Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P3642-cd

INTEGRATING DEPENDENCY ANALYSIS THROUGH STRUCTURAL EQUATION MODELING AND ARTIFICIAL NEURAL NETWORKS: A CASE STUDY IN THE MINING INDUSTRY

Pablo Viveros Gunckel
Universidad Técnica Federico Santa María, Departamento de Industrias, Chile.
E-mail: Pablo.viveros@usm.cl
Valentina Flores Castro
Universidad Técnica Federico Santa María, Departamento de Industrias, Chile.
E-mail: valentina.floresca@sansano.usm.cl
Fredy Kristjanpoller Rodriguez
Universidad Técnica Federico Santa María, Departamento de Industrias, Chile.
E-mail: fredy.kristjanpoller@usm.cl
Christopher Nikulin Chandia
Universidad Alberto Hurtado, Facultad de Ingeniería, Chile.
E-mail: cnikulin@uahurtado.cl
Rodrigo Mena Bustos
Universidad Técnica Federico Santa María, Departamento de Industrias, Chile.
E-mail: rodrigo.mena@usm.cl

Industry 4.0 technologies are revolutionizing industrial maintenance management, highlighting Machine Learning (ML) techniques as key tools to anticipate failures more efficiently. In this study, the dependencies between components of a crushing line of a mining company in Chile and the different types of failures are analyzed, using ML models and structural equation models (SEM), with the objective of determining which ML model best fits the data, providing reliable relationships, so that in future work these relationships can be used in failure prediction models. Both models complement each other, since it is currently recognized the importance of a comprehensive approach in the analysis of failure types, allowing to improve maintenance management by offering an alternative to reduce costs associated with maintenance and downtime. The main motivation is to increase the accuracy of early warning systems, supporting more informed decision making. ML models such as Random Forest (RF) and Artificial Neural Networks (ANN) are employed, whose dependency analyses have shown positive results in previous studies. In addition, Structural Equation Modeling (SEM) is integrated, which allows exploring the complex interrelationships between system variables and different types of faults. The models were evaluated using the confusion matrix, accuracy, precision, recall and F1 score, complemented by SEM-derived indicators that reinforce the validity of the results. ANN showed outstanding performance with an accuracy of 0.9926 and significant relationships according to SEM, whereas RF suffered from overfitting, limiting its applicability in SEM. This dependence analysis provides a novel approach, using two techniques that together provide a more robust analysis of dependence research and contributing to existing research in this field.

Keywords: *machine learning; structural equation model; dependency analysis.*

1. INTRODUCTION

The wear and tear of industrial assets directly affects production and the quality of the final product, which makes efficient maintenance a key factor for operational sustainability. In the mining industry, maintenance can represent more than 30% of operational costs (Pinciroli L., et al. 2023), so improving its efficiency not only reduces costs, but also optimizes equipment availability and reliability.

In this context, the adoption of advanced technologies, such as predictive maintenance based on Industry 4.0, has become an essential strategy for asset management. Tools such as Digital Twins allow simulating equipment behavior in real time, using operational data to anticipate failures and optimize maintenance planning (Wang S., et al. 2016). These digital solutions have proven to be highly effective in providing a more detailed view of asset condition, enabling more accurate and timely intervention.

Artificial Intelligence (AI) and ML have revolutionized the way industrial data is analyzed, enabling early detection of anomalies and prediction of failures before they occur. Unlike traditional corrective or preventive maintenance approaches, ML-based models can identify hidden patterns in large volumes of data, incorporating dependencies between critical variables to improve prediction accuracy and optimize maintenance strategies (Moglen R. L., et al. 2023).

In addition, SEM has been used extensively in studies of interdependence between multiple components, providing a more holistic view of the relationships between key variables in industrial systems.

This study focuses on the analysis of data from a crushing line in a mining operation, integrating information from a Distributed Control System (DCS), a Digital Twin and operational records. From this integration, the performance of different Machine Learning models will be evaluated to determine which one offers the best results in terms of accuracy and predictive capability. The main objective is to provide a solid basis for maintenance optimization using advanced technologies, thus contributing to the improvement of operational efficiency in the mining industry, where the results are expected to serve for the creation of an improved early failure prediction model.

