves that the VIF is at par with the threshold. Above Table 3 highlights the VIF values.

4.3 Structural Modeling

Structural Modeling is also called an inner model. Structural modeling in Multivariate analysis comprises mediation analysis through PLS-SEM (Partial Least Square-Structural Equation Modeling), moderation analysis [26]. Data Input is the mediator and it is hypothesized between the independent variable and dependent variable. Results of indirect effects in a model can yield the potency of the mediating effect. Following table 4 highlights the values on indirect effects

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (O/STDEV)	P Values
Task_complexity -> Data_Inputs -> Project_Performance	0.112	0.116	0.052	2.167	0.031
Challenge -> Data_Inputs -> Project_Performance	0.105	0.110	0.047	2.221	0.027
Project_Clarity -> Data_Inputs -> Project_Performance	0.328	0.328	0.083	3.958	0.000

Table 4 – Indirect Effects (Mediating Effects)

(Source: Generated from SmartPLS 3.3.3)

From Table 4 it is evident that Data Input mediates between the constructs efficiently. Based on the p-value, which is lesser than 0.05 it can be understood that all three independent variables require a mediator to achieve project performance (dependent variable).

4.4 Hypothesis Testing

The vital section in research is testing the claims or hunches or assumptions. Hypothesis testing is done for this respective research as this is more of quantitative research. A familiar or renowned technique to interpret the hypothesis testing is based on a p-value score. The P-value score is based on T-statistics. These both go hand-in-hand. P-value (probability value) has to be lesser than 0.05, which is 95% has positivity and 5% is doubtful. Thus, as long as 5% is in the doubtful scenario, the model is accepted.

5. Conclusion

Belt and Road Initiative (BRI) is a visionary project initiated by President Xi Jinping of China. After several challenges and constraints, this project has come to the construction stage in 2020. To amplify this project, there are strategies in inter-model logistics. However, using data from past trades in silk routes, there can be an optimized solution than inter-model [27]. If Big Data Analytics is used, this model will be more successful. However, this was merely the author's opinion before drafting this research manuscript.

After indulging in a data collection with 112 respondents who are involved with BRI, the data were analyzed and found that respondents abide that big data analytics will amplify the economic bottom line of the project and also enhance the performance. Big Data Analytics was not tested as an independent variable, whereas, it was placed as a moderator. Through moderation analysis, it was found that the p-value was 0.011 and t-statistics was 2.560, which is considered highly significant.

