# Time-series Dataset for Risk Assessment in Multi-echelon Supply Chain Networks

**Research Proposal** · April 2019

DOI: 10.17632/gystn6d3r4.2"



**5 authors**, including:

Grishma Saparia
SRM Institute of Science and Technology


Heerok Banerjee
Vrije Universiteit Brussel


Priyanshi Garg
SRM Institute of Science and Technology


V. Ganapathy
Saveetha School of Engineering, Savvetha Universtiy

# Time-series Dataset for Risk Assessment in Multi-echelon Supply Chain Networks

Grishma Saparia[1] *H. Banerjee* *P. Garg*

V. Ganapathy V. M. Shenbagaraman

## 1 Summary

The emergence of business intelligence has broadly reinforced the interplay among business entities in the digital realm. These AI tools are undoubtedly powerful in employing self-learning paradigms to facilitate convenient services but such tools still remain inadequate to reduce the impact of inevitable risks involved in business operations. In context to Supply Chain Risk Management (SCRM), it is therefore an open problem to eliminate and more significantly, optimize the impact of such risky operations that have been an evolutionary target for supply chain managers. This dataset is primarily targeted to support self-learning models and evolutionary algorithms for estimating non-linear functions and trends in supply chain systems. The Time-series sequence consists of 6 Lakh timesteps of seven attributes namely, Timestamp, RI_Supplier1, RI_Distributor1, RI_Manufacturer1, RI_Retailer1, Total_Cost, SCMstability_category. The time-series sequence is generated by simulating a multi-echelon supply chain network in MATLAB with three suppliers, one distributor, one manufacturer and one retailer. An arbitrary demand is introduced to the supply chain model and selective features such as total cost and risk index are calculated at each time-step. Section 3 describes the features and their mathematical formulation extensively. Finally, the obtained results are tabulated.

## 2 Scope of the Dataset

The purpose of this dataset is intended for academic purposes only. The dataset is not claimed to resemble the nature of a real-time supply chain networks. After a careful literature review, a set of succinctly defined mathematical models were selected to formulate the attributes and yield an empirical dataset. Hence, the dataset is generated based on limited assumptions on the underlying principles in supply chain networks.

This dataset is suitable for training machine learning models for time-series analyses and estimation of non-linear functions in risk-averse supply chain networks. It is thoroughly verified with absolute minimal outliers and could serve as an appropriate source for time-series analysis and evaluating non-linear auto-regressive models. As described in section 3, the calculated risk indices of supplier, distributor, manufacturer and retailer are non-linear in nature due to intended disruptions in the supply chain. Considering all the variables as cross-sectional data and performing trend analysis or regression analysis, a broader investigation on the nature of the risks involved in supply chain networks can be carried out.

# 3 Mathematical Modelling

In this section, a deterministic mathematical model for each attribute in the supply chain network is discussed.

For a given supply chain with 4 echelons: Supplier, Distributor, Manufacturer, Retailer and selective activities such as sourcing or supplying raw material to each component, assembling final products and delivering to destination markets, each component has its own cost, lead-time and associated risk.

1. The Risk Index (RI) is derived from the model proposed in [1], and can be mathematically formulated as:

$$RI_{supplier} = \sum_{i=1}^{n} \alpha s_{ij} \; . \; \beta s_{ij} \; . \; (1 - (1 - \prod_{j=1}^{m} P(\tilde{S}_{ij})) \tag{1}$$

   where, $\alpha s_{ij}$ is the consequence to the supply chain if the $i^{th}$ supplier fails,
   $\beta s_{ij}$ is the percentage of value added to the product by the $i^{th}$ supplier,
   $P(\tilde{S}_{ij})$ denotes the marginal probability that the $i^{th}$ supplier fails for $j^{th}$ demand,

   Similarly, the risk indices for the rest of the components can be calculated as:

$$RI_{distributor} = \alpha d_{risk_i} \; . \; \beta m_i \; . \; (1 - (1 - P(\tilde{M}_j))) \tag{2}$$

$$RI_{manufacturer} = \alpha m_{risk_i} \; . \; \beta m_i \; . \; (1 - (1 - P(\tilde{M}_j))) \tag{3}$$

$$RI_{retailer} = \alpha r_{risk_i} \; . \; \beta r_i \; . \; (1 - (1 - P(\tilde{R}_j))) \tag{4}$$

2. For each set of demand, the cumulative risk index of the supply chain network can be calculated as:

$$TRI = w_1 \; . \; RI_{supplier} + w_2 \; . \; RI_{distributor} + w_3 \; . \; RI_{manufacturer} + w_4 \; . \; RI_{retailer} \tag{5}$$

   where, $w_1, w_2, w_3, w_4$ are arbitrary weights such that $w_1 + w_2 + w_3 + w_4 = 1$

3. The risk fluctuation subjected to the supply chain network is generated by a sine-wave generator, and is mathematically gives as :

$$\tilde{\delta_r} = A. \; sin(\omega t + \phi) + B_i \tag{6}$$

   where, $\tilde{\delta_r}$ denotes the absolute value of risk fluctuation,
   A denotes the peak amplitude of disruption,
   $\omega$ denotes the angular frequency, t denotes time,
   $\phi$ denotes the phase and $B_i$ denotes a bias

4. The total supply chain cost can be calculated as [2]:

$$TC = \xi \times \sum_{i=1}^{N} ( \mu_i \, . \, \sum_{j=1}^{N_i} C_{ij} y_{ij}) \tag{7}$$

where, N is the number of components,
$\xi$ denotes the period of interest,
$\mu_i$ is the average demand per unit time,
$C_{ij}$ is the cost of the $j^{th}$ resource option for the $i^{th}$ component,
$y_{ij}$ is a binary variable denoting whether the $i^{th}$ component is a participant for the $j^{th}$ resource option

5. For each couple of normalized risk index and total cost of the supply chain, the main objective function Z is given as [2]:

$$Z = w_1 \, . \, TC_n \; + \; w_2 \, . \, TRI_n \tag{8}$$

where, $TC_n$ is the normalized total cost,
$TRI_n$ is the normalized total risk index,
$w_1, w_2$ are the weights ; $w_1 + w_2 = 1$

# 4 Citing the Dataset

The dataset is available online in Mendeley library.
To manually add the dataset in the bibiliography, use the following :
"Banerjee, Heerok; Saparia, Grishma; Ganapathy, Velappa; Garg, Priyanshi; Shenbagaraman, V. M. (2019), "Time Series Dataset for Risk Assessment in Supply Chain Networks", Mendeley Data, v2 http://dx.doi.org/10.17632/gystn6d3r4.2"

# References


[1] Neureuther, Brian D., and George Kenyon. "Mitigating supply chain vulnerability." Journal of marketing channels 16.3 (2009): 245-263.

[2] Mastrocinque, Ernesto, et al. "A multi-objective optimization for supply chain network using the bees algorithm." International Journal of Engineering Business Management 5.Godište 2013 (2013): 5-38.