2. BACKGROUND AND LITERATURE REVIEW

Understanding the theoretical foundations and previous research on dependencies, ML, and SEM is crucial for developing a robust analytical framework. This section reviews key concepts and relevant studies that support the proposed methodology.

2.1. Dependencies and interdependencies

The analysis of dependencies and interdependencies in complex systems is essential to understand their interactions and prevent cascading failures (Sun W., et al. 2022). Dependencies are unidirectional relationships, while interdependencies are bidirectional, with mutual impacts. Although they often improve efficiency, they also increase vulnerabilities during failures (Moglen R. L., et al. 2023).

Studies highlight that modelling these relationships reveals hidden connections and improves the ability to foresee contingencies and make strategic decisions (Zio E., et al. 2011) (Huang H., et al. 2024). Tools such as SEM provide a robust framework for quantifying and evaluating these interactions in critical systems.

2.2. Structural equation models

SEM models causal relationships between variables, overcoming limitations of traditional techniques by handling complex interactions and measurement errors (Kline R. B., 2016). It has been applied in thermal comfort, disaster management and soil moisture studies (Elnabawi M. H., et al. 2024) (Geddam S. M., et al. 2024) (Wang S., et al. 2023). In conjunction with ML, SEM configures causal relationships, while ML identifies nonlinear patterns, offering robust analysis for complex systems.

2.3. Machine learning and algorithms

ML identifies complex patterns and makes predictions, highlighting supervised learning for classification. The most prominent models in this field include Logistic Regression (LR), Support

Vector Machine (SVM), Decision Trees (DT), RF and ANN, with RF and ANN standing out for their accuracy (Siddique A. B., et al. 2024) (Goodfellow I., et al. 2016).

On the one hand, RF combines multiple decision trees to improve accuracy and reduce overfitting. It uses random samples of data to train each tree and majority or averaged voting for predictions. It is robust to missing data and provides information on feature importance, although it can be complex and computationally intensive (Liaw A., et al. 2002).

On the other hand, ANNs Inspired by the human brain, they model complex nonlinear relationships through layers of interconnected nodes. Their training adjusts weights through backpropagation and activation functions such as ReLU or sigmoid. They are versatile and effective for complex problems but have disadvantages such as opacity in their interpretation and high computational costs (Goodfellow I., et al. 2016).

The analysis focuses on RF and ANN to study dependencies between operational components and fault types, integrating the best model with validated hypotheses in a SEM framework to improve prediction accuracy.

The following section presents the methodology of the study given the literature review.

3. METHODOLOGY

This section describes the methodological approach followed in this study, detailing the process from data collection to the selection and evaluation of the ML-SEM dependency model that best represents the system under analysis. The primary objective of this methodology is to identify the key operational parameters influencing failure occurrences and to establish a structured dependency model that enhances interpretability and predictive capabilities in an industrial mining environment.

To achieve this, the methodology is divided into three main stages. First, data collection, preprocessing, and exploration are performed to ensure the quality and reliability of the dataset, which is crucial for obtaining meaningful insights. Then, two machine learning models, RF and ANN, are trained and evaluated to identify the most influential factors affecting system failures. The interpretability of these models is enhanced using Shapley Additive exPlanations (SHAP), allowing a deeper understanding of variable contributions. Finally, the results obtained from the ML models are integrated into a SEM framework, enabling a more comprehensive assessment of the dependencies between variables and validating the relationships through statistical fit indices.

By combining ML techniques with SEM, this approach not only enhances predictive accuracy but also provides a structured representation of the relationships between operational variables and failure types. This hybrid methodology bridges the gap between black-box predictive models and interpretable statistical frameworks, facilitating more informed decision-making in predictive maintenance strategies.

3.1. Data collection, preprocessing and exploration

Data collection is performed from the system to be studied, ensuring that data on operating parameters and types of failures are correctly extracted and retained. A comprehensive database is recommended to avoid bias.

Data preprocessing ensures data quality, consistency and usefulness, including filtering of numerical variables, elimination of erroneous data and normalization of variables.

In data exploration, patterns, relationships and anomalies are identified, determining the relevant variables and analyzing the frequency of failures to work with the most frequent ones, always prioritizing data quality.