Reference

- V. Raju, "Implementing Flexible Systems in Doctoral Viva Defense Through Virtual Mechanism," *Glob. J. Flex. Syst. Manag.*, vol. 22, no. 2, pp. 127–139, Jun. 2021, doi: 10.1007/s40171-021-00264-y.
- [2] C. Policy, "Belt and Road Initiative to boost world GDP by over \$7 trillion per annum by 2040 -CEBR." https://cebr.com/reports/belt-and-road-initiative-to-boost-world-gdp-by-over-7-trillion-perannum-by-2040/ (accessed Aug. 09, 2021).
- [3] S. Shahriar, *Literature Survey on the "Belt and Road" Initiative: A Bibliometric Analysis*. Northwest A&F University, China, 2019.
- [4] A. Petraška, K. Čižiūnienė, A. Jarašūnienė, P. Maruschak, and O. Prentkovskis, "Algorithm for the assessment of heavyweight and oversize cargo transportation routes," *Vilnius Gedim. Tech. Univ.*, vol. 18, no. 6, pp. 1098–1114, Nov. 2017, doi: 10.3846/16111699.2017.1334229.
- [5] V. Raju and S. P. Poh, "Sustainability in Performance Management through Supply Chain Management," *Int. J. Supply Chain Manag.*, vol. 8, no. 2, pp. 1085–1089, Apr. 2019, Accessed: Jul. 08, 2021. [Online]. Available: http://www.ojs.excelingtech.co.uk/index.php/IJSCM/article/view/2866.
- [6] R. Perez-Vega, V. Kaartemo, C. R. Lages, N. Borghei Razavi, and J. Männistö, "Reshaping the contexts of online customer engagement behavior via artificial intelligence: A conceptual framework," *J. Bus. Res.*, Nov. 2020, doi: 10.1016/j.jbusres.2020.11.002.
- [7] F. Visconti and J. M. de Paz, "Electrical Conductivity Measurements in Agriculture: The Assessment of Soil Salinity," in *New Trends and Developments in Metrology*, InTech, 2016.
- [8] G. Rampersad, "Robot will take your job: Innovation for an era of artificial intelligence," J. Bus. Res., vol. 116, pp. 68–74, Aug. 2020, doi: 10.1016/j.jbusres.2020.05.019.
- [9] V. Raju and S. Poh Phung, "Resilient Motive for Reviewing Literatures in Business Research: Exploring Process of Research Design," Apr. 2021. Accessed: Jul. 05, 2021. [Online]. Available: http://annalsofrscb.ro.
- [10] J. E. Adaletey, V. Raju, and S. P. Phung, "Role of stakeholder in revenue mobilization to alleviate poverty in Ghana using E-governance mechanisms," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 2S, 2018.
- [11] B. Kar, "Service Quality and SERVQUAL Model: A Reappraisal," Amity J. Oper. Manag. ADMAA, vol. 1, no. 2, pp. 52–64, 2016.
- [12] M. J. Alexander, E. Jaakkola, and L. D. Hollebeek, "Zooming out: actor engagement beyond the dyadic," J. Serv. Manag., vol. 29, no. 3, pp. 333–351, 2018, doi: 10.1108/JOSM-08-2016-0237.
- [13] M. Lee and L. J. Kray, "A gender gap in managerial span of control: Implications for the gender pay gap," Organ. Behav. Hum. Decis. Process., vol. 167, pp. 1–17, Nov. 2021, doi: 10.1016/j.obhdp.2021.06.001.
- [14] R. Flepp, P. Meier, and E. Franck, "The effect of paper outcomes versus realized outcomes on subsequent risk-taking: Field evidence from casino gambling," *Organ. Behav. Hum. Decis. Process.*, vol. 165, pp. 45–55, Jul. 2021, doi: 10.1016/j.obhdp.2021.04.003.
- [15] X. Guan, M. Liu, and Y. Meng, "A comprehensive ecological compensation indicator based on pollution damage – protection bidirectional model for river basin," *Ecol. Indic.*, vol. 126, p. 107708, Jul. 2021, doi: 10.1016/j.ecolind.2021.107708.
- [16] K. Plys, "Theorizing capitalist imperialism for an anti-imperialist praxis towards a rodneyan worldsystems analysis," J. World-Systems Res., 2021, doi: 10.5195/jwsr.2021.1022.
- [17] V. Raju and S. P. Phung, "Economic dimensions of blockchain technology: In the context of extention of cryptocurrencies," *Int. J. Psychosoc. Rehabil.*, vol. 24, no. 2, pp. 29–39, 2020, doi: 10.37200/IJPR/V24I2/PR200307.
- [18] M. R. Jalilvand, J. Khazaei Pool, H. Balouei Jamkhaneh, and R. A. Tabaeeian, "Total quality management, corporate social responsibility and entrepreneurial orientation in the hotel industry," *Soc. Responsib. J.*, vol. 14, no. 3, pp. 601–618, 2018, doi: 10.1108/SRJ-04-2017-0068.

- [19] N. Bin Chang, K. Bai, and C. F. Chen, "Integrating multisensor satellite data merging and image reconstruction in support of machine learning for better water quality management," *J. Environ. Manage.*, vol. 201, pp. 227–240, 2017, doi: 10.1016/j.jenvman.2017.06.045.
- [20] S. U. A. Rana and V. Raju, "Factors influencing glass ceiling focus on women administration in higher education in Malaysia," *Int. J. Eng. Adv. Technol.*, 2019, doi: 10.35940/ijeat.F1220.0986S319.
- [21] R. W. and F. D. D. Zhang, M, Griffeth, "Corporate social responsibility, continuous process improvement orientation, organizational commitment and turnover intentions," *Int. J. Qual. Reliab. Manag.*, vol. 25, no. 2010, 2011, doi: 10.1108/JHOM-09-2016-0165.
- [22] J. A. G. J. T. Quin, "Comparison of Transformational Leadership Practices: Implications for School Districts and Principal Preparation.," J. Leadersh. Educ., vol. 14, no. 3, pp. 71–85, 2015.
- [23] J. F. Hair Jr, G. T. M. Hult, C. Ringle, and M. Sarstedt, A primer on partial least squares structural equation modeling (PLS-SEM). Sage Publications, 2017.
- [24] J. F. Hair, "Successful Strategies for Teaching Multivariate Statistics," Proc. 7th Int. Conf. ..., pp. 1–5, 2006.
- [25] C. Ringle, S. Wende, and J. Becker, "Ringle, Christian M., Wende, Sven, & Becker, Jan-Michael. (2015). SmartPLS 3. Bönningstedt: SmartPLS. Retrieved from http://www.smartpls.com," *Retrieved from*. 2015.
- [26] K. A. Law, A. Bhaumik, P. Sun, and V. Raju, "Factors determining the relationship between superiors and their subordinates: Evaluating the trust factor in Chinese organizations," *Int. J. Psychosoc. Rehabil.*, vol. 24, no. 4, 2020, doi: 10.37200/IJPR/V24I4/PR201082.
- [27] J. Lu, L. Ren, C. Zhang, D. Rong, R. R. Ahmed, and J. Streimikis, "Modified Carroll's pyramid of corporate social responsibility to enhance organizational performance of SMEs industry," *J. Clean. Prod.*, vol. 271, p. 122456, 2020, doi: 10.1016/j.jclepro.2020.122456.