3.2. Dependency model training and evaluation

The data are used to train two models, RF and ANN, with the objective of identifying the most influential operating parameters in the types of failures. Mathematically RF is represented as follows:

$$\hat{f}(\mathrm{x}) = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \qquad (1)$$

Where B is the number of trees, T_b is the b-th tree and $\hat{f}(\mathrm{x})$ is the aggregate prediction.

On the other hand, ANN is represented as follows:

$$\mathrm{y} = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \qquad (2)$$

Where x_i are the inputs of the neuron, w_i are the weights associated with each input, b is the bias, which allows the activation function to be shifted, f is the activation function that introduces nonlinearity and finally, y is the output of the neuron.

After training, Shapley Additive exPlanations (SHAP) is used to decompose the predictions and understand the contribution of each variable. SHAP provides local interpretability, helping to identify the most relevant variables for each type of failure.

Finally, the models are evaluated using indicators such as the confusion matrix, accuracy, precision, recall and F1 score, with the confusion matrix showing the predictions against the actual classes.

3.3. SEM evaluation and integration

After evaluation of the ML models, the results obtained are fed into a SEM model. These results are structured so that each row represents an observation and each column an operational variable. The relationships specified in the model must be supported by the data, as they will be tested as hypotheses.

The SEM model is evaluated using several key indicators to measure its fit to the data. The most common ones are:

(i) Chi-square: Measures the relationship between categorical variables or model fit.

(ii) Comparative Fit Index (CFI): Measures how well the theoretical model fits the data compared to a null model. A value close to 1 indicates a good fit.

(iii) Tucker-Lewis Index (TLI): Compares the fit of the proposed model to the null model, penalizing more complex models. A value close to 1 is a good fit, and values between 0.90 and 0.95 are acceptable.

(iv) Root Mean Square Error of Approximation (RMSEA): Estimates the discrepancy per degree of freedom between the specified model and the perfect model. Values < 0.05 indicate a good fit, while values > 0.08 suggest a mediocre or poor fit.

(v) Standardized Mean Residuals (SRMR): Measures the average discrepancy between the observed and predicted correlations by the model. A value ≤ 0.08 is a good fit.

A combination of these indicators provides a complete evaluation of the model.

4. CASE STUDY

The system studied is a crushing line (Figure 1) of a copper mine located in northern Chile, which is composed of crusher No. 5, a CV-1C conveyor belt, the FE002 feeder and the CV-2C conveyor belt.

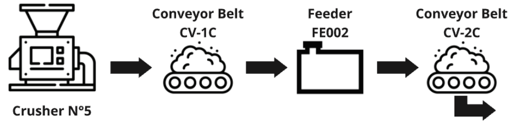

Fig. 1. Mining crushing line

The equipment in Figure 1 is connected to a DCS, which collects data to control industrial processes. This system is connected to a Digital Twin, which replicates the process in real time virtually. If there are significant discrepancies between reality and virtual, the information is sent to a maintenance database.

There is a unified database that includes maintenance records, DCS and the Digital Twin, with 547,953 historical data for a period from August 2020 to August 2021. Among the variables, there are 36 numeric, 6 qualitative and one time variable. Null columns and similar numerical variables are eliminated to avoid multicollinearity problems, and the data are standardized.

The 6 most relevant categorical variables are selected and the events that affect shredder No. 5 are observed (since this equipment is the one with the most data), where 87.5% of the data belong to scheduled maintenance. For the study, only unplanned events were selected: “Mechanical failures” and “Unscheduled operational stoppages”, leaving 31,834 records with 27 numerical and 2 categorical variables.

Machine learning models (RF and ANN) are trained using a multi-category classification approach. Relationships between variables are evaluated using SHAP values and entered a SEM model to validate hypotheses and understand causal relationships between components. The RF model is excellent for identifying nonlinear relationships and making accurate predictions, while the SEM model allows explicit representation of structural relationships between variables.

The ANN model follows a similar structure, with fully connected layers and ReLU activation function. Model performance is evaluated with classification metrics and further analysis is obtained using SHAP values for each fault type, which are then fed into the SEM model to obtain more detailed and structured explanations of the relationships between variables.

5. RESULTS

This section presents the results obtained from the different models, highlighting their classification performance, key influencing variables, and the fit of the SEM.

5.1. RF-SEM model results

With the training of the data using the RF model, a classification report was obtained with an accuracy of 0.9984 (see Table 1), indicating almost perfect classification. It is important to note that a class called "others" was included, which covers events that are neither mechanical failures nor unscheduled operational stops. This class, although not relevant to the study, is the majority and was kept to avoid losing valuable information.

Table 1. RF classification report

Accuracy:	0.9984			
Classification:	**Precision**	**Recall**	**F1-score**	**Support**
Mechanical Failures	1.00	1.00	1.00	372
Unscheduled Operational	1.00	0.98	0.99	176
Others	1.00	1.00	1.00	9003

The report also shows that the precision, recall (the proportion of true positives), and F1-score are almost perfect in all classes.

Next, using SHAP, the key numerical variables for mechanical failures and unscheduled operational stops are identified and presented in Table 2.

Table 2. RF hypotheses most influential variables for each event

Event	Most Influential Variables
Mechanical Failures	Resistance deviation + Ampere error Motor 1 + Total Tons DT + Performance Tons per Shift + Temperature Transmits
Unscheduled Operational	Resistance deviation + Ampere error Motor 1 + Total Tons DT + Performance Tons per Shift + Temperature Transmits

The model inputs are the dependent variables: unscheduled mechanical and operational failures, which are influenced by five independent variables. The SEM model is executed, and the results are shown in Table 3.

The test statistic for the User Model is 0.000 with 0 degrees of freedom, suggesting a perfect fit to the observed dataset, which could indicate overfitting. The CFI and TLI values of 1.000 also indicate a perfect fit, but this could also be problematic. Additionally, the RMSEA and SRMR values of 0.000 reinforce the idea of a perfect fit.

Perfect values in several indicators (RMSEA, CFI, TLI, SRMR) indicate overfitting, which could affect the model's ability to generalize to new datasets. Overfitting occurs when a model captures noise instead of relevant patterns, typically due to model complexity or a large sample size.

Table 3. SEM results of the RF model

Estimator	ML	
Optimization method	NLMINB	
Number of model parameters	13	
	Used	Total
Number of observations	24292	31834
Model Test User Model:		
Test statistic	0.000	
Degrees of freedom	0	
Model Test Baseline Model:		
Test statistic	2.776	
Degrees of freedom	11	
P-value	0.000	
User Model versus Baseline Model:		
Comparative Fit Index (CFI)	1.000	
Tucker-Lewis Index (TLI)	1.000	
Loglikelihood and Information Criteria:		
Loglikelihood user model (H0)	18.579	
Loglikelihood unrestricted model (H1)	18.579	
Root Mean Square Error of Approximation:		
RMSEA	0.000	
90 Percent confidence interval - lower	0.000	
90 Percent confidence interval - upper	0.000	
P-value H0: RMSEA ≤ 0.050	NA	
P-value H0: RMSEA ≥ 0.080	NA	
Standardized Root Mean Square Residual:		
SRMR	0.000	

5.2 ANN-SEM model results

When training the ANN model, the classification report presented in Table 4 was obtained, where it is observed that the accuracy is 0.9926, meaning that the model classified the events almost perfectly. It can also be seen that the class precision, recall, and F1-score are high values overall.

Table 4. ANN model results

Accuracy:	0.9926			
Classification:	**precision**	**recall**	**f1-score**	**support**
Mechanical Failures	1.00	0.97	0.99	372
Unscheduled Operational	0.98	0.66	0.79	176
Others	0.99	1.00	1.00	9003

SHAP is used to obtain the most important numerical variables for mechanical failures and unscheduled operational stoppages.

Table 5. ANN Hypotheses most influential variables for each event

Event	Most Influential Variables
Mechanical Failures	Ampere error Motor 2 + Ampere error Motor 1 + Temperature Transmite + Ampere Motor 1 DT + % of Voltage Utilization
Unscheduled Operational	Temperature Transmits + Ampere error Motor 1 + Ampere error Motor 2 + Performance Tons per Shift + Temperature Transmits 2

The SEM model aligns well with the hypothesis provided by the Neural Network model, with the algorithm successfully converging after 44 iterations and identifying the optimal parameters.

From Table 6, it can be seen that the statistical test for the fitted model is 173 with 4 degrees of freedom, indicating a difference between the model and the data. On the other hand, the Comparative Fit Index (CFI) is 0.939, indicating a fairly good fit. The TLI index, with a value of 0.772, is below 0.9, which is reasonable given that this index penalizes more complex models.

The RMSEA value is 0.036, which is below the typical threshold of 0.05, indicating a good fit of the model to the observed sample size. This is reinforced by the 90% confidence interval of [0.032, 0.041], which is also within acceptable values. On the other hand, the SRMR value is 0.012, which is significantly low, indicating an excellent fit of the model, as values below 0.08 are generally considered good.

Table 6. SEM results of the ANN model

Estimator	ML
Optimization method	NLMINB
Number of model parameters	13
Number of observations	31834
Model Test User Model:	
Test statistic	173
Degrees of freedom	4
P-value (Chi-square)	0.000
Model Test Baseline Model:	
Test statistic	2.797
Degrees of freedom	15
P-value	0.000
User Model versus Baseline Model:	
Comparative Fit Index (CFI)	0.939
Tucker-Lewis Index (TLI)	0.772
Loglikelihood and Information Criteria:	
Loglikelihood user model (H0)	25.563
Loglikelihood unrestricted model (H1)	25.649
Root Mean Square Error of Approximation:	
RMSEA	0.036
90 Percent confidence interval - lower	0.032
90 Percent confidence interval - upper	0.041
P-value H0: RMSEA ≤ 0.050	1.000
P-value H0: RMSEA ≥ 0.080	0.000
Standardized Root Mean Square Residual:	
SRMR	0.012

Although some indicators like the TLI suggest that the model is not completely perfect, other results like CFI, RMSEA, and SRMR point to the SEM model having an adequate fit. There may be room for improvement in the fit, but the results are good enough to consider that the model captures significant relationships between the variables. The causal relationships between the different variables are shown below in Figure 2.

The arrows show the direction of the relationships between the variables, and the coefficients indicate the magnitude of the effect. A negative coefficient (-0.08) between AD6E and Unscheduled Operations (UO) suggests that as AD6E increases, UO decreases. Negative coefficients indicate inverse relationships, while positive coefficients represent direct relationships. Additionally, the arrow between UO and Mechanical Failures (MF) with a coefficient of -0.06 indicates a small negative relationship, meaning that as UO increases, MF tends to decrease slightly.

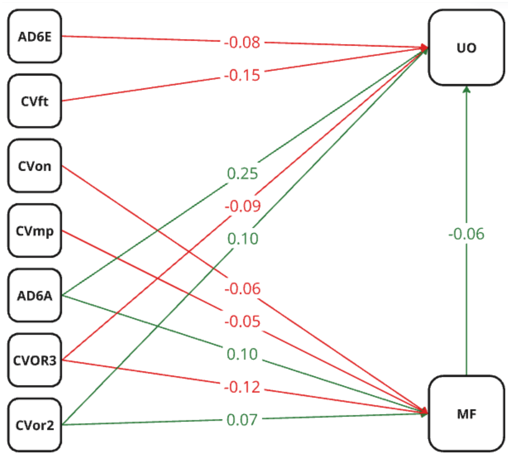

Fig. 2. Relation between variables

The meaning of the abbreviated variables is detailed in Table 7.

Table 7. Meaning variables

Abbreviation	Description Numeric Variable
AD6E	Temperature Transmits 2
CVft	Performance Tons per Shift
CVon	% of Voltage Utilization
CVmp	Ampere Motor 1 DT
AD6A	Temperature Transmits
CVOR3	Ampere Error Motor 1
CVor2	Ampere Error Motor 2

The results highlight significant differences in the performance of the RF-SEM and ANN-SEM models. While RF achieved near-perfect classification accuracy (0.9984), its SEM model exhibited signs of overfitting, as indicated by perfect fit indices (CFI, TLI, RMSEA, SRMR). This suggests that RF may be capturing noise rather than generalizable patterns. In contrast, ANN showed slightly lower classification accuracy (0.9926), but its SEM model demonstrated a more realistic fit (CFI = 0.939, TLI = 0.772, RMSEA = 0.036, SRMR = 0.012), indicating better generalizability. These findings suggest that while RF provides high accuracy, ANN-SEM may offer a more balanced trade-off between performance and model interpretability.

Future research should explore hybrid approaches.

6. CONCLUSIONS

The main objective of this research was to compare the performance of the RF and ANN machine learning models through SEM analysis, providing a more robust view of which model better fits the data. The SEM analysis played a key role in validating the dependency relationships identified by each model, revealing both the strengths and weaknesses of both approaches.

The RF model showed overfitting, which limited its ability to generalize to new data. This overfitting, confirmed through SEM, showed that the model was too closely fitted to the training data, capturing noise rather than meaningful relationships. This emphasizes the importance of using representative, high-quality data to prevent the model from capturing spurious patterns. The imbalance in the dataset, with only 1,944 instances of mechanical failures and unscheduled operational stops compared to 29,890 instances that were neither of these, underscores the need for diverse operational data in future research.

On the other hand, the ANN model showed superior performance, achieving 99.26% accuracy and effectively capturing the non-linear relationships of the system. The dependencies learned by the ANN model were validated through SEM, providing a solid foundation for its predictive capabilities.

This research demonstrates that ANN models outperform RF in identifying the key variables that affect failures in mining systems. By capturing complex dependencies, the ANN model provided more accurate predictions. This confirms the critical role of SEM in verifying the generalization and quality of machine learning model hypotheses, offering a valuable tool for ensuring the robust development of models.

In the future, the dependency relationships discovered are expected to be applied in predictive failure models, which could lead to improved early alerts and preventive measures.

References

Siddique A. B., N. Ali, M. Hamraz, S. Muhammad, M. R. Karim, S. U. Khan and S. A. Khattak (2024). *Prediction of magnetic nature of oxide compositions by using machine learning models*, Computational Condensed Matter.

Elnabawi M. H. and E. Jamei (2024). *The thermal perception of outdoor urban spaces in a hot arid climate: A structural equation modelling (SEM) approach*, Urban Climate, vol. 55, article 101969.

Pinciroli L., P. Baraldi, E. Zio (2023) *Maintenance optimization in industry 4.0*, Reliability Engineering & System Safety, vol. 234, p. 109204.

Wang S., J. Wan, D. Zhang, D. Li, and C. Zhang (2016). *Towards smart factory for industry 4.0: A self-organized multi-agent system with big data based feedback and coordination*, Computer Networks, vol. 101, pp. 158-168.

Moglen R. L., J. Barth, S. Gupta, E. Kawai, K. Klise, and B. D. Leibowicz (2023*). A nexus approach to infrastructure resilience planning under uncertainty*, Reliability Engineering & System Safety, vol. 230, p. 108931.

Sun W., P. Bocchini, and B. D. Davison (2022). *Overview of Interdependency Models of Critical Infrastructure for Resilience Assessment*, Natural Hazards Review, vol. 23, no. 1, p. 04021058.

Zio E. and G. Sansavini (2011). Modeling Interdependent Network Systems for Identifying Cascade-Safe Operating Margins, IEEE Transactions on Reliability, vol. 60, no. 1, pp. 94-101.

Huang H., Y. Tang, X. Shi, and X. Mao (2024). Dependency-Aware Neural Topic Model, Information Processing & Management, vol. 61, no. 1, p. 103530.

Kline R. B. (2016). Principles and practice of structural equation modeling (4th ed.), Guilford Press.

Geddam S. M. and C. A. Kiran (2024). Enhancing Disaster Management Effectiveness: An Integrated Analysis of Key Factors and Practical Strategies through Structural Equation Modeling (SEM) and Scopus Data Text Mining, Geohazard Mechanics, vol. 2, no. 2, pp. 95-107.

Wang S., R. Li, Y. Wu, and W. Wang (2023). Estimation of Surface Soil Moisture by Combining a Structural Equation Model and an Artificial Neural Network (SEM-ANN), Science of The Total Environment, vol. 876, p. 162558.

Goodfellow I., Y. Bengio, and A. Courville (2016). Deep Learning, MIT Press.

Liaw A. and M. Wiener (2002). Classification and regression by randomForest, R News, vol. 2, no. 3, pp. 18-22.